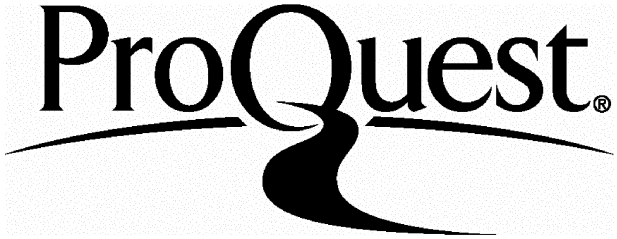
Mapping Studies of Mouse Chromosome Two

Lorna Wendy Pate

(University College London)

Thesis presented to the University of London for examination for the degree of Ph.D.

October 1994



To Tim, Mum and Dad

Acknowledgments

I acknowledge and thank Dr. Cathy Abbott for all the help, support and encouragement that she has given me over the years, without which this study would not have been possible.

Special thanks go to Professor Sue Povey for stepping in at a rather late stage, and taking over as my University of London supervisor.

Also, Many thanks to Dr. Margaret Fox for all the time, help and encouragement with the cell culture aspects of my study.

Thanks to Dr. Jonathan Wolfe for help in securing extended funding.

Thanks also to Alison Pilz and Stavros Malas for many happy times in the lab. Special thanks to Ali for all the technical help! A special thank you to Stavros for animated discussion, also for all the chocolate cake which sustained me during my final months!

Thanks also to Dr. C. Beechey at the MRC Radiobiology Unit, for supplying the translocation carrier mice. Also to Dr. P. Lalley and Dr. K. Fournier for supplying the EBS18Az and ABm hybrids, and to Dr. J. Thacker for supplying the V79TOR2 cells.

Many thanks to the Medical Research Council for supplying not only three years funding, but also for granting me an extension to complete my lab work.

Finally, all my thanks to my family and friends - Tim, for all his love and encouragement, without which I wouldn't be writing this. Mum and Dad, Richard, Mary and family for love and support, Kimberley, for being a great friend! Darren, Richard and Jo for reminding me that there life outside Genetics (and providing a suitable diversion!). Also, thanks to all my friends at UCL - too many to list!

Abstract

Mouse chromosome two harbours the interesting mutant genes ragged (*Ra*), wasted (*wst*), lethal spotting (*ls*) and ulnaless (*Ul*). These studies were intended to provide resources for the initiation of positional cloning projects to isolate one or more of the genes involved in these mutations. The first approach was the construction and characterization of two new panels of mouse:Chinese hamster somatic cell hybrids. the mouse parental cell line in each case being derived from spleen cells of mice carrying either the T(2;8)2Wa or the T(2;16)28H reciprocal translocations. These translocation breakpoints flank the region on distal mouse chromosome two in which ragged and wasted are thought to lie, and provide a means of defining the physical limits of the region. Forty four separate cell lines were produced, but no segregation of the translocation reciprocals in either panel was noted, despite subcloning. Forced antibody selection strategies were also investigated. The second aim of the project was to generate new molecular markers for mouse chromosome two by species specific Interspersed Repetitive Sequence Polymerase Chain Reaction (IRS-PCR) and Interspersed Repetitive Sequence to bubble (IRS-bubble) Polymerase Chain Reaction amplification of DNA from hybrid cells containing only mouse chromosome two. The PCR products were screened for the presence of putative CpG islands and microsatellite repeats. Fragments were then cloned and sequenced. PCR primers were designed from the sequence obtained, and variation was identified between *Mus musculus* and *Mus spretus*. Twelve new markers were produced. The new markers were then placed on the map of mouse chromosome two by interspecific backcross analysis. New markers showing linkage to ragged, wasted or ulnaless were then mapped in relation to these mutants by the use of interspecific and intersubspecific backcrosses in which the mutant was segregating.

Contents

Chapter 1 - General Introduction

Chapter two - Materials and methods

Chapter three - Somatic cell hybrid results

Chapter 4 - The generation of new molecular markers

Chapter 5 - Mapping results

Chapter 6 - Discussion

References

Appendix

List of Figures

List of tables

Abbreviations

ADOK - adenosine kinase
AMP - adenosine 5' monophosphate
ADP - adenosine 5' diphosphate
ATP - adenosine 5' triphosphate
APRT - adenine phosphoribosyl transferase
CMP - cytidine 5' monophosphate
CDP - cytidine 5' diphosphate
CTP - cytidine 5' triphosphate
dATP - deoxy adenosine 5' triphosphate
dCTP - deoxy cytidine 5' triphosphate
dGTP - deoxy guanosine 5' triphosphate
dTTP - deoxy thymidine 5' triphosphate
dUMP - deoxy uridine 5' monophosphate
FCS - fetal calf serum
GMP - guanosine 5' monophosphate
GDP - guanosine 5' diphosphate
GTP - guanosine 5' triphosphate
HPRT - hypoxanthine phosphoribosyl transferase
IFGT - irradiation fusion gene transfer
IRS-PCR - interspersed repetitive sequence polymerase chain reaction
IRS- bubble PCR - interspersed repetitive sequence to bubble polymerase chain reaction
IMP - inosine 5' monophosphate
LINE - long interspersed repetitive element
MEM - minimal essential medium
NOR - nucleolus organizer region
PIC - polymorphism information content
PRPP - phosphoribosyl pyrophosphate
sAMP - adenylosuccinic acid
SINE - short interspersed repetitive element
TMP - thymidine 5' monophosphate
TDP - thymidine 5' diphosphate
TTP - thymidine 5' triphosphate

Chapter 1

General introduction

1.1) The aims of the project

This project had three objectives :-

a) The development of a new panel of DNA markers for mouse chromosome two, particularly in the region of the interesting developmental mutants ragged (*Ra*), wasted (*wst*), lethal spotting (*ls*) and ulnaless (*Ul*).

b) The generation of two new panels of somatic cell hybrids from mouse parental cell lines carrying the reciprocal tranlocations T(2;8)2Wa and T(2;16)28H, to allow physical delineation of the region flanked by the translocation breakpoints.

c) The alignment of the genetic and physical maps of mouse chromosome two by the analysis of hybrid cells carrying the reciprocal translocations T(2;4)1Sn, T(2;4)13H, T(2;8)2Wa and T(2;16)28H.

Before we are able to consider these specific issues, we need to place this work in some perspective. To do this, we need to examine the role of the mouse in genetics; the methods used to generate genetic maps of the mouse genome; the general state of the map of mouse chromosome two and the characteristics of these mutants. In addition, strategies for cloning genes in mouse; the use of somatic cell hybrids and the characteristics of repetitive elements in the mouse genome will be considered.

1.2) The mouse in the field of genetics

The laboratory mouse has been one of the primary organisms used in the study of genetics for many decades. Fancier stocks were collected by people for their aesthetic value, and these collections and the breeding stocks established from them have provided a valuable resource for geneticists. The laboratory mouse has accumulated many mutations over the years, many of which have an effect on coat colour, and thus have been maintained in the fancier stocks. Constant inbreeding has helped to ensure that any new recessive mutation will

eventually appear in homozygous form in the stock. Mouse coat colour mutations are possibly the best understood examples, as these were among the first mutations to be actively selected for, by the mouse fancier community, since the early 19th century. Coat colour mutations in the mouse are usually correlated with other abnormalities, and thus can throw light on the cellular pathways, and the origins of different cell types throughout the development of the mouse (Jackson, 1991). Study of the dominant spotting (*W*) and steel (*Sl*) loci has elucidated the pathways leading to the migration and proliferation of pigment cells in the coat. Mice carrying these dominant mutations also manifest varying degrees of anaemia, showing the developmental connection between the pigment producing cells and the hematopoietic system. Similarly, the study of the piebald lethal (s^l) and lethal spotting (*ls*) loci has shown the connection between the cells responsible for pigmentation and those responsible for enervation of the murine colon, as both cell types originate from the neural crest.

On a more basic level, the mouse is an ideal organism for genetic research, because of its small size, its short generation times (a female mouse can produce 4 to 8 litters of 6 to 8 pups each in her lifespan, with a gestation period of only 19 to 20 days - Darling and Abbott, 1986), and the relatively small expense of maintaining mouse stocks. In the mouse, many individuals can be generated from one original set of parents, and specialised crosses can be set up to analyse more efficiently the inheritance and action of mutant genes.

Heterogeneity of the genetic background can cause problems in analysis in human populations, whereas in the mouse, the use of genetically homogeneous inbred strains can help to distinguish effects due to the genetic background from the genuine effects of the mutation. The phenotypes of many mouse mutations are affected by the genetic background. For example, in the case of the brachyury (*T*) mutation, the tail to body length ratio is dependent on the genetic background (Festing, 1979). Separate alleles of other mutations are in some cases indistinguishable on the same genetic background.

Tests for allelism can be carried out in the mouse by selected breeding experiments. Allelism between two mutations can be ruled out by the appearance of wild type progeny in a test cross between two double heterozygotes.

Discoveries made in the field of mouse genetics can elucidate the basis of mutation occurring in the human population. In addition, because of defined synteny relationships between mouse and human, genes mapped in the mouse can identify putative map positions in

the human genome. In some cases, the similarity in the effects of mutations of genes at a specific loci between mouse and man can identify possible mouse models of human disease. These models are important tools in the understanding of human disorders, and in the generation of new therapies. The use of the mouse provides easy access to any tissues which may be needed for analysis, at all stages of development, which is a considerable advantage. The increasing ease of homologous recombination (gene targeting) to generate new loss of function mutant alleles at specific loci can help to clarify the phenotypic effects of genes, allowing a greater understanding of the structure and function of the mammalian genome (Frohman and Martin, 1989). It is only a matter of time before mutations can be generated at any locus of interest, which will further increase our understanding of developmental processes in both mouse and man.

1.2.1) The origins of the modern mouse

The genus *Mus* is very complex (Bonhomme and Guenet, 1989) with over 60 forms of the basic *Mus* species *Mus musculus*.

The mouse is estimated to have diverged from other rodent populations, such as the rat (*Rattus* genera) between 8 and 12 million years ago, and from more closely related species such as the mouse-like *Nannomys*, *Coelomys* and *Pyromys* species approximately 5 to 6 million years ago. The modern mouse diverged from other mouse genera such as the Indian pygmy mice *Mus dunni* and *Mus booduga* between 1.5 and 3.2 million years ago (determined from taxonomy, morphology and more recently, DNA sequence analysis- Bonhomme and Guenet, 1989). Classification of the various *Mus* species has been greatly facilitated by biochemical analysis, such as isozyme typing, providing enough loci are typed.

The genus *Mus* can be broadly subdivided into five biochemical groups (Bonhomme *et al.*, 1984). These are :- *Mus musculus domesticus*, *Mus musculus musculus*, *Mus musculus molossinus* , *Mus musculus castaneus*, *Mus spretus*, *Mus spicilegus* (Southern), *Mus spicilegus* (Northern) and *Mus musculus bactrianus*. There are also the more diverged species *Mus caroli*, *Mus cervicolor* and *Mus dunni*. (For geographical distribution of species see figure 1).

Bonhomme *et al.* (1984) carried out extensive isozyme analysis to determine the relationships between the various *Mus* species and subspecies. The data were presented as a

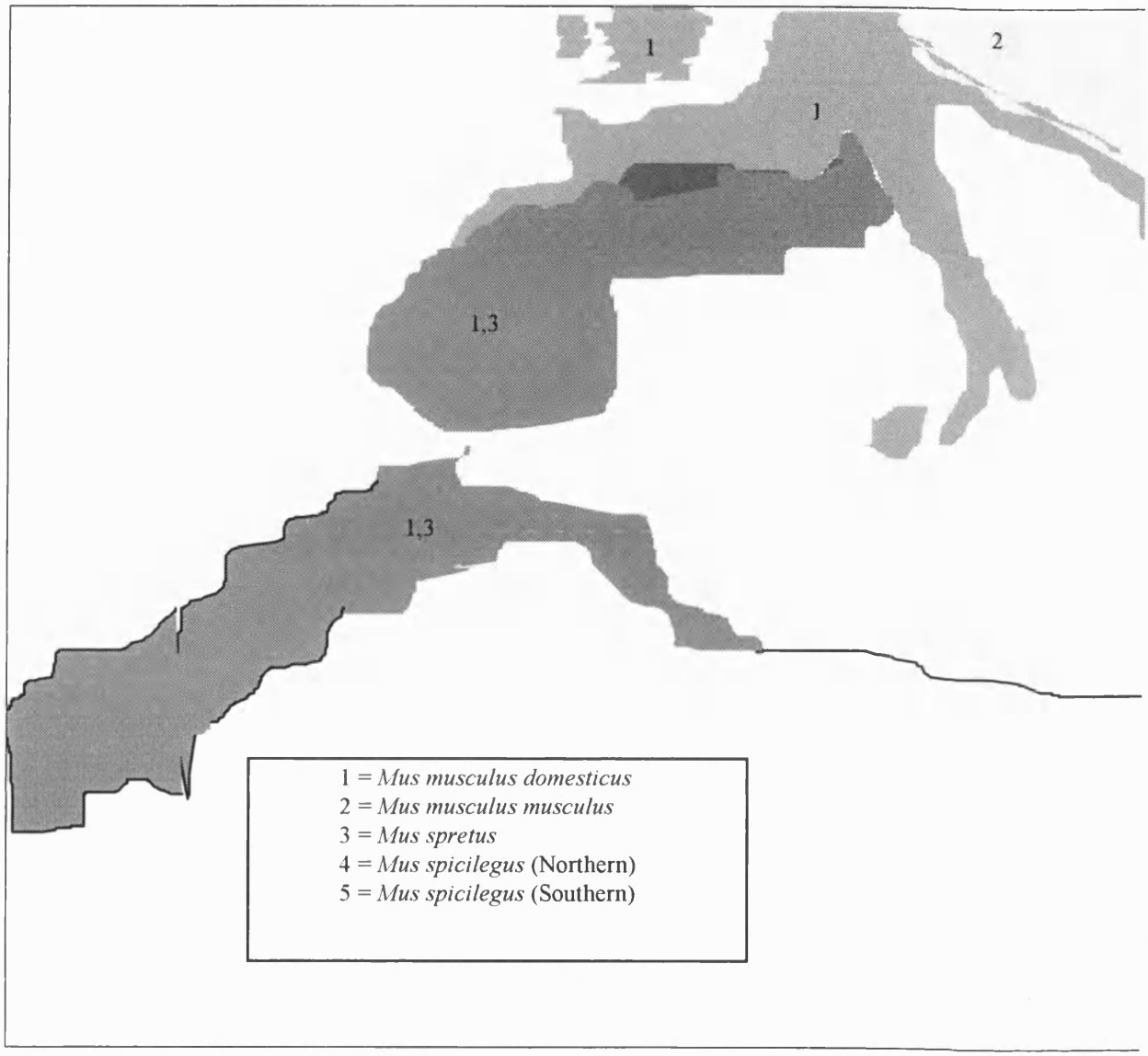
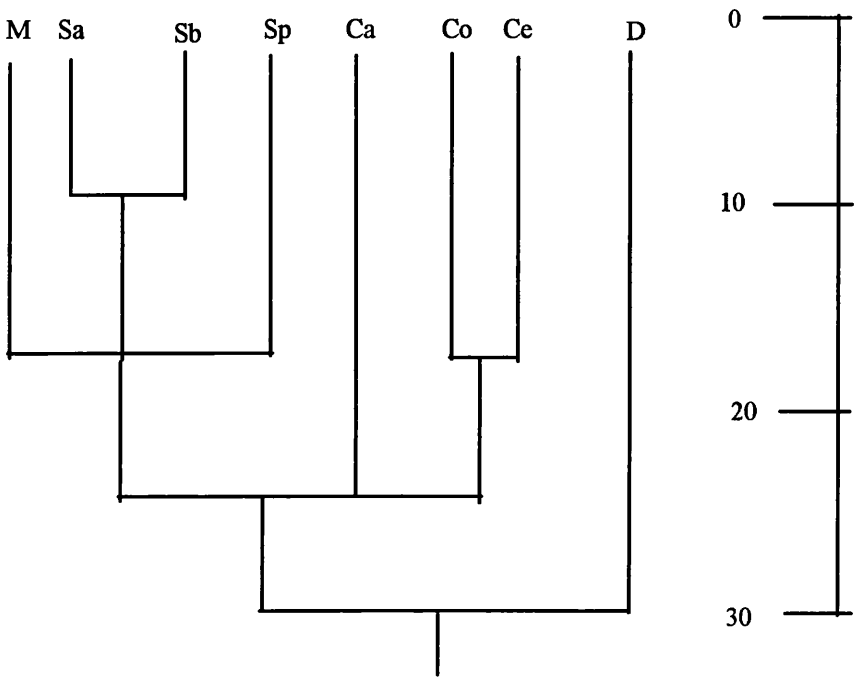


Figure 1: - The Geographical distribution of Mus species in Europe



M = *Mus musculus* complex species
Sa = *Mus spicilegus* (Northern)
Sb = *Mus spicilegus* (Southern)
Sp = *Mus spretus*
Ca = *Mus caroli*
Co = *Mus cookii*
Ce = *Mus cervicolor*
D = *Mus Dunni*

The scale refers to divergence times calculated by Nei's genetic index (Bonhomme et al., 1984)

Figure 2: - Phylogenetic tree of mouse species

phylogenetic tree, which is given in figure 2. These species are not isolated populations, gene flow between them does occur to varying degrees. For instance, *Mus musculus musculus* and *Mus musculus domesticus* interbreed along a narrow hybrid zone ranging from Denmark to Bulgaria (Bonhomme and Guenet, 1989). Similarly, *Mus musculus castaneus* interbreeds with *Mus musculus musculus* in Japan, and with *Mus bactrianus* in parts of China and Indonesia. *Mus musculus* species do not interbreed with *Mus spretus* in the wild, but can be induced to do so in the laboratory.

1.2.2) Inbred strains of mice

An inbred strain of mouse is one which has undergone over 20 generations of brother - sister mating, to produce a genetically homogeneous line. These inbred lines are predominantly descended from *Mus musculus domesticus*, but have traces of other *Mus* musculus species and even sequences descended from Asian mice such as the Japanese waltzing mouse (Bonhomme and Guenet, 1989). Ferris *et al.* (1982) presented data that suggest that many common inbred strains were derived from a single female, which may have lived as recently as 1920. This conclusion was reached by study of the sequences of mitochondrial (mt) DNA in the different inbred strains. Mitochondrial DNA is inherited in an predominantly maternal manner, and provides a good means of studying maternal lineages. The mtDNA of the inbred strains was found to be surprisingly similar, in the light of the many differences which are known to exist between them in nuclear DNA, indicating the common origin of many of the present inbred strains.

1.2.3) Mouse models of human disease

The similarities in phenotype between mouse mutations and certain human diseases has allowed the development of mouse models of human disease. The existence of these models is important for understanding of the mechanisms by which mutations cause their effects, and also in the development of new therapies for the human population. For a mutant to be classified as a mouse model for human disease, it must fulfil some preliminary requirements. These include phenotypic properties such as gross appearance, and age of onset, and genotypic properties such as mode of inheritance (Darling and Abbott, 1992). Certain mouse mutations have been proposed as models for human disorders on the basis of phenotypic similarities alone, such as the kidney disease locus (*kd*) on mouse chromosome 10. This has been proposed as a model for medullary cystic kidney disease in man (Darling and Abbott, 1992). On the

basis of known conserved synteny, this gene could be predicted to be in band 10q in man. It is not yet certain that the *kd* mutation in mouse maps within this group, although it does map close to it. Models are usually made on the basis of phenotypic similarities, and also from comparative mapping evidence, as was the case for small eye (*Sey*) which is a model for aniridia. Both disorders have been shown to be caused by changes in the *Pax6* gene (Glaser *et al.*, 1990, Hill *et al.*, 1991). Small eye is now known to be an exact homologue of aniridia, on the basis of mutations in the *Pax6* gene leading to clinically identical effects in both man and mouse. Models are not usually exact, there are usually differences in phenotype between mouse and human homologues. In some cases, mutations within specific genes appear to have different effects in mouse and man. For example, mutations in the hypoxanthine guanine phosphoribosyltransferase (HPRT) gene on the human X chromosome cause Lesch-Nyhan syndrome, characterised by compulsive self mutilation and spastic cerebral palsy, but mice with defective *Hprt* genes are asymptomatic in this respect (Finger *et al.*, 1988). Lesch-Nyhan patients also show defects in haematopoisis, and in testis development, these effects are again absent in the mouse (Ansell *et al.*, 1991). The differences in phenotype clearly point to differences in the purine and pyrimidine salvage pathways between man and mouse. Despite the differences, mouse models for human disease have a definite place in the future of disease prevention and treatment.

1.2.4) The use of mice carrying chromosomal rearrangements

In addition to the variety of mouse lines constructed by man, there exists another valuable resource for geneticists, which is found in nature. These are the mouse lines carrying chromosomal rearrangements. In recent times, the number of rearrangements has been added to by the production of deletions, inversions, insertions and translocations from mutagenized mice. The mutating agent is usually electromagnetic radiation (e.g. X-rays).

Translocations have many uses in mouse genetics. Initially, translocations were used to assign linkage groups to different chromosomes. For instance, the reciprocal translocation referred to at the time as T138Ca, involved linkage groups II and IX, and was known to involve chromosomes 9 and 17. A Robertsonian translocation Rb163H, involved linkage groups II and XII, and chromosomes 9 and 19. The common linkage group, II, must therefore be on chromosome 9. From this experiment, linkage groups IX and XII were assigned to chromosomes 17 and 19 respectively (Green, 1981). By these means, all linkage groups were assigned to chromosomes.

Translocations have also been of use in gene mapping. Snell (1945) mapped genes to chromosomes 2 and 4, by linking these genes to the T1Sn translocation breakpoints by virtue of male semi-sterility measured as a 50% decrease in litter size (the presence of a reciprocal translocation has no effect in balanced form in a carrier, but 50% of the gametes of such a mouse will be unbalanced, and therefore zygotes conceived with these gametes will not survive, hence the semi-sterility).

Translocations have uses in the study of imprinting. Cattanach (1986) examined the effects of uniparental disomy for mouse chromosome 2 by the use of mating schemes with mice carrying the T1Sn balanced translocation. These experiments have been repeated with the T2Wa breakpoint to narrow the region of imprinting down further (Cattanach *et al.*, 1992).

Translocations can be used to align the genetic and physical maps of chromosomes, They can also be used to separate two halves of a chromosome in somatic cell hybrids. Similarly, translocations involving chromosomes with selectable markers can allow the direct selection for chromosomal fragments which could not otherwise be selected for in cell culture (for instance, translocations involving the mouse chromosomes X or 11 can confer a selectable marker, hypoxanthine guanine phosphoribosyltransferase (HPRT) or thymidine kinase (TK) to the other translocated chromosome).

1.3) The use of the interspecific backcross in mouse genetics

Interspecific backcrosses are one of the most important resources in the field of mouse gene mapping. In other species, somatic and cytogenetic analysis provides a simple alternative, but in the mouse, identification of the chromosomes is more difficult. For this reason, genetic analysis in the mouse relies heavily on linkage experiments. In order for a gene to be mapped using linkage analysis by interspecific backcross, four conditions must be met. (1) Markers must be available that cross react between the two species to be used. (2) Linkage relationships must be maintained over considerable physical distances (the order of genes must be conserved between the species to a large degree). (3) Markers to be mapped must be variant between species to allow the differentiation of alleles from each species. (4) Knowledge of the evolutionary changes in chromosomal organisation which have occurred between the species must be available (Avner *et al.*, 1987). The requirement for these conditions to be met makes the use of species which are very diverged difficult, since there must also be some degree of fertility in the F1 offspring to allow the production of the N2. High resolution genetic mapping is dependent on the availability of a large number of DNA samples, and a large number of markers which are polymorphic between the parental species. Interspecific backcross analysis allows the construction of detailed linkage maps because of the potential availability of large numbers of backcross animals.

The basic form of the interspecific backcross is the mating of two individuals from closely related species to produce fertile F1 animals. The F1 animals are then mated back to one or other of the parental stocks (see figure 3). Recombination events are then analysed in the N2, and the gene order is determined by minimizing the number of double recombinants (it is generally taken that the presence of one chiasma reduces the probability of the occurrence of another in the same region of DNA by recombination interference). Each marker tested in the backcross has a specific segregation pattern. This, when compared to the segregation patterns for other markers (pedigree analysis) provides a localisation for the test locus, providing enough loci have previously been placed on the genetic map. In addition, linkage data from different backcrosses can be amalgamated to produce consensus linkage maps, as long as some loci are shared between them. This does have potential problems, as genetic distances between genes often vary between backcrosses. The mapping of a series of reference loci in each cross provides a partial solution to the alignment of the different maps.

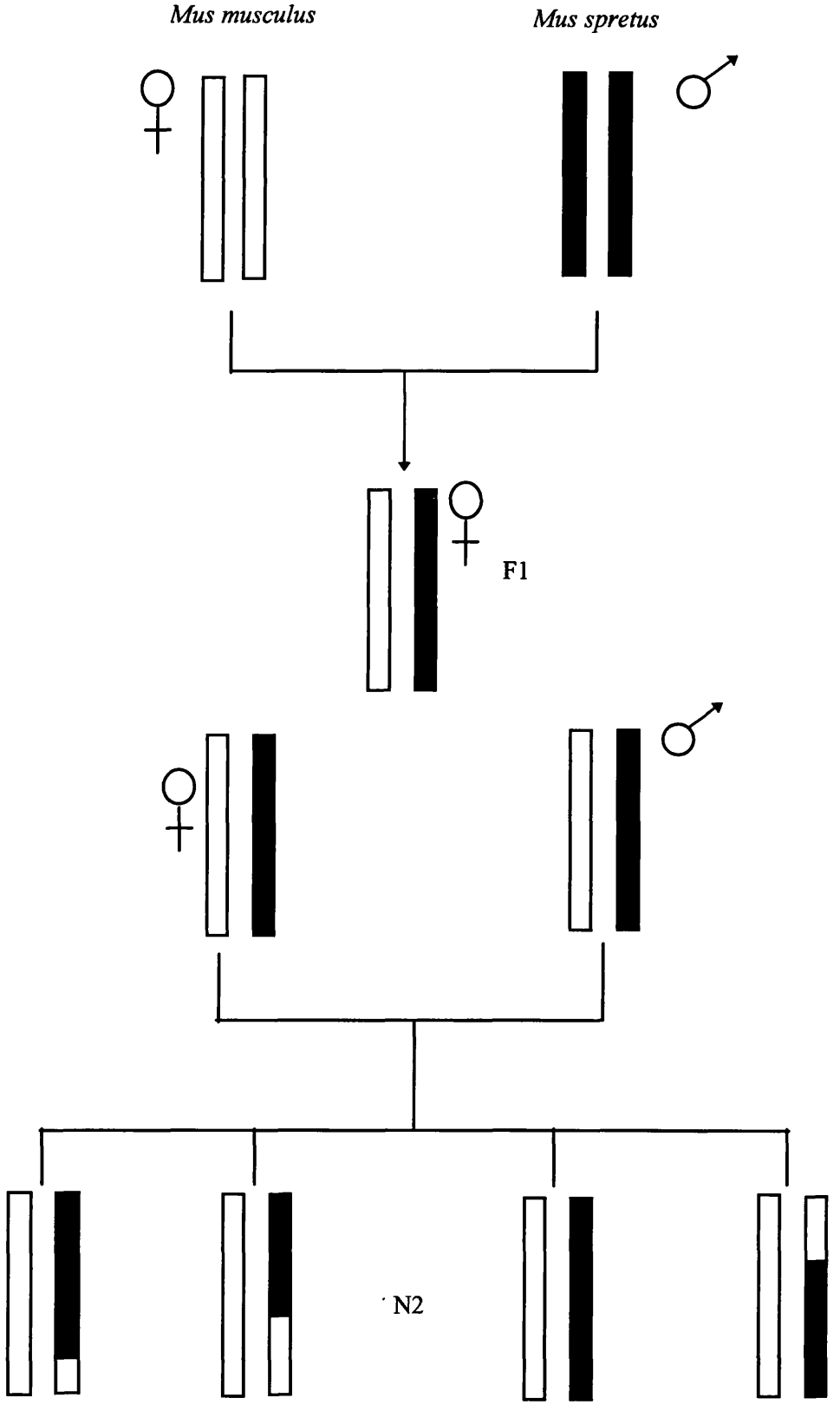


Figure 3: - The interspecific backcross

1.3.1) Choice of parental types for interspecific backcross.

The most important factors in the choice of parental species for interspecific backcross analysis are the ability of the species to produce fertile F1 animals (fertility of hybrids between mouse species and subspecies varies from complete fertility to complete sterility - Avner *et al.*, 1988) and the availability of variation between the species. The importance of each of these factors must be balanced. The use of species that are vastly diverged may provide an ample source of variation, but this is at the expense of the fertility of the F1 animals. There is evidence that sterility results from meiotic abnormality in the F1 (Hale *et al.*, 1993). *Mus spretus*, estimated to have diverged from *Mus musculus* approximately 3 million years ago (Bonhomme and Guenet, 1989), is the usual choice for the remaining parent in linkage studies of the laboratory mouse *Mus musculus* (for relatedness of mouse species, and geographical distribution see earlier section). *Mus spretus* has the advantage of having much sequence divergence from *Mus musculus*, whilst still being able to produce fertile hybrid female animals. Seldin *et al.* (1989) carried out comparisons of the genetic map of mouse chromosome 12 generated from intra and interspecific backcrosses, and concluded for chromosome 12 at least, that gene orders were identical for the two species. Additionally, no rearrangements of gene loci on mouse chromosome sixteen was noted between different species or strains in three interspecific, and five intersubspecific backcrosses, although sex-specific diferences in genetic distance were found to vary widely betwen the crosses (Reeves *et al.*, 1991). To date, the arrangement of all the *Mus musculus* and *Mus spretus* chromosomes has proved identical except for a small inversion on proximal chromosome 17, in the region of the *t* haplotype, in *Mus spretus*. This has also been confirmed cytogenetically. Any rearrangement not detected so far could only be on a scale below the level of cytogenetic resolution. The male F1 generation are sterile, but the females are fertile. This has the disadvantage that genetic distance can only be measured in the female map. This does not generally cause difficulties except in cases where genes in the pseudoautosomal region are under study. It is also likely that the *Mus spretus* genome might contain species specific recombinational hotspots, which would complicate genetic distance measurements. These are already known to exist in the genome of *Mus musculus* (Lyon, 1976). Another disadvantage lies in the difficulty in maintaining *Mus spretus* stocks. It is usual in *musculus / spretus* backcrosses to mate the F1 back to the *Mus musculus* parent since this is more efficient than the reciprocal cross. This does, however produce maps that are of the *Mus spretus* chromosome. Maps of the *Mus musculus* chromosome can be obtained by carrying out the reciprocal cross.

The use of other species or subspecies in interspecific backcrosses has some advantages over the use of *Mus spretus*. *Mus musculus castaneus* is being used increasingly in *Mus musculus* backcrosses, as this subspecies produces fertile animals of both sexes, whilst still retaining a reasonable amount of variation with inbred strains. Recently, backcrosses have been set up using the subspecies *Mus musculus molossinus* (Moriwaki *et al.*, 1993).

1.3.2) Gene - mutant correlation using interspecific backcrosses.

Mutations causing phenotypic effects in the mouse have been studied for many years. The cloning of the genes and the identification of the mutations therein, together with the comparison of the effects of the mutation with those produced from human disorders is an important step in the cloning and understanding of the underlying defect in human disease. The study of these disorders in the mouse allows a greater freedom to investigate possible therapies for the human population, as well as being inherently interesting.

An initial step to the understanding of any new mutant is to associate it with a particular chromosomal region. Once this has been acomplished, specific backcrosses can be set up, which segregate the mutation of interest (Copeland and Jenkins, 1991). The mice are then typed for the presence of the phenotypic effects of the mutation, and also for a selection of markers that also map to the region. The map position of the mutant gene can sometimes be refined by this method to such a point that possible candidate genes are identified, or that linkage of the mutant to closely associated flanking markers is established. Sometimes, close association between the mutant and a gene already mapped will allow identification of the primary genetic defect (see section 1.2.4). For instance, the jimpy (*jp*) mutation on the mouse X chromosome was shown to be tightly linked to the proteolipid binding protein (*Plp*) (Dautigny *et al.*, 1986). Subsequent studies showed that *jp* results from a mutation in the structural gene for *Plp* (Dautigny *et al.*, 1986).

If no candidate gene is evident, the use of interspecific backcrosses can identify closely linked flanking markers to provide a starting point for positional cloning experiments such as chromosome walking and jumping strategies. Possible problems with this approach often arise because of the reduction in fertility of some mutant genotypes. Also, in the case of true recessive mutants, heterozygotes cannot be identified without progeny testing. This can be

circumvented by the use of another very closely linked visible mutant in the backcross, in which case, heterozygotes will be marked, or by progeny testing.

1.3.3) Differences in recombination between the male and female genetic maps

There is considerable evidence that recombination rates differ between males and females in a given area of the genome. This has implications for interspecific backcross analysis, because the genetic distances will vary according to which parent the F1 were backcrossed - because of the sterility of *Mus musculus / Mus spretus* F1 hybrid males, genetic distance can only be measured in female meiosis.

It is generally taken that recombination is greatest in female meioses. Michie's studies on the linkage between vestigial tail and rex on chromosome 11 showed sex specific differences in recombination frequency that were statistically significant (Michie, 1954) with recombination being greatest in females. Subsequent studies (Dunn and Bennett, 1967) suggested that although overall recombination rates were greater in the female, chromosome specific variation was observed. For instance, in the case of most mouse chromosomes, female recombination rates are seen to be higher than male recombination rates, whereas in the case of chromosome 15, recombination rates were seen to be higher in the male. It is interesting to note that not all of the loci tested on the same chromosome showed this effect to the same degree. Evidence for the excess of female recombination is given by the observation that the average number of chiasmata is higher in females than in males - 2.8 chiasmata per bivalent in females, as compared to 2.4 chiasmata per bivalent in males (Crew and Koller, 1932).

1.3.4) Transmission distortion in interspecific backcrosses.

Transmission distortion refers to non-Mendelian segregation in *Mus musculus / Mus spretus* interspecific backcrosses due to preferential inheritance of either the musculus or spretus chromosomes or chromosomal subregions. The predicted frequency of parental alleles in backcross experiments by Mendelian segregation is 1:1. For certain chromosomes this is not the case. Transmission distortion occurs to differing degrees on different chromosomes in different backcrosses. Chromosomes 2, 4 and 10 and X have been associated with transmission distortion in various backcrosses (Siracusa *et al.*, 1989, 1992; Rowe *et al.*, 1994).

Siracusa *et al.* (1989) examined the transmission distortion (TD) on mouse chromosome two. The proportion of alleles transmitted from each parent was investigated at each locus in a (C57BL/6J X *M. spretus*) F1 X C57BL/6J cross. It was determined if the transmission differed significantly from the expected 1:1 ratio by chi squared test. It was concluded that significant differences from Mendelian ratios were found for each locus studied on mouse chromosome two with the exception of *Pck1*, the most distal locus studied. It was estimated that the distortion effect spans up to 20cM proximal to the β2 microglobulin (*β2M*) locus. The TD effect was examined further by Siracusa *et al.* (1991). The regions of mouse chromosome two that exhibited TD were refined, and the possibility of sex specific differences in TD was examined. In female N2 mice, loci from *Vim* to *Ada* were shown to exhibit TD, with the maximum distortion in favour of the Mus spretus allele being present between *Hox4.1* and *Fshb*. In N2 males, there appear to be two regions that show maximum levels of TD - the telomeric portion of the chromosome, in the region of Pck1, showed a preferential inheritance of *Mus spretus alleles*, whilst the centromeric portion of the chromosome, in the region of *Vim*, showed a preferential inheritance of the C57BL/6J allele. From this, it could be suggested that specific genes could be influencing the inheritance of parental chromosomes. However, data produced on larger numbers of animals using the European collaborative interspecific backcross (EUCIB) do not support these findings (A. Pilz and C. Abbott, personal communication). The EUCIB panels are of especial use not only in the high resolution mapping of loci, but in the identification of transmission distortion effects, because of the large sample size of the backcross (the European backcross collaborative group, 1994).

The biological causes of transmission distortion are yet to be determined. They will differ for different regions. It may be due to a) the differential production or viability of oocytes, b) differences in the abilities of sperm carrying different parental alleles to fertilise the oocyte, c) differential survival of N2 animals in utero or d) differential survival of N2 animals after birth. It is most likely that the genotype of an embryo could effect its survival due to allelic interactions at specific loci or groups of loci.

1.4) Overview of mouse chromosome two

Standard metaphase mouse chromosomes immobilised on slides and examined under a microscope allow the chromosome content of the cell to be analysed. Mouse cells contain a total of nineteen autosomes and the two sex chromosomes. All the mouse chromosomes are acrocentric, with the exception of the Y, which has a small short arm. They can be distinguished from each other on the basis of size and chromosome banding after modified Giemsa staining, although this is more difficult than in human cells because the mouse chromosomes are all fairly similar in size, and the banding patterns are more difficult to analyse. Until the advent of this form of chromosome banding carried out by Schnedl (1971), the individual mouse chromosomes could not be identified.

Mouse chromosome two is the second largest chromosome in the mouse karyotype, with a genetic length of 110cM (Green *et al.*, 1972). The chromosome has also been estimated to be 6.88% of the total female haploid length (Evans, 1989) taken from an average of the length of the chromosomes in ten mitotic cells from the liver of 12 day embryos. Mouse chromosome two has a characteristic banding pattern (see figure 4) in which the proximal region of the chromosome has two heavily staining bands at positions C1.1 and C1.3. Specialised staining techniques designed to reveal early and late replicating DNA reveal a different and more variable band pattern, since this method depends upon the timing of DNA replication in the individual cell.

A consensus linkage map of the loci mapped to mouse chromosome two has been put forward (Siracusa and Abbott 1993). This map is continually being added to as new loci are mapped. The consensus map (taken from Siracusa and Abbott, 1993), is pictured in figure 5. The map was created by integration of multipoint backcross analysis. The main body of the map was taken from the results of twelve backcrosses involving over 170 loci, with the remaining loci being mapped by other smaller backcrosses, two point crosses, cytogenetic analysis or by analysis of recombinant inbred strains. Mouse chromosome two is involved in several reciprocal translocations with other chromosomes, some of these are illustrated in figure 4.

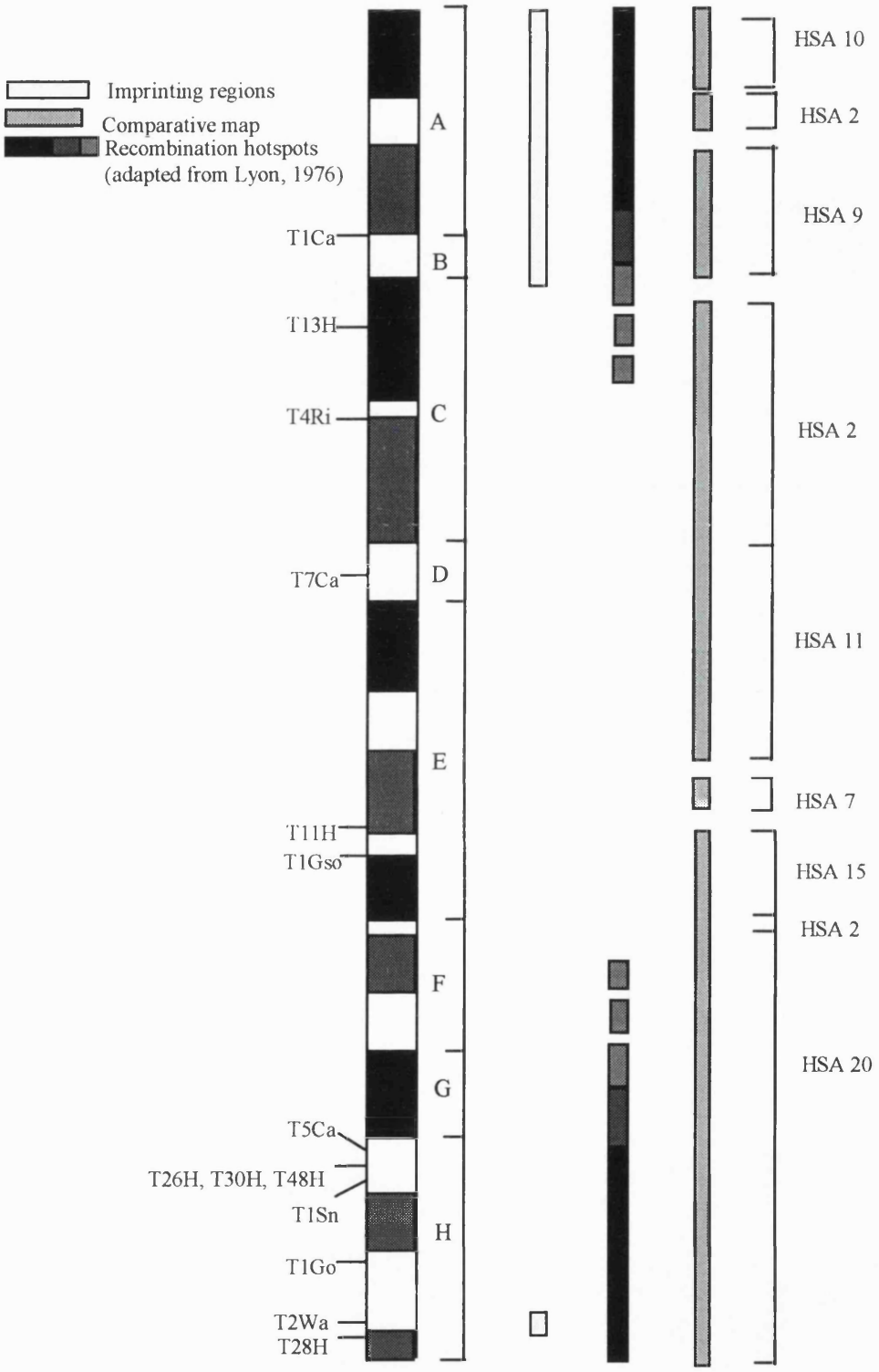


Figure 4: - Map of mouse chromosome two, showing recombination hotspots, comparative map and imprinting regions

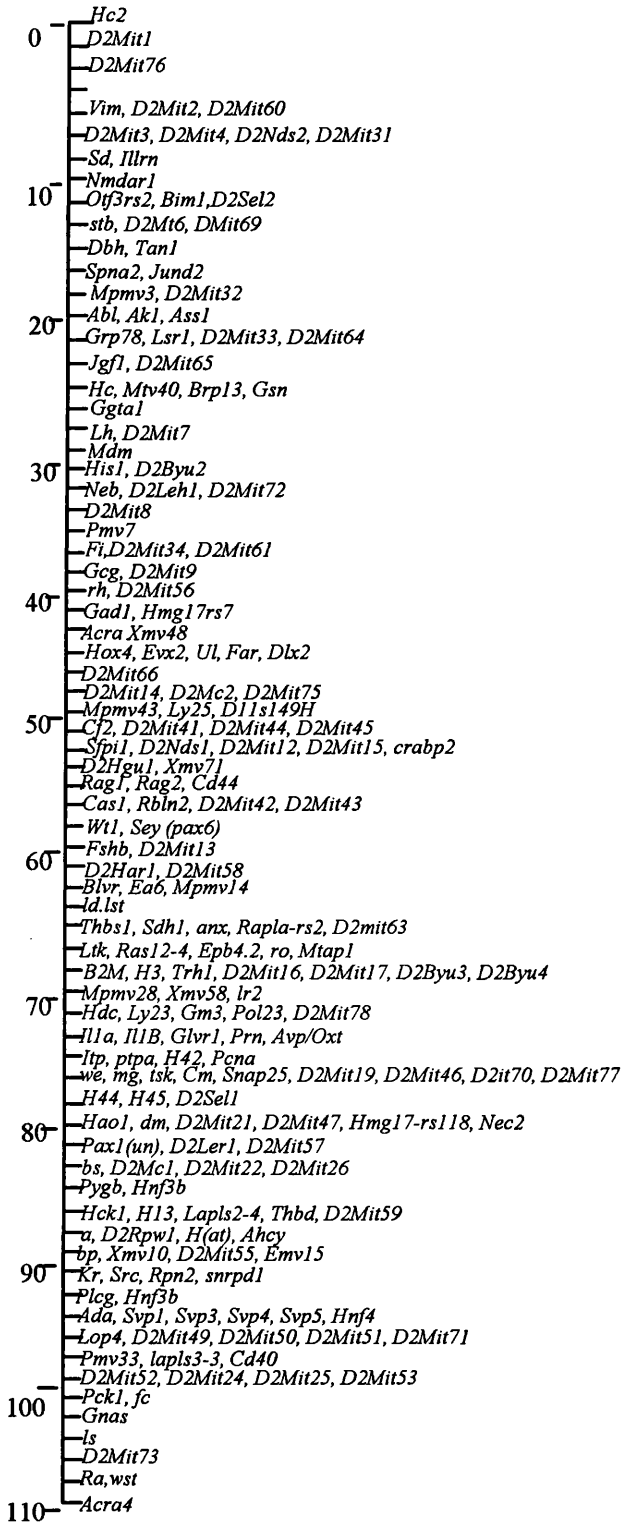


Figure 5: - Consensus linkage map of mouse chromosome two

1.4.1) Recombination hotspots on mouse chromosome two

There is considerable evidence to suggest that the pattern of recombination on certain mouse chromosomes is not uniform (see figure 4). Lyon (1976) carried out analysis of the distribution of genes on all the mouse chromosomes. If the frequency of recombination was equal over the length of the chromosome, the distribution of genes mapped by linkage analysis would be expected to be evenly spread. Lyon divided each mouse chromosome into sections roughly corresponding to 5cM intervals, and counted the number of genes assigned to each segment. Poisson analysis was then carried out to determine if any differences seen in the distribution of genes was statistically significant. On some mouse chromosomes, recombination events are more or less equally positioned along the chromosome, as reflected by the random distribution of genes. This is true for mouse chromosomes 1 and 9. Other chromosomes show altered recombination in some areas. These include mouse chromosomes 2, 4, 5, 6, 7,10, 11, 15, 17 and the X. Mouse chromosome two, for example, shows clustering effects located in the central portion of the chromosome, with genes widely spaced at both the telomeric and centromeric regions. Lyon (1976) stated that the small number of genes at the time of the analysis may cause discrepancies in the results, but this clustering effect is still apparent on mouse chromosome two with many loci now placed on the genetic map. This reflects possible recombination hotspots in these areas, suggesting that although gene orders calculated from the linkage data in these regions are correct, the physical distances are likely to be considerably smaller than calculated. Lyon (1976) suggested that these hotspots might be in part due to the preferential formation of chiasma in specific regions of the chromosome, such as preferred initiation of chiasma formation at the telomeres and centromeres of mouse chromosome two.

These findings were confirmed by Searle *et al.* (1979). Analysis was carried out of ten reciprocal translocation points on mouse chromosome two that were equally distributed along the chromosome (T24H, T11H, T30H, T26H, T5Ca, T1Sn, T28H, T1Ca, T13H and T7Ca). G-banding allowed more precise localisation of the translocation breakpoints. The study showed that although the order of the translocation breakpoints was conserved along the chromosome between linkage analysis and cytogenetic analysis, the genetic distances given by the two methods varied considerably. For instance, T7Ca, T24H and T11H map close together on the genetic map, but cytogenetic analysis shows that they are separated by a considerable distance. T1Sn and T28H are both located in band H on the cytogenetic map, but are separated by an estimated distance of 25 map units on the linkage map. Lawrie and Hulten (1992)

produced chiasma density and interference maps of mouse chromosome two for both male and female genomes. Chiasma positions were measured directly from spermatocytes and oocytes. Data were presented as vertical lines (representing individual chiasma) on a horizontal axis (representing mouse chromosome two). Chiasma clustering was taken to indicate recombination hotspots. Sex specific differences were noted, with clustering of chiasma in protelomeric regions in spermatocytes, and more proximal clustering in oocytes, with an absence of chiasma on the short arm of chromosome two. Estimates of genetic interference distances between chiasma were also given (Lawrie and Hulten, 1992).

These findings have important implications for cloning projects on mouse chromosome two. The ultimate goal of my research project was to add to the number of markers for mouse chromosome two, in the region of the ragged and wasted genes located in the region of the putative distal recombination hotspot.

1.4.2) The imprinting region on mouse chromosome two

It has been widely reported that contributions from both parental genomes are required for development of the embryos. In humans, hydatidiform moles result from the development of an androgenic pregnancy. Pregnancies resulting from the fusion of two maternal genomes and one paternal genome result in reasonable development of the embryo, but poor development of the extra embryonic tissues, whereas in the converse situation, foetal tissue is lacking, but the extra-embryonic tissues are well developed (Kajii and Ohama, 1977; Jacobs *et al.*, 1980).

Similar results have been seen in the mouse. Solter (1988) carried out experiments involving the artificial construction of androgenones and gynogenones. Similar results were seen to those observed in humans. Barton *et al.* (1984) carried out experiments using pronuclear transplantation, and analysed the resulting 'progeny' for embryonic and extra-embryonic development, confirming that both parental genomes were required for proper development of both the foetus and its associated tissues.

Many reports have suggested that differential methylation is the underlying causes of this effect. Reik *et al.* (1987) carried out a study which proved that a transgene inserted into the murine immunoglobulin *IgH* gene was differentially methylated according to the sex of the mouse through which it had been transmitted. Transmission from a female resulted in a heavily

methylated transgene, revealed by restriction digests with *Hpa II*, and the methylation sensitive restriction enzyme *MspI*. Inheritance through the male line resulted in hypomethylated DNA.

In the case of certain chromosomes, in both mouse and man, there is a requirement for the presence of genetic material originating from both parental genomes. This is known as imprinting. Several mouse chromosomes appear to have no imprinted regions, as defined by gross phenotypic effects, including mouse chromosomes 1, 4, 5, 9, 13, 14, and 15 (Lyon *et al.*, 1976; Cattanach and Kirk, 1985). This is based on uniparental disomy experiments which illustrate the requirement of whole chromosomes or chromosomal regions for correct development of the embryo. These experiments identify imprinting effects of genes or groups of genes that have a profound effect on the phenotype. Maternal disomy for chromosome 6 is lethal (Cattanach and Kirk, 1985). Mouse chromosome two is seen to have two regions of imprinting, at either end of the chromosome - see figure 4 (Cattanach and Kirk, 1985). Proximal mouse chromosome two (above the T13H breakpoint) shows imprinting for maternal duplication only. Distal mouse chromosome two shows imprinting effects for both maternal duplication/paternal deficiency and for paternal duplication/maternal deficiency (Cattanach and Kirk, 1985).

Cattanach and Kirk (1985) carried out a study in which mice heterozygous for the translocation T(2;11)30H were mated. Imprinting effects for chromosome 11 were noted, as was the case for distal mouse chromosome two. Mice with maternal duplications of distal mouse chromosome two have long, flat sided bodies, and are hypokinetic. These mice become almost totally immobile within a couple of hours of birth. Few of these mice survive beyond 24h. Paternal duplications produce the opposite phenotype. The mice have short square bodies with broad backs, and are hyperactive. The mice survive normally for several days, but fail to thrive.

Analysis of imprinting based on other translocation breakpoints has localised the region of imprinting to between the T1Sn and the T28H translocation breakpoints. Recent work has narrowed the imprinting region to between the T2Wa and T28H breakpoints (Cattanach *et al.*, 1992). The classic experiments carried out by Cattanach (1986) which determined that the proximal boundary of the imprinted region was the T1Sn breakpoint were repeated with respect to the T2Wa breakpoint. Peters *et al.* recently mapped several genes on distal mouse chromosome two relative to the T1Go, T2Wa and T28H translocation breakpoints, placing these translocations on the linkage map for the first time. This work

suggests that *Ada* lies distal to the T(2;8)2Wa translocation breakpoint, and is the most proximal gene in the distal imprinting region. The ragged, wasted and lethal spotting genes also lie in the distal imprinting region, although none show any evidence of imprinting (Peters *et al.*, 1994).

1.4.3) Comparative mapping of mouse chromosome two

Two or more genes which are located together in both mouse and man, regardless of gene order, are said to make up a region of conserved synteny. Genes which are present on equivalent chromosomes with the gene order also conserved are said to represent a conserved linkage. There are conditions which must be met before a gene in one species is considered to be homologous to a gene in another species. On the level of the gene product, these include molecular structure, biological function, subcellular location, substrate specificity or response to specific inhibitors (O'Brien, 1991a). There are many reasons for the study of comparative mapping. A gene which maps in the centre of a conserved linkage in the mouse can be placed with some confidence in the equivalent position in man, and vice versa. This has important implications for the mapping of disease genes. The map positions of mouse mutants affecting function or development can be used to predict the location of the gene in humans, for instance, the splotch (*Sp*) mutation on mouse chromosome one was found to be the murine homologue of Waardenburg syndrome type I (WS-1) (Chalepakis *et al.*, 1993). The co-segregation of *Sp* with the *Pax3* gene in the mouse (Skow *et al.*, 1988; Walther *et al.*, 1991) together with the correlation between the expression pattern of *Pax3* and the tissues affected by the *Sp* mutation proved that *Sp* was due to a mutation in *Pax3* (Epstein *et al.*, 1991a & b, 1993; Chalepakis *et al.*, 1994). Waardenburg syndrome type I was mapped to HSA 2q37, which shares conservation of synteny with proximal mouse chromosome 1 (Foy *et al.*, 1990). The similarities in clinical manifestations of WS-1 and *Sp* indicated that *Sp* was the murine counterpart of WS-1. Subsequent studies (Tassabehji *et al.*, 1992; Baldwin *et al.*, 1992) revealed mutations in the PAX 3 gene of WS-1 patients. The location of disease genes in man can be used to predict the map position of genes which may be potential models for human disorders. For instance, murine DNA markers corresponding to markers in the region of Down syndrome in man proved to be located in a region of mouse chromosome 16 syntenic with human chromosome 21q22. (Cheng *et al.*, 1988, Macdonald *et al.* 1988). The saturation of the genetic maps in both species with conserved markers would allow the prediction of the location of any gene mapped in one species in the other. The study of comparative mapping can also

cast light on the evolutionary relationships between the species, and the physical changes which may have occurred at the divergence of the species.

Mouse chromosome two has conserved regions with human chromosomes 2, 7, 9, 10, 11, 15 and 20 (O'Brien *et al.*, 1991). Figure 4 gives the arrangement of these conserved segments. The region of chromosome two from band A3 to B is homologous to HSA 9q33 to 9q34. Similarly, there is another conserved linkage corresponding to 2q14 to 2q33 located in bands C1 to D in the mouse. There is also a large conserved syntenic group on distal mouse chromosome two (bands F2 to H4) corresponding to 20p13 to 20q13.2. Pilz *et al.* (1992) placed the homologues of the *Pck* and *Acra4* genes on the human genetic map by somatic cell hybrid analysis. As *Acra4* is the most distal molecular marker placed on the map of mouse chromosome two, this extended the known conserved linkage of distal mouse chromosome two with HSA 20 to within 5cM of the telomere. It is highly likely that the whole of HSA 20 is represented in distal mouse chromosome two. The gene order appears to be conserved. It can therefore be postulated that the homologue of any gene mapping to HSA 20 will be located on distal mouse chromosome two.

1.5) Mouse mutant genes located on chromosome two

Mouse chromosome two harbours some interesting developmental mutants. (Green, 1989) Their chromosomal locations of these mutants are given in figure 6 (Siracusa and Abbott, 1993).

1.5.1) The ragged mutation

There are two alleles at the ragged locus, ragged (*Ra*) and ragged opossum (Ra^{op}).

1.5.1a) Ragged (*Ra*)

The first ragged mouse was described by Carter and Phillips (1952). This female mouse was derived from a stock that had been X-irradiated. Her father had undergone either a spontaneous mutation in one of his spermatogonia, or possessed a mosaic germ line on account of the irradiation. This female was mated to a wild type mouse and bore 47 young, of which 24 were ragged in phenotype. Ragged was shown to be a semidominant mutation (in cases of true dominance, heterozygotes are indistinguishable from homozygotes).

Ragged Heterozygotes (*Ra*/+)

Ragged heterozygotes are characteristic in appearance. The coat appears thinner than normal. The heterozygotes can be distinguished from normal littermates at a few days after birth because the vibrissae are noticeably shorter than normal. As development proceeds, the colour of these mice is considerably darker than normal littermates. By nine days of age, the coat of ragged heterozygotes is noticeably different from that of normal littermates. The altered texture of the coat is due to the absence of certain hair types. Mann (1963) carried out detailed analysis of the coat of ragged mice. Samples were taken at timed intervals both prenatally and after birth, and sectioned for microscopic analysis. The characteristic 'ragged' appearance of the coat was due to the presence of many guard hairs, at a higher density than in wild type mice in some areas. These tended to stick out from the skin, but an absence of auchenes was also noted, along with very few zigzags (undercoat), which tended to grow in clumps, mostly on the shoulders.

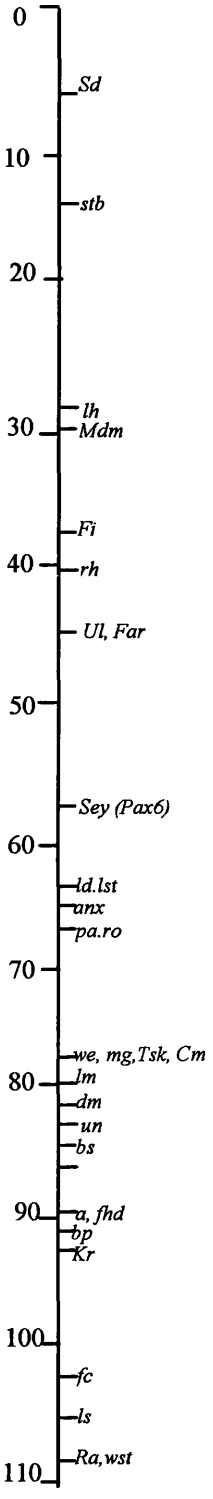


Figure 6: - Mutant genes on mouse chromosome two

The distribution of the other hairs in the coat is also patchy. Microscopic analysis of skin samples showed that the coat of ragged heterozygous individuals developed more slowly than that of normal mice. Auchene and zigzag follicles were seen to be present at birth, the lack of these in the adult coat was seen to be due to a failure of these follicles to differentiate (Mann, 1963). Tail rings are present, and the claws and eyelids are normal. Ragged mice are also consistently smaller than their normal littermates, but only slightly. Oedema is also seen in a proportion of heterozygous ragged mice. Crosses of +/+ X *Ra*/+ mice show that there is no decrease in viability associated with ragged in heterozygous form, as there is no statisticaly significant deviation from the expected ratio of ragged heterozygous mice to normal mice (Mann, 1963).

Ragged homozygotes (*Ra*/*Ra*)

The ragged gene has more profound effects in homozygous form. These mice are almost completely naked, with an absence of all types of hairs in the coat. Any hairs that are present are invariably abnormal. The follicles were still present, but failed to develop (Mann, 1963). As is the case for heterozygotes, the few hairs that develop are retarded in development compared to wild type mice. Ragged homozygotes also have reduced numbers of vibrissae (whiskers) in certain regions. Some of these mice also show generalised oedema, this is more pronounced than in the heterozygotes (Mann, 1963). It is possible that the oedema is due to imperfect action of the lymphatic system, endocrine abnormality or poor circulation. These mice do not survive to adulthood in most cases. It was shown by mating of *Ra*/+ X *Ra*/+ mice that ragged homozygotes have a reduction of viability of around 40% compared to normal mice, 20.4% of these prenatally. The ears and tails of ragged homozygotes are normally pigmented, and the claws are normal. A percentage of ragged heterozygotes and homozygotes show chylous ascites (the occurrence of a milky white fluid in the pleural and peritoneal cavities) in neonatal mice. Wallace (1972) carried out analysis of the genetic control of this in ragged mice. This condition is characterised by the occurrence of a milky white fluid in the peritoneal and pleural cavities of neonatal mice. Investigation of this has shown it to be due to inefficient drainage of the area by the lymphatic system. The initial study by Herbertson and Wallace (1964) showed chylous ascites to be present in 18% of ragged mice. Repeated backcrosses of these mice to wild type individuals resulted in a reduction in the incidence of chylous ascites so that by the fifth generation it was not observed. This suggests that the expression of chylous ascites is controlled by other modifying genes in the genome. Crosses back to a stock selected for the high incidence of chylous ascites showed a recurrence. Stocks

were also set up segregating five markers mapping to chromosome two (*Ra*, a^t, *we*, *fi* and *Sd*). The occurrence of chylous ascites in these stocks was analysed. Ragged was consistently associated with chylous ascites. Mice with the fidget (*fi*) mutation showed a higher incidence of chylous ascites. No effect on the occurrence of chylous ascites was seen from the agouti (*a*) or Danforth's short tail (*Sd*) loci. Wellhaarig (*we*) was shown to have an inhibiting effect. It was concluded that chylous ascites was a pleiotropic effect of the ragged mutation, which was enhanced or inhibited by other modifier genes, two of which were located on chromosome two. It was subsequently shown that another locus on chromosome one was also involved (Wallace, 1979).

1.5.1b) Ragged opossum (Ra^{op})

Green and Mann (1961) described a mutation which appeared in a litter of mice produced from a C57BL/6J X DBA/2J cross, at the Jackson laboratory, in 1957. This mutant was similar to the ragged mutation that had previously been described by Carter and Phillips (1954). Subsequent breeding tests proved that this mutation was semi-dominant. No study of this mutation in homozygous form has been done, since all these mice die in utero, thus all phenotypic data refers to the heterozygote.

Ragged opossum heterozygotes (Ra^{op}/+)

The initial observation of ragged opossum mice is that the coat is sparser than in wild type mice, the numbers of vibrissae and body size are also reduced. The reduction in coat thickness is due to an absence of many components of the coat, such as auchenes and zigzags. Some awls are present, and the number of guard hairs does not appear to be affected. Many ragged opossum mice are dead at birth. The female survivors are poor breeders but it is thought that the males may be normally fertile, as there is no apparent decrease in the litter size, or the frequency of litters of wild type females mated to heterozygous ragged opossum males (Green and Mann, 1961). There is a high occurrence of oedema. The skin of the oedematous mice has a stretched, glossy appearance. The onset of pigmentation is delayed in these mice because of the delay in the growth of the hairs of the coat. Ragged opossum mice also show occurrence of chylous ascites. Allelism was determined for ragged and ragged opossum by complementation analysis, since the compound heterozygote was seen to be lethal. and also by virtue of the very close linkage between the loci. Mann (1963) carried out studies to examine the effect of the mutation on the coat and lethality of the mutant, and to confirm the

allelism of ragged and ragged opossum. Sections of skin taken from ragged opossum mice at timed intervals during development show that although the pattern of hair development is the same as in wild type mice, the development of the follicles is retarded or arrested. There is a lack of vibrissae, and the only hair type that are fully represented are the guard hairs (tylotrichs). Ragged opossum was shown to have a more profound effect on the phenotype than ragged. *Ra*/+ X *Ra*/+ matings resulted in a reduction in viability of 20.4% (attributed to the homozygous progeny). *Ra*/+ X Ra^{op}/+ matings show a reduction of 32.5% and Ra^{op}/+ X Ra^{op}/+ matings a reduction of 49.6%, indicating that the heterozygotes also show a decrease in viability. Practically all the *Ra*/Ra^{op} mice, and all of the Ra^{op}/Ra^{op} mice die. Mann (1963) suggested that the similarity of the phenotypes of *Ra* and Ra^{op}, together with the fact that the compound heterozygote was inviable was further evidence of their likely allelism.

1.5.2) The wasted mutation (*wst*)

Wasted is a recessive mutation which arose spontaneously in a stock of HRS/J mice at the Jackson laboratory in 1972. The homozygotes can be recognised at 20 days of age by tremor and uncoordinated body movements. These mice develop progressive paralysis, and almost invariably die before 30 days of age. The *wst/wst* homozygotes are short lived, so the stock was crossed to a segregating stock of C3HeB/FeJ X C57BL/6J which improved the viability, by hybrid vigour (Shultz *et al.*, 1982). Subsequent studies showed the wasted homozygotes do not increase in body weight after 20 days (Lutsup and Rodriguez, 1989). Dissection of the animals reveals a marked lymphoid hypoplasia with significant reductions in the size of the spleen, lymph nodes and thymus. Degeneration of cells in the brain and spinal cord was also noted. Degeneration of Purkinje cells, with focal demyelination, in the cerebellar cortex have been seen to be degenerated in some cases, although this is disputed. Lutsep and Rodriguez (1989) discovered specific degeneration of cells in the anterior horn of the spinal cord, and also some abnormal vacuolated nerve cells. Also, fibrillary neurones (these neurones contain distinct 'thread-like bands' which interconnect to form networks within the cell) were more common in wasted mice than in wild type controls. Volume measurements of all types of neurones, both normal and abnormal were greater in wasted mice. Immunological staining of neurones from wst/wst mice with an antibody specific to proteins of the intermediate fibres in the cells revealed that the neurofilaments were highly phosphorylated (as the staining was seen to decrease following dephosphorylation). Shultz *et al.* (1982) put wasted forward as a model of the human condition ataxia telangiectasia. The mutant does share some similarities with the disease, such as the nerve degeneration and the ataxia. Experiments were carried out to

determine if there was any significant radiation sensitivity. A marked increase in spontaneous and γ-ray induced chromosomal abnormalities was noted, but no increased sensitivity to ultraviolet light. There was also no post-irradiation inhibition of DNA synthesis. These findings were also confirmed by several subsequent studies (Beechy and Searle, 1984, Inoue *et al.*, 1985). This therefore rules out wasted as a precise model for ataxia telangiectasia. The wasted mutation was also seen to lead to immunological abnormalities (Shultz *et al.*, 1982). Sections were taken from the gut of wasted homozygotes at various stages of the disease, and used for immunofluorescence experiments. Treatment of the sections with a goat anti-mouse antibody against IgA revealed that there were no plasma cells in the gut lining which were dedicated to IgA. No differences in gut morphology between wasted homozygotes and normal mice were seen. IgA plasma cells were totally absent from the entire small and large intestines. Further experiments proved that the absence of IgA was due to a failure of the cells to differentiate rather than a failure of the mature plasma cells to manufacture IgA. The number of IgA plasma cells was seen to decrease progressively from day 23 to day 27, until by 30 days no dedicated IgA cells were seen (Kaiserlian *et al.*, 1985). Interestingly, studies by the same group showed no deficiency of IgA in the serum of wasted homozygous mice (Kaiserlian *et al.*, 1985). Studies on the relative number of lymphoid cells in wasted animals showed a marked decrease in the numbers of all lymphoid cells, but the relative ratios of cells remained the same as in wild type animals. These immunological abnormalities are again in contrast to those observed in ataxia telangiectasia (Kaiserlian *et al.*, 1986). The relative numbers of B and T cells were conserved in wasted mice, in contrast to the situation in AT patients, where a specific depletion of T cells is noted. Other symptoms associated with AT in humans, such as elevated alpha-fetoprotein (AFP) levels and decrease in the circulating levels of the hormone thymulin were not seen in wasted mice (Kaiserlian *et al.*, 1986). Wasted mice have also been shown to be deficient in interleukin 2 (Kaiserlian *et al.*, 1986).

Wasted has also been put forward as a model of severe combined immune deficiency (SCID). This condition, in humans, is caused by an absence of adenosine deaminase (ADA) activity. Abbott *et al.* (1986) showed a decrease of ADA activity in wasted homozygotes with intermediate levels in heterozygotes. This was contradicted by Geiger and Nagy (1986) who found no deficiency of ADA in wasted homozygotes. Indeed, an increase in ADA activity was noted in whole brain, cerebellum and spleen. These differences could be attributed to the different methods of measurement of ADA activity, or the different genetic background of the mice used in each study, or by the fact that +/wst animals were used as controls in the

experiments of Geiger and Nagy (1986). These results were confirmed by Abbott *et al.* (1991) upon the finding that *Ada* mapped to band 2H3 on chromosome two by cytogenetic analysis. Linkage analysis showed that *Ada* mapped 13.8 +/- 2.7 cM proximal to wasted. Breeding tests were carried out, which ruled out allelism as 6 recombinants were found between *Ada* and *wst*.

Wasted does share similarities with other neurological disorders, such as motor neurone disease, and may be shown to be an animal model for this in the future (Lutsep and Rodriguez, 1989).

1.5.3) The lethal spotting mutation (*ls*)

Lethal spotting is a recessive mutation which arose in a subline of the inbred strain C57BL-a^t. Mice homozygous for this mutation show considerable white spotting. The eyes are usually black. The mutation resembles another spotting mutant, piebald lethal, with respect to coat colour, except that the ears and tails of *ls/ls* mice are less pigmented (Lane, 1966). These mice usually die before three weeks of age, with a grossly distended distal colon. Some homozygotes do however survive long enough to breed. Lane (1966) carried out experiments to determine the cause of the megacolon. Histological preparations were made from the colon of 12 day old *ls/ls* mice, and also from controls. The proximal section of the gut of *ls/ls* mice was found to be comparable to that of the normal controls, but the distal portion of the colon of *ls/ls* mice was found to be completely aganglionic. This connection between abnormal pigmentation and abnormal innervation of the colon is especially interesting because both cell types originate from the same region of the neural crest during development. This implies that the primary defect in lethal spotted mice arises from either a reduction of the neural crest cells, or a failure of these cells to migrate to the terminal portion of the chromosome.

A later study by Rothman and Gershon (1984) examined further the histological abnormalities in lethal spotted mice. The enteric nervous system (ENS) was examined from gut sections in adult *ls/ls* mice for the activity of the neurotransmitters acetylcholine (ACH - detected by acetylcholinesterase (AChE) activity), 5-hydroxytryptamine (5-HT) and serotonin. No differences in the uptake or activity of any of these neurological markers was noted in any region of the gastrointestinal tract except in the terminal portion of the distal colon. The distal colon of *ls/ls* animals was found to contain large numbers of loose bundles of nerve fibres, but no cell bodies were seen in the terminal 2mm of the colon (Rothman and Gershon, 1984). These neurones showed AchE activity. These neurones are abnormal with respect to the other

neurological markers tested in that no 5-HT was either taken up, or detected by immunological methods. The temporal appearance of the nerves in the gut of *ls/ls* mice was analysed, and found to be comparable with wild type mice (Rothman and Gershon, 1984). To determine whether progenitor nerve cells were present in the terminal 2mm of the gut, explants of tissue from this region were cultured. It is not possible to identify these precursor cells prior to their differentiation, but any neuronal growth in the tissue cultures must be derived from the tissue used to seed the plates. No evidence of the presence of neuronal precursor cells was found in the terminal 2mm of the colon (Rothman and Gershon, 1984). This lack of ganglia could be due to the inability of these precursor cells to migrate into this region of the colon. It is seen that some ganglia are present in an abnormal position in the region of the gut that lies just proximal to this - the cells are seen external to the longitudinal muscle, but nerve ganglia are not seen in this position in wild type mice. These cells could result from the growth of the cells that could not migrate to the terminal region of the gut. It is probable that the microenvironment of this region of the colon is defective and prohibits migration of neuronal precursors. Kapur *et al.* (1993) examined whether the primary defect in lethal spotting was intrinsic to the migrating neuronal precusors, or to some extrinsic effect of the cellular microenvironment in the terminal portion of the gut. Aggregation chimeric mice were generated from *ls/ls* (also carrying DβH-nlacZ transgene in homozygous form) and +/+ mice, and analysed for evidence of external pigment spotting, and the status of the terminal gut with respect to neuronal colonisation. Enteric neurones derived from *ls/ls* mice were identified by virtue of their blue colour following treatment with X-gal. This distinction was possible because of the targeted expression of the *lacZ* gene by the DβH-n promoter, in the transgene construct. The percentage of wild type neurones was determined by scintillation count analysis of Southern blotted DNA from tissue taken from the chimeric mice, hybridised to a *lacZ* probe. The study showed that in cases where there were greater than 20% wild type cells, neurones could colonise all parts of the gut, including the terminal 2mm. This indicates that the defect in *ls/ls* mice lies in the cellular microenvironment, and is not a function of the neurones themselves.

Lethal spotting has been put forward as an animal model of Hirschsprung's disease in humans. The aetiology of the syndrome is almost identical to the situation in *ls/ls* mice, in that the terminal portion of the bowel is aganglionic. One form of Hirschsprung's disease in humans has recently been shown to result from a mutation in the *Ret* proto-oncogene located on chromosome 10q11.2. The colonisation of the bowel by cells deriving from the neural crest is

extremely complex, and controlled by many genes. Thus, lethal spotting could still be a model for Hirchsprung's disease by its possible identity as the mouse homologue of another form of the disorder, or by having an indirect effect acting to modify one or more of the controlling genes.

1.5.4) The Ulnaless mutation (*Ul*)

Ulnaless (*Ul*) is a dominant or semi-dominant mutation that arose from irradiation experiments at the MRC Radiobiology unit at Harwell, in the 1960s. The mutation was first seen in the progeny of a female originating from a HT test stock, mated to 3H1 (C3H/HeH X 101/H)F1 male (Morris, 1967). The mutation causes abnormal development of the murine skeleton. The long bones of the legs are drastically shortened. Although the mutation was discovered in the late sixties, no thorough description had been made of the phenotypic effects of the gene, and no mapping data were available until 1990. Davisson and Cattanach (1990) carried out two point linkage experiments to assign a map position to the mutant. The segregation of the *Ul* mutation was compared to that of non-agouti (*a*), brachypodism (*bp*) and ragged (*Ra*). The mutation was shown to map to proximal mouse chromosome two, with loose linkage to *a* and *bp*, but no linkage to *Ra*, the most distal marker in the cross. Three point crosses were then carried out with the markers pallid (*pa*), *a* and the Robertsonian translocation Rb(2;8)Lub. Results from these crosses confirmed the proximal location of *Ul*, and placed the mutant a genetic distance of 18.4 +/- 3.9 map units proximal to *pa*. Davisson and Cattanach (1990) also carried out a more detailed analysis of the phenotypic effects of the mutation. They determined that there was a considerable reduction in length of the tibia and fibula, together with structural abnormalities. The tibia was bowed, and both tibia and fibula were rotated through 90 degrees. The flexibility of the metatarsals and metacarpels was also abnormal. These effects were only seen in heterozygotes, as the male *Ul*/+ mice do not breed. The sperm of these mice was found to be normally active, and of normal structure. The lack of vaginal plugs in the female mice housed with these males suggests that the infertility was due to an inability to mate rather that some dysfunction of the sperm. The percentage of *Ul*/+ mice was lower than expected in some crosses. This could be due to a slight reduction in viability in fetal and neonatal *Ul*/+ mice. Penetrance of the mutation appears to be total, making *Ul*/+ a useful marker for proximal mouse chromosome two. Examination of the bone histology in ulnaless mice reveals that although the bones are severely shortened, the bone morphology is normal (Davisson and Cattanach, 1990). Ulnaless could prove to be an animal model for some human mesomelic dwarfism syndromes, but this has yet to be established.

1.6) Cloning strategies in man and mouse

In order for the effects of any gene of interest to be analysed in detail, the gene must first be cloned. This enables the structure of the gene to be examined on a molecular level, and in addition, allows the detection of mutations or rearrangements which have effects on the phenotype.

Genes have been cloned in the past, on the basis of some knowledge of their function or the previous isolation of the gene product. The cloning of some genes has been facilitated by the occurrence of chromosomal rearrangements which have disrupted the function of the genes concerned. The breakpoints of the rearrangement point to the location of the gene. In the last two decades, techniques that allow the cloning of a gene with no prior knowledge of its function have been developed. These are known as positional cloning techniques, as they are based on knowledge of the map position of any gene of interest.

1.6.1) Cloning of genes on the basis of their function

Functional cloning of a gene relies upon the existence of knowledge of the gene product, its function, or similarities in the phenotypes it produces when mutated between species. These methods of gene cloning are the ones which have been historically used. The albino (*c*) locus was cloned (Kwon *et al.*, 1987) on the basis of some knowledge of its function, and predictions of the likely structure of the protein. Human melanocyte libraries were screened with hamster antibodies against the tyrosinase gene product. Of the clones that were sequenced, one showed characteristics of a copper based oxidoreductase, such as characteristic glycosylation sites, and copper binding motifs. Subsequent Southern blot analysis showed that the clone was very tightly linked to the albino locus. Another example of functional cloning is the cloning of the steel (*Sl*) gene. Upon the finding that the dominant spotting gene (*W*) was the *c-kit* oncogene (Geissler *et al.*, 1988), studies were initiated to discover the gene coding for its ligand. One candidate was the steel gene, mutations in which produced many phenotypic similarities with *W*. The protein product of the steel gene, the so called mast cell growth factor (MGF), which co-precipitates with the *c-kit* oncogene in response to a *c-kit* antibody, was shown to be the site of three independent mutations in steel mice (Zsebo *et al.*, 1990). Thus, the steel gene was cloned on the basis of its phenotypic similarity to the *W* gene. These methods,

however, require information on the structure or function of the gene to be cloned. This is not available in many cases.

1.6.2) Gene cloning on the basis of disruption of function

The occurrence of a chromosomal rearrangement which disrupts one or more genes is a great benefit for geneticists hoping to clone the gene. The physical boundaries of such a rearrangement provide an effective indicator of the position of the gene. Many of these rearrangements occur in nature, and many more result from human intervention in the form of mutagenesis experiments (such as X-ray irradiation of male mice), which produces not only point mutations, but also chromosomal rearrangements such as deletions, inversions, insertions and translocations. Genes cloned because of a fortuitous chromosomal rearrangement, either manmade or naturally occurring, include the limb deformity locus (*ld*), and the agouti locus (*a*) on mouse chromosome two. The limb deformity mutant is a spontaneous mutation which arose on two separate occasions (Cupp, 1962; Green, 1962). Woychik *et al.* (1985) generated the mutant some years later by the fortuitous insertion of a transgene into a region of the gene essential for its function. X-irradiation of male mice carried out by the same group (Woychik *et al.*, 1990) produced an insertion event which disrupted both the *ld* locus, and the *a* locus. Karyotype analysis of mutant mice revealed that a section of chromosome 17 had undergone an insertion into mouse chromosome two. A probe produced from sequences flanking the transgene insertion point (Woychik *et al.*, 1985) was used to examine gene expression in tissues from mutant mice. Transcripts of incorrect size were found in the brain and testes of mutant animals. The limb deformity locus was later shown to code for a variety of nuclear protein isoforms termed 'formins' (Woychik *et al.*, 1990; Vogt *et al.*, 1993). The same insertion allowed the cloning of a gene, mutations in which give rise to certain alleles at the agouti (*a*) locus (Bultman *et al.*, 1992). This gene has approximately 19 diferent alleles, arranged in a strict dominance hierarchy (Green, 1989). The gene was shown to contain four exons, three of which are coding. The gene product is an 131 amino acid protein, with characteristics indicative of secretory signal proteins (Bultman *et al.*, 1992; Michaud *et al.*, 1993; Miller *et al.*, 1993). The gene was shown to be rearranged in several agouti alleles (A^y, *a*, a^t, A^w) (Bultman *et al.*, 1994). It was postulted that this single gene was the site of mutation for all the agouti alleles, given the observation that the most dominant (lethal yellow A^y) and the most recessive (extreme non-agouti a^e) alleles in the series show alterations in the structure of this gene (Bultman *et al.*, 1992). Further evidence for the presence of a single gene, mutations in which are responsible for all the alleles at this locus, stems from the observation that

association of the *Xmv-10* gene with certain agouti alleles was due to close linkage, rather than true causality (Winkes *et al.*, 1994).

The occurrence of these chromosomal rearrangements so as to disrupt function with respect to specific genes is relatively rare, although animals with disruptions in genes can be generated by gene targeting, or by transgene insertions. Gene cloning by disruption of function has been useful for cloning some genes, but in the case of the majority of genes, no such rearrangement is available, and other methods must be employed.

1.6.3) Positional cloning of genes

Positional cloning of genes is possible even in the absence of any information on the gene to be cloned, except for its map position. Since for the majority of genes, the only information available includes the clinical manifestations, rough map positions, and perhaps some idea of the tissues which are affected, the ability to clone genes without detailed knowledge of gene structure or function is invaluable. The first stage in any positional cloning project is to construct a detailed genetic or physical map of the area in which the gene is thought to lie. From this map, the region which contains the gene is defined by flanking markers. This provides a measure of the absolute limits within which the gene must lie. In most cases, these are distances too large to look directly for candidate genes. The boundaries must be narrowed down further by techniques such as chromosome walking and jumping. Chromosome walking is the process of using the ends of DNA fragments as probes to identify those sequences which lie next to them. Chromosome walking can be carried out in two directions from the starting point, but is not an effective means of covering large distances. This is better done by chromosome jumping experiments (Collins and Weissman, 1984). This method has the advantage that larger distances can be covered, and each point on the jump can be used as the starting point for bidirectional chromosome walks. Another technology is the use of yeast artificial chromosomes (YACs) to clone and propagate large stretches of DNA (Burke *et al.*, 1987). If probes are available for markers closely linked to the mutant of interest, pools of YACS can be screened to identify clones that contain sequences originating from that region. Such a clone may in fortuitous circumstances contain the gene of interest, if linkage to the marker was very tight. Alternatively, an end probe from the positive YAC, can be used to isolate YACs in the pool which contain sequences which are contiguous with it (Coulson *et al.*, 1988). End probes may be generated from YACs by a number of procedures, such as inverse PCR (Silverman *et al.*, 1989) or by amplification of the DNA between YAC vector sequences

and repeat sequences in the insert (Nelson *et al.*, 1989). A useful technique is the use of specially constructed 'bubble' vectorettes, which allow the specific amplification of end sequences from YACs, by the use of an internal primer to the vectorette, and a primer specific to the YAC vector (Riley *et al.*, 1990). These technologies, together with ways of separating and visualising large DNA molecules such as pulsed field gel electrophoresis (PFGE) are powerful tools in the isolation of new genes (Schwartz and Cantor, 1984). PFGE allows the separation of large DNA fragments, produced by digestion with rare cutter restriction enzymes. By this method, it is possible to make long range restriction maps of the sequences surrounding the mutant gene, and to isolate the mutant and surrounding markers onto one or two restriction fragments. This is particularly useful, since linkage mapping can give inaccurate estimates of the distances involved if there are hotspots, or depressions of recombination. The PFGE map gives a definite measurement of physical distance. By the methods outlined, a contiguous range of fragments are produced (a contig).

Once a contig has been assembled, it can be scanned for expressed sequences. This can be done by many means, for instance, Southern blots of restriction digested DNA from several species (zoo blots) will identify areas of conservation which may correspond to coding regions (Monaco *et al.*, 1986). The contig can be digested with rare cutter enzymes to identify CpG islands (identified by clusters of rare cutter restriction sites) which are associated with the 5' ends of housekeeping genes and most tissue specific genes (Bird, 1986; Bird, 1987). Alternatively, cDNA libraries or Northern blots can be screened with whole phage or cosmid clones. This identifies any transcripts originating from the clones (Spies *et al.*, 1989). Over short scales, direct sequencing of clones and the identification of promoter sequences (Allen *et al.*, 1988), or open reading frames (Gray *et al.*, 1982), can prove useful in some cases. However, this is a very laborious task if many large clones are involved. A useful technique is exon amplification (exon trapping). This method relies upon *in vivo* selection for functional splice sites in genomic DNA. The technique, originally described by Duyk *et al.* (1990) requires specialised vectors, constructed with an intron of the HIV 1 *tat* gene, which is flanked by splice sites, and exons. The intron contains a multiple cloning site (MCS), and the SV40 early promoter. Genomic DNA is shotgun cloned into the MCS, and the resulting construct is transfected into COS cells, which direct transcription of the insert. The RNA message is then spliced using the donor splice site of the vector, and the splice acceptor site of the insert, to produce a heterogeneous fragment composed of the HIV-1 *tat* sequence joined to the captured exon. Primers specific to HIV-1 *tat* sequence are then used to amplify across the insert, which can then be cloned into a suitable plasmid vector and sequenced. Modern exon trapping vectors

incorporate sequences to circumvent the common problems encountered with the original exon trapping protocols, specifically, the background of non-recombinant molecules, which are preferentially amplified by the PCR step (being very abundant and of small size) and the problem of recombinant molecules containing no exogenous exons, due to splicing using a cryptic splice site in the vector (Church *et al.*, 1994). Exons recovered from these experiments can be used to identify cDNAs originating from expressed sequences in the contig. Once sequence is available for the mutant, comparisons can be made between mutated forms (affected individuals, mutant mice) and the wild type gene with an aim of identifying causative mutations.

An example of a gene cloned by positional cloning approaches is the short ear (*se*) gene on mouse chromosome 9 (Kingsley *et al.*, 1992). This mutation causes skeletal deformities, such as a reduction in the size of the ears, and deletion of pairs of ribs (Kingsley *et al.*, 1992). The gene was cloned by the production of a contig from an extended chromosome walk, originated from a point estimated to lie 200 to 300 kb from the short ear locus, as defined by deletion breakpoints (Kingsley, *et al.*, 1990). At each step in the walk, probes were produced, and used to analyse DNA from wild type and mutant mice, to discover if they were localised to deletions associated with various alleles of short ear. PCR primers were designed from sequences mapping within one such deletion. PCR screening of YAC libraries with these primers identified two clones, originating from this region were identified, and aligned to the chromosome walk, by cloning the ends of the YAC from λ libraries. Exon trapping experiments were carried out, which identified a fragment with homology to the *Bmp-5* gene (Kingsley *et al.* 1992). The expression of this gene was found to be aberrant in several alleles of short ear (such as se^{20Zb} and se^{36DSD}). In addition, the open reading frame of the *Bmp-5* gene was seen to be interrupted by a translocation which results in the short ear phenotype (Kingsley *et al.*, 1992). This data suggests that the *Bmp-5* gene is located at the short ear locus (Kingsley *et al.*, 1992).

1.6.4) The candidate gene approach

Genes cloned in the mouse by candidate gene approaches include the pallid and undulated mutations on mouse chromosome two. Recent work on the *Epb4.2* gene on mouse chromosome two has indicated that this gene may be a possible candidate for the pallid (*pa*) mutation (White *et al.*, 1992). This gene was mapped to mouse chromosome two, and was

found to cosegregate with the pallid gene. Southern analysis with an *Ebp4.2* cDNA probe showed differences in restriction band patterns between pallid and wild type mice, and studies on the expression of *Epb4.2* in pallid mice revealed mRNAs of abnormal size (White *et al.*, 1992). Similarly, the undulated (*un*) mutation on mouse chromosome two was shown by candidate gene approaches to be due to a mutation in the *Pax 1* gene. *Pax 1* was shown to map between the β2 microglobulin and the agouti loci on mouse chromosome two, with close linkage to undulated (Balling *et al.*, 1988). On the basis of the close linkage, *Pax 1* was suggested as a candidate gene for undulated. Direct sequencing of the *Pax 1* gene in undulated and wild type mice revealed a point mutation in a highly conserved region, which produced a Gly to Ser amino acid change (Balling *et al.*, 1988). This alteration was later shown to cause a reduction in the DNA binding specificity of *Pax 1* (Chalepakis *et al.*, 1991). Subsequent studies have shown that undulated extensive (un^{ex}) and undulated short tail (un^{s}) alleles are also associated with mutations in *Pax 1* (Balling *et al.*, 1992).

1.6.5) Differences in strategy between man and mouse

Although the techniques used in the actual cloning can be applied equally to man and mouse, there are certain differences in the initial steps. In man, studies to clone a mutated gene are initiated by pedigree analysis of disease families. This has some disadvantages. These families are usually small, and affected individuals sometimes do not breed. Linkage can be difficult to ascertain with such small sample sizes. Once linkage is established to a chromosome, markers originating from that chromosome are tested until a closest linkage is obtained. The construction of the genetic map by this method can be a long and difficult task. Problems are sometimes encountered where an individual vital to the analysis is dead or uncooperative. In the mouse, detection of linkage is much easier, as specialised backcrosses can be set up which segregate the mutant of interest. The large numbers of backcross progeny that can be generated allows the construction of high resolution linkage maps. However, cytogenetic analysis plays a much greater role in the cloning of human genes than in the cloning of mouse genes. New markers can be mapped quickly and efficiently in man by cytogenetic methods. This is more difficult in the mouse, because of the difficulty in identification of the mouse chromosomes, which are all acrocentric, and of similar sizes. Once the linkage map has been constructed, the positional cloning procedures outlined here can be equally applied to both man and mouse.

1.7) Somatic cell hybrids

A somatic cell hybrid is a cell line that originates from the fusion of two different cell types. These may be either intraspecific - two cell types from within the same species (Ephrussi *et al.*, 1964) or interspecific - two cell types originating from different species (Weiss and Ephrussi, 1965, 1966a). Although spontaneous cell fusion in mixed populations does occur, it is a relatively rare event. Somatic cell hybrids can be formed at greater frequency, by the use of inactivated *Sendai* virus (Poste, 1972), or by the use of polyethylene glycol (PEG) (Portecorvo, 1975). The fusion results from the prior formation of heterokaryon cells (cells with two nuclei inside one cell membrane) which may fuse to form a stable hybrid population. The frequency of heterokaryon formation is influenced by the choice of parental cell types, for instance, cells originating from different inbred strain of mice have been shown to produce fusion at different frequencies in somatic cell hybrid studies (Minna *et al.*, 1975). Comparisons of the ability to fuse of cells originating from different species have also been made (Okada and Tadokoro, 1963). The ratio of parental cell types in the fusion - preferably 1:1 (Klebe *et al.*, 1970: Davidson and Ephrussi, 1970), and the pH of the reaction - optimally between 7.6 to 8.0 (Croce et al., 1972) and the Ca^{+} ion and oxygen concentration in the fusion vessel (Okada, 1969) are also known to be important factors. It has been shown that the cells most likely to form heterokaryons and subsequently fuse to form hybrids are those which are undergoing synchronous DNA replication and cell division (Harris *et al.*, 1966). The mechanism is poorly understood. It has been suggested (Hosaka and Koshi, 1966) that agents such as inactivated *Sendai* virus act in four steps by a) adsorbing to the cells, and causing cellular agglutination, b) causing damage to the cell membrane at the site of adsorption, c) repair of the damaged sites, leading to interconnecting bridges between the cells d) the degeneration of the cell membrane between the interconnecting bridges, leading to the heterokaryon formation (see figure 7). Chemicals such as PEG may act by altering the properties of the cell membranes. The resulting fusion products are then cultured in selective media to permit the identification of the hybrid cells (Littlefield, 1964; Davidson and Ephrussi, 1965). The selection is usually on the basis of one parent cell being deficient in a metabolic enzyme of the salvage pathway. These cells can survive when the medium is supplemented with all required factors for metabolism, but is unable to grow in media containing agents which block the regular purine and pyrimidine biosynthesis pathways (Littlefield, 1964). Selection systems will be discussed in detail in the relevant section. Thus, the only cell types able to survive are hybrid cells, which have been 'rescued' by the presence of the enzyme from

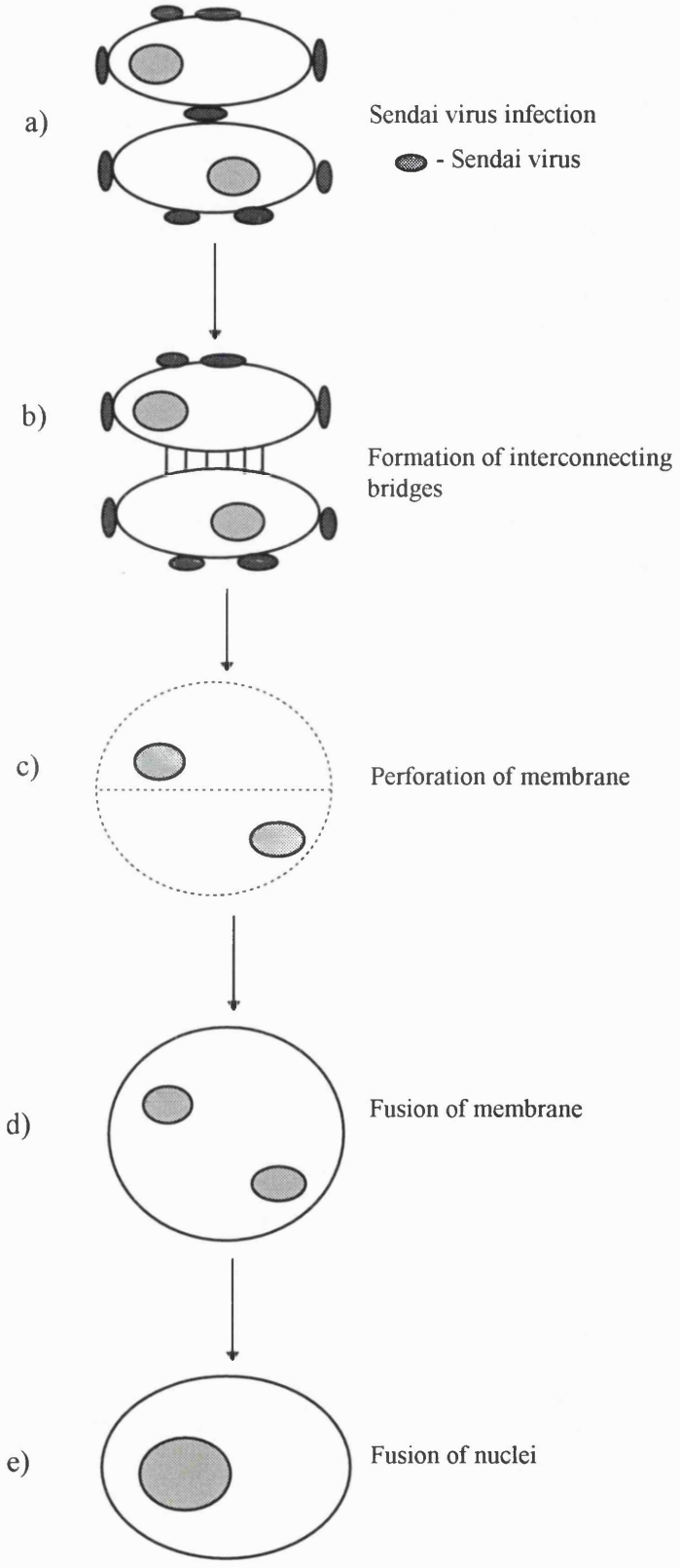


Figure 7: -Mechanism of cell fusion (adapted from Giles and Ruddle, 1973)

the other parent, and the non-deficient parental cell type, which is eventually overgrown in culture by the hybrid cells. The original cell fusion experiments all used cell types both derived from long term cultured lines (Scarletta *et al.*, 1967; Weiss and Green, 1966), but the use of primary cell lines as the non-deficient parent in the fusion has the advantage that these cell lines do not thrive in culture, and are quickly lost. Thus, the only cell types that are positively selected for are those formed from the fusion of the parental cell types (Nabholz *et al.*, 1969: Minna *et al.*, 1975). Also, these cell types are easily obtained, with no need for prior cell culture, and are diploid, with a regular karyotype. The resulting hybrids contain contributions from both parental genomes, which are both functionally active on the basis of colony morphology and isozyme activity (Weiss and Ephrussi, 1966b: Scarletta *et al.*, 1967). See figure 8 for a schematic diagram of the formation of somatic cell hybrids.

It was noted from the earliest experiments that the number of chromosomes in cells resulting from these fusions was significantly less than expected from the fusion of two cell lines with known modal chromosome numbers. This was seen to be due to loss of chromosomes from one species (Ephrussi *et al.*, 1964; Weiss and Ephrussi, 1965, 1966a). The mechanisms of chromosome loss are unknown, but non-disjunction, endoreplication and multipolar mitosis have all been suggested (Weiss and Ephrussi, 1966a; Handmaker, 1971). There are seen to be two patterns of chromosomal segregation a) extremely rapid segregation, occurring in the first 20 generations, probably in the first few cell divisions (White and Ephrussi, 1966a; Nabholz *et al.*, 1969), which then ceases, leaving a karyotypically stable cell population and b) a balanced chromosome complement from both species, with segregation occurring in a gradual fashion over several hundred generations. It has been noted that in the case of human : rodent hybrids, segregation occurs according to the first pattern, whereas in rodent : rodent hybrids, the second pattern is also common (Minna *et al.*, 1974, Minna, 1795). This may be due to differences in the mechanisms of segregation between the species, as although non-disjunction has been shown to play a role in segregation of rodent chromosomes, the same has not been proved for human chromosomes (Handmaker, 1971). Although the hybrids are karyotypically stable after approximately 20 generations, there is evidence that somatic cell hybrids may lose some stability in long term culture. Weiss and Green (1966) discovered that although their panel of human : rodent hybrids showed considerable karyotypic stability from 20 to 60 generations, after 85 generations, the number of human chromosomes declined by half, with 10 to 25% of hybrid lines losing all human chromosomes. In general, the early instability of the karyotype leads to heterogeneity in the hybrid population.

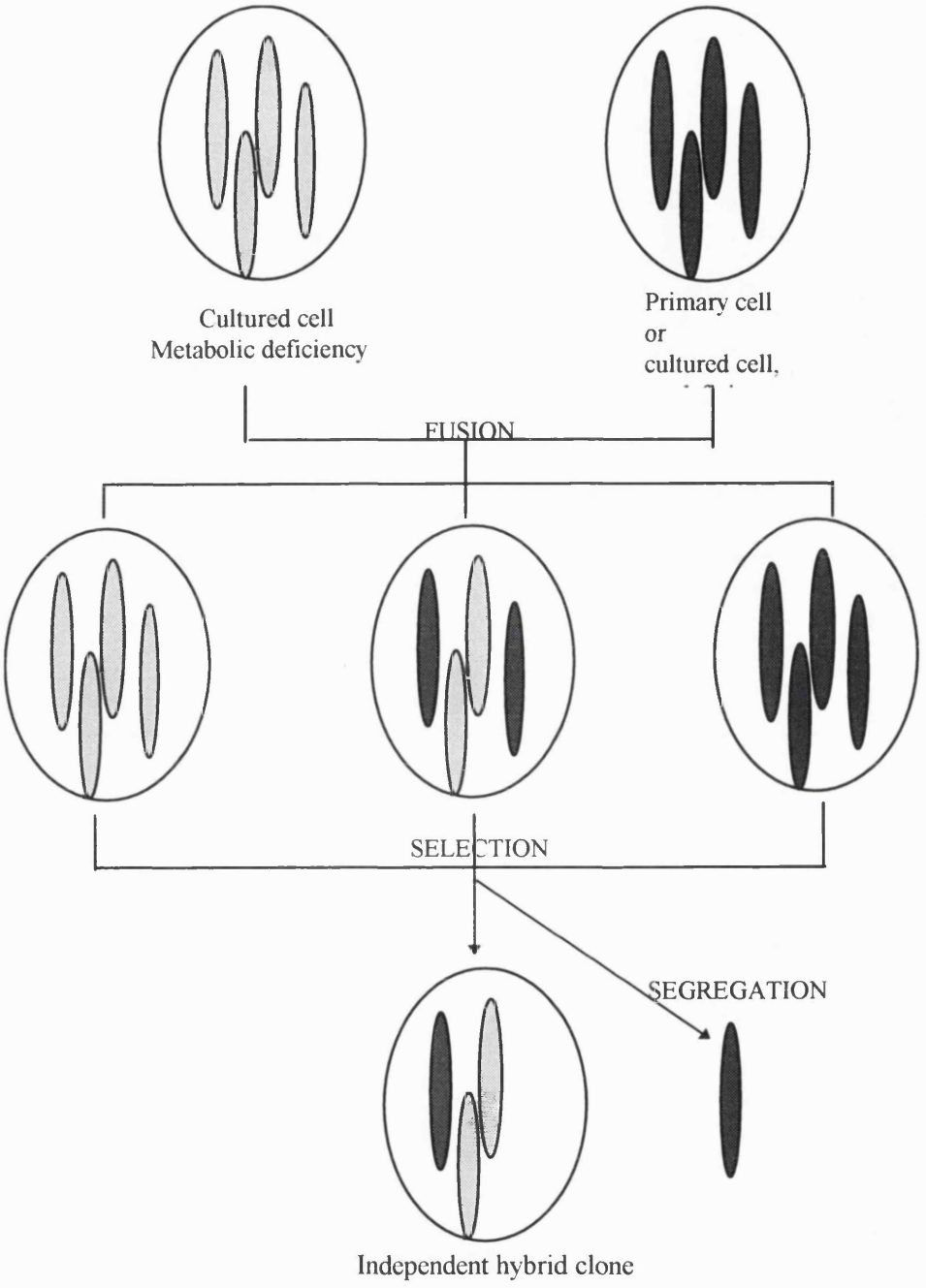


Figure 8:- The construction of somatic cell hybrids

In the case of rodent : human somatic cell hybrids, the vast majority of lines segregate human chromosomes (Weiss and Green, 1966; Nabholz *et al.*, 1967), whereas mouse : Chinese hamster somatic hybrids preferentially segregate mouse chromosomes (Scarletta *et al.*, 1966). There are exceptions to this rule, as panels of mouse : human hybrids have been reported which segregate mouse chromosomes (Weiss *et al.*, 1968, Jami *et al.*, 1971: Minna and Coon, 1974). Similarly, panels of mouse : Chinese hamster hybrids have been reported, which segregate hamster chromosomes (Handmaker, 1971). The determination of which parent undergoes segregation without outside intervention is not known. It has been suggested that the growth rate of the cells may be involved, with an inverse correlation between growth rate and segregation (Kao and Puck, 1970). This however does not hold for all panels (Handmaker, 1971). In cases where one parent is a primary cell line, the chromosomes of this are always segregated (Minna *et al.*, 1975). Where it is necessary to use two lines which have been maintained in long term culture in the fusion, X-irradiation or bromodeoxyuridine (BUdR) treatment of one parent results in the segregation of chromosomes of that parent. However, these treatments also lead to chromosomal breakage and rearrangement (Portecorvo, 1971).

If the segregation resulted entirely from mitotic 'accidents', it would be expected that the segregation of particular chromosomes would be random, with respect to each other. This is not the case. Norum and Migeon (1974) showed statistically significant alterations in the frequency of various unselected autosomes in their hybrid panels, and those of others (Allerdice *et al.*, 1973; Croce *et al.*, 1973). In the panel of Norum and Migeon (1974), the frequency of the Y chromosome was lowest, which is not unexpected as the human parent derived from a male, only one Y chromosome was present per cell, although since only 5/51 cells analysed possessed the Y, the effect cannot be entirely attributed to this.. The X chromosome was present at a higher than expected frequency, as was chromosome 12 (Norum and Migeon, 1974). This presumably indicates that strong selective pressures are acting on individual chromosomes in the system. This could be due, in part, to positive or negative effects on particular gene products (Norum and Migeon, 1974). This effect has also been noted for mouse chromosome 11, Y (negative selection) and 15 (positive selection) in mouse : Chinese hamster somatic cell hybrids (Francke *et al.*, 1977; Kozak and Ruddle, 1977).

1.7.1) Selective systems for somatic cell hybrids

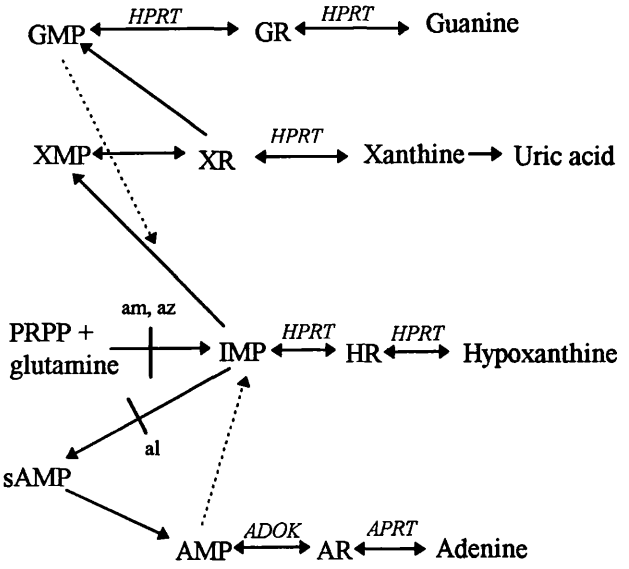
There are principally two methods of identifying hybrid cells from mixed populations of parental cells and hybrid cells. These are a) methods based on growth characteristics and/or

colony morphology of the hybrid clones and b) biochemical methods exploiting the sensitivity of one parental cell to chemicals in the growth medium.

1.7.1a) Hybrid selection by biochemical methods

Selection of hybrid cells in somatic fusion experiments relies upon the availability of mutations in one cell line which can be complemented by the introduction of a normal gene. There are many well characterized mutants of this sort available in Chinese hamster cell lines (reviewed in Puck and Kao, 1982). Some of these mutations are naturally occurring, many more are created by chemical treatment to knock out specific genes (Kao and Puck, 1974). Most of these mutants result from deficiencies in enzymes of the metabolic pathways. Mutant cells can be directly isolated in some cases. One method for this is the growth of mixed populations of deficient and non-deficient cells in media which will not sustain the growth of the deficient population. The culture is then treated with chemicals (such as bromodeoxyuridine (BUdR) which attack only fast growing cells. In this case, subsequent treatment with visible light kills the fast growing (non-deficient) cells, leaving the deficient auxotrophic mutant cells, which can be recovered by supplementing the growth media with the required nutrients (Puck and Kao, 1967). Temperature senstitivity mutations have also been used as mechanisms of selection (Thompson *et al.*, 1970).

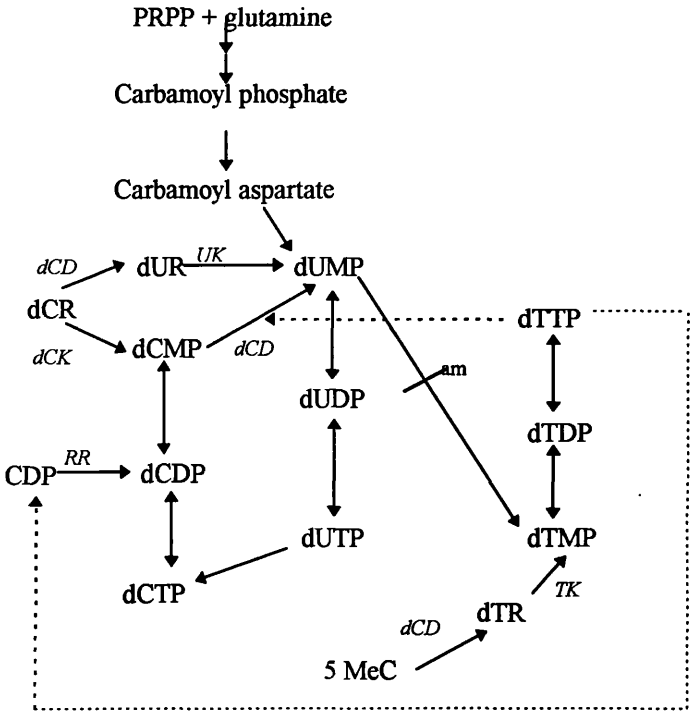
Mutant cells can thrive in culture provided that key products in the deficient enzyme pathways are provided. For instance, in selective systems relying on aminopterin imposed blocks on the purine biosynthesis pathways (see figure 9), the cell can synthesize all purines from IMP (inosine 5' monophosphate) provided all the necessary enzymes and metabolites are present. Cells in which blocks are imposed for the major IMP to AMP and GMP pathways can still synthesize guanine and adenine from hypoxanthine, provided hypoxanthine phosphoribosyltransferase (HPRT) is present, via the salvage pathway (Adams *et al.*, 1986). Specialized cell culture media have been designed, which allow only the growth of hybrid cells after fusion experiments. The selective media were designed primarily to isolate hybrid cells from the fusion products, but because they require the presence of specific enzymes, they also select for specific chromosomes. Similarly, the systems used to isolate enzyme deficient mutant cell lines can be used to provide negative selection against specific chromosomes. In the case of hypoxanthine phosphoribosyltransferase (HPRT), treatment of cell populations with 6-thioguanine or 8-azaguanine selects for loss of HPRT activity (Szybalski *et al.*, 1962; Littlefield, 1964), either by selecting for mutants at the *Hprt* locus in parental cell lines, or by



PRPP = 5' phosphoribosyl pyrophosphate, al = alanosine, am = aminopterin, az = azaserine, IMP = inosine 5' monophosphate, sAMP = adenylosuccininc acid, AMP = adenosine 5' monophosphate, AR = adenosine, *APRT* = adenosine phosphoribosyl transferase, HR = inosine, *HPRT* = hypoxanthine phosphoribosyl transferase, XMP = xanthine 5' monophosphate, XR = xanthosine, GR = guanosine, GMP = guanosine5' monophosphate, *ADOK* = adenosine kinase.

⋯⋯▸ negative feedback

Figure 9: - The purine bioynthesis pathways



CDP = cytidine 5' monophosphate, dCDP = deoxycytidine 5' diphosphate, dCMP = deoxycytidine 5; monophosphate, dCTP = deoxycytidine 5' triphosphate, dUMP = deoxyuridine 5' monophosphate, dUR = deoxyuridine, dCR = deoxycytidine, dTR = deoxythymidine, dTMP = deoxythymidine 5' monophosphate, dTDP = deoxythymidine 5' diphosphate, dTTP = deoxythymidine 5' triphosphate, *RR* = ribonucleoside diphosphate reductase, *dCD* = deoxycytidine deaminase, *dCK* = deoxycytidine kinase, *TK* = thymidine kinase, al = alanosine, *UK* = uridine kinase, 5MeC = 5' methyldeoxycytidine.

Figure 10: - The pyrimidine biosynthesis pathways

selecting for the loss of the X chromosome in hybrid cells. Similar chemicals exist to provide negative selection, and isolation of mutant parental cell lines for fusion experiments, in other systems (Kusano *et al.*, 1971; Chan *et al.*, 1975).

Similar selection systems have been designed to exploit blocks in the purine biosynthesis pathways (see figure 10) (Chan *et al.*, 1975; Dechamps *et al.*, 1975). The HPRT selection system also allows selection of hybrid cells containing the TK gene, as aminopterin also imposes blocks on the pyrimidine biosynthesis pathways (Littlefield,1964).

1.7.1b) Selective media

HAT selection

HAT medium is composed of Hypoxanthine (10^{-5} M), Aminopterin (3.2 X 10^{-6} M) and Thymidine (10^{-5} M) (Littlefield, 1964). The aminopterin in the medium imposes blocks on purine and pyrimidine biosynthesis, which prevent the *de novo* synthesis of nucleotides. Wild type cells can still produce the required nucleotides by the salvage pathways, because the enzyme Hypoxanthine phosphoribosyltransferase (HPRT) catalyses the reaction from hypoxanthine to IMP (see figure 11). $HPRT^-$ cells are unable to catalyze this reaction, and cannot survive in HAT medium. Similarly, because of the inhibition of dihydrofolate reductase, cells are unable to form thymidine 5' monophosphate from deoxy uridine 5' monophosphate, and thus cannot synthesize thymine. Wild type cells can however synthesize thymine from thymidine in the presence of the enzyme thymidine kinase (TK). TK^- cells are unable to survive in HAT medium. In addition to the blocking agent aminopterin, the medium supplies the substrates for HPRT (hypoxanthine) and TK (thymidine). Fusions between $TK^+/HPRT^-$ and $TK^-/HPRT^+$ cells which are subsequently grown in HAT media allow only the proliferation of cells which possess both of these enzymes. Thus, neither the $HPRT^-$ cells or the TK^- cells are able to survive, whereas the fused hybrid cells are 'rescued' by the endogenous HPRT or TK gene of the other fusion partner. It is not always necessary to select against both parental lines by biochemical means. A fusion between a primary cell line which is $HPRT^+$, and a permenent cell line which is $HPRT^-$ is much more common. Here again, the $HPRT^-$ cells cannot survive, but the fused cells are rescued, and frequently retain the fast growing properties of the $HPRT^+$ parent. The primary cells grow slowly if at all, and do not compete effectively with the hybrid cells. The use of HAT medium allows positive selection for the X chromosome (location of the *Hprt* gene) and mouse chromosome 11 (location of the

Tk gene). Methotrexate is occasionally used in place of aminopterin in HAT medium (it has the same effect), but it is less toxic.

Selective systems exist for chromosomes other than mouse chromosome 11 and the X. These are summarized in table 1 below.

mouse chromosome	Locus	positive selection	negative selection
2	FPGS	glycine, adenosine and thymidine deficient media (Fournier and Moran, 1982)	-
8	APRT	alanosine and adenine (Kusano *et al.*, 1971)	2-fluroadenine (Kusano *et al.*, 1971)
11	TK	HAT medium (Littlefield, 1964)	bromo deoxyuridine (Littlefield, 1964)
14	ADOK	alanosine, adenosine and uridine (Ruddle and Creagan, 1975)	MMPR (Thacker, 1980).
X	HPRT	HAT medium (Littlefield, 1964)	6-thioguanine, 8-azaguanine (Littlefield, 1964)
unmapped	serine hydroxymethylase	glycine deficient media (Jones *et al.*, 1972).	-
unmapped	deoxycytidine deaminase	HAM medium (Chan *et al.*, 1975)	5-bromodeoxy cytidine (Chan *et al.*, 1975)
unmapped	deoxycytidine kinase	HAT and deoxycytidine (Dechamps *et al.*, 1975)	cytosine arabinoside (Dechamps *et al.*, 1975)
unmapped	uridine kinase	adenosine and uridine (Medrano and Green, 1974)	flurouridine (Medrano and Green, 1974)

Table 1:- Mouse chromosomes for which selection systems are available (adapted from Ruddle and Creagan, 1975)

1.7.2) Methods of selecting chromosomes of choice in hybrid cells

The selective media described above provide selection systems for specific chromosomes. The ability to select for any particular chromosome depends upon the presence of a selectable marker on that chromosome, or the presence of a translocation involving the

chromosome of choice, and a chromosome harboring a selectable marker. This imposes restrictions on the chromosomes which can be specifically isolated in somatic cell hybrids. However, the presence on the cell surface of antigens can be used as a selective system, provided a suitable antibody against that antigen can be found (Chu and Powell, 1976). Genes coding for cell surface antigens are present on all mouse chromosomes, and this provides the basis of a method to select for any chromosome. This method depends upon the targeting of cell surface antigens with specific antisera, followed by complement mediated cell death (see figure 11). The complement cascade forms pores in the cell membranes at the site of the cell surface antigen (Watson *et al.*, 1987), leading to cell destruction because osmotic pressure forces water to enter the cell. The observation that complement mediated cell killing could be maintained *in vitro* (Oda and Puck, 1960) lead to the technique being used to map several antigens to specific chromosomes (Bai *et al.*, 1982; Goodfellow *et al.*, 1975; Puck *et al.*, 1971). Much use has been made of this method in the literature for mapping specific antigens, but the technique has also been applied with equal success to cell selection (Puck *et al.*, 1971). One requirement for this method is that the antibody can be made species specific. This is greatly aided by the use of polyclonal antibodies, which may contain recognition sites for several epitopes of the antigen - in this case, even if reaction occurs in both the parental strains, the reaction with one strain can be drastically reduced by incubation of the antibody with a lawn of the cells prior to treatment of the hybrids. This selectively removes any antibodies from the mix which cross react. In the absence of a suitable polyclonal antibody, monoclonal antibodies may be used if they are raised in a species close to the target species (S. Povey, personal communication). Oda and Puck (1960) established conditions to carry out complement mediated cell death *in vitro* , which will be further explored in the results section. It has been shown that the primary effect of the antibody treatment is to drastically reduce cell growth (Oda and Puck, 1960), so even in the absence of absolute cell death, selection may be carried out on the basis of the rate of cell growth of the hybrid population. After a single exposure to antibody and complement, the effects are irreversible (Oda and Puck, 1960; Pious et al., 1973), this removes the need for expensive continuous culture in the presence of antibody.

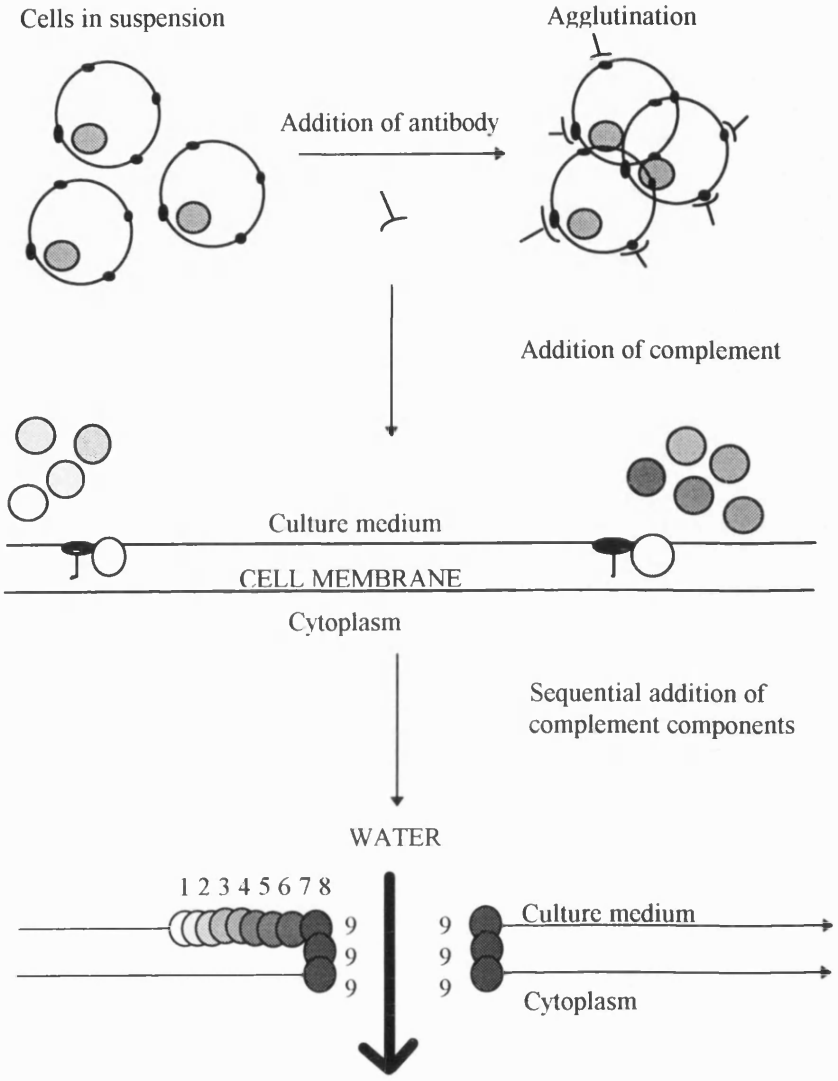


Figure 11: - Action of antibodies and complement on cells in culture

1.7.3) Characterization of somatic cell hybrids

The chromosome content of each hybrid clone must be assessed before it can be used for genetic analysis, also, the instability of the hybrid clones leads to the requirement for repeated characterization following each passage of cell growth Somatic cell hybrids may be characterized by karyotype analysis, isozyme analysis, or by polymerase chain reaction (PCR) techniques. Only characterization by DNA markers will be considered here in detail.

1.7.3a) Hybrid characterization by DNA markers

Somatic cell hybrids may also be characterized by correlating the presence of DNA markers or gene probes which have already been mapped to specific chromosomes with the chromosomes present in the hybrid clones (Creagan and Ruddle, 1975). This can be achieved by the detection of genes by using restriction fragment length variation (RFLV) analysis detected by Southern blotting or by species specific PCR analysis.

The use of the polymerase chain reaction (PCR) to characterize somatic cell hybrids has certain advantages over Southern blot or isozyme analysis. The procedure requires only small amounts of material, and is rapid and simple to perform. The technique involves the species-specific amplification of markers from specific chromosomes. PCR primers can be designed from untranslated regions of the gene, such as 5' and 3' untranslated regions, and intron sequences, to minimize cross reaction with the other parental species. Several markers can be typed per chromosome, which gives a rough estimation of the state of fragmentation of the chromosomes, provided not all fragments are retained. Panels of human specific (Abbott and Povey, 1991), and mouse specific (Abbott, 1992) PCR primers have been developed for use in human : mouse and mouse : Chinese hamster hybrid panels respectively. There also exists the possibility to test for more than one marker simultaneously in only one PCR reaction (multiplex PCR). This technique does, however, require alteration of the relative concentrations the reagents in the reaction (e.g. - primer concentration, $MgCl_2$ concentration, polymerase concentration) to obtain reproducible results, and also has the requirement that all primer pairs share similar annealing temperatures and ion concentration requirements. This technique has the disadvantages that it does not identify incidences of fragmentation if all fragments are retained in the hybrid cell, and also it cannot distinguish rearranged chromosomes from wild type chromosomes. The requirement for only small amounts of DNA

does however, mean that DNA supplies last longer, with the result that cells do not have to be cultured for as long, which reduces the chance of hybrid heterogeneity, and instability.

Hybrid clones can also be characterized by interspersed repetitive sequence PCR (IRS-PCR) - amplification of a particular chromosome, in a species-specific manner, produces a characteristic 'fingerprint'. Hybrid cells can be characterized by direct comparison of the profile produced by the hybrid, with positive controls (Bicknell *et al.*, 1991; Meese *et al.*, 1992; Benham *et al.*, 1992), or by the use of the resulting PCR products as FISH probes (Lichter *et al.*, 1990).

For complete characterization of new panels of somatic cell hybrids, or the recharacterization of clones following culture, a total estimate of the chromosomal content of each line can only be gained by a combination of PCR or isozyme analysis, and karyotypic analysis.

1.7.4) Types and uses of hybrid cells

Hybrid panels constructed as outlined above can be used to provide a rapid means of mapping new genes and DNA markers to specific chromosomes (Creagan and Ruddle, 1975; Kozak et al., 1975; Wilson *et al.*, 1987). These means are particularly useful, as they rely upon the availability of interspecies, rather than intraspecies,variation. New markers are tested on panels of somatic cell hybrids with defined chromosomal content, each marker is defined as being either '+' or '-' in each hybrid. The concordance and discordance of the new marker to each chromosome in the panel is calculated, the new marker will be seen to 'segregate' with a specific chromosome. The number of hybrid lines needed for such analysis is small - in ideal circumstances, given an optimal distribution of chromosomes in the clone panel, the number of separate lines is as low as five, for man and mouse, as this would produce 2^5 (32) individual combinations of chromosomes, enough to cover each of the chromosomes in man and mouse (Creagan and Ruddle, 1975). In most cases, more clones are needed, as the distribution of chromosomes in each clone is not ideal. Although many genes have been mapped in the mouse by interspecific backcross analysis, somatic cell genetics has allowed the positioning of some loci on the physical map. The alignment of the two maps will eventually give a complete picture of the distribution and organization of the mouse genome.

Somatic cell hybrids have also been widely used in the study of gene expression, by virtue of the fact that some control mechanism exists in hybrid cells to cause the cessation of differentiated function in hybrid cells formed by the fusion of a differentiated cell, and an undifferentiated cell. This phenomenon is known as extinction. Study of extinction in somatic cell hybrids can elucidate the processes which control gene expression in both man and mouse.

Parasexual methods of generating cell lines which originate from more than one source are not limited to the fusion of intact cells, hybrid cells can be made by 'fusing' cells from one species with naked DNA or isolated chromosomes from other species (Szybalska and Szybalski, 1962). Specialized 'microcell' hybrids can be constructed, which contain single chromosomes from one species on a background of chromosomes of another species (Fournier and Ruddle, 1977). One parental cell type can be treated with ionizing radiation prior to the fusion, yielding hybrids with only fragments from one species on a background of the other species (Goss and Harris, 1975). Finally, hybrid cells can be constructed from cell lines carrying chromosomal rearrangements (Kline et al., 1992). Each of these hybrid types has defined uses, which will be outlined later.

1.7.5) Somatic cell hybrids constructed with cell lines carrying translocations or deletions

Cell lines carrying chromosomal rearrangements in somatic fusion studies can be used to generate panels of somatic cell hybrids which can be used to define specific regions of the genome. For instance, a panel of somatic cell hybrids was constructed from individuals with specific rearrangements of chromosome 18 (Kline *et al.*, 1992). Cells from each individual were removed and transformed, and fused to a Chinese hamster cell line carrying a temperature sensitivity mutation (the cells were inviable at temperatures of 39 degrees or above, Cirullo *et al.*, 1983). The resulting hybrid clones were cytogenetically analyzed, and were shown to contain a range of deletions (Kline *et al.*, 1992).

Cell lines carrying reciprocal translocations to 17 or the X are useful for the creation of hybrid cells in which other chromosomes can be positively selected for, without the need for further specialized selection systems (Solomon *et al.*, 1976). Any chromosome involved in a translocation with a chromosome carrying a selectable marker can also be fixed in a hybrid population. These cell lines are of use in aligning the genetic and physical maps of particular chromosomes, as the genetic map is subject to interference from recombination hotspots and transmission distortion (see earlier sections). Genes can be assigned map positions with respect

to the chromosomal breakpoints involved in the translocation (as demonstrated by Searle *et al.*, 1979, in their studies of the colinearity of mouse chromosome two), this map can then be used to construct fixed reference points on the genetic map. Somatic cell hybrid panels constructed from translocation cell lines can be used as a means of defining small regions of the genome.

1.8) Repetitive elements in the mouse genome

The genome of higher eukaryotes is not comprised solely of coding DNA. Early DNA reassociation experiments have shown there to be three basic classes of DNA sequences, which have different frequencies and distributions in the genome, as well as different functions. DNA was disassociated, and the single strands were reassociated under appropriate conditions (Britten and Kohne, 1968). There is a highly repeated fraction, which reanneals the most rapidly, (Cot 1/2 = 0.03), comprised of the major and minor satellite units, a moderately repeated fraction comprised of numerous families of diverse repetitive elements and a single copy fraction which reanneals slowly (Cot 1/2 = 3000) comprised of the coding sequences.

1.8.1) The moderately repeated DNA fraction

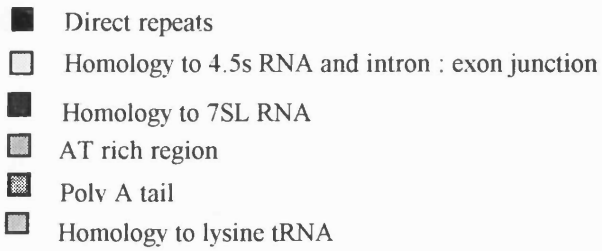
The moderately repeated fraction consists of the dispersed repetitive elements. These repeat units can vary from around 150 bp up to 7 kb. The shorter units are referred to as Short Interspersed Repetitive Elements (SINES), and the larger examples as Long Interspersed Repetitive Elements (LINES).

In the mouse, the major SINES are the B1 and B2 elements (present at 130,000 - 180,000 and 80,000 - 100,000 copies per genome respectively) and the major LINE is the L1 element (truncated forms are present at 100,000 copies per genome, full length copies at 10,000 copies per genome). It has been postulated (Jagadeeswaran *et al.*, 1981; Sharp, 1983) that these repeats arose by reverse transcription of RNA species followed by insertion into the genome. The B1 repeat is homologous to parts of the 7SL RNA, whilst the B2 repeat is similar to the lysine tRNA. Figure 12 illustrates the organization of the mouse B1, B2 and L1 repeats.

It is now also generally accepted that LINE elements derive from retroposons. It would appear that not all of the LINE elements in the mouse genome have the ability to retrotranspose, as reflected by the small number of sequence variants in the family. It has been postulated that only a small fraction have this ability (Edgell, 1994).

The B1 repeat

The B2 repeat



L1md repeat (Taken from Hastie, 1989)

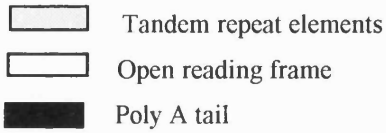


Figure 12: - Dispersed repetitive elements in the mouse genome

1.8.1a) The distribution of SINES and LINES in the genome

The composition of chromosomes is not uniform along their length. Differences exist in base composition (De la Chapelle *et al.*, 1973), replication timing (Goldman *et al.*, 1984; Somssich *et al.*, 1981), and the distribution of genes (Saccone *et al.*, 1992) and CpG islands (Burminster *et al.*, 1988). This heterogeneity of chromosome structure is manifested as chromosome banding patterns as visualized by several techniques, such as G, C, R, Q and T banding (Comings, 1978; Bickmore and Sumner, 1989; Craig and Bickmore, 1993). Chromosome bands are a manifestation of underlying structure differences along the chromosome. Banding techniques which rely on the same aspect of chromosome structure, such as Giemsa and Quinacrine banding, produce almost identical banding patterns. In addition, meiotic chromosomes also show transverse banding patterns (Ambros and Sumner, 1987). The number of bands produced is dependent on the degree of condensation of the chromosomes. The number of bands varies from around 300 to 800 bands as visualized on condensed metaphase chromosomes to around 2000 bands as seen on elongated prophase chromosomes (Yunis, 1981).

The segregation of SINES and LINES to different chromosome bands has been noted in the mouse genome (Soriano *et al.*, 1983; Boyle *et al.*, 1990). In situ hybridization of L1, B1 and B2 probes to mouse metaphase spreads reveals that the LINE sequences are present in the A+T rich, late replicating, light isochore fraction, producing a banding pattern highly suggestive of that produced by Giemsa (G) banding. There are exceptions, as reflected by the X chromosome, which exhibits no clear cut banding, showing heavy hybridization along its length (Boyle *et al.*, 1990). In this case, the exclusion of LINE sequences is less obvious, and it may prove that at least in the case of the X, SINE and LINE sequences are not always entirely mutually exclusive (Boyle *et al.*, 1990). The existence of an L1 sequence 'island' has also been reported on the mouse X chromosome (Nasir *et al.*, 1990; 1991). The intensities of the banding pattern produced by the L1 probe are also increased on mouse chromosome 7, indicating a high number of LINE sequences on this chromosome (Boyle *et al.*, 1990). The pattern produced by hybridization of the B2 probe to mouse metaphase spreads is very similar to that produced by Reverse (R) banding, except that chromosome 11 shows extreme banding, indicating that it is a rich source of B2 sequences (Boyle *et al.*, 1990). The pattern produced by the B1 probe is also reflective of R banding, although this is less clear-cut than the hybridization pattern for the B2 repeat. This may indicate that the distribution of the B1 repeat is closer to random than that of either the L1 or the B2 repeat (Boyle *et al.*, 1990).

1.8.2) Simple sequence repeats

There exists another distinct class of repeats in the mouse genome. These are composed of tandemly repeated tracts of two, three or four nucleotides. These are termed Simple Sequence Repeats (SSRs) or microsatellite repeats. This type of repeat was first identified by Miesfeld *et al.* (1981), when a tract of alternating G and T nucleotides was discovered in a region of mouse DNA that hybridised to the β Globin gene cluster. They carried out studies to determine if this sequence was conserved in other species. DNA hybridisation experiments revealed that similar repeats were present in the DNA of species as widely diverse as human, mouse, pigeon, *Xenopus* and *Drosophila*. Dinucleotide repeats have been shown to form the Z-DNA conformation in vitro under certain conditions (Hamada *et al.*, 1982). and to form four stranded complexes, even in the absence of any exogenous protein. The formation of these structures manifests as a decrease in mobility during electrophoresis. The structure of the complex was determined by electron microscopy (Gaillard and Strauss, 1994). Trinucleotide repeats have also been shown to be a preferential site for nucleosome assembly in the myotonic dystrophy gene. It has been suggested that tracts of CAG repeats may serve to repress transcription (Wang *et al.*, 1994).

1.8.2a) The evolution and distribution of Simple sequence repeats

Stallings *et al.* (1991) reviewed the distribution and evolution of (GT)n repeats in mammalian genomes. This confirmed the finding that simple sequence repeats were common in the genome of all eukaryotic cells, although they are more common in the genomes of mouse and human (100,000 copies and 50,000 copies per genome respectively) than they are in the genome of lower eukaryotes such as the yeast *Saccharomyces cerevisiae* (of the order of 100 copies). This may reflect the greater complexity of the genomes of the higher organisms. The group then carried out studies on the distribution of (CA)n repeats in approximately 3,700 human specific cosmids from a mouse : human somatic cell hybrid containing human chromosome 16. A combination of grid and Southern blot analysis led them to conclude that 63% of the human clones contain a (CA)n repeat (from this they calculated the average spacing of (CA)n repeats in the human genome as one (CA)n repeat approximately every 30 kb). They were able to estimate the percentage for the mouse genome also from this library, as they had access to a large number of mouse clones by virtue of the fact that the somatic cell

hybrid was on a mouse background. The percentage was found to be significantly higher for the mouse, 77.8% on grid hybridization alone. Human clones that were positive for heterochromatin were found to contain fewer repeats. From the results of Stallings *et al.*(1991), it was concluded that the average spacing of repeats in the human genome was one repeat every 28 kb, a figure comparing favourably with the experimental data. The figure for mice was estimated as one repeat every 18 kb, confirming the observation that (CA)n repeats are more common in the murine genome than they are in the human genome. Examination of the distribution and frequency of different types of trinucleotide repeat in the mouse genome reveals that the distribution of these repeats is not entirely random. Analysis of the positions of trinucleotide repeats in sequences submitted to the GenBank database show that in the mouse, (CAG)n and (CCA)n repeats were the most common (Stallings, 1994). (TAC)n, (GGC)n and (ATG)n repeats were found to be rare in the mouse genome (Stallings, 1994). Striking differences in the distributions of these repeats were also revealed. Despite their high frequency in the mouse genome, (CAG)n repeats were found to be entirely excluded from intron sequences, in a strand specific manner, as no reduction in the numbers of (CTG)n repeats (the opposite strand version of (CAG)n) was noted in introns. This implies that (CAG)n repeats are underrepresented in the hnRNA population of the cells (Stallings, 1994). The lack of (CAG)n repeats in intron sequences might be explained by the similarity of this repeat unit to the 3' consensus splice site CAGG. Some exclusion of (GGC)n repeats from intron sequences was also noted, although these repeats did prove to be biased towards 5' untranslated sequences (Stallings, 1994). No exclusion of any other type of repeat from intron sequences was noted. It appears that in the mouse, A rich simple sequence repeats occur commonly as tandem variations in the poly-A tract of SINE repeats. An association between SINE repeats and a proportion of (CA)n microsatellites has been noted in the pig (Wilke *et al.*, 1994) and the rat (Beckman and Weber, 1992). In man, no apparent bias of repeats to any particular region was found (Stallings *et al.*, 1991). The distribution and frequency of trinucleotide and tetranucleotide repeats in humans has also been studied by Edwards *et al.* (1991). The group studied the trinucleotide repeats (AAT)n and (CAG)n and the tetranucleotide repeats (AAGG)n, (AATG)n, (ACTC)n, (ACAG)n and (AGAT)n in a sample of unrelated blood donors representing four ethnic groups. It was concluded that of the classes tested, repeats are found in all regions of the genome, although tetranucleotide repeats are not found in coding regions. Analysis of the database suggests an overall frequency of 400,000 copies per genome, or one repeat every 10 kb for all possible forms of trinucleotide and tetranucleotide repeat in man. The study of human microsatellites reveals that in closely related species such as man and chimp some conservation of the location of the repeats was seen (Stallings *et al*, 1991). No

conservation of position has been seen between mouse and man, or mouse and rat.(Stallings *et al*, 1991).

Tautz and Renz (1984) carried out similar surveys on other forms of simple sequence repeats, (CA)n, (GA)n (A)n, and (G)n. DNA from diverse organisms was tested for hybridisation to probes corresponding to these repeat sequences, and also to a probe for the trinucleotide repeat (CAG)n. It was found that all these simple sequence repeats were represented to varying degrees in the genomes of all organisms studied (human, *Drosophila*, sea urchin, yeast, and the protozoan *Stylonychia*).

1.8.2b) Some Simple Sequence repeats have non random chromosomal distributions

Some simple sequence repeat units appear to be confined to specific regions of the chromosome. Nanda *et al.* (1990) determined the distribution of the repeats (GACA)n, (TCC)n, (CAC)n and (GATA)n. In situ hybridisation experiments were carried out with biotinylated probes to determine the organisation of the repeats on the chromosome. Surprising heterogeneities in the distribution of (GACA)n were found, in humans. This repeat appeared to map to the D and G chromosomes, the position of the nucleolus organiser regions, as well as to dispersed regions of the chromosomes. The hybridization sites corresponded with the rDNA positions of the NOR sites in a variety of primate species. Similar experiments were also carried out in the mouse. There are many (GACA)n repeats on the mouse Y chromosome (in contrast to humans) and mouse chromosome 17. A regular male specific pattern of hybridisation is observed with both the (GATA)n and(GACA)n probes for C3H mice, but this is not a constant feature in all mouse strains. There appears , however to be no hybridisation to the NOR bearing chromosomes. The hybridisation profile for the trimeric repeats appears to be more or less random (Nanda *et al.*, 1990).

1.8.2c) Simple sequence repeats are highly polymorphic

Shortly after it became known that simple sequence repeats were a ubiquitous feature of eukaryotic genomes, it was suggested that they might exhibit hypervariability. Levinson and Gutman (1987) suggested that slipped-strand mispairing might be a mechanism for generating polymorphism. This involves the faulty alignment of homologous sequences containing repetitive elements in the genome during DNA replication, to produce copies that have additional repeat units, and copies that are deleted by a corresponding amount. Evidence for this has been found by looking at the behaviour of synthetic oligomers, and also by observation of tandemly repeated moieties in the bacterial genome. Levinson and Gutman (1987) carried out experiments to examine frameshift mutations occurring in (CA)n repeat tract. These experiments determined that each frameshift was an independent effect. In all cases, these frameshifts originated from the loss or gain of an integral number of nucleotides (i.e. by two, or a multiple of two nucleotides). The distribution of frameshifts was seen to be biased towards the deletion or duplication of a single (CA)n repeat. The occurrence of frameshifts was seen to be a function of the length of the repeat unit - longer repeats appear to be rearranged more frequently (Levinson and Gutman, 1987).

Tautz (1989) examined whether simple sequence repeats were hypervariable. The species studied were human, *Drosophila melanogaster* and whale. Specific primers were designed which flanked the (CAG)n repeat in the *Drosophila Notch* gene, the (CA)n repeat in the intergenic region of the human δ and β globin genes, and a randomly isolated DNA clone containing a (CT)n dinucleotide stretch from the long finned pilot whale *Globicephala malaena*. Eleven independent *Drosophila* lines were analysed, and hypervarability of the repeat unit was observed. In humans, individuals from a three generation family, as well as unrelated individuals were tested. Hypervariability was observed, and observation of the allele distribution patterns in the family proved that the alleles were inherited in Mendelian fashion. The results from the long finned pilot whale proved that random cloned sections are also hypervariable.

Litt and Luty (1989) examined a (CA)n repeat in the fourth intron of the human cardiac actin gene. DNA was taken from 37 unrelated individuals, and also in three families with a total of 24 children. Again, hypervariability was observed. A total of 12 different alleles

was detected, and Mendelian co-dominant inheritance of five alleles was seen in the families. All allelic fragments contained even numbers of added or deleted nucleotides. It was postulated that the mechanism of polymorphism may be by unequal exchange during meiosis, or by slippage during DNA replication. Investigations of the variability of (CA)n repeats in the human genome were also carried out by Weber and May (1989). The frequency of(CA)n repeat tracts in the genome was estimated from the Genbank database. (CA)n repeat containing fragments were also specifically cloned from a chromosome 19 specific library, in order to determine the degree of variation of these repeats in the population. The number and sizes of alleles was determined by analysis of DNAs from many ethnic groups. Estimates of Polymorphism information content (PIC) were calculated. This was shown to vary between repeat elements, from values of 0.31 to 0.79, averaging around 0.55. Again, analysis of four three generation families showed these alleles to be inherited in a stable co-dominant fashion, concordant with Mendelian laws. To determine whether trinucleotide and tetranucleotide repeats also exhibited hypervariability, Edwards, *et al.* (1991) extracted DNA from blood samples taken from individuals of four large ethnic groups. Genotypes for the four groups were constructed from multiplex PCR with markers for a variety of tri and tetranucleotide repeats. Eighteen simple sequence repeat loci were analysed. Over half of these loci were found to be polymorphic, and for these, stable Mendelian inheritance was observed. There were marked differences between the four ethnic groups (White, Black, Hispanic and Asian) in heterozygosities and allele frequencies. Recently, it has been observed that the GC rich trinucleotide repeats are more variable than other microsatellite repeats. The stability of the repeat seems to be correlated with the repeat length, as small repeats are not as unstable. There are four basic forms of these repeats (Kuhl and Caskey, 1993). The trinucleotides with the smallest copy number are not variable. There is a second class of repeats of slightly longer length which are polymorphic but are stable between generations. The third class is those still larger repeats which are more variable. The fourth class includes repeats with vastly amplified copy numbers. Some of these repeats change in length by passing through a single meiosis. In certain human diseases this instability affects mitoses as well as meioses, resulting in individuals with different allele lengths at these loci in different tissue types, and even in different cells of the same tissue type. The cause of the unstability is not known.

Abbott and Chambers (1994) analysed (CAG)n trinucleotide repeats for evidence of repeat expansion in the mouse. This has been shown to be a new form of disease causing mutation in man. PCR primers were designed, which flanked sites known to contain (CAG)n repeats. The primers were then used to amplify DNA from different strains of mice. No

evidence for expansion of trinucleotide repeats, as indicated by 'smearing' of PCR products, was noted.

The mechanism by which microsatellite length variation arises is unclear. Three mechanisms have been suggested : unequal sister chromatid exchange, slippage during replication, and gene conversion events (Smith, 1976; Schlotterer and Tautz, 1990; Tchida and Iizuka, 1992). It has been shown that unequal exchange does not provide a strong mechanism for the generation of diversity in other forms of tandem repeats, such as minisatellites (Jeffreys *et al.*, 1994). Unequal exchange would be expected to produce equal numbers of deleted and duplicated alleles in minisatellites. However, an excess of duplication events is noted (Jeffreys *et al.*, 1994). Further evidence which suggests the generation of polymorphism in minisatellite repeats is not due to unequal exchange is demonstrated by the observation that polymorphic loci flanking repeat units show no evidence of unequal exchange (Jeffreys *et al.*, 1994). The model of Stephan and Cho (1994), for the generation of length variation in microsatellites, suggests that repeat length is influenced by the recombination rate. Recombination dependent mechanisms, such as unequal exchange, would be expected to produce long, heterogeneous repeats in areas of low recombination; however, there is some evidence to suggest that microsatellites may occur with equal frequency in areas of both high and low recombination, as reflected by their more or less random distribution (Stallings *et al.*, 1991). However, since there is so little information on areas of high and low recombination independent of microsatellite distribution, this is difficult to prove. Repeat units are often interrupted by other microsatellite sequences, or by random sequences (Stephan and Cho, 1994). This suggests that long range forces, such as unequal exchange, leading to homogeneity of repeat unit, are not acting on microsatellites. In addition, unequal crossing over is expected to occur with low frequency in members of families which are interspersed with other sequences (Dover, 1982). Slipped strand mispairing mechanisms are more likely to influence the structure of microsatellites. Experiments show that microsatellites evolve and change in length by replication slippage, when amplified in vitro, producing repeats of analogous structure and characteristics to those seen in vivo (Schlotterer and Tautz, 1990). Variation in simple sequence repeats may also be due to a similar mechanism to that which is occuring in the evolution of minisatellite repeats. This is more likely to be due to gene conversion events (Jeffreys *et al.*, 1994; Richards and Sutherland, 1994).

1.8.2d) The use of microsatellite based markers for gene cloning

The polymorphic nature of simple sequence repeats makes them excellent genetic markers. Microsatellite markers have been of much use in the cloning of both polygenic and single gene disorders. The use of PCR based microsatellite markers has provided a useful tool to investigate diseases under the control of many genes. Before microsatellite markers were widely available, the marker density of the genetic map was such that linkage of polygenic conditions to specific chromosomal regions was fairly difficult. The large numbers of microsatellite markers has allowed the analysis of polygenic conditions by exclusion mapping. This technique was recently used to analyze the polygenic control of autoimmune diabetes in the mouse (Ghosh *et al.*, 1993). Genes involved in this condition in the mouse have been associated with five chromosomes (1, 3, 9, 11 and17). A large of microsatellite based markers were analyzed for linkage to this disorder by the use of backcrosses, and allowed the exclusion of the entire mouse genome, except for small regions on chromosomes 1, 3, 9, 11 and 17. A total of ten loci on these chromosomes have been associated with the inheritance of diabetes in non-obese mice (Ghosh *et al.*, 1993).

1.9) The development of new molecular markers for mouse chromosome two

In 1991, when the study was commenced, the genetic map of mouse chromosome two consisted of a very limited number of molecular markers, clustered mostly in the central portion of the chromosome. The distal end of the chromosome containing ragged and wasted was especially underrepresented. The correlation of the physical map with the genetic map was also poorly defined because of the sparsity of markers. Given the state of affairs, it would be impractical to contemplate positional cloning of mutant genes. This was particularly true for ragged and wasted as there was no closely linked distal flanking marker. Cloning by candidate gene approaches was similarly hindered because of this lack of molecular markers. If the marker density of the map were increased, it would raise increase the probability that mutant genes could be cloned by positional cloning techniques. Similarly, an increase in the number of molecular markers would increase the value of using irradiation fusion gene transfer (IFGT) hybrid panels to define the physical map of the chromosome. Once characterized, these IFGT panels could be used as a source of additional markers for specific regions of the chromosome.

Thus, in the first instance, the aim of the project was the generation of new molecular markers for the whole of the chromosome, with a specific focus on the generation of markers mapping to distal mouse chromosome two.

A number of methods have been devised for generating chromosome specific libraries. These include: the use of somatic cell hybrids (described in detail below); flow sorted chromosomes (an approach which has had only limited success because of the uniformity in of mouse chromosome size); and microdissection of specific chromosomes or chromosome regions.

1.9.1) Isolation of chromosome specific libraries using somatic cell hybrids

As interspecific somatic cell hybrids tend to lose chromosomes from one or other of the two species used in the construction of the hybrids, it is possible to establish cell lines that retail single chromosomes of one species, on a genetic background of the other species. A number of techniques can be used to isolate either chromosomal specific clones or segments.

1.9.1a) Hybridization based methods for isolating chromosome specific clones

New DNA markers for specific chromosomes or chromosome regions can be generated by the construction of genomic libraries from sources such as monochromosomal somatic cell hybrids, IFGT hybrids, or microcell hybrids. For the generation of new chromosome specific mouse markers, libraries can be made from mouse:hamster monochromosomal hybrid lines. Because the hybrid cells contain the entire complement of hamster chromosomes in addition to the single mouse chromosome, the majority of clones produced will be derived from the hamster genome. The mouse clones can be identified using abundant species specific repetitive elements as probe (e.g. specifically designed oligonucleotides, or total mouse DNA with an appropriate competitor). Once identified, these mouse specific genomic clones can be screened with tissue specific cDNA libraries to ascertain whether they encode functional genes (Hochgeshwender *et al.*, 1989).

1.9.1b)Interspersed repetitive sequence polymerase chain reaction (IRS-PCR)

This technique allows the isolation of DNA fragments in a species specific manner, from the amplification of sequences between interspersed repetitive elements. These repetitive elements must be positioned in opposite orientations, within a suitable distance for PCR amplification (Nelson *et al.*, 1989). No knowledge of the DNA sequences located between the repeats is necessary for their isolation (see figure 13). PCR primers are designed from regions of the repeat unit showing the least interspecies conservation. Under stringent conditions, amplification of DNA from one species only is obtained. DNA from complex sources produces a smear of bands, but if the simplicity of the template is increased, by the used of prior restriction digestion of the template DNA (restricted *Alu*-PCR) (Guzetta *et al*, 1991), or by the use of hybrid clones containing limited numbers of chromosomes from one species (Cotter *et al.*, 1990; Herman *et al.*, 1991), the amplification yields a number of discrete bands, which can be isolated and cloned. Preliminary chromosomal locations can be assigned to IRS-PCR products by 'fingerprinting' experiments (i.e., the amplification of DNAs from a hybrid containing an intact chromosome or a hybrid containing only fragments of that chromosome). This will either confirm or exclude the location of the PCR product from that region dependent on the presence or absence of a shared band (Nelson *et al.*, 1989).

IRS-PCR primers have been developed to utilize the presence of *Alu* (Nelson *et al.*, 1989), and LINE repeats (Ledbetter *et al.*, 1990) in man. Primers have also been developed for

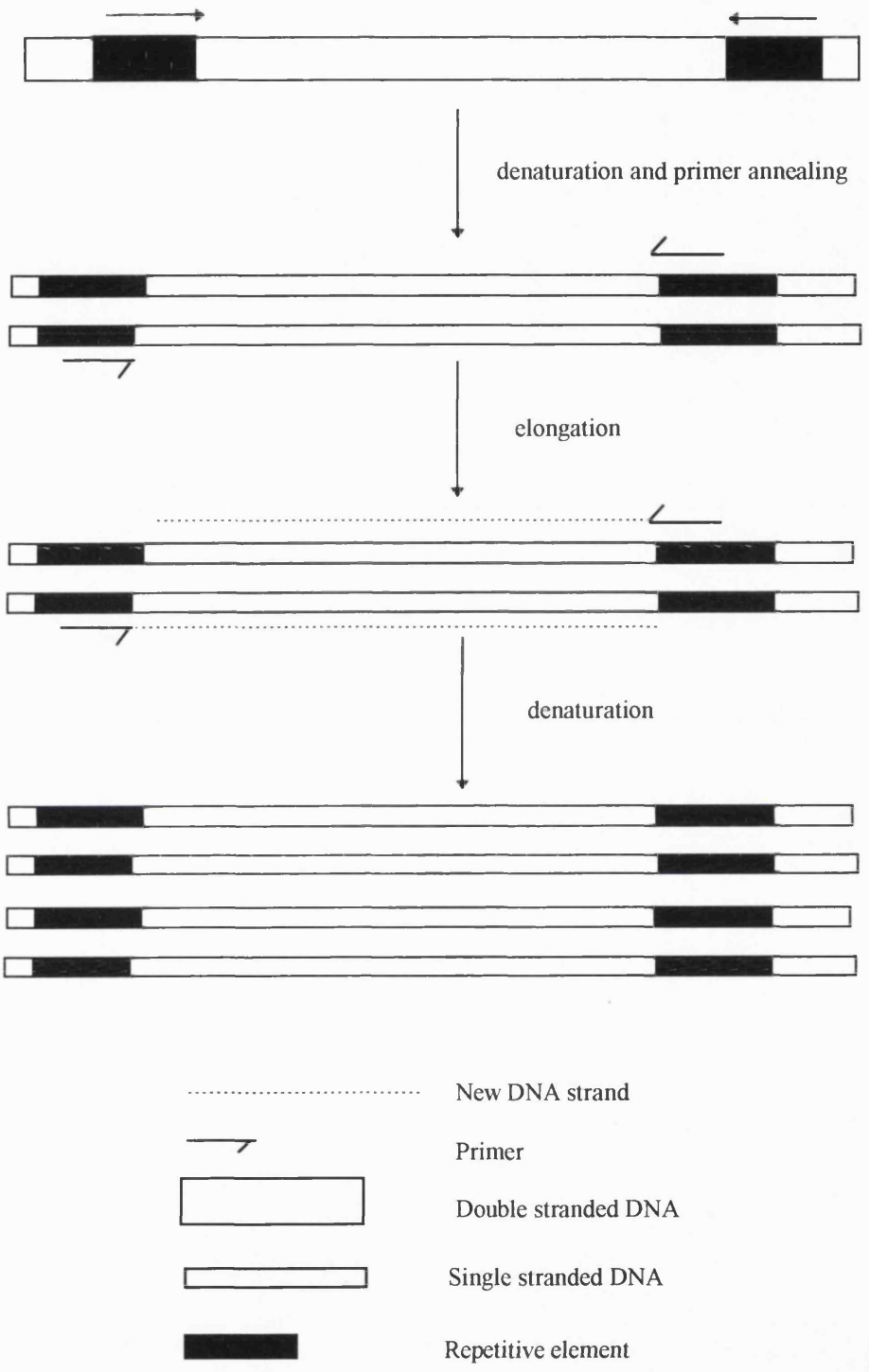


Figure 13: - Mechanism of IRS-PCR

the 3' ends of the B1 and B2 repeats (Simmler et al., 1991), and the 3' end of the L1md repeat (Irving and Brown, 1990) in the mouse. The 3' end is often the position of choice for the design of primers, as it is constant within species, but shows considerable interspecies variation, allowing species specific amplification (Irving and Brown, 1990).

1.9.1c) Interspersed repetitive sequence bubble PCR

One of the drawbacks of IRS-PCR is the requirement of repeat sequences to conform to certain conditions for amplification to occur - the repeat elements must be in the correct orientation, and also within an amplifiable distance on the chromosome. However, there is considerable evidence that the distribution of these repeats is not random on chromosomes - most SINE and LINE repeats are confined to different chromosome bands (Korenberg and Rykowski, 1988; Boyle *et al.*, 1990). In addition, evidence is also available which suggests that the orientation of these repeat units is not random - the orientation is strongly biased in favor of tandem arrays of repeat elements, particularly in cases where the repeat units are in close proximity to each other (Moyzis *et al.*, 1989). This leads to underrepresentation of certain regions of the genome in panels of markers generated by IRS PCR. Also, this reduces the total number of new markers that can be generated from a given chromosomal region by conventional IRS-PCR.

A recent technique which can circumvent some of these problems is interspersed repetitive element bubble PCR (IRE-bubble PCR). This technology is an adaptation of a technique originally developed for the isolation of end probes from yeast artificial chromosome clones (Riley *et al.*, 1990). Genomic DNA is first digested with a restriction enzyme, to generate fragments in the correct size range. Digested DNA is then ligated to a specialized 'vectorette' composed of two 50 bp oligonucleotides which have complementary 3' and 5' ends, but an uncomplementary center section. This yields a 'bubble' structure. Amplification is then carried out using one primer specific to a repeat element, and another specific to the 'bubble' section of the vectorette (see figure 14). Amplified fragments can be excised from the vectorette by digestion at the inbuilt restriction site, and cloning into a suitable plasmid for propagation and sequencing (Hunter *et al.*, 1993; Munroe *et al.*, 1994). This technique allows the cloning of sequences flanking all repeat elements, regardless of orientation or distribution.

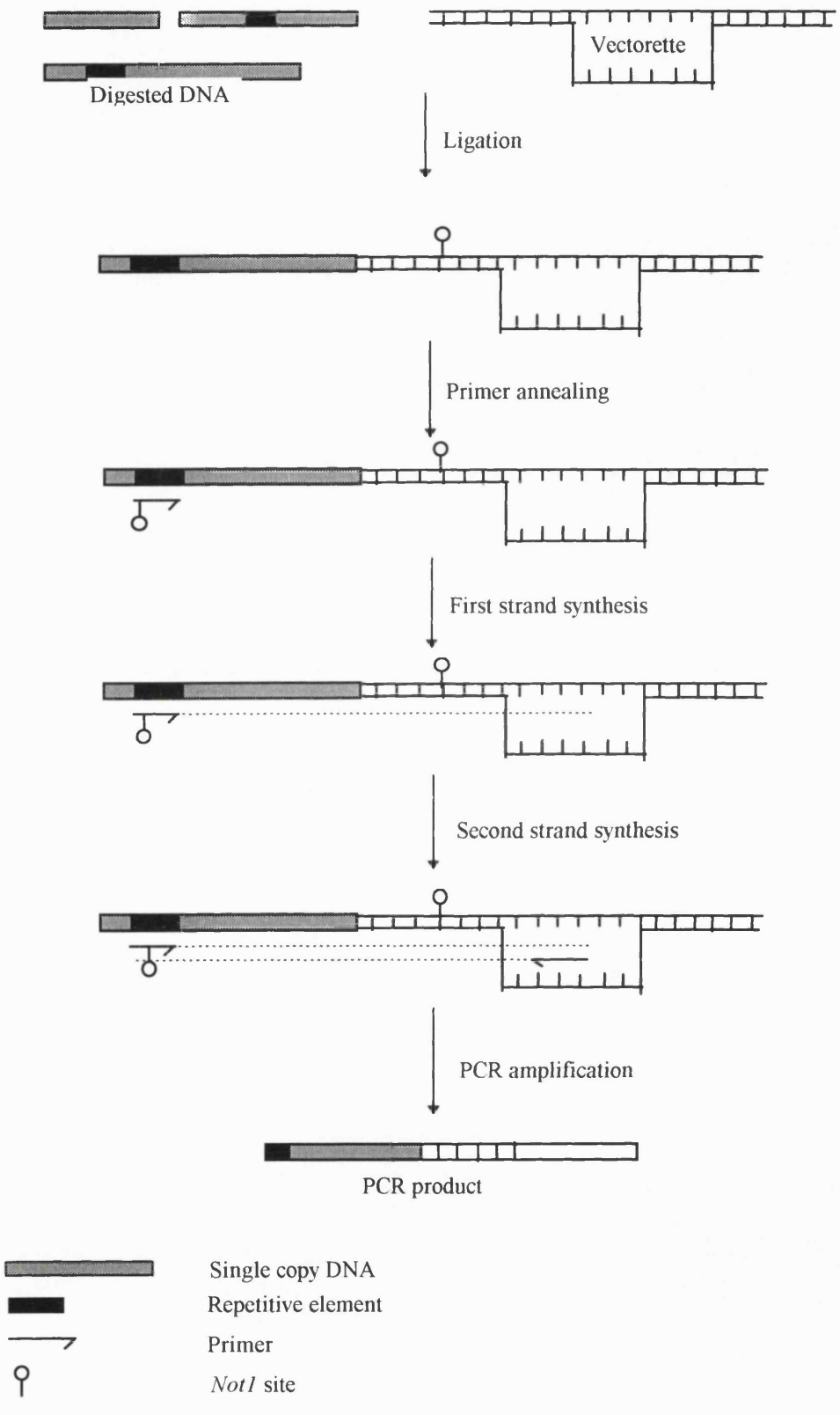


Figure 14: - IRS-bubble PCR

1.9.1d) Mapping of IRS-PCR and IRS-bubble PCR

New markers produced by these methods can in some cases be mapped directly by Southern blot analysis of backcross animals in the mouse, in the absence of internal repetitive segments. Polymorphism is generated by differences in the arrangement and distribution of interspersed repetitive elements on the *Mus musculus* and *Mus spretus* chromosomes (Cox *et al.*, 1991; Rikke and Hardies, 1991). The hybridization must be carried out in the presence of unlabelled mouse DNA, to reduce the signal due to the repetitive nature of the ends of fragments.

Alternatively, IRS-PCR products can be cloned, and sequenced. This technique removes problems caused by the repetitive nature of the clones, and allows screening for microsatellite sequences, and putative CpG islands. PCR primers can be developed to identify variation between species either from length differences of internal microsatellites (Cornall *et al.*, 1991), restriction fragment length variants (RFLVs) (Cole *et al.*, 1991), or the identification of single base pair changes by single stranded conformation polymorphisms (SSCPs) (Spinardi *et al.*, 1991). IRS-PCR products can also be mapped by direct visualization of their location by fluorescent in situ hybridization (FISH) analysis (Lichter *et al.*, 1990; Monaco *et al.*, 1991).

1.9.2) Differences in the organization of clones produced by IRS-PCR, IRS-bubble PCR and hybridization based techniques

The methods outlined above generate fragments which have different characteristics. Markers generated by the hybridization based method of Hochgeshwender (1989) (see section 1.9.2a) are identified by their repetitive content, and have a high content of interspersed repeats, such as L1 and B2. Markers identified by hybridization based techniques are usually of the order of 6 or 7 kilobases (although this is dependent on the choice of vector), and so should also contain fairly large amounts of single copy DNA. Markers generated by IRS-PCR or IRS-bubble PCR are smaller, usually ranging up to approximately 3 kilobases. The size of the fragments produced is dependent on the separation of the repeat elements in the genome, and on the limits imposed by the amplification. IRS-bubble PCR fragments are smaller than those produced by IRS-PCR (Hunter *et al.*, 1993; Munroe *et al.*, 1994), because only one repetitive element is needed. The size of markers produced by IRS-bubble PCR is also dependent upon the fragment size produced by restriction digestion, and the position of the

repeat unit in relation to the end of the fragment. Fragments produced by the three methods also exhibit differences in the types and distributions of repetitive elements. Although the fragments produced by hybridization based techniques also contain large amounts of repetitive DNA, the fragments produced by IRS-PCR and IRS-bubble PCR contain, by nature, highly repetitive ends, as well as any internal repeats. The markers produce by IRS-bubble PCR might be expected to give fewer problems with repetitive sequences than IRS-PCR clones, since only one end is repetitive.

The distribution of fragments produced on the chromosome may also be different across the techniques, because of the differences in the distribution of the different repeat elements. Due to the clustering of different types of interspersed repeats in different chromosome bands (Korenberg and Rykowski, 1988; Boyle *et al.*, 1990), fragments derived from L1 and B2 amplifications have a non-random distribution, whereas the fragments derived from hybridization based experiments may be expected to show a more random distribution, because they are derived from all types of interspersed repetitive element (when the distribution of L1 and B2 repeats is viewed together, their distribution is more or less random). If markers produced by IRS-PCR or IRS-bubble PCR can be mapped by their direct use as probes on Southern blots of interspecific backcross DNAs, this may circumvent problems with areas of the genome which are difficult to clone (which will be underrepresented in markers isolated from genomic libraries).

IRS-PCR and IRS-bubble PCR produces pools of fragments, which need to be subcloned, or isolated by other means in order for them to be sequenced. Markers produced by hybridization based methods are large, and may require subcloning to reduce the size of the insert, but they are composed of single fragments, except in cases of chimerism.

1.9.3) Experimental strategy

For the purposes of this study, new molecular markers were generated by two methods - I) IRS-PCR to allow species specific amplification of mouse fragments from a series of monochromosomal and microcell hybrids (Nelson *et al.*, 1989) and II) IRS-bubble PCR techniques to isolate fragments from regions of the genome in which IRS-PCR is not effective (Hunter *et al.*, 1993; Munroe *et al.*, 1994).

The choice of cell lines was based upon the fact that at the time of commencement of the study, no clones were available which contained smaller fragments of mouse chromosome two, such as irradiation-fusion gene transfer (IFGT) hybrid panels. The relative paucity of molecular markers with which to characterize new clones developed by this process, ruled out the construction of such a panel at this stage. The addition of large numbers of DNA markers to the map by the MIT laboratories has now made the construction of IFGT panels a more attractive proposition. The most appropriate cell lines were :-

a) the EBS18Az cell line (the kind gift of Dr. P. Lalley), which is essentially monochromosomal for the whole of mouse chromosome two, with the exception of a small fragment of proximal chromosome 15 (P. Lalley, personal communication) and a fragment of the Y chromosome which is undetectable by cytogenetic methods (A. J. Pilz, personal communication).

b) the microcell hybrid lines ABm 5, 7 and 11 (the kind gift of Dr. K. Fournier). These hybrids contain only mouse chromosome two. ABm 7 contains the proximal region of mouse chromosome two only, as determined by cytogenetic analysis (Fournier and Moran, 1980) and PCR analysis (N. Ray, personal communication). ABm 5 and 11 contain the majority of mouse chromosome two, although there may be some fragmentation and rearrangement (Fournier and Moran, 1980; C. Abbott and A. Pilz, personal communication).

New molecular markers were produced from the above hybrid lines, by either IRS-PCR or IRS-bubble PCR, and mapped by interspecific backcross, following the design of PCR primers. The efficiency of cloning inter-repeat sequences directly from pooled IRS-PCR products was compared with methods requiring their prior isolation and reamplification (see figure 15). Different methods of detecting variation (microsatellite analysis, RFLV analysis and SSCP analysis) between *Mus musculus* and *Mus spretus* were compared, and also any differences in the distributions of markers deriving from L1, B2 or L1/B2 IRS-PCR, and IRS-bubble PCR with a B2 primer were evaluated.

1.9.4) The generation of two new panels of somatic cell hybrids carrying different translocations of mouse chromosome two.

Somatic cell hybrids constructed from cell lines carrying reciprocal translocations have many uses in genetics. These hybrid cell lines allow the physical separation of two halves of a

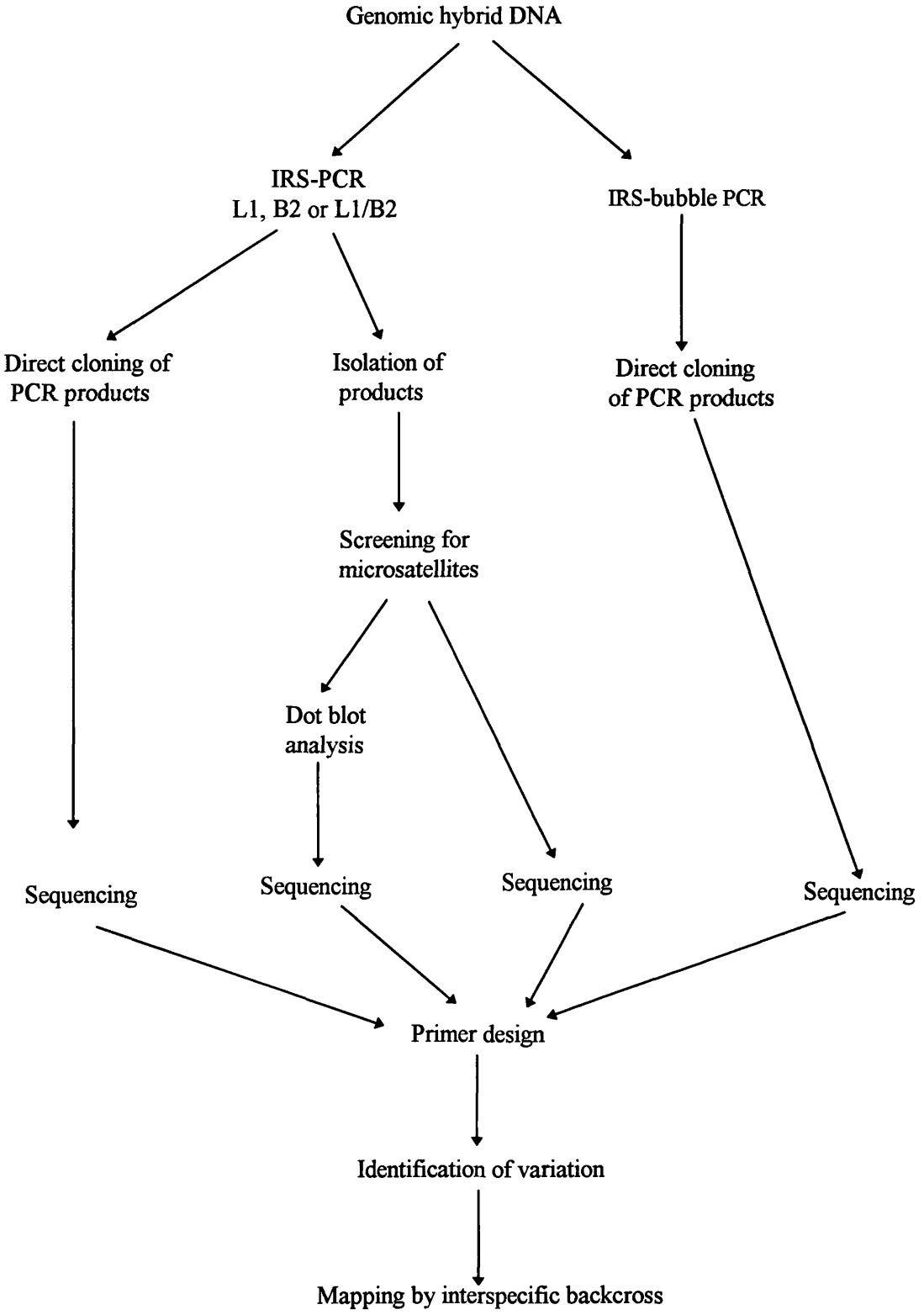


Figure 15:- Schematic diagram of experimental strategy

specific chromosome into different cells. This has uses in the alignment of genetic and physical maps, as outlined below, and also in the physical isolation of chromosomal regions. For example, the distal region of mouse chromosome two is the location of several developmental genes of interest, including ragged (*Ra*), wasted (*wst*) and lethal spotting (*ls*). The construction of these hybrid lines would allow the preliminary location of markers, without the need for lengthy interspecific backcross analysis, which need therefore be carried out only on markers which map to the required region, on the basis of analysis of the hybrid clones. These hybrid panels would in addition provide a valuable resource for the isolation of new markers from this region - for instance, markers generated from this region by IRS-PCR would give a shared band from amplification of clones containing the short translocation product of T2Wa, and the long translocation product of T28H (see figure 16). The construction of hybrid panels would also make possible the alignment of the genetic and physical maps of mouse chromosome two in this region, with the aim of defining more precisely the location of ragged and wasted with respect to the translocation breakpoints. Finally, the construction of these hybrid panels would provide a means of examining the distal imprinting region of mouse chromosome two, which would facilitate the identification of the gene or genes involved.

Two panels of somatic cell hybrids were constructed, one from spleen cells of mice carrying the T2Wa translocation, and one from spleen cells of mice carrying the T28H translocation, on a common background of the Chinese hamster cell line V79TOR2 (the kind gift of Dr. J. Thacker). Stability of hybrid clones has been seen to be a problem with somatic cell hybrids, especially in long term culture (Weiss and Green, 1966). Mouse : Chinese hamster somatic cell hybrids have been shown to be unstable and prone to fragmentation in our laboratory (A. J. Pilz, personal communication), in hybrid lines constructed from translocation carrier mouse parent cells fused to the Chinese hamster cell line A23. The cell line V79TOR2 (Thacker, 1980) has been shown to confer some stability to mouse : hamster hybrid lines constructed from it (Y. Boyd, unpublished results). Therefore, this was the cell line of choice for the Chinese hamster parent. Individual clones from the fusion were grown to large scale culture, and characterized by PCR using the primer pairs of Abbott (1992) and Dietrich *et al.* (1992).

As an additional strategy, in the event of inadequate segregation, subcloning techniques were developed. Clones showing evidence of heterogeneity for mouse chromosome two, as determined by variation in intensity of PCR intensity for markers on mouse

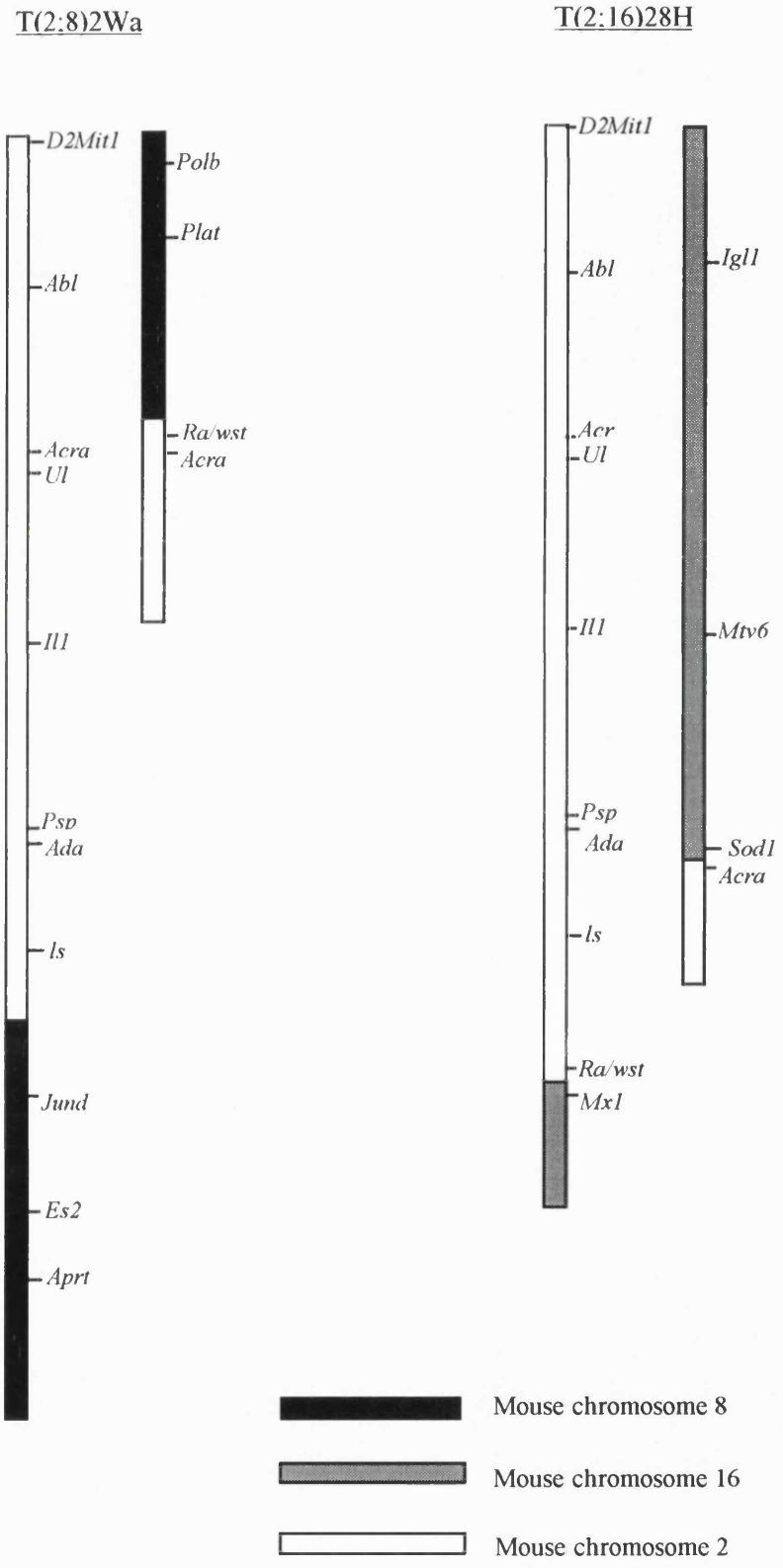


Figure 16:- To illustrate the appearance of T(2;8)2Wa and T(2;16)28H translocation chromosomes - genes close to the breakpoints are allocated preliminary map positions only

(Data for this figure were adapted from Mammalian Genome vol.4 , 1993)

chromosome two, were analyzed firstly for the presence of intact translocation products, and then subcloned by serial dilutions until an average of one cell per well was present on a microtitre plate. Individual colonies were grown up, and characterized for the presence of mouse chromosome two, by species specific PCR. As an alternative method, forced segregation was carried out by the use of an antibody specific to a cell surface antigen coded for by a gene on the proximal portion of mouse chromosome two. Conditions were derived for complement mediated cell killing of clones expressing this antigen.

1.9.5) The alignment of the genetic and physical maps of mouse chromosome two

Somatic cell hybrids carrying reciprocal translocations are of use in the alignment of genetic and physical maps - the genetic map is constructed by linkage analysis, whereas the physical map may be constructed from cytogenetic, somatic cell hybrid, or pulsed field gel electrophoresis or YAC contiging analysis. The two maps share the gene order, but the distance between genes can be distorted by the effects of recombination hotspots, transmission distortion effects, or by sex differences in the map distances. The allocation of markers to one side or other of a cytologically defined translocation breakpoint allows alignment of the genetic and physical maps. In order to align the genetic and physical maps of mouse chromosome two, the T(2;8)2Wa, T(2;16)28H, T(2;4)13H and T(2;4)1Sn translocation breakpoints were studied. Hybrid lines have already been set up which segregate T(2;4)1Sn and T(2;4)13H (A. J. Pilz, personal communication). The hybrid lines produced from the fusions described above were to be used to separate physically the two halves of mouse chromosome two, as defined by the T(2;8)2Wa and T(2;16)28H breakpoints. Initial studies were carried out to determine whether the translocation products were intact in the hybrid cell lines which were already available, T(2;4)1Sn and T(2;4)13H, by PCR analysis with markers which map to approximately evenly spaced regions of mouse chromosome two. Primer pairs were designed for genes which appear to map closely to one or other of the translocation breakpoints.

These studies are intended both to produce a panel of new molecular markers, which will help to increase the density of sequence tagged sites (STSs) for mouse chromosome two - this will facilitate the cloning of any gene of interest on mouse chromosome two, and to identify flanking markers for some developmental mutants on mouse chromosome two. New markers identified which appear to be linked to any of these genes will be mapped on specialized backcrosses, which segregate the mutant. Two backcrosses, produced in collaboration with the MRC Radiobiology unit, are currently available which segregate ragged

and wasted (*Mus musculus* X *Mus spretus*) and a third is currently being set up in Edinburgh, which segregates ulnaless (*Mus musculus* X *Mus musculus castaneus*). The construction of the hybrid panels will be of use in defining the genetic interval in which genes such as ragged (*Ra*) and wasted (*wst*) lie. The panels will allow the generation of new markers which are linked to these genes in a region specific manner, as well as providing a generalized mapping resource. Additionally, markers which are linked to these mutants may be produced in a region specific manner, by the use of IRS-PCR or bubble IRS-PCR on any segregated clones produced. In addition, these hybrid lines will provide a resource for the study of the imprinting region on distal mouse chromosome two. The alignment of the genetic and physical maps is of interest, in that it will allow a more accurate estimate of the physical distance which is indicated by the genetic distance between markers which closely flank developmental mutants. Taken as a whole, these studies are designed to provide a starting point for the positional cloning of ragged (*Ra*), wasted (*wst*), lethal spotting (*ls*) and ulnaless (*Ul*), as well as helping to increase the general marker density over the whole of mouse chromosome two.

Chapter 2

Materials and Methods

2.1) Suppliers and reagents

2.1.1) Suppliers

2.1.1a) Cell culture reagents

10X Eagle's MEM and 200mM glutamine were purchased from ICN Biomedicals Inc., U.S.A. Non-essential amino acids were purchased from Gibco BRL Life Technologies, Paisley, U.K. All antibiotics (penicillin, streptomycin and amphotericin B) were purchased from Sigma cell culture, St. Louis, U.S.A. Fetal calf serum was purchased from Advanced Protein Products Ltd., Brierley Hill, U.K.

All cell culture flasks, 24 well microtitre plates and 96 well microtitre plates were purchased from Gibco BRL Life Technologies (Nunc), Paisley, U.K. Cryotubes were purchased from Gibco BRL Life Technologies (Nunc), Paisley, U.K. Sterile universal tubes were purchased from Bibby Sterilin Ltd., Stone, U.K.

Rat anti-mouse Cd44 monoclonal antibodies (0.5mg/ml) were purchased from AMS Biotechnology (Europe) Ltd. Lyophilised guinea pig complement was purchased from ICN Biomedicals Inc., U.S.A., and was supplied in 1g aliquots, together with 1ml dilutant.

2.1.1b) PCR reagents

10X PCR buffer and *Taq* polymerase were obtained either from Advanced Biotechnologies Ltd., U.K., or from Promega, Madison, U.S.A. Oligonucleotides (including IRs-PCR primers) purchased were obtained either from Pharmacia Biotech, U.S.A. or from Oswel Oligonucleotides, Edinburgh, U.K. Ultrapure deoxynucleotides were obtained from Pharmacia Biotech, U.S.A. Additional oligonucleotides were obtained from Dr. Gabrielle Fisher at the HGMP Resource Center, Dr. C Abbott, and from Research Genetics Ltd.

2.1.1c) Gel reagents

Ultrapure agarose and low melting point agarose were obtained from Sigma Chemical Company, St. Louis, U.S.A. NuSieve 3:1 agarose was obtained from FMC Bioproducts, U.K.

2.1.1d) Restriction enzymes

All restriction enzymes were obtained from Gibco BRL Life Technologies, Paisley, U.K.

2.1.1e) Chemicals

All chemicals were obtained from BDH Analar Laboratory Supplies, Poole, U.K., unless otherwise stated.

2.1.1f) Radiochemicals

All radiochemicals (α^{32}P-CTP and γ^{32}P-ATP, α^{35}S-dATP) were obtained from Amersham International PLC, U.K.

2.1.1g) Radiolabelling and detection reagents

Random priming labelling kits ('Prime-it' version 2.0) were obtained from Stratagene, La Jolla, U.S.A. Polynucleotide kinase was obtained from Gibco BRL Life Technologies Ltd., Paisley, U.K. Sephadex G50 and G25 beads for probe purification were obtained from Sephadex Ltd., U.K. Microsatellite probes were obtained from British Biotechnologies, U.K., and from Oswel Oligonucleotides, Edinburgh, U.K. Hyperfilm double sided autoradiography film was obtained from Amersham Life Sciences, U.K.

2.1.1h) Membranes

Hybond N+ Hybridization filters were purchased from Amersham International PLC, U.K. Genescreen *Plus* hybridization filters were purchased from NEN research Products, Boston, U.S.A.

21.1I) Cloning Reagents

XL1-Blue F' competent cells and pBluescript KS+ phagemids were obtained from Stratagene Ltd, La Jolla, U.S.A. T4 DNA ligase and 5X buffer were obtained from Gibco BRL Life Technologies, Paisley, U.K. The Bubble PCR vectorette top and bottom strands (see table 6) were the kind gift of Dr. Cathy Abbott, as was the bubble vectorette PCR primer (see table 6).

2.1.1j) Sequencing reagents

Sequenase version 2.0 sequencing kits were purchased from United States Biochemicals, U.S.A. Sequagel 6.0 acrylamide mixture was obtained from National Diagnostics Ltd., U.K.

2.1.1k) Backcross DNAs

Panel 1 (BSB) and panel 2 (BSS) Jackson backcross DNAs (0.5μg/μl) were obtained from Dr. Lucy Rowe at the Jackson Laboratory, Bar Harbor, Maine, U.S.A. Ulnaless, Ragged and wasted backcross DNAs were the kind gift of Dr. Cathy Abbott.

2.1.2) Reagents

2.1.2a) Cell culture media

1X MEM - 1X Eagle's MEM, 1X Non-essential amino acids, 1mM sodium pyruvate, 10mM glutamine, 100u penicillin, 100μg streptomycin, 5ml HEPES buffer (1M HEPES, 0.25N NaOH per liter, The pH was adjusted to 8.0 with 10N NaOH), 1 to 4ml 5.6% $NaHCO_3$ pH 7.0, 395ml deionised distilled water (ddH_2O).

1X MEM-FCS - 1X MEM as above plus 10% (v:v) fetal calf serum.

1X MEM-FCS + HMT - 1X MEM-FCS as above, 1 X 10^{-4}M hypoxanthine, 1 X 10^{-5}M methotrexate and 1.6 X 10^{-5}M thymidine.

HANKS-B solution - stock : 1.1M NaCl, 0.04M KCl, 0.002M $Na_2HPO_4.12H_2O$, 0.04M glucose, 0.16g/L phenol red. Working stock : 90ml HANKS-B stock, 1080ml ddH_2O pH to 7.0 - 7.2 with 5.6% $NaHCO_3$.

Versene / trypsin solution - 100ml versene (0.005M EDTA in 1L HANKS-B working solution), 2.5ml 1% trypsin solution (0.02% EDTA, 2.5% trypsin) pH to 7.0 to 7.2 with 5.6% $NaHCO_3$.

Glycerol freezing medium - 72ml ddH_2O, 10ml glycerol, 8ml 10X Eagle's MEM, 1ml HEPES solution, 20ml fetal calf serum.

All solutions were sterilised by autoclaving at 121°C for 30 minutes. Heat labile reagents were sterilised by filtration through 0.2μm filters.

2.1.2b) DNA extraction reagents

TNES buffer - 50mM Tris-HCl, 100mM EDTA, 0.5% SDS, 400mM NaCl

2.6M NaCl

20mg/ml proteinase K solution.

2.1.2c) PCR reagents

Promega 10X PCR buffer - 500mM KCl, 100mM Tris-HCl pH 9.0 (at 25°C)

Advanced Biotechnologies 10X PCR buffer - 200mM $(NH_4)SO_4$, 200mM Tris-HCl pH 9.0 at 25°C, 0.1% (w:v) Tween.

Deoxynucleotide working stock - 2.0mM (composed of 2.0mM each of dATP, dGTP, dCTP and dTTP).

25mM $MgCl_2$ solution.

50% glycerol.

2.1.2d) Gel electrophoresis buffers

10X TAE buffer - 0.4M Tris-HCl pH 7.6, 0.05M sodium acetate, 0.01M EDTA.

5X TBE buffer - 0.45M Tris-HCl, 0.45M Orthoboric acid, 0.01M EDTA.

2.1.2e) Blotting and hybridization solutions

Church buffer - 7% SDS, 1mM EDTA, 0.5M $NaHPO_4$.

20X SSC - 3M NaCl, 1.3M Sodium citrate dihydrate.

TE buffer - 10mM Tris-HCl pH 7.5, 1mM EDTA.

20X SSPE - 3M NaCl, 0.2M Sodium phosphate, 0.02M EDTA.

Denaturing solution - 1.5M NaCl, 0.5M NaOH.

Neutralising solution - 1.5M NaCl, 0.5M Tris-HCl, pH 7.2, 0.001M EDTA.

Kinase buffer - 0.7M Tris-HCl pH 8.0, 0.1M $MgCl_2$, 50mM DTT.

2.1.2f) Cloning reagents

5 X T4 DNA ligase buffer - 250mM Tris-HCl pH 7.6 at 25°C, 50mM $MgCl_2$, 5mM ATP, 5mM DTT, 25% (w:v) polyethylene glycol 8000.

Ampicillin stock solution - 50mg/ml in distilled water. Filtered before use.

2.1.2g) SSCV analysis reagents

8% non-denaturing polyacrylamide gel : 8% acrylamide, 10% glycerol, 0.5X TBE (polymerise with 1:1 25% ammonium persulphate and TEMED).

loading buffer - 980µl deionised formamide, 20µl 0.5M EDTA, bromophenol blue.

Silver staining solution 1 - 10% ethanol, 0.5% acetic acid.

Silver staining solution 2 - 0.014M $AgNO_3$.

Silver staining solution 3 - 0.26M NaOH, 0.003M $NABH_4$, 4ml/L 37% formaldehyde.

2.1.2h) Preparation of backcross DNAs

The Jackson backcross DNAs were supplied at a concentration of 0.5µg/µl. The DNA was further diluted to 50ng/µl in TE buffer.

2.2) The construction and characterization of two new panels of somatic cell hybrids

Two new panels of mouse : Chinese hamster somatic cell hybrids were constructed and characterized. The hamster parent in each case was the cell line V79TOR2, which is MMPR resistant, and HAT sensitive (Thacker, 1980). This line was chosen for the hamster parent as it had conferred stability in some mouse : hamster hybrid panels (Y. Boyd, unpublished results). The mouse component derived from primary cells taken from mice carrying the reciprocal translocations T(2:8)2Wa (Searle, 1989) in one panel, and T(2:16)28H (Searle, 1989) in the other panel. The breakpoints on mouse chromosome two for these translocations are close together, flanking a region of chromosome which possibly contains the interesting developmental mutants ragged (*Ra*) and wasted (*wst*) (Siracusa and Abbott, 1993). The distal imprinting region on mouse chromosome two is also defined by the T2Wa and T28H translocation breakpoints (Cattanach, 1986; Cattanach *et al.*, 1992). The use of mice carrying these translocations in the hybrid panels provides a means of isolating the region of chromosome flanked by the breakpoints. The use of a primary cell line from these mice removed the need for an alternative selection system to prevent the survival of the murine homokaryons (Nabholz *et al.*, 1969; Minna *et al.*, 1975). The fusions were mediated by polyethylene glycol (PEG), and hybrid cells were selected for by the use of HMT medium. Individual clones were isolated, and characterized by polymerase chain reaction (Abbott and Povey, 1991; Abbott, 1992). Alternative systems for the production of segregation were examined : the use of subcloning techniques which exploit the inherent heterogeneity within the hybrid population, and the use of mouse specific antibodies as agents producing negative selection for particular chromosomes or chromosome regions.

2.2.1) The construction of the hybrid panels

2.2.1a) The preparation of the parental cell lines

The mouse parental cell lines for the somatic cell hybrid panels derived from mice carrying the T(2:8)2Wa and T(2:16)28H reciprocal translocations (see table 2). These mice were obtained from Mr. C. Beechey, at the MRC Radiobiology unit in Harwell. The spleens were dissected from freshly killed mice, and teased apart under 1X MEM to liberate the cells.

translocations	sex	age	comments
homozygous T(2;8)2Wa	male	32 days	also *c/c*
homozygous T(2;16)28H	female	53 days	-

Table 2 : Details on translocation carrier mice

The Chinese hamster parental cell line V79TOR2 (the kind gift of Dr. J. Thacker) was grown to a confluent monolayer in 1X MEM supplemented with 10% fetal calf serum (MEM-FCS). Just prior to fusion, the cells were washed with HANKS-B solution, and removed from the culture flasks with Versene / trypsin solution. This reaction was terminated with 1X MEM-FCS. The cells were counted using a haemocytometer. The volume of cell suspension yielding approximately one million cells was calculated using the formula :-

Number of cells per ml solution =

$$\frac{\text{Average number of cells in 4 counting chambers of haemocytometer}}{4} \times 10^5$$

The remainder of the cells were replaced into fresh 1X MEM-FCS medium, and cultured for storage.

2.2.1b) The fusion procedure

Approximately 1 X 10^6 Chinese hamster cells were mixed with cells deriving from one mouse spleen in a sterile universal bottle. The volume was adjusted to 20ml, with 1X MEM. The mixed cell suspension was pelleted by centrifugation at 1.5K for 5 minutes (all subsequent spins were under these conditions), and washed once with 10ml 1X MEM. The cell pellet was resuspended in 1ml of 50% polyethylene glycol (PEG) 6000, which had been previously heated at 100°C and cooled to 37°C. The pH of the PEG solution was adjusted to 7.0 - 7.2 with 5.6% $NaHCO_3$. The PEG-cell suspension was mixed well for exactly one minute, and the volume was then made up to 10ml with 1X MEM. The fused cells were then washed once more with 1X MEM, and resuspended in 20ml 1X MEM-FCS. This cell suspension was then transferred in either 1ml or 0.5ml aliquots (to ensure optimal spacing of individual colonies) into sterile 25cm^3 tissue culture flasks, and incubated at 37°C for 24 hours. After this time, the medium

was replaced with 1X MEM-FCS supplemented with HMT, and the cells were incubated at 37°C.

2.2.1c) The isolation and culture of individual clones

The cell lines were grown in 1 X MEM-FCS + HMT (occasionally supplemented with 2.5μg/ml amphotericin B to combat a latent fungal infection), at 37°C, until they reached a suitable size (colonies of approximately 2mm in diameter) for the isolation of the clonal cell lines. This was accomplished by the physical removal of each colony using a plugged Pasteur pipette, which had been heat treated to bend the tip, and seal the end. Each cell line was transferred to 25cm^3 tissue culture flasks containing 5ml 1X MEM-FCS + HMT. The cells were grown to confluent monolayers, at which point they were released from the base of the culture flask by Versene/trypsin solution (preceded by washing the cells in HANKS-B solution, as fetal calf serum inhibits the action of trypsin) and gentle agitation. The cell suspension was transferred to an 80cm^3 tissue culture flask. Growth rate, as measured by the amount of time taken to reach confluence from inoculation of an 80cm^3 tissue culture flask :-

very fast = 1-2 days to confluence.
fast = 2 - 4 days to confluence.
moderately fast - 4 - 6 days to confluence.
moderately slow = 7 - 9 days to confluence.
slow = 9 - 14 days to confluence.
very slow = more than 14 days to confluence.

Colony morphology and cell morphology were noted. Cells from at least two confluent 80cm^3 flasks were frozen in 1ml glycerol freezing medium at -175°C (in liquid nitrogen tanks) as a permanent resource. Individual hybrid cell lines were grown up to bulk culture (to fill an average of six 180cm^3 tissue culture flasks), and harvested by centrifugation, after three successive washes in 0.9% saline solution. Dry cell pellets were stored at -70°C until DNA preparation.

2.2.2) The characterization of the hybrid panels

DNA preparation was carried out according to the high salt extraction method outlined below. Characterization of the hybrid lines was achieved by polymerase chain reaction (PCR) using previously existing primer panels specific to mouse, (Love *et al.*, 1990; Abbott, 1992; Dietrich *et al.*, 1992) according to the conditions laid out in the publications, and also some newly designed primer pairs (designed for the alignment of the genetic and physical map). At least two markers per chromosome were used, to provide a rough estimate of the extent of chromosomal fragmentation. PCR products were analyzed on 1.2% agarose gels, stained with ethidium bromide (15μg per 30ml gel).

2.2.2a) DNA extraction procedure

The cell pellet was resuspended in 500μl TNES buffer, together with 10μl proteinase K (20mg/ml), and was incubated at 55°C overnight, or until all proteins had digested, as determined by the clarity of the solution. The NaCl concentration was adjusted to 1.5M and the solution was shaken, to precipitate all digested proteins. The DNA in solution was separated from the proteins by centrifugation at 13K for 10 minutes. DNA was precipitated from the supernatant by the addition of an equal volume of 100% ethanol. The DNA was removed by spooling onto a glass rod, or by centrifugation at 13k for 10 minutes. If the DNA to be used for amplification, it was immediately removed from the ethanol solution by spooling, as there is some evidence that prolonged exposure to ethanol reduces the ability of the DNA to undergo PCR amplification (Dr. C Abbott, personal communication). The DNA was redissolved in 100 to 500μl ddH_2O, to an approximate concentration of 1μg/μl.

2.2.2b) PCR characterization procedure

The amplification reactions were carried out using mouse specific PCR primers, designed from unconserved regions of genes, such as 3' and 5' untranslated, and intron sequences. This ensured that under stringent conditions, no signal originating from the hamster parent was produced. Any positive result was therefore due to the presence of the mouse gene in the hybrid cells. Mouse (DNA derived from T2Wa or T28H parent) and hamster (DNA derived from V79TOR2 cells) were co-amplified with the hybrids under test, as positive and

negative controls respectively. Primers for specific genes were obtained from Dr. C Abbott (Abbott, 1992), Dr. Gabrielle Fisher at the HGMP resource center (Love *et al*, 1990; Hearne *et al*, 1991) and from Research Genetics Ltd. (*D2Mit* primers; Dietrich *et al*, 1992). Additional primers (*Plf*, *Dfhr*, *Polb* s primer and *Plat* primers) were designed from sequences given in the literature (see table 3). At least two markers were tested for each chromosome, to allow a crude measure of the extent of chromosomal fragmentation within the clones. PCR reactions were carried out in a total volume of 100µl, unless otherwise stated. Reaction mixes contained 1X PCR buffer (Promega), 200µM dNTPs, $MgCl_2$ (the concentration of this was dependent on the requirements of the particular PCR primers), 200ng of mouse chromosome specific primers and 1µg hybrid DNA. See table 4 for additional primer details and conditions. The cycling conditions for the other primer pairs are as given in the literature. The amplification was mediated by 2u *Taq* polymerase. In each case, mouse and hamster positive and negative controls were included in the reaction. All PCR amplifications were carried out in Hybaid TR2 machines.

PCR products were visualised on 1.2% TAE agarose gels, stained with ethidium bromide (15µg per 30ml gel), and viewed under ultra-violet light.

2.2.3) The subcloning of the hybrid panels

As an alternative strategy to isolate segregate hybrid clones, subcloning techniques were devised. These techniques exploit the naturally occurring heterogeneity found in hybrid populations. Viewed under high magnification, chromosome preparations made from hybrid clones show high frequency of heterogeneity, in that not all cells in the clone contain all the chromosomes associated with that clone. The extent of heterogeneity varies between clones. This is to some extent a feature of the sensitivity of the characterization technique, as less sensitive techniques may not reveal the presence of chromosomes occurring at low levels in the hybrid clones. PCR characterization is sensitive enough to pick up a signal from picogram amounts of DNA, and therefore will identify chromosomes present at only low numbers within the clone. Heterogeneity of the hybrid clones is revealed as differences in the signal intensity after PCR amplification, provided that the reaction is in the linear phase. The clones produced in these experiments might be expected to show heterogeneity for some chromosomes in the population, as shown by variations in the intensity of PCR products for certain chromosomes compared with markers from other chromosomes in these clones, and also with the intensity of

locus name	chromosomal location	sequence	[$MgCl_2$]	conditions	product size
Acra4 (Designed by Dr. C Abbott).	2	5 ATG TCA GGC GTC CAG ATG AG 3' 5' GAC GAG AAG AAC CAG ATG AT 3'	1.5mM	a	220 bp
c-src (Sequence from Black, 1991).	2	5' AGA GGG GGA TGC TTC GCT GA 3 5' GAC CTG GTC TGG TGT AAC CC 3''	2.0mM	c	281 bp
Psp (Sequence from Shaw and Scheiber, 1986).	2	5' CAG TAC CCC ACC CTA GAT GG 3' 5' CTG ACT GTG CTG AGG AGG GT 3'	1.5mM	b	274 bp
Polb (Sequence from Yamaguchi *et al.*, 1987; Hearne *et al.*, 1991).	8	5' TTC CCA CAG TCA GTC ACT TA 3' 5' GGG AAA TGA TGT AAT TAT AAT CC 3' a/s oligo as designed by Hearne *et al.*, (1991)	3.0mM	a	225 bp
Plat (sequence from Rickles *et al.*, 1988)	8	5' GAA GAT GCG CCT AAA AGG CC 3' 5' GGC AGA CTT TGG TCT ACT CC 3'	3.0Mm	a	243 bp
Plf (Sequence from Linzer and Mordacq, 1987).	13	5' CTC ACA ATG TTC CTT GGG TG 3' 5' GCC ACA ACT ATG TCT TCC TC 3'	1.5mM	a	240 bp
Dfhr (Sequence from Chen *et al.*, 1984).	13	5' TAG CGT GAA GGC TGG TAG GA 3' 5' TAG CTT GGT CTC TGG CTG AG 3'	2.0mM	a	248 bp
Mx1 (Sequence from Hug *et al.*, 1988).	16	5' TCC TCC TGG AGA GAA GGG CA 3' 5' TGT AGG CTC AGC CCA CCA AG 3'	1.5mM	b	339 bp

PCR conditions :- **a**= 94°C - 15s, 55°C - 30s, 72°C - 30s, 30 cycles, **b** = 94°C - 30s, 55°C - 30s, 72°C - 30s, 30 cycles, **c** = 94°C - 30s, 57°C - 30s, 72°C - 30s, 30 cycles

Table 3:- PCR characterization primer sequences and conditions

Locus	Chromosomal location	Sequence	[$MgCl_2$]	Conditions	Product size
Jun (The kind gift of Dr. J. Friedman)	4	5' AGC AAC TTT CCT GAC CCA GAG G 3' 5' TTA AGA CTC CGC TAG CAC TCA CG 3'	1.0Mm	94° - 120s, 60° - 60s, 72° - 180s, 30 cycles.	420 bp.
Mos (The kind gift of Dr. J. Friedman)	4	5' TAC CAG GGT GTA AAC CGT CTG C 3' 5' TTC ATG TGA CAG CTG TGT CCC 3'	1.0Mm	94° - 120s, 60° - 60s, 72° - 180s, 30 cycles.	741 bp

Table 4:- Additional PCR primers for hybrid characterization

positive controls done in the same set of reactions. The clones chosen for subcloning experiments in these experiments were selected on the basis of faint positive PCR results for markers on mouse chromosome two. No reduction in signal was noted for other chromosomes. The integrity of the translocation products in these clones was first assessed by PCR characterization of the clones for larger numbers of markers mapping to mouse chromosome two. This mechanism cannot identify events of fragmentation in which all markers are retained, or rearangements which occur between markers.

IRS-PCR (using the L1 primer of Irving and Brown, 1991, see table 5) was carried out on one panel, to discover if any changes in chromosome content between the clones was reflected in the band patterns produced. This panel was also partially characterized for markers on chromosomes 1, 4, 6, 7 and 8, to provide evidence of the efficiency of the subcloning techniues for the production of segregation in somatic cell hybrids.

2.2.3a) Examination of integrity of translocation products in the hybrid cells

The cell lines T2Wa 1 and T28H 3b were selected for subcloning experiments. PCR reactions were carried out in 100µl reactions, as in the original characterization (see section 2.2.3b). PCR primer choice was arbitrary, but some attention was paid to the selection of primer pairs which gave strong, reproducible results. The products were visualised on 1.2% agarose gels, stained with ethidium bromide, and viewed under ultraviolet light.

2.2.3b) Subcloning techniques

Method 1:- Serial dilution. Cells were grown to confluence in $80cm^3$ tissue culture flasks. Cells were removed from the flasks using versene / trypsin solution after washing the cells with HANKS-B solution. The volume of the cell suspension was adjusted to 10ml wih 1X MEM-FCS + HMT. A 96 well microtitre plate was prepared, with 250µl 1 X MEM-FCS +HMT in each well. The first well on the top of the first row contained 250µl of the original cell suspension. 10µl serial dilutions were carried out from this well into the seven wells vertically below it. Eleven 10µl serial dilutions were then carried out from these eight wells horizontally across the plate (see figure 17). The inoculated plates were incubated at 37°C, with medium changes every few days. Any wells containing more than one colony were discarded. Individual colonies were removed from the microtitre plates when they reached an appropriate size, and

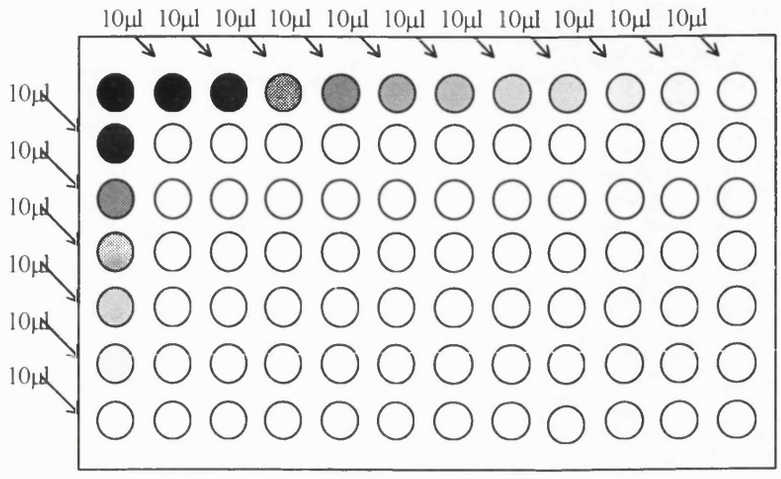


Figure 17: - Serial dilution method of subcloning

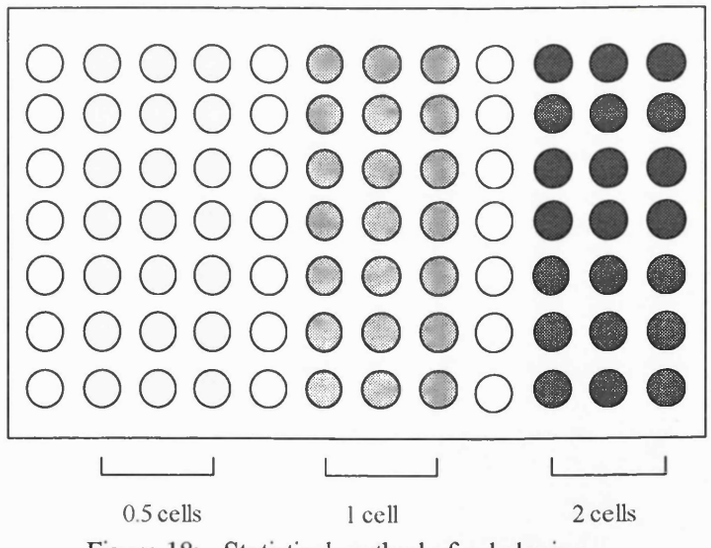


Figure 18: - Statistical method of subcloning

were transferred to $25cm^3$ tissue culture flasks, and grown up as in section 2.2.2c. The individual clones were then characterized for markers on mouse chromosome two, as in section 2.2.3b.

Method 2: - Statistical calculation of cell concentration. Cells were grown to confluence in $80cm^3$ tissue culture flasks, and released from the growth surface as above. The volume of the cell suspension was adjusted to 10ml. A 96 well microtitre plate was prepared, with 250µl 1X MEM-FCS +HMT in each well, as above. The number of cells per ml was calculated using a haemocytometer, and the volumes required to yield 0.5, 1 and 2 cells were calculated. These volumes of cell suspension were added to 30 wells each (see figure 18). The plates were incubated at 37°C. Any wells containing more than one colony were discarded. When the individual clones reached an appropriate size, they were transferred to $25cm^3$ tissue culture flasks, and cultured and characterized for markers on mouse chromosome two, as above.

2.2.3c) Evidence for segregation of other chromosomes

IRS-PCR was carried out with the mouse specific L1 primer of Irving and Brown (1991). Reaction mixes contained 1X PCR buffer (Promega), 200µM dNTPs, 1.5mM $MgCl_2$, 400ng of L1 primer and 0.5u *Taq* polymerase (Advanced Biotechnologies Ltd.) in a total volume of 100µl, under the conditions given in table 5. PCR products were visualised on 1.2% agarose gels, stained with ethidium bromide (35µg per 70ml gel), and viewed under ultraviolet light. The subclones were also partially characterized for one marker for chromosomes 1, 4, 6, 7 and 8, under the conditions stated in section 2.2.3b.

2.2.4) Forced segregation of mouse chromosomes by antibody selection

2.2.4a) Preparation of cell cultures, media and antisera

Prior to the experiment, mouse fibroblast 1R cells (the kind gift of Dr. M. Fox) and V79TOR2 cells were grown to confluence in 1X MEM-FCS, in $80cm^3$ tissue culture flasks. All media to be used in the experiments were heat inactivated at 65°C for 15 minutes. This serves to inactivate any endogenous complement and antisera in the fetal calf serum, which would otherwise interfere with the experiment. Due to the presence of sodium azide (which is highly cytotoxic) as a preservative in the antiserum, dialysis was carried out against 0.9%

saline at 4°C, for a total of 20 hours, with three succesive buffer changes. All manipulations of antisera were caried out under aseptic conditions. Toxicity tests were carried out with 5% antibody alone, and 25% complement alone, to ensure that any cell killing seen was the result of specific interactions, and not due to non-specific toxicity.

Mitotic cells were removed from the culture flasks, by gentle agitiation with 5ml HANKS-B solution alone. The use of mitotic cells ensures that all cells exposed to the antiserum have the ability to grow and divide. In order to produce a sufficiently low concentration of cells, the volume of the cell suspension was made up to 10ml with 1X MEM-FCS. One drop of this suspension (approximately 250μl) from a Pasteur pipette was found to give appropriate numbers of cells.

2.2.4b) Determination of optimal conditions for cell killing

Experiments were carried out in 24 well microtitre plates. 1ml heat treated 1X MEM-FCS was placed in each well. Two experiments were set up with mouse fibroblast cells. The first experiment tested the effect of the antibody concentration. Five wells were inoculated with 1R cells alone, as controls. Seven wells were each treated with 10% complement, and 1%, 2%, 3%, 5%, 10%, 15% and 20% antibody solution respectively, together with 1R cells. Another experiment was set up to examine the effect of the complement concentration. Five wells were again treated with 1R cells alone. Seven wells were treated with 5% antibody, and 1%, 2%, 3%, 10%, 15%, 20% and 25% complement respectively. The plates were incubated at 37°C, and examined after 48 hours. The number of cells in three representitive colonies in each well was averaged under a light microscope.

2.2.4c) Cross-reactivity tests

In order for the antibody selection system to be successful, the antibody must exibit species specificity. In order to determine the extent of species specificity for the particular antibody used in this experiment, complement titrations were set up. Experiments were carried out in 24 well microtitre plates. 1ml heat treated 1X MEM-FCS was placed in each well. Five wells were inoculated with V79TOR2 cells in the absence of antibody or complement, as growth controls. Five wells were treated with 0.001%, 0.01%, 10%, 20% and 50% complement respectiviely, in the presence of 5% antibody. Each treated well was inoculated

with 1 drop of 10ml V79TOR2 cell suspension (prepared in the same way as the 1R solution). Treated plates were incubated at 37°C, and examined after 24 and 48 hours. The number of cells in three representative colonies were counted under a light microscope, and an average taken.

2.2.5) The use of somatic cell hybrids carrying reciprocal translocations for the alignment of the genetic and physical maps of mouse chromosome two

For the purpose of the alignment of genetic and physical maps of mouse chromosome two, new PCR primers were designed from published sequence of genes located near to the T(2:4)1Sn, T(2:4)13H and T(2:16)28H translocation breakpoints, details of which are reviewed in Searle, 1989. The experiment was carried out using the T1Sn and T13H hybrid panels previously constructed by A. J. Pilz. The T28H hybrid panels were constructed for the purposes of this experiment. The PCR primers designed are intended for use on any future hybrid panels which segregate the T28H translocation products.

2.2.5a) Cell culture and DNA extraction

Prior to the experiment, T1Sn and T13H hybrids showing evidence of segregation (T1Sn 3-1, T13H 3-1 and T13H 10-1) were grown to bulk culture (an average of eight 180cm^3 tissue culture flasks), in 1X MEM-FCS + HMT. The cells were harvested by centrifugation, following sucessive washes in 0.9% saline solution. DNA was extracted as described in section 2.2.3a. The hybrids were recharacterized for markers on mouse chromosome two, to ensure that they had retained the chromosome two translocation products, as described in section 2.2.3b.

2.2.5b) Primer design

Sequence was obtained from publications for genes apparently mapping in the vicinity of the T1Sn and T28H translocation breakpoints. PCR primer pairs were designed from unconserved regions (3' or 5' untranslated sequence, intron sequence) of genes, to allow species specificity. PCR primers were synthesized for certain genes, and PCR conditions were established. See table 3 for primer sequences and details.

2.2.5c) Assessment of integrity of translocation chromosomes in T13H 3-1, T13H 10-1, T1Sn 3-1 and T1Sn 9-1

In order to test whether the translocation products were intact in the hybrid cells, as determined by PCR, characterization was carried out with markers mapping to mouse chromosome two (*D2Mit* primer pairs are described in Dietrich *et al.*, 1992, *D2Ucl* primer pairs were generated during the course of these studies, and are described in table 27. *Ada*, *Abl* and *Il1b* primers are described in Abbott, 1992, and *Acra4* primers are detailed in table 3) and chromosome four (*Mos* and *Jun* primers are described in table 4, details of *Mup*, *Lck* and *Orm* primers can be found in Love *et al.*, 1990). PCR was carried out in 100µl reaction volumes, according to the conditions given for each primer pair in the literature. Reaction mixes contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 200µM dNTPs, $MgCl_2$ (the concentration of this was dependent on the requirements of the particular PCR primers), 200ng each of mouse chromosome specific primers and 1µg hybrid DNA. The amplification was mediated by 2u *Taq* polymerase. In each case, mouse and hamster positive and negative controls were included in the reaction. All amplifications were carried out using Hybaid TR2 thermal cycling machines. PCR products were visualized on 1.2% agarose gels, stained with ethidium bromide (35µg per 70ml gel), and viewed under ultraviolet light.

2.3) The generation of new DNA markers for mouse chromosome two by IRS-PCR and IRS-bubble PCR

Interspersed repetitive sequence (IRS) PCR was carried out on a panel of microcell and monochromosomal somatic cell hybrids, containing various portions of mouse chromosome two. The fragments produced by IRS-PCR were examined for their suitability for use as hybridization probes on Southern blots, and examined for the presence of putative CpG islands. The fragment population was also screened for the presence of microsatellite repeats. The effectiveness of cloning directly from the products of the initial amplification was compared with that obtained by the cloning of individual fragments which had been isolated from low melting point agarose gels, and re-amplified before the cloning process. Hybridization techniques were employed to distinguish mouse clones from the background of hamster clones prior to sequencing. In addition, IRS-bubble PCR was carried out on a somatic cell hybrid which was essentially monochromosomal for mouse chromosome two. IRS-bubble PCR products were cloned in the same manner as IRS-PCR products, and sequenced. PCR primers were designed from the sequence obtained from IRS-PCR and IRS-bubble PCR products, for use as STS probes for interspecific backcross analysis.

The hybrid lines used for IRS-PCR were:-

a) the EBS18Az cell line (the kind gift of Dr. P. Lalley), which is essentially monochromosomal for the whole of mouse chromosome two, with the exception of a small fragment of proximal chromosome 15 (P. Lalley, personal communication) and a fragment of the Y chromosome which is undetectable by cytogenetic methods (A. J. Pilz, personal communication).

b) the microcell hybrid lines ABm 5, 7 and 11 (the kind gift of Dr. K. Fournier). These hybrids contain only mouse chromosome two. ABm 7 contains the proximal region of mouse chromosome two only, as determined by cytogenetic analysis (Fournier and Moran, 1980) and PCR analysis (N. Ray, personal communication). ABm 5 and 11 contain the majority of mouse chromosome two, although there is some fragmentation and rearrangement (Fournier and Moran, 1980; C. Abbott and A. Pilz, personal communication).

2.3.1) IRS-PCR amplification of inter-repeat sequences from mouse chromosome two.

IRS-PCR was carried out with L1, B2 or a combination of L1 and B2 primers, on DNA extracted from EBS18Az, ABm 5, ABm 7 and ABm 11. The reactions were carried out in 100µl volumes, in 1X PCR buffer, 1.5mM $MgCl_2$, 200µM dNTPs, 400ng primer and 1µg hybrid DNA. In the case of the L1/B2 amplifications, 200ng each primer were used. Mouse and hamster positive and negative controls were also amplified, as a means of assessing species specificity. The reactions were mediated by 2.5u *Taq* polymerase. The cycling conditions are given in table 5. The amplification products were visualized on 1.2% agarose gels, stained with ethidium bromide (35µg per 70ml gel), and viewed under ultraviolet light. Pools of fragments generated from IRS-PCR amplification were subjected to electrophoresis on 1.2% low melting point agarose midi gels (100ml capacity) for 18 hours at 15V/cm. Gels were run in 1X TAE buffer. Individual PCR products were excised from the gels under ultraviolet light, and redissolved in 500µl distilled water (ddH_2O). 5µl of this solution was then reamplified using the original primer, under the same conditions (given in table 5).

The L1/B2 and B2(*EcoR1*) reactions were repeated, in order for a comparison to be made between the efficiency of cloning directly from the initial PCR reactions, as compared to the efficiency of cloning from reamplified fragments, which had been isolated from low melting point agarose gels.

2.3.2) Use of IRS-PCR products directly as probes on Southern blots

2.3.2a) Southern blot procedure

DNAs from mouse, hamster, T1Sn 3-1 and T13H 3-1 were digested with *EcoR1* in 1X ReAct 3 buffer under the conditions supplied by the manufacturer. The digested DNA was electrophoresed for 18 hours on 1% agarose maxigels (capacity 300ml) at 15V/cm. Gels were run in 1X TAE buffer. The gels were stained with ethidium bromide (150µg per 300ml gel), and viewed under ultraviolet light. After electrophoresis, the gel was subjected to the following washes :- 10 minutes in 0.25N HCl, 30 minutes in denaturing solution and 45 minutes in neutralizing solution. The recipes for all solutions used in this experiment can be found in the appendix to this section. The gel was briefly washed with distilled water after the acid and alkali washes. The gel was blotted for 18 hours onto Genescreen Plus hybridization membrane,

Repeat	Sequence	Conditions
L1(*EcoR1*) (Irving and Brown, 1991, amended by the addition of *EcoR1* restriction site)	5' CCG AAT TCG GTA TGG GGG ACT TTT GGG AT 3' Designed from 3' end of L1 repeat, to ensure species specificity.	94° - 60s, 60° - 45s, 68° - 300s, 35 cycles.
B2(*Not1*) (Simmler *et al.*, 1991)	5' AAG TCG CGG CCG CTC TTC TGG AGT GTC TGA AGA 3' Designed from 3' end of B2 repeat, to ensure species specificity.	94° - 60s, 55° - 48s, 68° -300s, 40 cycles.
B2(*EcoR1*) (Simmler *et al.*, 1991)	5' CCG AAT TCT CTT CTG GAG TGT CTG AAG A 3' Designed from 3' end of B2 repeat, to ensure species specificity.	94° - 60s, 55° - 48s, 68° - 300s, 40 cycles.
L1(*EcoR1*)/B2(*EcoR1*) (Irving and Brown, 1991; Simmler *et al.*, 1991)	5' CCG AAT TCG GTA TGG GGG ACT TTT GGG AT 3' 5' CCG AAT TCT CTT TG GAG TGT CTG AAG A 3'	94° - 60s, 57° - 45s, 68° - 300s, 40 cycles.

Table 5:- Details of IRS-PCR primers and conditions

according to the manufacturers instructions. The transfer buffer was 10X SSC. After the transfer, the gel was washed in 0.4N NaOH for exactly one minute, and neutralized in 0.2M Tris-HCl, 1X SSC solution for 5 minutes. The DNA was fixed to the filter by baking in an 80°C oven for 2 hours.

2.3.2b) Oligolabelling procedure

25ng of IRS-PCR product was labelled with α^{32}P-dCTP by the random priming method. The 'Prime-it' version 2.0 oligolabelling kit was used, according to the manufacturers instructions. The probe was purified by centrifugation through Sephadex G50 beads at 3K for 5 minutes.

2.3.2c) Hybridization procedure

The filter was pre-hybridized at 65°C, for one hour, in 20ml hybridization solution (0.01M NaCl, 10% (w:v) dextran sulphate, 2% (w:v) SDS), in heat-sealed plastic bags. The probe was incubated with 4mg unlabelled sonicated mouse DNA for 4 hours prior to hybridization, as a means of competing out the repetitive sequences present in the probe. The boiled probe was added to the hybridization bag, along with 250µl sonicated salmon sperm DNA, to minimize background due to non-specific hybridization of the probe to the filter. The

filter was then incubated at 65°C for 18 hours. After the hybridization, the filter was washed twice in 2X SSC at room temperature, to remove the majority of the unbound probe, and once for 30 minutes in 2X SSC/ 1% SDS, at 65°C, to remove non-specifically bound probe. The filter was dried briefly, and exposed to Hyperfilm autoradiography film at -70°C for 24 hours, in the presence of an intensifying screen.

2.3.3) Screening for the presence of putative CpG islands

It has been noted that the dinucleotide CpG, although underrepresented in the genome as a whole, is often associated with the 5' ends of housekeeping and some tissue specific genes. Clusters of hypomethylated CpG dinucleotides, known as CpG islands, are useful markers for the presence of genes in cloned DNA (Bird, 1986; Bird, 1987). The presence of CpG islands is revealed by clustering in DNA of rare cutter restriction sites (with high G+C content in the recognition site), such as *Not1*and *EagI*. The restriction enzyme *HpaII* has a GC rich recognition site, and thus *HpaII* cuts frequently in CpG islands. In order to identify any putative CpG islands present in the fragments produced by IRS-PCR, each fragment was subjected to restriction digestion with *Not1* and *HpaII* under the manufacturer's instructions. The digested fragments were electrophoresed on 3% 3:1 NuSieve agarose gels, stained with ethidium bromide (15μg per 30ml gel), and viewed under U.V. light.

2.3.4) Screening for the presence of microsatellites

In order to screen the isolated IRS-PCR products for the presence of microsatellite repeats, the fragments were first affixed to hybridization membrane, as described in section 2.3.2a, with the exception that the fragments were electrophoresed on 1.2% agarose midigels (capacity 100ml) rather than 1.2% agarose maxigels (capacity 300ml) prior to the transfer process. These gels were run for 18 hours at 15V/cm. Oligonucleotide probes corresponding to a number of dinucleotide and trinucleotide repeat units were radioactively labelled, by the end labelling process.

2.3.4a) End-labelling procedure

Oligonucleotide probes were end-labelled with γ^{32}P-dATP, mediated by polynucleotide kinase. Labelling reactions contained 10 pmol oligonucleotide, 1X kinase buffer (see appendix

to section 2.2), 1µl γ^{32}P-dATP and 10u polynucleotide kinase in a total volume of 10µl. This mixture was incubated at 37°C for 1 hour. The resulting probe was purified by centrifugation through Sephadex G25 beads, with 3X SSC as an inert carrier solution.

2.3.4b) The hybridization procedure

The IRS-PCR fragments were screened with the following probes : -$(CA)_{15}$, $(GA)_5$, $(CAT)_5$, $(GTT)_5$, $(AGA)_5$, $(CCA)_5$, $(CGA)_5$, $(AAT)_5$, $(CCG)_5$ and $(TCC)_5$. Each dinucleotide repeat oligonucleotide is capable of identifying four repeats (e.g. the $(CA)_{15}$ probe identifies (CA)n, (AC)n, (TG)n and (GT)n). The trinucleotide repeat probes are each capable of identifying six repeat units (e.g. the $(CAG)_5$ probe identifies (CAG)n, (AGC)n, (GCA)n, (CTG)n, (TGC)n and (GCT)n repeats). Melting temperatures (Tm) were calculated for each oligonucleotide, using the formula given below :-

$$Tm = \frac{81.5 + 16.6 \log[Na^+] + 0.41 (\%C+G)}{L} - \frac{600}{L}$$

where Tm = melting temperature and L = the length of the oligonucleotide.

The Tm values for each oligonucleotide are given in table 6. The hybridization procedure (adapted from Cornall *et al.*, 1991) was carried out in Church buffer (see appendix to this section) for 3 hours, at 5°C below the calculated Tm value for each oligonucleotide (see table 6). After the hybridization, the filters were washed four times. The first three washes were for 15 minutes, in 6X SSC/ 0.1% SDS, at 5°C below Tm, The final wash was for 2 minutes exactly, in 6X SSC/ 0.1% SDS, at Tm. The filters were briefly air-dried, and exposed to Hyperfilm autoradiography film at -70°C for 24 to 72 hours, depending on the signal from the filters, as assessed by Geiger-Muller counter. The same set of filters were used for each hybridization, the blots being stripped by treatment of the filters with boiling 0.1% SDS, and monitored to ensure adequate removal of probe.

2.3.5) The generation of new DNA markers by IRS-bubble PCR

2.3.5a) The production of the EBS18Az-bubble vector constructs

IRS-bubble PCR was carried out only on EBS18Az DNA, as this hybrid was the only source of an intact chromosome two, with no apparent rearrangements. 10μg DNA from this hybrid was first digested with 30u *HaeIII*, in order to produce a range of fragments, some of which contained internal repeat sequences. These fragments were ligated to bubble vectorettes, annealed by boiling for 5 minutes, followed by slow cooling to 35°C (see table 7).

2.3.5b) The IRS-bubble PCR reaction

IRS-bubble PCR reactions were carried out in a total volume of 25μl. The reaction mix contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 1.5mM $MgCl_2$, 200μM dNTPs, 50ng *Not1* bubble primer (Hunter *et al.* 1994), 50ng B2*Not1* primer (Simmler *et al.*, 1991) and 1μl EB18Az:bubble template. The reaction was mediated by 2u *Taq* polymerase. The conditions for this reaction are given in Hunter *et al.*, 1994. PCR products from this reaction were cloned into the *Not1* site of pBluescript KS+ as described below.

Microsatellite	Tm
$(CA)_{15}$	77.5°C
$(GA)_{15}$	77.5°C
$(CGA)_5$	63.8°C
$(CCA)_5$	63.8°C
$(GTT)_5$	50.2°C.
$(CAT)_5$	50.2°C.
$(AAT)_5$	36.5°C.
$(CCG)_5$	77.5°C.
$(AGA)_5$	50.2°C.
$(TCC)_5$	63.8°C

Table 6: -Tm values for microsatellite repeats

	Sequence
Not1 bubble vector (top strand) (Hunter *et al.*, 1994)	5' AAG GAG AGG ACG CTG TCT GTC GAA GGT AAG GAA CGG ACG AGA GAA GGG AGA G 3'
Not1 bubble vector (bottom strand) (Hunter *et al.*, 1994)	5' CTC TCC CTT CTG GCG GCC GCA GTT CGT CAA CAT AGC ATT TCT GTC CTC TCC TT 3'
Not1 bubble primer (Hunter *et al.*, 1994)	5' GCG GCC GCA GTT CGT CAA CAT AGC ATT TCT 3'

Table 7:- Bubble PCR vectorette and bubble primer sequences

2.3.6) The cloning of IRS-PCR and IRS-bubble PCR fragments

The cloning of IRS-PCR products was either carried out from individual fragments, isolated from low melting point agarose gels, or directly from total IRS-PCR products. The cloning of IRS-bubble PCR fragments was carried out directly from the original PCR products.

2.3.6a) The host bacterial cell - XL1-BLUE MRF'

The host cells used in the cloning process were XL1-BLUE MRF'. These cells are a restriction deficient strain of Escherichia coli, and are deficient for all known restriction systems, although the modification systems are intact. In addition, XL1-BLUE MRF' cells are endonuclease deficient and recombination deficient, ensuring stability of any inserts cloned with this system. This bacterial strain contains the F' episome, which contains a tetracycline resistance gene, and also fragments of the bacterial *lacZ* gene. This, in the presence of vector plasmids containing no insert, on a medium containing the lactose analog X-gal, and the lactase inducer IPTG, produces colonies which are blue.

2.3.6b) The production of super-competent cells

The use of the procedure outlined below allows the production of cells which are 'supercompetent', and produce transformation frequencies far in excess of those produced by conventional approaches.

An culture of XL1-BLUE MRF' cells was grown to lag phase, in the presence of tetracycline. This culture was used to inoculate 200ml sterile LB medium, supplemented with 12.5μg tetracycline, which was then incubated at 37°C until the optical density was 0.45 at a wavelength of 550nm, as measured by spectrometer. All pipette tips, and reaction tubes used in this experiment were sterile, and had been maintained at -20°C overnight, and transfered to a 4°C cold room imediately prior to the procedure. The bacterial culture was divided into eight aliquots, each of these was treated in the same manner. The cells in each tube were pelleted by centrifugation at 4°C for 15 minutes at 2.5K. The cell pellet was washed in 10ml TFB1 solution, the cells were left in contact with this for 15 minutes. The cells were again collected by centrifugation, and resuspended in 1ml TFB2 solution. The resulting 8ml suspension of competent cells were divided into 200μl aliquots, and stored at -70°C.

2.3.6c) The cloning vector - pBluescript SK+

pBluescript SK+ is a phagemid. This vector is essentially a plasmid with a phage origin. This vector contain a multiple cloning site (containing 21 unique restriction sites), flanked by the T3 and T7 RNA polymerase promoter sequences. These provide an ideal target for sequencing primers. The plasmid also contains the terminal portion of the *lacZ* gene, within which the multiple cloning site has been positioned. Therefore, when transformed into a bacterial strain which contains all the other essential components of the *lacZ* gene, and grown on a substrate containing the lactose analog X-gal, and the lactase inducer IPTG, the colonies produced have a blue colour, except where the *lacZ* gene fragment in the plasmid has been interrupted by the inclusion of an insert into the multiple cloning site. The phagemid contains an ampicillin antibiotic resistance gene, for use as a selectable marker. DNA from pBluescript SK+ can be recovered either as single or double stranded DNA.

2.3.6d) The preparation of the plasmid vector

XL1-BLUE MRF' cells containing pBluescript SK+ phagemids were streaked out onto an agar plate, containing 50μg ampicillin per ml. Four separate colonies were picked from this plate, with a sterile toothpick, and transferred to 5ml LB broth, containing 50μg ampicillin per ml. This solution was cultured for 18 hours. plasmid DNA was extracted from the bacterial cells by the Promega 'wizard' miniprep kit, according to the manufacturers instructions. The resulting plasmids were resuspended in 50μl TE buffer.

10μg plasmids (at 200ng per μl) were then treated with 50u *EcoR1*, in a total volume of 100μl. This restriction enzyme has a unique restriction site within the multiple cloning site. The digested plasmids were then treated with calf intestinal alkaline phosphatase (100u) at room temperature for 30 minutes, to remove the 5' phosphate molecules, and thus prevent self-ligation of the vector. Both restriction digestion and alkaline phosphatase treatment were carried out in One-Phor-All buffer, as supplied by Pharmacia Biotech, U.S.A.

The restricted, phosphatase treated plasmids were then purified by the Promega 'Magic' PCR preps kit, according to the manufacturers instructions. The plasmids were then redissolved in 50μl TE buffer. This kit was designed for the purification of PCR products prior to cloning, but was found to be adequate for the purification of small plasmids, such as pBluescript SK+.

2.3.6e) The preparation of the IRS-PCR or IRS-bubble PCR products for cloning

In the case of fragments that had been isolated from low melting point agarose gels in the initial stages of the experiment, 1μg reamplified fragments were digested with 50u *EcoR1*, in a volume of 50μl overnight, at a temperature of 37°C. The digested DNA was then purified using the Promega 'Magic' PCR preps kit, according to the manufacturers instructions. In the case of IRS-PCR fragments to be cloned directly from the initial PCR product mixture and IRS-bubble PCR fragments, the entire PCR reaction was digested with 50u and 20u *EcoR1* in volumes of 100μl and 20μl respectively, and then purified in the same manner. These fragments were resuspended in 50μl and 20μl of TE buffer respectively.

2.3.6f) The ligation reaction

The ligation reactions were carried out in a total volume of 30µl. Experiments were carried out to determine the optimal ratio of insert to vector. For some fragments, ratios of 2:1 were adequate, but for the majority of fragments, 4:1 gave the best results. The reaction mixtures contained 1X T4 DNA ligase buffer (as a source of necessary ions, and ATP), 200ng purified digested and phosphatase treated pBluescript SK+ and 400ng - 800ng purified restricted IRS-PCR or IRS-bubble PCR DNA, and 10u T4 DNA ligase. The reactions were incubated at 15°C for 18 hours, after which they were maintained at -20°C until transformation.

2.3.6g) The transformation procedure

100µl XL1-BLUE MRF' competent cells per transformation reaction were thawed on ice, and then added to the 30µl ligation reaction. The cell suspension was mixed well, and maintained on ice for 45 minutes. After this time, the cells were heat-shocked by exposure to a temperature of 43°C for 2 minutes. This served to increase the permeability of the bacterial cell wall to the foreign insert DNA. The cell suspension was then returned to the ice for 5 minutes. The uptake of the ligated DNA by the cells was then complete. 500µl LB broth, supplemented with 5µl 1M $MgSO_4$, was then added. This mixture was then incubated with constant agitation at 37°C for one hour, to allow expression of the ampicillin resistance proteins. The transformants were plated onto LB agar megaplates containing ampicillin at a working concentration of 50µg/ml, X-gal at a working concentration of 80µg/ml, and 20mM IPTG. Where replica plates were to be taken, the transformants were plated onto nylon filters, but where no further screening was to take place, the transformants were plated directly onto the agar. The plates were incubated in an inverted position, at a temperature of 37°C for 18 hours.

2.3.7) Detection of microsatellite containing clones

2.3.7a) Replica plating technique

In cases where the insert fragment was judged to contain a microsatellite repeat on the basis of the hybridization studies, positive clones were identified by replica screening with the relevant oligonucleotide probe. The original transformation plates were used to prepare

duplicate replica lifts, onto Hybond N+ hybridization membranes, which were placed on LB-ampicillin plates (as used in the original transformations). The plates were then incubated at 37°C, in an inverted position, to allow growth of the transferred clones. The master plates were also incubated for a period of one hour, to regenerate the original colonies. After this time, colonies were clearly visible on the replica plates. The plasmid DNA was released by treatment of the filters for 7 minutes with denaturing solution, and 6 minutes with neutralizing solution soaked on 3MM filter paper pads (See section 2.1 for details of solutions). The filters were then washed in 2X SSC to remove any traces of agar. The DNA was fixed to the filters by U.V. crosslinking for 15 seconds followed by baking in an 80°C oven for 2 hours.

2.3.7b) The detection of positive clones by hybridization

The oligonucleotide probes were end-labelled with γ^{32}P-dATP, as in the original experiments. For dinucleotide repeats, the hybridization solution was 1.2X SSPE, 10% dextran sulphate (w:v) and 2X SSC. Because of the lower hybridization temperatures required for the trinucleotide repeats with lower (A+T) contents (see table 8), the SSPE concentration was increased to 4X SSPE, for the detection of trinucleotide repeats. This allowed an increase in

Microsatellite	$(CA)_{15}$	$(GA)_{15}$	$(GTT)_5$	$(CAT)_5$
Tm for 4X SSPE (°C.)	-	-	58.21	58.21
Tm for 1.2X SSPE (°C.)	68.0	68.0	39.0	39.0

Table 8: - Tm values for (CA), (GA), (GTT) and (CAT) repeats in 1.2X SSPE and 4X SSPE

temperature to a level where the SDS remained in solution. This reduced the level of the background considerably. The Tm values calculated for the repeats tested are given in table 8. In order to reduce the number of hybridizations required for clones containing more than one type of microsatellite repeat, probes with identical Tm values were pooled. Hybridizations was carried out at 5°C below Tm, for a period of 3 hours. Filters were washed three times at 5°C below Tm, as in the original experiments, and once at Tm, for 2 minutes exactly. The washes were carried out in 2X SSPE/ 0.1% SDS for the dinucleotide repeats, and in 4X SSPE/0.1% SDS for the trinucleotide repeats. The filters were dried briefly, and exposed to Hyperfilm

autoradiography film for 24 to 72 hours, depending on the intensity of the signal, as determined by Geiger Muller counter.

2.3.8) Isolation of individual clones, and extraction of plasmid DNA

Individual white colonies (as the blue colonies represented vector plasmids with no insert) were picked from the growth surface with sterile toothpicks, and were transferred to sterile culture tubes containing 5ml LB, and 250μg ampicillin. These cultures were incubated at 37°C for 18 hours, with constant agitation.

When the growth was complete, a proportion of the culture was removed for the preparation of a glycerol stock (cultured cells were mixed with 15% glycerol (v:v), and stored at -70°C, as a permanent resource). The remainder of the cells were pelleted by centrifugation for 15 minutes at 2.5K, at a temperature of 4°C. The plasmids were released from the bacterial cells by alkaline mediated cell lysis (according to the procedures outlined in the Promega 'Magic' minipreps kit) followed by removal of the cell debris by centrifugation. The plasmids were purified by means of the Promega 'Magic' minipreps kit, and resuspended in 80μl TE buffer. Later minipreps were carried out to the same procedure, but using the Promega 'Wizard' minipreps kit.

2.3.9) The development of techniques to distinguish between mouse and hamster clones

Although the IRS-PCR primers were apparently species specific, low level amplification of hamster sequences might be expected to occur. This suggests that a proportion of the DNA fragments produced will be of hamster, rather than mouse, origin. Initially, total mouse DNA labelled by the random primer method with α^{32}P-dCTP was used to probe plasmids immobilized on nylon filters. This hybridization was carried out in 5X SSPE, 10% dextran sulphate (w:v) and 2% SDS, at a temperature of 68°C, in the presence of 2mg unlabelled hamster DNA, in a total volume of 15ml. This reduced any hybridization signal due to the hamster repetitive sequences.

2.3.9a) Dot blot techniques - preparation of filters

Dot blot techniques were also used to identify hamster clones. Samples of mouse and hamster DNA were heated to 95°C, an equal volume of 20X SSC was then added. 1.0μg aliquots of heated mouse and hamster DNA were then alternately spotted onto strips of hybridization membrane, which had been prewetted with 10X SSC. The position of the hamster dots was marked. The membrane was then placed on a pad of filter paper soaked in neutralization solution (Hybond N+) for 5 minutes. After this time, the membrane was transferred to a similar pad soaked in neutralizing solution (Hybond N+) for 1 minute. The DNA was fixed to the filter strips by U.V. treatment for 15 seconds, followed by baking in an 80°C oven for 2 hours.

2.3.9b) Preparation of probes

Fragments cloned into the plasmid were released by restriction digestion (*EcoR1* digestion in the case of L1, L1/B2 and B2(*EcoR1*) fragments, *Not1* in the case of B2(*Not1*) and IRS-bubble PCR fragments). The digestion was carried out according to the instructions given by the manufacturer.

Digested fragments were electrophoresed on 0.8% agarose gels, in 1 X TBE buffer. The gels were run at 250V/cm, for about one hour. The gels were stained with ethidium bromide, and viewed under U.V. light. Individual inserts were excised from the gels using sterile scalpels, and were purified by centrifugation through glass wool, to remove traces of agarose. The final concentration was approximately 10ng/μl. 25ng insert was labelled to high specific activity with α^{32}P-dCTP by the random priming method, with the 'prime-it' version 2.0 kit, according to the manufacturers directions.

2.3.9c) Hybridization techniques

The probes were boiled for 5 minutes, in order to render them single stranded. They were then added to heat-sealed plastic bagging, containing the membrane strips bearing mouse and hamster DNA. The hybridization solution was 5X SSPE, 10% (w:v) dextran sulphate, 2% SDS. The hybridization was carried out at 68°C for 18 hours. The filters were washed for 30 minutes, at 68°C in 2X SSPE/0.1% SDS, and then exposed to Hyperfilm autoradiography film

at -70°C for 24 to 48 hours, depending on the intensity of signal, as determined by Geiger Muller counter. The position of the hamster spots was marked on the film, with the point of a scalpel.

2.3.10) Sequencing reactions

The sequence of the insert was determined by double-stranded dideoxy sequencing. The Sequenase version 2.0 sequencing kit was used for all reactions. 5-8μg plasmid DNA was first denatured by treatment with 0.1 volume of 2N NAOH/ 2mM EDTA for 30 minutes, in a total volume of 90μl. The DNA was then precipitated by 2 to 4 volumes of 100% ethanol, following adjustment of the ion concentration with 0.1 volume 3M sodium acetate. The DNA was pelleted by centrifugation, and washed once with 70% ethanol. The pellet was dried at 95°C for 5 minutes, and then resuspended in 7μl deionised distilled H_2O.

The DNA was then sequenced using primers designed from the T3 or T7 RNA promoter genes present in the vector, or from the vector specific KS primer (see table 9 below).

Sequencing primer	Sequence
KS	5' AAT AAC CCT CAC TAA AG 3'
T7	5' GCC TAA TAC GAC TCA CTA TA 3'
T3	5' CGA GGT CGA CGG TAT CG 3'

Table 9:- Details of sequencing primers

These primers flank the site of insertion. The sequencing reactions contained 1X Sequenase buffer (200mM Tris-HCl pH 7.5, 100mM $MgCl_2$, and 250mM NaCl), 0.5pmol/μl T3 or KS primer and 7μl denatured DNA from the denaturation reaction, in a total volume of 10μl. To this, the labelling reaction mix of 0.01M DTT, 1:10 diluted labelling mix, and 0.5μl α^{35}S-dATP was added, together with 3.25u of Sequenase version 2.0 T7 DNA polymerase and 0.6u inorganic pyrophosphatase (to eliminate the sequence dependent reversal of the polymerase reaction by pyrophosphorolysis - this can lead to weakening of the signal from certain areas of the insert). The labelling reaction was carried out for 5 minutes, at room temperature. The reaction was then terminated by the addition of 3.5μl of the reaction to each of four tubes containing the termination mixture containing 80μM each deoxynucleotide, and

8μM of one dideoxynucleotide in 50mM NaCl. The termination was carried out for 15 minutes, at 42°C. The reactions were then completed by the addition of 4μl 95% formamide, 20mM EDTA, 0.05% bromophenol blue and 0.05% xylene cyanol. The sequencing reactions were carried out according to the instructions supplied with the Sequenase kit. All reagents were maintained on ice. For clones without microsatellite repeats, only the ends of the fragments were sequenced. Where microsatellites were present, sequence was obtained from either side of the repeat. Only partial sequence was obtained for clones. When it was necessary to obtain sequence located further into the clone (for example when repeat sequences were not reachable from one side of the clone, due to their position), a single primer was designed, from which further sequencing reactions were carried out. Sequencing reactions were electrophoresed at 50W for 3 hours. The gels were dried at 80°C, under vacuum, and then exposed to Hyperfilm autoradiography film for 1 to 7 days, depending on the intensity of the signal.

2.4) Primer design and the detection of variation between *Mus musculus* and *Mus spretus*

New markers were generated for mouse chromosome two, by IRS-PCR and IRS-bubble PCR. Sequence information obtained from experiments described in section 2.3 was used to design new PCR primers which could then be used to map the DNA fragments by interspecific backcross analysis. Conditions were established for these primers, and variation between *Mus musculus* and *Mus spretus* was detected by the following techniques:-

I) The manipulation of the PCR to produce *Mus musculus* specific amplification.
II) The detection of length variation deriving from the presence of microsatellite repeats.
III) The detection of length variation deriving from polypurine or polypyrimidine repeats originating from the degeneracy of the 3' end of the SINE and LINE repeats by radioactive PCR.
IV) The detection of RFLVs between *Mus musculus* and *Mus spretus*.
V) The detection of single stranded conformation variations (SSCVs) between *Mus musculus* and *Mus spretus*.

2.4.1) The production of new STSs for loci on mouse chromosome two

2.4.1a) Primer design

Sequence information for the DNA fragments cloned in section 2.3 was used to design new PCR primers. In the case of sequence containing microsatellite repeats, PCR primers were designed which flanked the repeat unit, from sequence taken from one end of the insert. For all other fragments, whether of IRS-PCR or IRS-bubble PCR origin, PCR primers were designed from sequence from both ends of the insert.

Before the new PCR primers were purchased, they were first checked for repetitive DNA content against the GenBank database, using the 'Blast' program offered by the computing services at the HGMP (Altschul *et al.*, 1990).

2.4.1b) The establishment of amplification conditions for new PCR primers

Upon the arrival of new primer pairs, they were divided into 100µl aliquots, as a precaution against contamination (since the primers were designed from plasmid sequence, the presence of airborne plasmid posed a contamination problem, despite measures to contain this). These included the preparation of reaction mixtures and the execution of the PCR in a separate room to that used for plasmid manipulation, and the use of dedicated pipettes, tips and tubes. Occasional cases of contamination were treated by restriction digestion of either primers or master mixes, given the presence of an appropriate restriction site in the plasmid, or by U.V. treatment of primers or master mixes). PCR was carried out in 50µl reaction volumes. The reactions contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 200µM dNTPs, 100ng each primer, and 50ng mouse DNA, for a number of different magnesium concentrations, commonly 1.0mM, 1.5mM, 2.0mm, 2.5mM, 3.0mM, 3.5mM and 4.0mM. The reaction was mediated by 0.5u *Taq* polymerase. The reactions were carried out under standard PCR conditions of 55°C annealing and 72°C elongation. Magnesium titrations were also carried out with 50ng hamster DNA, under the same conditions.

If these conditions did not yield results, the annealing temperature and the length of time at each step was adjusted. 10% glycerol or 10% DMSO were added to the reaction mixture to improve the quality of the PCR product if necessary. Under conditions where more than one band was produced, the stringency of the reaction was increased until this was at a minimal level. The addition of 10% glycerol was used as a means of reducing the number of bands produced. Sequences of each primer pair, and the amplification conditions are given in table 10. PCR products were visualized on 2% agarose gels stained with ethidium bromide (35µg per 70ml gel), and viewed under U.V. light.

2.4.2 Methods of detecting interspecies variation

Once PCR conditions were established which reproducibly amplified DNA from *Mus musculus*, experiments were commenced to identify variation between *Mus musculus* and *Mus spretus* in the PCR product produced by each pair of primers. This was accomplished by the methods outlined below.

2.4.2a) The manipulation of the PCR to produce species specific amplification

Species specific amplification occurs when the sequences of the *Mus musculus* and *Mus spretus* DNA fragments differ sufficiently to prevent annealing of a primer designed from a *Mus musculus* clone to anneal to the *Mus spretus* sequence, or if there is a large insertion between the priming sites in *Mus spretus*. The exact sequence differences between the species cannot be determined without direct sequencing of the homologous fragment, but primer sequences designed from *Mus musculus* sequence will fortuitously fail to amplify *Mus spretus* DNA in approximately 20% of cases (this value is estimated from the primer pairs designed by Love *et al.*, 1990 and Hearne *et al.*, 1991). Under conditions where the *Mus musculus* locus is amplified to a much greater degree than the *Mus spretus* locus, the stringency of the PCR reaction was increased (by increasing the annealing temperature, decreasing the elongation time or decreasing the magnesium concentration in the reaction mixture) until no amplification was obtained from *Mus spretus* DNA. This allowed the fragment to be mapped on an interspecific DNA panel backcrossed to *Mus spretus*. The PCR reactions were carried out in a volume of 50µl. The reaction mixture contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 200µM dNTPs, 100ng each primer, 10ng DNA (from either *Mus musculus* or *Mus spretus*), and a magnesium concentration appropriate to each primer pair. The reaction was mediated by 0.5u *Taq* polymerase. PCR products were displayed on 2% agarose gels. Gels were electrophoresed in 1X TBE buffer, at 200V/cm, stained with ethidium bromide (35µg per 70ml gel) and viewed under U.V. light.

2.4.2b) Variation due to length differences originating from the presence of microsatellite repeats

Where primer sequences were taken from sequences flanking a microsatellite repeat, amplification across the repeat unit detected length variation between species in the majority of cases. This was due to differences in the copy number of the basic repeat between *Mus musculus* and Mus spretus. These differences in allele length can sometimes be seen even on low percentage agarose gels, but in the majority of cases, higher percentage gels can identify variants separated by upwards of six base pairs. PCR reactions were carried out in 50µl or 25µl volumes, according to the requirements of the primers. Reaction mixtures contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 50ng (for 25µl reactions) or 100ng (for 50µl reactions) each PCR primer, 200µM dNTPs, 10ng DNA and a magnesium concentration

appropriate to the primers. PCR reactions were mediated by 0.25u (for 25µl reactions) or 0.5u *Taq* polymerase (for 50µl reaction). Amplifications were carried out according to the conditions stated for each pair of primers in table 10. PCR products were visualized on gels of an appropriate agarose percentage to distinguish the variation - gel concentrations varied from 3% to 6%, according to the separation of the alleles. However, due to the difficulty of melting and pouring higher percentage gels, markers producing very small variations in product size were mapped by radioactive PCR, as described in section 3.4.3d.

3.4.3c) Identification of restriction fragment length variants (RFLVs)

Since the majority of fragments cloned did not contain microsatellites, attempts to identify RFLVs were undertaken. PCR primers were designed from the ends of the insert, ideally to amplify products of 1.0 to 1.5kb, which would be more likely to contain restriction sites which might not be common to both *Mus musculus* and *Mus spretus*. RFLVs can be generated by mechanisms such as insertion or deletion events in one species, leading to differences in the size of fragments produced, or by point mutation events, which disrupt restriction sites in one species.

PCR amplifications were carried out in 50µl reaction volumes, containing 1X PCR buffer (Advanced Biotechnologies Ltd.), 100ng each PCR primer, 200µM dNTPS, 10ng *Mus musculus* or *Mus spretus* DNA, and magnesium according to the requirements of each primer pair. The reaction was mediated by 0.5u *Taq* polymerase. Following verification that the samples had amplified, restriction digestions comparing *Mus Musculus* and *Mus spretus* PCR products were set up in 20µl volumes, with 10u frequent cutter restriction enzyme, such as *RsaI*, *DdeI*, *HinfI*, *HincII*, *HaeIII*, *AluI*, *Tru9AI*, *XbaI MspI*, *TaqI* and *MboI*. If no RFLV was detected with these enzymes, others such as *PstI*, *EcoRV*, *BamH1*, *HindIII*, *BglII*, *SstI*, *HpaII* or *KpnI* were tested. All restriction digestions were carried out at a temperature of 37°C, with the exception of *TaqI*, which was carried out at 65°C. All digestions were carried out for 18 hours.

Digested PCR products were visualized on 4% agarose gels, stained with ethidium bromide (35µg per 70ml gel) and viewed under U.V. light.

2.4.2d) The detection of small length differences by radioactively labelled PCR

PCR products which appear to show no length variation between *Mus musculus* and *Mus spretus* DNAs on even high percentage agarose gels can often be shown to be variant on polyacrylamide gels. The greater resolution allowed by these gels allows length differences of only one base pair to be detected. In order to produce the maximum resolution, 6% polyacrylamide sequencing gels were used to investigate whether there was any length variation between the two species. The use of these gels requires the use of different means of visualizing the products from those used so far, because of the differences in the set-up of the electrophoresis equipment, and in the thickness of the gels - whereas the average agarose gel is approximately 0.5 to 1cm in depth, and therefore easy to handle without breakage, polyacrylamide sequencing gels are only of the order of a millimeter thick, and very fragile. Therefore, the PCR products were radioactively labelled with α^{32}P-dCTP during the course of the amplification. This was achieved by the inclusion of α^{32}P-dCTP into the reaction mixture. The concentration and composition of the nucleotide mix was adjusted to compensate for this. Each reaction was carried out in a 10µl reaction volume. The reaction mixture contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 200µM each dATP, dTTP and dGTP, 10µM dCTP, 0.36µl of a 1:75 dilution of stock α^{32}P-dCTP, 5ng *Mus musculus* or *Mus spretus* DNA, an appropriate magnesium concentration for the primers used and 0.25u *Taq* polymerase. The cycling conditions were identical to those used for non-radioactive PCR, and are given in table 10.

Once variation had been detected between *Mus musculus* and *Mus spretus*, new molecular markers were mapped by interspecific backcross analysis.

2.4.2e) Single stranded conformation variation (SSCV) analysis

Single stranded conformation variation (SSCV) arises by virtue of point mutations, or small rearrangements, which alter the conformation (and thus the motility) of single stranded DNA under non-denaturing gel electrophoresis. (Orita *et al.*, 1898a; 1989b). Although the technique was originally carried out on digested genomic DNA, the same methods can be employed to detect single base differences in PCR products amplified from different species, such as *Mus musculus* and *Mus spretus*.

PCR reactions were carried out according to the conditions defined for each primer pair. Standard amplifications were carried out in 50µl reactions. The reaction mixture contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 200µM dNTPS, 100ng each primer, magnesium concentrations appropriate to each primer pair and 10ng DNA. The PCR products were electrophoresed on 8% polyacrylamide slab gels, containing 10% glycerol, and 0.5X TBE. The electrophoresis was carried out at 200V for 12 - 16 hours, depending on the size of the PCR product.

DNA was visualized by silver staining. The gel was first washed in 10% ethanol, 0.5% acetic acid solution, for two 3 minute periods. The gel was then treated with 1g/L silver nitrate solution for 10 minutes, with agitation. The gel was then briefly rinsed twice with distilled water, to remove excess silver nitrate. A small amount of stain was then added, which was then poured off, and the gel again rinsed with distilled water. Finally, the gel was exposed to the stain for 15 to 20 minutes, which highlighted homoduplex and heteroduplex DNA.

2.5) Determination of the map position of new molecular markers from mouse chromosome two.

2.5.1) Mapping of new markers to mouse chromosome two

2.5.1a) The Jackson Laboratory interspecific backcrosses

Two panels of DNA from interspecific backcross animals were obtained from the Jackson laboratory. Both panels involved the laboratory inbred strain C57BL/6J, and *Mus spretus*. Panel one (BSB) consists of 94 N2 offspring from the C57BL/6J X *Mus spretus* F1, backcrossed to C57BL/6J. Panel two (BSS) consists of 94 N2 animals from the C57BL/6J F1, backcrossed to *Mus spretus* (Rowe *et al.*, 1994). These panels have been characterized for a number of loci on all mouse chromosomes, and thus are useful for the initial localization of new markers. The resolution obtainable from these crosses has been estimated at between 1 and 5cM (Rowe *et al.*, 1994). New molecular markers obtained from the IRS-PCR and IRS-bubble PCR cloning experiments were first mapped on one or other of these backcrosses to establish linkage to mouse chromosome two. The choice of which panel a new marker was mapped on largely depended on which panel illustrated the variation between C57BL/6J and *Mus spretus* alleles in the clearest manner. In the case of markers with which no amplification of *Mus spretus* DNA was noted, there was no choice but to map on panel two (BSS). In this case, all backcross animals possessed *Mus spretus* alleles, only the C57BL/6J allele was segregating. The existing marker *D2Mit52* was also mapped on panel two, as a reference locus for the construction of the map.

2.5.1b) PCR amplification of backcross DNAs

PCR reactions were carried out in a total volume of 50µl. Reaction mixtures contained 1X PCR buffer (Advanced Biotechnologies Ltd.), 200µM dNTPS, 100ng each primer, magnesium concentrations appropriate to the primer pair in question (see table 10 for primer details), and 50ng backcross DNA from the required panel. In cases where variation was to be detected by radioactive PCR followed by analysis on sequencing gels, the reaction mix was altered to include α^{32}P-dCTP, as outlined in section 2.1.4d. All reactions were carried out on Hybaid Omnigene thermal cyclers, with the exception fo the radioactive PCRs, which were carried out in Hybaid TR2 thermal cyclers, for safety reasons. Variation was detected as outlined in section 2.4 (see table 11 for details of techniques used), and backcross animals were scored as either BB or BS in the BSB cross, or SS or BS in the BSS cross. The data was then

submitted to the Jackson database for detailed analysis. The data was analysed by the Map Manager program (Manly, 1993).The program was originally written for the analysis of data from recombinant inbred lines, but has been adapted for the analysis of interspecific backcross data. The program calculates first the probability that any two loci are linked, and then calculates map order and map distance. The map order is determined by minimising the number of double recombinants required to produce a particular haplotype, and the map distance is calculated from the number of recombination events between any two adjacent loci. Sufficient loci have now been placed on the maps that definate haplotypes have now emerged for each animal, and double recombination events are easy to distinguish. New loci can usually be allocated rough map positions by comparison of the segregation patterns produced from the new marker under test, and existing markers already mapped to mouse chromosome two. Print outs were received, giving exact positions for new markers, together with details on the presence of transmision distortion, or other phenomenon.

2.5.2) Detailed mapping of new markers mapping in areas of interest

New molecular markers mapping to regions of interest on mouse chromosome two (such as the areas where the genes for the ragged, wasted and ulnaless mutations are thought to lie) were mapped on specialized backcrosses segregating the mutant of interest. Three such backcross panels are available at present, segregating ragged, wasted and ulnaless respectively.

2.5.2a) The ragged backcross

The ragged backcross was set up from a *Ra*/+ heterozygote X *Mus spretus* interspecific cross (Abbott *et al.*, 1994). The *Ra*/*Mus spretus* female F1 animals were then backcrossed to a mouse of the strain C3H/HeH. This backcross produced (at the time of writing) 166 N2 animals, of genotype *Ra*/C3HHeH or *Mus spretus*/C3HHeH (see figure 19a). These two types of progeny are easy to distinguish from each other - *Ra*/C3HHeH animals have the distinctive 'ragged' coat of ragged heterozygotes whereas C3HHeH/*Mus spretus* animals are completely wild type. This backcross has been typed for several loci on mouse chromosome two, especially in the region of the ragged mutation.

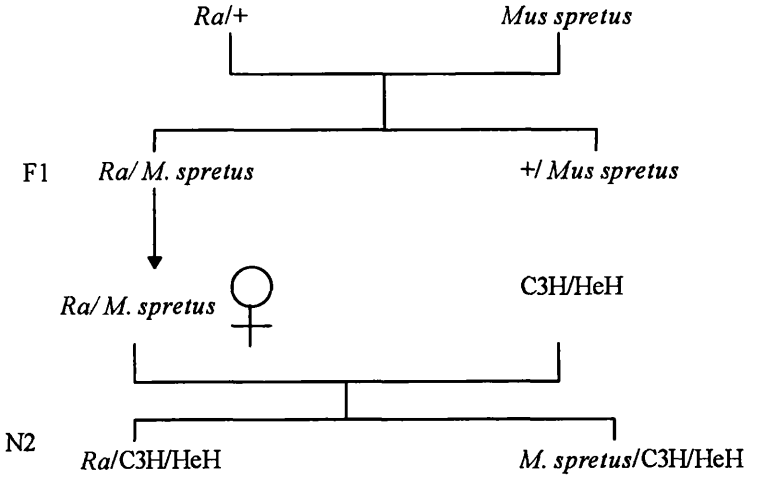


Figure 19a: - The ragged backcross

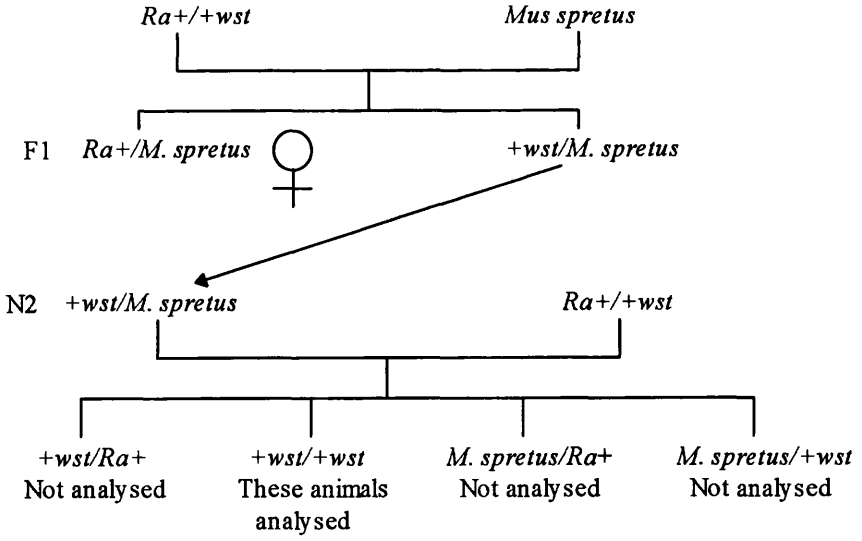


Figure 19b: - The wasted backcross

2.5.2b) The wasted backcross

The wasted backcross posed an additional problem, in that wasted heterozygous mice are indistinguishable from the wild type. To circumvent this problem, this backcross was maintained on a background also segregating ragged. The parental animals were *Ra+ / +wst* and *Mus spretus*. This produces *Ra+ / Mus spretus* and *+wst / Mus spretus* F1 animals. The *+wst/Mus spretus* F1 females (identified by their non-ragged coat) are then backcrossed to *Ra+ / +wst*. This produces the following offspring:- a) *+wst/+wst*, b) *+wst/Mus spretus*, c) *Ra+/+wst* and d) *Ra+/Mus spretus* N2 animals (Abbott *et al.*, 1994)(see figure 19b). Only animals which were visibly *wst/wst* homozygotes (on the basis of tremor, ataxia and weight loss) were included in the analysis, as *wst* heterozygotes were indistinguishable from wild type. This method of identifying wasted heterozygotes is made possible because of the close linkage between ragged and wasted (Peters *et al.*, 1994). This backcross yielded 94 animals (at the time of writing), and has also been typed for a variety of markers on distal mouse chromosome two (Abbott *et al.*, 1994).

2.5.2c) The ulnaless backcross

The parental strains of the ulnaless backcross were *Ul/+* and the *Mus musculus castaneus* strain CAST/Ei, rather than *Mus spretus*. Phenotypically ulnaless F1 female offspring Ul/CAST/Ei were then backcrossed to the inbred strain CBA. This backcross produced *Ul*/CBA and CAST/Ei/CBA N2 offspring (see figure 19c). The background of the ulnaless stock was very heterogeneous, containing C3H, CBA, 101, and a mutation testing stock (Dr. C. Abbott, personal communication). The interspecific backcross generated 93 N2 animals. These animals have been characterized for markers in the *D2Mit11* to *D2Mit30* region, within which the ulnaless mutation is located.

2.5.2d) Mapping new molecular markers on the ragged, wasted and ulnaless backcrosses

Markers previously mapped on the Jackson backcross panels to the region in which ragged, wasted and ulnaless are thought to lie were then used to ampify DNA from the relevant backcross. It was judged unnecessary to type the entire backcross for markers already mapped on the Jackson crosses - only animals showing evidence of recombination events in the region of the chromosome in which the new marker was located were typed. Markers lying in the *D2Mit11* - *D2Mit30* region of mouse chromosome two were typed on the ulnaless cross,

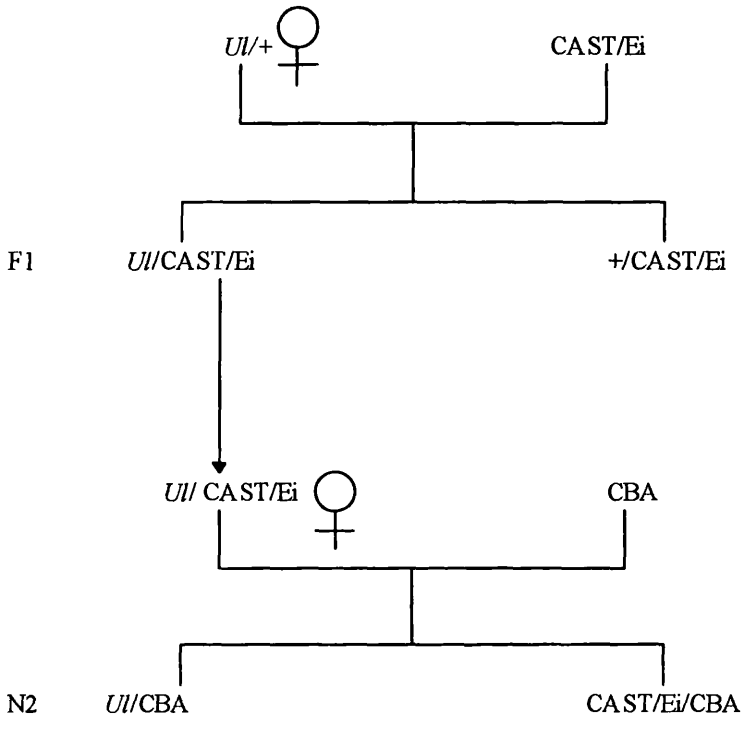


Figure 19c: - The ulnaless backcross

whereas critical recombinants from the *Gnas* - *tel* region were typed for the ragged and wasted backcrosses.

PCR reactions were carried out for each new marker according to the conditions established for mapping on the Jackson backcross. For markers to be mapped on the ulnaless backcross, a source of variation between CBA and CAST/Ei was first determined, as in section 2.4.

Reactions were carried out in 50µl total volume, with the exception of the ulnaless cross, in which reactions were carried out in a total volume of 30µl. All reactions were carried out in Hybaid Omnigene thermal cyclers.

Data were analysed by similar methods to those employed in the Map Manager program (Manly, 1993). Several loci had already been placed on all the individual backcrosses, and new markers were allocated map positions on the basis of the minimum number of recombination events required to generate a particular haplotype. Genetic distance between loci was calculated as a recombination percentage, and standard error was calculated as a function of the square root of 'p' multiplied by 'q' over the square root of 'N', where P was the recombination frequency (expressed as a decimal), $P + q = 1$, and N was the number of animals analysed.

Chapter 3

Results

3.1.1) The construction of two new panels of somatic cell hybrids

Two new panels of mouse:Chinese hamster somatic cell hybrids were constructed. The Chinese hamster parent was the $HPRT^-$ cell line V79TOR2, the mouse parent (spleen cells) in the panels carried the T(2;8)2Wa translocation in one panel, and the T(2;16)28H translocation in the other panel. 44 individual clones were isolated from the two panels :- 25 from the T(2;8)2Wa panel, and 19 from the T(2;16)28H panel. The characteristics of each clone are given in tables 10a and 10b.

3.1.1a) Cell morphology, colony morphology and growth rate of isolated hybrid cell lines

As can be seen from tables 10a and 10b, the colony morphology, cell morphology and growth rates of the clones differed from line to line. Some of the hybrid lines (T28H-2, T28H-5, T2Wa-11a) produced showed similarities in the cell morphology to the hamster parental cell line V79TOR2, which also showed spindle-shaped cells. Other lines (T28H-20, T2Wa-15, T2Wa-16) showed more similarities to the rounder cells derived from mouse spleen. Cells of other lines appeared to have a morphology all of their own (T28H-9, T28H-18b, T2Wa-14). These lines showed characteristics of both parental lines. Similarly, the colony structure and growth rate seemed to resemble one or other parent, or a combination of the two. The hamster cell line showed very fast growth, with little evidence of contact inhibition - indeed, these cells continued to grow until all nutrients had been exhausted in the medium, at which point they detached from the base of the flask, and died. This trait was also seen in several of the hybrid lines, notably T28H-6 and T2Wa-5. No estimate of growth was available for the mouse parental cell line, as it was derived from primary cells. however, the growth rates of primary cells in culture are known to be slow. Some hybrid lines showed extremely slow growth (T2Wa-1, T2Wa-15), in many cases not surviving the transfer to fresh culture flasks following isolation. These cell lines may have showed many similarities with the mouse parental cell line, perhaps because they retained the majority of the mouse chromosomes.

Clone name	colony structure	cell morphology	growth rate
T28H-1	Open, loose association	large cells, rounded	moderately fast
T28H-2	sparse colonies, dense association of cells	small spindle-shaped cells	slow
T28H-3	/	/	Cells did not survive transfer from parent flask
T28H-3b	Dense colonies, but haphazard structure	small round cells	fast
T28H-4	loose association of cells in colony, rounded shape	small rounded cells	fast
T28H-5	colonies are circular, structure is loose.	elongated, spindle-shaped cells	fast
T28H-5b	round colony structure	rounded cells	moderately fast
T28H-6	rounded colonies	small, round cells	very fast
T28H-7	rounded colonies	cells less rounded	moderately fast
T28H-8	rounded colonies, with loose structure	rounded cells	moderately fast
T28H-8b	loosely circular, loose association of cells	cells less rounded	moderately slow
T28H-9	colonies roughly circular	elongated, tapered cells	moderately fast
T28H-10	small, round colonies	rounded cells	slow
T28H-11	sprawling colonies, rounded structure	rounded cells	moderately fast
T28H-12	/	/	Cells did not survive transfer from parent flask

Table 10a:- Characteristics of each hybrid clone isolated from the T(2;16)28H panel

T28H-13	Discrete, rounded structure	small, round cells	moderately slow
T28H-14	small, rounded	small cells, less rounded	moderately fast
T28H-15	rounded colonies	small, tapered cells	moderately fast
T28H-16	rounded colonies	small, tapered cells	moderately slow
T28H-17	small, rounded	large, rounded cells	slow
T28H-18a	/	/	cells did not survive transfer from parent flask
T28H-18b	large, rounded colonies	large, tapered	fast
T28H-19	rounded colonies	large cells	slow
T28H-19b	not enough cells to determine	large, rounded	very slow
T28H-20	small round colonies, looses association between cells	small, rounded cells	fast
T28H-21	large, round colonies	large, round cells	very fast

Table 10a (cont.):- Characteristics of each hybrid clone isolated from the T(2;16)28H panel

clone name	colony structure	cell morphology	growth rate
T2Wa-1	small, round colonies, loose association between cells	small, rounded cells	very slow
T2Wa-1b	small, rounded colonies	small, rounded cells, dense structure	very slow
T2Wa-2	large, round colonies, loose structure	large, spindle-shaped cells	moderately fast
T2Wa-3	large, round colonies	large, tapered cells	fast
T2Wa-4	very dense, rounded colonies	rounded cells	moderately fast
T2Wa-5	discrete, round colonies	small, rounded cells	very fast
T2Wa-6	small, scattered colonies, borders not well defined	small, rounded cells	moderately fast
T2Wa-7	/	/	cells did not survive transfer from parental flask
T2Wa-8	small, round colonies	small, rounded cells	moderately slow
T2Wa-9	small, sparse colonies	small cells	moderately slow
T2Wa-10	ill-defined boundaries to colonies	small, rounded cells	slow
T2Wa-11a	small, round colonies	small, spindle-shaped cells	fast
T2Wa-11b	small, discrete colonies, dense structure	small, round cells	fast
T2Wa-11c	loose association of cells, roughly circular	very small, round cells	fast

Table 10b:- Characteristics of each clone isolated from the T(2;8)2Wa hybrid panel

T2Wa-12	/	/	cells did not survive transfer from parental flask
T2Wa-13	small colonies	small, round cells	slow
T2Wa-14	small, dense colonies	large, round cells	moderately fast
T2Wa-15	very small colonies	small, round cells	very slow
T2Wa-16	small, discrete colonies	small, round cells	very slow

Table 10b (cont.):- Characteristics of each clone isolated from the T(2;8)2Wa hybrid panel

3.1.2) The characterization of the hybrid panels

The new hybrid lines were grown to bulk culture, and DNA was extracted from them, as outlined in the methods chapter. Only 11 clones from the T2Wa panel, and 15 clones from the T28H panel were cultured in this manner, and characterized. The remaining clones were not characterized either because they did not survive in culture, or because they harbored heavy fungal infection, and it was judged to be an unnecessary risk to the other clones to maintain them in culture. The new panels of somatic cell hybrids were characterized using the mouse specific PCR primers outlined in the materials and methods. Each line was characterized for at least two markers per chromosome, where available. The translocation chromosomes were characterized in more detail. An illustration of the results from the PCR hybrid characterization is shown in figure 20. Mouse positive controls and hamster negative controls were carried out in each case. Presumed heterogeneity in the hybrid population (as assessed by reduction in the amount of PCR product for a set of markers on a particular chromosome, relative to the other chromosomes in the same hybrid) was noted in some cases. Heterogeneity for the translocation chromosomes was also seen, notably in T28H-3b, and T2Wa-1. This reduction in PCR signal intensity was not dramatic, but does indicate differences in the chromosome content of cells within the clone. The chromosome content of each hybrid cell line is given in tables 11a and 11b.

3.1.2a) Analysis of segregation patterns of the two new panels

It can be seen that segregation has occurred to a large degree in some hybrid clones, and not in others. For the purposes of the analysis of the overall segregation rate, chromosomes

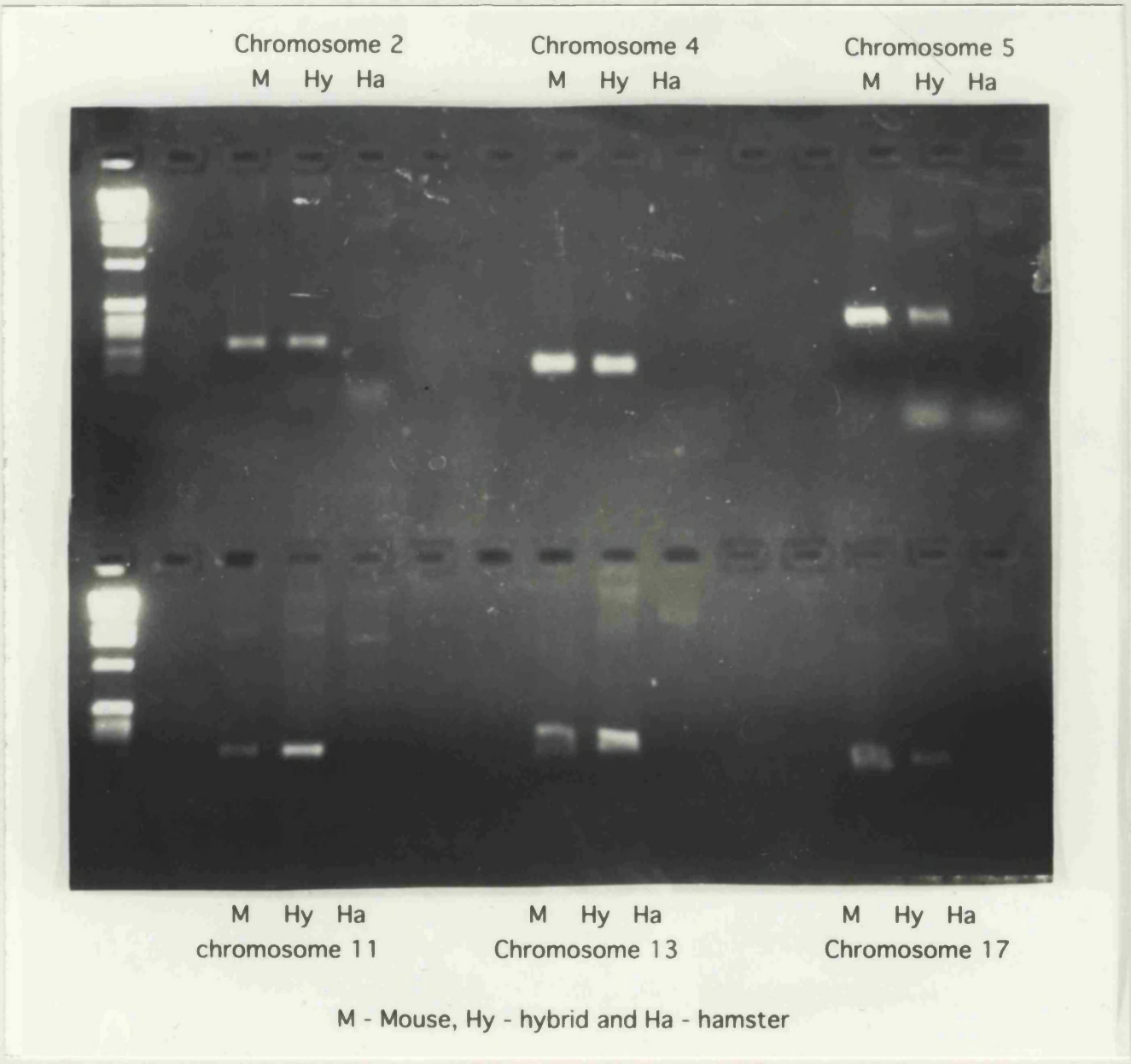


Figure 20: - The characterization of hybrid lines by species specific PCR

The figure shows the results obtained from species and chromosome specific amplification of DNA from mouse cells (positive control, marked 'M'), hybrid test line (T2Wa 1, marked 'Hy') and hamster cells (negative control, marked 'Ha'). It can be seen that no amplification was noted in the hybrid line for any of the markers tested, therefore, the signals detected in the hybrid cell lines must derive from the mouse chromosome. The markers tested were *Ada* (chromosome 2), *Hist2* (chromosome 3), *Mup* (chromosome 4), *Tyr* (chromosome 11), *Dhfr* (chromosome 13) and *Tcr* (chromosome 17).

cell line	Chromosome																		
	1	2:8	3	4	5	6	7	8:2	9	10	11	12	13	14	15	16	17	18	19
1	+	+	+	+•	-	+•	+	+#	+	-	+#	-	+	-	+	+	+	-	+
2	+	+•	+	+	+	+	+	+•	+	+	+	+	+	+	+	+	+	+	+
3	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+•	+	+	+
4	+	+•	-	-	-	+•	-	+•	+	+	-	-	-	+	-	+	-	-	-
5	-	+	+	+	+	+	-	+	+	+	-	+	+	-	+•	+	+	+	+
6	+	-	-	+	-	+	-	-	-	+	+	+	-	+	-	+	-	-	-
8	+	+	+	+	+#	+	+	+	-	+	-	+	+	+	+	+	+	+	+
10	+	+	+	-	+	+	+	+	+	+	+	+•	+	+	+	+	+	+	-
11a	+	+	+	+	+	+	+	+	+	+	+	+•	+	+	+	+	+	+	+
11b	+	+	+	+	+	+	+•	+	+	-	+•	+•	+•	-	+	+	+	+	+
14	-	-	+	-	+•	-	-	-	-	-	+	+•	-	+•	+	-	+	-	-

= extensive heterogeneity in the hybrid clones, as reflected by a decrease in PCR intensity relative to markers on other chromosomes in the same clone, and to other clones carrying that chromosome.

* = fragmentation of the hybrid clone reflected by non-concordant segregation of markers mapping to the same chromosome.

Table 11a:- Characterization results for the T2Wa panel

cell line	Chromosome																		
	1	2:16	3	4	5	6	7	8	9	10	11	12	13	14	15	16:2	17	18	19
1	+	+	+	+	+	+	+•	+	+	+	+	+	+	+	+	+	+	+	+
3b	+	+#	+	+	+	+	+	+	+	+	+	+•	+	+	+	+#	+	+	+
4	+	+	+	+	+	+	+	+	-	+	-	+	+	+	+#	+	-	+	+
5	+	+	+	+•	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
6	-	-	+	-	-	+	-	+	-	+	+	+	-	-	-	-	+	-	-
7	+	+	+•	+	+	+	+	+	+	+	+	+	+	+	+•	+	+	+•	-
8	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+•	+
9	+	+	+	+	+	+	+	+	+	-	+	+	+	+	+	+	+	+	+
11	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
14	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+•	+
15	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
16	+	+	+	+	+	+	+	+	+	-	+	-	+	+	-	+	+	+	-
17	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+
19	+	+	+	+	+	+	+	+	+	+	+#	+	+	+•	+	+	+	+	+
21	+•	+	+	+•	+	+	+	+	-	+	+#	+	+	+	+	+	+	+	-

= extensive heterogeneity in the hybrid clones, as reflected by a decrease in PCR intensity relative to markers on other chromosomes in the same clone, and to other clones carrying that chromosome.

* = fragmentation of the hybrid clone reflected by non-concordant segregation of markers mapping to the same chromosome.

Table 11b:- characterization results for the T28H panel

showing evidence of at least one marker were classified as '+'. The occurrence of fragmentation of particular chromosomes within the panels was not taken into account at this stage. Table 12a gives the frequency (expressed as a percentage) of retention of each mouse chromosome or chromosome fragment (with the exception of the translocation chromosomes which were analysed separately) in each panel, and in the combined panels. The primary observation is the drastic variation in the overall percentage of segregation that has occurred in each panel. The mean of the percentage retention of chromosomes or chromosome fragments in the T2Wa panel is 76%, whereas the mean of the percentage retention of chromosomes or chromosome fragments in the T28H panel is 91%. This difference is highly significant, at $P = <0.001$, tested by t test (t is taken as the difference between the two means divided by the standard error of the difference between the means; Strickberger, 1985). The standard error between the difference of the means was calculated as $^{s}x\text{-}y = 6.4$). The reason for this difference is unclear. Two alternative mechanisms of segregation have been put forward:-
a) extremely rapid segregation, occurring in the first 20 generations, probably in the first few cell divisions, which then ceases, leaving a karyotypically stable cell population and b) a balanced chromosome complement from both species, with segregation occurring in a gradual fashion over several hundred generations (White and Ephrussi, 1966a; Nabholz *et al.*, 1969). It may be that the T2Wa panel is segregating according to the first pattern, whereas the T28H panel is segregating according to the second pattern. However, the proportion of chromosomes which have segregated in either panel (24% for T2Wa, 9% for T28H) is low, even though there is some evidence that mouse:Chinese hamster hybrids do not segregate to the same degree as do human:rodent hybrids (Dr. C. Abbott, personal communication). Despite the overall high degree of retention, there is variation in the amount of segregation within each panel. Although on the whole, both panels show low levels of segregation, several clones do show evidence of reasonably high levels of segregation. For instance, T2Wa-4 and T2Wa-6 have each lost approximately half of their mouse chromosomes, whereas T2Wa-14 and T28H-6 have each lost slightly more than half of their mouse chromosomes.

Specific chromosomes appear to be retained at differing frequencies in the hybrid population. These differences can be examined for statistical significance by comparison of the percentage of retention of each chromosome (or chromosome fragment) to the expected value (i.e. - the mean percentage of chromosome retention) by χ^2 analysis (χ^2 was taken as the sum of the χ^2 for the deviation from the expected value for the presence of the chromosome, and the deviation from the expected value for the absence of the chromosome). Tables 12b and 12c

Chromosome	Percent of cell lines positive for chromosome (T2Wa)	Pecent of cell lines positive for chromosome (T28H)	Percent of cell lines positive for chromosome (overall)
1	82%	87%	85%
3	82%	100%	90%
4	73%	93%	85%
5	73%	93%	85%
6	91%	100%	96%
7	64%	93%	81%
8	-	100%	100%
9	73%	80.%	77%
10	73%	86%	81%
11	73%	93%	85%
12	82%	93%	89%
13	73%	93%	85%
14	73%	93%	85%
15	82%	87%	85%
16	91%	-	91%
17	82%	93%	89%
18	64%	93%	81%
19	64%	73%	69%

Table 12a (cont.):- The relative percentage of each chromosome or chromosome fragment retained in the hybrid panels.

give the χ^2 analysis for the T2Wa panel and the T28H panel respectively. Chromosomes showing significant deviations from the expected values are marked in bold type. In the T2Wa panel, chromosomes 6 and 16 were significantly overrepresented at $P = <0.01$ and $P = <0.001$ respectively. Chromosomes 7, 18 and 19 were significantly under-represented, at $P = <0.01$ in all cases. In the case of the T28H panel, chromosomes 3, 6 and 8 were significantly over-

Chromosome	Deviation from mean (76%)	χ^2	P
1	6	1.58	0.20
3	6	1.58	0.20
4	3	0.40	0.50
5	3	0.40	0.50
6	**15**	**11.48**	**<0.001**
7	**12**	**7.63**	**<0.01**
9	3	0.40	0.50
10	3	0.40	0.50
11	6	1.58	0.20
12	3	0.40	0.50
13	3	0.40	0.50
14	3	0.40	0.50
15	6	1.58	0.20
16	**15**	**11.48**	**<0.001**
17	6	1.58	0.20
18	**12**	**7.63**	**<0.01**
19	**12**	**7.63**	**<0.001**

Table 12b: -χ^2 analysis to identify significant deviations from the mean segregation rate for the T2Wa panel

Chromosome	Deviation from mean (91%)	χ^2	P
1	4	2.24	0.15
3	**9**	**8.36**	**<0.01**
4	2	0.25	0.60
5	2	0.25	0.60
6	**9**	**8.36**	**<0.01**
7	2	0.25	0.60
8	**9**	**8.36**	**<0.01**
9	**11**	**15.10**	**<0.001**
10	5	2.24	0.15
11	2	0.25	0.60
12	2	0.25	0.60
13	2	0.25	0.60
14	2	0.25	0.60
15	5	2.24	0.15
17	2	0.25	0.60
18	2	0.25	0.60
19	**18**	**39.38**	**<0.001**

Table 12c:- χ^2 analysis to identify significant deviations from the mean segregation rate for the T28H panel

represented at P = <0.01 in each case. The under-representation of chromosomes 9 and 19 were both highly significant, at P = <0.001 each.

Overall, taken as a whole, the two panels indicate chromosome 19 and 9 as the chromosome least represented, being present in only 63% and 77% respectively of the clones tested. The most frequently retained chromosomes were chromosome 6, present in 96% of clones tested, and chromosome 16, present in all the hybrid lines tested.

3.1.2b) The segregation of the translocation chromosomes

The results from table 11a and 11b show that there is no hybrid with conclusive evidence of the loss of one intact translocation chromosome from either panel (N.B. the nomenclature used 2:8 implies that the proximal portion of the chromosome consists of material from chromosome 2, whilst the distal portion of the chromosome originated from chromosome 8). The translocation chromosomes were present in all clones isolated from the T28H panel with the exception of T28H-6, and all the clones in the T2Wa panel, with the exception of T2Wa-2, T2Wa-4, T2Wa-6 and T2Wa-14. Unfortunately, T2Wa-2 and T2Wa-4 also show evidence that the translocation chromosomes are fragmented. To discover the extent of fragmentation, further characterization studies were carried out. The results are presented in tables 13a and 13b. The results are disappointing, in that whilst for T2Wa-4 at least the point of breakage appears to be in the chromosome 8 portion of 8:2, and could theoretically be a candidate for mapping studies, the reciprocal translocation 2:8 also appears to be fragmented at both ends. In the case of T2Wa-2, the point of breakage on 8:2 is not certain. It could be present in either the chromosome 8 or the chromosome 2 derivative. The 2:8 chromosome is also fragmented.

The high proportion of co-segregation was noted in both subcloning experiments also (see later section). The deviation from the mean retention rates for the translocation chromosomes taken individually is not statistically significant in either panel, but the absolute co-segregation is highly significant at $P = <0.001$ for both panels. In the absence of selective pressures acting for the retention of both translocation chromosomes, or against segregation, it could be expected that the frequency of co-retention (+/+) would be $(0.76)(0.76) = 0.58$ (or 58%), for the T2Wa panel. The frequencies for co-loss (-/-) of the translocation chromosomes and the frequency for segregation (-/+ and +/-) of each product were calculated in the same manner. The fragmented T2Wa translocation chromosomes have been left out of the analysis.

	Translocation chromosome 2:8
D2Mit6	+
D2Mit8	+
D2Mit11	+
D2Mit14	+
Il-1b	+
Psp	+
Ada	+
D2Mit24/D2Mit25 *	+
Aprt	-

	Translocation chromosome 8:2
Polb	+
Plat	+
*Acra4**	-

Table 13a:- Translocation chromosome characterization results for T2Wa 2

	Translocation chromosome 2:8
D2Mit6	-
D2Mit8	-
D2Mit11	-
D2Mit14	+
Il-1b	+
Psp	+
Ada	+
D2Mit24/D2Mit25 *	-
Aprt	-

	Translocation chromosome 8:2
Polb	+
Plat	-
*Acra4**	-

Table 13b:- Translocation chromosome characterization results for T2Wa 4

N.B. - Markers marked with '*' were allocated to one or other reciprocal, on the basis of apparent map position. The true location is uncertain at the present time.

Tables 14a and 14b give the χ^2 analysis for significance of the co-segregation of the translocation reciprocals. The lack of intact segregated translocation chromosomes in either panel is very highly significant, at P = <0.001. Presumably, there must be a selective advantage

	+/+	+/-	-/+	-/-
Observed value	78%	0%	0%	22%
Expected value	58%	18%	18%	6%
χ^2	6.6	17.2	17.2	44.9

Total χ^2 = 85.91, P = <0.001

Table 14a: -Deviation from expected segregation patterns for T2Wa panel

	+/+	+/-	-/+	-/-
Observed value	93%	0%	0%	7%
Expected value	83%	7%	7%	1%
χ^2	1.2	6.3	6.3	27.0

Total χ^2 = 40.87, P = <0.001

Table 14b: - Deviation from expected segregation patterns for T28H panel

for the retention of both reciprocals, or a pressure against segregation for T(2;8)2Wa and T(2;16)28H. Evidence for partial segregation for markers on the translocation products was only noted in two cases; T2Wa-2 and T2Wa-4. The presence of a selective disadvantage to the segregation is substantiated by the observation that the response of the cells to segregation of these chromosomes is fragmentation, presumably to remove from the cell the chromosome regions on which are the focus of the selection. Fragmentation has been shown to have occurred in both T2Wa-2 and T2Wa-4. The pressure could be a positive pressure for the retention of both reciprocal chromosomes, but is more likely to be a negative pressure against the presence of only one reciprocal, in view of the fragmentation. The fact that partial segregation has been noted in the T2Wa panel but not the T28H panel is probably a reflection of the differing segregation rates for the panels.

3.1.2c) The extent of chromosomal fragmentation in the non-translocation chromosomes in the new hybrid panels

It can be seen from tables 11a and 11b that there is some chromosomal fragmentation in both of these hybrid panels. It is difficult to obtain an accurate picture of the extent of chromosomal fragmentation from PCR characterization results alone, as chromosome

rearrangements or fragmentation events will not be identified by this method, unless one or more of the markers tested was deleted or lost. For a more complete estimate of chromosome fragmentation, karyotypic analysis is required. However, a proportion of the chromosomal fragmentation events are represented by the reproducible non-concordant segregation of molecular markers mapping to the same chromosome by PCR characterization. Anomalous results of this type were recharacterized to rule out the possibility of false positives (originating from PCR contamination) or false negatives (originating from failure of PCR amplification). From these results alone, it can be seen that several chromosomes are more frequently fragmented. In the T2Wa panel, chromosome 12 appears to be extensively fragmented, with 36% of the clones containing chromosome 12 in fragmented form. Both of the translocation chromosomes (2:8 and 8:2) show evidence of fragmentation in 18% of cases each. Chromosome 6 also shows fragmentation in 18% of cases. The only chromosomes appearing to be intact (on the basis of PCR characterization results) in all clones containing them are chromosomes 1, 3, 10, 17, 18 and 19. There appears to be less fragmentation in the T28H panel, as estimated from PCR characterization results. Several chromosomes show evidence of chromosomal fragmentation. Chromosome 18 shows non-concordant segregation patterns in 20% of the hybrid clones. Chromosome 4 shows fragmentation, as do chromosomes 1, 3, 7, 11, 12, 14 and 15. The remaining chromosomes, including the translocation chromosomes, appear to be intact, on the basis of this method. Chromosome 11 is thought to be relatively more unstable in some hybrid lines (Dr. C. Abbott, personal communication). Although chromosome 11 does show evidence of fragmentation in both panels, the extent is not remarkable.

The levels of chromosome fragmentation are likely to be much higher than anticipated. It was originally expected that the clones would be useful as a generalized mapping panel, but efforts to map markers blind gave no correlation between the predicted result from the hybrids, and the known chromosome location. This suggests that there is a high degree of chromosome fragmentation or rearrangement, which was not highlighted by the characterization results.

3.1.3) Subcloning experiments on T2Wa-1 and T28H-3b

Due to the failure of the initial experiment to produce adequate segregation of the translocation chromosomes in either hybrid panel, subcloning studies were undertaken, in an attempt to produce segregation of one or both panels. Two hybrid lines were selected, on the basis of their apparent heterogeneity for markers on one or other translocation chromosome,

and the intact nature of the translocation chromosomes was determined as far as possible by further PCR characterization.

3.1.3a) The determination of the extent of fragmentation of T2Wa-1 and T28H-3b

At least two hybrids (T2Wa-1 and T28H-3b) appear to have intact (as determined by PCR), if unsegregated, translocation chromosomes. The clones do show considerable heterogeneity for these chromosomes in the hybrid populations, as shown by variations in the intensity of PCR products for these chromosomes, compared with markers from other chromosomes in these clones, and also with the intensity of positive controls done in the same set of reactions. This makes these clones a good substrate for subcloning studies. Two new panels of subclones were generated from T2Wa-1 and T28H-3b, after verification of the intact nature of the translocation chromosomes. These panels were first characterized for markers originating from the translocation chromosomes. The T28H-3b panel was also typed partially for markers from selected other chromosomes, to illustrate the use of the subcloning technique.

The characterization results for T2Wa-1 and T28H-3b are given in table 15a and 15b. As far as can be determined by these methods, the translocation chromosomes appear to be intact in both cell lines.

3.1.3b) The subcloning of T2Wa-1 and T28H-3b

Both hybrid cell lines were subcloned by both method 1 and method 2 outlined in section 2.2.3b of chapter 2. For each panel, the subclones generated by each technique were combined to form one panel. Method 2 proved to be more effective at generating clones originating from single cells than method 1. The serial dilution tended to result in either many cells per well, or no cells per well. Only wells containing colonies originating from single cells were analyzed. In contrast to the original panels, no major differences in cell or colony morphology, or growth rate were noted within the subclone population for either panel. This could be taken to indicate the higher level of similarity between separate subclones within the panel.

In total, 9 subclones were isolated from T28H-3b, and 12 subclones were isolated from T2Wa-1. DNA was extracted from each of these lines as described in 2.2.2a, and the subclones were characterized in the same manner as the parental hybrid lines. The T2Wa-1 panel was

characterized only for markers mapping to the translocation chromosomes. The T28H-3b panel was partially characterized for several other chromosomes in addition to the translocation chromosomes. The characterization results for the translocation chromosomes are given in table 16a and 16b.

No evidence of complete segregation of the translocation chromosomes was noted in either panel. Extensive fragmentation of the 8:2 chromosome had occurred - 50% of the recovered subclones contained a fragmented derivative of the 8:2 chromosome. Karyotypic analysis may also highlight even higher levels of chromosomal rearrangement. This may reflect some inherent advantage in cells harboring genes located on either proximal mouse chromosome 8 or distal mouse chromosome two (some response against some unplanned positive selection pressure for this chromosome). These hybrids presumably contain the full complement of Chinese hamster chromosomes, the mouse content should not in theory have

	Translocation Chromosome 2:8
D2Mit6	+
D2Mit8	+
D2Mit11	+
D2Mit14	+
Il-1b	+
Psp	+
Ada	+
D2Mit24/D2Mit25 *	+
Aprt	+

	Translocation chromosome 8:2
Polb	+
Plat	+
*Acra4**	+

Table 15a:- Translocation chromosomes of T2Wa 1

	Translocation chromosome 2:16
D2Mit6	+
D2Mit8	+
D2Mit11	+
D2Mit14	+
Il-1b	+
Psp	+
Ada	+
D2Mit24/D2Mit25 *	+

	Translocation chromosome 16:2
Igl1	+

Table 15b:- Translocation chromosomes of T28H 3b

subclone	2:8	8:2
1(1)	+*	+
1(2)	+	+
1(3)	+	+
1(4)	+*	+
1(5)	+*	+
1(6)	+*	+
1(7)	+	+
1(8)	+*	+
1(10)	+*	+
1(11)	+	+
1(12)	+	+
1(13)	+	+

Table 16a:- Subclone translocation chromosome characterization results (T2Wa 1 panel)

Subclone	2:16	16:2
3b(1)	+	+
3b(2)	+	+
3b(3)	+	+
3b(4)	+	+
3b(5)	+	+
3b(6)	+	+
3b(8)	+	+
3b(9)	+	+
3b(10)	+	+

Table 16b:- Subclone translocation chromosome characterization results (T28H 3b) panel

much effect on the phenotype of the subclones produced. However, some mouse chromosomes obviously confer an advantage or disadvantage to the recipient cell, as reflected in the non-random pattern of segregation. Also, the effect of specific mouse chromosomes can be seen by observation of the differing growth rates, cell and colony morphology of the initial cell lines produced from the fusion, as seen in tables 10a and 10b. Alternatively, the high levels of fragmentation associated with the 8:2 chromosome both in the subclone panel, and also reflected to a lesser degree in the initial clones may be due to the presence of a fragile site or sites (however, there was no consistent site of fragmentation, as determined by PCR characterization).

In order to determine whether segregation of other chromosomes has occurred, two methods were employed. Firstly, IRS-PCR was carried out on selected subclones from the T28H-3b panel. Secondly, partial characterization for other chromosomes was carried out. Figure 21 shows the L1 band patterns produced from IRS-PCR on T28H-3b2, 3b4, 3b8 and 3b9. As can be seen, there are differences in the band patterns produced. T28H-3b9 appears to be lacking several bands present in the other subclones. Similarly, the very strong band present at approximately 240bp in T28H-3b2 is seen in the other subclones at a much lower level. This may reflect segregation in process in the other subclones. Closer examination of the figure reveals bands present in 3b4 and 3b8 which are not present in either 3b2 or 3b9. These

differences in L1 band patterns presumably reflect differences in the chromosome content of the subclones. These differences can also be illustrated by partial PCR characterization. With this in mind, the T28H-3b subclones were characterized for chromosomes 1, 4, 6, 7 and 8. The characterization results are given in table 17. Differences in mouse chromosome content were reflected, even with a very partial typing. Chromosome 8 appeared to have been very heterogeneous in the hybrid population - 3/7 subclones tested showed no sign of this chromosome. Chromosome 7 also proved to be fairly heterogeneous. It is interesting to note that the subclone that showed the most differences in L1 band pattern (T28H-3b9) also showed the most evidence of chromosomal loss on the basis of the partial characterization, having lost chromosomes 4, 7 and 8. It can be assumed that further characterization would reveal more evidence of chromosome loss. These results indicate that whilst the translocation chromosomes did not segregate, the subcloning method is valid as a technique, since it produced segregation of other chromosomes.

	Chromosome 1	Chromosome 4	Chromosome 6	Chromosome 7	Chromosome 8
T28H 3b(1)	+	+	+	+	-
T28H 3b(2)	+	+	+	-	+
T28H 3b(3)	+	+	+	+	+
T28H 3b(4)	+	+	+	+	-
T28H 3b(8)	+	+	+	+	+
T28H 3b(9)	+	-	+	-	-
T28H 3b(10)	+	+	+	+	+

Table 17:- Partial characterization of T28H 3b subclones

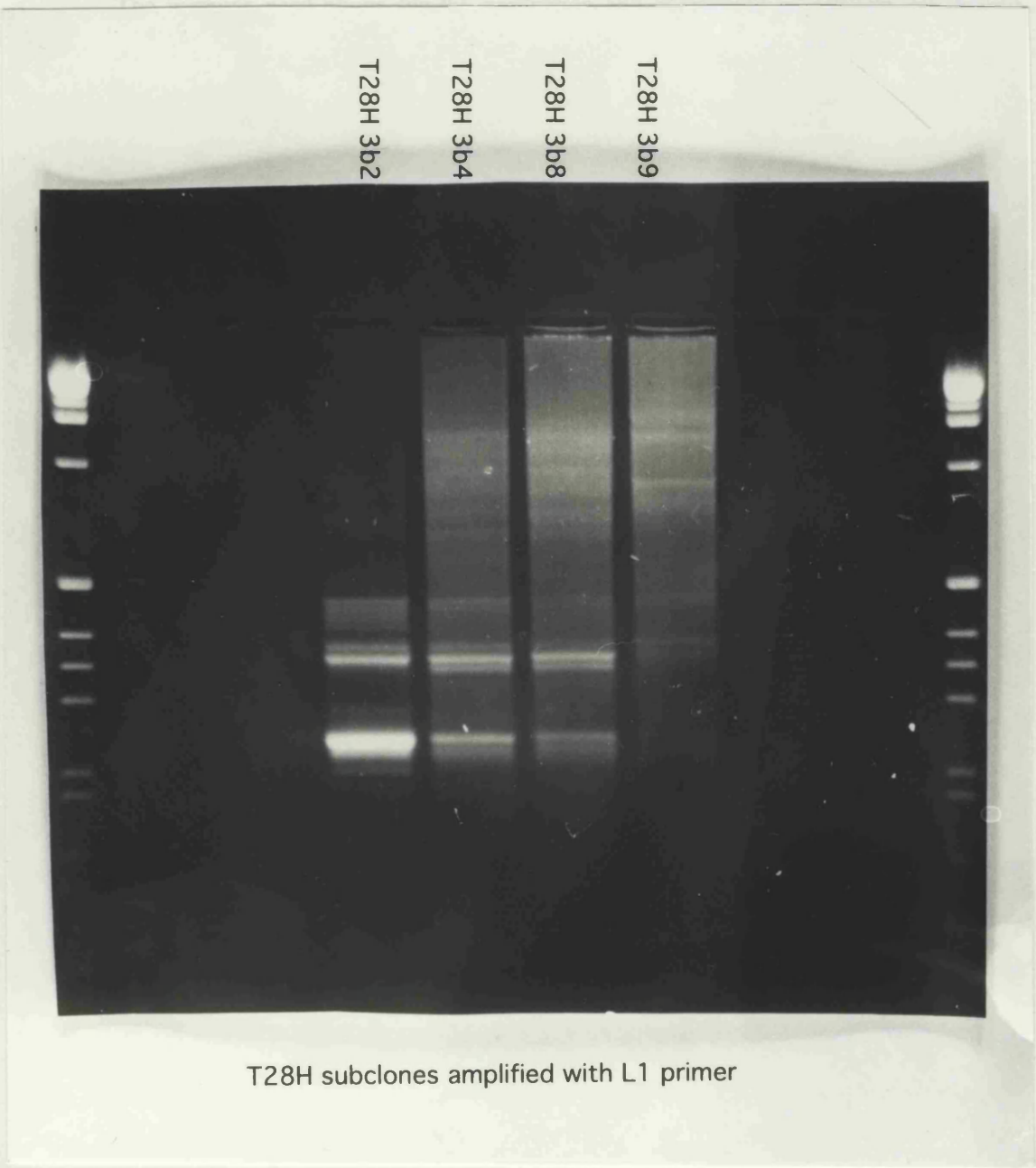


Figure 21:- The L1 band pattern produced from selected T28H-3b subclones

This figure illustrates the results acheived from amplification of DNA from four separate T28H-3b subclones with an L1 primer. This primer is species specific, and thus detects only mouse sequences.Note the differences in band pattern produced by different hybrid subclones. These band patterns were found to be reproducible, indicating genuine differences in mouse chromosome content between the cell lines. The loss of specific bands represent the loss of specific chromosomes in the hybrid cells. A smear was noted when mouse DNA was amplified, whilst the hamster control showed no product (not shown).

3.1.4) Studies on antibody selection

The methods used above for the production and subcloning of somatic cell hybrid panels rely upon the chance segregation of chromosomes of interest. In the light of the lack of segregation of either hybrid panel, by the initial fusion and subsequent segregation, or by the subcloning of hybrid clones, experiments were undertaken to produce the forced selection of the translocation chromosomes. This was to be accomplished by means of the complement mediated cell killing effect of cell surface antigen targeting antibodies. A monoclonal antibody against the Cd44 cell surface antigen was used to target cells expressing this protein on their surface. This antibody is one of very few cell surface antibodies available for mouse chromosome two, which would be appropriate for the selection of translocation chromosomes in both hybrid panels, selecting for 2:8 in the case of the T2Wa panel, and 2:16 in the case of the T28H hybrid panel.

In order for this selection system to be effective, the target must not only be expressed on the cell surface, but it must also be expressed in the appropriate cell type. For these reasons, the Cd44 antigen was judged to be an appropriate choice for a selection target. The experiment also requires that the antibody is species specific. In this set of experiments, the conditions required for complement mediated cell killing of mouse fibroblasts was investigated, and since no species specificity details were available from the manufacturers, reactivity of the antibody against V79TOR2 cells was determined.

3.1.4a) Determination of optimal conditions for complement mediated cell killing

The optimal conditions for complement mediated cell killing of mouse fibroblast (1R) cells *in vitro* were examined. Complement and antibody titrations were set up, to investigate the relative requirement for each component. The effect of the treatment was measured by the average number of cells present in three representative colonies in each well. Cell counts were taken under light microscope 48 hours after treatment. These figures were compared with those obtained from untreated wells. The results from this experiment are given in table 18 below.

I) Antibody titration (10% complement)

[antibody]	0%	1%	2%	3%	5%	10%	15%	20%
% survival	100%	5.4%	4.8%	5.4%	4.1%	5.4%	4.8%	4.1%

II) Complement titration (5% antibody)

[complement]	0%	1%	2%	3%	10%	15%	20%	25%
% survival	100%	63%	36%	19.8%	9.9%	6.3%	5.4%	4.5%

Table 18:- Antibody and complement titrations for cell killing of mouse fibroblast (1R) cells

The conclusions that can be drawn from these experiments are that whilst the antibody concentration is not critical within the 1 to 20% range, the complement concentration is of immense importance. This is not wholly unexpected, since for the complement mediated cell lysis, only one antibody molecule is required, whereas for the cell killing to be effective, all components of the complement cascade must be present at sufficient concentrations. These results are also illustrated in graphical form in figures 22 and 23. The small numbers of cells remaining after selection could be explained by a small population which did not express the Cd44 antigen. The complement titrations were also maintained at 37°C without antibody or complement, to assess whether or not these remaining cells could recommence cell growth, or if they were dead. No signs of cellular reproduction were noted, even after 10 days. Indeed, any cells present tended to die, even with daily replacement of new medium. Experiments were also carried out to determine the effect of 50% complement, together with 5% antibody. Under these conditions, no settlement of inoculated cells at all was noted. These results suggest that high concentrations (25 - 50%) of complement are adequate to bring about the non-reversible cessation of cell growth, even in the presence of relatively small amounts of antisera.

3.1.4b) Cross reactivity of Cd44 monoclonal antibody with V79TOR2 cells

Once the optimal conditions for complement mediated cell killing of mouse cells had been established, it was necessary to determine the effects of these conditions on V79TOR2 cells. The resistance of the Chinese hamster cells to these conditions is of course an important

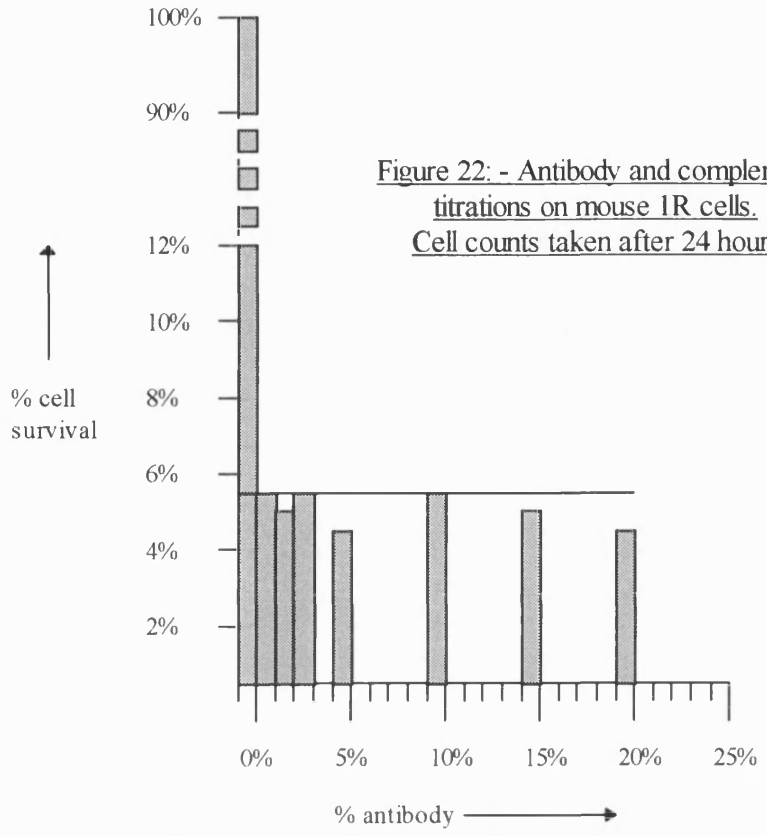
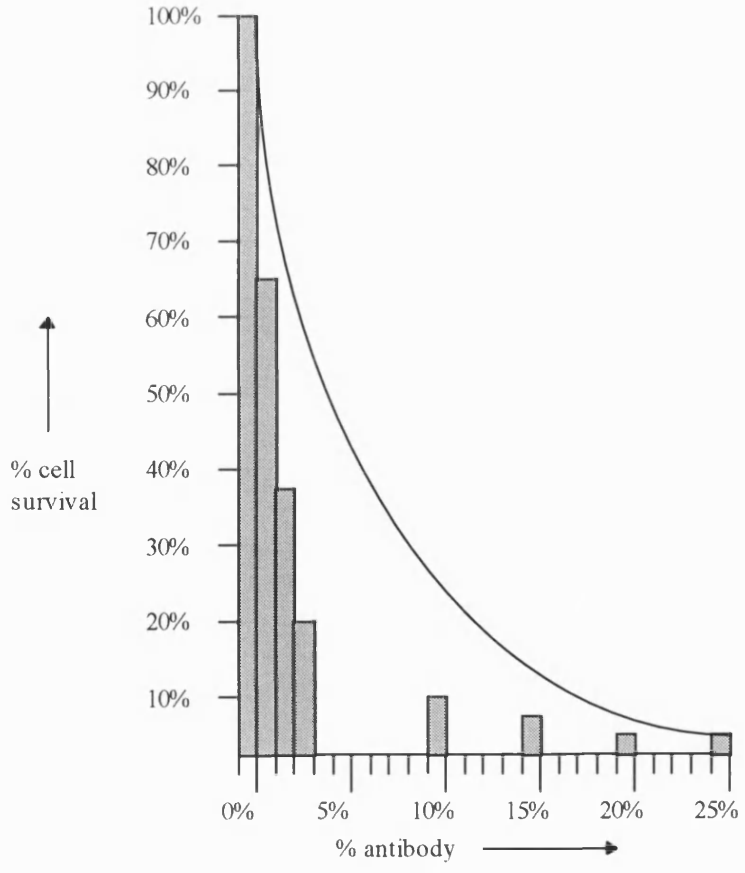


Figure 22: - Antibody and complement titrations on mouse 1R cells.
Cell counts taken after 24 hours



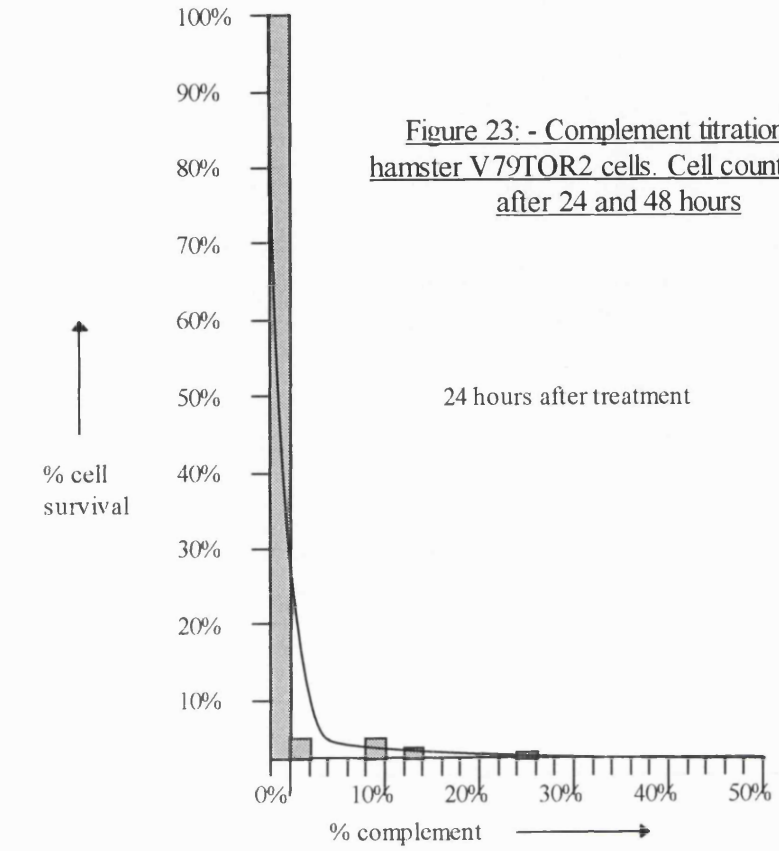
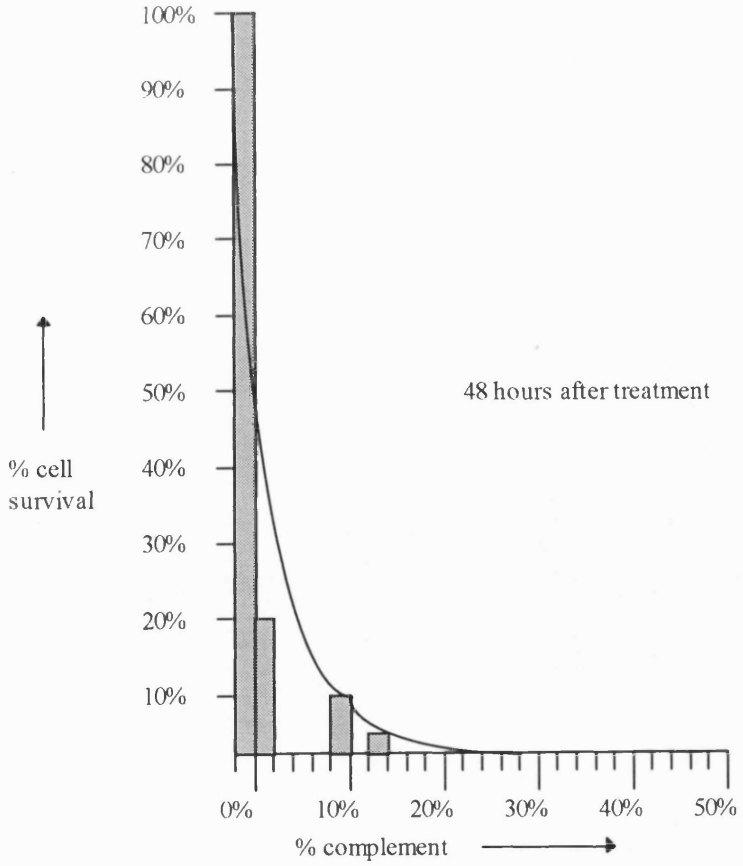


Figure 23: - Complement titrations on hamster V79TOR2 cells. Cell counts taken after 24 and 48 hours



prerequisite to the use of this system for the selection of hybrid cells. No antibody titration was carried out, as the former experiment established the complement concentration as the limiting factor. Complement titrations were carried out in the presence of 5% antibody, as described in section 2.2.4c of the materials and methods. Cell counts were taken as above after 24 and 48 hours. The results are presented in table 19 below.

24 hours

[complement]	0%	0.1%	10%	12.5%	25%	50%
% survival	100%	1%	1.2%	0.4%	0.2%	0%

48 hours

[complement]	0%	0.1%	10%	12.5%	15%	50%
N° of cells	100%	20%	10%	2%	0%	0%

Table 19:- Cross reactivity studies of Cd44 monoclonal antibody

These results are also presented graphically in figure 23. The hamster cells are seen to be equally sensitive (if not more so) than the mouse cells, and so are not suitable for their use as selective agents for the T2Wa and T28H somatic cell hybrid panels. The choice of antisera was limited by the paucity of available antibodies against cell surface antigens on mouse chromosome two. The antibody used in these experiments was a monoclonal antibody, which also imposed limitations. The use of a polyclonal antibody (i.e.- a mixture of antibodies, each recognizing a separate epitope on the same target) would have allowed the adsorption of hamster cross-reactive elements, which may have removed the hamster cross reactivity. This could be accomplished by the pre-exposure of the antisera to lawns of hamster cells, prior to their use as selective agents. By these means, antibodies with fairly substantial cross reactivity can be used as selective agents. However, this approach is not appropriate to a monoclonal antibody. The equally strong response of the hamster cells to the treatment precluded the use of the antibody/complement mixture at some threshold level, above which the mouse cells were sensitive, but the hamster cells were not.

3.1.5) The use of hybrid cells carrying translocations for the alignment of the genetic and physical map of mouse chromosome two

In the absence of any appropriate cell lines from the T2Wa and T28H panels for the alignment of the genetic and physical maps of mouse chromosome two, the T1Sn and T13H hybrid panels constructed by A. J. Pilz were examined for their suitability for the experiment (these hybrids were formed from the fusion of cells from mice carrying either the T1(2;4)Sn or T(2;4)13H translocation with the hamster cell line A23). In order for the hybrid lines to be appropriate for the alignment of the maps, they must fulfill two criteria. Firstly, the hybrid panel must contain lines which have segregated one half of the translocation away from the other half. Secondly, the translocation chromosomes must be intact. The first requirement was fulfilled by the existence of the T1Sn 3-1, T13H 3-1 and T13H 10-1 lines. T1Sn 3-1 contains a 2:4 derivative chromosome of the T1(2:4) reciprocal translocation. T13H 3-1 contains the 4:2 derivative of the T(2:4)13H translocation. T13H 10-1 contains the 2:4 derivative of the T(2:4)13H translocation.

The suitability of these hybrid lines was next assessed by investigating whether or not the translocation chromosomes were intact, by detailed characterization of these lines for markers on mouse chromosome two and mouse chromosome 4. The hybrid T1Sn 9-1 was also characterized. The characterization results are given in tables 20a - 20d (N.B.- the characterizations of the 4:2 translocation chromosome of T1Sn 3-1, the 2:4 translocation chromosome of T13H 3-1 and the 4:2 translocation chromosome of T13H 10-1 had been previously carried out by A. J. Pilz).

Examination of the tables 20a to 20d reveal that in all cases, the translocation chromosomes of T1Sn 3-1. T13H 3-1 and T13H 10-1 were extensively fragmented, as were both translocation products of T1Sn 9-1. This result, although disappointing, is not wholly unexpected. These hybrids have been shown to contain the remaining mouse chromosomes in highly fragmented form (A. J. Pilz, personal communication). In the light of the fragmented state of the T2Wa and T28H somatic cell hybrids constructed for the purposes of this thesis, it could be interpreted that hybrids containing mouse chromosomes in reciprocally translocated form might be more susceptible to chromosomal rearrangement.

The fragmented state of the translocation chromosomes of the T1Sn and T13H hybrids, and the unsegregated (and fragmented) state of the T2Wa and T28h hybrid panels

renders them inappropriate for the alignment of the genetic and physical maps of mouse chromosome two.

	Translocation chromosome 2:4
D2Mit1	+
D2Mit3/D2Mit4	nt
D2Ucl11	-
D2Ucl12	-
Abl	-
D2Mit7	+
D2Mit11	-
D2Mit12	-
Il-1b	-
D2Mit22	-
D2Mit27	+
D2Ucl14	+

	Translocation chromosome 4:2
Mos	-
Mup	+
Orm	+
Jun	nt
Lck	nt
Ada	-
D2Mit52	-

nt = not tested

Table 20a:- Characterization results for T1Sn 9-1

	Translocation chromosome 2:4
D2Mit1	-
D2Mit3/D2Mit4	nt
D2Ucl11	+
D2Ucl12	+
Abl	+
D2Mit7	+
D2Mit11	nt
D2Mit12	+
Il-1b	+
D2Mit22	-
D2Mit27	+
D2Ucl14	-

	Translocation chromosome 4:2
Mos	-
Mup	-
Orm	-
Jun	-
Lck	-
Ada	-
D2Mit52	-

nt = not tested

Table 20b:- Characterization results for T1Sn3-1

	Translocation chromosome 2:4
D2Mit1	-
D2Mit3/D2Mit4	-
Abl	-
Lck	-

	Translocation chromosome 4:2
Mos	-
Mup	-
Orm	-
D2Mit7	-
D2Mit11	-
D2Mit12	-
Il-1b	+
D2Mit22	nt
D2Mit27	nt
D2Ucl14	nt
Ada	-
D2Mit52	nt

nt = not tested

Table 20c:- Characterization results for T13H 3-1

	Translocation chromosome 2:4
D2Mit1	+
D2Mit3/D2Mit4	-
Abl	-
Lck	-

	Translocation chromosome 4:2
Mos	-
Mup	-
Orm	-
D2Mit7	--
D2Mit11	-
D2Mit12	-
Il-1b	-
D2Mit22	-
D2Mit27	-
D2Ucl14	nt
Ada	-
D2Mit52	-

nt = not tested

Table 20d:- Characterization results for T13H 10-1

Chapter 4

The generation of new molecular markers from mouse chromosome two

4.1) The production of new markers by IRS-PCR or IRS-bubble PCR

Three different methods were employed to generate new molecular markers for mouse chromosome two. These were a) IRS-PCR followed by the isolation of individual amplified fragments, b) IRS-PCR followed by the cloning of the entire PCR product, and c) IRS-bubble PCR. IRS-PCR was carried out on DNA from the ABm hybrids and EBS18Az, using either L1 or B2 primers, specific to mouse repeats, or a combination of L1 and B2 primers. IRS-bubble PCR was carried out only on EBS18Az, as this was the only hybrid considered to have a truly intact chromosome 2. The IRS-bubble PCR reactions were carried out using the using the B2 primer alone. Isolated fragments were screened for microsatellite repeats, and for putative CpG islands.

Cloned fragments were sequenced, and PCR primers designed from each insert. PCR products from *Mus musculus* and *Mus spretus* were screened, in order to identify inter-species variation. Markers showing evidence of differences between the two species were mapped by interspecific backcross analysis.

4.1.1) IRS-PCR amplification of DNA from a series of microcell and monochromosomal hybrids

The L1, B2 or L1/B2 amplification yielded a series of bands in each case. Figure 24 shows a typical range of IRS-PCR products, amplified with a B2 primer. The figure shows the characteristic smear produced from the amplification of total mouse DNA, and the smaller number of bands produced from a hybrid clone with limited mouse DNA content. No amplification was noted in the hamster negative control (not shown). The PCR products range from approximately 700bp to approximately 2kb. This is in the usual range of IRS-PCR amplification, although fragments of up to 3kb are not uncommon. Amplification with the L1 primer yields a similar spread of fragments. Amplification with L1 and B2 primers together

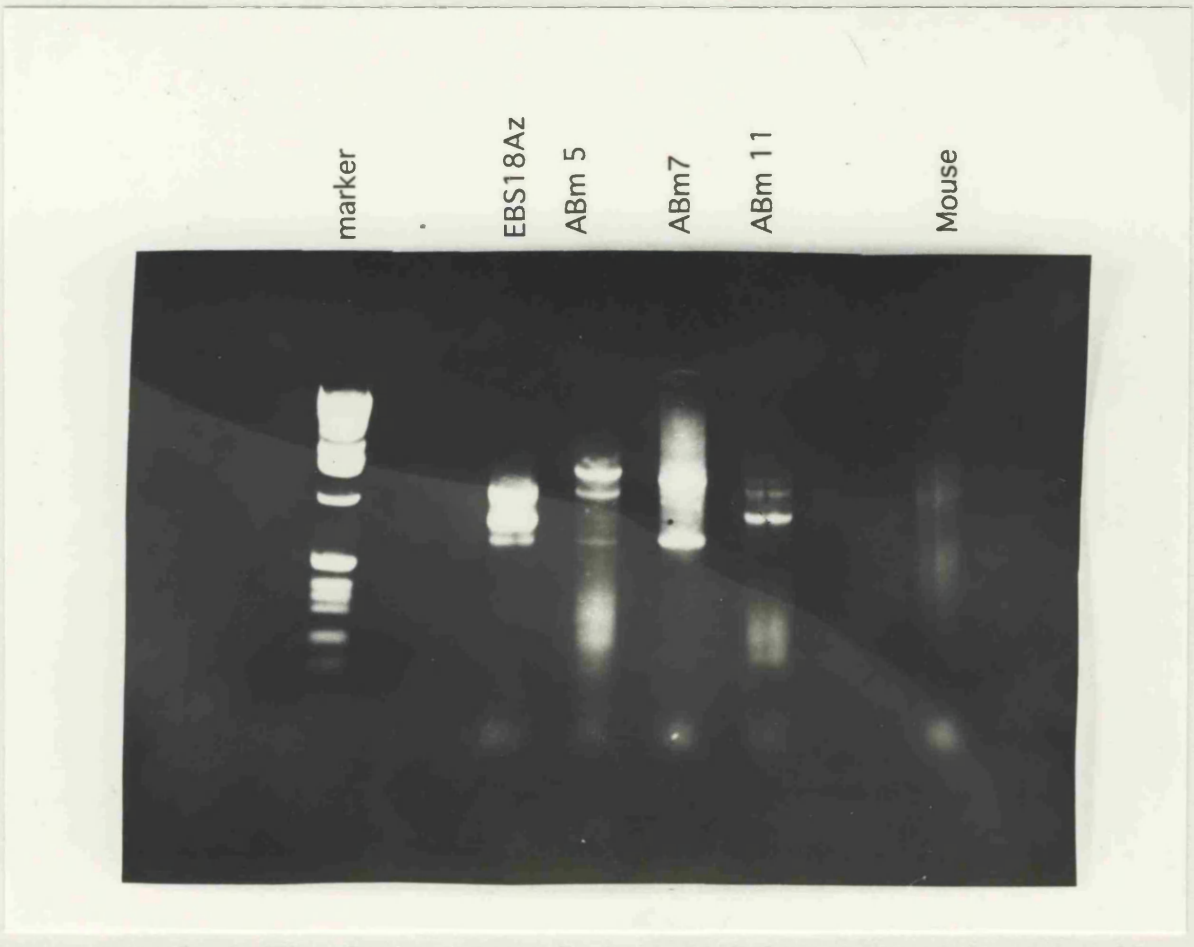


Figure 24 - Band patterns produced by IRS-PCR amplification of hybrid DNA with B2 oligonucleotide

This figure shows the band patterns produced from IRS-PCR amplification of genomic DNA from several sources with a B2 oligonucleotide. This primer is species specific under stringent conditions, no ampification was noted in the hamster controls (not shown). Note the complex band pattern produced from amplification of genomic mouse DNA, compared with the defined bands obtained from the amplification of hybrid DNAs, with reduced mouse chromosome content. The size range of the fragments was determined to lie between approximately 500bp and 3kb.

yields PCR products that are smaller, ranging from approximately 200bp up to about 1.5kb. Some indication of the chromosomal location of the IRS-PCR fragments can be obtained by comparison of the patterns produced from ABm 5 and 7, and ABm 11 and EBS18Az, as the first two lines contain only the proximal portion of mouse chromosome two. However, this is not an effective means of determining the chromosomal location of the fragments, due to the rearranged state of the ABm hybrids, and the inclusion of fragments originating on chromosomes other than chromosome two in EBS18Az. The amplification dynamics of each fragment must also be considered.

4.1.2) The isolation of individual IRS-PCR products from agarose gels.

Individual IRS-PCR products were excised from low melting point agarose gels, as described in section 2.3.2. Despite careful excision of the individual bands, subsequent amplification yielded a pool of fragments. No reduction in complexity was seen by varying the amount of template, or adjustment of the PCR conditions. The pattern of bands was much simplified from the initial reaction, and in the majority of cases consisted of the intended target, and a few subsidiary bands. Many of these bands were shared between pools of fragments derived from the same PCR reaction. Tables 21a to 21d give the hybrid of origin of each pool of isolated bands.

4.1.2a) Determination of the suitability of the IRS-PCR products for use as probes on Southern blots

L1-5, L1-6 and L1-7 were selected for use as probes on Southern blots, as these samples consisted of one very strong primary band, and only one or two much weaker secondary bands. These fragments were radioactively labelled by random priming, and used to probe blots of *Mus musculus* and *Mus spretus* DNA, amplified with an L1 primer, and also blots of mouse and hamster genomic DNA (Brooks-Wilson *et al.*, 1990; Meese *et al.*, 1992). In neither case was an unequivocal single copy signal obtained. The levels of background were so high that no single copy signal was ever seen, even in the presence of very high (up to 4mg) levels of competition with sonicated mouse DNA. This is illustrated by figure 25. Efforts to isolate regions of single copy DNA (by restriction digestion of fragments, followed by Southern blot with total mouse DNA) from within the PCR products yielded similarly poor results: these probes were also very repetitive. It is uncertain whether the background signal results simply

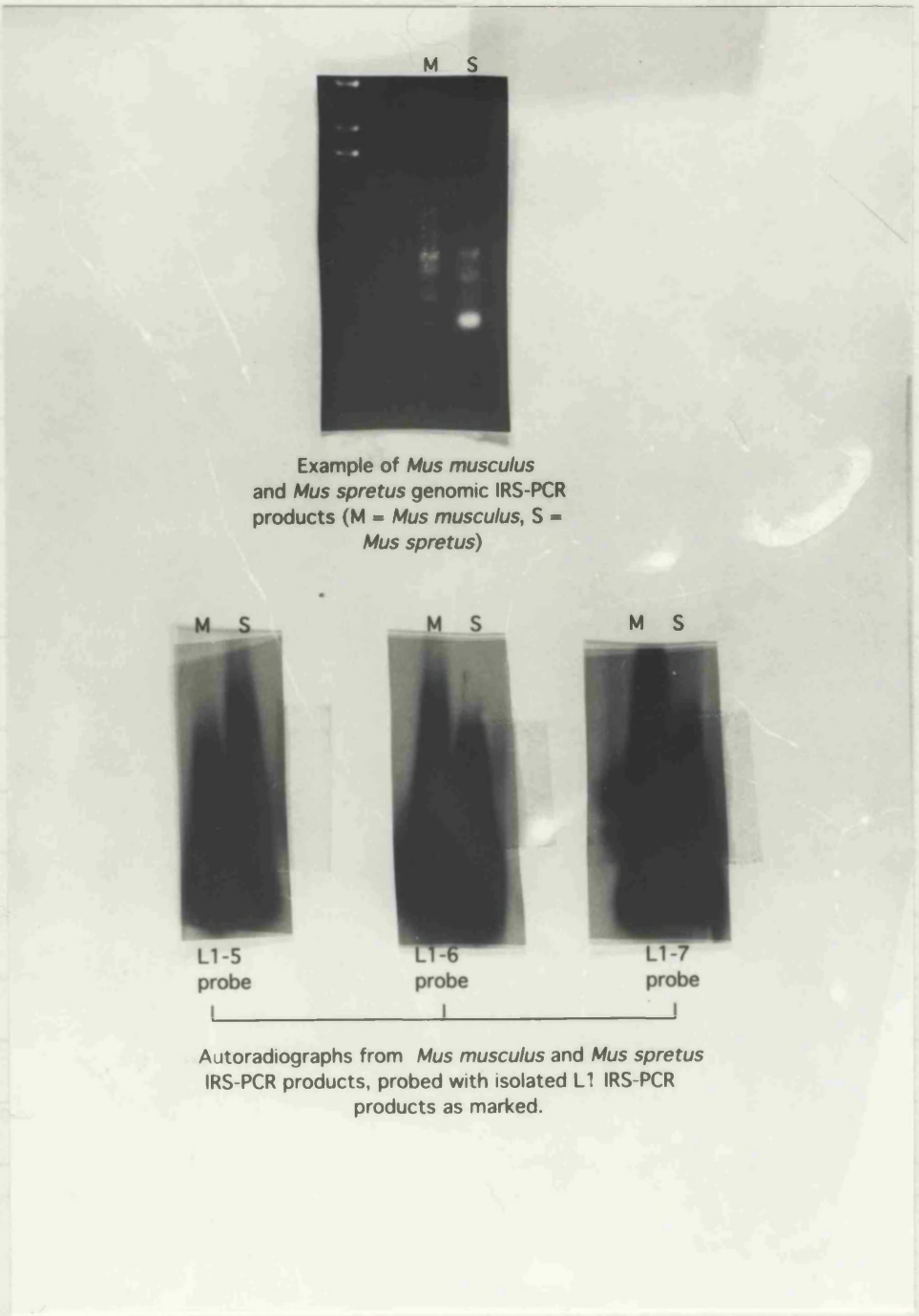


Figure 25 - Southern blotting experiments using isolated L1 IRS-PCR products

IRS-PCR amplification of *Mus musculus* and *Mus spretus* DNAs produces a characteristic band pattern. M = Mus musculus, and S = Mus spretus. Examples of the use of isolated IRS-PCR products as probes on Southern blots of the band patterns shown is illustrated above. rather than producing a single copy signal, hybridization occured in a non-specific fashion, even in the presence of heavy competition. The three L1 probes used, L1-5. L1-6 and L1-7,did not show any appreciable differences in the specificity of hybridization.

L1 isolated bands	Hybrid of origin
L1-1	EBS18Az
L1-2	EBS18Az
L1-3	EBS18Az
L1-4	EBS18Az
L1-5	EBS18Az
L1-6	ABm 5
L1-7	ABm 5
L1-8	ABm 5
L1-9	ABm 7
L1-10	ABm 7
L1-11	EBS18Az

Table 21a:- Origin of L1 IRS-PCR products

B2*EcoR1* isolated bands	Hybrid of origin
B2E-1	EBS18Az
B2E-2	ABm 11
B2E-3	ABm 5
B2E-4	ABm 5
B2E-5	ABm 5
B2E-6	ABm-7
B2E-7	ABm 7

Table 21b:- Origin of B2*EcorR1* IRS-PCR products

B2*Not1* isolated bands	Hybrid of origin
B2N-1	EBS18Az
B2N-2	EBS18Az
B2N-3	EBS18Az
B2N-4	EBS18Az
B2N-5	ABm 11
B2N-6	ABm 11
B2N-7	ABm 5
B2N-8	ABm 5
B2N-9	ABm 7

Table 21c:- Origin of *B2Not1* IRS-PCR products

L1/B2 isolated bands	Hybrid of origin
L1/B2 1	ABm 11
L1/B2 2	ABm 5
L1/B2 3	ABm 7
L1/B2 4	ABm 7
L1/B2 5	ABm 7

Table 21d:- Origin of L1/B2 IRS-PCR products

from the repetitive ends of the fragments, or from internal repetitive sequences, but since efforts to isolate internal single copy fragments were not successful, it is likely the overall nature of the fragments is repetitive, and therefore the fragments were judged unsuitable for use on Southern blots.

4.1.2b) Screening for putative CpG islands

Clustering of *NotI* and *HpaII* sites in the genome is often indicative of CpG islands. Therefore, since CpG islands are often associated with the 5' end of genes, all IRS-PCR products were digested with *HpaII* and *NotI*, in order to determine whether they contained fragments of genes. Although some isolated *HpaII* sites were discovered, there was no evidence of clustering, and no *NotI* sites were present (data not shown). From these results, it was concluded that no fragment contained the 5' end of a gene that as associated with a CpG island.

4.1.2c) Screening for the presence of microsatellites

IRS-PCR products immobilized on filters were screened for the presence of microsatellites. Figure 26 shows the results of one such experiment. Table 22 shows the presence and distribution of microsatellite repeats within the IRS-PCR product population. 13 / 32 (41%) fragments (or pools of fragments) showed evidence of containing microsatellite repeats. This figure is high at first glance, but the repeats appear to be distributed such that several fragments appear to contain more than one type of microsatellite - L1-7 contains a (CA)n repeat, a (GA)n repeat and a (GTT)n repeat. L1-1 contains both a (CA)n repeat, and a (CAG)n repeat. Only 7 out of 32 (22%) different fragments contain microsatellites. The number of fragments proving positive for each microsatellite is low, as expected, considering the relatively small size of the fragments. Given that the distribution of (CA)n repeats in mouse is approximately 1 repeat every 18kb (Stallings *et al.*, 1991), it would be unexpected to find more than three or four (CA)n repeats in the population of IRS-PCR fragments. The presence of microsatellite repeats is useful for the development of new molecular markers, as there is a high chance of any PCR product produced from primers flanking such a repeat showing length variation between mouse species.

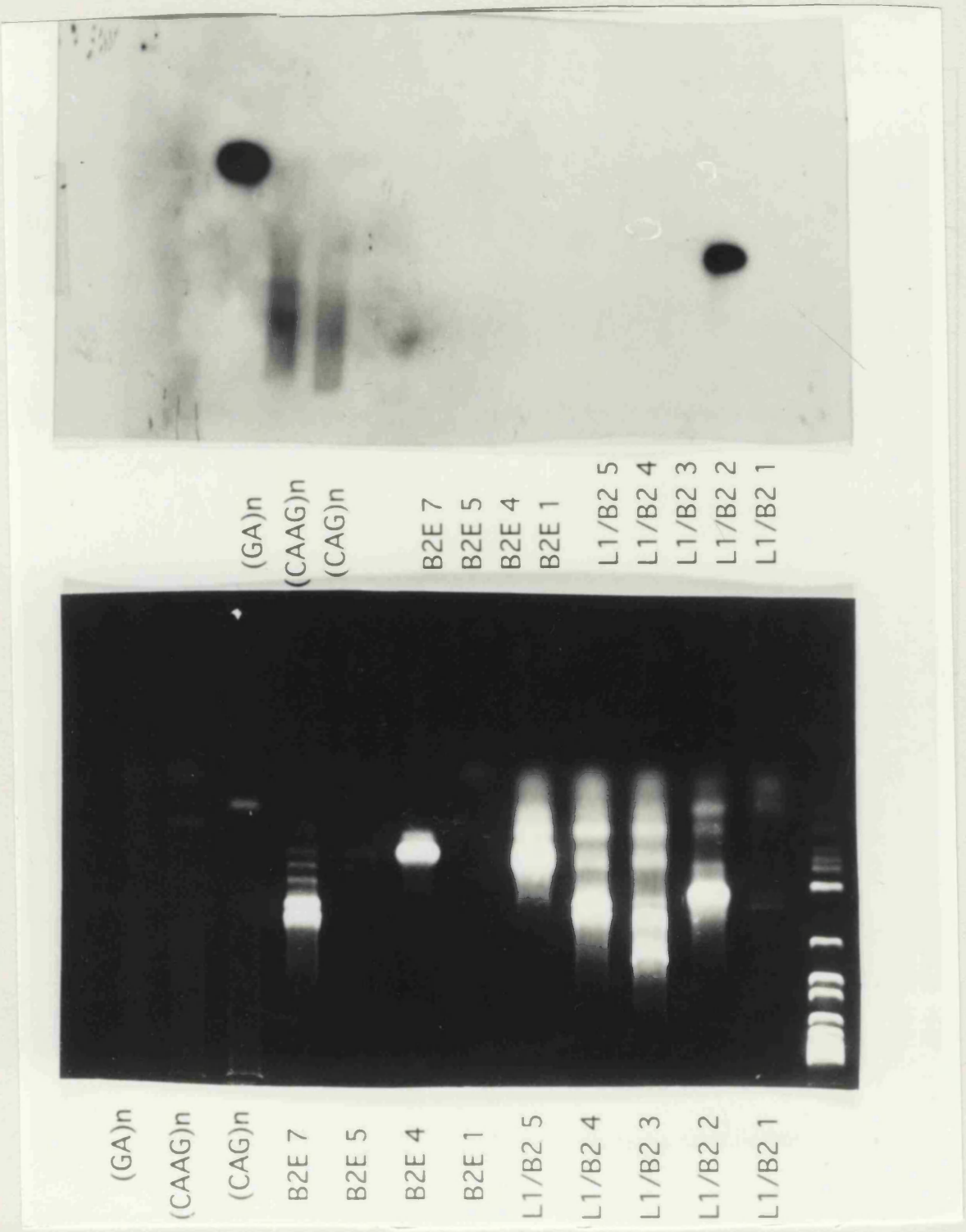


Figure 26 - Screening for the presence of microsatellites in isolated IRS-PCR products.
The top figure shows a gel photograph of selected L1/B2 and B2E IRS-PCR products electrophoresed on 1.2% low melting point agarose, together with positive hybridization controls. The bottom figure shows an autoradiograph obtained by Southern blot of the above bands, hybridized to a $(GA)_{15}$ repeat probe. the autoradiograph and the photograph shown here are not to the same scale. For the purposes of identification of positive fragments, the photograph and the autoradiograph were to the same scale. Superposition of the photograph and the autoradiograph allowed the identification of positive fragments.

Repeat unit	Positive fragments
(CA)n	3 (One band each L1-1, L1-7, B2N-2/B2N4)
(GA)n	2 (One band L1/B2-2, one band L1-7)
(CAT)n	3 (2 bands common to L1-2/L1-2/L1-4, One band in L1-1)
(GTT)n	3 (One band L1-2/L1-3, one band L1-7, one band B2E-1)
(CGA)n	no positive bands
(CCA)n	no positive bands
(AAT)n	1 (L1-9 - very faint)
(CCG)n	no positive bands
(AGA)n	1 (L1/B2-3)
(TCC)n	no positive bands

Table 22:- Distribution of microsatellites within IRS-PCR product population

4.1.3) IRS-bubble PCR amplification of fragments

The IRS-bubble PCR reaction was carried out on DNA isolated from the hybrid EBS18Az. This hybrid was the only one selected for IRS-bubble PCR because it was judged to be the only cell line showing no evidence of deletions or rearrangements. The amplification was carried out with the *Not1* bubble vectorette of Hunter *et al.* (1994), and the B2*Not1* primer of Simmler *et al.* (1991).

The amplification yielded smears rather than defined bands, indicating a larger number of different PCR products. The amount of ligated vectorette required for optimal amplification was investigated. Amplifications were set up using 0.5µl, 1µl, 2µl and 5µl aliquots of ligated EBS18Az-*Not1* vectorette bubble construct. No variation in amplification efficiency as reflected by the amount or spread of PCR products was noted for any amount of the construct. Both the conditions set out in Hunter *et al.* (1994) and a set of modified conditions were employed. The second set of conditions (in which the annealing temperature and elongation temperatures were adjusted to correspond to the original conditions required for the B2 primer, see table 5) although slightly more stringent, were found to produce no relative reduction in the intensity or spread of the PCR product produced. These conditions were thereafter used, to reduce any co-amplification of hamster sequences.

4.1.4) The cloning of IRS-PCR products from isolated bands

Individual IRS-PCR products were cloned from the simplified pools of fragments isolated from low melting point agarose gels. These pools were cloned into the *EcoR1* site of pBluescript SK+, and selected by blue-white screening. Colonies picked from the plates were miniprepped using the commercially available Promega 'magic' minipreps kit, or in later days, the Promega 'wizard' minipreps kit. Clones produced from fragments found on the primary screening to contain microsatellite repeats underwent secondary screening for the presence of these repeats prior to the miniprep procedure.

The cloning experiments were first carried out on the PCR products isolated from the L1 and B2 amplifications, as these sets of PCR products contained the most microsatellites. A total of eighteen fragments were cloned from the L1 amplification, and fifteen from the B2 amplifications. A large number of clones (20/34 - 59%) produced were derived from hamster, rather than mouse, sequences. This was determined at first only after subsequent sequencing

and amplification with new PCR primers, obviously a time consuming and expensive enterprise. This unexpected finding could be explained by a reduction in species-specificity due to the reamplification step in the cloning procedure. The hamster background probably arose at this point, since the original PCR reactions showed no sign of amplification of hamster DNA. In order to identify putative mouse clones prior to the sequencing reactions, dot-blotting techniques were employed.

4.1.4a) Secondary screening of microsatellite containing clones

DNA fragments judged to contain microsatellites were cloned first. Priority was given to pools of fragments judged to contain more than one type of repeat. The presence of other bands in every pool of fragments necessitated replica plating techniques to identify the particular clone containing the microsatellite. Colony lifts and hybridizations were carried out as described in section 2.3.7a. The use of hybridization solutions containing higher concentrations of SSPE (4% rather than 1.2%) produced lower levels of background, due to the higher hybridization temperatures allowed (see table 8). Figure 27 shows the results of a typical replica plating experiment.

Alignment of the orientation spots on the autoradiograph allowed the isolation of individual colonies containing microsatellites. These colonies were isolated and cultured as described in section 2.3.8. *HindIII* digested plasmids isolated from these colonies were fixed to filters and rehybridized to the probe. An example of this is shown in figure 28. The gel photograph given in this figure is shown larger than scale.

Clones proving positive on both the replica plating and plasmid hybridization were selected for sequencing.

4.1.4b) Methods to distinguish mouse clones from hamster clones

One problem that was frequently encountered during the course of the cloning experiments, was the high background of hamster clones, even following amplification with a 'mouse specific' primer. The use of these primers showed no apparent amplification of hamster sequences from the initial PCR reactions, as observed by the lack of any appreciable PCR product on the gels. However, the reduction in complexity of the PCR templates during the re-amplification reactions. The scale of this problem was not anticipated, although some

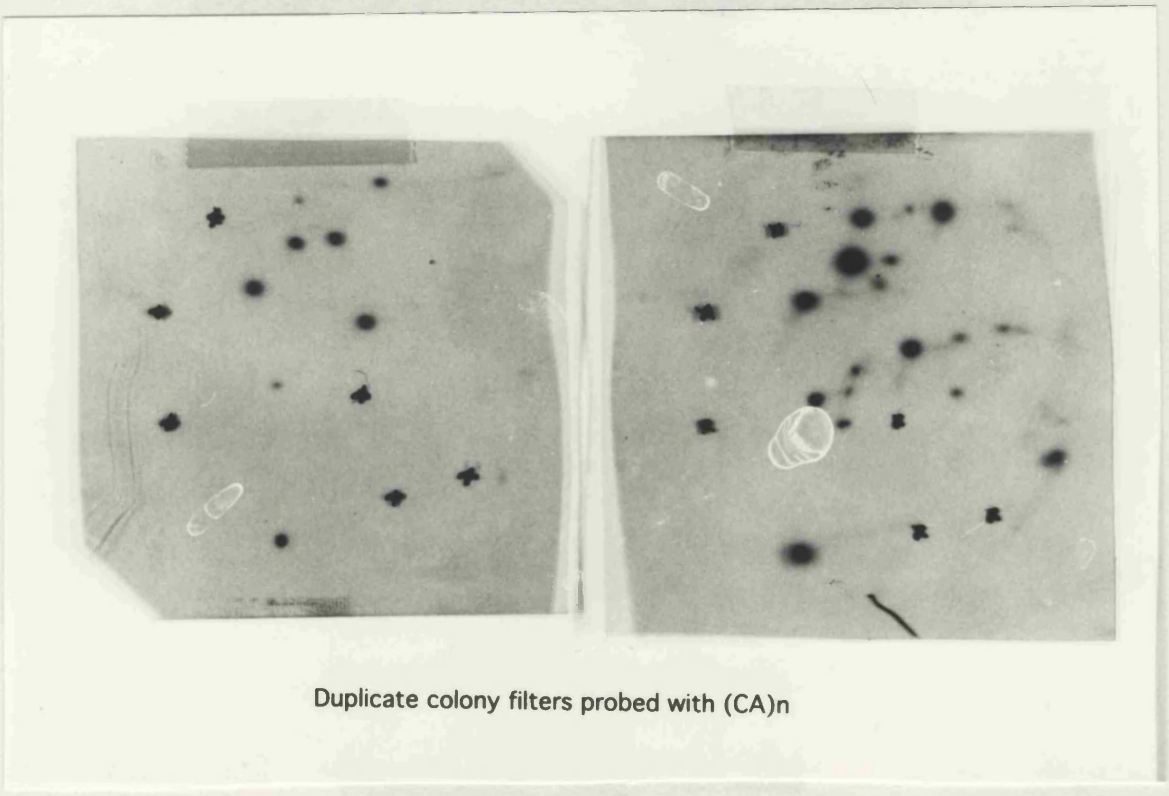


Figure 27: - Replica plating screen for (CA)n positive colonies

This figure shows the results of an experiment in which replica lifts were taken of colonies produced from a transformation were hybridized to a $(GT)_{15}$ repeat probe. The filters were aligned by use of the marks placed on the film prior to exposure (marked as small crosses). Colones producing hybridization signals on both autoradiographs were taken as representing true positives.

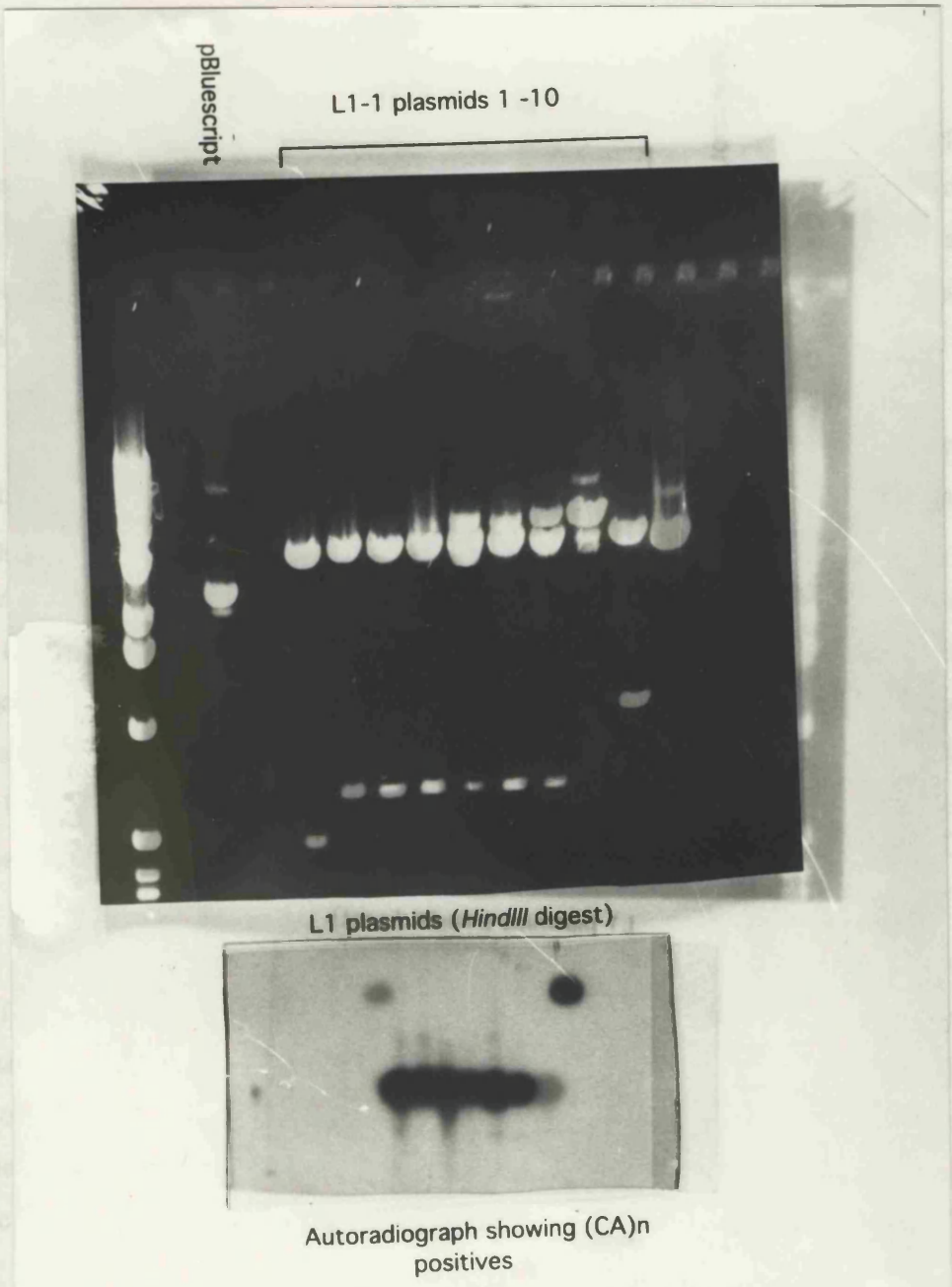


Figure 28: - Identification of plasmids containing microsatellites

Colonies giving rise to positive hybridization signals on the replica lift experiments occasionally did not contain repeats. Before cloning, the (CA)n positive status of each plasmid (extracted from such colonies) was verified by Southern blotting of plasmid DNAs, and rehybridization to the $(GT)_{15}$ repeat probe. Genuine positives were confirmed by this method. The autoradiograph and photograph here are not shown to scale, but for the identification of positives on the filters, the pictures were to scale. Superposition of the photograph and the autoradiograph identified plasmids containing microsatellite sequences.

breakthrough of hamster sequences was expected. However, five of the first seven pairs of primers produced were derived from hamster rather than mouse sequences. .

In order to minimize the wastage of time and resources, two techniques were employed to identify fragments of mouse origin. Firstly, Southern blots of plasmid DNA extracted from white colonies picked from the plates were probed with total mouse DNA, in the presence of unlabelled hamster DNA, to reduce the signal from the repetitive ends of the fragments. This technique did not prove useful, as the background hybridization was above an acceptable level.

The second technique was the use of dot blots of mouse and hamster DNA, probed with the DNA inserts of individual plasmids. This technique proved highly effective, able to distinguish between mouse and hamster clones. No clone identified as being of mouse origin proved to derive from hamster sequences upon subsequent analysis. Figure 29 illustrates one such experiment. The figure shows the results from two identical strips, dot-bloted with alternate spots of mouse and hamster DNA, probed with different inserts, one originating from a hamster clone, and one originating from a mouse clone. It can be seen that clones derived from hamster sequences detected all six dots, whilst mouse clones detect only three dots. Clones derived from vector alone do not detect any dots.

Tables 23a and 23b illustrate the species of origin of each clone from the L1 and B2 amplifications. Most of the microsatellite sequences discovered were unfortunately present in hamster clones. The L1 amplification was the most effective, with 44% of clones deriving from the mouse. The B2 amplifications produced only 33% mouse fragments. For this reason, PCR products were not cloned from the isolated fragments for the remainder of the B2 fragments, or for the L1/B2 amplifications.

4.1.5) The cloning of IRS-PCR products directly from the initial PCR reactions

The L1 and B2 primers used in these experiments have been found to be species specific in other labs (Irving and Brown, 1991; Simmler *et al.*, 1991). One possible cause of the failure of species specificity could be the inclusion of an amplification step after the isolation of individual bands. In order to test this as a possibility, experiments were undertaken to clone the remaining B2 and L1/B2 products directly from the initial IRS-PCR reactions. If the background of hamster clones derives from this second amplification step, the use of entire PCR products, with no requirement for reamplification should eliminate the problem.

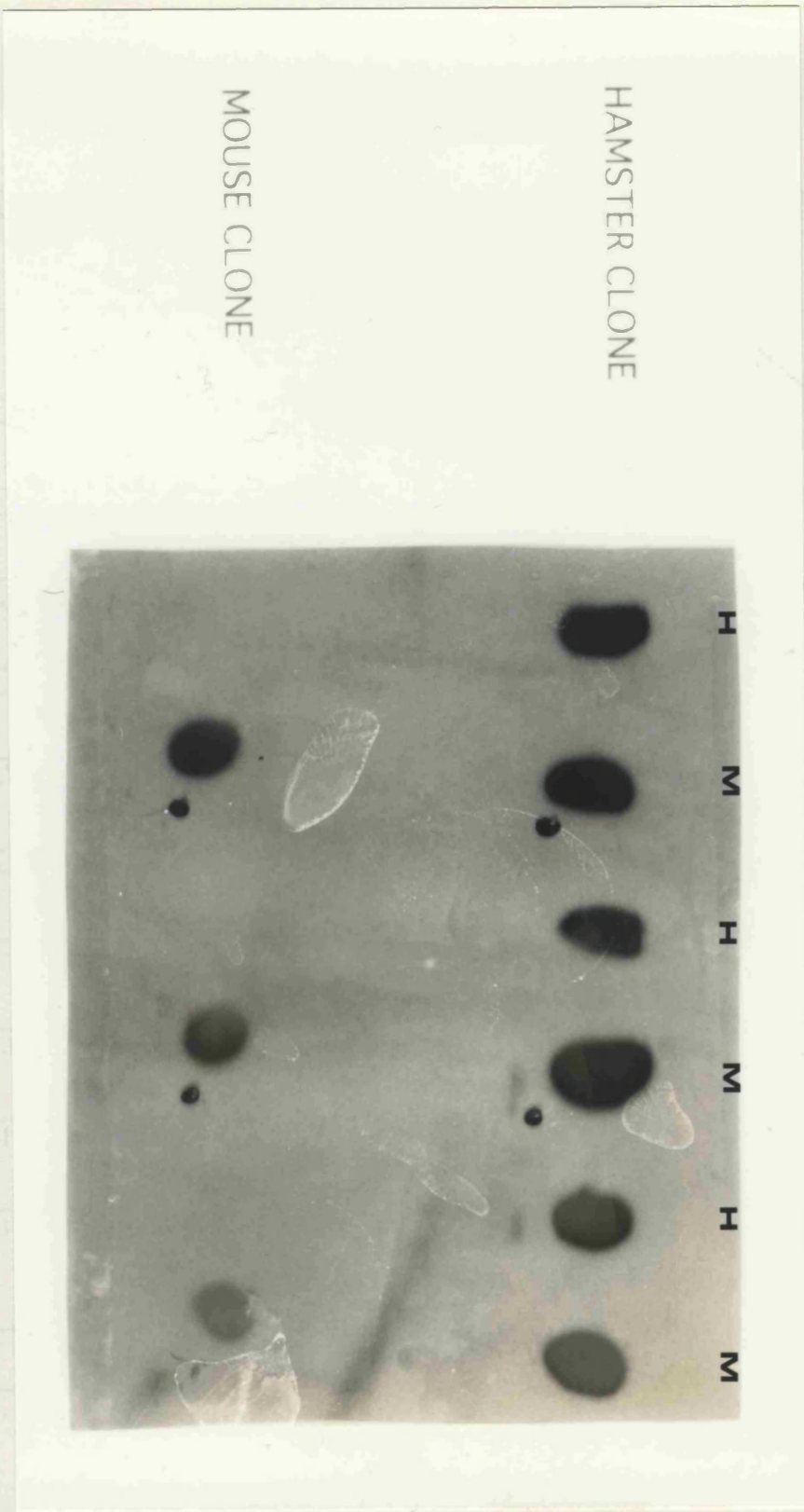


Figure 29: - Identification of mouse plasmids by dot blot analysis

The figure shows the results of probing two identical strips of alternate mouse and hamster DNA spots with two different probes originating from plasmid inserts, one derived from the hamster, and one derived from the mouse. It can be seen that the hamster clone hybridizes to both mouse and hamster spots, whilst the mouse clone hybridizes to the mouse spots alone. This method proved to be a reliable means of identifying plasmids containing mouse sequences.

L1 fragment pool	Total number of fragments cloned	Number of clones of mouse origin	Number of clones of hamster origin
L1-1	4	2	2
L1-2	-	-	-
L1-3	-	-	-
L1-4	-	-	-
L1-5	3	2	1
L1-6	1	0	1
L1-7	2	0	2
L1-8	1	1	0
L1-9	2	1	1
L1-10	1	0	1
L1-11	4	2	2
	Total 18	Total 8	Total 10

(L1-1, L1-2, L1-3 and L1-4 share all fragments in common, therefore only L1-1 was cloned.)

Table 23a:- Species of origin of each L1 clone

B2 fragment pool	Total number of fragments cloned	Total number of fragments of mouse origin	Total number of fragments of hamster origin
B2E-1	3	2	1
B2E-2	1	0	1
B2E-3	5	1	4
B2E-5	1	0	1
B2E-6	4	1	3
B2N-4	1	1	0
	Total	Total	Total
	15	5	10

Note- The remainder of the B2 fragments were not cloned because of the high degree of hamster clones

Table 23b:- The species of origin of each B2 clone

B2 and L1/B2 amplifications were carried out with the B2*EcoR1* primer of Simmler *et al.* (1991) and the L1 primer of Irving and Brown (1991), adapted by the addition of an *EcoR1* site at the 5' end. Mouse positive controls and hamster negative controls were also included in the reactions, to assess the degree of hamster amplification. Having established that no amplification was detected in the hamster samples, and any PCR products produced were likely to derive from the mouse, PCR products were cloned directly into pBluescript SK+, as described in section 2.3.6, and recombinant clones were detected by blue:white selection with IPTG and X-gal. Recombinant clones were grown and plasmids were isolated from them, as described in the materials and methods chapter. There was a very high level of clones consisting of vector alone, appearing as definite white colonies on the plates. The L1/B2 amplification was more successful, three separate mouse clones were obtained, from a total of 14 plasmids isolated. No hamster clones at all were seen.

4.1.6) The cloning of IRS-bubble PCR products

IRS-bubble PCR products were cloned directly from the PCR reactions as described in the materials and methods chapter. From these reactions, a total of 24 fragments were cloned. Table 24 gives the relative number of mouse clones produced. The remaining 12 plasmids consisted of vector alone.

	Total number of hamster clones	Total number of mouse clones	Number of different mouse clones
IRS bubble PCR products	0	12	5

Table 24:- Nature of each IRS-bubble PCR clone produced

Twelve of the clones produced did derive from mouse. Again, there was no incidence of any hamster background. There appeared to be either some predominance of certain fragments in the IRS-bubble PCR reaction mixture, or some fragments were considerably easier to amplify, as of these fifteen clones, only five separate sequences were obtained.

4.1.7) Generation of sequence information for each mouse clone

Each clone proved to be of mouse origin was sequenced as described in section 2.3.9 of the materials and methods. Figure 30 shows sequence information obtained for a clone containing a (CA)n repeat.

The sequence for each clone was not determined for the entire insert. For clones which showed evidence of microsatellites, sequence was determined from either side of the repeat unit. The majority of the remaining clones were only sequenced for approximately 150bp after the repetitive section of the insert. This region could be easily recognized from the gels, by the presence of a long tract of (A)n or (T)n. The ends of these repeats are degenerate, and therefore all sequences obtained were checked against the GenBank database for homology with mouse repetitive elements. Where chimerism was suspected (i.e.- *EcoRI* restriction digestion indicated that the plasmid contained two ligated mouse inserts), or where subsequent amplification with the primers designed proved difficult, one primer was used to sequence further into the insert.

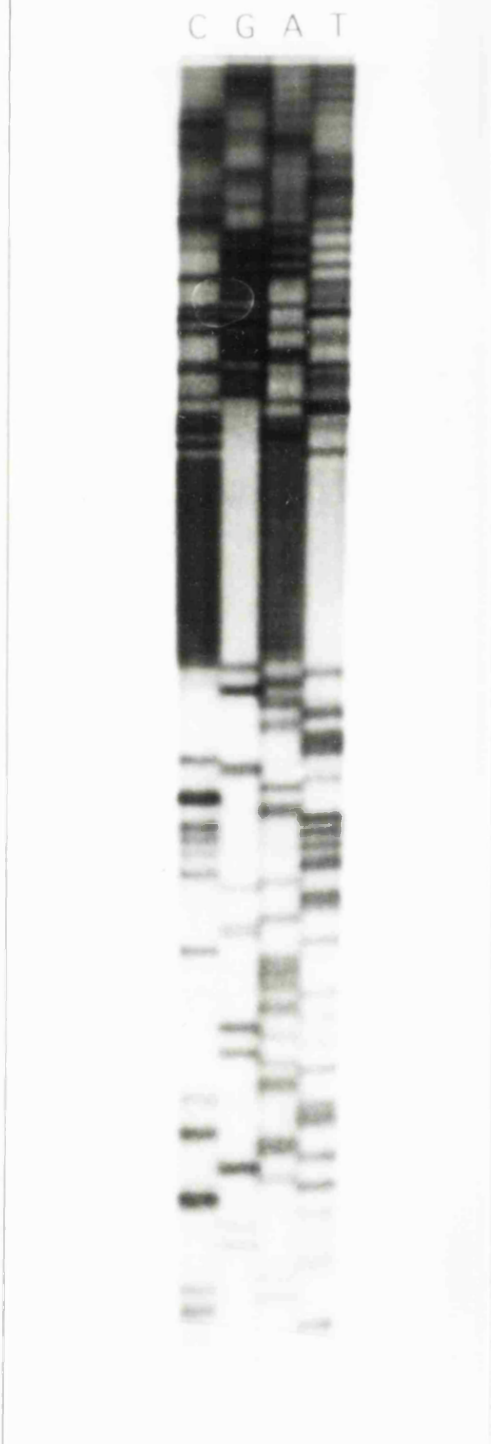


Figure 30: - Sequence derived from a clone containing a (CA)n repeat

This figure shows sequence information obtained from an insert within wich a (CA)n repeat is present. The (CA)n repeat is seen to be distinctive, and easily recognised. In this case, sequence 5' to the repeat would be used to desugn a 5' primer. 3' sequence was obtained by longer gel runs, or by the use of the 5' primer as a new sequencing primer.

In some cases, no sequence was obtained for a clone, either because it was unstable, and lost during cloning (as in the case of B2E 1-8), or because it was not possible to sequence it, since it was rearranged, and had deleted the priming sites. One of the IRS-bubble PCR products (bubble 17) fell into this category. Table 25 gives the fate of each mouse clone produced.

Clone	Fate	Designated name
L1-1(2)	cloned, sequenced and mapped	*D2Ucl13*
L1-1(11)	cloned, sequenced and mapped	*D2Ucl12*
L1-5(1)	cloned, sequenced and mapped	*D2Ucl16*
L1-5(6)	cloned and sequenced	-
L1-8(2)	Fragment identical to L1 5(1)	-
L1-9(5)	cloned, sequenced and mapped	*D2Ucl19*
L1-11(2)	cloned, sequenced and mapped	*D2Ucl14*
L1-11(5)	cloned, sequenced and mapped	*D2Ucl15*
B2E 1(8)	lost during cloning	-
B2E-1(2)	cloned, sequenced and mapped	*D2Ucl11*
B2E-3(8)	Fragment identical to L1 5(1)	-
B2E-6(4)	Fragment identical to L1-5(1)	-
B2N-4(9)	cloned and sequenced	-
Bubble 1	cloned, sequenced and mapped	*D2Ucl17*
Bubble 2	cloned and sequenced	-
Bubble 5	cloned, sequenced and mapped	*D2Ucl18*
Bubble 14	cloned and sequenced - no variation between *Mus musculus* and *Mus spretus*	*bubble 14*
Bubble 17	unstable and rearranged	-
L1/B2-7	cloned and sequenced	-
L1/B2-11	cloned, sequenced and mapped	*D2Ucl20*
L1/B2 12	cloned, sequenced and mapped	*D2Ucl31*

Table 25:- The fate of each mouse clone

Chapter 5

5.1) The determination of map position for newly generated molecular markers from mouse chromosome two

PCR primers designed from sequence obtained in the previous section were screened for variation between *Mus musculus* and *Mus spretus*. Once suitable variation had been discovered, the new markers were mapped on the Jackson laboratory interspecific backcrosses. Markers mapping in region of ragged, wasted or ulnaless were further mapped on specialized backcrosses, in which the specific mutant was segregating.

5.1.1) Establishment of primer conditions

New pairs of primers were first amplified under varying magnesium concentrations of 1.0, 1.5, 2.0, 2.5, 3.0, 3.5 and 4.0mM magnesium per reaction, in the presence of either mouse or hamster DNA. Occasionally, plasmid samples were also amplified in order to determine the exact size of the product, or to check that the primers could amplify DNA under optimal conditions if amplification did not occur on genomic DNA. All precautions were taken to avoid contamination of reagents or equipment with plasmid. If suitable conditions were difficult to establish, glycerol (10 or 20%) or dimethyl sulphoxide (10%) were included in the reactions. The use of glycerol served to reduce the complexity of the PCR product, so that a single band was produced in cases where the absence of glycerol led to the production of multiple products. The inclusion of DMSO appears to increase the amplification of target sequences. The mechanisms by which glycerol or DMSO act are unclear. Several pairs of primers were problematic, in that multiple products were produced under all conditions tried. In these cases, further primers were designed in an attempt to produce a single copy product. Fragments L1-5(6), B2N 4(9), L1-9(5), L1/B2-7 and bubble 2 were especially troublesome in this respect. One explanation for the difficulty in obtaining single PCR products could be the inherent repetitive nature of the cloned fragments. Although all primers were screened for homology to repetitive sequences prior to synthesis, the degenerate nature of the extreme 3' end of both the L1 and B2 repeat units could lead to slight homology not being identified in the database.

Table 26 gives the sequences and amplification conditions for all new markers which produced reliable and reproducible amplification. In the majority of cases, amplification produced a single band in *Mus spretus*. Other cases (such as *D2Ucl17* and *D2Ucl18*) produced

Marker	Origin	Sequence	[$MgCl_2$]	Conditions	Other requirements	Allele sizes
D2Ucl11	B2	5' GGG CAC CAG ACT TCA TTA TG 3' 5' TGC AAG TGC CCT ATT TCC CA 3'	2.0mM	94°C - 30s, 55°C - 30s, 72° - 30s, 35 cycles.	10% glycerol	*Mus musculus* allele = 170bp Mus spretus = 1 or 2 bases larger.
D2Ucl12	L1	5' ACT CAT TGA GAC CAC AGT CA 3' 5' TAT TGA CCA CCT CTG TAT CC 3'	1.5mM	94°C - 30s, 55°C - 45s, 72°C - 60s, 35 cycles.	10% glycerol	*Mus musculus* allele = 231bp *Mus spretus* allele < 231bp.
D2Ucl13	L1	5' GAA GCT CAC TTA AAC TGT GC 3' 5' CGA TGG TAT TAG ACT CTG TC 3'	2.0mM	94°C - 30s, 58°C - 30s, 72°C - 30s, 35 cycles	none	*Mus musculus* allele = 396bp *Mus spretus* fails to amplify.
D2Ucl14	L1	5' AGT GAC CTC CTC ATC AAT GC 3' 5' CTA CCC AGT AGG TAC CAT AG 3'	2.0mM	94°C - 30s, 55°C - 30s, 72°C - 30s, 30 cycles.	none	*Mus musculus* allele = 100bp *Mus spretus* fails to amplify.
D2Ucl15	L1	5' AAC AGT GTG CTT TTC CAT AC 3' 5' ACT TTT GGG ATA GCA TTG GA 3'	1.5mM	94°C - 30s, 55°C - 30s, 72° - 30s, 35 cycles.	none	*Mus musculus* and *Mus spretus* give 210bp band (other faint bands also present).
D2Ucl16	L1	5' AAT CTG CTC TGA CAG TTG TG 3' 5' TGG CAA ACA GTC AGT TTC AG 3'	1.5mM	94°C - 30s, 57°C - 30s, 72°C - 30s, 30 cycles.	none	*Mus musculus* allele = 144bp, *Mus spretus* allele is about 5 bases larger.
D2Ucl17	IRS-bubble	5' TTG GGA CCA TAG TCA CCC AG 3' 5' CAG CAG TCA GGG TTG GAT GA 3'	1.5mM	94°C - 15s, 61°C - 30s, 72°C - 30s, 30 cycles.	none	*Mus musculus* DNA yields three bands of 298bp, 250bp and 220bp. *Mus spretus* DNA does not yield the 220bp band.

Table 26 :- New molecular markers generated for mouse chromosome two by IRS-PCR and IRS-bubble PCR

D2Ucl18	IRS-bubble	5' CGA TCA GTC TCA CAG TAG CT 3' 5' CAC AAC TGT CAG AGC AGG AT 3'	3.0mM	94°C - 15s, 58°C - 30s, 72°C - 30s, 30 cycles.	10% glycerol	*Mus musculus* DNA yields three bands of 1.3kb, 500bp and 450bp. *Mus spretus* DNA yields three bands of 1.0kb, 500bp and 390bp.
D2Ucl19	L1	5' AGT AGG AGC GGA CAC CAT CA 3' 5' ACC TTA AGG GCA GTG GGG AA 3'	1.0mM	94°C - 15s, 59°C - 30s, 72°C - 30s, 30 cycles.	none	*Mus musculus* and *Mus spretus* DNA both yield 1.1kb product.
D2Ucl20	L1/B2	5' AAG TAC TTC TTG AGA GCC AG 3' 5' CCC AAA CAG CTT CCT AGA AA 3'	1.5mM	94°C - 15s, 55°C - 30s, 72°C - 30s, 30 cycles.	none	*Mus musculus* and *Mus spretus* both yield 240bp product.
D2Ucl31	L1/B2	5' CAC AGA TCT TTA CCA GGC AT 3' 5' GGA GAG TCT AAG ACA GCT AC 3'	1.25mM	94°C - 15s, 56°C - 30s, 72°C - 30s, 30 cycles.	none	165bp product amplified from *Mus musculus* DNA, no amplification in *Mus spretus*
Bubble 14	IRS-bubble	5' CCT GCT CTG ACA GTT GTG TG 3' 5' AGC TAC TGT GAG GCT GAT GC 3'	1.0mM	94°C - 30s, 50°C - 30s, 72°C - 30s, 30 cycles.	10% glycerol	Primers ampify a 95bp product in both *Mus musculus* and *Mus spretus*. No RFLV or SSCP.

Table 26 (cont.) :- New molecular markers generated for mouse chromosome two by IRS-PCR and IRS-bubble PCR

small numbers of subsidiary bands. Amplification with *D2Ucl17* produced three closely spaced bands in *Mus musculus*. Amplification with *D2Ucl18* primers produced three widely spaced bands also. In these cases, amplification of plasmid DNA indicated the correct PCR product. The number of PCR products produced and their sizes are also indicated in table 26.

Once suitable conditions had been established, PCR was carried out on DNA from C57BL/6J and *Mus spretus* parental types from the Jackson laboratory interspecific backcross panels.

5.1.2) The detection of variation between mouse species

Five methods of detecting variation between *Mus* species were employed, as outlined in the materials and methods. Figure 31 gives examples of each technique. The choice of method was largely governed by the ease of detection, and the rapidity with which data could be

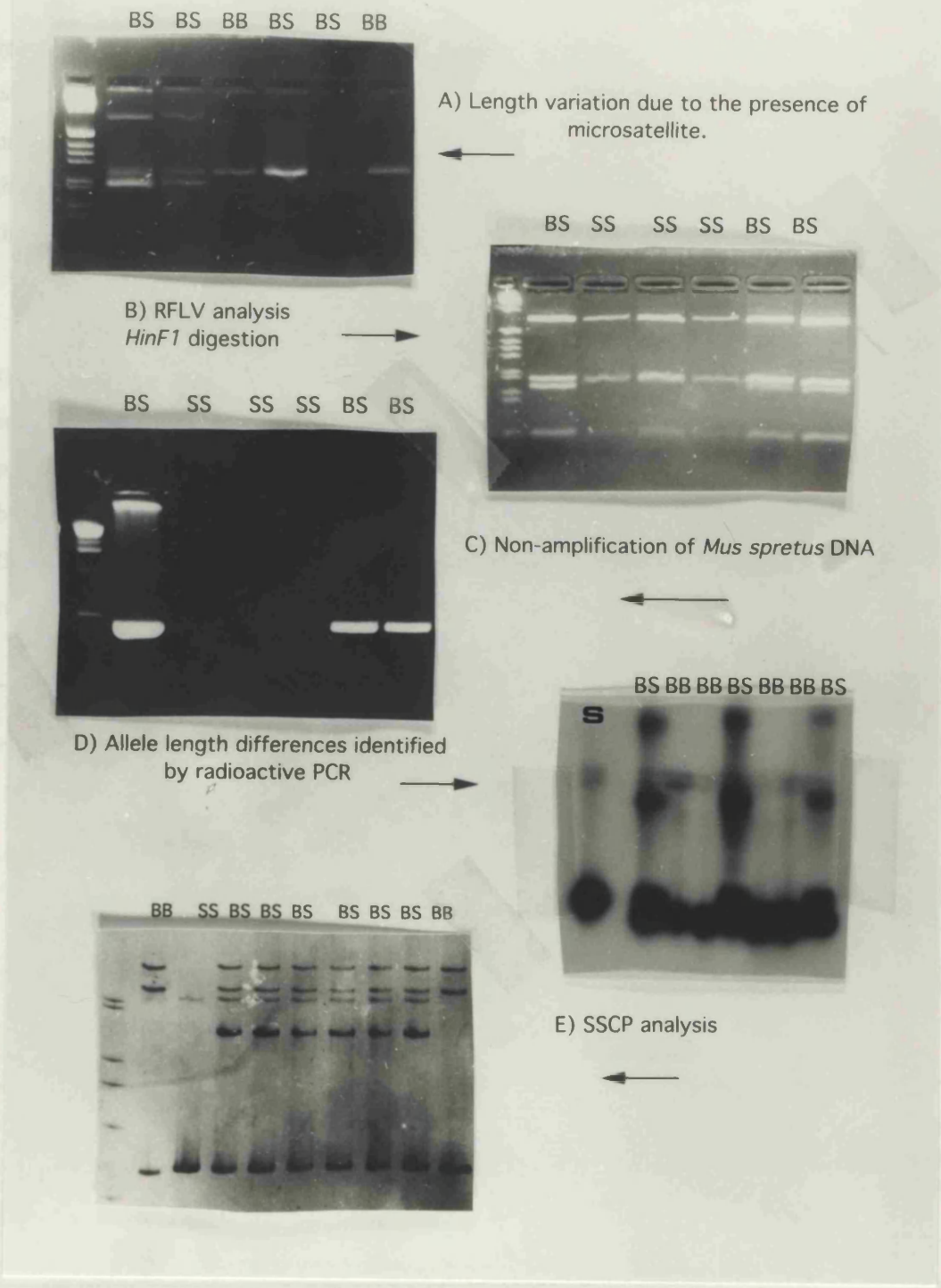


Figure 31: - Five methods by which variation was detected

Variation between *Mus musculus* and *Mus spretus* (or *Mus musculus* and *Mus musculus castaneus*) was detected by five different methods in these experiments. The figure above shows an example of each method. Homozygotes in the BSB panel are denoted BB, heterozygotes as BS (A and D). In the BSS panel, homozygotes are denoted SS, and heterozygotes as BS (B, C and E). In all cases except C, heterozygotes are identified by the presence of two bands. Because the homozygous SS alleles are not amplified by method C on the BSS cross, heterozygous animals are easily characterized by one band only. In the case of method E, heterozygotes are typed by the presence of the prominent heteroduplex band, running at approximately 496bp. In the case of E, BB and SS refer to the parental types.

generated. In cases where no amplification of *Mus spretus* DNA was obtained, these markers could be mapped directly on the BSS (panel two) Jackson laboratory interspecific backcross panel. In these cases, no DNA controls were always included into the set of reactions, to ensure that positive signals were not due to contamination. *D2Ucl13* and *D2Ucl 14* were mapped in this way.

In cases where the marker produced PCR products from both *Mus musculus* and *Mus spretus*, other means of detecting variation, and thus mapping the marker had to be used. In some cases, the presence of microsatellites, either intentionally cloned or not, caused the PCR products from each species to differ in size. If this difference in size was sufficiently great, the bands could be resolved on high percentage agarose gels (3 - 6%). *D2Ucl12* was mapped in this way. The correlation between microsatellite length and variability was also examined. The results are given in table 27. 55% of the markers contained microsatellites, these were mainly

Locus	microsatellite	Length variation
D2Ucl11	$(T)_{15}$ interrupted	PCR product from *Mus musculus* and *Mus spretus* varies by 1-2bp.
D2Ucl12	$(CA)_{23}$	size variants noted between several species, and inbred strains
D2Ucl14	$(CAT)_8$ interrupted	*Mus spretus* allele does not amplify
D2Ucl16	$(T)_{16}$ interrupted	PCR product varies by approximately 3-4bp between *Mus musculus* and *Mus spretus*
D2Ucl17	$(AAAC)_5$	*Mus spretus* allele does not amplify
D2Ucl19	Two closely placed repeats: $(T)_{14}$ and $(G)_7$	no variation

Table 27: - Correlation of microsatellite repeat length and variability

mononucleotide tracts. From the comparison of repeat length and variability, the most variable repeat unit was also the unit with the longest tract of uninterrupted repeats (*D2Ucl12* - $(CA)_{23}$). Other markers were mapped by virtue of small length differences arising from small insertions

and deletions, or from the presence of mononucleotide repeats, as are common in the vicinity of the 3' ends of SINES and LINES. Three fragments contained such variable repeats: (*D2Ucl11* - $(T)_{15}$, *D2Ucl16* - $(T)_{16}$ and *D2Ucl19*- $(T)_{14}$ and $(G)_7$). All three of these fragments contained tracts of T nucleotides, which were distinct from the A+T rich tail of the 3' end of the L1 repeat. The first two repeats contained a one nucleotide interruption, but both were variable. The size difference between the alleles was much smaller than seen for dinucleotide or trinucleotide repeats, presumably due to the smaller size of the repeat unit. Both these sources of variation were only detectable after radioactive labelling, and electrophoresis on polyacrylamide sequencing gels. This technique proved time consuming, and was only used as a last resort. The third repeat, *D2Ucl19*- $(T)_{14}$ and $(G)_7$, was not variable. Neither of the separate repeat units were interrupted. There appears to be some correlation between the 'purity' of the repeat, and the variability, and also between the length of uninterrupted repeats. It was difficult to draw many conclusions as to the proportion of microsatellites that were variable, since the remaining two markers (*D2Ucl14* and *D2Ucl17*) did not produce ampification in *Mus spretus*. However, of the four microsatellites which did amplify DNA of both species, 75% were variable.

In some cases, length differences were visible in PCR products in the absence of microsatellites. This could have arisen from deletions or insertions into the DNA from one or other species. This was noted in the case of *D2Ucl18*. In cases such as *D2Ucl18*, where the primers produced multiple products, the extra bands were not polymorphic in the majority of cases, although two of the three bands produced by *D2Ucl18* did show variation between *Mus musculus* and *Mus spretus* - the second band showing variation did not appear to map to chromosome two. The amplification of well defined extra bands could indicate the presence of another locus with high homology to the cloned locus in the genome. In no case did the presence of these bands interfere with the typing of the backcross progeny.

Where no length difference visible on either agarose or polyacrylamide was present, PCR products from *Mus musculus* and *Mus spretus* were screened with a panel of restriction enzymes. The probability of finding polymorphic restriction sites is related to the size of the PCR product. For this reason, primers were designed to amplify as large a fragment as possible. Nevertheless, restriction sites are found at a lower frequency in smaller PCR products so these were also screened for the presence of RFLVs if necessary. Several enzymes proved useful for the detection of RFLVs, namely *Hinf1*, *Dde1*, *Rsa1*, *HaeIII* and *Taq1*. For instance

D2Ucl15 was mapped by virtue of an *RsaI* variation, and *D2Ucl19* was mapped by virtue of a *HinfI* variation.

In cases where no other source of variation could be exploited, or in the case of very small PCR products, variation was detected by single stranded conformation variation analysis. The optimal size range for SSCV analysis is between 150 and 300bp (Hayashi, 1992; Sheffield *et al.*, 1993). PCR products in this size range were mapped directly from the SSCV gels. Larger products may be restriction digested prior to the gel electrophoresis, to liberate a smaller fragment. A range of gel conditions was used to identify variation. *D2Ucl20* was mapped by virtue of a single stranded conformation variant detected on a 10% glycerol, 0.5X TBE, 8% acrylamide gel. The development of a prominent heteroduplex band was of great assistance in typing this marker.

The order of choice of the methods was a) direct typing by virtue of non-amplification of *Mus spretus* DNA. This method is more error prone than the other methods. No DNA controls were used as a means of identifying contamination, but occasional failure to amplify in one tube led to some false negative results. To circumvent this, negative samples were re-amplified. b) the detection of length differences on agarose gels, c) SSCV analysis, d) RFLV

Molecular marker	Mapped on panel....	Method of mapping
D2Ucl11	BSB	radioactive PCR
D2Ucl12	BSB	length differences visible on 4% agarose
D2Ucl13	BSS	non-amplification of *Mus spretus* allele
D2Ucl14	BSS	non-amplification of *Mus spretus* allele
D2Ucl15	BSB	mapped by *RsaI* variation
D2Ucl16	BSB	radioactive PCR
D2Ucl17	BSS	non-amplification of *Mus spretus* allele
D2Ucl18	BSB	length differences visible on 2% agarose
D2Ucl19	BSS	mapped by *HinFI* variation
D2Ucl20	BSS	SSCV analysis
D2Ucl31	BSS	non-amplification of *Mus spretus* allele

Table 28:- Methods by which new molecular markers were mapped.

analysis and e) the detection of small length differences by radioactive PCR followed by electrophoresis on 4% polyacrylamide sequencing gels. Table 28 gives details of the method by which each new marker was mapped.

The use of these five methods proved very effective in detecting variation between *Mus spretus* and *Mus musculus*. In only one case (bubble 14) out of 12 was no source of variation identified. This marker was placed onto mouse chromosome two by virtue of its amplification of DNA from the ABm hybrids, which only contain this chromosome. The lack of any variation might indicate the presence of a conserved segment of the mouse genome. Marker bubble 14 also amplified hamster DNA, but produced multiple bands.

5.1.3) Mapping new molecular markers by interspecific backcross analysis

Two interspecific backcross panels were available for the interspecific backcross analysis of new markers for mouse chromosome two. These were the BSB panel (panel 1) - (C57BL/6J X *Mus spretus)* F1 X C57BL/6J and the BSS panel (panel 2) - (C57BL/6J X *Mus spretus*) F1 X *Mus spretus*.

The choice of which panel was used to map the new markers was largely governed by the ease of typing the different sources of variation - in the case of markers typed by non-amplification of the *Mus spretus* allele, there was no choice but to map on the BSS panel. In some cases, the variation was only visible on one panel, such as in the case of *D2Ucl19*, where the restriction enzyme digested only the PCR product from *Mus musculus* DNA. This marker similarly had to be typed on the BSS panel, where the segregation of the *Mus musculus* allele could be followed. Table 28 gives the panel on which each new marker was mapped.

The segregation of one or other of the alleles was followed after amplification of the chosen panel. Anomalous typings were repeated. Common sources of error were found to be failure of amplification of one or other allele, or failure of the restriction digestion. The segregation patterns of all the new markers are given in table A1 in the appendix. The data generated from the typing experiments were then sent to the Jackson Laboratory at Bar Harbor, U.S.A. for analysis. The map positions were generated by the Map Manager program (Manly, 1993). The Jackson laboratory also carried out analysis for transmission distortion, map

distance with standard error and LOD scores for linkage to the nearest markers on the map. It was usually possible to detect linkage to chromosome two by comparison of the segregation pattern of a new marker, compared to that of other markers on the chromosome, from printouts sent from the Jackson.

Figures 32 and 33 give the proposed maps of mouse chromosome two generated from panel 1 and panel 2 respectively. Figure 34 gives the composite map of mouse chromosome two generated by the combination of the data from both mapping panels. This map was generated by comparison of the map positions of the new markers relative to each other. Figures 35 and 36 give the pedigree analysis for panel 1 and panel 2 respectively. Seven additional markers were included in the pedigree analysis of the BSB panel, to act as reference loci. The additional loci were mapped by other investigators (A. Pilz, personal communication) The map produced from analysis was : - *D2Mit1* - (16.2 +/- 3.8cM) - *Dbh* - (2.12 +/- 1.5cM) - *Spna2/D2Ucl11* -(2.25 +/- 1.6cM) - *D2Ucl16* -(18.1 +/- 4.0cM) - *D2Ucl12* - (6.45 +/- 2.1cM) - *Acra* - (4.32 +/- 2.1cM) - *D2Ucl15* - (13.96 +/-3.6cM) - *Il1b* - (30.95 +/- 3.6cM) - *Gnas* - (2.13 +/- 1.4cM) - *D2Mit74/D2Uc18*. In the case of the BSS panel, three additional markers were included in the pedigree analysis to act as reference loci. With the exception of D2Mit52, the other reference loci were mapped by others (A. Pilz, personal communication). The map produced from the analysis was : - *Acra* - (11.7 +/- 3.3cM) - *D2Ucl19* - (12.8 +/- 3.4cM) - *Il1b* - (3.18 +/- 1.8cM) - *D2Ucl13* - (8.5 +/-2.9cM) -*D2Ucl17* - (4.3 +/-2.1cM) - *D2Ucl14/D2Ucl20* - (1.06 +/- 1.1cM) - *D2Ucl31* - (6.38 +/- 2.5cM) - *D2Mit52*.

All new markers produced were found to map to mouse chromosome two. This is fairly unexpected, since a proportion of mouse chromosome 15 was also present in EBS18Az, and a proportion of the IRS-PCR products will have derived from this.

5.1.3a) Analysis of the distribution of the new *D2Ucl* markers on mouse chromosome two

Whilst the number of new markers generated in this study remains fairly small, some interesting features of the distribution of these new markers on mouse chromosome two have emerged. Analysis of the composite map reveals that the overall spread appears fairly homogeneous and no clustering of markers in the central portion of the chromosome is noted. This clustering has been seen in most backcrosses, and is due to the higher recombination rates at the ends of the chromosomes. The Jackson backcross maps overall do show this typical clustering of markers in the middle of the chromosome, but there is no evidence of the *D2Ucl*

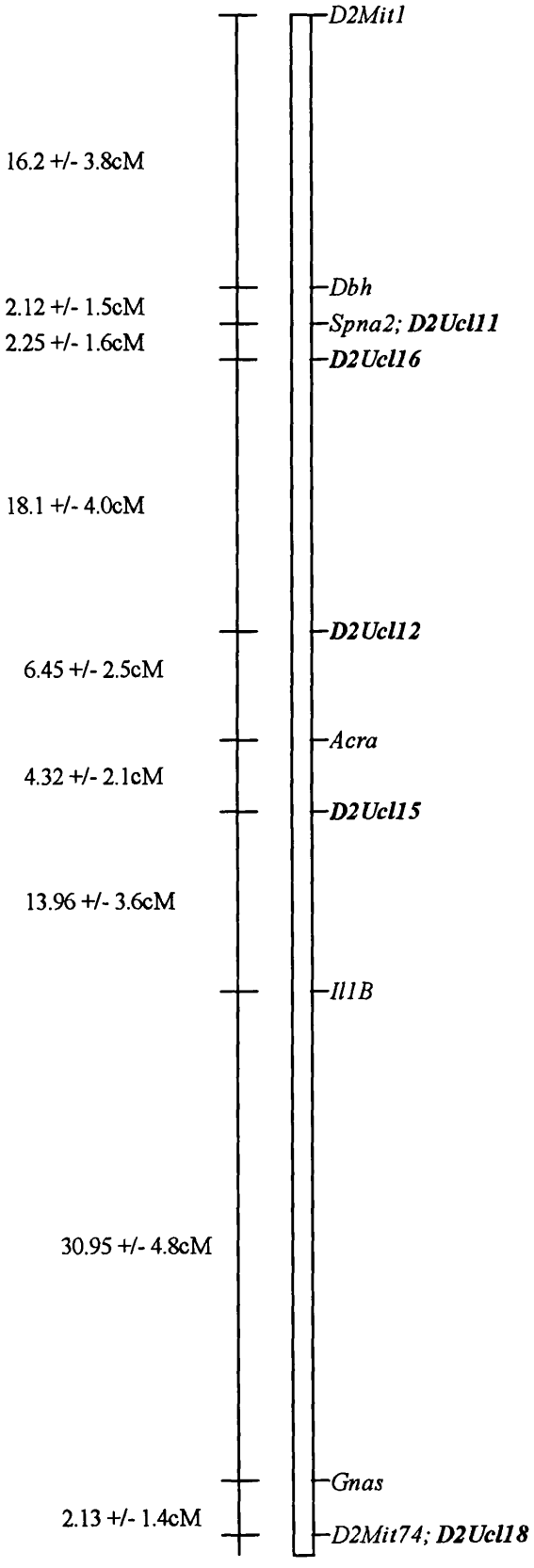


Figure 32: - Proposed map of mouse chromosome two including new *D2Ucl* loci (marked in bold) - panel one

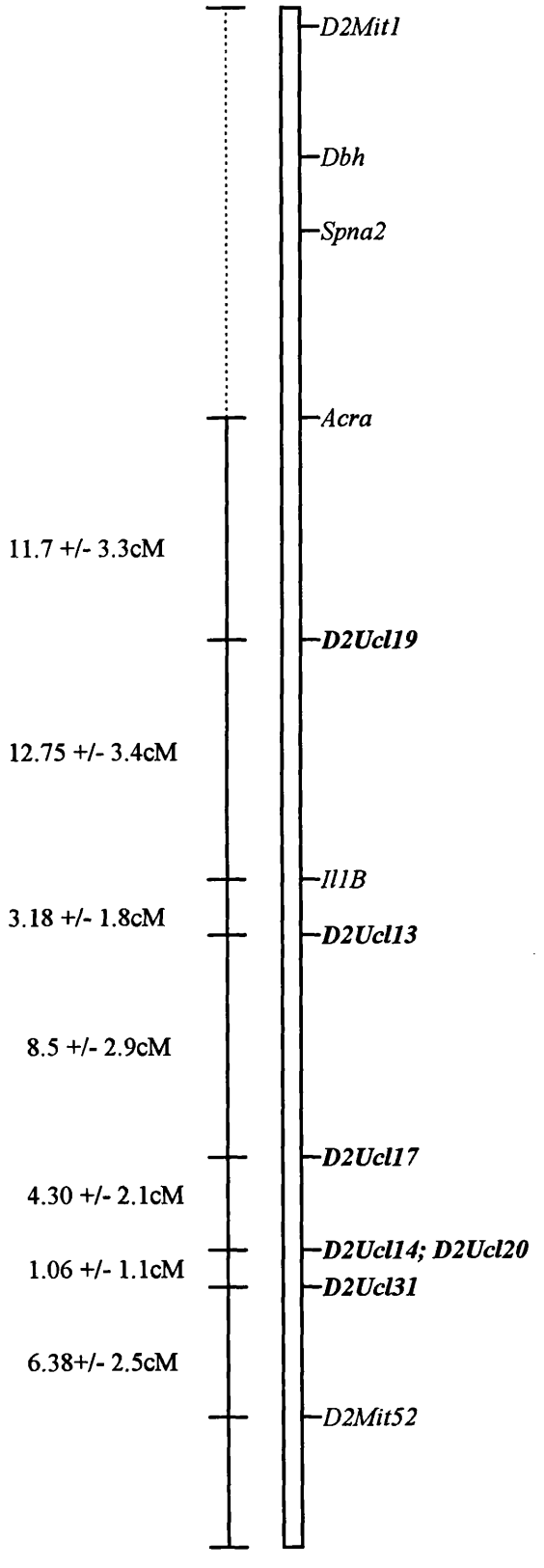


Figure 33: - Proposed map of mouse chromosome two including new *D2Ucl* loci (marked in bold) - Panel two

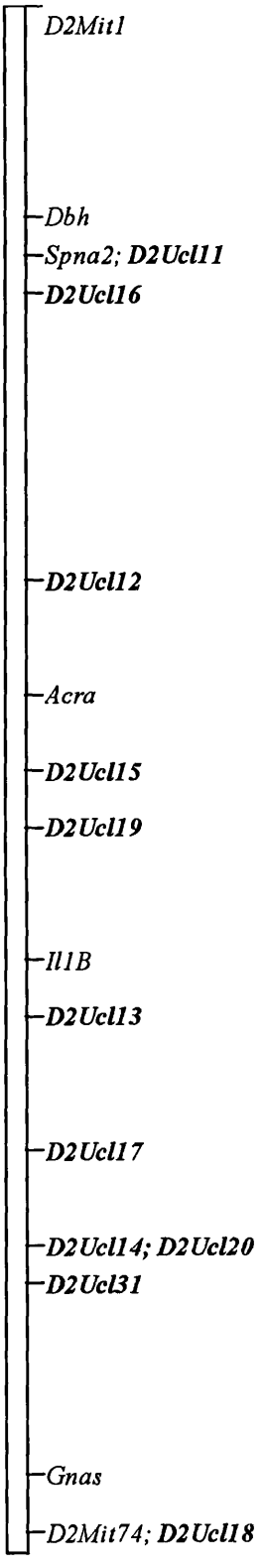


Figure 34: - Proposed composite map of mouse chromosome two

markers generated in this study fitting this pattern. Instead, there is evidence of clustering of these markers in the distal portion of the chromosome, particularly in bands G and H. Indeed, three of four markers mapping in this region map very close together in the region surrounding *Xmv10*. The remaining marker maps approximately 4.3cM proximal to this. This finding is unexpected, especially in view of the known recombination hotspot in this region (Lyon, 1976, Searle *et al.*, 1979). Those markers which co-segregate have all been analysed to ensure that they do not represent the same fragment, by direct comparison of sequence, and by amplification of the plasmid containing the insert from which one pair of primers were generated with the primers derived from the others. All primers were found to derive from different IRS-PCR products.

There are several explanations for this finding. Firstly, it could be a reflection of the distribution of L1 and B2 repeats in this area of the genome - it may be that distal mouse chromosome two is rich in SINES and LINES. Two of these markers are derived from L1 amplification, one from L1/B2 amplification and one from B2-bubble amplification. Alternatively, and probably more likely, is that the amplification reactions favoured the IRS-PCR products produced from this region. Preferential amplification of a subset of IRS PCR products has been noted in the IRS-bubble PCR reactions carried out in this study, in which only 5/15 fragments were different. The exact reasons why these distally located fragments has been favoured is not easy to determine. It may be that the spacing of repeat units in this region is such that the PCR products produced were of smaller size, and thus were more easily amplified. It could be that the amplification dynamics of the reaction were such that the these fragments rapidly overtook the reaction in the first few cycles to the exclusion of many others. It is also possible that the original hybrid population had some degree of heterogeneity for the proximal portion of mouse chromosome two. If the proximal region was only present in a proportion of cells, whereas the distal portion was present to a larger degree, it would be expected that a larger proportion of markers originating from this hybrid would map distally.

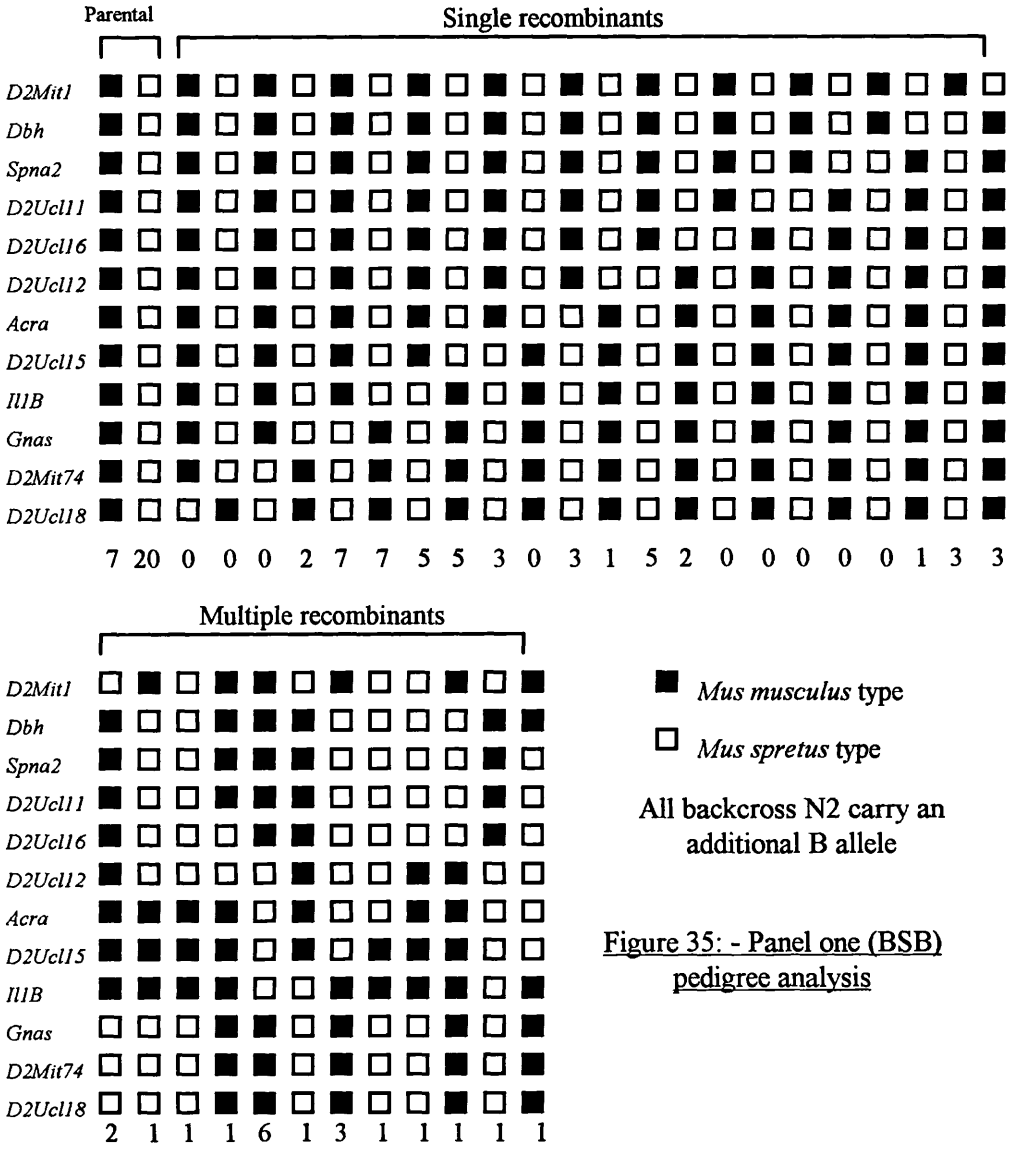


Figure 35: - Panel one (BSB) pedigree analysis

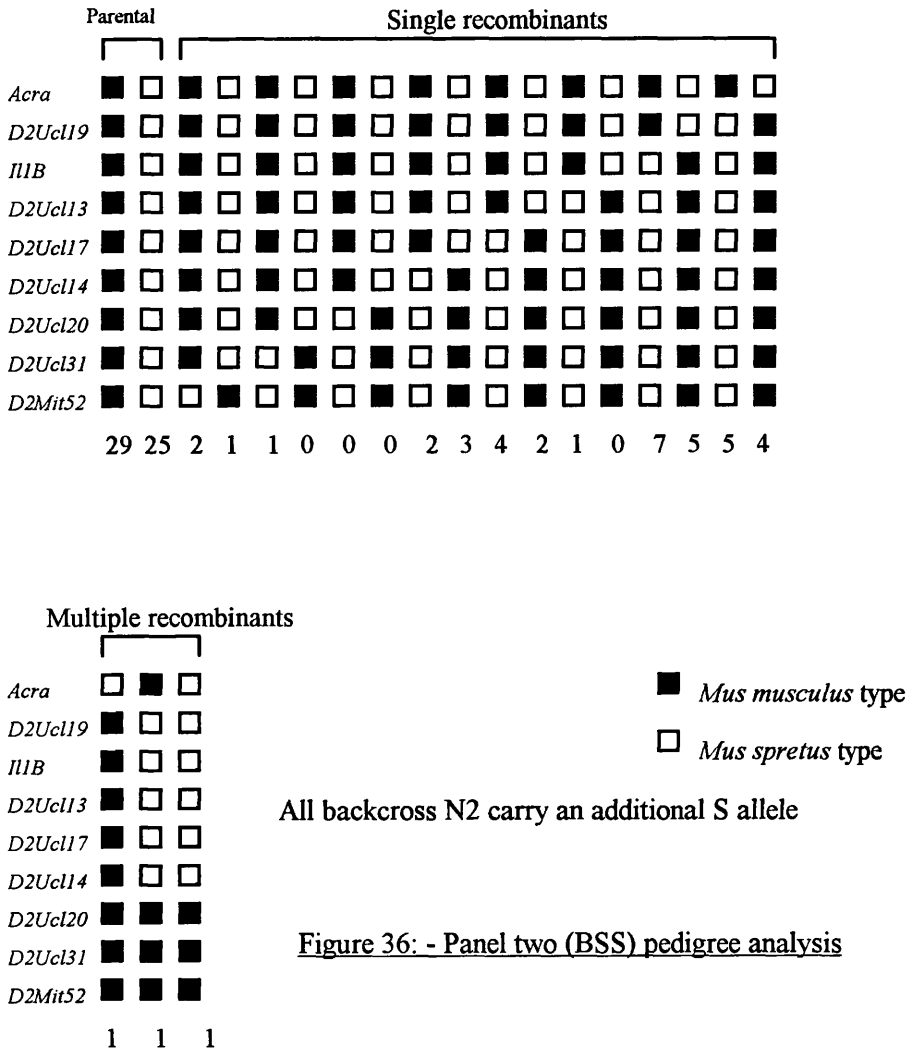


Figure 36: - Panel two (BSS) pedigree analysis

5.1.3b) Analysis of transmission distortion

Transmission distortion for distal mouse chromosome two has been noted on the BSB panel in the Jackson laboratory backcrosses (Rowe *et al.*, 1994). Tables 29a and 29b shows the χ^2 and P values for each new *D2Ucl* locus on mouse chromosome two, calculated from the raw data provided by the Jackson laboratory. As described by Rowe *et al.* (1994), no transmission distortion is noted for loci on panel 2 (BSS). Significant distortion is noted in favour of the *Mus spretus* allele for loci on mouse chromosome two, in the BSB panel. Two loci, *D2Ucl12* and *D2Ucl15* show evidence of transmission distortion, in favour of the *Mus spretus* allele :- P= 0.03 for *D2Ucl12*, and P=0.01 for *D2Ucl15*. P values under 0.05 are generally taken as being significant. Loci showing significant transmission distortion are highlighted in bold type on table 29a. At this stage, it is not determined which gene or genes are causing the transmission distortion effects.

5.2) The determination of map position for *D2Ucl12, D2Ucl19* and *D2Ucl18* relative to ragged, wasted and ulnaless.

Four of the new *D2Ucl* markers produced were found to map in the region of ragged (*Ra*), wasted (*wst*) or ulnaless (*Ul*), and were chosen for further analysis on the specialized backcrosses.

5.2.1) The ragged backcross

One marker, *D2Ucl18*, mapped in the same region of mouse chromosome two as ragged, according to the Jackson Laboratory interspecific backcross panel BSS. Figure 37 gives the map position of this marker on the ragged backcross. The map produced was : - *Il1b* - (37 +/- 3.6cM) - *Gnas* - (1.21 +/- 0.85cM) - *D2Mit73* - (1.21 +/- 0.85cM) - *D2Ucl18* - (1.21+/- 0.85cM) - *D2Mit74/Acra4/Ra*. Figure 38 gives the pedigree analysis for the ragged backcross. The segregation pattern for this backcross is given in table A2 in the appendix.

D2Ucl18 maps an estimated 1.21cM from ragged. This rules out *D2Ucl18* as a closely linked flanking marker, although it could be useful as an STS for YAC analysis of this region.

Locus	Observed B alleles	Observed S alleles	Expected number of each for 1:1 ratio	Deviation from expected for B alleles	Deviation from expected for S alleles	χ^2	P (1 degree of freedom)
D2Ucl11	45	49	47	-2	+2	0.096	0.80
D2Ucl16	44	50	47	-3	+3	0.266	0.75
D2Ucl12	**36**	**58**	**47**	**-11**	**+11**	**4.691**	**0.03**
D2Ucl15	**34**	**60**	**47**	**-13**	**+13**	**6.649**	**0.01**
D2Ucl18	40	54	47	-7	+7	1.798	0.15

Table 29a: - χ^2 analysis for BSB panel

Locus	Observed B alleles	Observed S alleles	Expected number of each for 1:1 ratio	Deviation from expected for B alleles	Deviation from expected for S alleles	χ^2	P (1 degree of freedom)
D2Ucl19	50	44	47	+3	-3	0.266	0.65
D2Ucl13	46	48	47	-1	+1	0.011	0.90
D2Ucl17	46	48	47	-1	+1	0.011	0.90
D2Ucl14	48	46	47	+1	-1	0.011	0.90
D2Ucl20	48	46	47	+1	-1	0.011	0.90
D2Ul21	47	47	47	0	0	-	-

$$\chi^2 = \text{Sum } \frac{[(\text{obs} - \text{exp})]^2}{\text{exp}}$$

(Strickberger, 1985)

Table 29b:- χ^2 analysis for BSS panel

Transmission distortion has also been analysed in this backcross. Table 30 shows the χ^2 and P values for loci on the ragged backcross. No significant transmission distortion was noted for any of the loci tested, although *Il1b* was only just not significant at P = 0.06, with a slight excess of alleles originating from *Mus musculus*.

5.2.2) The wasted backcross

D2Ucl18 was also placed on the map of distal mouse chromosome two by analysis of the wasted backcross. Figure 40 gives the map position of this marker from the wasted backcross. The map produced was: - *Il1b* - (22.3 +/- 4.3cM) - *Gnas* - (1.1 +/- 1.1cM) - *D2Mit73/D2Ucl18* - (1.1 +/- 1.1cM) - *D2Mit74* - (1.1 +/- 1.1cM) - *Acra4* - (3.2 +/- 1.8cM) - *wst*. The pedigree analysis of the wasted backcross is given in figure 39. The segregation pattern for this backcross are given in table A3 in the appendix.

This finding places *D2Ucl18* approximately 5.5cM proximal to *wst*. The results of this backcross are interesting, because of the apparent distance of wasted and *Acra4*. *Ra* and *wst* are very closely linked (Peters *et al.*, 1994). It is highly unexpected to find *Acra4* and *wst* separated by an estimated 3.2 +/- 1.8cM on the wasted backcross, since ragged cosegregates with *Acra4* on the ragged backcross.

Transmission distortion could not be analysed for this cross, because only the *wst/wst* homozygotes were included in the backcross.

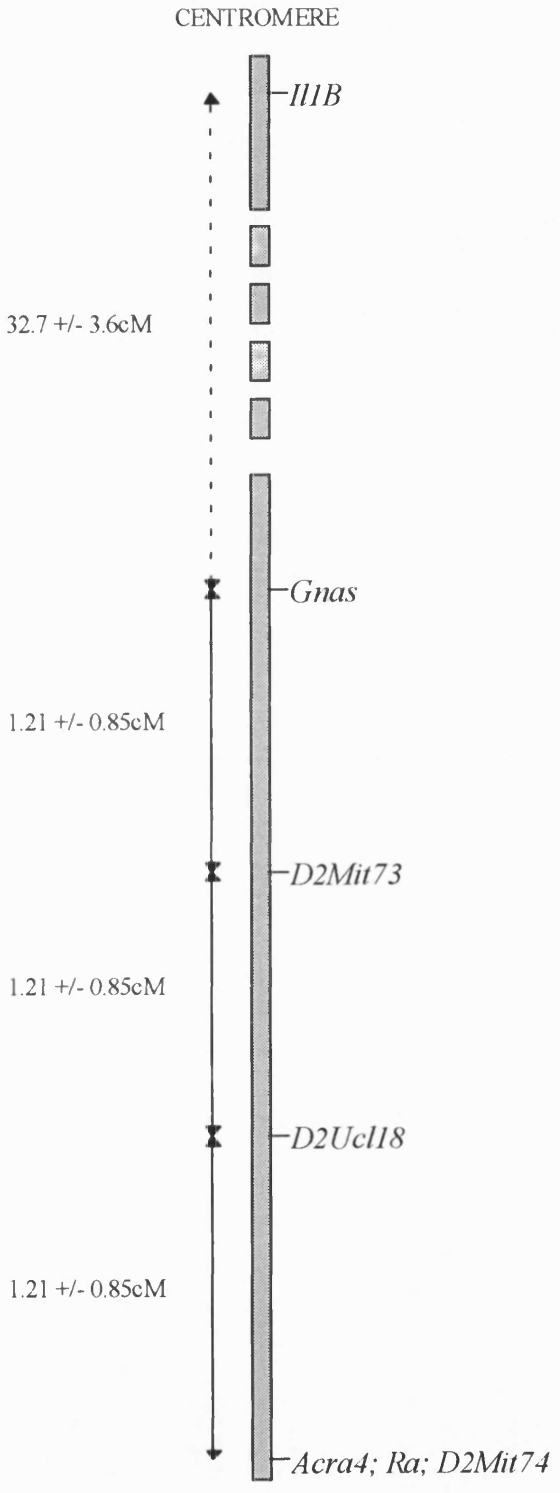


Figure 37: - *D2Ucl18* mapped relative to ragged

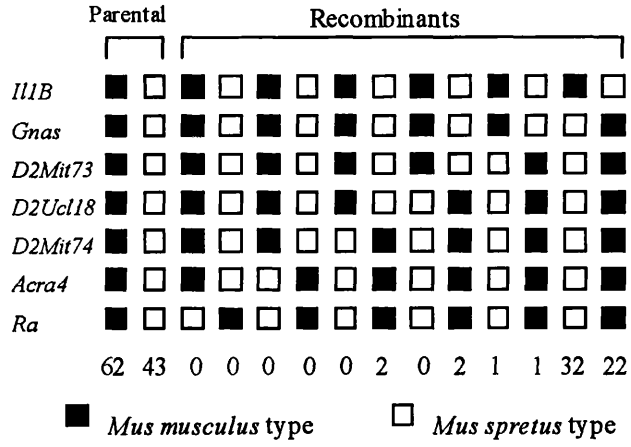


Figure 38: - Pedigree analysis for the ragged backcross

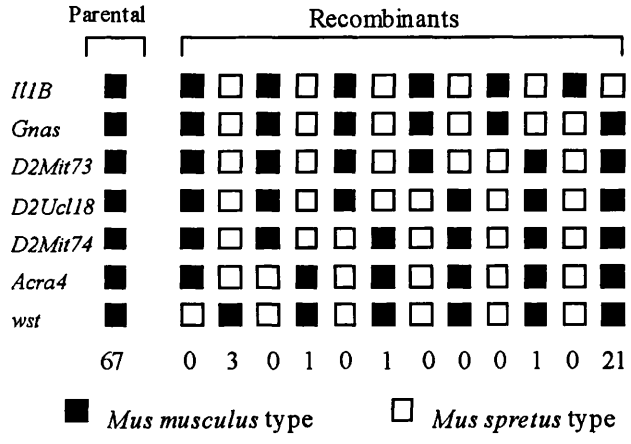


Figure 39: - Pedigree analysis for the wasted backcross

Locus	Observed *Mus musculus* alleles	Observed S alleles	Expected number of each for 1:1 ratio	Deviation of *Mus musculus* alleles from expected	Deviation of S alleles from expected	χ^2	P
Il1b	95	70	82.5	+12.5	-12.5	3.490	0.06
Gnas	85	80	82.5	+2.5	-2.5	0.097	0.93
D2Mit73	85	80	82.5	+2.5	-2.5	0.097	0.93
D2Ucl18	87	78	82.5	+4.5	-4.5	0.388	0.55
D2Mit74	88	77	82.5	+5.5	-5.5	0.606	0.40
Acra4	89	76	82.5	+6.5	-6.5	0.873	0.45
Ra	89	76	82.5	+6.5	-6.5	0.873	0.45

$$\chi^2 = \text{Sum} \frac{[(\text{obs} - \text{exp})]^2}{\text{exp}}$$

(Strickberger, 1985)

Table 30: - χ^2 analysis for the ragged backcross

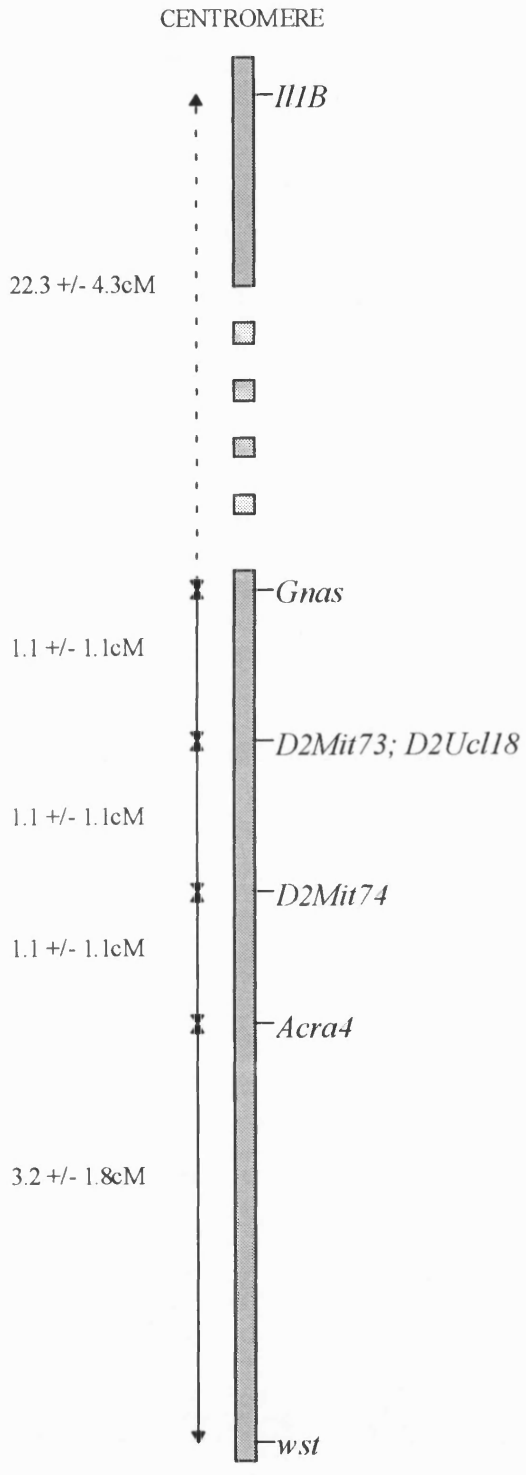


Figure 40: -*D2Ucl18* mapped relative to wasted

5.2.3) The ulnaless backcross

In the case of the ulnaless cross, markers judged to lie in the *D2Mit30* to *D2Mit11* region of mouse chromosome two (by comparison of maps produced by the Jackson Laboratory interspecific backcross, and the consensus map of mouse chromosome two) were analysed. By comparison of the consensus linkage map, and the maps produced from the Jackson Laboratory, three markers were judged to lie in this region: -*D2Ucl12*, *D2Ucl15* and *D2Ucl19*. Of these, *D2Ucl12* and *D2Ucl19* showed suitable variation between *Mus musculus* and CAST/Ei - a size difference in the case of *D2Ucl12*, and a *HinF1* RFLV in the case of *D2Ucl19*. No variation was found between *Mus musculus* and CAST/Ei for *D2Ucl15*.

Figure 41 shows the map produced of mouse chromosome two, by analysis of the ulnaless backcross. Complete pedigree analysis for the ulnaless cross is given in figure 42. The map orders and distances produced were: telomere - *D2Mit30* -(8.6 +/- 2.9cM)- *D2Ucl19* - (3.2 +/- 1.8cM) - *Cf2/D2Mit14* - (2.2 +/- 1.5cM) - *Ul/+* - (2.2 +/- 1.5cM) - *D2Ucl12* - (1.1 +/- 1.0cM) - *D2Mit11* - centromere. Only the two *D2Ucl* loci were typed during the course of these experiments. The other loci were typed elsewhere (Dr. C. Abbott, personal communication). The raw data for the ulnaless backcross are given in table A4 in the backcross. This map places *D2Ucl12* 2.2 +/- 1.5cM distal to the ulnaless gene. This distance represents approximately 4 megabases, and is therefore too large a distance to make *D2Ucl12* a suitable flanking marker for positional cloning experiments. However, this marker could be useful as an STS for any YAC clones produced in the future. The other marker, *D2Ucl19,* maps considerably more proximal to ulnaless. Ulnaless maps close to *Acra* on the consensus map. *D2Ucl15* maps closer to *Acra* than either *D2Ucl12* or *D2Ucl19*. If a source of polymorphism could be determined for this marker, it is likely that it would map in fairly close proximity to ulnaless, and could possibly serve as a flanking marker for positional cloning experiments.

The presence of transmission distortion effects in this backcross were studied by χ^2 analysis. Table 31 gives the χ^2 and P values for this backcross. No transmission distortion effects were noted.

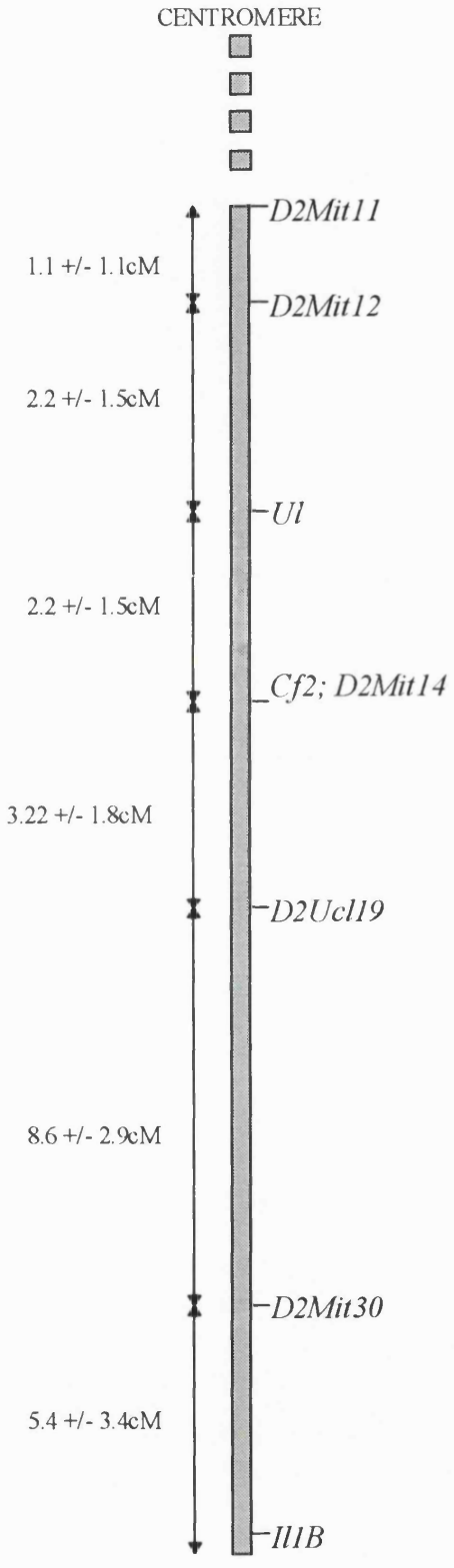


Figure 41:- *D2Ucl12* and *D2Ucl19* mapped relative to ulnaless

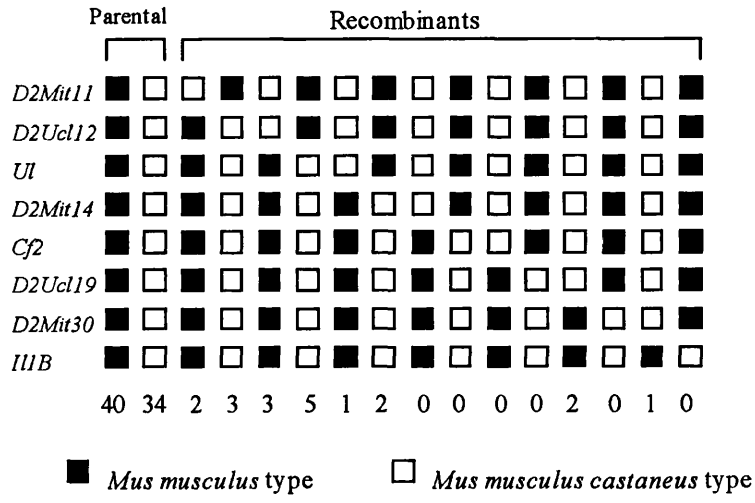


Figure 42: - Pedigree analysis for the ulnaless backcross

Locus	Observed number of *Mus musculus* alleles	Observed number of C alleles	Expected number of each for 1:1 ratio	Deviation from expected for *Mus musculus* alleles	Deviation from expected for S alleles	χ	P
Il1b	46	35	40.5	+5.5	-5.5	1.235	0.28
D2Mit30	12	18	15	-3	+3	0.833	0.45
D2Ucl19	10	11	10.5	-0.5	+0.5	-	-
Cf2	9	15	12	-3	+3	1.042	0.30
D2Mit14	13	21	17	-4	+4	1.441	0.22
Ul	45	47	46	-1	+1	0.011	0.90
DUcl12	12	10	11	+1	-1	0.045	0.80
D2Mit11	49	44	46.5	+2.5	-2.5	0.172	0.70

$$\chi^2 = \text{Sum } \frac{[(\text{obs} - \text{exp}) - 0.5]^2}{\text{exp}}$$

(Strickberger, 1985)

Table 31: - χ^2 and P values for the ulnaless backcross

5.2.4) Estimated positions of new *D2Ucl* markers on the cytogenetic map

Since SINES and LINES have been found to map to different chromosome bands (see introduction), it would be expected that markers derived from IRS-PCR products would show a similar segregation to G positive or G negative chromosome bands. The estimated positions of the new loci on the cytogenetic map were generated as follows. Markers already placed on the cytogenetic map were used as reference points. The map position of these markers and the new markers on the consensus genetic map was compared. The cytogenetic map positions of new markers were only predicted for markers which were closely linked to points on the cytogenetic map. It was assumed that in the case of very closely linked markers such as *Abl* and *D2Ucl11*, the position of *D2Ucl11* on the map could be predicted by *Abl*. Markers which did not co-segregate with existing markers were allocated map positions according to the cytogenetic map positions of flanking loci. The predicted positions of the new *D2Ucl* loci on the cytogenetic map is given in figure 43. *D2Ucl11* and *D2Ucl16* are closely linked to *Abl* (2.1 +/- 1.5cM and 1.1 +/- 1.1cM on the Jackson backcross panel 1, respectively) , which has been placed in band

B on the cytogenetic map. From this, it was estimated that *D2Ucl11* and *D2Ucl16* would also lie in band B. Similarly, *D2Ucl14*, *D2Ucl20* and *D2Ucl31* all lie in close proximity to *Emv15*, placed in band H1 on the cytogenetic map. *D2Ucl14* and *D2Ucl20* co-segregate with *Emv15* according to the consensus map, *D2Ucl31* is separated from it by approximately 1.1 +/- 1.1cM on the Jackson backcross panel 2. It could therefore be postulated that *D2Ucl14*, *D2Ucl20* and *D2Ucl31* all map to band H1. Figure 43 illustrates the band allocations of the *D2Ucl* markers. *D2Ucl16*, *D2Ucl12*, *D2Ucl15*, *D2Ucl19*, *D2Ucl13*, *D2Ucl14* were derived from L1 repeats, and be expected to map in the Giemsa dark bands. *D2Ucl11*, *D2Ucl17*, *D2Ucl18* were derived from B2 repeats and would be expected to map in the Giemsa light bands. *D2Ucl20* and *D2Ucl31* derive from both L1 and B2 repeats, and would be expected to map to junction regions. Overall, the results do fit the pattern, with some discrepancies, notably *D2Ucl16* and *D2Ucl14*. The other L1 based repeats appear to map to Giemsa dark bands. All the B2 based repeats except *D2Ucl18* map to Giemsa light regions. The L1/B2 derived markers, *D2Ucl20* and *D2Ucl31* appear to map in the Giemsa light regions, but they could lie on either the proximal or distal boundaries, in a region where both SINE and LINE repeats are present at fairly high frequency.

The existence of the anomalous cytogenetic typings could be explained by the fact that although the different Giemsa bands are depleted in either SINE or LINE repeats, there are nevertheless some repeats of each type present in all bands. It could be that again, certain fragments, derived from the uncommon L1 repeats in the Giemsa light bands are preferentially amplified for some reason in the PCR reaction.

It must also be noted that these cytogenetic locations are very rough estimates, and there are bound to be errors. The only way to be certain of the cytogenetic location would be to place the new markers on the cytogenetic map by FISH analysis. *In situ* hybridization is complicated in the mouse due to the difficulty in identifying specific mouse chromosomes. In addition, the IRS-PCR and IRS-bubble PCR products produced were in all likelihood too small for use as cytogenetic probes. For these reasons, and for reasons of time constraints, this experiment was not attempted.

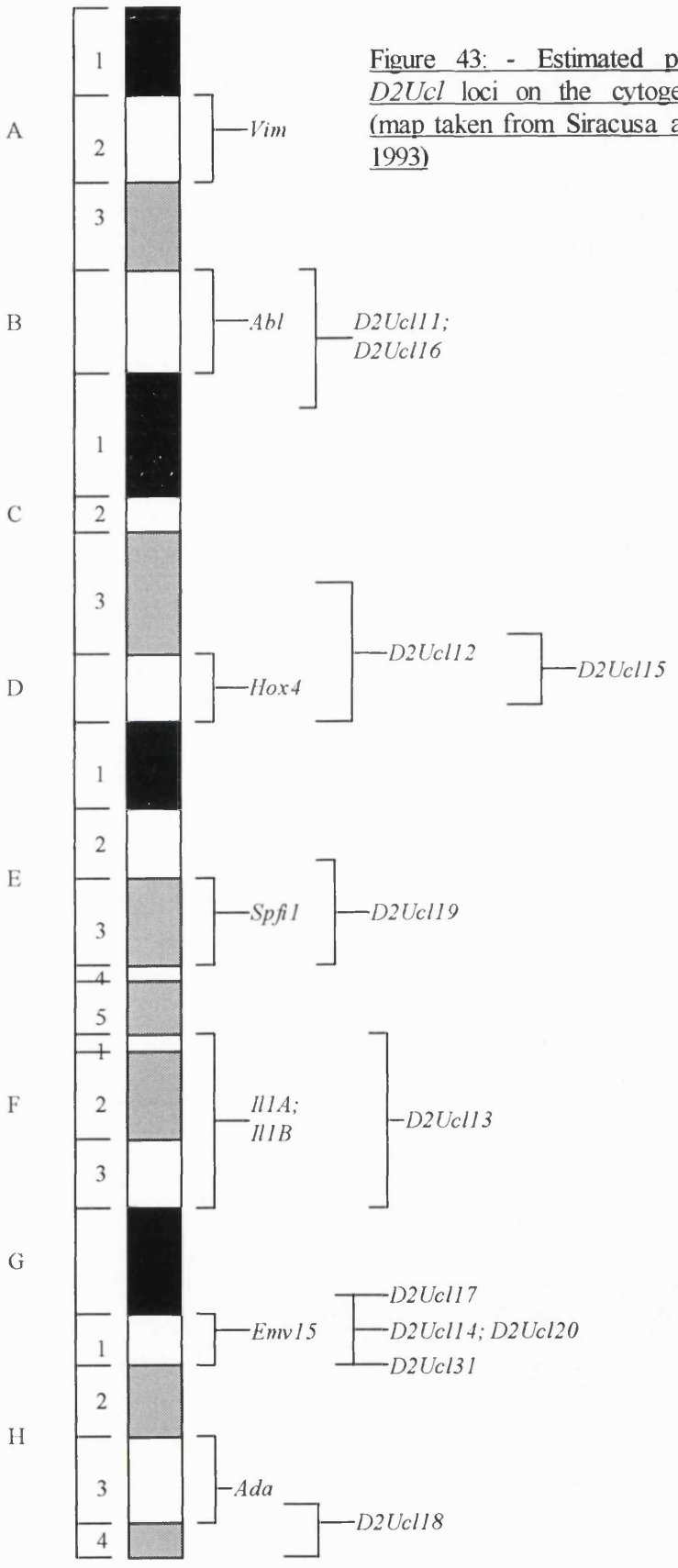


Figure 43: - Estimated positions of *D2Ucl* loci on the cytogenetic map (map taken from Siracusa and Abbott, 1993)

Chapter 6

Discussion

6.1) The construction and characterization of two new panels of somatic cell hybrids

6.1.1) The characteristics of the hybrid panels

Two new panels of mouse Chinese hamster somatic cell hybrids were constructed and characterized. The mouse parent in each case was homozygous for either the T(2;8)2Wa or the T(2;16)28H reciprocal translocations.

The first observation to be considered is the fact that colony morphology, cell morphology and growth rates are widely varied within a single panel. This observation could be expressed as some hybrids being more 'mouse type' and some hybrids being more 'hamster type'. This variation in gross features of the cell, such as growth rate and colony and cell morphology of the cell population is presumably due to the presence or absence of specific mouse chromosomes, since the hamster chromosome content could be expected to be approximately equal in each case. There may be some correlation between the cell morphology and the growth rate: - cells which are similar in structure to the mouse parent (i.e. small, rounded cells) seem to be slow growing in the majority of cases, although there are exceptions. This is not surprising, since the mouse cell line used in the fusion was a primary cell line, and not well maintained in culture. Cells which are reminiscent of the hamster parent (spindle shaped) may share the properties of the hamster parent which allow it to grow well in culture. This suspicion is given weight by the observation that in most cases, cell lines showing similar structure to the hamster parent grow well in culture, although again, there are exceptions. There appears to be no correlation between the colony morphology and the cell morphology. Cell lines in some cases show characteristics of both parents. Characteristics such as contact inhibition are also widely variant in the hybrid population. The hamster parent showed little evidence of contact inhibition, growing until all nutrients were exhausted from the medium, and then detaching from the base of the culture flask. There appeared to be no sign of a cessation of growth before this point. Some hybrid clones also showed this characteristic, specifically T28H-6 and T2Wa-5.

Wih regard to the segregation rates of the panels, the overall rate of retention is high. There are considerable differences in the mean segregation rates between panels. The reasons for this are unclear. It may be that the mouse cell population in one panel was at a different stage in the cell cycle from the other panel. It has been observed that the fusion of a cell in M to one in G1 can force the second cell into premature mitosis, before the chromatids have been replicated, resulting in an overall loss of chromosomes (Ruddle, 1972). There are differences in the percentage retention for specific chromosomes, for the panels as a whole, indicating that non-random segregation has occurred. The lack of intact segregated translocation chromosomes in either panel is significant. Presumably, there must be a selective advantage for the retention of both reciprocals, or a pressure against segregation for T(2;8)2Wa and T(2;16)28H, although the lack of segregation must be at least partially attributed to the overall low segregation in the panels.

There appears to be little correlation between the loss of a large number of mouse chromosomes and the growth rate. It has been suggested that the loss of large numbers of mouse chromosomes might act to remove some of the constraining elements, which mitigate against the rapid growth and loss of contact inhibition of the transformed parental cell (Ruddle, 1972). Some hybrid cell lines (T28H-6, T2Wa-4) show very rapid growth, and loss of a large number of mouse chromosomes. Other hybrids however, show similar loss of the majority of the mouse chromosomes, but no corresponding increase in growth rate (T2Wa-14). Other hybrids show very fast rates of growth, but very little loss of mouse genetic material (T2Wa-5). It is more likely that the increase in cell growth rate is correlated with the presence or loss of specific mouse genes, and therefore specific mouse chromosomes. The loss of large numbers of mouse chromosomes would of course increase the probability that these chromosomes would not be retained in the hybrid clone. For instance, T2Wa-1, T2Wa-6 and T2Wa-14 all show considerable loss of chromosomes, but a slow growth rate. Presumably, these hybrids have retained the mouse chromosome or chromosomes required to quench the rapid growth ability shown by the hamster parental cell. The only chromosome held in common between these hybrids is chromosome 11. Therefore, it could be proposed that a gene or genes on chromosome 11 might be having growth suppressant effects on the cell. Other hybrids showing evidence of the presence of chromosome 11 are T2Wa-4, T2Wa-5, T2Wa-8 and T2Wa-11b. Of these, T2Wa-4, T2Wa-11b and T2Wa-5 show fast growth, and T2Wa-8 shows moderately fast growth. This could be due to the loss of the chromosome segment containing the growth suppressant genes from these hybrid lines, considering the extent of fragmentation apparent in the panels, or moderation of the effects of one gene by others. In the T28H panel, T28H-6

shows very fast growth, yet retains chromosome 11. The factors controlling the growth rate in hybrid cells are likely to be very complex, and under multiple genetic controls. It is therefore not unexpected to find such a complicated picture amongst these hybrid clones. The presence of certain chromosomes alone may not be sufficient to repress the growth promoting effects of the hamster genes, it is almost certain that other moderating genes play their part. The presence of a high degree of fragmentation makes the correlation of growth rate to specific chromosomes very difficult, since the presence of chromosomal fragments other than those detected in each hybrid clone cannot be ruled out.

Both hybrid panels show evidence of fragmentation, as indicated by non-concordant segregation of markers mapping to the same chromosome in the panels. This is not a complete measure, as chromosomes which have undergone fragmentation without loss of the products would appear as intact. The slightly lower apparent degree of fragmentation in the T28H panel may reflect the lower segregation rate, in that fragmentation products were less likely to be lost, and so attached to Chinese hamster chromosomes. This would give positive characterization signals, which would be taken as the chromosomes being intact. Alternatively, the lower segregation rate could be due to a very high degree of fragmentation, with subsequent attachment of the fragments to the hamster chromosomes, which are not segregating. Evidence for this theory is provided by the observation that the degree of fragmentation is much greater than expected from the PCR results, as reflected by the failure of the panel to predict correct chromosome locations for blind-mapped markers. In actual fact, the T2Wa panel could have a lower degree of fragmentation than the T28H panel, if the fragments were truly lost from the cells, rather than attached to the host chromosomes. There is no doubt that the panels are much more fragmented than indicated by PCR analysis. This fragmentation of chromosomes in panels of somatic cell hybrids derived from translocation mouse stocks has been noted before. The T13H and T1Sn panels constructed by A. J. Pilz also show a high degree of chromosome fragmentation both in the other chromosomes (A. Pilz, personal communication) and also the translocation chromosomes (this study).

The Chinese hamster parental cell line was chosen for its apparent ability to increase the stability of hybrid lines constructed from it. The use of this cell line seemed to have no effect on the stability of the chromosomes, which are extensively fragmented. The use of more plastic embryonic cells rather than spleen cells might have acted to increase segregation and stability in the panels. However, once the segregation (and fragmentation) were complete, the hybrids were stable, as reflected by the lack of alteration in mouse chromosome content over

many cell generations, and passages in culture. The stabilizing effects of V79TOR2 could also partially explain the low overall segregation rates in the panels, but not the apparent difference in segregation rates.

6.1.2) Subcloning experiments on the T2Wa and T28H hybrid panels

The subcloning experiments were carried out in order to generate segregated subclones from clones from both panels which showed evidence of heterogeneity. Both methods employed generated colonies originating from single cells, but the serial dilution was found to result in many wells with more than one cell, or no cells at all. For these reasons, most of the subclones produced originated from the statistical calculation of the cell volume required to give one cell, half a cell or two cells per well.

The results demonstrate the effectiveness of the subcloning technique. Both the L1 PCR experiments and the partial characterization reveal differences in mouse chromosome content between individual subclones originating from the same parental hybrid line. The technique relies upon the presence of chromosomal heterogeneity within the hybrid population in a clone. The subcloning technique is fairly sensitive to even low levels of heterogeneity in the hybrid population, as in the T28H-3b subclone panel, chromosomes 4, 7 and 8 were found to have segregated. Heterogeneity for these chromosomes was not noted in the original PCR characterization. Their segregation in the subclone panels might be due to further segregation of the chromosomes during the longer periods in culture. However, such large scale loss is unlikely to be due entirely to this, since the overall stability of hybrid cells is such that significant alterations in the segregation of larger numbers of chromosomes take place only after considerable time in culture (Weiss and Green, 1966).

No segregation of the translocation products was noted for either the T2Wa panel, or the T28H panel from the subcloning experiment. This lack of segregation is unexpected, the degree of heterogeneity estimated to exist within cells in both the T2Wa-1 clone and the T28H-3b clone is well within the limits for the subcloning experiment, as demonstrated by the segregation of chromosomes 7 and 8 in the hybrid panel. This is further evidence for the action of a selective pressure against the segregation of the translocation products. There is a very high degree of fragmentation of the 8:2 reciprocal (50%) in the T2Wa 1 subclone panel (this chromosome was judged to be intact in T2Wa-1), which could reflect the response of the cell to pressure to segregate. The 2:8 chromosome was not found to be fragmented in any case, nor

were the 2:16 and 16:2 derivatives in the T28H panel. This selective pressure is thus probably strongest in the case of the 8:2 chromosome. The pressures must also be present at a lesser degree for the other translocation products, because they do not segregate. The basis of this selective pressure is uncertain, but may demonstrate a requirement for mouse chromosome 2, mouse chromosome 16 and mouse chromosome 8 in the hybrid cells for some reason. This has not been noted on other panels. It is interesting to note that in the T28H panel, chromosome 8 is also significantly over-represented, and in the case of the T2Wa panel, chromosome 16 is seen to be significantly over-represented. The high level of heterogeneity for the translocation products (both 2:16 and 16:2 in T28H-3b, 8:2 in T2Wa-1) would at first appear to be at odds with the presence of a selective pressure for retention of both the translocation chromosomes. However, this could be explained if the translocation chromosomes which were undergoing segregation in a proportion of the cells in the clone were fragmented, and thus stabilized that way, as in T2Wa-2 and T2Wa-4.

These results also indicate a degree of low level heterogeneity for many chromosomes within the clones. This is reflected by the segregation of chromosomes which were not judged to be heterogeneous in the T28H-3b subclone panel. Only a few chromosomes were seen to be very heterogeneous as seen from the characterization tables (tables 11a and 11b). However, the occurrence of segregation for other chromosomes in the T28H-3b subclone population indicates that these chromosomes must have been at least low level heterogeneous in T28H-3b.

6.1.3) Antibodies as selective agents

These experiments determined that the addition of antibodies and complement could reliably and reproducibly direct cell killing by the complement cascade. The critical factor was found to be the complement concentration. This was presumably because of the complexity of the reaction. The role of the antibody has only one stage, the binding to the Cd44 antigen on the mouse cell membranes. The role of the complement is much more complicated, requiring the recognition of the antibody : antigen complex, and then the sequential addition of all the necessary complement subunits (which must all be present at adequate concentration) to form a pore in the cell membrane. It was therefore not surprising that the complement is the critical component. The relationship between the complement concentration and the cell killing is fairly simple. The addition of more complement at first caused a rapid reduction in cell numbers, which then reached a plateau giving only small reductions in the cell killing per extra percent complement. This suggests that the system becomes saturated with all the complement that is

needed, with any further additions were being in excess. This can be seen from the shape of the curve produced from the graph given in figure 22.

The same experiments carried out on the V79TOR2 Chinese hamster cells indicate that this cell line was equally susceptible to the action of complement. This is because the antibody must have recognized the epitope on the Chinese hamster Cd44 antigen. This is the reason that a polyclonal antibody would have been preferable. The use of such a mixture of antibodies would have allowed the adsorption of the hamster cross reactivity by prior growth on a lawn of hamster cells. Such an antibody was not available. The dynamics of the reaction appear to be similar to that seen in the original experiments with mouse 1R cells.

The presence of significant cross reactivity with hamster cells precluded the use of this system as a selective method for the T2Wa and T28H hybrid panels. However, in the light of the apparent selective pressures for the retention of both reciprocal translocation products, it is likely that this experiment would also have produced fragmentation in the cell lines. The technique in itself is potentially very useful. Given the existence of an appropriate antibody, and the existence of appropriate genes expressed on the cell surface, it is theoretically possible to select for any chromosome of interest. Since the means now exist to tailor make any antibody required, and there is no shortage of cell surface antigens, this is a real possibility. Not only does the technique have the potential to provide negative selection against particular chromosomes, it also has the possibility of providing a means of isolating physically the two forms of cell in a heterogeneous population such as clone T28H-3b. This could be accomplished by conjugating the antibody to a magnetic bead. Instead of using the complement mediated cell killing pathway, it would then be possible to use a magnetic cell sorter to separate the cells containing the antigen from those which did not.

6.1.4) The alignment of the genetic and physical map of mouse chromosome two

The alignment of the genetic and physical maps of mouse chromosome two would be an important step to aid the positional cloning efforts for genes on distal mouse chromosome two. The presence of the distal recombination hotspot makes any real estimate of physical distance from genetic distance difficult. An accurate physical distance is of course invaluable for assessment of the usefulness of flanking markers as staring points for chromosome walks, and other methods of positional cloning. The mapping of genes mapping close to physical breaks in the chromosome (such as translocation breakpoints) relative to those breaks would be

an important step in the alignment of the maps. Thus, studies were undertaken to characterize the translocation chromosomes in hybrids known to have segregated them. The results produced indicate that the translocation chromosomes are extensively fragmented in the T1Sn and T13H hybrid panels, as they are in the T2Wa panel. This is not unexpected, given the highly fragmented state of the other chromosomes in the clones used. This is perhaps a similar phenomenon to that which is potentially occurring in the T2Wa and T28H hybrid panels, and reflects some selective pressure for the retention of both reciprocals.

6.2) The generation of new molecular markers for mouse chromosome two by IRS-PCR and IRS-bubble PCR

New DNA markers for mouse chromosome two were generated by three different methods. Firstly, IRS-PCR was carried out on DNA from sources containing only mouse chromosome two, or mouse chromosome two plus small fragments of other chromosomes. These fragments were then either cloned directly from the PCR reaction mixes, or were first isolated on low melting point agarose gels, and then cloned. The final technique involved the use of IRS-bubble PCR, to generate DNA fragments, which were then cloned directly from the PCR reactions without prior isolation.

6.2.1) The advantages and disadvantages of the various cloning methods

All three methods had some advantages. The relative merits and shortcomings of each technique are laid out below.

6.2.1a) IRS-PCR followed by isolation of individual bands

The first method of IRS-PCR amplification, followed by the isolation of individual PCR products on low melting point agarose gels, had many advantages. Firstly, the method allowed the examination of the fragments on an individual level, prior to cloning: it was possible to screen for the presence of putative CpG islands, and microsatellite repeats. The presence of CpG islands would suggest the presence of the 5' end of a gene. The presence of microsatellites facilitates the identification of variation between the parental backcross animals. Also, had the fragments been less repetitive, it would have been possible to use them directly as probes on Southern blots, without the need for cloning. Many different fragments were produced from a single amplification, there was only one fragment which was duplicated out of 10 different fragments, although various portions of this fragment was present in four of the plasmids isolated, including one B2 inter-repeat product. The disadvantage of this technique was the high level of hamster fragments which were co-amplified. In some cases, the hamster fragments appeared to be in the majority in the reaction. This is highly unexpected in the light of the apparent species specificity of all the primers used. This could be due to the amplification of very small amounts of hamster DNA in the original reaction, which were then amplified further during the next amplification step. This hypothesis is substantiated by the

observation that no hamster fragments were noted amongst the clones isolated directly from the initial PCR reactions, either from the IRS-PCR reactions, or from the bubble IRS-PCR fragments. The numbers of clones produced by the L1/B2 amplification are too small to be absolutely sure that no hamster clones would have been present. However, in the case of the IRS-bubble PCR, using the B2 primer, no hamster clones were noted out of a total of twelve. Methods were available to circumvent the hamster breakthrough, although these did add a further two or three days to the amount of time needed for the analysis of the clone.

6.2.1b) IRS-PCR followed by cloning directly from the amplification reactions

The use of the initial PCR reactions rather than the isolated bands as the cloning substrate had different advantages. The removal of the need for the isolation of the bands prior to cloning substantially decreased the time and effort required to clone the fragments. However, this was counteracted somewhat by the limited number of clones produced. The B2 amplification yielded no mouse fragments at all, and the L1/B2 amplification only yielded three separate fragments, although these were all different. The small number of clones produced required the need for more than one cloning experiment. However, no breakthrough of hamster clones was noted. Although the number of clones produced was very small, given the proportion of hamster fragments occurring in the prior experiments, it would have been expected that one of the fragments at least would have been hamster, if this was as much of a problem with this technique. The numbers are too small to be statistically significant, but are somewhat substantiated by the results from the IRS-bubble PCR, which also produced no hamster fragments, out of a larger number of clones. Another problem with this technique was the high level of vector colonies produced.

6.2.1c) IRS-bubble PCR

There were also certain advantages to the use of the IRS-bubble PCR technique. Firstly, the use of this technique produced larger numbers of clones, none of which were of hamster origin. The B2 primer used gave only 33% mouse clones in the original experiments. It would be expected that if there was any significant hamster co-amplification, hamster clones would have been noted in this number (12) of fragments. From these results, it was concluded that the high hamster background observed in the original experiments was due to the re-amplification step. Another advantage lies in the expectation that this method would be likely to allow the production of markers from areas under-represented in SINES and LINES, since only

one repeat is needed. Many SINE and LINE repeats are present in tandem arrangement on the chromosome (Moyzis *et al.*, 1989), from which conventional IRS-PCR would not yield results. IRS-bubble PCR should allow the generation of markers from these regions, which would not be reached by IRS-PCR. There were two apparent disadvantages to this technique. Firstly, there were significant numbers of plasmids containing no inserts (approximately half). This occurred with two separate preparations of vector, and may be due to the lower proportion of inserts in the ligation reactions in this method, and in the former method, compared to the use of isolated and re-amplified PCR products. Secondly, there was a limited number of different fragments produced. Out of a total of twelve clones isolated, only five were different. This was due to the amplification of five DNA fragments, which for some reason were highly over-represented in the fragment population. This over-representation of these fragments could have occurred during several stages of the experiment. Possibly, the ligation of the vectorette unit and the *HaeIII* digested EBS18Az DNA produced only these limited number of fragments. However, how this could have occurred is not easy to visualize, since all the targets were presumably equally likely to have been incorporated into the construct. It is more likely that the dynamics of the PCR reaction were such that only a limited number of fragments were amplified. There could be many reasons for this. Perhaps the size or conformation of the fragments was such that their amplification was favoured.

6.2.2) The determination of variation between mouse species

After the determination of suitable PCR conditions for the new markers, the PCR products were analyzed for variation between *Mus musculus* and *Mus spretus*. Of the methods employed, non-amplification of *Mus spretus* DNA was the most commonly used, in 4/11 cases. SSCV analysis was used in 1/11 case, radioactive PCR in 2/11 cases, RFLV analysis in 2/11 cases and length variation in 2/11 cases. The proportion of primer pairs not amplifying *Mus spretus* DNA was higher than expected: 36% rather than the expected 20%. The expected figure was derived from the primers of Love *et al.* (1991) and Hearne *et al.* (1992). The reason that the figure produced in this study is higher is unclear. It may be that the method of generating the markers (IRS-PCR or IRS-bubble PCR) produced fragments which were intrinsically less conserved than the microsatellite markers of Love *et al.* (1991) and Hearne *et al.* (1992) for some reason. It may be that the markers by chance happen to amplify regions of the genome which have insertions between *Mus musculus* and *Mus spretus*, although the probability of this happening in 36% of cases is fairly low. Alternatively, it could be a feature of the primer design, although the primers did have G+C contents varying from 7/20 to 12/20.

In the case of the observation of length differences, either detected by agarose gel electrophoresis or by polyacrylamide gel electrophoresis, in only three of the fragments could the difference be attributed to the presence of microsatellite repeats. However, since the entire sequence of the inserts was not determined, it is possible that other microsatellites existed internally to the regions sequenced. Of the microsatellites which amplified both *Mus musculus* and *Mus spretus* DNA, all but one were polymorphic. It is also possible that small insertions or deletions in the *Mus spretus* chromosome were present, which altered the size of the PCR product. Some of these alterations in size maybe also derived from the differing lengths of polypurine or polypyrimidine tracts near the extreme end of the poly A tail of the SINE or LINE repeat. Although not actually part of the repeat, these short tracts of mononucleotide repeats were fairly common, and were noted on many sequencing gels.

RFLVs were useful in two cases. The presence of the restriction sites is a function of the length of the PCR products, RFLVs were more common in large fragments. In cases where markers were derived from the extreme ends of the insert, and the products were large, RFLVs were easy to find. For smaller fragments, many enzymes had to be tried, before one was successful. Out of preference, restriction enzymes were chosen which cut the PCR products from both species (such as was the case for *D2Ucl15*), to eliminate the problem of false typings caused by lack of digestion. This was not always possible, where only one species was cut, anomalous typings were double checked (as was the case for *D2Ucl19*).

Where smaller fragments generated identical sized products, variation was determined by SSCV analysis. This proved a rapid and easy way to map the markers. The presence of heteroduplex bands in the case of *D2Ucl20* allowed a double check for the conformation variant.

Overall, 92% of new markers were variant between *Mus musculus* and *Mus spretus* by one of the methods outlined above. The remaining one marker was not variant under any conditions tried. This marker may represent an area that is conserved between *Mus musculus* and *Mus spretus*, perhaps a part of a gene. The results from these experiments indicate that although the presence of microsatellites is an advantage in the identification of variation between species, the need for screening programs to specifically clone them may not always be necessary for the generation of markers for mapping purposes. Of course, should markers prove interesting, problems may arise in cases where variation is present between *Mus*

musculus and *Mus spretus*, but not between *Mus musculus* and *Mus musculus castaneus*, as was the case for *D2Ucl15*, which could not be mapped on the ulnaless backcross, due to a lack of variation. An added problem could arise should markers mapped on the basis of non-amplification of *Mus spretus* DNA be needed to mapped on specialized backcrosses, in which the backcross parent is *Mus musculus* (such as the ragged and wasted backcrosses). The mapping of these markers under these conditions would necessitate the design of more primers, and a new search for variation.

6.2.3) The problems encountered during the cloning and mapping experiments

During the course of these experiments, several problems were highlighted, which hampered the progress. There were several stages at which problems were encountered. Firstly, during the amplification and cloning of the fragments, the high hamster background posed a considerable problem. This was counteracted by the use of dot-blotting techniques, to identify mouse clones prior to sequencing. Also a problem at this stage (even with different techniques) was the high level of plasmids consisting of vector alone which came through the blue:white screening as recombinant colonies. Problems also arose at this stage with the large number of identical clones produced from the IRS-bubble PCR, and to a much lesser extent, the L1 and B2 PCR reactions. Further problems were encountered with some instability of clones - two clones were impossible to maintain or sequence.

After the sequencing reactions, problems arose from the difficulty in defining suitable PCR conditions for the new primers designed. On some occasions, no amplification at all was achieved. This could be due to rearrangement of the plasmid DNA relative to the PCR target sequence, or undetected chimerism (i.e. - the presence of more than one insert) in the plasmid. In some cases, no amplification was seen for primer pairs when all these possibilities were ruled out. There appeared to be a rough correlation between clones which were difficult to sequence, and sequences which were difficult to amplify. In other cases, no specificity of PCR reaction could be achieved, amplification yielding ladders of bands which were not eliminated despite the use of all combinations of conditions, and several pairs of primers. This could be due to the inherent repetitive nature of the fragments produced, although the clones were screened for homology to repetitive elements prior to primer design. It could be that the fragments were only low level repetitive, which was not highlighted in the database searches, but that nevertheless caused problems in the PCR. This undetected repetitive nature could be

partially due to the higher degeneracy of both L1 and B2 repeats at the 3' end. All primers were designed from sequences after the AT rich tail.

Given the definition of adequate amplification conditions, problems arose with difficulty in identifying variation between *Mus musculus* and *Mus spretus*. This was not a large problem, only 1/12 fragments proved non variant. This marker, Bubble-14, was tested with 24 separate restriction enzymes, and different SSCV conditions. The marker was finally linked to mouse chromosome two by amplification of DNAs from the ABm series of microcell hybrids.

6.3) The mapping of new molecular markers for mouse chromosome two on the Jackson Laboratory interspecific backcrosses

New molecular markers derived from IRS-PCR and IRS-bubble PCR amplification of EBS18Az, ABm5, ABm 7 and ABm 11 hybrid DNAs were placed on the linkage map of mouse chromosome two by analysis of interspecific backcrosses.

6.3.1) Map positions of new molecular markers.

The first observation to be made was that despite the fact that EBS18Az (from which most of the new markers were derived) contained fragments of other chromosomes, all markers produced originated from mouse chromosome two. This is an unexpected finding, as a fragment of mouse chromosome 15 was also present in this hybrid. It would have been expected that at least one of the new markers produced from this hybrid would derive from mouse chromosome 15 (All 12 markers produced were in fact derived from this hybrid).

All new markers except one (for which no source of variation could be found) were placed on the map of mouse chromosome two by analysis of the Jackson Laboratory interspecific backcrosses panels 1 (BSB) and 2 (BSS). The distribution of the new loci on mouse chromosome 2 would appear to differ from the distribution of other markers derived from (CA)n microsatellites. No clustering of new markers was noted in the central portion of the chromosome, but some clustering of new markers was noted more distally, in the region around *Xmv10,* on panel 2. Although the number of markers produced is too small to say categorically that the distribution of the IRS-PCR and IRS-bubble PCR markers is different to that of microsatellite based markers, other workers in the laboratory have noted significant clustering of similar numbers of microsatellite based markers in the central portion of the chromosome (S. Malas, personal communication). This perhaps suggests that the use of IRS-PCR and IRS-bubble PCR to generate DNA fragments produces markers with a different distribution from that produced by the sole use of microsatellite based systems.

The estimated positions of the new markers on the cytogenetic map shows a rough compartmentalization of L1 based markers to Giemsa dark bands, and B2 based markers to Giemsa light bands, although there are exceptions to both cases. This to be expected, given that the distribution of SINES and LINES is different in the two types of chromosome bands. The presence of some L1 based markers mapping in Giemsa light bands, and some B2 based

markers mapping in Giemsa dark bands is also not totally unexpected, since the compartmentalization is not absolute.

6.3.2) Refinement of map position for interesting markers

Markers mapping to the area of mouse chromosome two containing ragged, wasted and ulnaless were judged as deserving of further interest. Markers mapping around lethal spotting would have also been analyzed further, if a suitable backcross had been available. One marker, *D2Ucl18*, mapped in the region of ragged and wasted. This marker was mapped relative to both mutants by analysis of either the ragged or wasted backcrosses. *D2Ucl18* was found to map 1.21 +/- 0.85cM proximal to ragged, and 5.4 +/- 2.3cM proximal to wasted, ruling it out as a closely linked flanking marker for the initiation of positional cloning projects. No transmission distortion was noted for the ragged or wasted backcrosses, but in the case of *Il1b* on the ragged backcross, the χ^2 test indicated that the excess of ragged alleles was only just not significant.

Three markers mapped in the region of ulnaless, *D2Ucl12*, *D2Ucl15* and *D2Ucl19*. No source of variation between *Mus musculus* and *Mus musculus castaneus* could be determined for *D2Ucl15*, but the remaining two markers were mapped relative to ulnaless on the ulnaless backcross. *D2Ucl12* was found to map 2.2 +/- 1.5cM proximal to ulnaless, whilst *D2Ucl19* mapped 5.4 +/- 2.3 cM distal to ulnaless. neither of these markers appeared to lie in the candidate region. *D2Ucl15* could be a better candidate, mapping between *D2Ucl12* and *D2Ucl19*, and may be worth further study in the future. No transmission distortion was noted for any loci on this cross.

D2Ucl18 could also be closely linked to lethal spotting. This mutant lies centrally within the *Gnas* - *D2Mit73* region on the consensus linkage map of mouse chromosome two. Analysis of the ragged and wasted backcross places *D2Ucl18* 1.21 +/- 0.85cM proximal to *D2Mit73*, and 1.21 +/- 0.85cM distal to *Gnas* (i.e.- roughly equidistant between these two markers). *D2Ucl18* could turn out to be a closely linked marker for lethal spotting, and thus may be worthy of future study. This marker could not be mapped relative to the mutant, because no suitable backcross was available.

.4) Overall progress of the project and future prospects

The primary aim of the project was to develop resources with which to initiate positional cloning projects to isolate the ragged, wasted and/or ulnaless mutations. This was to be accomplished by two complementary strategies: the development of two new panels of somatic cell hybrids, and the production of new molecular markers for mouse chromosome two, in order to increase the marker density of the genetic map. The results generated in this study by these strategies have not allowed a significant step forward in the progress towards the cloning of ragged, wasted or ulnaless, although some progress has been made in the increased marker density of mouse chromosome two.

The construction of the translocation hybrid panels would have provided a means of isolating genetically and physically a small area of distal mouse chromosome two, in which ragged and wasted were originally thought to lie. Ragged and wasted are very closely linked (Peters *et al.*, 1994). The development of these panels would have allowed the production of markers in a region specific manner by IRS-PCR, and also provided a method of easily identifying whether or not new markers deriving elsewhere were likely to be closely linked to ragged and wasted. In the light of the lack of segregation of either reciprocal translocation product, in either panel rendered this strategy unworkable.

The generation of new molecular markers yielded a total of twelve new markers for mouse chromosome two. Eleven of these have been placed on the genetic map. The remaining marker showed no variation, and so was not able to be mapped by interspecific backcross analysis. Instead, this marker was mapped to mouse chromosome two by amplification of the locus in the ABm hybrids, containing only mouse chromosome two. The new markers are distributed along the length of mouse chromosome two, and provide a small but significant increase in marker density on the genetic map. However, only one new marker, *D2Ucl18*, was derived from the putative region containing ragged, if not wasted. This is not unexpected, given that the starting material was an entire mouse chromosome two. Three markers were isolated from the *D2Mit11* - *D2Mit30* region, in which ulnaless lies.

6.4.1) Further work on the hybrid panels

With regard to the hybrids, the blind-mapping experiments indicate that the extent of fragmentation is much greater than detected by the PCR. The low segregation rate could be a reflection of heightened fragmentation, together with the fragments generated joining to the hamster chromosomes, which were not segregating. In order to define the extent of the fragmentation, and the fate of the fragments, detailed karyotyping or *in situ* hybridization should be carried out. *In situ* hybridization could be done using labelled IRS-PCR products to hybridize to normal mouse spreads. This might allow the easier identification of individual chromosome fragments, as the hybridization would be to an intact chromosome. This could be identified by its G-banding pattern in the hands of an experienced mouse cytogeneticist. Alternatively, spreads could be made of the hybrid cells, whih could then be probed with totla mouse DNA. This would allow an estimation of the number of fragments present in each hybrid line. The identification of the exact mouse chromosome content of each hybrid line would allow their use as a generalized mapping panel, provided the chromosomal stability in culture was adequate. In addition, detailed karyotypic and cytogenetic analysis of the translocation products present in T2Wa-2 and T2Wa-4 should shed light on their usefulness for the original purposes.

The apparent selective pressures acting against segregation of the translocation products is also worthy of further study. It is entirely possible that this finding is merely a characteristic of the panels generated in this study, since other panels involving the T1Sn and T13H translocations have segregated the translocation chromosomes, although some fragmentation was also apparent (A. Pilz, personal communication; this study). The reluctance of the cells to segregate the reciprocals could be investigated either by maintaining selected cell lines in long term culture (there is some evidence that after sufficient times in culture, cells become progressively more unstable (Weiss and Green, 1966), or by further forced segregation experiments, with a different antibody. By these means, the use of T2Wa-1 and T28H-3b as substrates for further subcloning or forced segregation could be more accurately evaluated.

Given more time and resources, an alternative strategy, in retrospect, would be to carry out afresh the construction of the hybrid panels. In these circumstances, since the conditions for optimal cell killing *in vitro* by the complement cascade have been established, further antibodies could be tested. Given the existence of an appropriate antibody, forced selection of specific chromosomes after fusion becomes a real possibility.

The antibody treatment could either be applied to the T(2;8)2Wa or T(2;16)28H hybrid population after the selective action of HMT, or synchronous with it. In order to obtain both reciprocals, an alternative strategy to complement mediated cell killing could be used. Should the antibody be conjugated to a magnetic bead, the cell population could be sorted using a magnetized cell sorter. Whether or not this experiment would produce segregation without fragmentation is not determined, but studies to determine the presence and nature of the selective pressures should elucidate this matter.

In order to produce more segregation in general, different hamster parents, or different mouse tissues could be used. If the high degree of retention of mouse chromosomes was due to the fragmentation, and subsequent fixation to the hamster cells, the use of a different mouse tissue (fibroblasts, liver cells, embryo cells) may act to stabilize the system. The low segregation rates for both panels could be due to the apparent stabilizing effect of V79TOR2. High rates of segregation must lead to some instability in the cell, so the use of a plastic cell line such as embryo cells, balanced by the stabilizing influence of V79TOR2 could feasibly act to increase segregation without too much loss of cell stability. This might in itself increase the suitability of the hybrids as generalized mapping panels, as well as increasing the probability that the translocation chromosomes themselves would segregate. The fragmentation could be due to some feature of the mouse parent. Little information is available to determine whether the use of translocation cell lines has any effect on stability, and whether or not this is related to the tissue origin of the cell line.

6.4.2) The generation of further molecular markers

During the course of these studies, some understanding of the problems and pitfalls associated with the cloning methods has been achieved. Given the development of means to identify the species of origin of clones, the high background of hamster clones generated by some of the methods can be circumvented. The isolation of individual bands has many advantages, not least that ability to detect the presence of putative CpG islands and microsatellites. The isolation of individual bands also increases the number and type of fragments generated from a single cloning experiment. However, the detection of variation between *Mus musculus* and *Mus spretus* in random fragments has not posed a problem. Whether the step of microsatellite detection is necessary is debatable. The use of PCR products without isolation of individual bands would allow the generation of a higher proportion of

mouse fragments, since the background is eliminated without the re-amplification step. The apparent predominance of certain fragments in the initial PCR reactions would have to be investigated further. Adjustment of the quantity of DNA in the amplifications, or the PCR conditions could increase the efficiency of the reactions. Similarly, the ratio of vector to insert, to allow higher concentrations of target DNA could be investigated. The dephosphorylation of insert rather than vector could allow higher concentrations of insert DNA without an increase in chimerism. Whether this would result in a higher proportion of vector colonies appearing as white under the selective conditions used is uncertain. The high background of apparent recombinant vector colonies was found to pose a considerable problem in the experiments in this study.

In order to generate a higher proportion of markers from distal mouse chromosome two, it would now be possible to construct of an irradiation fusion gene transfer hybrid panel, to reduce the amount of mouse chromosome two as starting material. Such a panel could be constructed from EBS18Az, or from ABm11, although the integrity of ABm 11 is debatable. This was not a viable strategy at the time the project was started, because of the paucity of markers with which to characterize such a panel. Now, the greatly increased marker density for mouse chromosome two (due to the efforts of the Whitehead Laboratory) would allow the accurate characterization of such a panel. Suitable markers for the characterization of such a panel would be *D2Mit74* and *D2Mit266* for the ragged - wasted region. There are at least thirty *D2Mit* markers in the ulnaless region which show variation between *Mus musculus* and *Mus musculus castaneus* (C. Abbott, personal communication), and many more which would be of use for hybrid studies. These include *D2Mit66, D2Mit128, D2Mit14 and D2Mit159* for the ulnaless region. The use *of D2Mit11, D2Mit159, D2Mit66, D2Mit41, D2Mit254, D2Mit190, D2Mit305, D2Mit209, D2Mit283, D2Mit169, D2Mit71, D2Mit52, D2Mit147* and *D2Mit266* as characterization markers would allow the identification of useful hybrid lines.

Should any of the proposed experiments to produce segregation of the T(2;8)2Wa and T(2;16)28H hybrid panels bear fruit, the generation of new molecular markers in a much more region-specific manner becomes feasible. Hybrids which have segregated the reciprocal translocations could be used as a preliminary method to identify fragments originating from the region between the translocation breakpoints. Chromosome two specific fragments (produced by the methods described earlier) could be used as probes for Southern blot analysis of DNA from hybrids showing segregation of the translocation products. This would necessitate the identification of single copy sequences within each fragment in some cases, as in these studies

IRS-PCR products were found to be too repetitive for use as probes. This strategy would indicate which clones were worthy of further analysis. This would reduce the workload to produce large numbers of markers from the region between the breakpoints. Perhaps more importantly, the use of IRS-PCR on these hybrids, and comparison of the band patterns produced would allow the generation of fragments specifically from this region. This would require a fairly low mouse chromosome content for success, but the complexity of the PCR product could be reduced by restricted *Alu* PCR (Guzzetta *et al.*, 1992) if necessary. This involves the digestion of the genomic DNA with frequent cutter restriction enzymes before the IRS-PCR amplification. The use of these methods would reduce the need to produce IFGT hybrid lines. The distal chromosomal location of ragged and wasted precludes the use of a set of recombinant interspecific backcross animals to position new markers with respect to the mutants (to date, there are no distal flanking markers to either gene). Markers which fit with the established pedigree analysis could be typed on the recombinant animals alone, but markers which show anomalous typings must be tested on the entire backcross. Hybrids containing this region alone, or segregated T2Wa and T28H hybrid panels would be pose a considerable advantage. A combination of these methods would improve the probability of producing markers close enough to ragged and wasted to allow the initiation of positional cloning experiments to isolate the genes involved. Markers showing linkage to the mutants described earlier would be localized precisely in relation to the mutant by use of the interspecific backcrosses described earlier.

6.4.3) Positional cloning strategies

In the event that flanking markers are generated which are sufficiently close to the mutants of interest to initiate positional cloning projects, strategies can be undertaken to isolate the genes involved. Many proximal flanking markers to ragged and wasted are available, such as *D2Mit266*, *D2Mit74*, and *Acra4*. Any of these markers could be used as a proximal starting point for chromosome walks. The telomeric location of the mutants also opens up the possibility of the use of a telomere specific primer together with either *D2Mit266* or *D2Mit74*, or a SINE or LINE based primer, to generate fragments from the band H4 - tel region. Such clones might possibly contain ragged and wasted. The first step in the positional cloning of ragged, wasted or ulnaless would be to isolate contiguous fragments of the region between the two closest flanking markers. This could be accomplished by the use of YAC vectors, which in fortuitous circumstances could include both marker loci and mutant genes. This would be a relatively rare event. Should the YAC not be of sufficient size to contain the mutant, the ends

of the YAC could then be isolated, in order to identify further YACs to extend the contig. This would be a form of chromosome walk, each step would be oriented and aligned by restriction mapping. Such a contig could also be constructed from phage or cosmid clones. The next step would be to identify coding sequences in the contig. This could be a laborious prospect, depending on the size of the contig. The load could be perhaps reduced by screening of cDNA libraries from different tissues and times in development, to identify clones expressed in the tissues affected by the mutation - hair follicles and skin in ragged; spleen, thymus and brain in wasted, an embryonic cDNA library in the case of ulnaless. Coding sequences could also be identified by the methods described in the introduction, such as exon trapping. Each piece of coding sequence could then be used to identify cDNAs from the affected tissues, or the appropriate time in development. The cDNAs could be compared with genomic sequence, to identify inrton:exon baoundarues, and control elements. If many genes genes are present in the region, this portion of the experiment can be very longwinded.

Once the entire sequence of each gene in the contig has been captured, sequence information can be generated from each one. Once sequence is available, it can be examined for homology with the known sequence of genes for which the function is known. This may identify possible functions for the cloned gene. Alternatively, predicted amino acid sequences from all reading frames could be generated, which might also shed light on function - for instance, the gene mutated in ragged could have characteristics of an ion channel, in light of the chylous ascites and oedema which occurs in homozygotes, and occasionally in heterozygotes.

Conclusive proof of the cloned gene as the site of mutation in ragged, wasted, lethal spotting or ulnaless would be the discovery of altered sequence in mutant mice compared to wild type, either in the coding sequence, or in the control elements. If such changes were present in both alleles of ragged (*Ra* and Ra^{op}), for instance, that would be a strong indication that the cloned gene was the gene mutated in ragged. This gene would presumably have effects on hair development, and also lymph drainage. Homologous genes from different species could be sequenced, with a view to determining which regions of the gene are highly conserved, and thus critical for function. Experiments could be undertaken to induce mutations into wild type genes, by homologous recombination, or to attempt to restore function by the use of a transgene. The mode of inheritance of the mutant could give clues as to the form of the mutation. Recesive mutations are often associated with loss of function mutations, where the levels of a particular gene product required for normal development fall below some threshold level (such as the albino (*c*) mutation - Kwon *et al.*, 1987) Dominant mutations often cause

altered function, altered expression or gain of function of a partcular gene product (such as the lethal yellow (A^y) allele in the agouti allelic series - Bultman *et al.*, 1994). Recessive mutations are sometimes caused by changes in enzyme structure or function, whereas dominant mutations can be associated with defects in structural components (Vogel, 1984), although there are many exceptions. On a molecular level, recessive mutations sometimes derive from small rearrangements, or point mutations, which either abolish transcription of the gene, or alter critical regions of the gene product. Dominant mutations have been associated with gross alterations such as deletions, splice site mutations or inversions, for example many alleles of the dominant spotting (*W*) locus are associated with deletions (Geissler *et al.*, 1988). Again, there are many exceptions. The dominant ulnaless mutation was radiation induced, and so probably constitutes a relatively large chromsomal rearrangement, such as a deletion, although this was not apparent on cytogenetic analysis (Davisson and Cattanach, 1990). Ulnaless could be a mutation in a pattern forming gene, such as the *Hox* or *Pax* family. Candidate genes in the region of ulnaless (on the basis of the consensus map) include the *Hoxd* gene cluster, *Exv2*, *Dlx2* and *Mdk* (from examination of the consensus map; Siracusa and Abbott, 1993). Lethal spotting (*ls*) has been shown to be due to a defect in the microenvironment rather than a cell intrinsic change (Kapur *et al.*, 1993). This raises the possibility that *ls* is the ligand for some undefined receptor, which influences the migration of enteric neurones. Wasted was a spontaneous (probably point) mutation which may have arisen in a gene involved in phosphorylation and development of neurofilaments, given the extensive crosslinking in the proximal portion of the cell body and axon noted by Lutsep and Rodriguez (1989). The sharp developmental profile of this mutant possibly implicates a gene which is inactive in early foetal and neonatal life. The degeneration of cells in the nervous system may reflect a mutation causing apoptosis (or programmed cell death). Cell adhesion molecules have recently been implicated in apoptosis and wasted may fall into this category. The semi-dominant ragged mutation was spontaneous. There are few clues to the nature of this alteration. The close proximity of *Acra4* to ragged makes this gene a possible candidate, although no gross alteration in the amount or size of *Acra4* mRNA could be detected by RT-PCR (data not shown). The phenotype of the ragged mutation possibly implicates a defective ion channel, or altered lipid metabolism, given the presence of lipids in the chylous fluid, and the reduced amounts of subcutaneous fat in these animals (Herbertson and Wallace, 1964).

Once genes responsible for mouse mutants have been identified, studies can be initiated to clone the human homologue. This may lead to the designation of new mouse models for

human disorders, and give clues to the properties of the genes likely to be involved in the disease.

References

Abbott, C., Skidmore, C., Searle, A. and Peters, J. Deficiency of adenosine deaminase in the wasted mouse. *P.N.A.S.* **83** p693 - 695 (1986).

Abbott, C., West, L., Povey, S., Jeremiah, S., Murad, Z., DiScipio R. and Fey G. The gene for human complement component C9 mapped to chromosome 5 by polymerase chain reaction.*Genomics* **4** p606 -609 (1989).

Abbott, C., Evans, E., Burtenshaw, M., Ball, S., Skidmore, C., Jones J. and Peters J. Adenosine deaminase, *Ada*, is in mouse chromosome 2H3, and is not allelic with wasted, *wst*. *Biochem. Genet.* **29** p537 - 544 (1991).

Abbott, C. Characterization of mouse-hamster somatic cell hybrids by PCR: A panel of mouse-specific primers for each chromosome.*Mammalian Genome* **2** p106 - 109 (1992).

Abbott, C. and Povey, S. Development of human chromosome specific PCR primers for characterization of somatic cell hybrids.*Genomics* **9** p73 - 77 (1991).

Abbott, C., Malas, S., Pilz, A., Pate, L., Ali, R. and Peters, J. Linkage mapping around the ragged (*Ra*) and wasted (*wst*) loci on distal mouse chromosome two. *Genomics* 20 p94 - 98 (1994).

Abbott, C. and Chambers, D. Analysis of CAG trinucleotide repeats from mouse cDNA sequences. *Ann Hum. Genet.* **58** p87 - 94 (1994).

Adams, R., Knowler J., and Leader D. In *The Biochemistry of the Nucleic Acids* (tenth edition) **p120-133.** Cambridge university press (1986).

Allen, N., Cran, D., Barton, S., Hettle, S., Reik, W. and Surani, M-A. Transgenes as probes for active chromosomal domains in mouse development. *Nature* **333** p852 - 855 (1988).

Allerdice, P., Miller, O., Pearson, P., Klein, G. and Harris, H. Human chromosomes in 18 man-mouse somatic hybrid cell lines analysed by quinacrine fluorescence. *J. Cell. Sci.* **12** p809 - 830 (1973).

Altschul, S., Gish, W., Miller, W., Myers, E. and Lipman, D. A basic local alignment search tool. *J. Mol. Biol.* **215** p403 - 410 (1990).

Ambros, P. and Sumner, A. Correlation of pachytene chromosomes and metaphase bands of human chromosomes, and distinctive properties of telomeric regions.*Cytogenet. Cell Genet.* **44** p223 - 228 (1987).

Ansell, J., Samuel, K., Whittingham, D., Patek, C., Hardy, K., Handyside, A,. Jones, K., Muggleton-Harris, A., Taylor, A. and Hooper, M. Hypoxanthine Phosphoribosyl transferase deficiency, haematopoiesis and fertility in the mouse. *Development* **112** p489 - 498 (1991).

Armour, J., Patel, I., Thein, S-L., Fey M. and Jeffreys A. Analysis of somatic mutations at human minisatellite loci in tumors and cell lines. *Genomics* **4** p328 - 334 (1989).

Atchley, W. and Fitch, W. Gene trees and the origins of inbred strains. *Science* **254** p554 - 558 (1991).

Avner, P., Amar, L., Arnaud, D., Hanauer, A. and Cambrou, J. Detailed ordering of markers localising to the Xq26 - Xqter region of the human X chromosome by the use of an interspecific *musculus* X *spretus* mouse cross. *P.N.A.S.* **84** p1628 - 1633 (1987).

Avner, P., Amar, L., Dandolo,L. and Guenet, J-L. Genetic analysis of the mouse using interspecific backcrosses.*Trends Genet.* **4** p18 - 23 (1988).

Bai, Y., Sheer, D., Hiorns, L., Knowles R. and Tunnacliffe, A. A monoclonal antibody recognising a cell surface antigen coded for by a gene on human chromosome 17. *Ann. Hum. Genet.* **46** p337 - 347 (1982).

Baldwin, C., Hoth, C., Amos, J., Da-Silva, E-O. and Milunsky, A. An exonic mutation in the HuP2 paired gene domain causes Waardenberg's syndrome. *Nature* **35** p637 - 639 (1992).

Balling, R.,. Deutsch, U., and Gruss, P. Undulated , a mutation affecting the development of the mouse skeleton has a mutation in the paired box of *Pax 1. Cell* **55** p531 - 535 (1981).

Balling, R., Lau, C., Dietrich, S., Wallin, J. and Gruss, P. Development of the skeletal system. *Ciba Foundation Symposium* **165** p132 - 140 (1992).

Barton, S., Surani, M. and Norris,M. Role of paternal and maternal genomes in mouse development. *Nature* **311** p374 - 376 (1984).

Beckmann, J. and Weber, J. Survey of human and rat microsatellites. *Genomics* **12** p 627 - 631 (1992).

Beechey, C. and Searle, A. - Radiosensitivity of wasted in male germ cells. *Mouse News Lett.* **70** p78 - 78 (1984).

Bender, W., Spierer, P. and Hogness, D. Chromosomal walking and jumping to isolate DNA from the *Ace* and *Rosy* loci and the bithorax in Drosophila melanogaster. *J. Mol. Biol.* **168** 17 - 33 (1983).

Benedetto, M., Anzai, Y. and Gordon, J. Isolation and analysis of the mouse genomic sequence encoding Cu^{2+}-Zn^{2+} superoxide dismutase. *Gene* **99** p191-195 (1991).

Benham, F., Hart, K., Crolla, J., Bobrow, M., Francavilla, M. and Goodfellow, P. A method for generating hybrids containing nonselected fragments of human chromosomes.*Genomics* **4** p509 - 517 (1989).

Benham, F. and Rowe, P. Use of *Alu*-PCR to characterize hybrids containing multiple fragments, and to generate new Xp21.3 - Xp22.2 markers.*Genomics* **12** p368 - 376 (1992).

Bennett, K., Hill, R., Pietras, D., Woodworth-Gustai, M., Kane-Haas, C., Houston, J., Heath, J. and Hastie, N. Most highly repeated dispersed DNA families in the mouse genome. *Mol. Cell. Biol.* **4** p1561 - 1571 (1984).

Bernardi, G., Olofsson, B., Filipski, J., Zenal, M., Salinas, J., Cuny, G., Meunier-Rotival, M. and Rodier, F. The mosaic genome of warm blooded vertebrate. *Science* **228** p953 - 958 (1985).

Bernardi, G. The isochore organization of the human genome. *Annu. Rev. Genet.* **23** p637 - 661 (1989).

Bickmore, W., Maule, J., Van-Heyningen, V. and Porteous, D. Long range structure of HRAS1 selected transgenomes. *Somat. Cell Mol. Genet.* **15** p229 - 235 (1989).

Bickmore, W. and Sumner, A. Mammalian chromosome banding - an expression of genome organization.*Trends Genet.* **5** p144 - 148 (1989).

Bicknell, D., Markie, D., Spurr, N. and Bodmer, W. The human chromosome content in human X rodent somatic cell hybrids analysed by a screening technique using *Alu*-PCR.*Genomics* **10** p186 - 192 (1991).

Bird, A. CpG-rich islands and the function of DNA methylation.*Nature* **321** p209 - 213 (1986)

Bird, A. CpG islands as gene markers in the vertebrate nucleus. *Trends Genet.* **3** p342 - 347 (1987).

Black, D. Does steric interference between splice sites block the splicing of a short *c-src* neuron-specific exon in non-neuronal cells? *Genes Dev.* **5** p389 - 402 (1991).

Bobrow,M and Cross, J. Differential staining of human and mouse chromosomes in interspecific cell hybrids. *Nature* **251** p77 - 79 (1974).

Bonhomme, F., Catalan, J., Britton-Davidian, J., Chapman, V., Moriwaki, K., Nevo, E. and Thaler, L. Biochemical diversity and evolution in the Genus *Mus*. *Biochem. Genet* **22** p275 - 303 (1984).

Bonhomme, F. and Guenet, J-L. The wild house mouse and it's relatives : In *Genetic variants and strains of the laboratory mouse* (2nd edition) Edited by Lyon and Searle **p649 - 662** (1989).

Boyle, A., Ballard, G. and Ward, D. Differential distribution of long and short interspersed element sequences in the mouse genome : Chromosome karyotyping by fluorescence *in situ* hybridization. *P.N.A.S.* **87**, p7757 - 7761 (1990).

Breen, M., Deakin, L., macdonald, B., Miller, S., Sibson, R., Tarttelin, E., Avner, P., Bourgade, F, Guenet, J.L., Montagutelli, X., Poirier, C., Simon, D., Tailor, D., Bishop, M., Kelly, M., Rysavy, F., Rastan, S., Norris, D., Shepherd, D., Abbott, C., Pilz, A., Hodge, S., Jackson, I., Boyd, Y., Blair, H., Maslen, G., Todd, J., Reed, P., Stoye, J., Ashworth, A., McCarthy, L., Cox, R., Shalkwyk, L., Lehrach, H., Klose, J., Gangadharan, U. and Brown, S. Towards high resolution maps of the mouse and human genomes - a facility for ordering markers to .0.1cM resolution. *Human Mol. Genet.* **3** p621-627 (1994).

Britten, R. and Kohne, D. Repeated sequences in DNA.*Science* **161** p529 - 540 (1968).

Britten, R. Evidence that most human *Alu* sequences were inserted in a process that ceased about 30 million years ago. *P.N.A.S.* **91** p 6148 - 6150 (1994).

Brooks-Wilson, A., Goodfellow, P., Povey, S., Nevanlinna, H., J. de Jong, P., Goodfellow, P. Rapid cloning and characterization of new chromosome 10 DNA markers by *Alu*-element mediated PCR. *Genomics* **7** p614 - 620 (1990).

Buck, D., Larrick, J., Raubitschek, A., Truitt, K., Senyr, G., Wang, J. and Dyer, B. Production of human monoclonal antibodies. In *Monoclonal antibodies and functional cell line: progress and applications* Ed. R. Kennett, K. Bechol and T. McKearn. pp275 - 309. Plenum press, New york/London (1984).

Bultman, S., Michaud, E. and Woychik, R. Molecular characterisation of the agouti locus.*Cell* **71** p1195 - 1204 (1992).

Bultman, S., Klebig, M., Michaud, E., Sweet, H., Davisson, M. and Woychk, R.. Molecular analysis of reverse mutations from nonagouti (*a*) to black-and-tan (a^t) and white-bellied agouti (A^w) reveals alternate forms of agouti transcripts. *Genes Dev.* **8** p481-490 (1994).

Burke, D., Carle, G. and Olson, M. Cloning of large exogenous DNA into yeast by means of yeast artificial chromosome vectors. *Science* **236** p806 - 811 (1987).

Burmeister, M., Monaco, A., Gillard, E., Van Ommenm, G-J., Affara, N., Ferguson-Smith, M., Kunkel, L. and Lehrach, H. A 10 megabase physical map of human Xp21 includung the Duchenne muscular dystrophy gene. *Genomics* **2** p189 - 202 (1988).

Cacheiro, N., Rutledge, J., Cain, K., Cornett, C. and Generoso, W. Cytogenetic analysis of malformed mouse fetuses derived from balanced translocation heterozygotes. *Cytogenet. Cell Genet.* **66** p139 - 148 (1994).

Carter, T. and Phillips, R. Ragged, a semi-dominant coat texture mutant. *J. Hered.* **45** p151 - 154 (1954).

Caspersson, T., Zech, L., Johansson,C. and Modest, E. Identification of human chromosomes by DNA binding fluorescent agent. *Chromosoma* **30** p215 - 227 (1970).

Cattanach, B. and Kirk, M. Differential activity of maternally and paternally derived chromosomal regions in mice. *Nature* **315** p496 - 498 (1985).

Cattanach, B. - Parental origin effects in mice. *J. Embryol. Exp. Morphol.* **97**(suppl.) p137 - 150 (1986).

Cattanach, B., Ball, S., Beechey, C., Blair, H., Boyd, Y., Burtenshaw, M., Evans, E., Fisher, G., George, A., Glenister, P., Holt, J., Jones, S., Laval, S., Maslan, G., Kirby, M., Lyon, M., Peters, J., Rasberry, C., Reed, R., Searle, A., Taylor, B., Tease, C. and Williamson, P. Further delimitation of the distal chromosome 2 imprinting region. *Mouse. News Lett.* p82 - 85 (1992).

Chakapalis, G., Fritsch, R., Fickenscher, H., Deutsch, U., Goulding, M. and Gruss, P. The molecular basis of the undulated / *Pax 1* mutation. *Cell* **66** p873 - 884 (1991).

Chalepakis, G., Stoykova, A., Wjinholds, J., Tremblay, P. and Gruss, P. Pax: Gene regulators in the developing nervous system. *J. Neurobiol.* **24** p1367 - 1384 (1993).

Chalepakis, G., Goulding, M., Read, A., Strachan, T. and Gruss, P. Molecular basis of splotch and Waardenburg *Pax3* mutations. *P.N.A.S.* **91** p 3685-3689 (1994).

Chan, T., Long, C. and Green, H. A human-mouse somatic cell hybrid line selected for human deoxycytidine deaminase. *Somat. Cell Genet.* **1** p81 - 90 (1975).

Chen, M., Shimada, T., Davis Moulton, A., Cline, A., Humphries, K., Maizel, J. and Nienhuis, A. The functional human dihydrofolate gene. *J. Biol. Chem.* **259** p3933-3943 (1984).

Chen, T. and Ruddle, F. Karyotype analysis utilising differentially stained constitutive heterochromatin of human and murine chromosomes. *Chromosoma* **34** p51 - 72. (1971).

Chu, E. and Powell, S. Selective systems in somatic cell genetics. *Adv. Hum. Genet.* **7** p189 - 259 (1976).

Church, D., Stotler, C., Rutter, J., Murrell, J., Trofatter, J. and Butler, A. Isolation of genes from complex sources of mammalian genomic DNA using exon amplification. *Nature Genetics* **6** p98 - 105 (1994).

Cirullo, D., Arredondo-Vega, F., Smith, M. and Wasmuth, J. Isolation and characterization of interspecific heat-resistant hybrids between a temperature sensitive Chinese hamster cell asparaginyl-tRNA synthetase mutant and normal human leucocytes : assignment of human asnS gene to chromosome 18. *Somat.Cell Genet.* **9** p215 - 233 (1983).

Cole, C., Goodfellow, P., Bobrow, M. and Bentley, D. Generation of novel sequence tagged sites (STSs) from discrete chromosomal regions using *Alu*-PCR. *Genomics* **10** p816 - 826 (1991).

Collins, F. and Weissman, S. Directional cloning of DNA fragments at a large distance from an initial probe : a circularization method. *P.N.A.S.* **81** p6812 - 6816 (1984).

Comings, D. Mechanisms for chromosome banding and implications for chromosome structure. *Ann. Rev. Genet.* **12** p25 - 46 (1978).

Copeland, N. and Jenkins, N. Development and applications of a molecular linkage map of the mouse genome. *Trends Genet.* **7** p113 - 118 (1991).

Cornall, R., Aitman, T., Hearne, G. and Todd, J. The generation of a library of PCR analysed microsatellite variants for genetic analysis of the mouse genome. *Genomics* **10** p874 - 881 (1991).

Cotter, F., Hampton, G., Nasipuri, S., Bodmer, W. and Young, B. Rapid isolation of human chromosome specific DNA probes from a somatic cell hybrid. *Genomics* **7** p257 - 263 (1990).

Coulson, A., Waterson, R., Kiff, J., Sulston, J. and Kohara, Y. Genome linking with yeast artificial chromosomes. *Nature* **335** p184 - 186 (1988).

Cox, D., Burmeister, M., Price, E., Kim, S. and Myers, R. Radiation hybrid mapping : A somatic cell genetic method for constructing high resolution maps of mammalian chromosomes *Science* **250**, p245 - 250 (1990).

Cox, R., Copeland, N., Jenkins, N. and Lehrach, H. Interspersed repetitive element polymerase chain reaction product mapping using a mouse interspecific backcross. *Genomics* **10** p375 - 384 (1991).

Craig, J. and Bickmore, W. Chromosome bands - flavours to savour. *Bioessays* **15** p349 - 354 (1993).

Creagan,R. and Ruddle, F. The clone panel: A systematic approach to gene mapping using interspecific somatic cell hybrids. *Cytogenet. Cell Genet.* **14** p282 - 286 (1975).

Crew, F. and Koller, P.-The sex incidence of chiasma frequency and genetical crossing-over in the mouse. *J. Genet.* **26** p359 - 383 (1932).

Croce, C., Koprowski, H. and Eagle, H. Effect of environmental pH on the efficiency of cellular hybridization. *P.N.A.S.* **69** p1953 - 1956 (1972).

Croce C., Knowles, B. and Koprowski, H. Preferential retention of human chromosome C-7 in human (thymidine kinase deficient) mouse hybrid cells. *Exp. Cell Res.* **82** p457 - 461. (1973).

Cuny, G., Soriano, P., Macaya, G. and Bernardi, G. The major components of the mouse and human genomes. *Eur. J. Biochem.* **115** p227 - 233 1981).

Cupp, M. Notes on identity of mutants. *Mouse News Lett.* **26** p51 - 51 (1962).

Darling, S. and Abbott, C. Mouse models of human single gene disorders I - non-transgenic mice. *Bioessays* **14** p359 - 366 (1992).

Dautigney, A., Mattei, M-G., Morello, D., Alliel, P., Pham-Dinh, D., Amar, L., Arnaud, S., Simon, D., Mattei, J., Guenet, J-L., Jolles, P. and Avner, P. The structural gene coding for myelin-asociate proteolipi proein is mutated in *jimpy* mice. *Nature* **321** p867 - 869 (1986).

Davidson, R. and Ephrussi, B.- Factors influencing the "effective mating rate" of mammalian cells. *Exp. Cell. Res.* **61** p222 - 226 (1970).

Davidson, R. and Ephrussi, B. A selective system for the isolation of hybrids between L cells and normal cells. *Nature* **205** p1170 - 1171 (1965).

Davisson, M. and Cattanach, B. The mouse mutation ulnaless on mouse chromosome 2. *J. Hered.* **82** p151 - 153 (1990).

Deininger, P. and Daniels, G. The recent evolution of mammalian repetitive elements. *Trends Genet.* **2** p76 - 80 (1986).

Dietrich, W., Katz, H., Lincoln, S., Shin, H-S., Friedman, J., Dracopoli, N. and Lander, E. A genetic map of the mouse suitable for typing interspecific backcrosses. *Genetics* **2** p423 - 447 (1992).

De la Chapelle, A., Schroder, J., Kajsa-Selander, R. and Stentrand, K. Differences in DNA composition along mammalian metaphase chromosomes. *Chromosoma* **42** p365 - 382 (1973).

Dover, G. Molecular drive - a cohesive mode of species evolution. *Nature* **299** p111 - 117 (1982).

Dunn, L. and Bennett, D. Sex differences in recombination of linked genes in animals. *Genet. Res.* **9** p211 - 220 (1967).

Duyk, G., Kim, S., Myers, R. and Cox, D. Exon trapping : A genetic screen to identify candidate transcribed sequences in cloned mammalian genomic DNA. *P.N.A.S.* **87** p8995 - 8999 (1990).

Edgell, M. Tracing transposable elements. *Nature Genetics* 7 p 120 - 121 (1994).

Edwards, A., Civitello, A., Hammond, H. and Caskey, C-T. DNA typings and genetic mapping with trimeric and tetrameric repeats. *Am. J. Hum. Genet.* **49** p746 - 756 (1991).

Ephrussi, B., Scarletta, L., Stenchever, M. and Yoshida, M. Hybridization of somatic cells *in vitro* .*Symp. Intern. Soc. Biol.* **3** p13 - 25 (1964).

Ephrussi, B. and Weiss, M. Interspecific hybridization of somatic cells. *P.N.A.S.* **53** p1040 - 1042 (1965)

Epstein, D., Vekemans, M. and Gros, P. Splotch (SP^{2H}), a mutation affecting development of the mouse neural tube shows a deletion within the paired homeobox domain of *Pax3*. *Cell* **67** p767 - 774 (1991a).

Epstein, D., Malo, D., Vekemans, M. and Gros, P. Molecular characterization of a deletion encompassing the splotch mutation on mouse chromosome 1. *Genomics* **10** p89 - 93 (1991b).

Epstein, D., Vogan, K., Trasler, D. and Gros, P. A mutation within exon 3 of the *Pax3* gene produces abberently spliced mRNA transcripts in the splotch mutant mouse. *P.N.A.S.* **90** p532 - 536 (1993).

Evans, E. Standard karyotype of the mouse. In *Genetic variants and strains of the laboratory mouse* (2nd edition) Eds. Lyon and Searle, Oxford university press, p**576 - 581** (1989).

Evans, J. and Palmiter, R. Retrotransposition of a mouse L1 element.*P.N.A.S.* **88** p8792 - 8795 (1991).

Ferris, S., Sage, R. and Wilson, A. Evidence from mtDNA sequences that common laboratory strains of mice are descended from a single female. *Nature* **295** p163 - 165 (1982).

Festing, M. *Inbred Strains in biomedical research*, Macmillan Press **p62 - 141** (1979).

Finger, S., Heavens, R., Sirinathsinghji, D., Kuehn, M. and Dunnett, S. Behavioral and neurochemical evaluation of a transgenic mouse model of Lesch-Nyhan syndrome. *J. Neurol. Sci.* **86** p203 - 215 (1988).

Fournier, R.E.K, Juricek, D. and Ruddle, F. Somatic cell genetic nalysis of transgenome integration. *Somatic Cell Genet.* **5** p1061 - 1077 (1979).

Fournier, R.E.K. A general high efficiency procedure for production of microcell hybrids. *P.N.A.S.* **78** p6349 - 6353 (1981).

Fournier, R.E.K. and Moran, R. Complementation mapping in microcell hybrids: Localization of *FPGS* and *AK-1* on *Mus musculus* chromosome 2. *Somatic Cell Genet.* **82**, p69 - 84 (1982).

Fournier, R.E.K. and Ruddle, F. Micorcell mediated transfer of murine chromosomes into mouse, Chinese hamster and human somatic cells. *P.N.A.S.* **74** p319 - 323 (1977).

Foy, C., Newton, V., Wellesly, D., Harris, R. and Read, A. Assignment of the gene for Waardenberg syndrome type I to human chromosome 2q37 and possible homology to the splotch mouse. *Am. J. Hum. Genet.* **46** p1017 - 1023 (1990).

Francke, U., Lalley, P., Moss, W., Ivy, J. and Minna, J. Gene mapping in *Mus musculus* by interspecific cell hybridization: assignment of the genes for tripeptidase-1 to chromosome 10, dipeptidase-2 to chromosome 18, acid phosphatase-1 to chromosome 12, and adenylate kinase to chromosome 2. *Cytogenet. Cell Cenet.* **19** p57 - 84 (1977).

Frohman, M. and Martin, G. Cut, Paste and Save : New approaches to altering specific genes in mice. *Cell* **56** p145 - 147 (1989).

Gaillard, C. and Strauss, F. Association of poly (CA). poly (GT) DNA fragments into four-stranded complexes bound by HMG1 and 2. *Science* **264** p 433 - 436 (1994).

Gale, G., Ostrander, W. and Atkins, L. Effects of alanosine on purine and pyrimidine synthesis. *Biochem. Pharmacol.* **17** p1823 - 1832 (1968).

Gebhard, W., Meitinger, T., Hochtl, J. and Zachau, H. A new family of interspersed repetitive DNA sequences in the mouse genome. *J. Mol. Biol.* **157** p453 - 471 (1982).

Geiger, J. and Nagy, J. Lack of adenosine deaminase deficiency in the mutant mouse wasted. *Febs Letts.* **208** p431 - 434 (1986).

Geissler, E. and Ryan, M. The dominant white spotting (W) locus of the mouse encodes the c-kit protooncogene. *Cell* **55** p185 - 192 (1988).

Ghosh, S., Palmer, S., Rodrigues, N., Cordell, H., Hearne, C., Cornall, R., Prins, J., McShane, P., Lathrop, G., Peterson, L., Wicker, L. and Todd, J. polygenic control of autoimmune diabetes in nonobese diabetic mice. *Nature Genetics* **4** p404 - 409 (1993).

Giles, R. and Ruddle, F. Production and characterization of proliferating somatic cell hybrids. In *Tissue culture; methods and applications* page **475 - 501** (1973)

Glaser, T., Walton, S. and Maas, R. Genomic structure, evolutionary conservation and aniridia mutations in the human *PAX3* gene. *Nature Genetics* **2** p232 - 239 (1992).

Goldman, M., Holmquist, G., Gray, M., Caston, L. and Nag, A. Replication timing of genes and middle repetitive sequences. *Science* **224** p686 - 692 (1984).

Goodfellow, P., Jones, E., Van Heyningen, V., Soloman, E., Bobrow, M., Miggiano, V. and Bodmer, W. The β2 microglobulin gene is on chromosome 15 and not in the HL-A region. *Nature* **254** p267 - 267 (1975).

Goodfellow, P. and Pritchard, C. Chromosome fragmentation by chromosome mediated gene transfer. *Cancer Surv.* 7 p251 - 265 (1988).

Goss, S. and Harris, H. New method for mapping genes in human chromosomes. *Nature* **255** p680 - 684 (1975).

Graham, F. and Van Der Eb, A. A new technique for the assay of infectivity of human adenovirus 5 DNA. *Virology* **52** p456 - 467 (1973).

Gray, M., Colot, H., Guarante, L. and Robach, M. Open reading frame cloning : identification, cloning and expression of open reading frame DNA. *P.N.A.S.* **79** p6598 - 6602 (1989).

Grey, J., Dean, P., Fuscoe, J., Peters, D., Trask,B., Van Der Engh, G. and Van Dilla, M. High-speed chromosome sorting. *Science* **238**, p323 - 329 (1987).

Green, M. Linkage data. *Mouse News Lett.* **26** p34 - 34 (1962).

Green, M., Demant, P., Egorov, I., Gruneberg, H., Hutton, J., Kondo, K., Lyon, M., Roderick, T., Sabourdy, M., Schmidt, R., Searle, A. and Staats, J. Standard karyotype of the mouse *Mus musculus* : Committee on standard genetic nomenclature for mice. *J. Hered.* **62** p69 - 72 (1972).

Green - In *methods of animal experiments Vol VI.* Edited by Gay, Academic Press (1981).

Green, E. and Mann, S. Opossum - a semi-dominant lethal mutation. *J. Hered.* **52** p223 - 227 (1961).

Gross, D. and Garrard, W. The ubiquitous potential Z-DNA forming sequence of eukaryotes, (dT-dG).$(dC\text{-}dA)_n$ is not detectable in the genomes of eubacteria, archebacteria or mitochondria. *Mol. Cell Biol.* **6** p3010 - 3013 (1986).

Guzetta, V., Montes De Oca-Luna, R., Lupski, J. and Patel, P. Isolation of region specific and polymorphic markers from chromosome 17 by restricted *Alu* polymerase chain reaction. *Genomics* **9** p31 - 36 (1991).

Hale, D., Washburn, L. and Eicher, E. Meiotic abnormalities in hybrid mice of the C57BL/6J x *Mus spretus* cross suggest a cytogenetic basis for Haldane's rule of hybrid sterility. *Cytogenet Cell genet.* **63** p221 - 234 (1993).

Hamada, H. and Kakanunga, T. Potential Z-DNA forming sequences are highly dispersed in the human genome. *Nature* **298** p396 - 398 (1982).

Hamada, H., Petrino, M. and Kakanunga, T. A novel repeat element with Z-DNA forming potential is widely found in evolutionarily diverse eukaryotic genomes. *P.N.A.S.* **79** p6465 - 6469 (1982).

Hamada, H., Seidman, M., Howard, B. and Gorman, C. Enhanced gene expression by the poly (dT-dG).poly (dC-dA) sequence. *Mol. Cell Biol.* **4** p2622 - 2630 (1984).

Handmaker, S. Cytogenetic analysis of a Chinese hamster-mouse hybrid cell. *Nature* **233** p416 - 419 (1971).

Harris, H., Watkins, J., Ford, C. and Schoefl, G. Artificial heterokaryons of animal cells from different species. *J. Cell Sci.* **1** p1- 30 (1966).

Harris, H. Protein poymorphism in man. *Can. J. Genet. Cytol.* **13** p381 - 396 (1971).

Hastie, N. Highly repeated DNA families in the genome of *Mus musculus* : in *Genetic variants and strains of the laboratory mouse* (2nd edition). Edited by Lyon and Searle. Oxford university press **p559 - 573** (1989).

Hayashi, K. PCR-SSCP: A simple and sensitive method for detection of mutations in the genomic DNA. *PCR Methods. Appl.* **1** p 34 - 38 (1992).

Hearne, C., McAleer, M., Love, J., Aitman, T., Cornall, R., Ghosh, S., Knight, A., Bas Prins, J., Todd, J. Additional microsatellite markers for mouse genome mapping. *Mammalian Genome* **1** p273-282 (1991).

Heller, D. and Arnheim, N. Structure and organization of the highly repeated and interspersed 1.3 Kb *EcoR1 - Bgl-II* sequence family in mice. *N.A.R.* **8** p5031 - 5042 (1980).

Herbertson, B. and Wallace, M. Chylous ascites in newborn mice. *J. Med. Genet.* **1** p10 - 23. (1964).

Herman, G., Berry, M., Munro, E., Craig, I. and Levy, E. The construction of human somatic cell hybrids containing portions of the mouse X chromosome and their use to generate DNA probes via interspersed repetitive sequence polymerase chain reaction. *Genomics* **10** p961 - 970 (1991).

Herman, G., Nadeau, J. and Hardies, S. Dispersed repetitive elements in mouse genome analysis. *Mammalian Genome* **2** p207 - 214 (1992).

Hill, R., Favor, J., Hogan, B., Ton, C., Saunders, G., Hanson, I., Prosser, J., Jordan, T., Hastie, N. and Van-Heyningen, V. Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354** p522 - 525 (1991).

Hilwig, I. and Gropp, A. Staining of constitutive heterochromatin in mammalian chromosomes with a new fluorochrome. *Exp Cell Res* **75** p122 - 126. (1972).

Hochgeshwender, U., Sutcliffe, J. and Brennan, M. Construction and screening of a genomic library specific for mouse chromosome 16. *P.N.A.S* **86**, p8482 -8486 (1989).

Holmes, S., Dombroski, B., Krebs, C., Boehm, C. and Kazazian, H. A new retrotransposable human L1 element from the *LRE2* locus on chromosome 1q poduces a chimeric insertion. *Nature Genetics* **7** p143 - 148 (1994).

Holmquist, G. Chromosome bands, their chromatin flavors, and their functional features. *Am. J. Hum. Genet.* **51** p17 - 37 (1992).

Hosaka, Y. and Koshi, Y. Electron microscope study of cell fusion by HVJ virions. *Virology* **34** p419 - 434 (1968).

Hug, H., Costas, M., Staeheli, P., Aebi, M. and Weissman, C. organization of the murine *Mx* gene and characterization of its interferon- and virus-inducible promotor. *Mol. Cell Biol.* **8** p 3065-3079 (1988).

Hunter, K., Watson, M., Rochelle, J., Ontiveros, S., Munroe, D., Seldin, M. and Housman, D. Single strand conformation polymorphism (SSCP) mapping of the mouse genome: Integration of the SSCP, microsatellite and gnen maps of mouse chromosome 1. *Genomics* **18** p510 - 519 (1993).

Inoue, T., Aikawa, K., Tezuza, H., Kada, T. and Schultz, L. Effect of DNA damaging agents on isolated spleen cells and lung fibroblasts from the mouse mutant 'wasted', a putative animal model for ataxia telangiectasia. *Cancer Res.* **46** p3979 - 3982 (1985).

Irving, N. and Brown, S. Mouse chromosome specific markers generated by PCR and their mapping through interspecific backcross. *Genomics* **11** p679 - 686 (1991).

Jackson, I. Mouse coat colour mutations : A molecular genetic resource that spans the centuries. *Bioessays* **13** p439 - 446 (1991).

Jagadeeswaran, P., Forget, B. and Weissman, S. Short interspersed repetitive DNA elements in eukaryotes - transposable DNA elements generated by reverse transcription of RNA pol III transcripts. *Cell* **26** p141 - 142 (1981).

Jeffreys, A., Wilson, V. and Thein, S-L. Hypervariable 'minisatellite' regions in human DNA. *Nature* **314** p67 - 73 (1985).

Jeffreys, A., Wilson, V., Kelly, R., Taylor, B. and Bulfield, G. Mouse DNA 'fingerprint' analysis of chromosomal location and germ line stability of hypervariable loci in recombinant inbred strains. *N.A.R.* **15** p2823 - 2836 (1987).

Jeffreys, A., Royle, N., Wilson, V. and Wong, Z. Spontaneous mutation rates to new length alleles at tandem repetitive hypervariable loci in human DNA. *Nature* **332** p278 - 281 (1988).

Jeffreys, A., Tamaki, K., Macleod, A., Monckton, D., Neil, D. and Armour, J. Complex gene conversion events in germline mutation at human minisatellite loci. *Nature Genetics* **6** p136 - 145 (1994).

Jones, C., Wuthier, P., Kao, F-T. and Puck, T. Genetics of somatic mammalian cells. XV. Evidence for linkage between human genes for lactase dehydrogenase B and serine hydroxymethylase. *J. Cell. Physiol.* **80** p291 - 297 (1972).

Jones, C., Patterson, D. and Kao, F-T. Assignment of the gene coding for Phosphoribosylglycineamide Formyltransferase to human chromosome 14. *Somatic Cell Genet.* 7 p399 - 409 (1981).

Julier, C., Gouyon, B., Georges, M., Guernet,, J-L., Nakamura, Y., Avner, P. and Lathrop, G. Minisatellite linkage maps in the mouse by cross hybridization with human probes containing tandem repeats. *P.N.A.S.* **87** p4585 - 4589 (1990).

Kaiserlian, D., Delacroix, D. and Bach, J. The wasted mutant mouse I. An animal model of secretory IgA deficiency with normal serum IgA. *J. Immunol.*. **135** p1126 - 1132 (1985).

Kaiserlian, D., Savino, W., Uriel, J., Hassid, J., Dardenne, M. and Bach, J. The wasted mutant mouse II. Immunological abnormalities in a mouse described as a model for ataxia telangiectasia. *Clin. Exp. Immunol.* **63** p562 - 569 (1986).

Kao, F-T. and Puck, T. Genetics of somatic mammalian cells : Linkage studies with human - Chinese hamster cell hybrids. *Nature* **228** p329 - 332 (1970).

Kapur, R., Yost, C. and Palmiter,R. Aggregation chimeras demonstrate that the primary defect responsible for aganglionic megacolon in lethal spotted mice is not neuroblast autonomous. *Development* **117** p993 - 999 (1993).

Kazazian, H., , Wong, C., Youssoutian, H., Scott, A., Phillips, D. and Antonarakis, S. Haemophilia A resulting from a *de novo* insertion of L1 sequences represents a novel mechanism of mutation in man. *Nature* **332** p164 - 166 (1988).

Kelly, R., Bulfield, G., Collick, A., Gibbs, M. and Jeffreys, A. Characterization of a highly unstable mouse minisatellite locus. Evidence for somatic mutation during early development. *Genomics* **5** p844 - 856 (1989).

Kingsley, D., Rinchik, E., Russell, L., Ottiger, H., Sutcliffe, J., Copeland, N., and Jenkins, N. Genetic ablation of a mouse gene expressed specifically in brain. *EMBO J.* **9** p395 - 399 (1990).

Kingsley, D., Bland, A., Grubber, J., Marker, P., Russell, L., Copeland, N., and Jenkins, N. The mouse short ear skeletal morphogenesis locus is asociated with defects in a bone morphogenetic member of the TGFβ superfamily. *Cell* **71** p399-410 (1992).

Kingsmore, S., Giros, B., Suh, D., Bieniarz, M., Caron, M. and Seldin, M. Glyine receptor β subunit gene mutation in spastic mouse associated with LINE-1 element insertion. *Nature Genetics* **7** p136 - 142 (1994).

Klebe, R., Chen, T. and Ruddle, F. Controlled production of proliferating somatic cell hybrids. *J. Cell. Biol.* **45** p74 - 62 (1970).

Kleijer, W., Hoeksema, J., Sluyter, M. and Bootsma, D. Effects of inhibitors of repair of DNA on normal human and xeroderma pigmentosum cells after exposure to X-rays and ultraviolet radiation. *Mut. Res.***17** p385 - 394 (1973).

Kline, A., Rojas, K., Mewar, R., Moshinsky, D. and Overhauser, J. Somatic cell hybrid deletion map of human chromosome 18. *Genomics* **13** p1 - 6 (1992).

Klobutcher, L. and Ruddle, f. Phenotype stabilization and integration of transferred material in chromosome mediate gene transfer. *Nature* **280** p657 - 660 (1979).

Korenberg, L. and Rykowski, M. Human genome organization: *Alu*, LINES, and the molecular structure of metaphase chromosome bands. *Cell* **53** p391 - 400 (1988)

Kozak, C., Nichols, E. and Ruddle, f. Gene linkage analysis in the mouse by somatic cell hybridization : Assignment of Adenine phosphoribosyl transferase to chromosome 8 and α-galactosidase to the X chromosome. *Somatic Cell Genet.* **1** p371 - 382 (1975).

Kramerov, D., Lekakh, I., Samarina, O. and Ryskov, A. The sequences homologous to major interspersed repeats B1 and B2 of mouse genome are present in mRNA and small cytoplasmic poly A^+ RNA. *N.A.R.* **10** p7477 - 7491 (1982).

Kreyev, A., Markushevk, T., Kramerov, D., Ryskov, A., Skryabin, K., Bayev, A. and Georgiev, G. Ubiquitous transposon-like repeats B1 and B2 of the mouse genome, B2 sequencing. *N.A.R.* **10** p7461 - 7475 (1982).

Kuhl, D. and Caskey, C-T. Trinucleotide repeats and genome variation. *Genetics and Development* **3** p404 - 406 (1993).

Kusano, T., Long, C. and Green, H. A new reduced human-mouse somatic cell hybrid containing the human gene for adenine phosphoribosyltransferase. *P.N.A.S* **68** p82 - 86 (1971).

Kwon, B., Haq, A., Pomerantz, S. and Hlaban, R. Isolation and sequence of a cDNA clone for human tyrosinase that maps at the mouse c- albino locus. *P.N.A.S.* **84** p7473 - 7477 (1987).

Lane, P. Association of megacolon with two recessive spotting genes in the mouse. *J. Hered.* **57** p29 - 31 (1966).

Larsen F., Gundersen, G., Lopez, R. and Prydz, H. CpG islands as gene markers in the human genome.*Genomics* **13** p1095 - 1107. (1992).

Lawrie, M. and Hulten, M. Sex differences in hotspots of meiotic recombination in mice : In the report on the first international workshop on chromosome 9. *Am. J. Hum. Genet.* **56** p167 - 167 (1992).

Ledbetter, S., Nelson, D., Warren, S. and Ledbetter, D. Rapid isolation of DNA probes within specific chromosome regions by interspersed repetitive element polymerase chain reaction. *Genomics* p475 - 481 (1990).

Leuders, K. and Paterson, B. A short interspersed repetitive element near some mouse structural genes. *N.A.R.* **10** p7715 - 7729 (1982).

Levenberg, B., Melnick, I. and Buchanan, J. Biosynthesis of the purines XV: The effects of aza-serine and 6-diazo-5-oxo-l-norleucine on inosinic acid biosynthesis *de novo*. *J. Biol. chem.* **225** p163 - 176 (1957).

Levinson, G. and Gutman, G. High frequencies of short frameshifts in poly CA/TG tandem repeats borne by bacteriophage M13 in Escherichia coli K12. *N.A.R.* **15** p5323 - 5338 (1987).

Lewis, W., Srinivasan, P., Stokoe, N. and Siminovitch, L. Parameters governing the transfer of genes for thymidine kinase and dihydrofolate reductase into mouse cells using metaphase chromosomes or DNA. *Somatic Cell Genet.* 6 p333 - 347 (1980).

Lichter, P., Ledbetter, S., Ledbetter, D. and Ward, D. Fluorescence in situ hybridization with *Alu* and L1 polymerase chain reaction probes for rapid characterization of human chromosomes in hybrid cell lines. *P.N.A.S* **87** p6634 - 6638 (1990).

Linzer, D. and Mordacq, J. Transcriptional regulation of proliferin gene expression in response to serum in transfected mouse cells. *Embo J.* **6** p2281-2288 (1987).

Litt, M. and Luty, J. A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *Am. J. Hum. Genet.* **44** p397 - 401 (1989).

Littlefield, J. Selection of hybrids from matings of fibroblasts in vitro, and their presumed recombinants. *Science* **145** p709 - 710 (1974).

Loeb, D., Padgett, R., Hardies, S., Shehee, W-R., Comer, M., Edgell, M. and Hutchison, C. The sequence of a large L1md element reveals a tandemly repeated 5' end and several features found in retrotransposons. *Mol. Cell Biol.* **6** p168 - 182 (1986).

Love, J., Knight, A., McAleer, M. and Todd, J. Towards construction of a high resolution map of the mouse genome using PCR-analysed microsatellites. *N.A.R.* **18** p 4123-4130 (1990).

Lutsep, H. and Rodriguez, M. Ultrastructural, morphometric and immunocytochemical study of anterior horn cells in mice with 'wasted' mutation. *J. Neuropath. Exp. Neurol.* **48** p519 - 533 (1989).

Lyon, M. Distribution of crossing-over in mouse chromosomes. *Genet. Res.* **28** p291 - 299 (1976).

Lyon, M., Ward, H. and Simpson, G.A genetic method for measuring non-disjunction in mice with Robertsonian translocations. *Genet. Res.* **26** p283 - 295 (1976).

Mann, S. The phenogenetics of hair mutants in the mouse : Opossum and Ragged. *Genet. Res.* **4** p1 - 11 (1963).

Manuelidis, L. and Ward, D. Chromosomal and nuclear distribution of the *Hind III* 1.9 kb human DNA repeat segment. *Chromosoma* **91** p28 - 38 (1984).

Marmur, Rownd and Schildkraut - *N.A.R.* **1**/231 (1963).

McBride, O. and Ozer, H. Transfer of genetic information by purified metaphase chromosomes. *P.N.A.S.* **70** p1258 - 1262 (1973).

McNeill, C. and Brown, R. Genetic manipulation by means of microcell mediated transfer of normal human chromosomes into recipient mouse cells. *P.N.A.S.* **77** p5394 - 5398 (1980).

Medrano, L. and Green, H. A uridine kinase-deficient mutant of 3T3 and a selective method for cells containing the enzyme. *Cell* **1** p23 - 26 (1974).

Meese, E., Meltzer, P., Ferguson, P. and Trent, J. *Alu*-PCR: Characterization of a chromosome 6 specific hybrid mapping panel and cloning of chromosome specific markers. *Genomics* **12** p549 - 554 (1992).

Meunier-Rotival, M., Soriano, P., Cuny, G., Strauss, F. and Bernardi, G. Sequence organization and genomic distribution of the major family of interspersed repeats of mouse DNA. *P.N.A.S.* **79** p355 - 359 (1982).

Michaud, E., Bultman, S., Stubbs, L. and Woychik, R. The embryonic lethality of homozygous lethal yellow mice (A^y/A^y) is asociated with the disruption of a novel RNA-binding protein. *Genes. Dev.* **7** p1203-1213 (1993).

Michie, D. Genetical studies with vestigial tail mice I : The sex difference in crossing-over between vestigial-tail and rex. *J. Genet.* **53** p270 - 279 (1955).

Midgeon, B. and Miller, C. Human-mouse somatic cell hybrids with single human chromosome (group E): link with thymidine kinase activity. *Science* **162** p1005 - 1007 (1968).

Miesfield, R., Krystal, M. and Arnheim, N. A member of a new repeated sequence family which is conserved throughout eukaryotic evolution is found between the human δ and β globin genes. *N.A.R.* **9** p5931 - 5947 (1981).

Miller, M., Duhl., D., Cordes, S., Ollmann, M., Winkes, B. and Barsh, G. Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice containing the lethal yellow mutation. *Genes Dev.* 7 p 454-467 (1993).

Minna, J., Gazdar, A., Iverson,G.M., Marshall, T., Stromberg,K. and Wilson, S. Oncornavirus expression in human x mouse hybrid cells segregating mouse chromosomes. *P.N.A.S.* **71** p1695 - 1700 (1974).

Minna, J., Marshall, T. and Schaffer-Berman, P. Chinese hamster X mouse hybrid cells segregating mouse chromosomes and isozymes. *Somatic cell genet.* **1** p355 - 369 (1975).

Minna, J. and Coon H. Human x mouse hybrid cells segregating mouse chromosomes and isozymes. *Nature* **252** p401 - 404 (1974).

Monaco, A., Neve, R., Colletti-Feener, C., Bertelson, C., Kurnit, D. and Kunkel, L. Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* **323** p646 - 650 (1986).

Moriwaki, K. Gene mapping based on Inter-subspecific (*domesticus/molossinus*) backcross : In the abstracts from the *Seventh International Mouse Genome Conference* **p3 -3**(1993).

Morris, T. *Mouse News Lett.* **36** p34 - 34 (1967).

Moser, F., Dorman, B. and Ruddle, F. Mouse-human heterokaryon analysis with a 33258 Hoechst Giemsa technique. *J. Cell Biol.* **66** p676 - -680 (1975).

Mouchiroud, D., D'Onofrio, G., Aissani, B., Macaya, G., Gautier, C. and Bernardi, G. The distribution of genes in the human genome. *Gene* **100** p181 - 187. (1991).

Moyzis, R., Buckingham, J., Cram, L-S., Dani, M., Deavan, L., Jones, M., Meyne, J., Ratliffe, R. and Wu, J-R. A highly conserved repetitive DNA sequence, $(TTAGGG)_n$ present at the telomeres of human chromosomes. *P.N.A.S.* **85** p6622 - 6626 (1989).

Moyzis, R., Torney, D., Meyne, J., Buckingham, J., Wu, J-R., Burks, C., Sirotkin, K. and Goad, W. The distribution of interspersed repetitive sequences in the human genome. *Genomics* **4** p273 - 289 (1989).

Munroe, D., Haas, M., Bric, E., Whitton, T., Aburatani, H., Hunter, K., Ward, D. and Housman, D. IRE-bubble PCR: A rapid method for efficient and representative amplification of human genomic DNA sequences from complex sources. *Genomics* **19** p506 - 514 (1994).

Nabholz, M., Miggiano, V. and Bodmer, W. Genetic analysis with human-mouse somatic cell hybrids. *Nature* **223** p358 - 363 (1969).

Nakamura, Y., Leppert, M., O' Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E. and White, P. - Variable number of tandem repeats (VNTR) markers for human gene mapping : *Science* **235** p1616 - 1622 (1987).

Nanda, I., Deubelbeiss, C., Guttenbach, M., Epplen, J. and Schmidt, M. Heterogeneities in the distribution of $(GACA)_n$ simple sequence repeats in the karyotypes of primates and mouse. *Hum. Genet.* **85** p187 - 194 (1990).

Nasir, J., Fisher, E., Brockdorff, N., Disteche, C., Lyon, M. and Brown, S. Unusual molecular characteristics of a repeat sequence island located within a Giemsa-positive band on the mouse X chromosome. *P.N.A.S.* **87** p399 - 403 (1990).

Nasir. J., Maconochie, M. and Brown, S. Co-amplification of L1 line elements with localised low copy repeats in Giemsa dark bands: Implications for genome organization. *N.A.R.* **19** p3255 - 3260 (1991).

Nelson, D., Ledbetter, S., Corbo, L., Victoria, M., Ramirez-Soliz, R., Webster, T., Ledbetter, D. and Caskey, C-T. *Alu*-polymerase chain reaction: A method for rapid isolation of human specific sequences from complex DNA sources. *P.N.A.S.* **86** p6686 - 6690 (1989).

Nesbitt, M. and Francke, U. A system of nomenclature for band patterns of mouse chromosomes. *Chromosoma* **41** p145 - 158. (1973).

Nordheim, A. and Rich, A. The sequence (dC-dA).$(dG\text{-}dT)_n$ forms negatively supercoiled plasmids. *P.N.A.S.* **80** p1821 - 1825 (1983).

Norum, R. and Migeon, B. Non-random loss of human markers from man-mouse somatic cell hybrids. *Nature* **251** p72 - 74 (1974).

Nottenberg, C., Rees, G. and John, T. Isolation of mouse Cd44 cDNA: structural features are distinct from the primate cDNA. *P.N.A.S* **86** p8521-8525 (1989).

O'Brien, R. and Graves, J. Report of the committee on comparative gene mapping. *Cytogenet. Cell. Genet.* **58** p1124 - 1151 (1991).

O'Brien, R., Womack, J., Lyons, L., Moore, K., Jenkins, N. and Copeland, N. Anchored reference loci for comparative mapping in mammals. *Nature Genet.* **3** p103 - 112 (1993).

Oda, M. and Puck, T. The interraction of mammalian cells with antibodies. *J. Exp. Med.* **113** p599 - 613 (1961).

Okada, Y and Tadokoro, J. The distribution of cell fusion capability among several cell strains or cells caused by HVJ. *Exp. Cell. Res.* **32** p417 - 430 (1963).

Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. *Genomics* **5** p 874 - 879 (1989a).

Orita, M., Iwahana, H., Kanazawa, H., Hayashi, K. and Sekiya, T. Detection of polymorphisms of human DNA by gel electrophoresis as single stranded conformation polymorphisms. *P.N.A.S* **86** p 2766 - 2770 (1989b).

Peters, J., Beechey, C., Ball, S. and Evans, E. Mapping studies of the distal imprinting region of mouse chromosome two. *Genet. Res.* **63** p 169 - 174 (1994).

Pietras, D., Bennett, K., Siracusa, L., Woodworth-Gutai, M., Chapman, V., Gross, P., Kane-Haas, C. and Hastie, N. Construction of a small *Mus musculus* repetitive DNA library - identification of a new satellite sequence in *Mus musculus*. *N.A.R.* **11** p6965 - 6983 (1980).

Pilz, A., Willer, E., Povey, S., and Abbott, C. The genes coing for phosphoenolpyruvate carboxykinase-1 (PCK-1) and neuronal nicotinic acid acetylcholine receptor α4 subunit (CHRNA 4) map to human chromosome 20 extending the known region of homology with mouse chromosome 2. *Ann. Hum. Genet.* **56** p289 - 293 (1992).

Pious, D., Hawley, P. and Forrest, G. Isolation and characterization of HLA varients in cultured human lymphoid cells. *P.N.A.S* **70** p1397 - 1400. (1973).

Portecorvo, G. Induction of directional chromosome elimination in somatic cell hybrids. *Nature* **230** p367 - 369 (1971).

Portecorvo, G. Production of mammalian somatic cell hybrids by means of polyethylene glycol treatment. *Somatic Cell Genet.* **4** p397 - 400 (1975).

Porteous, D., Morten, J., Cranston, G., Fletcher, J., Mitchell, A., Van-Heyningen, V., Fantes, J., Boyd, Y. and Hastie, N. Molecular and physical rearrangements of human DNA in HRAS1 selected chromosome mediated transfectants : *Mol. Cell Biol.* **6** p2223 - 2232 (1986).

Porteous, D. Chromosome mediated gene transfer : a functional assay for complex loci and an aid to human gene mapping. *Trends Genet.* **3** p177 - 181 (1987).

Poste, G. Mechanisms of virus-induced cell fusion. *Int. Rev. Cytol.* **33** p157 - 252 (1972).

Prescott, D., Myerson, D. and Wallace, j. Enucleation of mammalian cells with cytochalasin B. *Exp. Cell Res.* **71** p480 - 485 (1977).

Pritchard, C. and Goodfellow, P. Investigation of chromosome mediated gene transfer using the *HPRT* region of the human X chromosome as a model : *Genes. Dev.* **1** p172 - 178 (1987).

Puck, T. Wuthier, P., Jones, C. and Kao, F. Genetics of somatic mammalian cells: Lethal antigens as genetic markers for study of human linkage groups : *P.N.A.S.* **68** p3102-3106 (1971).

Puck, T. and Kao,F. - Genetics of somatic mammalian cells.V. Treatment with 5-bromodeoxyuridine and visible light for isolation of nutritionally deficient mutants : *P.N.A.S* **58** p1227 - 1234 (1967).

Puck, T. and Kao, F-T. Somatic cell genetics and its applications to medicine. *Ann. Rev. Genet.* **16** p225 - 271 (1982).

Reeves, R., Crowley, M., Moseley, W. and Seldin, M. Comparison of interspecific to intersubspecific backcrosses demonstrates species and sex specific differences in recombination frequency on mouse chromosome 16. *Mammalian Genome* **1** p158-164 (1991).

Reid, D. and Parsons, P. Sex of parent and variation of recombination with age in the mouse. *Heredity* **18** p107 - 108 (1963).

Reik, W., Collick, A., Norris, M., Barton, S. and Surani, M-A. Genomic imprinting determines methylation of parental alleles in transgenic mice. *Nature* **328** p248 - 251 (1987(.

Richards, R. and Sutherland, G. Simple sequence repeat DNA is not replicated simply. *Nature Genetics* **6** p114 - 116(1994).

Rickles, R., Darrow, A. and Strickland, S. Molecular cloning of complementary DNA to mouse tissue plasminogen activator mRNA and its expression during F9 teratocarcinoma call differentiation. *J. Biol. Chem.* **263** p1563-1569 (1988).

Rikke, B. and Hardies, S. LINE-1 repetitive DNA probes for species specific cloning from *Mus spretus* and *Mus domesticus* genomes. *Genomics* **11** p895 - 904 (1991).

Riley, J., Butler, R., Ogilvie, D., Finniear, R., Jenner, D., Powell, S., Anand, R., Smith, J. and Markham, A. A novel, rapid method for the isolation of terminal sequences from yeast artificial chromosome (YAC) clones.*N.A.R.* **18** p2887 - 2890 (1990).

Roderick, T., Ruddle, F., Chapman, V. and Shows, T. Biochemial polymorphisms in feral and inbred mice (*Mus musculus*). *Biochem. Genet.* **5** p457 - 466. (1971).

Rogers, J. - Origin and evolution of retroposons : *Int. Rev. Cytol.* **93** p-187 - 279 (1985).

Rothman, T. and Gershon, M. Regionally defective colonization of the terminal bowel by precursors of enteric neurones in lethal spotted mice. *Neurosci* **12** p1293 - 1131 (1984).

Rowe, L., Nadeau, J., Turner, R., Frankel, W., Letts, V., Eppig, J., Ko, M., Thurston, S, and Birkenmeier, E. Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. *Mammalian genome* **5** p253 - 274 (1994).

Royle, N., Clarkson, R., Wong, Z. and Jeffreys, A. Clustering of hypervariable minisatellites in proterminal regions of human autosomes. *Genomics* **3** p352 - 360 (1988).

Ruddle, F.- Linkage studies employing mouse-man somatic cell hybrids. *Fed. Proc.* **30** p921 - 925 (1971).

Ruddle, F. Linkage analyis using somatic cell hybrids. *Adv. Hum. Genet.* **3** p173 - 235 (1972).

Ruddle, F. and Creagan, R. Parasexual approaches to the genetics of man. *Ann. Rev. Genet.* 9 p407 - 487 (1975).

Ruoslahti, E. and Reed, J. Anchorage dependence, integrins and apoptosis. *Cell* **77** p477 - 478 (1994).

Sanford, J., Forrester, L., Chapman, V., Chandley and Hastie, N. Methylation patterns of repetitive DNA sequences in the germ cells of *Mus musculus* . *N.A.R.* **12** p2823 - 2836 (1984).

Santos, J., Cole, Y. and Pellicer, A. Phylogenetic relationships amongst laboratory and wild origin *Mus musculus* strains on the basis of genomic DNA RFLPs. *Mammalian Genome* **4** p485 - 492 (1993).

Saxon, P., Srivatsan, E., Leipzig, V., Sameshima, J. and Stanbridge, E. Selective transfer of individual human chromosomes to recipient cells. *Mol. Cell. Biol.* **5** p140 - 146 (1985).

Scarletta, L., Rushforth, N. and Ephrussi, B. Isolation of hybrids between somatic mouse and Chinese hamster cells. *Genetics* **57** p107 - 122 (1967).

Schaffer, J., Sutton, R., Bewley, G. Isolation of a cDNA clone for murine catalase and analysis of an acatalasemic mutant. *J. Biol. Chem.* **262** p12908 - 12911 (1987).

Schlotterer, C. and Tautz, D. Slippage synthesis of simple sequence DNA. *N.A.R.* 20 p211-215 (1992).

Schmidt, C. and Jelinek, W. The *Alu* family of dispersed repetitive sequences. *Science* **216** p1065 - 1070 (1982).

Schnedl, W. The karyotype of the mouse. *Chromosoma* **35** p111 - 116 (1971).

Schultz, L., Sweet, H., Davisson, M. and Coman, D. 'Wasted' a new mutant of the mouse with abnormalities characteristic of ataxia telangiectasia. *Nature* **297** p402 - 404 (1982).

Schwartz, D. and Cantor, C. Separation of yeast chromosomal sized DNA by pulsed field gradient gel electrophoresis. *Cell* **37** p67 - 75 (1984).

Scott, A., Schmeckpeper, B., Abdelrazik, M., Comey, C., O'Hara, B., Rossiter, J., Cooley, T., Heath, P., Smith, K. and Margolet, L. Origin of the human L1 elements: proposed progenitor genes deduced from a consensus DNA sequence. *Genomics* **1** p112 - 125 (1987).

Searle, A., Beechey, C., Eicher, E., Nesbitt, M. and Washburn, L. Colinearity in the mouse genome: A study of mouse chromosome . *Cytogenet. Cell Genet.* **23** p255 - 263 (1979).

Seldin, M., Howard, T. and Eustachio, P. Comparison of linkage maps of mouse chromosome 12 derived from laboratory strain intraspecific and *Mus spretus* interspecific backcrosses. *Genomics* **5** p24 - 28 (1989).

Sharp, P. Conversion of RNA to DNA in mammalian *Alu*-like elements and pseudogenes. *Nature* **301** p471 - 472 (1981).

Shaw, P. and Schibler, U. Structure and expression of the parotid secretory protein gene of the mouse. *J. Mol. Biol.* **192** p567-576 (1986).

Sheffield, V., Beck, J., Kwitek, A., Sandstrom, D. and Stone, E. The sensitivity of single-stranded conformation polymorphism analysis for the detection of single base substitutions. *Genomics* **16** p 325 - 332 (1993).

Siden, T., Hoglund, M. and Rohme, D. Construction of a microcell hybrid panel containing different *Neo* gene insertions in mouse chromosome 17 used for chromosome mediated gene transfer. *Somat. Cell Mol. Genet.* **15** p245 - 253 (1989).

Silverman, G., Ye, R., Pollock, K., Sadler, J.E. and Korsmeyer, S. Use of yeast artificial chromosome clones for mapping and walking within human chromosome segment 18q21.3. *P.N.A.S.* **86** p7485 - 7489 (1989).

Simmler, M-C., Cox, R. and Avner, P. Adaptation of the interspersed repetitive sequence polymerase chain reaction to the isolation of mouse DNA probes from somatic cell hybrids on a hamster background. *Genomics* **10** p770 - 778 (1991).

Singer, M. and Skowronski, J. Making sense of LINES - long interspersed repetitive elements in mammalian genomes - *Trends Biochem.* **10** p119 - 122 (1985).

Siracusa, L., Buchberg, A., Copeland, N. and Jenkins, N. Recombinant inbred strain and interspecific bckcross analysis of molecular markers flanking the murine agouti coat colour locus. *Genetics* **122** p669 - 679 (1991).

Siracusa, L., Alvord, W., Bickmore, W., Jenkins, N. and Copeland, N. Interspecific backcross mice show sex-specific differences in allelic inheritance. *Genetics* **128** p813 - 821 (1991).

Siracusa, L. and Abbott, C. Mouse chromosome 2. *Mammalian Genome* **4** ps31 - s46 (1993).

Skow, L., Donner, M., Huang, S-M., Gardner, J., Taylor, B., Beamer, W. and Lalley, P. Mapping of the mouse gamma crystallin genes on mouse chromosome 1. *Biochem. Genet.* **26** p557 - 570 (1988).

Smith, G. Evolution of repeated DNA sequences by unequal crossover. *Science* **191** p528-535 (1976).

Snell, G. An analysis of translocations in the mouse. *Genetics* **31** p157 - 180 (1946).

Soloman, E., Bobrow, M., Goodfellow, P., Bodmer, W., Swallow, D., Povey, S. and Noel, B. Human gene mapping using an X/autosome translocation. *Somatic. Cell Genet.* **2** p125 - 140 (1976).

Solter, D. Differential imprinting and expression of maternal and paternal genomes. *Ann. Rev. Genet.***22** p127 - 146 (1988).

Sommsich, I., Hameister, H. and Winking, H. The pattern of early replicating bands in the chromosomes of the mouse . *Cytogenet. Cell Genet.* **30** p222 - 231 (1981).

Soriano, P., Meunier-Rotival, M. and Bernardi, G. The distribution of interspersed repeats is non-uniform and conserved in the mouse and human genomes. *P.N.A.S.* **80** p1816 - 1820 (1983).

Spies, T., Blancke, G., Bresnahan, M., Sands, J. and Strominger, J. A new cluster of genes within the human major histocompatability complex. *Science* **243** p214 - 217 (1989).

Spinardi, L., Mazars, R. and Theillet, C. Protocols for an improved detection of point mutations by SSCP. *N.A.R* **19** p4009 - 4009 (1991).

Staeheli, P. and Sutcliffe, G. Identification of a second interferon-regulated murine *Mx* gene. *Mol. Cell Biol.* **8** p4524-4528 (1988).

Stallings, R., Ford, A., Nelson, D., Tourney, D., Hildebrand, C. and Moyzis, R. Evolution and distribution of $(GT)_n$ repetitive sequences in mammalian genomes. *Genomics* **10** p807 - 815 (1991).

Stallings, R. Distribution of trinucleotide repeats in different categories of mammalian genomic sequence: Implications for human genetic dieases. *Genomics* **21** p 116 - 121 (1994).

Stephan, S., and Cho, S. Possible role of natural selection in the formation of tandem repetitive non-coding DNA. *Genetics* **136** p-333-341 (1994).

Strickberger, M. In *Genetics* (Third edition) Macmillan publishing company, New York, pp**126 - 146** and pp**243 - 250** (1985).

Stubblefield, E. DNA synthesis and chromosomalmorphology of Chinese hamster cells cultured in media containing N-deacytyl-N-methylcolchicing (colchemid). In *Cytogenetics of cells in culture*, Ed. Harris (Academic Press, New York) Vol 3 **p223 - 248** (1964).

Sumner, A., Evans, H. and Buckland, R. New technique for distinguishing between human chromosomes. *Nature New Biol.* **232** p31 - 32 (1971).

Sutton, W. and McCallum, M. Related satellite DNAs in the genus *Mus*. *J. Mol. Biol.* **71** p633 - 656 (1972).

Szybalska, E. and Szybalski, W. Genetics of human cell lines IV. DNA-mediated heritable transformation of a biochemical trait. *P.N.A.S.* **48** p2026 - 2034 (1962).

Tachida, H., and Iizuka, M. Persistance of repeated sequences that evolve by replication slippage. *Genetics* **131** p 471-478 (1992).

Tassabehji, M., Read, A., Newton, V., Patton, M., Gruss, P., Harris, R. and Strachan, T. Mutations in the *PAX3* paired box gene causing Waardenberg's syndrome type I and II. *Nature. Genet.* **3** p26 - 30 (1993).

Tautz, D. Hypervariability of simple sequences as a general source of polymorphic DNA markers. *N.A.R.* **17** p6463 - 6471 (1989).

Tautz, D. Genomic fingerprints made simple. *Bioessays* **12** p44 - 46 (1990).

Tautz, D. and Renz, M. Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *N.A.R.* **12** p4127 - 4138 (1984).

Taylor, B. and Meyer, H. Mapping the adrenal lipid depletion gene of the AKR/J strain. *Genetic Res.* **26** p307 - 321 (1976).

Timsit, Y., Vilbois, E. and Moras, D. Base-pairing shift in the major groove of $(CA)_n$ tracts by B-DNA crystal structures. *Nature* **354** p167 - 170 (1991).

Thacker, J. Resistance to methyl mercaptopurine in cultured hamster cells. Preliminary characterization of resistant cells and conditions affecting their selection in quantitative mutation studies. *Mut. Res.* **74** p37 - 59 (1980).

Thompson,L., Mankovitz, R., Baker, R., Till, J., Siminovitch, L. and Whitmore, G. Selective and non-selective isolation of temperature sensitive mutants of L cells. *P.N.A.S.* **66** p377 - 384 (1970).

Treco, D. and Arnheim, N. The evolutionarily conserved repetitive sequence d(TG.AC)$_n$ promotes reciprocal exchange and generates unusual recombinant tetrads during yeast meiosis. *Mol. Cell Biol.* **6** p3935 - 3947 (1986).

Ullu, E., Murphy, S. and Melli, M. Human 7SL RNA consists of a 140 nucleotide middle repetitive sequence inserted in an *Alu* sequence. *Cell* **29** p195 - 202 (1982).

Ullu, E. and Tschudi, C. *Alu* sequences are processed 7SL RNA genes. *Nature* **312** p171 - 172 (1984).

Van Arsdell, S., Denison, R., Bernstein, L., Weiner, A., Manser, T. and Gesteland, R. Direct repeats flank the small nuclear RNA pseudogenes in the human genome. *Cell* **26** p11 - 17 (1981).

Vogel, F. Gene or point mutations. In *Mutations in man.* Ed. G. Obe, Springer-verlag Berlin Heidelberg pp101 - 127 (1984).

Vogt, T., Jackson-Grunsby, L., Rush, J. and Leder, P. Formins - phosphoprotein isoforms encoded by the mouse limb deformity locus. *P.N.A.S.* **90** p5554 - 5558 (1993).

Vorlickova, M., Kypr, J., Stokrova, S.and Sponar, J. A Z-like form of poly (dA-dC).poly (dG-dT) in solution? *N.A.R.* **10** p1071 - 1080 (1982).

Wallace, M. Analysis of genetic control of chylous ascites in ragged mice. *Heredity* **43** p9 - 18 (1979).

Walter, M., Spillett, D., Thomas, P., Weissenbach, J. and Goodfellow, P. A method for constructing radiation hybrid maps of whole genomes. *Nature Genetics* 7 p22-28 (1994).

Walther, C., Guenet, J-L., Simon, D., Deutsch, U., Jostes, B., Goulding, M., Plachov, B., Balling, R. and Gruss, P. Pax - a murine multigene family of paired-box containing genes. *Genomics* **11** p424 - 434 (1991).

Wang, Y., Amirhaeri, S., Kang, S., Wells, R. and Grifith, J. Preferential nucleosome assembly at DNA triplet repeats from the myotonic dystrophy gene. *Science* **265** p669 - 671 (1994).

Weber, J. and May, P. Abundant class of DNA polymorphisms which can be typed using the polymerase chain reaction. *Am. J. Hum. Genet.* **44** p388 - 396 (1989).

Weiss, M., Ephrussi, B. and Scarletta, l. Loss of T antigen from somatic hybrids between mouse cells and SV40-transformed human cells. *P.N.A.S* **59** p1132 - 1135 (1968).

Weiss, M. and Ephrussi, B. Studies of interspecific (rat x mouse) somatic hybrids. 1) Isolation, growth and evolution of the karyotype. *Genetics* **54** p1095 - 1109 (1966a).

Weiss, M. and Ephrussi, B. Studies of interpspecific (rat x mouse) somatic hybrids. II) Lactase dehydrogenase and β–glucuronidase. *Genetics* **54** p1111 - 1122 (1966b).

Weiss, M. and Green, H. Human-mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes. *P.N.A.S.* **58** p1104 - 1111 (1967).

Weller, P., Jeffreys, A., Wilson, V., and Blanchetot, A. Organization of the human myoglobin gene. *EMBO J.* **3** p439 - 446 (1984).

White, R., Peters, L., Adkinson, L., Korsgren, C., Cohen, C. and Lux, S. The murine pallid mutation is a platelet storage pool disease associated with the protein 4.2 (pallidin) gene. *Nature Genetics* **2** p 80 - 83 (1992).

Wilke, K., Yung M., Chen, Y. and Geldermann, H. Porcine (GT)n sequences: structure and association with dispersed and tandem repeats. *Genomics* **21** p 63 -70 (1994).

Wilson, D., Woodard, D., Sandler, A., Erickson,, J. and Gurney, A. Provisional assignment of the gene for uridine monophosphatase (Umph-2) to mouse chromosome 11. *Bioch. Genet.* **25** p1 - 6 (1987).

Winkes, B., Ollmann, M. and Barsh, G. Association of *Xmv10* and the non-agouti (*a*) muatation explained by close linkage rather than causality. *Mammalian Genome* **5** p 3 - 10 (1994).

Worton, R., Duff, C. and Flintoff, W. Microcell mediated cotransfer of genes specify methotrexate resistance, emetine sensitivity, and chromate resistance with Chinese hamster chromosome 1. *Mol. Cell Biol.* 1 p330 - 335 (1981).

Woychik, R., Stewart, T., Davis, L., Eustachio, P. and Leder, P. An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. *Nature* **318** p36 - 40 (1985).

Woychik, R., Genfroso, W., Russell, L., Cain, K., Cacheiro, N., Bultman, S., Selby, P., Dickinson, M., Hogan, B., and Rutledge, J. Molecular and genetic characterization of a radiation-induced structural rearrangement in mouse chromosome 2 causing mutations at the limb deformity and agouti loci. *P.N.A.S.* **87** p2588 - 2592 (1990).

Wyman, A. and White, R. A highly polymorphic locus in human DNA. *P.N.A.S.* **77** p6754 - 6757 (1980).

Yoshiura, K., Kubota, T., Soejima, H., Tamura, Izumikawa, I., Niikawa, N. and Jinno, J. A comparison of GC content and the proportion of *Alu/Kpn1* repetitive sequences in a single dark- and light- band region from a human chromosome. *Genomics* **20** p243 - 248 (1994).

Yunis, J., Roldan, L., Yasmineh, W. and Lee, J. Staining of satellite DNA in metaphase chromosomes. *Nature* **231** p532 - 533 (1971).

Yunis, J. Mid-prophase human chromosomes. The attainment of 2000 bands : *Hum. Genet.* **56** p293 - 298 (1981).

Zsebo, C., Williams, D., Geisler, E., Broudy, V., Martin, F., Atkins, H., Hsin, S., Birkett, N., Ouino, K., Murdock, D., Jacobson, F., Langley, K., Smith, K., Takeishi, T., Cattanach, B.,

Galli, K. and Suggs, S. Stem cell factor is encoded at the steel locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor. *Cell* **63** p213 - 224 (1990).

Appendix

Animal	*D2Ucl 11*	*D2Ucl 12*	*D2Ucl 13*	*D2Ucl 14*	*D2Ucl 15*	*D2Ucl 16*	*D2Ucl 17*	*D2Ucl 18*	*D2Ucl 19*	*D2Ucl 20*	*D2Ucl 31*
1A	B	B	B	B	B	B	B	B	B	B	B
1B	S	S	S	S	S	S	S	S	S	S	S
1C	B	B	B	B	B	B	B	S	B	B	B
1D	S	S	S	S	S	S	S	B	S	S	S
1E	S	S	S	S	S	S	S	S	S	S	S
1F	B	S	B	B	S	B	B	S	B	B	B
1G	S	S	S	S	S	S	S	S	S	S	S
1H	B	B	S	S	B	B	S	S	B	S	S
2A	S	S	B	B	S	S	B	S	B	B	B
2B	S	S	B	B	B	S	B	S	B	B	B
2C	S	S	S	S	S	S	S	S	B	S	S
2D	B	S	S	S	B	S	S	B	S	S	S
2E	S	S	S	S	S	S	S	S	S	S	S
2F	B	B	B	B	B	B	B	B	B	B	B
2G	S	S	B	B	S	S	B	S	B	B	B
2H	B	B	S	B	S	B	B	S	S	B	B
3A	B	B	S	S	B	B	S	S	S	S	S
3B	S	S	B	B	S	S	B	S	B	B	B
3C	S	S	S	S	S	S	S	S	S	S	S
3D	B	S	B	B	S	B	B	S	B	B	B
3E	B	B	B	B	B	B	B	S	B	B	S
3F	S	S	B	B	S	S	B	S	B	B	B
3G	B	B	B	B	B	B	B	S	B	B	B
3H	B	S	S	S	S	B	S	S	S	S	S
4A	B	S	S	S	S	B	S	B	B	S	S
4B	S	B	S	S	B	S	S	B	S	S	S
4C	S	S	S	S	S	S	S	S	S	S	S
4D	B	B	S	S	S	B	S	S	S	S	S
4E	S	S	B	B	S	S	B	S	B	B	B
4F	B	B	S	S	B	B	S	S	B	S	S
4G	B	B	S	S	B	B	S	B	S	S	S
4H	S	S	S	S	B	S	S	B	S	S	S
5A	B	S	S	S	S	B	S	B	S	S	S
5B	S	S	B	B	S	S	B	B	B	B	B
5C	S	S	B	B	S	S	B	B	B	B	B
5D	B	S	S	B	S	B	S	B	S	B	B
5E	S	S	S	S	S	S	S	B	B	S	S
5F	S	S	S	S	S	S	S	B	S	S	S
5G	S	S	B	S	S	S	S	S	B	S	S
5H	B	S	S	S	S	B	S	B	S	S	S
6A	B	B	S	S	B	B	S	S	B	S	S
6B	S	S	B	B	S	S	B	S	S	B	B
6C	S	S	S	S	S	S	S	S	S	S	S
6D	S	S	S	S	S	S	S	S	S	S	S
6E	B	S	S	S	S	B	S	S	S	S	S
6F	S	S	S	S	S	S	S	B	S	S	S
6G	B	B	S	B	B	B	B	B	S	B	B
6H	B	B	B	S	B	B	S	S	B	S	S
7A	B	S	B	B	S	B	B	B	B	B	B
7B	B	B	S	S	B	B	S	S	B	S	S
7C	B	B	S	S	B	B	S	S	S	S	S

Appendix 1: - Segregation data for new *D2Ucl* loci on the Jackson backcrosses panels one B(BSB) and two (BSS)

7D	S	S	B	B	S	S	B	S	B	B	B
7E	B	B	S	S	B	B	S	B	S	S	S
7F	B	B	S	B	B	B	B	S	S	B	B
7G	B	B	B	B	B	B	B	S	B	B	B
7H	B	B	B	B	B	B	B	S	B	B	B
8A	S	S	S	S	S	S	S	B	S	S	S
8B	B	B	S	S	B	B	S	S	S	S	S
8C	B	B	S	S	S	B	S	S	S	S	S
8D	S	S	B	B	S	S	B	B	B	B	B
8E	S	S	B	B	S	S	B	S	B	B	B
8F	S	S	B	B	B	S	B	S	S	B	B
8G	S	S	B	S	S	S	B	S	B	S	S
8H	B	B	B	S	B	B	S	B	B	S	S
9A	S	S	S	S	S	S	S	S	S	S	S
9B	S	S	B	B	S	S	B	S	B	B	B
9C	B	S	B	B	S	B	B	B	B	B	B
9D	B	S	B	B	S	B	B	S	B	B	B
9E	S	S	B	B	S	S	B	S	B	B	B
9F	S	S	B	B	S	S	B	B	B	B	B
9G	B	B	B	B	S	B	B	S	B	B	B
9H	S	B	S	S	B	S	S	S	S	S	S
10A	S	S	B	B	S	S	B	B	B	B	B
10B	S	S	B	B	S	S	B	B	B	B	B
10C	B	B	S	S	B	B	S	B	S	S	S
10D	S	B	S	B	B	S	S	B	S	B	B
10E	S	S	B	B	S	S	B	B	B	B	B
10F	S	S	B	S	S	S	S	S	B	S	S
10G	B	S	S	S	S	B	S	S	S	S	S
10H	S	S	B	B	S	S	B	S	B	B	B
11A	B	B	S	B	B	B	B	B	B	B	B
11B	S	S	S	S	S	S	S	S	S	S	S
11C	S	S	B	B	S	S	B	B	B	B	B
11D	S	S	B	B	S	S	B	B	B	B	B
11E	B	B	B	B	B	B	B	B	B	B	B
11F	B	B	B	B	B	B	B	B	B	B	B
11G	B	B	S	S	B	B	S	S	B	S	S
11H	S	S	B	B	S	S	B	B	B	B	B
12A	B	B	S	B	S	B	S	S	S	B	B
12B	S	S	B	B	S	S	B	B	S	B	B
12C	S	B	B	B	B	S	B	B	B	B	B
12D	S	S	S	S	S	S	S	B	S	S	S
12E	B	B	B	B	B	B	B	B	S	B	B
12F	B	B	S	S	S	B	S	S	S	S	S
12G	B	B	S	S	B	B	S	B	S	S	S
12H	S	S	B	B	S	S	B	B	S	B	B

Appendix 1 (cont.): - Segregation data for new *D2Ucl* loci on the Jackson backcrosses panels one (BSB) and two (BSS)

Animal number	*Gnas*	*D2Mit73*	*D2Ucl18*	*D2Mit74*	*Acra4/Ra*
1	S	-	S	S	S
2	Ra	-	Ra	Ra	Ra
3	Ra	-	Ra	Ra	Ra
4	Ra	-	Ra	Ra	Ra
5	S	-	S	S	S
6	S	-	S	S	S
7	Ra	-	Ra	Ra	Ra
8	S	-	S	S	S
9	S	S	Ra	Ra	Ra
10	S	-	S	S	S
11	Ra	Ra	Ra	Ra	Ra
12	S	-	S	S	S
13	S	S	S	S	S
14	Ra	-	Ra	Ra	Ra
15	S	Ra	Ra	Ra	Ra
16	Ra	-	Ra	Ra	Ra
17	S	-	S	S	S
18	Ra	Ra	Ra	Ra	Ra
19	S	S	S	S	S
20	Ra	-	Ra	Ra	Ra
21	Ra	-	Ra	Ra	Ra
22	Ra	Ra	Ra	Ra	Ra
23	S	-	S	S	S
24	S	-	S	S	S
25	S	-	S	S	S
26	Ra	-	Ra	Ra	Ra
27	Ra	-	Ra	Ra	Ra
28	Ra	-	Ra	Ra	Ra
29	S	-	S	S	S
30	S	-	S	S	S
31	S	S	S	S	S
32	S	-	S	S	S
33	S	-	S	S	S
34	Ra	-	Ra	Ra	Ra
35	Ra	-	Ra	Ra	Ra
36	Ra	-	Ra	Ra	Ra
37	Ra	-	Ra	Ra	Ra
38	Ra	-	Ra	Ra	Ra
39	S	-	S	S	S
40	S	S	S	S	S
41	Ra	-	Ra	Ra	Ra
42	S	S	S	S	S
43	Ra	-	Ra	Ra	
44	Ra	-	Ra	Ra	Ra
45	Ra	-	Ra	Ra	Ra
46	Ra	Ra	Ra	Ra	Ra
47	S	S	S	S	S
48	S	S	S	S	S
49	S	S	S	S	S

Table A2 : - Segregation data for the ragged backcross

50	S	S	S	Ra	Ra
51	Ra	-	Ra	Ra	Ra
52	Ra	Ra	Ra	Ra	Ra
53	Ra	-	Ra	Ra	Ra
54	Ra	-	Ra	Ra	Ra
55	Ra	-	Ra	Ra	Ra
56	S	S	S	S	S
57	S	S	S	S	S
58	Ra	-	Ra	Ra	Ra
59	Ra	-	Ra	Ra	Ra
60	Ra	Ra	Ra	Ra	Ra
61	S	S	S	S	S
62	S	S	S	S	S
63	S	-	S	S	S
64	Ra	-	Ra	Ra	Ra
65	Ra	-	Ra	Ra	Ra
66	Ra	-	Ra	Ra	Ra
67	S	S	Ra	Ra	Ra
68	S	S	S	S	S
69	S	-	S	S	S
70	Ra	-	Ra	Ra	Ra
71	Ra	Ra	Ra	Ra	Ra
72	S	-	S	S	S
73	Ra	-	Ra	Ra	Ra
74	Ra	-	Ra	Ra	Ra
75	S	-	S	S	S
76	Ra	Ra	Ra	Ra	Ra
77	Ra	-	Ra	Ra	Ra
78	S	-	S	S	S
79	Ra	-	Ra	Ra	Ra
80	Ra	-	Ra	Ra	Ra
81	Ra	-	Ra	Ra	Ra
82	S	S	S	S	S
83	S	S	S	S	S
84	Ra	-	Ra	Ra	Ra
85	Ra	Ra	Ra	Ra	Ra
86	S	-	S	S	S
87	Ra	-	Ra	Ra	Ra
88	Ra	-	Ra	Ra	Ra
89	Ra	-	Ra	Ra	Ra
90	S	-	S	S	S
91	S	-	S	S	S
92	Ra	-	Ra	Ra	Ra
93	S	S	S	S	S
94	Ra	-	Ra	Ra	Ra
95	S	S	S	S	S
96	Ra	-	Ra	Ra	Ra
97	Ra	-	Ra	Ra	Ra
98	S	S	S	S	S
99	S	S	S	S	S
100	S	-	S	S	S
101	S	-	S	S	S

Table A2 (cont.): - Segregation data for the ragged backcross

102	S	S	S	S	S
103	S	Ra	Ra	Ra	Ra
104	S	S	S	Ra	Ra
105	S	S	S	S	S
106	S	S	S	S	S
107	S	S	S	S	S
108	Ra	Ra	Ra	Ra	Ra
109	Ra	Ra	Ra	Ra	Ra
110	Ra	Ra	Ra	Ra	Ra
111	Ra	Ra	Ra	Ra	Ra
112	S	S	S	S	S
113	Ra	Ra	Ra	Ra	Ra
114	S	S	S	S	S
115	Ra	Ra	Ra	Ra	Ra
116	S	S	S	S	S
117	Ra	-	Ra	Ra	Ra
118	S	S	S	S	S
119	S	S	S	S	S
120	S	S	S	S	S
121	Ra	-	Ra	Ra	Ra
122	S	S	S	S	S
123	S	S	S	S	S
124	Ra	Ra	Ra	Ra	Ra
125	Ra	-	Ra	Ra	Ra
126	Ra	-	Ra	Ra	Ra
127	S	S	S	S	S
128	S	S	S	S	S
129	S	S	S	S	S
130	Ra	-	Ra	Ra	Ra
131	S	-	S	S	S
132	S	-	S	S	S
133	S	-	S	S	S
134	Ra	-	Ra	Ra	Ra
135	Ra	-	Ra	Ra	Ra
136	Ra	Ra	Ra	Ra	Ra
137	S	-	S	S	S
138	Ra	-	Ra	Ra	Ra
139	Ra	Ra	Ra	Ra	Ra
140	S	S	S	S	S
141	S	-	S	S	S
142	Ra	-	Ra	Ra	Ra
143	Ra	Ra	Ra	Ra	Ra
144	Ra	-	Ra	Ra	Ra
145	Ra	-	Ra	Ra	Ra
146	Ra	Ra	Ra	Ra	Ra
147	S	S	S	S	S
148	S	S	S	S	S
149	Ra	Ra	Ra	Ra	Ra
150	S	S	S	S	S
151	Ra	Ra	Ra	Ra	Ra
152	Ra	Ra	Ra	Ra	Ra
153	Ra	Ra	Ra	Ra	Ra

Table A2 (cont.): - Segregation data for the ragged backcross

154	S	-	S	S	S
155	Ra	Ra	Ra	Ra	Ra
156	S	S	S	S	S
157	Ra	-	Ra	Ra	Ra
158	Ra	-	Ra	Ra	Ra
159	Ra	S	S	S	S
160	S	S	S	S	S
161	S	S	S	S	S
162	Ra	-	Ra	Ra	Ra
163	S	S	S	S	S
164	Ra	-	Ra	Ra	Ra
165	S	-	S	S	S

Table A2 (cont.): - Segregation data for the ragged backcross

Animal number	*Gnas*	*D2Mit73*	*D2Ucl18*	*D2Mit74*	*Acra4*
1	S	S	S	S	S
2	W	W	W	W	W
3	S	S	S	S	S
4	W	W	W	W	W
5	W	W	W	W	W
6	W	W	W	W	W
7	W	W	W	W	W
8	W	W	W	W	W
9	W	W	W	W	W
10	W	W	W	W	W
11	W	W	W	W	W
12	W	W	W	W	W
13	W	W	W	W	W
14	W	W	W	W	W
15	W	W	W	W	W
16	W	W	W	W	W
17	W	W	W	W	W
18	W	W	W	W	W
19	W	W	W	W	W
20	W	W	W	W	W
21	W	W	W	W	W
22	W	W	W	W	W
23	W	W	W	W	W
24	W	W	W	W	W
25	W	W	W	W	W
26	W	W	W	W	W
27	W	W	W	W	W
28	W	W	W	W	W
29	W	W	W	W	W
30	W	W	W	W	W
31	W	W	W	W	W
32	W	W	W	W	W
33	W	W	W	W	W
34	W	W	W	W	W
35	W	W	W	W	W
36	W	W	W	W	W
37	W	W	W	W	W
38	W	W	W	W	W

39	W	W	W	W	W
40	W	W	W	W	W
41	W	W	W	W	W
42	W	W	W	W	W
43	W	W	W	W	W
44	W	W	W	W	W
45	W	W	W	W	W
46	W	W	W	W	W
47	W	W	W	W	W
48	W	W	W	W	W
49	W	W	W	W	W
50	W	W	W	W	W
51	W	W	W	W	W
52	W	W	W	W	W
53	S	S	S	W	W
54	W	W	W	W	W
55	W	W	W	W	W
56	W	W	W	W	W
57	W	W	W	W	W
58	W	W	W	W	W
59	W	W	W	W	W
60	W	W	W	W	W
61	W	W	W	W	W
62	W	W	W	W	W
63	W	W	W	W	W
64	W	W	W	W	W
65	W	W	W	W	W
66	W	W	W	W	W
67	W	W	W	W	W
68	W	W	W	W	W
69	W	W	W	W	W
70	W	W	W	W	W
71	W	W	W	W	W
72	W	W	W	W	W
73	S	W	W	W	W
74	W	W	W	W	W
75	W	W	W	W	W
76	W	W	W	W	W
77	W	W	W	W	W
78	W	W	W	W	W
79	W	W	W	W	W
80	W	W	W	W	W
81	W	W	W	W	W
82	W	W	W	W	W
83	W	W	W	W	W
84	S	S	S	S	S
85	W	W	W	W	W
86	W	W	W	W	W
87	W	W	W	W	W
88	W	W	W	W	W
89	W	W	W	W	W
90	W	W	W	W	W
91	S	S	S	S	W
92	W	W	W	W	W
93	W	W	W	W	W
94	W	W	W	W	W

Table A3 (cont.): - Segregation data for the wasted backcross

Animal number	*D2Mit30*	*D2Ucl19*	*Cf2*	*D2Mit14*	*Ul*	*D2Ucl12*	*D2Mit11*
1	C	C	C	C	C	C	C
2	C	C	C	C	C	C	C
3		C	C	C	C	C	C
4	C	C	Ul	Ul	Ul	Ul	Ul
5		Ul	Ul	Ul	Ul	Ul	Ul
6		Ul		Ul	Ul	Ul	Ul
7		Ul	Ul	Ul	Ul	Ul	Ul
8		Ul			Ul		Ul
9					Ul		Ul
10					Ul		Ul
11				C	C		C
12	Ul	C	C	C	C	C	C
13					Ul		Ul
14	Ul		Ul	Ul	Ul		Ul
15	Ul	Ul	C		C	C	C
16					Ul		Ul
17					Ul		Ul
18					Ul		Ul
19	C		C	C	C		C
20					C		C
21					C		C
22	Ul		Ul	Ul	Ul		Ul
23					Ul		Ul
24	C		C	C	C		C
25					C		C
26					Ul		Ul
27	C		C	C	C		C
28					C		C
29					Ul		Ul
30					Ul		Ul
31					C		C
32					C		C
33					C		C
34	C	C	C	C	C	Ul	Ul
35					Ul		Ul
36					Ul		Ul
37	C	Ul	Ul	Ul	Ul	Ul	Ul
38	Ul			Ul	Ul		Ul
39					C		C
40					Ul		Ul
41					C		C
42					C		C
43	Ul	C	C	C	C	C	C
44					Ul		Ul
45					Ul		Ul
46					Ul		Ul
47					C		C

Table A4: - Segregation data for the ulnaless backcross

48					Ul		Ul
49					Ul		Ul
50					Ul		Ul
51					C		C
52	Ul	C	C	C	C	C	C
53					C		C
54					C		C
55	Ul	Ul	C	C	C	C	C
56					C		C
57					Ul		Ul
58					Ul		Ul
59					Ul		Ul
60					C		C
61		Ul	Ul	Ul	Ul	Ul	Ul
62					Ul		Ul
63	C	C	Ul	Ul	Ul	Ul	Ul
64					C		C
65					C		C
66					C		C
67	Ul	C		C	C	C	C
68					Ul		Ul
69					Ul		Ul
70					Ul		Ul
71					Ul		Ul
72	C	Ul	Ul	Ul	Ul	Ul	Ul
73					C		C
74		C	C	C	C	Ul	Ul
75	Ul	Ul	C		C	C	C
76					C		C
77		C	C	C	C	Ul	Ul
78					C		C
79					Ul		Ul
80					Ul		Ul
81					Ul		Ul
82					Ul		Ul
83					Ul		Ul
84	Ul			Ul	Ul		Ul
85	C			C	C		C
86	Ul			Ul	Ul		Ul
87	C				C		C
88	C			C	C		C
89	C			C	C		C
90	C			C	C		C
91	C			C	C		C
92	C			C	C		C
93	C			C	C		C

Table A4 (cont.): - Segregation data for the ulnaless backcross

Sequence information for mouse clones

All primer sequences are underlined. Sequences are as read, except where marked with a row of asterixes, where primer sequences are complementary and reversed.

D2Ucl11

KS sequence

TTCATATCCATAGAATAATTCTTTTTGTAAAAGGATTTATTTATATATTACATATATCAAT
GTTCTGCCTGCACGATACCTGCAGTCCAGAAAAGGGCACCAGACTTCATTATGGATGGTT
ATGAGCCACCATGTGGTTGCTGGGAATTGAACTAGGACCTCTGGAAGAATAGCAGTACTC
TTAACATCTGAACATCTCTCTCAGCCTGAATAATAA$(T)_{11}$C$(T)_4$AAGATACATTGGGAAATA
****
GGGCACTTGCAGGTGTTGATGGGTAAGAAGGCAATGGAGGAAATATACAATG
****

D2Ucl12

T3 sequence

CCCATCTAAAATTCTTAAGTCCAGATGTTTTTAGCACAGAATTCACCATACTTTCAAAGAA
GACTAATAACAATACAGACAATTTCACAAAACAGAAACAAAGGAACATTTCCAAACTTAC
TCATTGAGACCACAGTCACCCAAATAACGTAAACCACAAATACGTAAACCACAAAGAATC
AATAAAG

KS sequence

AATCCTTTCCATATAGGATATTGGTGTGAAAATACTCAATAAAACATTCACAAACTGAAC
TCAAGAA$(CA)_{23}$TGCACACAAATCATCCACCATGATCAGTAAGCTTCCACCCAGGATACAG
AGGTGGTCAATATATGAATCTCATCACTGATTCTA

D2Ucl13

T3 sequence

AAGAAGCTCACTTAAACTGTGCTTTTCTGCTGTGCGGCATATTTACTTGTCTGTACTTCAC
CATTAACGATTTCATAGCATACACATACAGGCTTTGGTGATACACGGTATGGAGTTTAAT
GAGTGACTACTTTACTTATTAGTTTCAATATACTTTGGACTACTTAAGG

KS sequence

TTGTGCATCTTTAGAGCCTCCCATCGATTCGTGTCGATGGTATTAGACTCTGTCCGTGTTG
TGATAGTGTGTGGCACTTACTTGTAGACACACCCTAGGACATCTGTCCCAAACTAAGCAC
TTGTCCCACACTGTGCTTGAACTTTCATGTCAAGCTTCTACCGGTAAAGCCAGTTTCGAGT
CCTTCTTCCTCCATCCAATTCACAAATACATAGCTGTTATTGTAGTCT

D2Ucl14

T3 sequence

TTTTCCCATAGCATTTGTAAATGCAGAAAAATACTAATAAAAAAATATGATGTGTTCTCAT
CTTGGAGACAGAAATAAAGTGACCTCCTCATCAATGCCGTAT(CAT)$_3$CGTAGTCGTTCGTT
TAT(CAT)$_5$TATTTCTACCCTAATTACTATGGTACCTACTGGGTAGT
****

D2Ucl15

T3 sequence

GGGACTTTTGGGATAGCATTGGAAATGTAATTGAGGGAAAATACGTAATAAAAAAGATAT
TTTAAAAAAGAGACAAAAAATGGGGGGAGAGCCTCGAACTCATTGGCAGAGGAGATGAC
TTCTGCATTGGACAAGAGCTCAAGTACTAAATTACAATTAATAGG

KS sequence

CCCAAATAACAGTGTGCTTTTCCATACATAGTCAAAACAGCTACATATATCCATTTATTAA
TTCTTAATTTTATACTTGAGCTCTGTCAATCATAAAGTCATCTCCTCTGCCAATATTCACTC
TCCCATTTCTCTTTTTCTCTTTTTTAAAATATCTTTTATTTCATTTCCTCAATT

D2Ucl16

T3 sequence

GGGATAGCATTGGATATGTAATGTTGAAAAATACCTAATTAAAAAAATCAAAAATGAAA
AAAAAATAGAGTTATAAAGATATGGCAAACAGTCAGTTTCAGCAACTGTGAGACTGATTG
CCACACCAGCCAGAAGGAGAGGAGTCTGCAGAGAGCAAGCTCCAGCTCAAACTAAAGGA
AGGTAACATAGATA

KS sequence

AAAAATATTTACCCTGATGTTGTGTATTTGATAAATCTGCTCTGACAGTTGTGTGTCATCT
TATTTCTGTTGTTTCCTTTTAGTTTGAGCTGGGAGCTTGCACACAGCAGACTCTCTTCTCTG
CTCCTCTGGCATCAGCTCACAGTTGCTCAAACGATCTGTTTTTGCACATTATACTCTA(T)$_7$C
CT(T)$_9$

D2Ucl17

T3 sequence

CCATATGACTGCAACATATTTATTTGGGACCATAGTCACCCAGGGAAGACAACTCAATTA
AAACAACAA(CAAA)$_4$TCAAAAAATCTAGCCCAGGCTACTCCATTCGATCCCG

KS sequence

TACAAATCGGAGTTTATATTTTTCAGCAGTCAGGGTTGGATGACATACTGACCTCATCCAC
TTTGTCCATCTAGATTCAGTA

D2Ucl18

T3 sequence

TTTAGTTTGAGCTCCCACTTGCTCTCTGCAGACTCCTCTTCTCTGGCATCAGTCTCACAGT
AGCTCAACTGACT

KS sequence

GTAACATAGATAGATGACACACACAACTGTCAGAGCAGGATTATCAATACACAACATCAG
GTAATATTTTTTCCTCTT

D2Ucl19

T3 sequence

AATTTTTTTTGGTTTTTTTTTGAATCATACTCCAGGCCAATTGGCCATAGACTTATCCAAAT
CTCCATTTCTGCCCCCCCCACCTTGCAGTAGGAGCGGACACCATCACATCTGTC$(T)_{14}$GGT$(G)_7$
TGTGCACTGAA

KS sequence

GAAGTAGAGAAGCAGAGGCACACCAAGGAGGGCCTCTCTGTCTAAATAAAAGCCGACCT
TAAGGGCAGTGGGGAACCATACCCTTTACCAGGGAATGAACAC

D2Ucl20

T3 sequence

AAGTATAAATGAGGAAAAAAAACCTAATAAAAAATATTAAAAAAGAAAAAGAAAGAAAT
ATGAAGAGATGATCATCAAGAAGTACTTCTTGAGAGCCAGAAACTGAACTTGCTGTATTT
CTTAGTTGTATAATCTTATACAGCTGTTTGAAAAGAACCAGTGTCTGTAATGTTTTGCATT
AGTTATCTATCTCTGAACATACTTAA

KS sequence

CAAATCACTTAGTATTCCAGAAGACAAATTCTCCCCAATACCCCAAACAGCTTCCTAGAA
AACAGTAGTATTAATACTACAGAAATCAAATACTAAGAAAACTAAGAAAAGTAGTATTA
GAGTACCAGTGTTCTGGCTCT

D2Ucl31

T3 sequence

GGGGGACTTTTGGGATAGCATTGAAAATGTAAATGAGGAAAATACCTAATTAAAAAAAA
AAGACATTTAAAAAGAAAAGAAAAGAAAAAACACAGATCTTTACCAGGCATCAGAAGAC
TTACTCTTGTTTTCTTCTGAGTCTTAGTTCTTGTAAGCCTTTTTAAAAGGATACTGAGAGAT
ACGTGTTTGGAACCTATTAGTTTTTGCCTCTTACTTCCATTCATAACGTAGCTGTCTTAGA
****

CTCTCCAGAAGAGAAT
****

Bubble 14

T3 sequence

AAAAAAGAGGGAAAAAATATTTACCCTGATGTTGTGTATTTGATAAATCCTGCTCTGACA
GTTGTGTGTCATCTTATTCTATTGTTTCCTTTCCTTTAGTGAGCTGGGACTTAGTGTGTGCT
GT

KS sequence

AAACAGTCAGTTTCAGCTACTGTGAGGCTGATGCCAGAGAAGCTGTTGTTGTGGAGTCTG
CAGAGAGCAAGCTCCCATAGCTCAAAATCAAAAGGAAAGGAAACAATAGAATAAGATGT
CTCTCTTCTGTACGTGACGGATTTATCAAATA