To the Graduate Council:

I am submitting herewith a dissertation written by Ethan A. Carver entitled "Comparative Mapping for Positional Cloning and Defining Homology Region Borders between the Mouse and Human Genomes.." I have examined the Final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biomedical Sciences.

Lisa Stubbs

Lisa Stubbs, Major Professor

We have read this dissertation
and recommend its acceptance:

Debbie Johnson

Mary Ann Handel

Edward J. Michaud

John Eppig Wilkinson

Accepted for the Council:

Associate Vice Chancellor and
Dean of The Graduate School

COMPARATIVE MAPPING FOR POSITIONAL CLONING AND DEFINING HOMOLOGY REGIONS BETWEEN THE MOUSE AND HUMAN GENOMES

A Dissertation Presented for the Doctor of Philosophy Degree

The University of Tennessee, Knoxville

Ethan A. Carver

May 1999

Dedication page

This dissertation is dedicated to my family. My wife, Tonya, who has been remarkably patient and understanding throughout this endeavor. My parents, Glenn and Audrey Carver, who instilled in me a curiosity for the world in which we live. And Michael, for being more than my brother, for being my friend.

ACKNOWLEDGEMENTS

I would like to express my thanks and appreciation to innumerable people who have helped me, both personally and professionally, to achieve this goal. I would like to especially thank my major advisor, Lisa Stubbs, for giving me encouragement, instruction, and guidance throughout my graduate studies. Also, I would like to thank my committee members; Mary Ann Handel, Dabney Johnson, Edward Michaud, and Erby Wilkinson, for their encouragement, advice, and constructive criticisms during the course of these studies. I am very grateful to Cymbeline Culiat for her guidance, instruction, and help during these projects too numerous to detail. I would like to thank Waldrico Generoso, who created the 1Gso mutant, for thought-provoking discussions, and for providing a different perspective on my research. Also, I am indebted to K.T. Cain, who maintained most of these animals, and taught me much about the mouse in general. I would also like to thank members of both Oak Ridge National Laboratory and Lawrence Livermore National Laboratory for their technical support and friendship; Mark Shannon, Joomyeong Kim, Johannah Doyle, Nestor Cacheiro, Linda Ashworth, Anne Olsen, Mitch Walkowicz, Estela Generoso, Michael Mucenski, Rick Woychik, Anne Bergman, Beverly Stanford, Teresa Sobhani, John Geisler, Bill Dunn, and the mouse caretakers at both facilities.

ABSTRACT

The goal of this dissertation project was to utilize genetic and physical mapping methods as a means to define genomic homology between human and mouse genomes, as well as to use this information to define functional relationships between the two species. The comparative mapping studies were designed to expand upon the current knowledge of comparative mapping and homology regions between mouse and man, and to begin to study the homology region borders. Positional cloning research was initiated to localize the translocation breakpoints in a mutant mouse associated with a neurological defect as a first step toward isolation of genes that could be involved in the phenotype of this animal.

Comparative mapping of human chromosome 19 and related regions of the mouse genome represents one major focus of this research. Human chromosome 19 was a good target for comparative studies due to the extensive physical mapping of the chromosome, and availability of conserved, mapped gene markers to use for these studies. Comparative studies involved both the 19q- and 19p-arms of the chromosome, and helped lay the foundation of a chromosome-wide comparative map. One region, 19p13.1, was investigated in detail. This region is shown to be prone to rearrangements during evolution, as indicated by the homology groups associated within both the mouse and human genomes. These studies revealed the need for more fine mapping of the genome for both species, demonstrating that examination of specific homology groups at a higher resolution reveals that both similarities and surprising differences between related mouse and human regions. These studies provided data suggesting that repeated sequences are associated with at least some homology region borders, a concept that may serve as a guide for future comparative mapping between human, mouse, and other species.

One application of comparative mapping is its ability to link functional information derived from mouse mutations to specific genes and disease within the human genome. As part of this effort, research focused upon localizing and physical mapping a region

surrounding the breakpoint in 1Gso, a mouse mutation associated with a reciprocal translocation. Homozygosity for the translocation causes developmental lethality. Animals that are heterozygous display neurological defects including the inability to swim and abnormal startle responses. A genetic map of the region surrounding one of the translocation breakpoints was established on mouse chromosome 2. Fluorescence *in situ* hybridization (FISH) techniques were used to localize translocation breakpoints to a small interval less than 1cM in length. This region shows remarkable linkage conservation with a well-characterized region of human chromosome 11p13, which harbors genes responsible for the WAGR syndrome (Wilms tumor, aniridia, genitourinary malformations, and mental retardation). Comparative mapping information was used to narrow down the translocation interval, and establish a contig consisting of yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), and P1-derived artificial chromosomes (PACs), which covered approximately 600kb throughout the breakpoint region. This allowed the breakpoint to be isolated within a single PAC clone of 150 kb in length.

Together these studies have set the stage for future investigations of genes located in chromosome 19, and for the cloning of gene(s) associated with the 1Gso phenotype in mice.

TABLE OF CONTENTS

SECTION	PAGE
PART I: UTILITY OF THE MOUSE/HUMAN COMPARATIVE MAPS	1
CHAPTER 1. INTRODUCTION	2
History of mapping in mouse	4
History of mapping in humans	6
Mouse backcross explanation	8
Human metric map of chromosome 19	14
Merging of human and mouse maps	14
REFERENCES	24
PART II. COMPARATIVE MAPPING OF HUMAN CHROMOSOME 19p AND RELATED REGIONS OF THE MOUSE GENOME	38
CHAPTER 1. INTRODUCTION	39
REFERENCES	43
CHAPTER 2. PUBLICATIONS DESCRIBING THE COMPARATIVE MAPPING OF HUMAN	

CHROMOSOME 19p AND RELATED
REGIONS OF THE MOUSE GENOME 46

SECTION 1. LOCATION OF MOUSE AND
HUMAN GENES CORRESPONDING TO
CONSERVED CANINE OLFACTORY RECEPTOR
GENE SUBFAMILIES 47

Abstract 49

Introduction 50

Materials and Methods 51

Interspecific Backcross Mapping 51

Human Mapping 51

Molecular Probes 52

Results 52

Murine genes related to CfOLF1 and CfOLF2
subfamilies are clustered in Mmu2 . . . 52

Locations of mouse genes related to CfOLF3
and CfOLF4 subfamilies 58

Physical localization of CfOLF4-related genes
in human Chromosome 19p 64

Discussion 67

Acknowledgments 70

References 71

SECTION 2. PHYSICAL MAPPING OF EMR1
AND CD97, TWO DISTANTLY RELATED GENES,

TO HUMAN CHROMOSOME 19 AND ASSIGNMENT OF
Cd97 TO MOUSE CHROMOSOME 8 77
Acknowledgments 86
References 87

SECTION 3. COMPARATIVE MAPPING
OF HUMAN 19p13.1 SURROUNDING THE *NOTCH3*
GENE AND OLFACTORY GENE CLUSTERS:
DELINEATION OF HOMOLOGY GROUPS ON
MOUSE CHROMOSOMES 8, 10, AND 17 90
Introduction 92
Materials and Methods 95
Molecular probes 95
Isolation of olfactory receptor-specific probes. . . 95
Mouse genetic mapping 98
DNA sequence analysis 98
Results 99
Mapping of *NOTCH3* and surrounding
olfactory receptors in mouse. 99
Sequence and classification of 19p13.1
Oflr genes 104
Discussion 105
Acknowledgments 109
References 110

PART III. LOCALIZATION AND CLONING OF THE MMU 3 TRANSLOCATION BREAKPOINT IN THE 1GSO MUTANT MOUSE 116

CHAPTER 1. INTRODUCTION 117

Translocations can disrupt the structure and/or expression of critical genes 124

DNA rearrangements provide visible access to mutant genes 125

CHAPTER 2. MATERIALS AND METHODS 129

Interspecific backcross mapping 129

Genomic DNA isolation 129

Mouse metaphase chromosome preparation 129

Fluorescence *in situ* hybridization 130

Single-primer PCR to generate endclones 130

Yeast Artificial Chromosome experiments 131

Transformation associated recombination (TAR) cloning . 131

CHAPTER 3. THE 1GSO PROJECT 133

SECTION 1. CONSTRUCTION OF A BACKCROSS MAP OF MMU 2 AND MMU 14 133

SECTION 2. USING FLUORESCENCE *IN SITU* HYBRIDIZATION TO LOCALIZE THE TRANSLOCATION TO A SMALL INTERVAL ON MOUSE CHROMOSOME 2 . 141

Use of FISH to narrow down the location of the translocation breakpoint site 141
FISH strategy to map translocation breakpoints . . . 142
FISH mapping of the translocation breakpoint 147

SECTION 3. PHYSICAL LOCALIZATION OF THE BREAKPOINT ON MMU 2 157
Isolation of smaller clones to further narrow the breakpoint region 162
Isolation of breakpoint containing clones 168
FISH analysis on chromosome 14 171
Genes excluded by the FISH analysis on both Mmu 2 and Mmu 14 172

SECTION 4. OBSERVATIONS ON THE 1GSO MUTANT PHENOTYPE AND GENETIC STUDIES 179
Detailed studies of the phenotype of 1Gso animals 179
Complementation studies involving 1Gso mice and animals carrying Sey^{Dey}, a deletion spanning the breakpoint on Mmu2 181

SECTION 5. CONSTRUCTION OF A COSMID LIBRARY OF 1GSO/+ DNA AND TAR CLONING TO GAIN ACCESS TO THE BREAKPOINT 185

SUMMARY 190

REFERENCES 191

PART IV: CONCLUSIONS 206

CHAPTER 1. INTRODUCTION 207

SECTION 1. THE COMPARATIVE MAP OF CHROMOSOME 19p AND RELATED REGIONS OF THE MOUSE GENOME 209

Summary of research 209

Conclusions and future directions 212

State of sequence generated and analyzed 215

SECTION 2. CONCLUSIONS OF THE 1GSO POSITIONAL CLONING PROJECT AND FUTURE DIRECTIONS 217

Summary of research 217

Conclusions and future directions 218

Sequencing across the breakpoint region in mouse chromosome 2. 218

Types of mutations associated with translocations and their possible relevance to the *diver* phenotype . 222

Types of mutations associated with translocations . . . 222

1. Null mutations 222

2. Production of truncated proteins 223
3. Overexpression or expression of genes at ectopic sites . 224
4. Association of the translocation with other, complex rearrangements 225
5. Disruption of genes on both chromosome 2 and 14 . . 226
Dissecting out the contributions of candidate genes . . 227
Potential types of genes involved in the 1Gso phenotype, as seen by other murine phenotypes . 228
1. *Pallid* 228
2. Targeted deletions of dopamine receptors 229
3. Neurotrophic factors and a targeted deletion of the Bdnf gene in mice 229
A hypothesis regarding the molecular basis of the 1Gso mutation and correlation to its phenotype . . 230
Future applications of translocation cloning in the context of the completed Genome project . . 231
Final summation 232
REFERENCES 234

APPENDIX 246

VITA 251

LIST OF TABLES

TABLE PAGE

PART II. COMPARATIVE MAPPING OF HUMAN CHROMOSOME 19p AND RELATED REGIONS OF THE MOUSE GENOME

CHAPTER 2. SECTION 1.

Table 1. Description of probes and variant fragments. 55

Table 2. Protein sequence species search. 60

CHAPTER 2. SECTION 3.

Table 1. Description of probes and variant fragments. 97

PART III. LOCALIZATION AND CLONING OF THE MMU 3 TRANSLOCATION BREAKPOINT IN THE 1GSO MUTANT MOUSE

Table 1. Clones and probes used to construct a backcross map on mouse chromosome 2. . . . 138

Table 2. Clones and probes used to construct a backcross map on mouse chromosome 14. . . . 140

Table 3. Probes and genomic clones used for FISH analysis. . . . 144

LIST OF FIGURES

FIGURE PAGE

PART I: UTILITY OF THE MOUSE/HUMAN COMPARATIVE MAPS

Figure 1. Mouse Interspecific Backcross. 11

Figure 2. Generation of a genetic linkage map. 13

Figure 3. Mouse-Human comparative map. 18-19

PART II. COMPARATIVE MAPPING OF HUMAN CHROMOSOME 19p AND RELATED REGIONS OF THE MOUSE GENOME

Chapter 1.

Figure 1. Human chromosome 19 and related regions of the mouse genome. 41

Chapter 2. Section 1.

Figure 1. Segregation patterns and predicted map positions of murine genes related to canine CfOLF1, CfOLF2, and CfOLF3 subfamilies. . 57

Figure 2. Map positions of two murine CfOLF4-related loci. . . 63

Figure 3. Location of CfOLF4-related genes on human chromosome 19p. 66

Chapter 2. Section 2.

Figure 1. Location of *EMR1* and *CD97* genes on human chromosome 19p13. 82

Figure 2. Segregation pattern of *Cd97* and surrounding genes mapping to mouse chromosome 8. . . 85

Chapter 2. Section 3.

Figure 1. Segregation patterns of *Notch3* and surrounding olfactory receptor genes mapping to mouse chromosome 17 and 10. 101

Figure 2. Location of genes in the human metric map, and their locations in the mouse genome. . . 103

Figure 3. Location of genomic regions in human and mouse subject to evolutionary rearrangements. . . 107

PART III. LOCALIZATION AND CLONING OF THE MMU 3 TRANSLOCATION BREAKPOINT IN THE 1GSO MUTANT MOUSE

Figure 1. Giemsa-band karyotype of 1Gso female mouse. . . 119

Figure 2. 1Gso "diver" phenotype. 121

Figure 3. 1Gso escape response. 123

Figure 4. Backcross Map of mouse chromosome 2 and 14. . . 134

Figure 5. FISH analysis of translocation mutations. . . . 146

Figure 6. *Hoxd4* and *Pax6* FISH analysis. 149

Figure 7. Summary of FISH steps. 151

Figure 8. Mouse and human homology surrounding the WAGR region. 154

Figure 9. FISH analysis between *Hoxd4-Wt1* and *Wt1-Pax6*. . . 156

Figure 10. YAC D0214D FISH analysis. 159

Figure 11. Restriction map of YAC D0214D. 161

Figure 12. FISH analysis of *Rcn* and *Pax6*. 164

Figure 13. Chromosome walk onto wrong chromosome. . . 167

Figure 14. Physical Map of the 1Gso breakpoint region on mouse chromosome 2. 170

Figure 15. Overlapping clones used in FISH. 174

Figure 16. The region of the translocation breakpoint in relation to the physical map. 176

Figure 17. Chromosome 14 FISH analysis. . . . 178

Figure 18. Compound mutant, 1Gso and Sey^{Dey} , and the corresponding region in the mouse genome. . . . 183

Figure 19. TAR cloning vector and strategy. 188

PART IV: CONCLUSIONS

Chapter 1. Section 1.

Figure 1. Human chromosome 19 and homology regions with the mouse genome. 211

Chapter 1. Section 2.

Figure 2. Analysis of genomic DNA of 1Gso and background animals. 221

PART I

UTILITY OF THE HUMAN/MOUSE COMPARATIVE MAPS

CHAPTER 1.
INTRODUCTION

The research projects described here utilized comparative mapping of genomic homology regions in mouse and human genomes and took advantage of these relationships to localize a translocation breakpoint in the mouse. Both series of experiments were a part of a laboratory-wide effort to establish an interspecific backcross panel and comparative maps within several regions of the human and mouse genomes. Efforts to extend these genetic maps further in specific regions, and to apply them in a positional cloning project in the mouse are described.

This dissertation is divided into four parts to present the data in an organized and readable manner. The first Part introduces the concepts of genetic mapping in both mouse and humans, as well as the establishment of the comparative maps in both species.

Part II describes studies that extended the comparative maps of human chromosome 19 and related regions of the mouse genome. The comparative analysis of human chromosome 19p illustrates the similarities and differences that have arisen in two closely related regions during mammalian evolution. The fine-mapping of conserved genes from human chromosome 19 in the mouse genome (Stubbs *et al.*, 1996b; Carver and Stubbs, 1997; Kim *et al.*, 1997; Carver *et al.*, in preparation), illustrates the need for closer investigation of homology groups that have been established between the two species. Although grossly similar, human chromosome 19p regions differ significantly in detailed structure compared to the mouse genome. In the course of these studies, two small homology segments were mapped, of which one had not been previously described. This region will be an interesting test case for comparative sequencing because it includes a region that contains multiple mouse homology regions, several different gene families, and known human disease genes. Knowledge of the degree to which the structure of this

region differs between the species should be important for development of disease models and other functional studies in mice.

Part III describes the evolution of a positional cloning study centered on the 1Gso translocation mutant. The unusual and interesting phenotype of this mutant mouse is detailed. The positional localization of the 1Gso breakpoint in mouse relied upon the establishment of a backcross map of the translocation interval to aid in narrowing down the breakpoint to a small region on mouse chromosome 2. A fluorescence *in situ* hybridization (FISH) strategy was utilized to bracket the translocation between sets of mapped markers, and to localize the region to a small interval that could be approached with conventional positional cloning techniques. This region shows linkage conservation to human chromosome 11p13, and we exploited this conservation to help delineate the region further in the mouse. A chromosome walk was initiated and a contiguous array (contig) of overlapping clones was built that covered the breakpoint interval on mouse chromosome 2. FISH was then employed to bracket the breakpoint within a single clone. The clone will be the starting point to begin to search for potential candidate genes.

Part IV summarizes and gives future goals for each research series. The comparative mapping of human chromosome 19 and the homology region borders are discussed, as are theories regarding mechanisms for genome divergence (Carver and Stubbs, 1997). Also the possible next stages of this project are envisioned and discussed.

The likely outcome, potential candidate gene(s) associated with the phenotype and future directions for the 1Gso positional cloning project are considered. The relevance of the two different projects and their reliance on the comparative maps between human and mouse as well as the future utility of the comparative maps and sequence analysis concludes the section.

These maps and comparisons between mouse and human genomes are extremely powerful tools in positional cloning, clinical diagnosis, evolutionary biology, and many

other applications. The mapping techniques employed today differ greatly in their methodologies, but all retain their origins in the genetics that has been established over the last century.

History of mapping in mouse

The basis for mapping genes in any genome has its beginnings with the basic laws of inheritance established by Gregor Mendel (1865). These revolutionary ideas were not widely known during Mendel's lifetime, but were rediscovered in early 1900. The idea of "heritable units" which could determine a dominant or recessive trait, and their segregation and independent assortment from other heritable units are the basis for modern genetics.

With the use of different coat colors in mouse fancy strains, research quickly affirmed that inheritance in animals conformed to the newly rediscovered laws (Castle, 1903, and Cuénot, 1905). The mouse as a genetic mapping resource began with this research in the first decade of the twentieth century. The use of various phenotypes in the mouse was the starting point for segregation studies and linkage analysis. Pink-eyed dilute coat color mice were bred with albino mice and linkage between the two coat color genes was established (Haldane *et al.*, 1915). Early work with visible phenotypes utilized mice in mating strategies to determine if the traits would segregate together, indicating linkage or close proximity to each other, or if they would segregate independently of each other, meaning that they were not closely associated with each other. This established a linkage map based purely on the analysis of existing phenotypes in mice. As each new mutation was discovered, it was bred to strains with other phenotypic markers. Further breeding was done to ascertain whether the new mutation showed linkage to any of the previously characterized markers. As linkage groups became known, they were given group designations, and additional breeding experiments were done to refine these map

positions. The existence of the "heritable unit" was inferred, without regard to its physical structure. The establishment of these types of maps preceded the discovery that DNA was the heritable source of the phenotypes (Avery, MacLeod, and MacCarty, 1944; Hershey and Chase, 1952), the structure of DNA itself (Watson and Crick, 1953), or the elucidation of the genetic code (summarized CSHL symposium, 1967).

The map positions of many different mouse phenotypes established eighteen autosomal linkage groups by the beginning of the 1970s (Green, 1971). These linkage groups had no relationship to their physical placement within specific chromosomes. It was not until the early 1970s that the first assignment of an autosomal linkage group (XII) was made to a mouse chromosome (19) (Eicher, 1971).

In the 1970-80s, two different developments allowed for the construction of a mouse genetic map that covered the entire genome based entirely on DNA marker loci, and not upon linkage of visible phenotypes. First was the discovery of restriction fragment length polymorphisms (Kan and Dozy, 1978), and their subsequent use in mice (Arnheim *et al.*, 1982; Bennett *et al.*, 1982). These polymorphisms were initially difficult to use as a mapping resource due to the common ancestry of the traditional inbred strains of mice that were used at the time. The second significant event in mouse genome mapping was the generation and use of the interspecific backcross (IB). A backcross between two distinct species of mice, *Mus musculus* and *Mus spretus*, was generated (Avner *et al.*, 1988), which carried DNA polymorphisms for nearly every DNA probe tested. The high degree of sequence variation existing between the parental species allowed the segregation patterns of a majority of DNA loci to be followed in the backcross progeny. Backcrosses have become the major system which is used to map markers in the mouse (Copeland *et al.*, 1993; Dietrich *et al.*, 1995 and 1996). Since the advent of the IB methodologies, new polymorphic markers have been added to the mouse genetic map at an astounding rate.

The mouse has become the model of choice for comparative mapping for a number of reasons. Mice are well suited to genetic research because of their small size and docility, short developmental time to become breeding adults, and their overall similarity, both physiologically and genetically, to man. The mouse has therefore been used historically for linkage analysis. It was and still is the best-mapped mammalian genome besides human. Therefore, direct comparison between the two species has been possible for many years and relationships between the mouse and human genomes are well researched.

In parallel, within the last twenty years, there have been an increasing number of advances in the ability of researchers to study and manipulate the mouse embryonic genome in ways to better understand specific genes and their biological function. These reasons epitomize why the mouse has become the major mammalian model with which to study mammalian biology.

History of mapping in humans

During the same period of history, some of the earliest linkage mapping in humans was of genes on the X chromosome in relation to recessive disorders such as color blindness (Wilson *et al.*, 1911). The first autosome linkage, Lutheran blood group vs. secretor factor, was discovered in humans during the early 1950s (Mohr, 1951). Autosomal linkage was demonstrated by use of the sib-pair method, based upon the concept that if two loci were linked, siblings will fail to show random association of a trait determined by the loci. The correct chromosome number in man was discovered in the mid-1950s (Tijio and Levan, 1956; Ford and Hamerton, 1956). Finally, the Duffy blood group locus was assigned to human chromosome 1 by the association of the phenotype to a cytogenetically visible marker near the centromere of chromosome 1 within a pedigree derived from the Donahue family (Donahue *et al.*, 1968).

Human-mouse cell hybrids have been used as mapping reagents since the late 1960s (Weiss and Green, 1967), due to the differential loss of human chromosomes from these cells. Linkage between enzyme phenotypes could be inferred from concordant segregation of single human chromosomes within these hybrid cells. Sub-chromosomal localization could be achieved through the use of hybrid cell lines carrying only particular regions of a human chromosome. These methods were used with great success to map genes in humans and to assign the human linkage groups to their appropriate chromosomes (reviewed by Ruddle 1972 and 1981).

Mapping in humans progressed in concert with the mouse, and when developments occurred which aided analysis in one species, they were quickly adapted to work in the other species. Development of linkage maps in humans and mice was initially hampered due to the lack of available markers, but with the advent of recombinant DNA technology, this limitation was overcome. The first DNA polymorphism, detected by restriction enzyme *Hpa*I, was closely linked to the B-globin gene (Kan and Dozy, 1978). Restriction polymorphisms illustrated that DNA sequence was variable and that differences in DNA sequence could be detected by size fractionation on gel electrophoresis.

Construction of a human linkage map by using restriction fragment length polymorphisms (RFLPs) was proposed in the early 1980s (Botstein *et al.*, 1980). The use of random, single-copy DNA probes to detect restriction polymorphisms in DNA, and the ability to trace these changes in a human pedigree, allowed researchers to localize DNA sequences within the genome, and to check their segregation in relationship to a known inherited trait. This made possible the genetic localization of a trait without direct access to the responsible gene's DNA, and revolutionized the construction of linkage maps for both mouse and man. Detailed pedigree analysis was shortly accomplished in families exhibiting genetic disorders such as Duchenne muscular dystrophy (Murray *et al.*, 1982),

Huntington disease (Hodge *et al.*, 1980), Wilsons disease (Frydman *et al.*, 1985), and many others.

Linked DNA markers were then used as a predictive marker for the trait, and this fact initiated a new era in disease gene cloning. The isolation of genes based upon their position within the genome by "reverse genetics", or positional cloning was possible. The first unknown gene cloned by a positional cloning approach was the Duchenne muscular dystrophy gene (Monaco *et al.*, 1986).

Because of use of family linkage studies, chromosome studies, somatic cell hybrid studies, and molecular polymorphism methodologies, construction of the linkage map in human began to progress at an almost logarithmic rate from the late 1960s into the early 1990s (McKusick, 1991). These maps have become the basis for the generation of a complete physical map of the human genome (Cohen *et al.*, 1993; Schuler *et al.*, 1996; Bentley *et al.*, 1998); the ultimate goal is the complete sequence of the genome by the year 2003 (Collins *et al.*, 1998).

Mouse backcross explanation

The research described in the next two chapters relied upon the use of interspecific backcross mapping to localize markers from the metric physical map of human chromosome 19 to different regions of the mouse genome. These two different types of maps are described concisely in the ensuing paragraphs to facilitate a better understanding of the maps themselves.

Briefly, two different species of mouse, *Mus musculus* and *Mus spretus*, were used to generate the IB panel we have used in the Stubbs laboratory. Different subspecies can also be used, such as *Mus musculus castaneus* and *Mus musculus musculus*; however, the frequency at which polymorphic sequences are found to distinguish these two subspecies is much lower. *Mus spretus* is one of the most distantly related *Mus* species that will

breed with *Mus musculus* and produce fertile F_1 hybrid animals (Avner *et al.*, 1988). These two species of animals are mated together to create an F_1 hybrid generation. The F_1 females (the males are sterile and cannot be used) are then "backcrossed" to male mice of the *Mus musculus* species. The backcross to *Mus musculus* animals reduces the complexity of the genomes in F_2 animals and allows the recombination events that took place in the F_1 hybrid animals during gametogenesis to be observed easier (Figure 1).

Genomic DNA is then isolated from each backcross progeny mouse and digested with an array of restriction enzymes. Each set of digests is Southern-blotted (Southern, 1975) and restriction fragment length variance is used to differentiate between the *Mus musculus* allele and the *Mus spretus* allele for each gene or marker (Figure 1). Since the *Mus musculus* allele is present in all animals, due to the contribution of the *Mus musculus* male in the backcross step, the presence/absence of the *Mus spretus* alleles are scored. By following the inheritance patterns of a number of these animals, one can trace closely linked markers together (Figure 2). In the Stubbs' laboratory and many others, these patterns are analyzed by a computer program, the Map Manager data analysis program (Manly, 1993), which suggests linkage relationships and calculates genetic distances and standard errors using statistical methods (Silver, 1985). The mouse backcross panel [(129/Rl-*p* c^{ch}/*p* c^{ch} x *M. spretus*) X 129/Rl-*p* c^{ch}/*p* c^{ch}] used in our laboratory for mapping markers on chromosome 7, was generated originally in Dr. Rinchik's laboratory (Johnson, 1990). A second IB system [(C3Hf/R1-$Mgf^{SI\text{-}ENUR_8}$/+ x *M. spretus*) X C3Hf/R1], was originally generated by Dr. Rinchik's group, and has been typed for over 400 conserved gene markers distributed throughout all 19 mouse autosomes

Figure 1. Mouse Interspecific Backcross.

This figure illustrates the chromosomes at each of the steps in generating a panel of backcross animals to use for mapping markers within the mouse genome. The segregation of a single chromosome is followed to simplify the figures, with the understanding that similar occurrences are undergone with all other chromosomes in the genome. *Mus musculus* and *Mus spretus* animals are bred together to generate an F_1 hybrid animal. The males of this generation are sterile, and can not be used. During gametogenesis, crossovers between the two chromosomes occur, and an exchange of DNA takes place. The hybrid females are then mated to *Mus musculus* males. These animals are then used to prepare DNA, which has been digested with various restriction enzymes, size fractionated by gel electrophoresis and Southern blotted. Probes are then prepared, and hybridized to a panel of DNA with an enzyme that shows a restriction fragment length variation between the two species.

Backcross matings

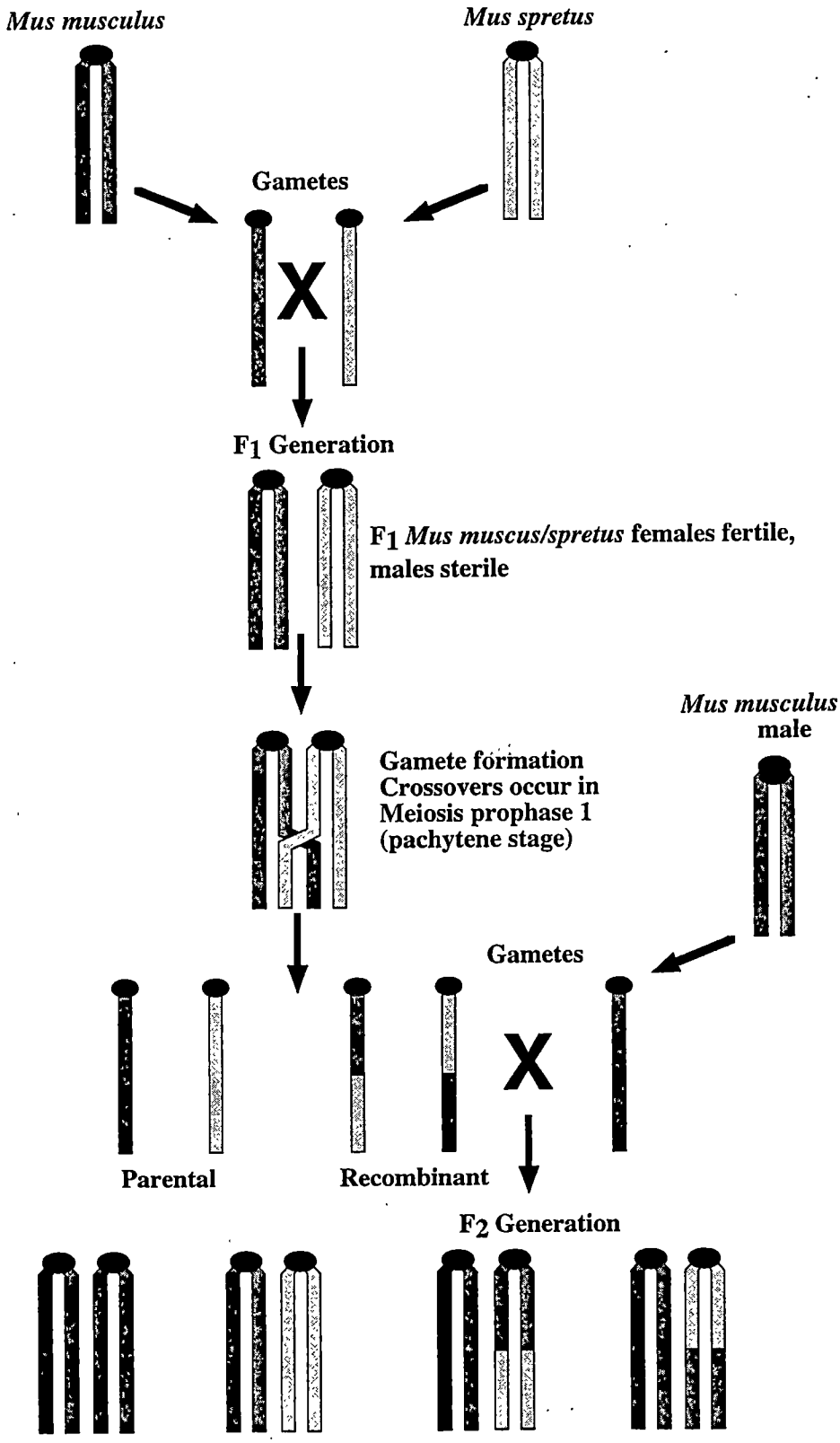
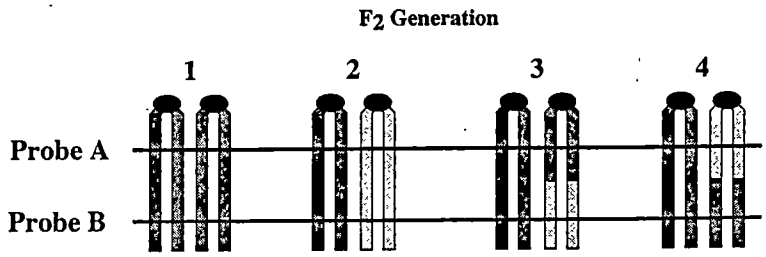


Figure 2. Generation of a genetic linkage map.

This figure illustrates the use of the backcross animals to generate a genetic linkage map. At the top, probes are hybridized to panels of DNA that correspond to different F_2 generation animals. Depending upon where the probe localizes in the genome, a different segregation pattern will emerge, as illustrated at the bottom of the figure. This segregation pattern can be compared to patterns formed by other probes. If the markers show the same pattern of inheritance for a majority of animals, then linkage between the markers can be inferred. By using multiple animals, the statistics involved in this analysis become more significant.

Backcross Analysis



A Genetic Linkage Map is based on the frequency of coinheritance of two or more "genetic markers".

These "genetic markers" are screened by their ability to detect RFLVs that occur between the *Mus musculus* and the *Mus spretus* species of mice.

	1	2	3	4
Probe A	M	S	M	S
Probe B	M	S	S	M

Crossovers will occur throughout the genome. By doing multiple backcrosses, and aquiring many F_2 animals, the statistics involved with this type of analysis become more significant.
Each area of a given chromosome will develop its own "pattern" of crossovers which are used to localize markers to a specific position within the genome.

and the X-chromosome (Stubbs *et al.*, 1996a; Doyle *et al.*, 1996; Kim *et al.*, 1997; Carver *et al.*, 1998 and submitted and unpublished; Stubbs and colleagues, unpublished).

Human metric map of chromosome 19

Most of our comparative mapping studies were based on human chromosome 19 (Stubbs *et al.*, 1996; Shannon *et al.*, 1996; Carver *et al.*, 1998). A physical metric map of human chromosome 19 has been generated (Ashworth *et al.*, 1995). This map is based on the assembly of sets of overlapping cosmids into contigs. These contiguous sets were generated by an automated fingerprinting method (Carrano *et al.*, 1989), and distances and order were established by using fluorescence *in situ* hybridization techniques (Brandriff *et al.*, 1991; Trask *et al.*, 1993). The contigs have been joined together to cover most of the chromosome. The chromosome is about 64 Mb in length, and over 450 different markers including genes, STSs, cDNAs, and others, have been localized to the chromosome 19 physical map. A large part of mouse chromosome 7 shows conserved synteny of homologous genes linked on human chromosome 19q (Stubbs *et al.*, 1996b; Shannon *et al.*, 1996; Kim *et al.*, 1997). Human chromosome 19p is comprised of several small regions of conserved synteny, including regions related to mouse chromosomes 8, 9, 10, and 17. Because homology region relationships in the 19p region were not well established, this region was a primary focus of subsequent mapping studies.

Merging of human and mouse maps

The field of comparative gene mapping has gained considerable momentum in recent years, because findings indicate that genes closely linked in one species also tend to be closely linked in the genomes of other vertebrates. This is generally the case for genes that are physically located close to each other, but during evolution some rearrangement of the genome has occurred. During evolution and speciation in mammals, the genome

appears to have remained roughly equivalent in size and gene content, while the order of genes has been disturbed due to rearrangements in the genome, including deletions, inversions, insertions, and transposition events.

Nevertheless, within large conserved blocks of DNA, the chromosomes of different vertebrate species appear to be organized in a very similar manner. Even though conservation of genomic organization in different mammals has been postulated since the early 1900s (Haldane, 1927), conservation of synteny in mouse and human was not formally illustrated until much later (Ohno, 1967; Minna *et al.*, 1976; Lalley *et al.*, 1978). With the number of markers available at that time, it seemed that at the gross level, the two genomes were remarkably similar in structure and organization (Nadeau and Taylor, 1984).

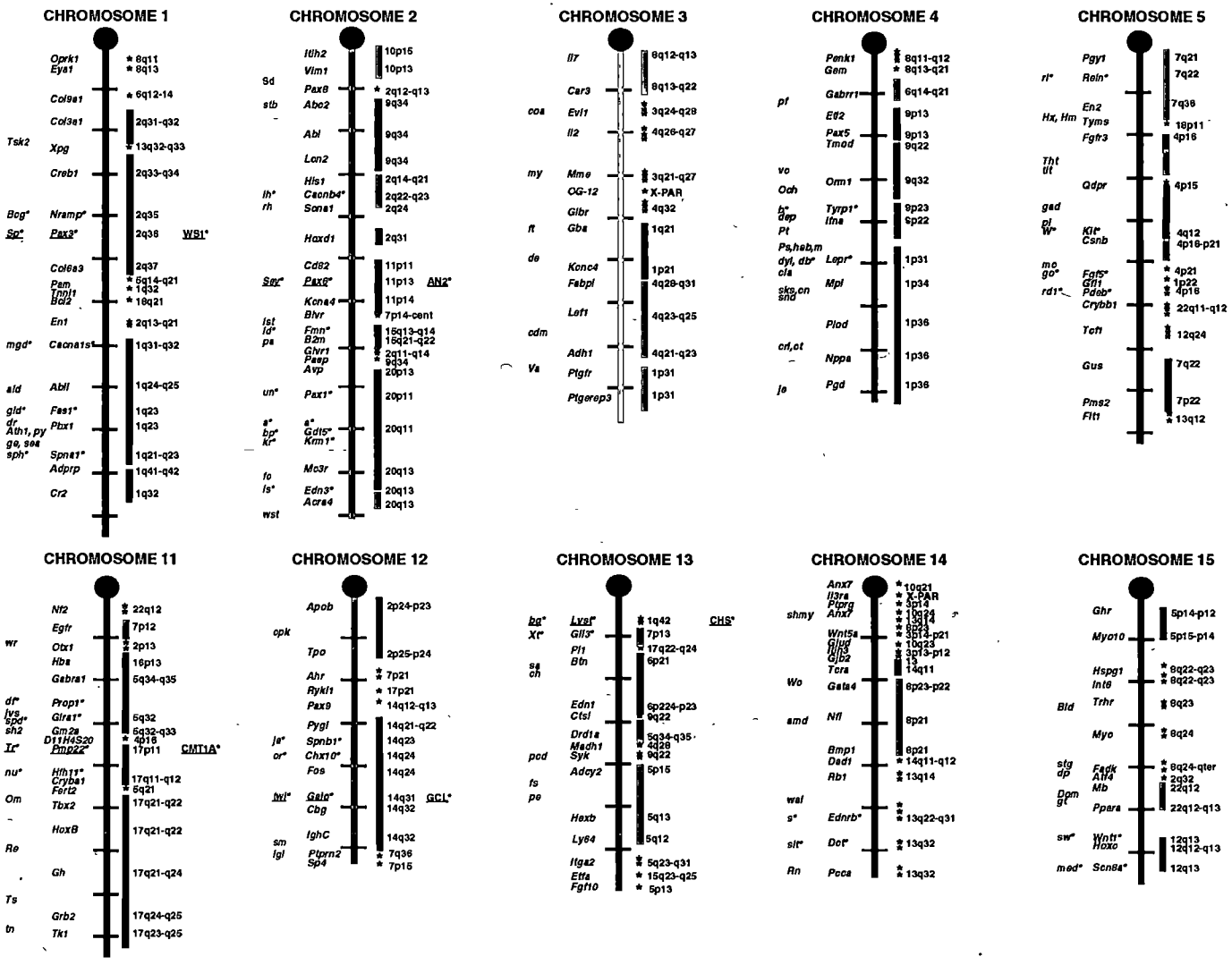
It has since become apparent that rodent chromosomes have undergone an unusually high number of genomic rearrangements per unit of evolutionary time compared to other mammals (Graves, 1996). Therefore, mouse chromosomes do not provide the simplest models for human comparison. Even so, there are considerable amounts of conservation in gene content and order, and the depth of the available maps allows good resolution of the different homology groups. Although significant conservation has been observed throughout other mammalian orders, and surprisingly to a great extent in the Japanese pufferfish (*Fugu rubripes*) and zebrafish (Schofield *et al.*, 1997; Miles *et al.*, 1998; Davidson *et al.*, 1999), no two vertebrate genomes have been studied in as much depth as those of the human and laboratory mouse.

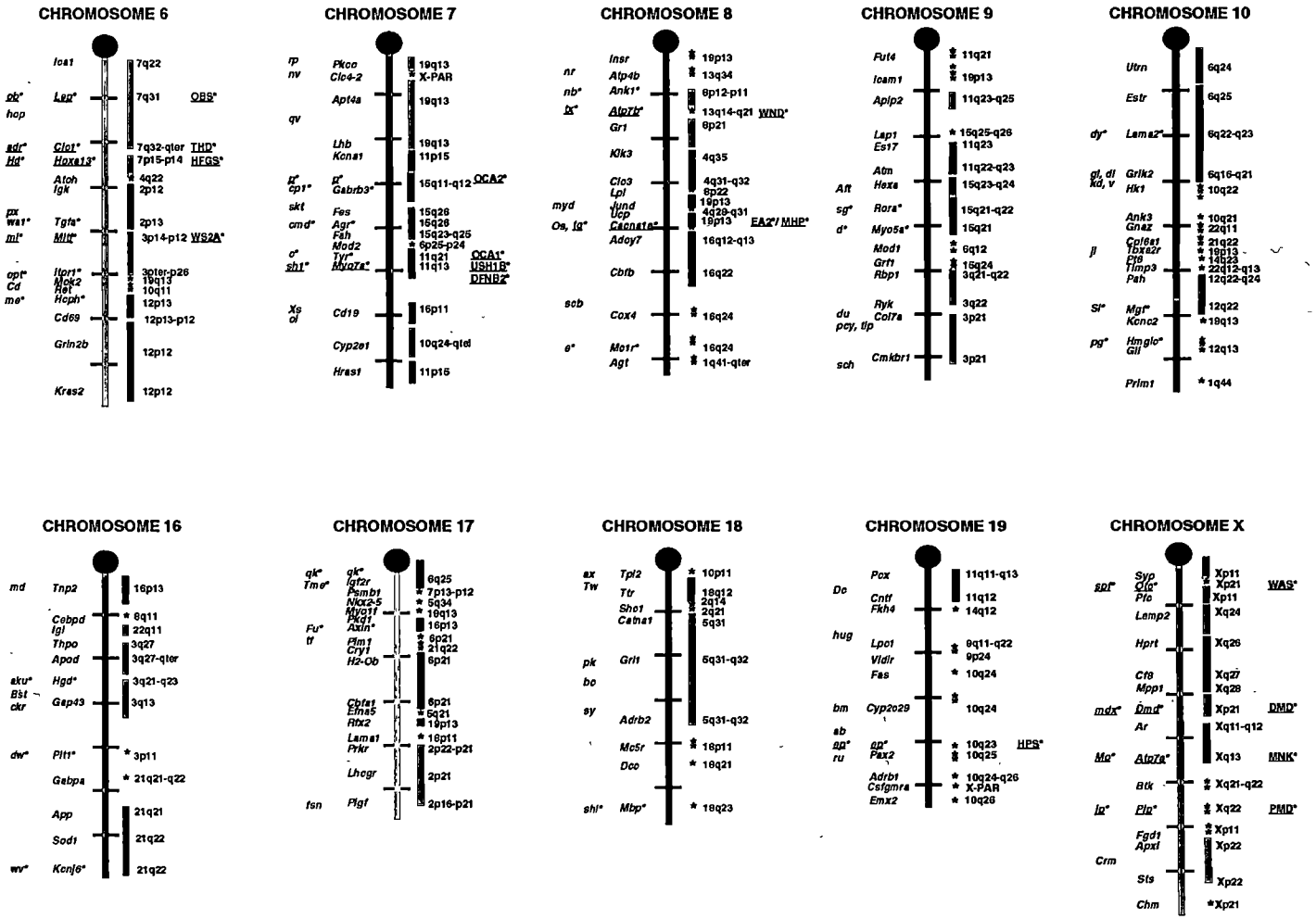
The many genetic and physical mapping studies have confirmed the idea that human and mouse chromosomes closely resemble each other within homologous segments of varying sizes (reviewed by Copeland *et al*, 1993). At this time, there are approximately 200 linkage groups conserved between the two species (Figure 3; DeBry and Seldin, 1996; Eppig and Nadeau, 1995; Carver and Stubbs, 1997).

Figure 3. Mouse-Human Comparative map.

A map of the mouse genome and corresponding human homology regions (updated from Carver and Stubbs, 1997). A map of the mouse genome is illustrated, with human homology regions indicated to the right of each map. The lengths of the chromosomes are delineated by crosshatch bars, which are spaced at 10 cM intervals along the mouse map. Human homology segments are labeled with the cytogenetic bands to which they map. Conserved markers are placed to the left of the mouse map to provide anchor points for each homologous segment. Solid bars indicate solid homology regions, defined by numerous conserved markers that have been mapped with good accuracy in both species. Hatched bars denote regions to which fewer markers have been mapped and for which human and/or mouse map assignments have been established with less accuracy. Asterisks indicate the approximate positions of single markers that have been mapped in both species. Several asterisks together indicate that a small number of markers have been mapped to that region, but those markers are known to map close together and/or are mapped at low levels of accuracy. A selection of locus symbols corresponding to mouse mutations that are known to map within a particular homology region is included at far left of the map in italics. An asterisk denotes that the human gene or mouse mutation has been cloned, while underlining indicates that human diseases and murine model, aligned in parallel positions on the map, are known to be associated with orthologous genes. Positions of additional mouse mutations, many of which are not cloned or solidly anchored to the human map, are also listed. On-line resources which aided the construction of Figure 3 included: (MGD: The Mouse Genome Database http://www.informatics.jax.org/mgd.html), (Davis Human/Mouse Homology Map http://www.ncbi.nlm.nih.gov/Homology/), and (OMIM -- Online Mendelian Inheritance in Man http://www.ncbi.nlm.nih.gov/Omim/). Information regarding phenotypes, cloned

genes, and other mutations that are not represented here can be obtained directly from these sources.


CHROMOSOME 1
Oprk1 8q11
Eya1 8q13
Col9a1 6q12-14
Col3a1 2q31-q32
Tsk2 Xpg 13q32-q33
Creb1 2q33-q34
Bcg* Nramp* 2q35
Sp* Pax3* 2q36 WS1*
Col6a3 2q37
Pam 6q14-q21
Tnnt1 1q32
Bcl2 18q21
En1 2q13-q21
mgd* Cacna1s* 1q31-q32
ald Abl1 1q24-q25
gld* Fes1* 1q23
dr Ath1, py ge, sea sph* Pbx1 1q23
Spna1* 1q21-q23
Adprp 1q41-q42
Cr2 1q32
CHROMOSOME 2
Ilh2 10p15
Vim1 10p13
Sd Pax8 2q12-q13
stb Abo2 9q34
Abl 9q34
Lcn2 9q34
His1 2q14-q21
lh* Casnb4* 2q22-q23
rh Sons1 2q24
Hoxd1 2q31
Cd82 11p11
Sey* Pax6* 11p13 AN2*
Kcna4 11p14
Bhrf 7p14-cent
Ist Id* Fmn* 15q13-q14
ps B2m 15q21-q22
Ghrl 2q11-q14
Paep 9q34
Avp 20p13
un* Pax1* 20p11
a* bp* kr* Gdf5* Krm1* 20q11
fo Mc3r 20q13
ls* Edn3* 20q13
Acra4 20q13
wst
CHROMOSOME 3
Il7 6q12-q13
8q13-q22
Car3
cos Evi1 3q24-q28
Il2 4q26-q27
my Mme 3q21-q27
OG-12 X-PAR
Glbr 4q32
ft Gba 1q21
de Konc4 1p21
Fabpl 4q28-q31
Lef1 4q23-q25
cdm
Adh1 4q21-q23
Va Ptgfr 1p31
Ptgerep3 1p31
CHROMOSOME 4
Penk1 8q11-q12
Gem 8q13-q21
Gabrr1 6q14-q21
pf
E6l2 9p13
Pax5 9p13
Tmod 9q22
vo
Och Orm1 9q32
b* dep Tyrp1* 9p23
Ifna 9p22
Pt
Ps,heb,m
dyl, dbr* Lepr* 1p31
cla
sks,cn snd Mpl 1p34
Plod 1p36
crl,ot
Nppa 1p36
je Pgd 1p36
CHROMOSOME 5
Pgy1 7q21
rl* Reln* 7q22
En2 7q36
Hx, Hm Tyms 18p11
Fgfr3 4p16
Tht
tlt
Qdpr 4p15
gad
pl W* Kit* 4q12
Csnb 4p16-p21
mo go* 4p21
Fgf5* 1p22
Gfi1
rd1* Pdeb* 4p16
Crybb1 22q11-q12
Tcfl 12q24
Gus 7q22
Pms2 7p22
Flt1 13q12
CHROMOSOME 11
Nf2 22q12
Egfr 7p12
wr Otx1 2p13
Hba 16p13
Gabra1 5q34-q35
df* Prop1* 5q32
lvs spd* Glra1* 5q32-q33
sh2 Gm2a 4p16
D11H4S20
Tr* Pmp22* 17p11 CMT1A*
nu* Hfh11* 17q11-q12
Cryba1 6q21
Om Fert2 17q21-q22
Tbx2
HoxB 17q21-q22
Re
Gh 17q21-q24
Ts
Grb2 17q24-q25
tn Tk1 17q23-q25
CHROMOSOME 12
Apob 2p24-p23
opk
Tpo 2p25-p24
Ahr 7p21
Ryk1 17p21
Pax9 14q12-q13
Pygl 14q21-q22
Je* Spnb1* 14q23
or* Chx10* 14q24
Fos 14q24
twl* Gata* 14q31 GCL*
Cbg 14q32
sm IghC 14q32
lgl Ptprn2 7q36
Sp4 7p15
CHROMOSOME 13
bg* Lyst* 1q42 CHS*
Xt* Gli3* 7p13
Plt 17q22-q24
sa Btn 6p21
ch
Edn1 6p24-p23
Ctsl 9q22
Drd1a 5q34-q35
Madh1 4q28
pod Syk 9q22
Adcy2 5p15
fs
pe
Hexb 5q13
Ly64 5q12
Itga2 5q23-q31
Etfa 15q23-q25
Fgf10 5p13
CHROMOSOME 14
Anx7 10q21
Il3ra X-PAR
Ptprg 5p14
shmy Anx7 10q24
5p14
5p13
Wnt5a 5p13-p21
Glud 10q23
Il5 3p13-p12
Gjb2 13
Tcra 14q11
Wo Gata4 6p23-p22
amd Nfl 6p21
Bmp1 8p21
Dad1 14q11-q12
Rb1 13q14
wal
s* Ednrb* 13q22-q31
slt* Dcf* 13q32
Rn Pcca 13q32
CHROMOSOME 15
Ghr 5p14-p12
Myot0 5p15-p14
Hspg1 8q22-q23
Int6 8q22-q23
Bld Trhr 8q23
Myo 8q24
stg dp Fadk 8q24-qter
Alf4 2q32
Dpm Mb 22q12
gt Ppara 22q12-q13
sw* Wnt1* 12q13
Hoxc 12q12-q13
med* Scn8a* 12q13


CHROMOSOME 6
Ica1 7q22
ok* Lez* 7q31 OBS*
hop
sdr* Clo1* 7q32-qter THD*
Hd* Hoxa13* 7p15-p14 HFGS*
Atoh 4q22
Igk 2p12
px 2p13
wa1* Tgfa*
ml* Mlt* 3p14-p12 WS2A*
opt* Itpr1* 3pter-p26
Cd Mok2 19q13
me* Ret 10q11
Hoph* 12p13
Cd69 12p13-p12
Grin2b 12p12
Kras2 12p12
CHROMOSOME 7
rp Pkcc 19q13
nv Clc4-2 X-PAR
Apt4a 19q13
qv
Lhb 19q13
Kcna1 11p15
tt* Gabrb3* 15q11-q12 OCA2*
cpt* tt*
skt Fes 15q26
cmd* Agr* 15q26
Fah 15q23-q25
Mod2 6p25-p24
c* Tyr* 11q21 OCA1*
sh1* Myo7a* 11q13 USH1B* DFNB2*
Xs Cd19 16p11
ol
Cyp2e1 10q24-qter
Hras1 11p15
CHROMOSOME 8
Insr 19p13
nr Atp4b 13q34
nb* Ank1* 8p12-p11
lx* Atp7b* 13q14-q21 WND*
Gr1 8p21
Klk3 4q35
Clo3 4q31-q32
Lpl 8p22
myd Jund 19p13
Ucp 4q28-q31
Os, lg* Cacna1a* 19p13 EA2* MHP*
Adoy7 16q12-q13
Cbfb 16q22
scb
Cox4 16q24
e* Mo1r* 16q24
Agt 1q41-qter
CHROMOSOME 9
Fut4 11q21
Icam1 19p13
Aplp2 11q23-q25
Lap1 15q25-q26
Es17 11q23
11q22-q23
Atn
Aft Hexa 15q23-q24
sg* Rora* 15q21-q22
d* Myo5a* 15q21
Mod1 6q12
Grf1 15q24
Rbp1 3q21-q22
du Ryk 3q22
pcy, tip Col7a 3p21
sch Cmkbr1 3p21
CHROMOSOME 10
Utrn 6q24
Estr 6q25
dy* Lama2* 6q22-q23
gl, dl Grik2 6q16-q21
kd, v Hk1 10q22
Ank3 10q21
Gnaz 22q11
Cpl6a1 21q22
Tbxa2r 19p13
jl Ptb 14q23
Timp3 22q12-q13
Pah 12q22-q24
Sl* Mgf* 12q22
Kcnc2 19q13
pg* Hmgic* 12q13
Gli
Prlm1 1q44
CHROMOSOME 16
md Tnp2 16p13
Cebpd 8q11
Igl 22q11
Thpo 3q27
Apod 3q27-qter
aku* Hgd* 3q21-q23
Bst
ckr Gap43 3q13
dw* Plt1* 3p11
Gabpa 21q21-q22
App 21q21
Sod1 21q22
wv* Kcnj6* 21q22
CHROMOSOME 17
qk* qk* 6q25
Tme* Igf2r 7p13-p12
Psmb1 5q34
Nlgx2-5 16q13
Myg1f 16p13
Fu* Pkd1 Axin* 8p21
tf Pim1 21q22
Cry1 8p21
H2-Ob
Cbfa1 8p21
Ets5 5q21
Rtx2 19p13
Lama1 16p11
Prkr 2p22-p21
Lhcgr 2p21
fsn Pigf 2p16-p21
CHROMOSOME 18
ax Tpi2 10p11
Tw Ttr 18q12
Sho1 2q14
Cathe1 2q21
5q31
pk Grl1 5q31-q32
bc
sy
Adrb2 5q31-q32
Mc5r 18p11
Dcc 18q21
shl* Mbp* 18q23
CHROMOSOME 19
Pcx 11q11-q13
Do Cntf 11q12
Fkh4 14q12
hug Lpo1 9q11-q22
Vldlr 9p24
Fas 10q24
bm Cyp2o29 10q24
ab
ap* 10q23 HPS*
ru Pax2 10q25
Adrb1 10q24-q26
Csfgmra X-PAR
Emx2 10q26
CHROMOSOME X
spf* Syp Xp11
Ofd* Xp21 WAS*
Pfc Xp11
Lamp2 Xq24
Hprt Xq26
Cf9 Xq27
Mpp1 Xq28
mdx* Dmd* Xp21 DMD*
Ar Xq11-q12
Ma* Atp7a* Xq13 MNK*
Btk Xq21-q22
lg* Plp* Xq22 PMD*
Fgd1 Xp11
Apxl Xp22
Crm
Sts Xp22
Chm Xp21

Even with this large number of homology groups, new ones are still being discovered as both genomes are mapped with increasing resolution (Stubbs *et al.*, 1996b; Dinulos *et al.*, 1996; Watkins-Chow *et al*, 1997; Mao *et al.*, 1998; EAC *et al.*, in preparation). However, the relationships between syntenically homologous mouse and human regions are not always as straightforward as they may seem on the grosser level. Some of these differences have been illustrated most clearly where initially large regions of conserved synteny were thought to exist. In fact, when compared at a more detailed level, these regions were found to be subject to internal rearrangements that were not previously recognized (Stubbs *et al.*, 1996b; Fernandes, *et al.*, 1998, Cole *et al.*, 1999; Carver *et al.*, in preparation). This illustrates the need for further research, as well as the understanding that present comparative maps might not be completely correct. Since synteny is used to predict disease models, search for candidate genes, and as the basis for constructing physical maps, the need to chart these small but significant rearrangements remains high.

Comparative maps have also been instrumental in the cloning of human disease genes based on the discovery of a mouse mutant or gene, and comparing its phenotype and location to the human genome (reviewed by Bedell *et al.*, 1997; Montague *et al.*, 1997). Diseases which have been cloned in this manner include piebaldism (Giebel and Spritz, 1991), Waardenburg syndrome (Tassabehji *et al.*, 1992), Chediak-Higashi syndrome (Barbosa *et al.*, 1996), and many others.

The cloning of the Chediak-Higashi syndrome and the *beige* mouse mutation is one of the best examples of the influence of comparative maps in associating mouse models with human disease, and the use of a model organism to expedite the cloning of human disease-causing genes. The features of Chediak-Higashi syndrome and *beige* mouse mutants include decreased pigmentation of hair and eyes as well as giant intracellular vesicles, which arise from lysosomes. These similar features were the first indication that the two syndromes could be related (Lutzner *et al.*, 1967). Later, the association between

the two was strengthened by the mapping of a gene from human chromosome 1q to mouse chromosome 13 closely linked to the *beige* mutant, as well as the linkage of the Chediak-Higashi syndrome to human chromosome 1q42 (Jenkins *et al.*, 1991; Fukai *et al.*, 1996). By fine mapping the interval on mouse chromosome 13, a contig of YACs was developed for the region. From this contig, a gene was isolated that was responsible for the mouse *beige* mutation (Barbosa *et al.*, 1996). The sequence data from the mouse gene was then used to analyze patients with Chediak-Higashi syndrome for mutations, and several different mutations were subsequently discovered (Barbosa *et al.*, 1996), proving that the mouse was an excellent model for the disease itself, as well as a critical tool to isolate the affected gene in humans.

The maps of the mouse and human genomes have grown in size from the initial linkage groups to the extensive physical and genetic maps of today. Consequently, the power of prediction in any homology region has increased tremendously with the levels of resolution inherent in a particular set of mapping data. One region in human chromosome 1 and related regions of the mouse genome has revealed a large region of conserved synteny in which gene order, content, and spacing are nearly identical between the two species (Oakey *et al.*, 1992). This still remains one of the best examples of such a high degree of conservation. Another well-characterized homology region is between human chromosome 21, the region implicated in Down syndrome, and corresponding regions on mouse chromosomes 10 and 16 (Cabin *et al.*, 1998; Cole *et al.*, 1999). These regions show remarkable conservation of gene order and content, and will facilitate the continuing development of mouse models of the disease, as well as the elucidation of genes which have a direct impact on the complex phenotype associated with Down syndrome.

In contrast, some regions of the human and mouse genomes are very different. Some genes have been shown to be present in one but not the other species. For example,

apolipoprotein(a) was first reported in humans (Berg, 1963), and has been seen only in Old World primates (Guo *et al.*, 1991) and one species of hedgehog (Laplaud *et al.*, 1988). The gene has not been found in New World primates, or other mammalian species, suggesting that the gene has evolved recently. In humans, increased risk for atherosclerosis and its manifestations--myocardial infarction, stroke, and restenosis--is associated with elevated plasma levels of Lp(a) (Dahlen *et al.*, 1975). Lp(a), is thought to be a surrogate for ascorbate in humans and other species that do not synthesize vitamin C (reviewed in Rath and Pauling, 1990). Since apolipoprotein(a), which is a component of Lp(a), is not present in most mammalian species, there were no convenient animal models to study the function of this gene. However, transgenic mice expressing human apolipoprotein(a) are more likely than wild-type mice to develop lipid-staining lesions in the aorta, and it was shown that apolipoprotein(a) co-localizes with this lipid deposition in the arterial walls (Lawn *et al.*, 1992). This highlights one of the differences seen between mouse and human genomes, and how some genes which can cause disease might be absent from one genome or the other. Also, the development of the transgenic mice demonstrated an elegant means to create animal models for disease genes of this type. Therefore, the study of differences between the genomes will be of great interest.

Besides the presence or absence of a gene(s) from one genome or the other, some genes have been repositioned within the genome, an event that could have a profound effect on the inheritance and function of the gene. Rearrangements involving genes located on the X chromosome had been thought to be deleterious (Ohno, 1967). Surprisingly, four genes found on the human X chromosome are located in autosomes in the mouse. These include *CSF2RA*, *IL3RA*, *CLCN4*, and *SHOX*, (also called *PHOG*), which map to autosomes Mmu19, Mmu14, Mmu7, and Mmu3, respectively in the mouse (Figure 3; Disteche *et al.*, 1992; Palmer *et al.*, 1995; Rao *et al.* 1997, Ellison *et al.*, 1997; reviewed by Blaschke and Rappold, 1997). The *Clcn4* gene is present on the X

chromosome in humans and in wild *Mus spretus* mice, but is located on chromosome 7 in the laboratory mouse, *Mus musculus* (Palmer *et al.*, 1995). This is due to a very recent transposition of the gene. Therefore, one must be aware that differences in comparison maps can also occur, not just between distantly related species such as human and mouse, but among closely related species as well.

Comparative maps have been used for a variety of positional cloning projects, have validated mouse models for human disease, and have provided the first glimpse into the genome evolution that has taken place between these two mammalian species. Although very few physically mapped mouse and human regions have been compared, the existing studies provide intriguing insights to genome evolution. In addition to rearrangements, recent duplications and deletions of genes could define some of the differences between the two species. Genes that have been added or deleted as mice and humans have evolved are likely to have played an important role in the bifurcation and subsequent evolution of these two species. Therefore, further resolution of the comparative maps is still important; the ultimate resolution will be the comparison of the two completely sequenced genomes. The next two chapters rely on comparative mapping, both to increase our understanding of the comparative maps themselves by examining regions at greater resolution, and to apply these maps to a positional cloning project to localize a translocation breakpoint in mice which display a neurological defect.

REFERENCES

Ashworth, L.K., Batzer, M.A., Brandriff, B., Branscomb, E., deJong, P., Garcia, E., Garnes, J., Gordon, L., Lamerdin, J.E., Lennon, G., Mohrenweiser, H., Olsen, A., Slezak, T., and Carrano, A.V. (1995). A metric physical map of human chromosome 19. *Nature Genet* **11**:422-427.

Arnheim, N., Treco, D., Taylor, B., and Eicher, E.M. (1982). Distribution of ribosomal gene length variants among mouse chromosomes. *Proc Natl Acad Sci* **79**:4677-4680.

Avery, O.T., MacCleod, C.M., and MacCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types. *J Exp Med* **79**:137-158.

Avner, P., Amar, L., Dandolo, L., and Guenet, J.L. (1988). Genetic analysis of the mouse using interspecific crosses. *Trends Genet* **4**:18-23.

Barbosa, M.D., Nguyen, Q.A., Tchernev, V.T., Ashley, J.A., Detter, J.C., Blaydes, S.M., Brandt, S.J., Chotai, D., Hodgman, C., Solari, R.C., Lovett, M., and Kingsmore, S.F. (1996). Identification of the homologous beige and Chediak-Higashi syndrome genes. *Nature* **382**:262-265.

Bedell, M.A., Largaespada, D.A., Jenkins, N.A., and Copeland, N.G. (1997). Mouse models of human disease. Part II: recent progress and future directions. *Genes Dev* **11**:11-43.

Bennett, K.L., Lalley, P.A., Barth, R.K., and Hastie, N.D. (1982). Mapping the structural genes coding for the major urinary proteins in the mouse: combined use of recombinant inbred strains and somatic cell hybrids. *Proc Natl Acad Sci* **79**:1220-1224.

Bentley, D.R., Pruitt, K.D., Deloukas, P., Schuler, G.D., and Ostell, J. (1998). Coordination of human genome sequencing via a consensus framework map. *Trends Genet* **14**:381-384.

Berg, K., (1963). A new serum type system in man; the Lp system. *Acta Pathol Microbiol Scand* **59**:369-382.

Blaschke, R. J., and Rappold, G.A. (1997). Man to Mouse - Lessons learned from the distal end of the human X chromosome. *Genome Res* **7**:1114-1117.

Botstein, D., White, R.L., Skolnick, M., and Davis, R.W. (1980). Construction of a human linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet* **32**:314-331.

Brandriff, B.F., Gordon, L.A., and Trask, B.J. (1991). A new system for high- resolution DNA sequence mapping in interphase pronuclei. *Genomics* **10**:75-82.

Cabin, D.E., McKee-Johnson, J.W., Matesic, L.E., Wiltshire, T., Rue, E.E., Mjaatvedt, A.E., Huo, Y.K., Korenberg, J.R., and Reeves, R.H. (1998). Physical and comparative mapping of distal mouse chromosome 16. 5 p5. *Genome Res* **8**:940-950.

Carrano, A.V., Lamerdin, J., Ashworth, L.K., Watkins, B., Branscomb, E., Slezak, T., Raff, M., de Jong, P.J., Keith, D., McBride, L., Meister, K.S., and Kronick, M. (1989). A high-resolution, fluorescence-based, semiautomated method for DNA fingerprinting. *Genomics* **4**:129-136.

Carver, E. A., Issel-Tarver, L., Rine, J., Olsen, A., and Stubbs, L. (1998). Location of mouse and human genes corresponding to conserved canine olfactory receptor gene subfamilies. *Mammal Genome* **9**:349-354.

Carver, E. A., and Stubbs, L. (1997). Zooming in on the Human-Mouse Comparative Map: Genome Conservation Re-examined on a High-Resolution Scale. *Genome Res* **7**:1123-1137.

Carver, E.A., Doyle, J., Olsen, A.S., Ashworth, L., and Stubbs, L. (1999). Comparative mapping of human 19p13.1 surrounding the NOTCH3 gene and olfactory gene clusters: delineation of homology groups on mouse chromosomes 8, 10, and 17. Manuscript in preparation.

Castle, W.E. (1903). Mendel's law of heredity. *Proc Am Acad Arts Sci* **38**:535-548.

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

Cole, S.E., Wiltshire, T., Rue, E.E., Morrow, D., Hieter, P., Brahe, C., Fisher, E.M., Katsanis, N., Reeves, R.H. (1999). High-resolution comparative physical mapping of

mouse chromosome 10 in the region of homology with human chromosome 21. *Mamm Genome* **10**:229-234.

Collins, F.S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., Walters, L. and the members of the DOE and NIH planning groups. (1998). New goals for the U.S. Human Genome Project: 1998-2003. *Science* **282**:682-689.

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

Cuénot, L. (1905). Les races pures et leur combinaisons chez les souris. *Arch Zool Exp* **3**:223-232.

Dahlen, G., Berg, K., Gillnas, T., and Ericson, C. (1975). Lp(a) lipoprotein/pre-beta1-lipoprotein in Swedish middle-aged males and in patients with coronary heart disease. *Clin Genet* **7**:334-341.

Davidson, A.J., Postlethwait, J.H., Yan, Y.L., Beier, D.R., van Doren, C., Foernzler, D., Celeste, A.J., Crosier, K.E., and Crosier, P.S. (1999). Isolation of zebrafish gdf7 and comparative genetic mapping of genes belonging to the Growth/Differentiation factor 5, 6, 7 subgroup of the TGF-beta superfamily. *Genome Res* **9**:121-129.

DeBry, R.W., and Seldin, M.F. (1996). Human/mouse homology relationships. *Genomics* **33**:337-351.

Dietrich, W.F., Copeland, N.G., Gilbert, D.J., Miller, J.C., Jenkins, N.A., and Lander, E.S. (1995). Mapping the mouse genome: current status and future prospects. *Proc Natl Acad Sci* **92**:10849-10853.

Dietrich, W.F., Miller, J., Steen, R., Merchant, M.A., Damron-Boles, D., Husain, Z., Dredge, R., Daly, M.J., Ingalls, K.A., O'Connor, T.J., Evans, C.A., DeAngelis, M.M., Levinson, D.M., Kruglyak, L., Goodman, N., Copeland, N.G., Jenkins, N.A., Hawkins, T.L., Stein, L., Page, D.C., and Lander, E.S. (1996). A comprehensive genetic map of the mouse genome. *Nature* **380**:149-152.

Disteche, C.M., Brannan, C.I., Larsen, A., Adler, D.A., Schorderet, D.F., Gearing, D., Copeland, N.G., Jenkins, N.A., and Park, L.S. (1992). The human pseudoautosomal GM-CSF receptor alpha subunit gene is autosomal in mouse. *Nature Genet* **1**:333-336.

Dinulos, M.B., Bassi, M.T., Rugarli, E.I., Chapman, V., Ballabio, A., and Disteche, C.M. (1996). A new region of conservation is defined between human and mouse X chromosomes. *Genomics* **35**:244-247.

Donahue, R.P., Bias, W.B., Renwick, J.H., and McKusick, V.A. (1968). Probable assignment of the Duffy blood group locus to chromosome 1 in man. *Proc Natl Acad Sci* **61**:949-955.

Doyle J., Hoffman S., Ucla C., Reith W., Mach B., and Stubbs L. (1996). Locations of human and mouse genes encoding the RFX1 and RFX2 transcription factor proteins. *Genomics* **35**:227-230.

Eicher, E.M. (1971). The identification of the chromosome bearing linkage group XII in the mouse. *Genetics* **69**:267-271.

Ellison, J.W., Wardak, Z., Young, M.F., Robey, P.G., Webster, M., and Chiong, W., (1997). PHOG, a candidate gene for involvement in the short stature of Turner syndrome. *Hum Mol Genet* **8**:1341-1347.

Eppig, J.T., and Nadeau, J.H. (1995). Comparative maps: the mammalian jigsaw puzzle. *Curr Opin Genet Dev* **5**:709-716.

Fernandes, M., Lespinasse, F., Rotomondo, F., Poirier, C., Guenet, J.L., Gaudray, P., and Carle, G.F. (1998). Comparative mapping of two adjacent regions of MMU19 with their human counterpart on HSA11q13. *Cytogenet Cell Genet* **81**:237-246.

Ford, C.E., and Hamerton, J.L. (1956). The chromosomes of man. *Nature* **178**:1020-1023.

Frydman, M., Bonne-Tamir, B., Farrer, L. A., Conneally, P. M., Magazanik, A., Ashbel, S., and Goldwitch, Z. (1985). Assignment of the gene for Wilson disease to chromosome 13: linkage to the esterase D locus. *Proc Nat Acad Sci* **82**:1819-1821.

Fukai, K., Oh, J., Karim, M.A., Moore, K.J., Kandil, H.H., Ito, H., Burger, J., and Spritz, R.A. (1996). Homozygosity mapping of the gene for Chediak-Higashi syndrome to chromosome 1q42-q44 in a segment of conserved synteny that includes the mouse beige locus (bg). *Am J Hum Genet* **59**:620-624.

Giebel, L.B., and Spritz, R.A. (1991). Mutation of the KIT (mast/stem cell growth factor receptor) protooncogene in human piebaldism. *Proc Natl Acad Sci* **88**:8696-8699.

Guo, H.C., Michel, J.B., Blouquit, Y., and Chapman, M.J. (1991). Lipoprotein(a) and apolipoprotein(a) in a New World monkey, the common marmoset (Callithrix jacchus). Association of variable plasma lipoprotein(a) levels with a single apolipoprotein(a) isoform. *Arterioscler Thromb* **11**:1030-1041.

Graves, J.A. (1996). Mammals that break the rules: genetics of marsupials and monotremes. *Annu Rev Genet* **30**:233-260.

Green, M.C. (1971). Linkage map of the mouse. In Whitten, W.K.(ed): 42nd Annual Report. Bar Harbor, ME, The Jackson Laboratory, p56.

Haldane, J.B.S. (1927). The comparative genetics of color in rodents and carnivora. *Biol Rev Camb Philos Soc* **2**:199-212.

Haldane, J.B.S., Sprunt, A.D., and Haldane, N.M. (1915). Reduplication in mice. *J Genet* **5**:132-135.

Hershey, A.D., and Chase, M. (1952). Independent functions of viral protein and nucleic acid in growth of bacteria. *J Gen Physiol* **36**:39-56.

Hodge, S. E., Spence, M. A., Crandall, B. F., Sparkes, R. S., Sparkes, M. C., Crist, M., and Tideman, S. (1980). Huntington disease: linkage analysis with age-of-onset corrections. *Am J Med Genet* **5**:247-254.

Jenkins, N. A., Justice, M. J., Gilbert, D. J., Chu, M.-L., and Copeland, N. G. (1991). Nidogen/entactin (Nid) maps to the proximal end of mouse chromosome 13 linked to beige (bg) and identifies a new region of homology between mouse and human chromosomes. *Genomics* **9**:401-403.

Johnson, D.K. (1990). Reverse genetics and genes that control development: initial molecular characterization of two deletion complexes in mouse Chromosome 7 by random-clone mapping and genetic linkage analysis. Doctoral Thesis, University of Tennessee, Knoxville, TN.

Kan, Y.W., and Dozy, A.M. (1978). Polymorphism of DNA sequence adjacent to human beta-globin structural gene: relationship to sickle mutation. *Proc Natl Acad Sci* **75**:5631-5635.

Kim, J., Ashworth, L., Branscomb, E., and Stubbs, L. (1997). The human homolog of a mouse-imprinted gene, Peg3, maps to a zinc finger gene-rich region of human chromosome 19q13.4. *Genome Res* **7**:532-540.

Lalley, P.A., Minna, J.D., and Francke, U. (1978). Conservation of autosomal gene synteny groups in mouse and man. *Nature* **274**:160-163.

Laplaud, P.M. Beaubatie, L., Rall, S.C. Jr, Luc, G., and Saboureau, M. (1988). Lipoprotein[a] is the major apoB-containing lipoprotein in the plasma of a hibernator, the hedgehog (Erinaceus europaeus). *J Lipid Res* **29**:1157-1170.

Lawn, R.M., Wade, D.P., Hammer, R.E., Chiesa, G., Verstuyft, J.G., and Rubin, E.M. (1992). Atherogenesis in transgenic mice expressing human apolipoprotein. *Nature* **360**:670-672.

Lutzner, M.A., Lowrie, C.T., and Jordan, H.W. (1967). Giant granules in leukocytes of the beige mouse. *J Hered* **58**:299-300.

Manly, K.F. (1993). A MacIntosh program for the storage and analysis of experimental genetic mapping data. *Mammal Genome* **4**:303-313.

Mao, N-C., Steingrimsson, E., DuHadaway, J., Wasserman, W., Ruiz, J.C., Copeland, N.G., Jenkins, N.A., and Prendergast, G.C. (1999). The Murine Bin1 Gene Functions Early in Myogenesis and Defines a New Region of Synteny between Mouse Chromosome 18 and Human Chromosome 2. *Genomics* **56**:51-58.

McKusick, V.A. (1991). Current trends in mapping human genes. *FASEB J* **5**:12-20.

Miles, C., Elgar, G., Coles, E., Kleinjan, D.J., van Heyningen, V., and Hastie, N. (1998). Complete sequencing of the Fugu WAGR region from WT1 to PAX6: dramatic compaction and conservation of synteny with human chromosome 11p13. *Proc Natl Acad Sci* **95**:13068-13072.

Minna, J.D., Lalley, P.A., and Francke, U. (1976). Comparative mapping using somatic cell hybrids. *In Vitro* **12**:726-733.

Mohr, J. (1951). A search for linkage between the Lutheran blood group and other hereditary characters. *Acta Pathol Microbiol Scand* **28**:207-210.

Monaco, A.P., Neve, R.L., Colletti-Feener, C., Bertelson, C.J., Kurnit, D.M., and Kunkel, L.M. (1986). Isolation of candidate cDNAs for portions of the muscular dystrophy gene. *Nature* **322**:646-650.

Montague, C.T., Farooqi, I.S., Whitehead, J.P., Soos, M.A., Rau, H., Wareham, N. J., Sewter, C.P., Digby, J. E., Mohammed, S. N., Hurst, J. A., Cheetham, C. H., Earley, A. R., Barnett, A. H. , Prins, J. B., and O'Rahilly, S. (1997). Congenital leptin deficiency is associated with severe early-onset obesity in humans. *Nature* **387**:903-908.

Murray, J. M., Davies, K. E., Harper, P. S., Meredith, L., Mueller, C. R., and Williamson, R. (1982). Linkage relationship of a cloned DNA sequence on the short arm of the X chromosome to Duchenne muscular dystrophy. *Nature* **300**:69-71.

Nadeau, J.H., and Taylor, B.A. (1984). Lengths of chromosomal segments conserved since divergence of man and mouse. *Proc Natl Acad Sci* **81**:814-818.

Oakey, R.J., Watson, M.L., and Seldin, M.F. (1992). Construction of a physical map on mouse and human chromosome 1: comparison of 13 Mb of mouse and 11 Mb of human DNA. *Human Mol Genet* **1**:613-620.

Ohno, S. (1967). Sex chromosomes and sex -linked genes. Berlin, New York; Springer Verlag.

Palmer, S., Perry, J., and Ashworth, A. (1995). A contravention of Ohno's law in mice. *Nature Genet* **10**:472-476.

Rao, E., Weiss, B., Fukami, M., Rump, A., Niesler, B., Mertz, A., Muroya, K., Binder, G., Kirsch, S., Winkelmann, M., Nordsiek, G., Heinrich, U., Breuning, M.H., Ranke, R.B., Rosenthal, A., Ogata, T., and Rappold, G.A. (1997). Pseudoautosomal deletions encompassing a novel homeobox gene cause growth failure in idiopathic short stature and Turner syndrome. *Nature Genet* **16**:54-63.

Rath, M., and Pauling, L. (1990). Hypothesis: lipoprotein(a) is a surrogate for ascorbate. *Proc Natl Acad Sci* **87**:6204-6207.

Ruddle, F.H. (1972). Linkage analysis using somatic cell hybrids. *Adv Hum Genet* **30**:173-235.

Ruddle, F.H. (1981). A new era in mammalian gene mapping: somatic cell genetics and recombinant DNA methodologies. *Nature* **294**:115-120.

Schofield, J.P., Elgar, G., Greystrong, J., Lye, G., Deadman, R., Micklem, G., King, A., Brenner, S., and Vaudin, M. (1997). Regions of human chromosome 2 (2q32-q35) and mouse chromosome 1 show synteny with the pufferfish genome (Fugu rubripes). *Genomics* **45**:158-167.

Schuler, G.D., Boguski, M.S., Stewart, E.A., Stein, L.D., Gyapay, G., Rice, K., White, R.E., Rodriguez-Tome, P., Aggarwal, A., Bajorek, E., Bentolila, S., Birren, B.B., Butler, A., Castle, A.B., Chiannilkulchai, N., Chu, A., Clee, C., Cowles, S., Day, P.J., Dibling, T., Drouot, N., Dunham, I., Duprat, S., East, C., Edwards, C., Fan, J-B., Fang, N., Fizames, C., Garrett, C., Green, L., Hadley, D., Harris, M., Harrison, P., Brady, S., Hicks, A., Holloway, E., Hui, L., Hussain, S., Louis-Dit -Sully, C., Ma, J., MacGilvery, A., Mader, C., Maratukulam, A., Matise, T.C., McKusik, K.B., Morissette, J., Mungall, A., Muselet, P., Nusbaum, H.C., Page, D.C., Peck, A., Perkins, S., Piercy, M., Qin, F., Quackenbush, J., Randby, S., Rief, T., Rozen, S., Sanders, C., She, X., Silva, J., Slonim, D.K., Soderlund, C., Sun, W-L., Taber, P., Thangarajah, T., Vega-Czarny, N., Vollrath, D., Voyticky, S., Wilmer, T., Wu, X., Adams, M.D., Auffray, C., Walter, N.A.R., Brandon, R., Dehejia, A., Goodfellow, P.N., Houlgatte, R., Hudson Jr., J.R., Ide, S.E., Iorio, K.R., Lee, W.Y., Seki, N., Nagase, T., Ishikawa, K., Nomura, N., Phillips, C., Polymeropoulos, M.H., Sandusky, M., Schmitt, K., Berry, R., Swanson, K., Torres, R., Venter, J.C., Sikela, J.M., Beckman, J.S., Weissenback, J., Myers, R.M., Cox, D.R., James, M.R., Bentley, D., Deloukas, P., Lander, E.S., and Hudson, T.J. (1996). A gene map of the human genome. *Science* **274**:540-546.

Shannon, M., Ashworth, L.K., Mucenski, M.L., Lamerdin, J.E., Branscomb, E., and Stubbs, L. (1996). Comparative analysis of a conserved zinc finger gene cluster on human chromosome 19q and mouse chromosome 7. *Genomics* **33**:112-120.

Silver, J. (1985). Confidence limits for estimates of gene linkage analysis of recombinant inbred strains. *J Heredity* **76**:436-440.

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

Stubbs, L., Carver, E., Ashworth, L., Lopez-Molina, L. (1996a). Location of the DBP transcription factor gene in man and mouse. *Mammal Genome* **7**:65-67.

Stubbs, L., Carver, E.A., Shannon, M.E., Kim, J., Geisler, J., Generoso, E.E., Stanford. B.G., Dunn, W.C., Mohrenweiser, H., Zimmermann, W., Watt, S.E., and Ashworth, L.K. (1996b). Detailed Comparative Map of Human Chromosome 19q and Related Regions of the Mouse Genome. *Genomics* **35**:499-508.

Tassabehji, M., Read, A.P., Newton, V.E., Harris, R., Balling, R., Gruss, P., and Strachan, T. (1992). Waardenburg's syndrome patients have mutations in the human homologue of the Pax-3 paired box gene. *Nature* **355**:635-636.

Tijio, J.H., and Levan, A. (1956). The chromosome numbers of man. *Hereditas* **42**:1-6.

The Genetic Code. (1967). *Cold Spring Harbor Symp. Quant. Biol.*, vol. 31. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Trask, B., Fertitta, A., Christensen, M., Youngblom, J., Bergmann, A., Copeland, A., de Jong, P., Mohrenweiser, H., Olsen, A., Carrano, A., and Tynan, K. (1993). Fluorescence in situ hybridization mapping of human chromosome 19: cytogenetic band location of 540 cosmids and 70 genes or DNA markers. *Genomics* **15**:133-145.

Watkins-Chow, D. E., Buckwalter, M.S., Newhouse, M.M., Lossie, A.C., Brinkmeier, M.L. and Camper, S.A. (1997). Genetic mapping of 21 genes on mouse chromosome 11 reveals disruptions in linkage conservation with human chromosome 5. *Genomics* **40**:114-122.

Watson, J.D., and Crick, F.H.C. (1953). Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. *Nature* **171**:737-738.

Weiss, M.C., and Green, H. (1967). Human-mouse hybrid cell lines containing partial complements of human chromosomes and functioning human genes. *Proc Natl Acad Sci* **58**:1104-1111.

Wilson, E.B. (1911). The sex chromosomes. *Arch Mikrosk Anat Entwicklunsmech* **77**:149-271.

PART II

COMPARATIVE MAPPING OF HUMAN CHROMOSOME 19p AND RELATED REGIONS OF THE MOUSE GENOME

CHAPTER 1.
INTRODUCTION

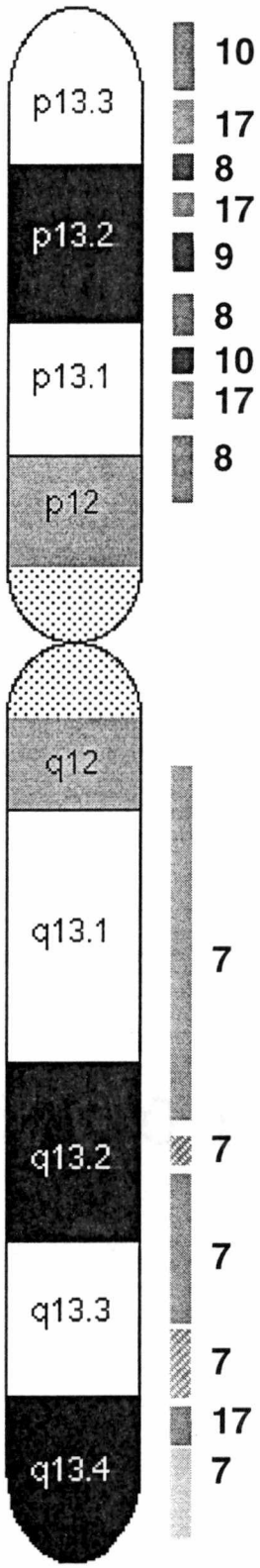
In this section, we use a mouse genetic map and a human physical map, as discussed in Part I, to make comparisons between human chromosome 19 and related regions of the mouse genome. The purpose of these studies was to identify and explore homology regions between human chromosome 19 and the mouse. Human chromosome 19 has homology regions to 5 different mouse chromosomes (Figure 1). Comparative mapping studies have occurred on both arms of the chromosome (Doyle *et al.*, 1996; Doyle *et al.*, 1997; Stubbs *et al.*, 1996a and b; Shannon *et al.*, 1996; Kim *et al.*, 1997; Carver *et al.*, 1998). The q-arm of human chromosome 19 has extensive homology to mouse chromosome 7 and a small region of homology to mouse chromosome 17 (Figure 1). The region of Chr 19q/Mmu7 homology is fairly well conserved, with large blocks of conserved synteny existing between the two species, despite some rearrangements (Stubbs *et al.*, 1996b; Kim *et al.*, 1997). The p-arm of human chromosome 19 has undergone many more evolutionary rearrangements, as indicated by the multiple regions of mouse-human homology that are present (Figure 1). The region of 19p13.1 was of particular interest because it contains genes associated with human disease, including CADASIL, episodic ataxia, familial hemiplegic migraine, in addition to several different types of gene families (Ashworth *et al.*, 1995; Joutel *et al.*, 1996; Ophoff *et al.*, 1997). Therefore, the comparative mapping of this region could give better insight into mouse disease models, and chromosomal evolution.

The ability to link mouse mutations accurately to the human 19p-homology regions, as well as the human diseases to the mouse, depends on the development of a very precise comparative map.

Figure 1. Human chromosome 19 and related regions of the mouse genome.

This figure depicts the cytogenetic band pattern of human chromosome 19 and related regions of the mouse genome, and illustrates the amount of divergence that has occurred between the mouse and corresponding human regions during evolution. The left side of the figure is a drawing of the karyotype of human chromosome 19 with the different regions of the chromosome labeled. To the right are the corresponding mouse homology regions with the mouse chromosomes indicated to the far right.

CHROMOSOME 19



Initial experiments seemed to indicate that this region had solid homology with mouse chromosome eight, near known mouse mutations such as oligosyndactylism and tottering (Davidson *et al.*, 1989; Dietrich *et al.*, 1996; Doyle *et al.*, 1997). However, preliminary data suggested that extensive rearrangements distinguished this human region from counterparts in mouse. Comparative mapping within an ~2Mb interval of human chromosome 19p13.1 has defined the borders of four distinct homology regions. We have extended, and added new genes to a recently discovered segment of homology with mouse chromosome 17 (Gao *et al.*, 1998). Also, we established a new homology group from human 19p13.1 to mouse chromosome 10. Our data show that the rearrangements have occurred at or near clustered gene families. These sections are a collection of papers that have been published, submitted for publication, or are in preparation, which describe the comparative mapping of human chromosome 19 between mouse and man, and illustrates a correlation between homology region borders and the localization of family gene clusters (reviewed in Carver and Stubbs, 1997).

REFERENCES

Ashworth, L.K., Batzer, M.A., Brandriff, B., Branscomb, E., deJong, P., Garcia, E., Garnes, J., Gordon, L., Lamerdin, J.E., Lennon, G., Mohrenweiser, H., Olsen, A., Slezak, T., and Carrano, A.V. (1995). A metric physical map of human chromosome 19. *Nature Genet* **11**:422-427.

Carver, E. A., Issel-Tarver, L., Rine, J., Olsen, A., and Stubbs, L. (1998). Location of mouse and human genes corresponding to conserved canine olfactory receptor gene subfamilies. *Mammal Genome* **9**:349-354.

Carver, E.A., and L. Stubbs. (1997). Zooming in on the human-mouse comparative map: genome conservation re-examined on a high-resolution scale. *Genome Res* **7**:1123-1137.

Davisson, M.T., Fox, R.R., Roderick, T.H., Sweet, H.O., and Spencer, C. (1989). Esterase-1 tightly linked to Os. *Mouse News Lett* **84**:90.

Dietrich, W.F., Miller, J., Steen, R., Merchant, M.A., Damron-Boles, D., Husain, Z., Dredge, R., Daly, M.J., Ingalls, K.A., O'Connor, T.J., Evans, C.A., DeAngelis, M.M., Levinson, D.M., Kruglyak, L., Goodman, N., Copeland, N.G., Jenkins, N.A., Hawkins, T.L., Stein, L., Page, D.C., and Lander, E.S. (1996). A comprehensive genetic map of the mouse genome. *Nature* **380**:149-152.

Doyle, J., Hoffman, S., Ucla, C., Reith, W., Mach, B., and Stubbs, L. (1996). Locations of human and mouse genes encoding the RFX1 and RFX2 transcription factor proteins. *Genomics* **35**:227-230.

Doyle, J., Ren, X., Lennon, G., and Stubbs, L. (1997). Mutations in the Cacnl1a4 calcium channel gene are associated with seizures, cerebellar degeneration, and ataxia in tottering and leaner mutant mice. *Mammal Genome* **8**:113-20.

Gao, X., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., and Gridley, T. (1998). Assignment of the murine Notch2 and Notch3 genes to chromosomes 3 and 17. *Genomics* **49**:160-161.

Joutel, A., Corpechot, C., Ducros, A., Vahedi, K., Chabriat, H., Mouton, P., Alamowitch, S., Domenga, V., Cecillion, M., Marechal, E., Maciazek, J., Vayssiere, C., Cruaud, C., Cabanis, E-A., Ruchoux, M., Weissenbach, J., Bach, J., Bousser M., and Tourneir-Lasserve, E. (1996). Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383**:707-710.

Kim, J., Ashworth, L., Branscomb, E., and Stubbs, L. (1997). The human homolog of a mouse-imprinted gene, Peg3, maps to a zinc finger gene-rich region of human chromosome 19q13.4. *Genome Res* **7**:532-540.

Ophoff, R.A., Terwindt, G.M., Vergouwe, M.N., van Eijk, R., Oefner, P.J., Hoffman, S.M., Lamerdin, J.E., Mohrenweiser, H.W., Bulman, D.E., Ferrari, M., Haan, J., Lindhout, D., van Ommen, G.J., Hofker, M.H., Ferrari, M.D., and Frants, R.R. (1996).

Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. *Cell* **87**:543-52.

Shannon, M., Ashworth, L.K., Mucenski, M.L., Lamerdin, J.E., Branscomb, E., and Stubbs, L. (1996). Comparative analysis of a conserved zinc finger gene cluster on human chromosome 19q and mouse chromosome 7. *Genomics* **33**:112-120.

Stubbs, L., Carver, E., Ashworth, L., Lopez-Molina, L. (1996a). Location of the DBP transcription factor gene in man and mouse. *Mammal Genome* **7**:65-67.

Stubbs, L., Carver, E.A., Shannon, M.E., Kim, J., Geisler, J., Generoso, E.E., Stanford, B.G., Dunn, W.C., Mohrenweiser, H., Zimmermann, W., Watt, S.E., and Ashworth, L.K. (1996b). Detailed comparative map of human chromosome 19q and related regions of the mouse genome. *Genomics* **35**:499-508.

CHAPTER 2. PUBLICATIONS DESCRIBING THE COMPARATIVE MAPPING OF HUMAN CHROMOSOME 19p AND RELATED REGIONS OF THE MOUSE GENOME

SECTION 1. LOCATION OF MOUSE AND HUMAN GENES CORRESPONDING TO CONSERVED CANINE OLFACTORY RECEPTOR GENE SUBFAMILIES

LOCATION OF MOUSE AND HUMAN GENES CORRESPONDING TO CONSERVED CANINE OLFACTORY RECEPTOR GENE SUBFAMILIES

[1,2,6]Ethan A. Carver, [3,5] Laurie Issel-Tarver, [3]Jasper Rine, [4]Anne S. Olsen, and [1,6,7]Lisa Stubbs

[1]Life Sciences Division, Oak Ridge National Laboratory, P.O. Box 2009, Oak Ridge, TN 37831-8077, USA

[2]University of Tennessee-Oak Ridge School of Biomedical Sciences, Oak Ridge National Laboratory, P.O. Box 2009, Oak Ridge, TN 37831-8077, USA

[3]University of California, Division of Genetics, Department of Molecular and Cell Biology, Berkeley, California 94720, USA

[4]Human Genome Center, Lawrence Livermore National Laboratory, P.O. Box 808, L-452, Livermore, CA 94550, USA

[5]Present Address: Acacia Biosciences, 4136 Lakeside Dr., Richmond, CA 94806

[6]Present Address: Human Genome Center, Lawrence Livermore National Laboratory, P.O. Box 808, L-452, Livermore, CA 94550, USA

[7]Corresponding author

tel. (925) 422-8473

fax. (925) 422-2282

email: Stubbs5@llnl.gov

Published in *Mammalian Genome* (1998) May; 9(5): 349-354.

ABSTRACT

Olfactory receptors are G protein-coupled, seven transmembrane domain proteins that are responsible for binding odorants in the nasal epithelium. They are encoded by a large gene family, members of which are organized in several clusters scattered throughout the genomes of mammalian species. Here we describe the mapping of mouse sequences corresponding to four conserved olfactory receptor genes, each representing separate, recently identified canine subfamilies. Three of the four canine genes detected related gene clusters in regions of mouse Chromosomes (Chrs) 2, 9, and 10, near previously mapped mouse olfactory genes, while one detected a formerly unidentified gene cluster located on mouse Chr 6. In addition, we have localized two human gene clusters with homology to the canine gene, CfOLF4, within the established physical map of Chromosome 19p. Combined with recently published studies, these data link the four conserved olfactory gene subfamilies to homologous regions of the human, dog and mouse genomes.

INTRODUCTION

Olfaction is very important to animals, playing a central role in food acquisition, social interactions, and evaluation of environmental cues. The sense of smell is thought to be triggered by the binding of a distinct odorant to a specialized receptor molecule in the nasal epithelium. Olfactory receptor (OLFR) genes comprise a large multigene family, each member of which encodes a seven transmembrane domain protein coupled to a G-protein (Buck and Axel, 1991). Genes representing four canine olfactory receptor gene subfamilies have recently been isolated and characterized (Issel-Tarver and Rine, 1996). Subsequent studies localized human orthologs of the four canine OLFR genes to chromosomes 7, 11, and 19 (Issel-Tarver and Rine, 1997).

In this paper, we describe the genetic localization of counterparts of these canine olfactory receptor gene subfamilies in the mouse, and define the position of two related human subfamilies within the established physical map of Chromosome 19p. These studies identified murine genes for CfOLF1 and CfOLF2 closely linked on mouse Chromosome 2 (Mmu2). In addition, canine CfOLF3 subfamily-specific sequences detected a new OLFR gene-containing region on Mmu6. In contrast to other canine OLFR gene probes, CfOLF4 subfamily probes detected a relatively large number of widely dispersed loci in the mouse. Two CfOLF4-related loci were assigned to proximal Mmu9 and central Mmu10. The proximal Mmu9 location, within an established region of 19p13.2/Mmu9 homology, correlates well with the position determined in this study to contain one group of human CfOLF4-related genes. A second cluster of CfOLF4-related genes was assigned to 19p13.1, but the position of potential mouse orthologs of this more proximal cluster of human OLFR genes could not be established with certainty.

Comparison between canine, human, and mouse sequences of olfactory receptor genes, combined with these comparative mapping data, provide a first glimpse at the similarities and differences existing between gene clusters located at syntenically

homologous sites of the dog, human, and mouse genomes. These comparisons provide a basis for further research into the evolution and function of mammalian olfactory gene clusters.

MATERIALS AND METHODS

Interspecific Backcross Mapping

To localize mouse sequences related to the four groups of canine olfactory receptor genes, we followed the segregation of restriction fragment length variants (RFLVs) detected by subfamily-specific probes in a subset of the progeny of a previously-described *Mus musculus* X *Mus spretus* interspecific backcross [IB; (C3Hf/Rl-*Mgf* $Sl2ENURg_{/+}$ X *M. spretus*) X C3Hf/Rl), Doyle *et al.*, 1996; Stubbs *et al.*, 1996]. Southern blots carrying restriction digests of DNA samples prepared from each of 118 backcross progeny were hybridized with radiolabled probes as previously described (Stubbs *et al.*, 1990), with the mouse probes hybridized at higher stringencies (hybridizations in the presence of 50% formamide at 42°C; washed at 0.2 X SSC at 65°C), and canine markers hybridized under conditions of lower stringency (hybridized in buffer containing 45% formamide at 42°C; washed at 1 X SSC at 65°C). Map Manager data analysis program (Manly, 1993) was used to store and analyze mapping data, with standard errors derived using standard statistical methods (Silver, 1985).

Human Mapping

The CfOLF4 subfamily was represented by a 900 bp PCR product that spans most of the length of a canine olfactory receptor gene present on canine cosmid "COS12" (Issel-Tarver and Rine, 1996). The probe was hybridized to high density arrays of human Chromosome 19-specific cosmids (Olsen *et al.*, 1994). After hybridization, filters were then washed in several steps of increasing stringency, to a final wash of 0.1 X SSC, 0.1%

SDS at 65 °C. The relative positions of cosmid contigs in this library have been established by high resolution fluorescent *in situ* hybridization (FISH) procedures, as previously described (Brandriff *et al.*, 1994; Ashworth *et al.*, 1995).

Molecular Probes

All markers analyzed in this study and variant fragments used to assign them to mouse chromosome maps are described in Table 1. The oligonucleotide primers used to derive *Sfpi1* by PCR are listed: (5' ACCAGTTCCTGTTGGACCTG 3') and (5' CTTCACCTTCTTGACCTCGC 3') (*Sfpi1*; Smith *et al.*, 1993). PCR reactions used 0.2 U Taq polymerase (Perkin-Elmer/Cetus) in standard, supplied buffers, and were amplified with the following program in a GeneAmp 9600 thermocycler (Perkin-Elmer/Cetus): 95°C, 4 min.; (94°C, 1 min.; 55 °C 1 min.; 72°C, 2 min.) X 30 cycles; 72°C, 5 min. Comparisons with related OLFR gene sequences from several species were carried out using GCG software, Version 8.1 (Genetics Computer Group).

RESULTS

Murine genes related to CfOLF1 and CfOLF2 subfamilies are clustered in Mmu2

To determine in mouse the positions of orthologs of the canine genes CfOLF1, CfOLF2, CfOLF3, and CfOLF4, we followed the segregation of RFLVs detected by subfamily-specific probes in a *Mus musculus* X *Mus spretus* interspecific backcross (Doyle *et al.*, 1996; Stubbs *et al.*, 1996). In general, the number of mouse restriction fragments found to cross-hybridize with a particular canine gene probe was not altered with more stringent hybridization and washing conditions, indicating that selected subfamily members with roughly similar levels of sequence homology were detected. Since the OLFR genes are known to be small and intronless, the number of restriction fragments detected by each subfamily-specific probe provides a good indicator of the

number of genes within a particular group (Issel-Tarver and Rine, 1996). To maximize the number of independent family members that could be mapped in each IB hybridization experiment, we chose enzyme digests that produced the largest numbers of RFLVs for each probe, and traced the segregation of each variant fragment independently (see Table 1).

CfOLF1 and CfOLF2 probes each detected 6-8 DNA fragments in restriction enzyme-digested mouse DNA, consistent with the number of subfamily members present in the dog genome (Issel-Tarver and Rine, 1996). All CfOLF1 and CfOLF2 RFLVs mapped together in this IB system, indicating that, as observed in human and dog, these two OLFR gene subfamilies are tightly linked in mouse (Figure 1A; Issel-Tarver and Rine, 1997). The two mouse OLFR gene clusters also cosegregated with three unrelated loci, including *Cf2* and *Sfpi1*, which had been localized to central Mmu2 in previous studies (Friezner-Degen *et al.*, 1990; Paul *et al.*, 1991), and *Acp2,* encoding acid phosphatase 2 (Pohlmann *et al.*, 1988), the position of which is reported here for the first time. These data position the CfOLF1 and CfOLF2 gene subfamilies within a region of solid genomic homology to human Chromosome 11p11-p12 (Nguyen *et al.*, 1990; Glaser *et al.*, 1985). Human *HsOLF1* and *HsOLF2* have been assigned to the centromeric region of human Chromosome 11q (*HsOLF1* or *OLF1* accession number U56420, Issel-Tarver and Rine, 1997), suggesting that the homology region extends across the centromere of human Chromosome 11.

Linkage of CfOLF1- and CfOLF2-related genes to *Sfpi1* also places the orthologs of the two canine OLFR subfamilies in the region to which a mouse olfactory receptor gene cluster, *Olfr4*, had been mapped in previous studies (Sullivan *et al.*, 1996). The relationship between CfOLF1, CfOLF2, and *Olfr4* subfamilies was investigated by comparing the sequences of the genes within the mouse and canine clusters. Although DNA sequence is available for only portions of each mouse gene, certain general

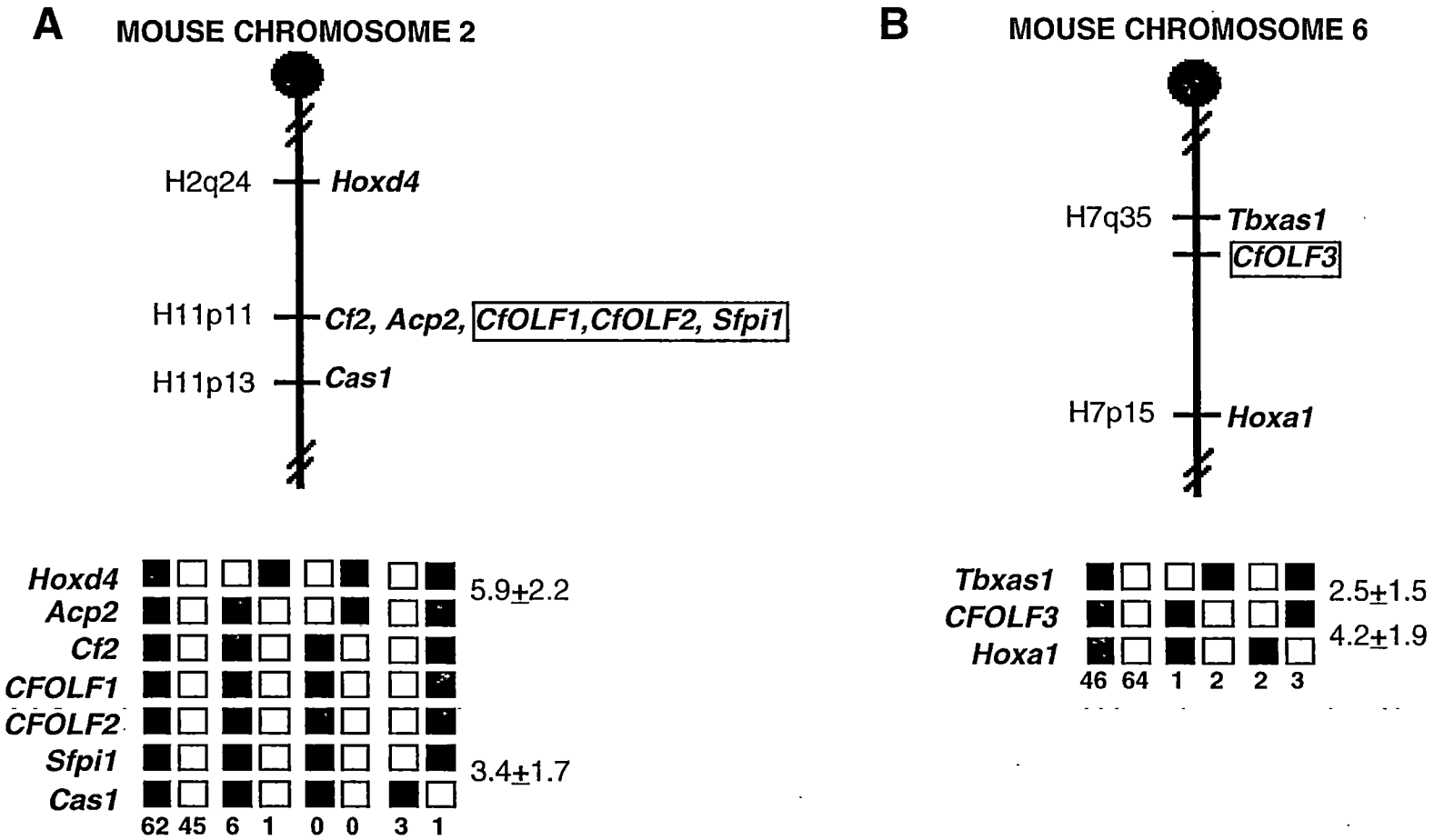
Table one. Description of probes and variant fragments.

For cloned probes, the type of clone used (c=cDNA, g=genomic) and species of origin are listed (H=human; M=mouse; D=dog). The sizes of variant fragments used to follow the segregation of each marker in *Mus musculus* (M) or *Mus spretus* (S) DNA, and the restriction enzyme used to generate those fragments, are also shown. References describing the probes used and source are listed at right (ATCC=American Type Culture Collection, Rockville, MD; EST=Expressed Sequence Tags, available from the I.M.A.G.E. Consortium).

Chr.	Locus	Type/ Species	M(kb)	S(kb)	Enzyme	Reference
2	*Hoxd4*	g/M	2.8	2	*Eco*RI	ATCC#63103 Featherstone et al., 1988
2	*Cf2*	c/M	1.9	3.5	*Taq*I	Friezner-Degen et al., 1990
2	*Acp2*	c/M	2	1.8	*Taq*I	ATCC#59900 Pohlmann et al. 1988
2	CfOlf1	c/D	6.2	4.4	*Pvu*II	Issel-Tarver and Rine, 1996
2	CfOlf2	c/D	6.2	7.7	*Taq*I	Issel-Tarver and Rine, 1996
2	*Sfpi1*	g/M	10	17	*Bam*HI	Smith et al., 1993
2	*CAS1*	c/H	1.3	1.4	*Taq*I	ATCC#57354 Quan et al., 1986
6	*Tbxas1*	c/M	2.4	6.5	*Taq*I	EST 875135 Marra et al., 1996
6	CfOLF3	c/D	6.5	2.8	*Eco*RI	Issel-Tarver and Rine, 1996
6	*HOXA1*	c/H	18.5	13	*Eco*RV	EST 530888 Hillier et al., 1995
9	*LDLR*	g/H	5.3	3.5	*Taq*I	ATCC#57004 Yamamoto et al., 1984
9	*Icam*	c/M	5.3	7.5	*Hinc*II	ATCC#63043 Horley et al., 1989
9	CfOlf4rs1	c/D	2.8	2.4	*Eco*RV	Issel-Tarver and Rine, 1996
9	*HEXA*	c/H	3.4	2.6	*Hinc*II	ATCC#57352 Korneluk et al., 1986
10	*Myb*	c/M	3.1	4.4	*Taq*I	Mucenski et al., 1991
10	CfOlf4rs2	c/D	3.5	3.2	*Eco*RV	Issel-Tarver and Rine, 1996
10	*PFKL*	c/H	2.6	4.2	*Eco*RI	ATCC#61648 Levanon et al., 1989
10	*Igtb2*	c/M	9.8	17	*Eco*RV	ATCC#63056 Wilson et al., 1989

Figure 1. Segregation patterns and predicted map positions of murine genes related to canine CfOLF1, CfOLF2, and CfOLF3 subfamilies.

Sequences were analyzed in 118 progeny of an interspecific backcross system IB; ([(C3Hf/Rl-*Mgf*^{Sl2ENUR8}/+ X *Mus spretus*) X C3Hf/Rl], Doyle *et al.*, 1996; Stubbs *et al.*, 1996) typed for markers derived from all mouse chromosomes. Probes and variant fragments used to follow each locus are described in Table 1. The partial chromosome linkage maps showing the location of each CFOLF gene in relation to linked genes appear above each chromosome segregation pattern. Recombination distances between loci are shown to the right between pairs of rows representing adjacent genes, in centimorgans, with standard errors. The positions of the human homologs of all loci are given, when known, at the left of the chromosome map. Each column represents a type of parental or recombinant chromosome; the number of backcross animals inheriting a particular type of chromosome is listed at the bottom of each column. A white box represents the inheritance of a *Mus spretus* allele for a gene, while a black box denotes the presence of a C3Hf/Rl allele for the same gene. A) The position of CfOLF1 and CfOLF2 related sequences on mouse chromosome 2, within a homology region of human 11p is shown. B) The position of CfOLF3 related sequences is localized on mouse chromosome 6 in a region of human 7q homology. These data can be found in Mouse Genome Database (www.jax.org) under data accession number J:44485.


A
MOUSE CHROMOSOME 2
H2q24
Hoxd4
H11p11
Cf2, Acp2, CfOLF1,CfOLF2, Sfpi1
H11p13
Cas1
Hoxd4
Acp2
Cf2
CFOLF1
CFOLF2
Sfpi1
Cas1
62 45 6 1 0 0 3 1
5.9±2.2
3.4±1.7
B
MOUSE CHROMOSOME 6
H7q35
Tbxas1
CfOLF3
H7p15
Hoxa1
Tbxas1
CFOLF3
Hoxa1
46 64 1 2 2 3
2.5±1.5
4.2±1.9

conclusions can be drawn. One gene in the *Olfr4* cluster, designated K17 (Sullivan *et al.*, 1996), was found to be most similar in amino acid sequence to canine CfOLF1 and CfOLF2 (Issel-Tarver and Rine, 1996), and is more closely related to CfOLF2 than CfOLF1 (Table 2). However, the "variable" portion of CfOLF1, corresponding to transmembrane domains four and five of the receptor protein, shares 90% amino acid sequence identity with the same region of the gene's known human homolog (*OLF1*; Issel-Tarver and Rine, 1997). The relatively low degree of conservation observed between mouse K17 and canine CfOLF1 or CfOLF2, and the striking similarity existing between dog CfOLF1 and its human counterpart, indicates that although K17 represents the closest known murine relative of CfOLF1 and CfOLF2, true orthologs of the sequenced canine genes have not been cloned in mice.

Locations of mouse genes related to CfOLF3 and CfOLF4 subfamilies

The CfOLF3 subfamily probe detected four restriction fragments in mouse DNA digests, indicating that a similar number of related genes exist in human, dog and mouse. One of the four CfOLF3-related fragments produced a variant fragment that could be traced in *Eco*RI-digested DNA samples, and permitted us to localize CfOLF3-related sequences to proximal Mmu6 (Figure 1B). Although no olfactory receptor genes have yet been assigned to this mouse region, the position is consistent with the assignment of CfOLF3-related sequences to the syntenically homologous region of human Chromosome 7q35 (Issel-Tarver and Rine, 1997).

The CfOLF4 probe detected a much larger collection of independent restriction fragments in human and dog DNA samples than did other subfamily representatives, indicating that this subfamily contains a relatively large number of genes in both species (Issel-Tarver and Rine, 1996).

Table Two. Protein sequence species search.

Each CfOLF sequence corresponding to amino acid residues 101-250 were used to search the Swissprot database (Bairoch and Boeckmann, 1993) to identify proteins from other species that had a significant degree of identity with the CfOLF proteins. Under each CfOLF type are the best returns for different species including chicken, rat, human, and mouse. The percent of amino acid sequence identity of each protein to its corresponding CfOLF type is shown as well as its accession number and reference.

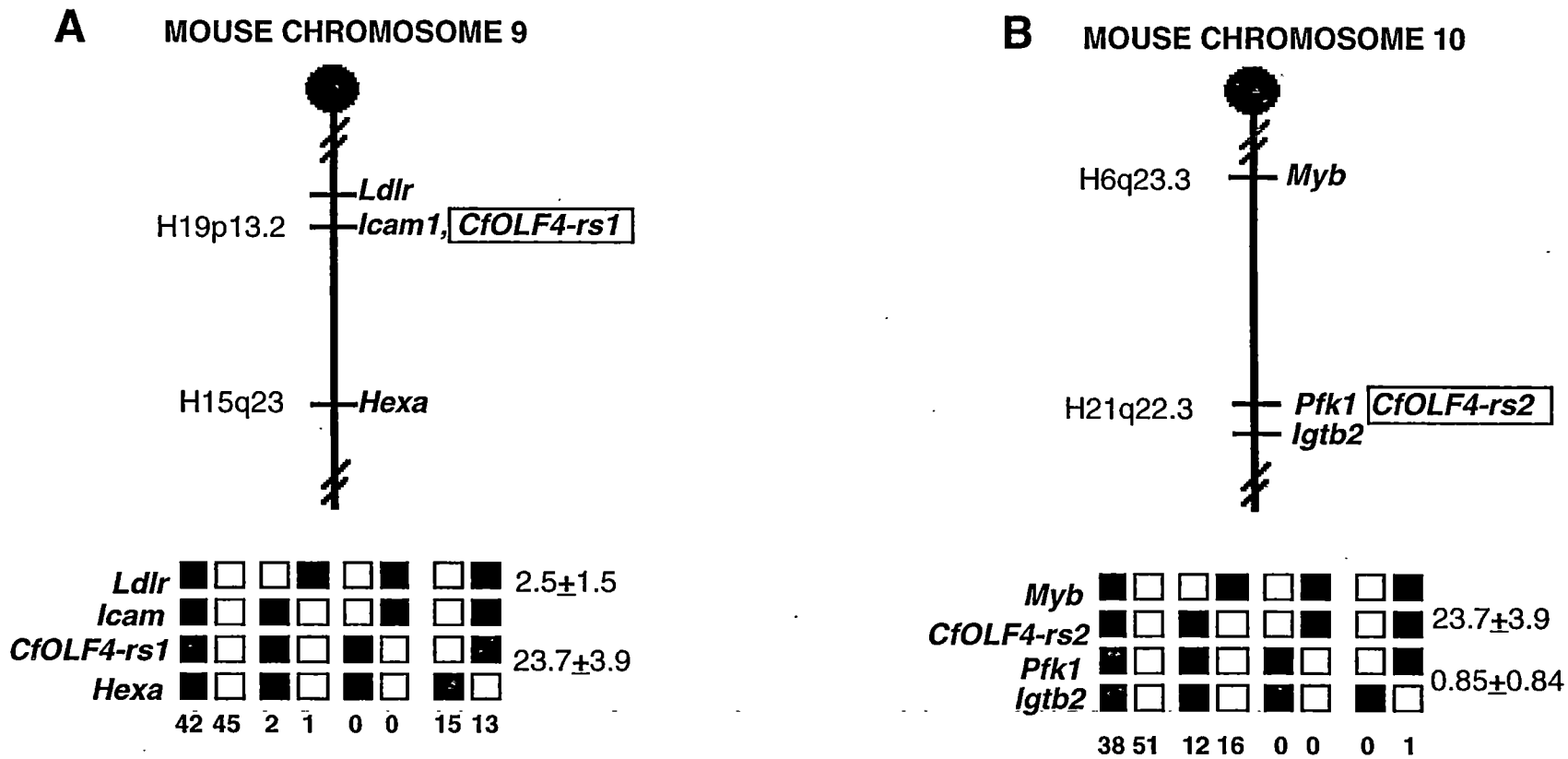
	Amino Acid Identity %	Accession #	Reference
CfOLF1			
Human *OLF1*	90.0	U56420	Issel-Tarver and Rine, 1996
Mouse *Olfr4*-K17	45.5	U28769	Sullivan et al., 1996
Chick Olf4	48.3	P37070	Nef et al., 1996
Human *Ol1d*	47.3	P47885	Ben-Arie et al., 1994
CfOLF2			
Mouse *Olfr4*-K17	52.7	U28769	Sullivan et al., 1996
Chick Olf5	48.3	P37071	Nef et al., 1996
Human *Ol1e*	47.3	P47881	Ben-Arie et al., 1994
CfOLF3			
Human *OLF3*	88.7	U56421	Issel-Tarver and Rine, 1996
Mouse *Olf7*	46.3	P34984	Ressler et al., 1993
Human *OlFJ*	47.3	P30954	Parmentier et al., 1992
CfOLF4			
Mouse *Olfr8*-M64	55.4	U28781	Sullivan et al., 1996
Rat Olf3	68.7	P23265	Buck and Axel, 1991
Human *Ol1J*	52.0	P47890	Ben-Arie et al., 1994
Mouse *Olf7*	43.6	P34984	Ressler et al., 1993

A similarly large number of restriction fragments was also observed in mouse. Most of the 20-25 fragments detected with the CfOLF4 probe in various restriction digests did not vary in size in *Mus musculus* and *Mus spretus* DNA samples, and it was possible to map only two variant fragments with certainty. One variant CfOLF4-related locus (*CfOlf4-rs2*) was found to be located on Mmu10, near *Pfkl* and *Itgb2* in the central region of the chromosome (Figure 2B). This central Mmu10 region is comprised of several small blocks of mouse-human genomic homology, corresponding to human chromosomes 21q22.3, 22q11.1, 10q21, or 19p13.3 (Taylor *et al.*, 1996; Debry and Seldin, 1996), and for this reason the location of related human genes cannot be predicted with any certainty. A mouse olfactory receptor gene cluster, *Olfr8*, has been mapped to this same Mmu10 region in previous reports (Sullivan *et al.*, 1996). The relationship between CfOLF4 and *Olfr8* gene family members is underscored by similarity in amino acid sequence with an *Olfr8* cluster member, M64 (Table 2). This level of similarity may indicate a relationship between CfOLF4 and M64, although the genes are probably not orthologous.

Unlike mouse sequences detected by CfOLF1, CfOLF2, or CfOLF3 probes, fragments recognized by CfOLF4 did not appear to be clustered at a single site in the mouse genome. Another variant CfOLF4-related fragment traced in this IB system (*CfOlf4-rs1*) was mapped to the proximal region of Mmu9, tightly linked to markers *Ldlr* and *Icam1* (Figure 2A). This murine region also contains a previously identified mouse olfactory receptor gene, *Olfr2* (Copeland *et al.*, 1993). Since the DNA sequence available for *Olfr2* is derived only from the 5' untranslated region (Chess *et al.*, 1994) and untranslated regions are generally not expected to be conserved between mouse and dog, it is difficult to assess the potential relationship between these genes. However, the close linkage of *CfOlf4*-rs1, *Ldlr,* and *Icam1* is consistent with earlier studies that have localized human CfOLF4-related genes to human Chromosome 19p13.2 (Issel-Tarver and Rine, 1997).

Figure 2. Map positions of two murine CfOLF4-related loci.

The CfOLF4-related sequences were analyzed in the same interspecific backcross system that was described in Figure 1. The CfOLF4 probe and variant fragments used to follow each locus are indicated in Table 1. The partial chromosome linkage maps showing the position of each CfOLF4-related sequence in relation to linked genes appear above each chromosome segregation pattern. Recombination distances, positions of the human homologs, and columns are all arranged similarly to Figure 1. A) *CfOLF4-rs2* maps to mouse chromosome 9 in a region of human 19p13.2 homology. B) *CfOLF4-rs1* is located on mouse chromosome 10 in a region with several human homology groups.


A
MOUSE CHROMOSOME 9
Ldlr
H19p13.2
Icam1, CfOLF4-rs1
H15q23
Hexa
Ldlr
Icam
CfOLF4-rs1
Hexa
2.5±1.5
23.7±3.9
42 45 2 1 0 0 15 13
B
MOUSE CHROMOSOME 10
H6q23.3
Myb
H21q22.3
Pfk1
Igtb2
CfOLF4-rs2
Myb
CfOLF4-rs2
Pfk1
Igtb2
23.7±3.9
0.85±0.84
38 51 12 16 0 0 0 1

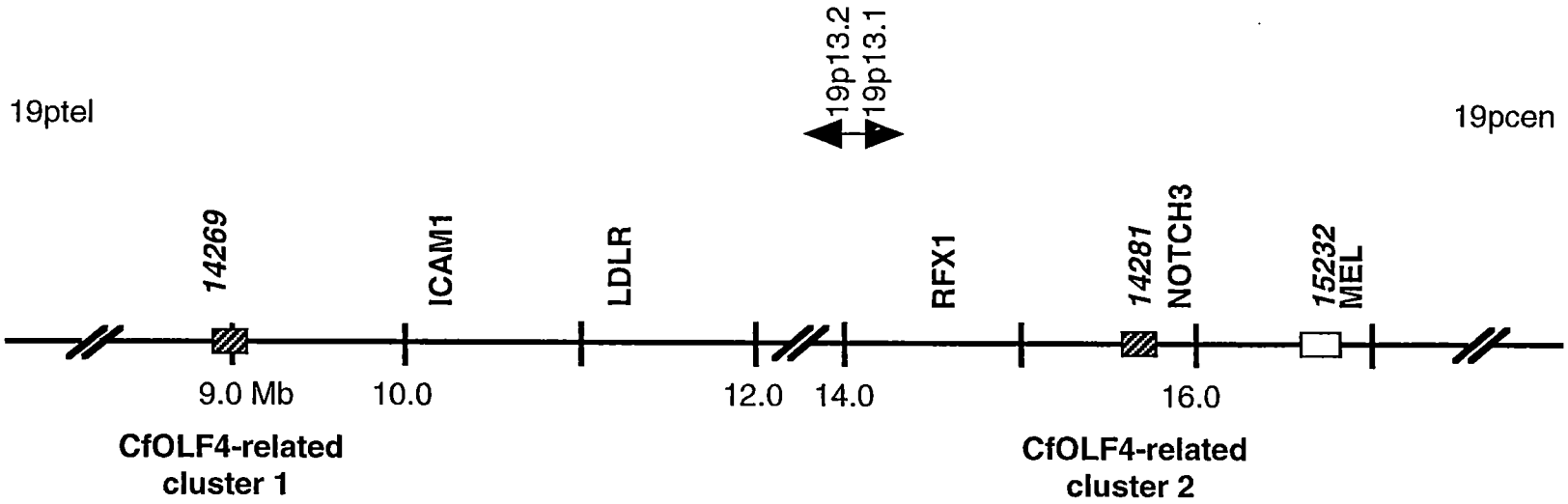
Physical localization of CfOLF4-related genes in human Chromosome 19p

Genes related to all four canine subfamilies have been mapped in human using somatic cell hybrids, and some were located on human Chromosome 19 (Issel-Tarver and Rine, 1997). To better define the position(s) of potential CfOLF4-gene homologs, we hybridized high density filter arrays of ordered human Chromosome 19 cosmids (Ashworth *et al.*, 1995) with the CfOLF4 probe. At least twenty cosmids were clearly detected in these experiments, even under conditions of high stringency. The positive clones mapped to two locations on the Chromosome 19 metric map (Ashworth *et al.*, 1995), both of which were previously identified as containing olfactory receptor genes (R.R. Reed, unpublished). Eight strongly hybridizing cosmids were distributed over three adjacent 19p13.2 cosmid contigs spanning a total distance of approximately 200 kb, suggesting the presence of a minimum of three distinct genes. The human 19p13.2 CfOLF4-related cosmids are located approximately 1.2 Mb distal of ICAM1 and 2.2 Mb distal of LDLR, respectively (CfOLF4-related cluster 1; Figure 3), consistent with the close linkage between CfOLF4-related genes to *Ldlr* and *Icam1* determined in mouse (Figure 2A). The assignment of CfOLF4-related sequences to 19p13.2 and Mmu 9 extends the length of this established region of conserved mouse-human synteny from approximately 1 Mb to a region at least 2 Mb in length (Figure 3).

The remaining twelve CfOLF4-related cosmid clones were distributed throughout a second ca. 200 kb region located in p13.1 (CfOLF4-related cluster 2; Figure 3). The positive cluster of cosmids was positioned by these studies approximately 15 Mb from the 19p telomere, and therefore approximately 6 Mb proximal of the 19p13.2 OLFR cluster. The distribution of these positive clones throughout the overlapping cosmids also suggested that they represent at least three different genes.

Figure 3. Location of CfOLF4-related genes on human chromosome 19p.

An ~8Mb section of the p13.2-13.1 region of chromosome 19 is represented. The scale below the line indicates distance from p-telomere in Mb, as determined by a combination of high resolution pronuclear FISH and restriction mapping (updated from Brandriff *et al.*, 1994, and Ashworth *et al.*, 1995). The locations of several genes mapped to this region are indicated above the line. Numbers in italics indicate the locations of three OLFR-containing cosmids on the pronuclear FISH map (Brandriff *et al.*, 1994). Hatched boxes indicate the two clusters of CfOLF4-related genes in p13.2 and p13.1. The open box indicates an OLFR gene cluster in p13.1 with no detectable relationship to CfOLF4.


19ptel
19p13.2
19p13.1
19pcen
14269
ICAM1
LDLR
RFX1
14281
NOTCH3
15232
MEL
9.0 Mb
10.0
12.0
14.0
16.0
CfOLF4-related cluster 1
CfOLF4-related cluster 2

None of the CfOLF4-related cosmid sequences were located in a nearby 19p13.1 region, located between human *NOTCH3* and *MEL* genes, previously noted to contain a cluster of OLFR sequences (R.R. Reed, unpublished; Ashworth *et al.*, 1995); genes in this most proximal 19p13.1 cluster are therefore probably derived from a different OLFR subfamily.

DISCUSSION

These studies have defined the map position of five independent mouse olfactory receptor genes corresponding to the closest murine relatives of dog gene subfamilies, CfOLF1, CfOLF2, CfOLF3, and CfOLF4. Our results demonstrate that CfOLF1- and CfOLF2-related genes are clustered in mouse, as they are in the dog and human genomes. The position determined for CfOLF1- and CfOLF2-related sequences suggest that these genes may be members of the previously mapped mouse *Olfr4* cluster, although the low level of sequence similarity between mouse and canine genes suggests that true murine orthologs of CfOLF1 and CfOLF2 have yet to be discovered. CfOLF3 sequences detected a mouse gene located in proximal Mmu6, consistent with the assignment of a human CfOLF3-related gene cluster in the syntenically homologous region of human Chromosome 7q35 (Issel-Tarver and Rine, 1997); these data define a new location for OLFR genes in the mouse. CfOLF4 subfamily probes detect 20 or more independent loci in mouse, human, and dog and are located in at least two distinct regions in mouse: proximal Mmu9 and central Mmu10, respectively. The similar map position of CfOLF4-related genes and *Olfr2* in proximal Mmu9, and previous reports assigning the *Olfr8* gene cluster to central Mmu10, suggest the potential familial relationship between these previously mapped mouse genes and CfOLF4. Human physical mapping studies have identified two independent clusters of CfOLF4-related sequences, each spanning approximately 200 kb and located 5 Mb apart in 19p13.1 and 19p13.2, respectively. The

19p13.2 gene cluster is located near *LDLR* and *ICAM1*, in agreement with the tight linkage of *Ldlr, Icam1,* and *CfOLF4-rs1* in proximal Mmu9. CfOLF4-related sequences were also found between *MEL* and *RFX1* in 19p13.1 in a region with extensive homology to central Mmu8 (Doyle *et al.*, 1996; DeBry and Seldin, 1996). However, mouse genes representing potential orthologues of these human genes have not been discovered in Mmu8. CfOLF4-related genes are numerous, widely distributed, and difficult to map (Issel-Tarver and Rine, 1996; this study), therefore additional murine CfOLF4-related clusters are likely to be found in future studies.

Protein sequence searches identified olfactory receptor proteins from chicken that are as similar to the canine genes as are proteins predicted from most of the available mammalian OLFR gene sequences (Table 2). This fact indicates that at least some OLFR genes are highly conserved between species, and underscores the likelihood that true murine counterparts of the four sequenced canine genes have not been identified.

At present, OLFR gene cluster organization is not well understood; specific clusters may be comprised of closely related genes, or they may contain members representing a selection of different sequence subfamilies. Genes comprising one of the best characterized OLFR gene clusters, located in human Chromosome 17p13.3, have been shown to represent two distinct human subfamilies. However, closer examination of this 350 kb region has demonstrated that genes of like sequence are grouped together within two smaller subclusters (Ben-Arie *et al.*, 1994). Members of at least one conserved OLFR subfamily (represented by CfOLF4 sequences) are dispersed at several sites in humans, dogs, and mice. Data presented here support the notion that similar genes are clustered in at least some OLFR loci in the mouse, since multiple, independently-traced RFLVs detected by two conserved canine probes (fragments detected with both CfOLF1 and CfOLF2) mapped to a single chromosomal site.

Together with previously published reports, our results confirm the close conservation of four OLFR gene subfamilies in mouse, human, and dog and provide a basis for functional studies of these genes in the three species. With the organization, evolution, expression control, and neurobiology of odorant perception still only poorly understood, isolation and mapping of OLFR genes in the mouse provides a first step toward elucidating the functions of these genes in the context of a powerful *in vivo* model system.

ACKNOWLEDGMENTS

We would like to thank Xiaojia Ren and Beverly Stanford for assistance with mouse gene mapping, Susan Allen for hybridization screening of the human Chromosome 19 cosmid library, and Mark Shannon for critical reading of the manuscript. Our gratitude is also extended to Sandra Friezner-Degen for the kind gift of clone MC3 (*Cf2*), and Mike Mucenski and Hsi-Hsien Lin for the gift of the *Myb* probe. L.I-T was supported by a Howard Hughes Medical Institute Predoctoral Fellowship and a Berkeley Fellowship. This work was supported by grants from the U.S. DOE under contract nos. DE-AC05-96ORO22464 (Oak Ridge National Laboratory) and W-7405-ENG-48 (Lawrence Livermore National Laboratory).

REFERENCES

Ashworth, L.K., Batzer, M.A., Brandriff, B., Branscomb, E., deJong, P., Garcia, E., Garnes, J., Gordon, L., Lamerdin, J.E., Lennon, G., Mohrenweiser, H., Olsen, A., Slezak, T., and Carrano, A.V. (1995). A metric physical map of human chromosome 19. *Nature Genet* **11**:422-427.

Bairoch, A., and Boeckmann, B. (1993). The SWISS-PROT protein sequence data bank, recent developments. *Nucleic Acids Res* **21**:3093-3096.

Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D.H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H., and North, M.A. (1994). Olfactory receptor gene cluster on human chromosome 17: possible duplication of an ancestral receptor repertoire. *Hum Mol Genet* **3**:229-235.

Brandriff, B.F., Gordon, L.A., Fertitta, A., Olsen, A.S., Christensen, M., Ashworth, L.K., Nelson, D.O., Carrano, A.V., and Mohrenweiser, H.W. (1994). Human chromosome 19p: a Fluorescence in situ hybridization map with genomic distance estimates for 79 intervals spanning 20 Mb. *Genomics* **23**:582-591.

Buck, L. and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**:175-187.

Chess, A., Simon, I., Cedar, H., and Axel, R. (1994). Allelic inactivation regulates olfactory receptor gene expression. *Cell* **78**:823-834.

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

DeBry, R.W., and Seldin, M.F. (1996). Human/mouse homology relationships. *Genomics* **33**:337-351.

Doyle, J., Hoffman, S., Ucla, C., Reith, W., Mach, B., and Stubbs, L. (1996). Locations of human and mouse genes encoding the RFX1 and RFX2 transcription factor proteins. *Genomics* **35**:227-230.

Featherstone, M.S., Baron, A., Gaunt, S.J., Mattei, M.G., and Duboule, D. (1988). Hox-5.1 defines a homeobox-containing gene locus on mouse chromosome 2. *Proc Natl Acad Sci* **85**:4760-4764.

Friezner-Degen, S.J., Schaefer, L.A., Jamison, C.S., Grant, S.G., Fitzgibbon, J.J., Pai, J.A., Chapman, V.M., and Elliott, R.W. (1990). Characterization of the cDNA coding for mouse prothrombin and localization of the gene on mouse chromosome 2. *DNA Cell Biol* **9**:487-498.

Glaser, T., Gerhard, D., Jones, C., Albritton, L., Lalley, P., and Housman, D. (1985). A fine structure deletion map of chromosome 11p. *Cytogenet. Cell Genet* **40**:643.

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,Parsons, J., Rifkin, L., Rohlfing, T., Tan, F.,

Trevaskis, E.,Waterston, R., Williamson, A., Wohldmann, P., and Wilson, R. (1995). WashU-Merck EST Project. Unpublished.

Horley, K.J., Carpenito, C., Baker, B., and Takei, F. (1989). Molecular cloning of murine intercellular adhesion molecule (ICAM-1). *EMBO J* **8**:2889-2896.

Issel-Tarver, L. and Rine, J. (1996). Organization and expression of canine olfactory receptor genes. *Proc Natl Acad Sci* **93**:10897-10902.

Issel-Tarver, L. and Rine, J. (1997). The evolution of mammalian olfactory receptor genes. *Genetics* **145**:185-195.

Johnson, D.K. (1990). Reverse genetics and genes that control development: initial molecular characterization of two deletion complexes in mouse Chromosome 7 by random-clone mapping and genetic linkage analysis. Doctoral Thesis, University of Tennessee, Knoxville, TN.

Korneluk, R.G., Mahuran, D.J., Neote, K., Klavins, M.H., O'Dowd, B.F., Tropak, M., Willard, H.F., Anderson, M.J., Lowden, J.A., and Gravel, R.A. (1986). Isolation of cDNA clones coding for the alpha-subunit of human beta-hexosaminidase. Extensive homology between the alpha- and beta-subunits and studies on Tay-Sachs disease. *J Biol Chem* **261**:8407-8413.

Levanon, D., Danciger, E., Dafni, N., Bernstein, Y., Elson, A., Moens, W., Brandeis,M., and Groner, Y. (1989). The primary structure of human liver typephosphofructokinase and its comparison with othertypes of PFK. *DNA* **8**:733-743.

Manly, K.F. (1993). A MacIntosh program for the storage and analysis of experimental genetic mapping data. *Mammal Genome* **4**:303-313.

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin,J.,Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R., and Waterston, R. (1996). The WashU-HHMI mouse EST project. Unpublished.

MacDonald, G., Chu, M.L., and Cox, D.R. (1991). Fine structure physical mapping of the region of mouse chromosome 10 homologous to human chromosome 21. *Genomics* **11**:317-323.

Mucenski, M. L., McLain, K., Kier, A. B., Swerdlow, S. H., Schreiner, C.M., Miller, T. A., Pietryga, D. W., Scott, W. J., and Potter, S. S. (1991). A functional c-myb gene is required for normal murine fetal hepatic hematopoisis. *Cell* **65**:677-689.

Nef, S., Allaman, I., Fiumelli, H., De Castro, E., and Nef, P. (1996). Olfaction in birds: differential embryonic expression of nine punative odorant receptor genes in the avain olfactory system. *Mech Dev* **55**:65-77.

Nguyen, V.C., Ray, D., Gross, M.S., de Tand, M.F., Frezal, J., Moreau-Gachelin, F. (1990). Localization of the human oncogene SPI1 on chromosome 11, region p11.22. *Hum Genet* **84**:542-546.

Olsen, A., Teglund, S., Nelson, D., Gordon, L., Copeland, A., Georgescu, A., Carrano, A., and Hammarstrom, S. (1994). Gene Organization of the Pregnancy-Specific Glycoprotein Region on Human Chromosome 19: Assembly and Analysis of a 700 kb Cosmid Contig Spanning the Region. *Genomics* **23**:659-668.

Parmentier, M., Libert, F., Schurmans, S., Schiffmann, S., Lefort, A., Eggerickx, D., Ledent, C., Mollereau, C., Gerard, C., Perret, J., Grootegoed, A., and Vassart, G. (1992). Expression of members of the putative olfactory receptor gene family in mammalian germ cells. *Nature* **355**:453-455.

Paul, R., Schuetze, S., Kozak, S.L., Kozak, C.A., and Kabat, D. (1991). The Sfpi-1 proviral integration site of Friend erythroleukemia encodes the ets-related transcription factor Pu.1. *J Virol* **65**:464-467.

Pohlmann, R., Krentler, C., Schmidt,.B., Schroder, W., Lorkowski, G., Culley, J., Mersmann, G., Geier, C., Waheed, A., Gottschalk, S., Grzeschik, K-H., Hasilik, A., and von Figura, K. (1988). Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment. *EMBO J* **7**:2343-2350.

Quan, F., Korneluk, R.G., Tropak M.B., and Gravel, R.A. (1986). Isolation and characterization of the human catalase gene. *Nucleic Acids Res* **14**:5321-5335.

Silver, J. (1985). Confidence limits for estimates of gene linkage analysis of recombinant inbred strains. *J Heredity* **76**:436-440.

Smith, M.W., Clark S.P., Hutchinson, J.S., Wei, Y.H., Churukian, A.C., Daniels, L.B., Diggle, K.L., Gen, M.W., Romo, A.J., Lin, Y., Selleri, L., McElligott, D.L., and Evans, G.A. (1993). A sequence-tagged site map of human chromosome 11. *Genomics* **17**:699-725.

Stubbs, L., Carver, E., Ashworth, L., and Lopez-Molina, L. (1996). Location of the DBP transcription factor gene in man and mouse. *Mammal Genome* **7**:65-67.

Stubbs, L., Huxley, C., Hogan, B., Evans, T., Fried, M., Duboule, D., and Lehrach, H. (1990). The HOX-5 and surfeit gene clusters are linked in the proximal portion of mouse chromosome 2. *Genomics* **6**:645-650.

Sullivan, S.L., Adamson, M.C., Ressler, K.J., Kozak, C.A., and Buck, L.B. (1996). The chromosomal distribution of mouse odorant receptor genes. *Proc Natl Acad Sci* **93**:884-888.

Taylor, B.A., Burmeister, M., and Bryda, E.C. (1996). Mouse Chromosome 10. *Mammal Genome* **6**:S190-200.

Wilson, R.W., O'Brein, W.E., and Beaudet, A.L. (1989). Nucleotide sequence of the cDNA from the mouse leukocyte adhesion protein CD18. *Nucleic Acids Res* **17**:5397.

Yamamoto, T., Davis, C.G., Brown, M.S., Schneider, W.J., Casey, M.L., Goldstein, J.L., and Russell, D.W. (1984). The human LDL receptor: a cysteine-rich protein with multiple Alu sequences in its mRNA. *Cell* **39**:27-38.

SECTION 2. PHYSICAL MAPPING OF *EMR1* AND *CD97*, TWO DISTANTLY RELATED GENES, TO HUMAN CHROMOSOME 19 AND ASSIGNMENT OF *Cd97* TO MOUSE CHROMOSOME 8

Physical mapping of *EMR1* and *CD97*, two distantly related genes, to human chromosome 19 and assignment of *Cd97* to mouse chromosome 8

[1,2]Ethan A. Carver, [3]Jörg Hamann, [1]Anne S. Olsen, [1,4]Lisa Stubbs

[1]Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, P.O. Box 808, L-452, Livermore, CA 94550, USA

[2]University of Tennessee-Oak Ridge School of Biomedical Sciences, Oak Ridge National Laboratory, 1060 Commerce Park, Oak Ridge, TN 37830-8026, USA

[3]CLB and Laboratory for Experimental and Clinical Immunology, Academic Medical Centre, University of Amsterdam, Plesmanlaan 125, 1066 CX Amsterdam, The Netherlands

[4]Corresponding author

tel. (925) 422-8473

fax. (925) 422-2282

email: Stubbs5@llnl.gov

Running title: *EMR1, CD97,* and *Cd97* mapping

Submitted: *Mammalian Genome*

EMR1 and *CD97* belong to the EGF-TM7 family of nonclassical seven-span transmembrane receptors, members of which are expressed primarily in the immune system. Here we describe the physical mapping of both *CD97* and *EMR1* in humans and the mapping of *Cd97* in mouse. *CD97* and *EMR1* are linked on human chromosome 19p13, but are located on separate chromosomes in mouse. Interestingly, *CD97* and *EMR1* are located near a second pair of distantly related immune system genes, *RFX1* and *RFX2*, respectively, in both human and mouse. These data suggest that the two pairs of genes may represent conserved remnants of an ancient regional chromosomal duplication that occurred before the divergence of mouse and human from a common vertebrate ancestor.

The membrane-spanning region of *CD97* and *EMR1* is homologous to the secretin receptor superfamily. Unlike this group of peptide hormone receptors, EMR1 and CD97 have extended extracellular portions that possess several EGF domains at the N-terminus (EGF-TM7 family) (McKnight *et al.*, 1998). The EGF domain region can function as a ligand-binding site, as demonstrated by the interaction of CD97 with CD55 (Hamann *et al.*, 1996). The EMR1 and CD97 proteins show a high degree of structural similarity, despite the fact that they share only about 31% amino acid sequence identity. The similarity between these two genes suggests that *CD97* and *EMR1* may have arisen as an ancient duplication of a common ancestral gene. *CD97* was originally assigned to human chromosome 19p13.12-p13.2 by fluorescence *in situ* hybridization (FISH) (Hamann *et al.*, 1995). Both FISH and Southern analysis of a panel of somatic cell hybrids were used to assign *EMR1* to 19p13.3 (Baud *et al.*, 1995). *Emr1* has been mapped to distal mouse chromosome 17, within a cluster of genes whose counterparts are located in human 19p13.3 (Lin *et al.*, 1997; McKnight *et al.*, 1997). These data suggest that the two related genes are not clustered, but are separated by a significant distance in the human genome. However, physical positions of *CD97* and *EMR1* have not been confirmed. We therefore

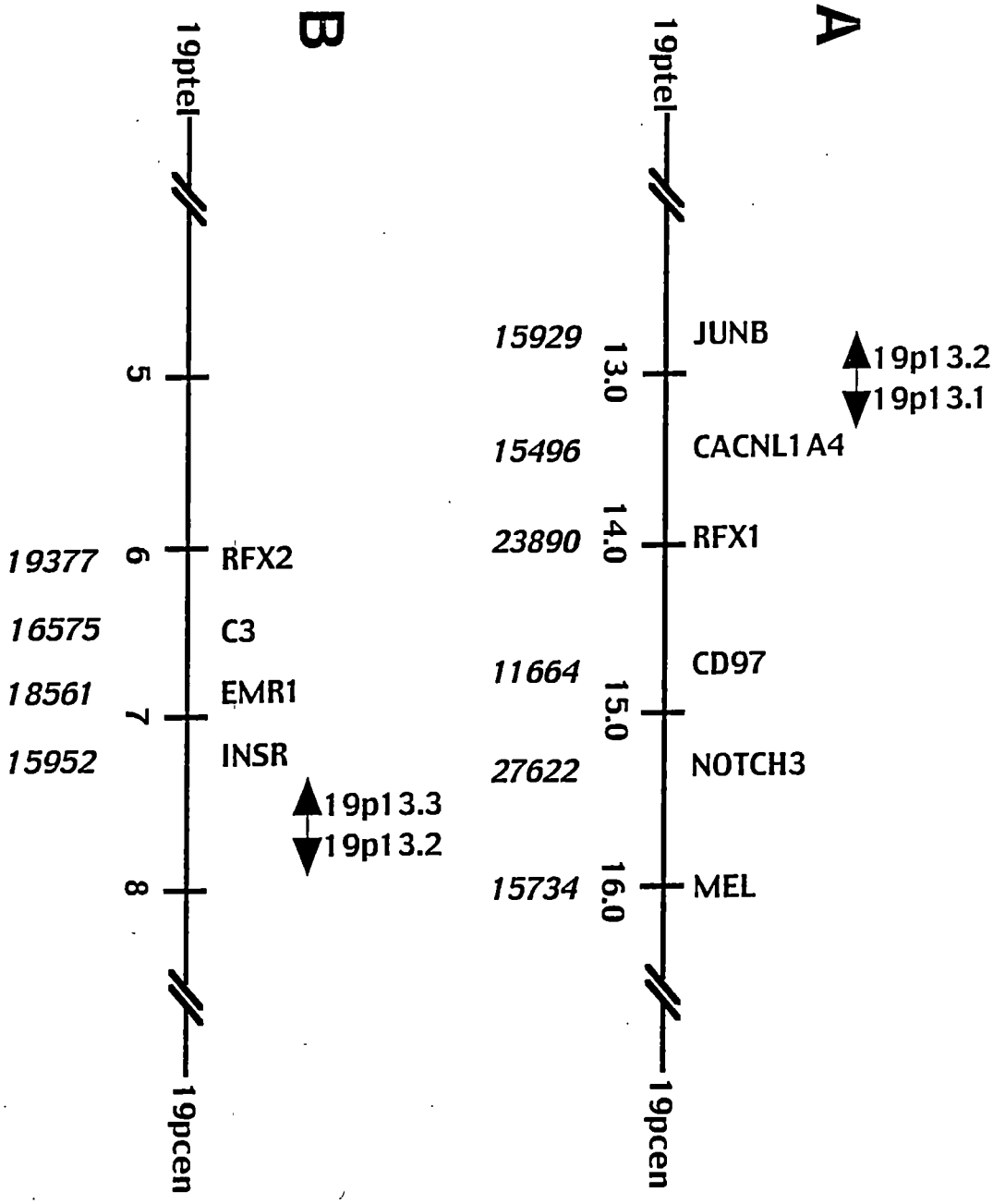
set out to demonstrate the positions of *CD97* and *EMR1* on the established physical map of human chromosome 19p (Ashworth *et al.*, 1995) and to determine the genetic localization of *Cd97* in mouse.

In order to localize the *CD97* and *EMR1* genes within the physical map of chromosome 19, human probes representing the two genes were radioactively labeled and hybridized to colony filters of a human chromosome 19 cosmid library, as previously described (Olsen *et al.*, 1994). *CD97* was localized using a 1495 bp human PCR product amplified from the CD97 cDNA with primers CD97.11-tgagccacaacaacaccaagg [nt 1140-1160 of the (+) strand] and CD97.20-gatgggacaagcgcccttagtc [nt 2613-2634 of the (-) strand] (GenBank accession #X84700; Hamann *et al.*, 1995). *EMR1* was mapped using a 201 bp human genomic PCR product generated from the *EMR1* cDNA 3' UTR sequence with primers EMRI3primeF-aaccctctggggaagaatgt [nt 2820-2839 of the (+) strand] and EMRI3primeR-caacaagggcaggagaaaac [nt 3001-3020 of the (-) strand] (GenBank accession #X81479; Baud *et al.*, 1995). The relative positions of positively hybridizing cosmid contigs were established by high resolution FISH, as previously described (Ashworth *et al.*, 1995; Brandriff *et al*, 1994). The human *CD97* probe identified several overlapping cosmids located in 19p13.1, between the *RFX1* and *NOTCH3* genes in the established chromosome 19 map (Figure 1A). The EMR1 probe detected two cosmid clones, 31568 and 34349; positive hybridization of both clones was confirmed by PCR with *EMR1* primers. This localized the human *EMR1* gene to chromosome 19p13.3, within a cosmid contig located approximately 1 Mb away from *RFX2* (Figure 1B).

These results confirm that the *CD97* and *EMR1* genes are separated by ~7.5 Mb in chromosome 19p. Interestingly, *CD97* and *EMR1* are within 0.7 and 0.5 Mb from *RFX1* and *RFX2*, respectively (Pugliatti *et al.*, 1992). *RFX1* and *RFX2* encode site-specific DNA binding proteins (Reith *et al.*, 1994), and are also likely to have arisen from a common ancestral gene sequence in distant evolutionary time. The low sequence identity

Figure 1. Location of *EMR1* and *CD97* genes on human chromosome 19p13.

Two ~4 Mb sections of chromosome 19p are depicted. Scales below the lines indicate distance from p-telomere in Mb, as determined by both high-resolution pronuclear FISH and restriction mapping (updated from Ashworth *et al.*, 1995; Brandriff *et al.*, 1994). The human *CD97* and *EMR1* genes were initially mapped to 19p13.2-p13.12 and 19p13.3, respectively (Hamann *et al.*, 1995; Baud *et al.*, 1995). In order to place both genes on a high resolution map of chromosome 19, gene-specific probes were radioactively labeled and hybridized to colony filters of a human chromosome 19 cosmid library, as previously described (Olsen *et al.*, 1994). The relative positions of positively hybridizing cosmid contigs were established by high-resolution FISH procedures, as previously described (Ashworth *et al.*, 1995; Brandriff *et al.*, 1994). **A.)** The locations of several genes surrounding *CD97* in this region are indicated above the line, along with the corresponding cosmid identities indicated below the line. **B.)** The locations of several genes mapping to the same region as *EMR1* are indicated as detailed above.


A
19ptel
15929
JUNB
13.0
15496
CACNL1A4
23890
RFX1
14.0
11664
CD97
15.0
27622
NOTCH3
15734
MEL
16.0
19pcen
19p13.2
19p13.1
B
19ptel
5
19377
RFX2
6
16575
C3
18561
EMR1
7
15952
INSR
8
19pcen
19p13.3
19p13.2

and high degree of structural similarity that characterizes CD97 and EMR1 is also characteristic of RFX1 and RFX2, which share 44% amino acid sequence identity. Sequence data suggest that both gene pairs have diverged as long as 250 million years ago (data not shown).

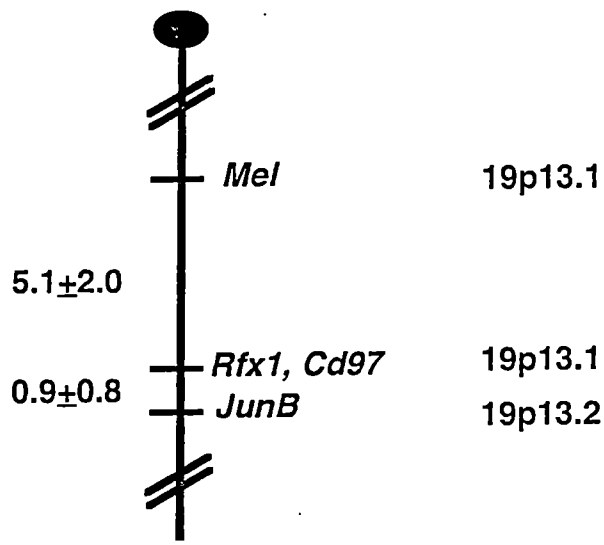
Although it is known that *Emr1* and *Rfx2* are tightly linked in mouse (Lin *et al.*, 1997), the chromosomal position of *Cd97* has not been determined. To localize this mouse gene, we followed the segregation of variant restriction fragments in the progeny of a *Mus musculus* X *Mus spretus* interspecific backcross (Doyle *et al.*, 1996; Stubbs *et al.*, 1996). As probe, a 1636 bp PCR product was amplified from the murine *Cd97* cDNA with primers mCD97.1-ctgtccctgatggtgaaggag [nt 1104-1124 of the (+) strand] and mCD97.8-cctgcctcgtgtaccaggcag [nt 2719-2739 of the (-) strand] (GenBank accession #Y18365; Hamann *et al.*, submitted). *Cd97* was mapped to central mouse chromosome 8, between *Rfx1* and *Mel*, a position concordant with the position determined for the human gene (Figure 2). Since *Cd97* and *Notch3* are located on chromosomes 8 and 17, respectively, (Gao *et al.*, 1998; EAC, manuscript in preparation), these two markers, separated by less than approximately 800 kb in human chromosome 19p, flank the borders of at least two distinct regions of human-mouse syntenic homology (Figure 1A).

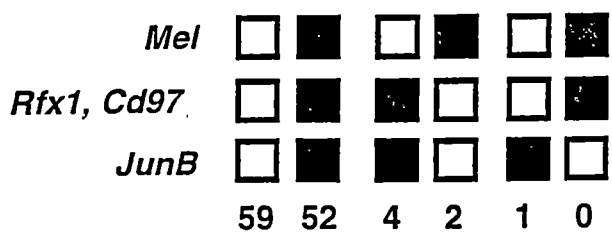
These results confirm that the physical linkage of *CD97* and *RFX2* genes, on the one hand, and *EMR1* and *RFX1* loci on the other, is conserved in human and mouse. These data also suggest an ancient duplication event containing these genes. This study provides new clues to the evolution of the EGF-TM7 family and insight into the sequences located on human chromosome 19.

Figure 2. Segregation pattern of *Cd97* and surrounding genes mapping to mouse chromosome 8.

To localize the mouse ortholog of the human *CD97* gene, we followed the segregation of variant restriction fragments detected by a murine *Cd97* probe in 118 progeny of a previously-described *Mus musculus* X *Mus spretus* interspecific backcross [(C3Hf/Rl-*Mgf* ^{Sl2ENURg}/+ X *Mus spretus)* X C3Hf/Rl; Doyle *et al.*, 1996; Stubbs *et al.*, 1996)]. This backcross has previously been typed for multiple markers, mapping to all mouse chromosomes, utilizing the Map Manager data analysis program (Manly, 1994) with standard errors derived using standard statistical methods (Silver, 1985). Southern blots carrying restriction digests of DNA samples prepared from each of the backcross progeny were hybridized with radiolabled probes as previously described (Stubbs *et al.*, 1990). A partial chromosome 8 linkage map showing the location of *Cd97* in relation to linked genes appears above the chromosome segregation pattern. Recombination distances between loci are shown to the left of the chromosome in centimorgans. The positions of the human homologs of each locus are given at the right of the chromosome. The segregation patterns of *Cd97* and flanking loci are shown at the bottom. Each column represents a type of parental or recombinant chromosome. The number of backcross animals inheriting a particular type of chromosome is listed at the bottom of each column. White boxes represent the inheritance of a *Mus spretus* allele for a particular set of loci, while black boxes denote the presence of C3Hf/Rl alleles for the same genes.

MOUSE CHROMOSOME 8





ACKNOWLEDGMENTS

We thank Anca Georgescu, Beverly Stanford and Estela Generoso for excellent technical help, and Mark Shannon for critical review of this manuscript. This work was supported by grants from the U.S. DOE under contract no. W-7405-ENG-48 (Lawrence Livermore National Laboratory).

REFERENCES

Ashworth, L. K., Batzer, M. A, Brandriff, B., Branscomb, E., de Jong, P., Garcia, E., Garnes, J., Gordon, L., Lamerdin, J. E., Lennon, G., Mohrenweiser, H., Olsen, A., Slezak, T., and Carrano, A. V. (1995). A metric physical map of human chromosome 19. *Nature Genet* **11**:422-427.

Baud, V., Chissoe, S. L., Viegas-Pequignot, E., Diriong, S., N'Guyen, V. C., Roe, B. A., and Lipinski, M. (1995). EMR1, an unusual member in the family of hormone receptors with seven transmembrane segments. *Genomics* **26**:334-344.

Brandriff, B. F., Gordon, L. A., Fertitta, A., Olsen, A. S., Christensen, M., Ashworth, L. K., Nelson, D. O., Carrano, A. V., and Mohrenweiser, H. W. (1994). Human chromosome 19p: a fluorescence in situ hybridization map with genomic distance estimates for 79 intervals spanning 20 Mb. *Genomics* **23**:582-591.

Doyle, J., Hoffman, S., Ucla, C., Reith, W., Mach, B., and Stubbs, L. (1996). Locations of human and mouse genes encoding the RFX1 and RFX2 transcription factor proteins. *Genomics* **35**:227-230.

Gao, X., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Gridley, T. (1998). Assignment of the murine Notch2 and Notch3 genes to chromosomes 3 and 17. *Genomics* **49**:160-161.

Hamann, J., Eichler, W., Hamann, D., Kerstens, H.M., Poddighe, P.J., Hoovers, J.M., Hartmann, E., Strauss, M., and van Lier, R.A. (1995). Expression cloning and

chromosomal mapping of the leukocyte activation antigen CD97, a new seven-span transmembrane molecule of the secretion receptor superfamily with an unusual extracellular domain. *J Immunol* **155**:1942-1950.

Hamann, J., Vogel, B., van Schijndel, G. M., and van Lier, R. A. (1996). The seven-span transmembrane receptor CD97 has a cellular ligand (CD55, DAF). *J Exp Med* **184**:1185-1189.

Lin, H.H., Stubbs, L.J., and Mucenski, M.L. (1997). Identification and characterization of a seven transmembrane hormone receptor using differential display. *Genomics* **41**:301-308.

Manly, K. F. (1993). A MacIntosh program for the storage and analysis of experimental genetic mapping data. *Mammal Genome* **4**:303-313.

McKnight, A. J., Macfarlane, A. J., Seldin, M. F., and Gordon, S. (1997). Chromosome mapping of the Emr1 gene. *Mammal Genome* **8**:946.

McKnight, A.J. and Gordon, S. (1998). The EGF-TM7 family: unusual structures at the leukocyte surface. *J Leukoc Biol* **63**:271-280.

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

Pugliatti, L., Derre, J., Berger, R., Ucla, C., Reith, W., and Mach, B. (1992). The genes for MHC class II regulatory factors RFX1 and RFX2 are located on the short arm of chromosome 19. *Genomics* **13**:1307-1310.

Reith, W., Ucla, C., Barras, E., Gaud, A., Durand, B., Herrero-Sanchez, C., Kobr, M., and Mach, B. (1994). RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins. *Mol Cell Biol* **14**:1230-1244.

Silver, J. (1985). Confidence limits for estimates of gene linkage analysis of recombinant inbred strains. *J Hered* **76**:436-440.

Stubbs, L., Carver, E., Ashworth, L., and Lopez-Molina, L. (1996). Location of the DBP transcription factor gene in man and mouse. *Mammal Genome* **7**:65-67.

Stubbs, L., Huxley, C., Hogan, B., Evans, T., Fried, M., Duboule, D., and Lehrach, H. (1990). The HOX-5 and surfeit gene clusters are linked in the proximal portion of mouse chromosome 2. *Genomics* **6**:645-650.

SECTION 3. COMPARATIVE MAPPING OF HUMAN 19p13.1 SURROUNDING THE NOTCH3 GENE AND OLFACTORY GENE CLUSTERS: DELINEATION OF HOMOLOGY GROUPS ON MOUSE CHROMOSOMES 8, 10, AND 17

Comparative mapping of human 19p13.1 surrounding the *NOTCH3* gene and olfactory gene clusters: delineation of homology groups on mouse chromosomes 8, 10, and 17

[1,2]Ethan A. Carver, [2,3]Johannah Doyle, [1]Anne S. Olsen, [1]Linda Ashworth, [1,4]Lisa Stubbs

[1]Human Genome Center, Lawrence Livermore National Laboratory, P.O. Box 808, L-452, Livermore, CA 94550, USA

[2]University of Tennessee-Oak Ridge School of Biomedical Sciences, Oak Ridge National Laboratory, P.O. Box 2009, Oak Ridge, TN 37831-8077, USA

[3]Present Address: National Human Genome Research Institute, National Institutes of Health, 49 Convent Drive, Bethesda, MD 20892-4431

[4] Corresponding author

tel. (925) 422-8473

fax. (925) 422-2282

email: Stubbs5@llnl.gov

Running title: Comparative mapping of human 19p and mouse genomes

Manuscript in Preparation

INTRODUCTION

In this study, we describe the comparative mapping of a ~2 Mb region of human chromosome 19p13.1 and related regions of the mouse genome. The *NOTCH3* (Notch gene homolog 3) gene, mutations in which are associated with stroke and dementia (Joutel *et al.*, 1996), maps to human chromosome 19p13.1, a region with extensive homology to mouse chromosome 8 (Ashworth *et al.* 1995). However, recent studies have localized the *Notch3* gene to mouse chromosome 17 (Gao *et al.*, 1998), suggesting that the homology relationships within this region might be more complex than originally assumed. Two clusters of olfactory receptor genes (Olfr) have been mapped to human 19p13.1 using conserved Olfr family probes (Ashworth *et al.*, 1995), and recent studies have indicated that the two clusters are comprised of members of two distinct Olfr gene subfamilies (Carver et al., 1998). The human metric map and emerging DNA sequence data were used to generate conserved gene probes surrounding *NOTCH3* to delineate the organization of this region in the mouse genome.

Olfactory receptor genes are relatively small, with most or all of the coding sequence contained within a single exon. The genes encode seven transmembrane domain proteins which share characteristic sequence motifs including a central variable region that corresponds to a putative ligand binding site (Buck and Axel, 1991). *Olfr* genes are often found in tandem arrays, with closely related genes clustered together (Ben-Arie *et al.*, 1994, Sullivan *et al.*, 1996; Carver *et al.*, 1998). By searching human 19p13.1 clone contigs, we identified cosmids surrounding *NOTCH3* that contained sequences that hybridized positively to the consensus olfactory receptor gene probe. We generated specific human olfactory receptor sequences using PCR amplification techniques. Sequence comparisons of segments generated from different 19p13.1 cosmids and known gene sequences confirmed the presence of two distinct olfactory receptor gene subfamilies. PCR fragments representing each of the two subfamilies were mapped in mouse, using a

well-typed interspecific backcross (IB) (Stubbs *et al.*, 1996; Doyle *et al.*, 1996a, 1997; Kim *et al.*, 1997; Carver *et al.*, 1998). Sequences detected by one subfamily probe are clustered together adjacent to *Notch3* in the mouse, confirming the presence and extending the length of a small segment of homology between human chromosome 19p13.1 and central Mmu17. The second cluster of 19p13.1 *Olfr* genes was known from previous studies to be similar in sequence to a large, complex Olfr subfamily related to canine gene, CfOLF4 (Carver *et al.*, 1998). Members of this canine Olfr gene subfamily map to at least one other site in human 19p13.2, while related mouse genes are located in at least two regions in the mouse, including proximal Mmu9 (in a region related to 19p13.2) and central Mmu10 (Carver *et al.*, 1998). Since no syntenic homology between human 19p13.1 and central Mmu10 had been documented, the evolutionary relationship between the related Olfr gene clusters remained uncertain. Examination of the human DNA 19p13.1 sequences permitted us to identify a unique gene, *SLC1A6* (Solute carrier family 1, member 6), located between *NOTCH3* and the more distal of the two 19p13.1 *Olfr* gene clusters. *Slc1a6* co-segregated with the *Olfr* genes in Mmu10, confirming the existence of a new region of 19p13.1 homology in the mouse genome.

In human chromosome 19p13.1, a group of single copy genes is organized as follows: telomere- *CD97* (CD97 antigen), *SLC1A6*, *NOTCH3*, and *MEL* (Cell line NK14 derived transforming oncogene) –centromere (Ashworth *et al.*, 1995). *Cd97* and *Mel* are also linked together on mouse chromosome 8, although the homologous segments containing these two genes are separated by an insertion of genes related to human chromosome 4q. Therefore, two relatively large gene-rich segments whose linkage is conserved in human 19p13.1 and Mmu8 have been interrupted in both rodent and human lineages by insertion of unrelated genomic material. These and other data suggest that this region may be especially prone to rearrangements during the evolution of both mouse and man. Homology region borders defining the four distinct segments of mouse syntenic homology

within the 6 Mb 19p13.1 region are occupied by densely packed members of olfactory receptor, cytochrome P450, zinc-finger-containing genes and other gene families, providing further evidence of the potential relationship between evolutionary instability and the presence of locally repeated DNA sequences (Stubbs *et al.*, 1996; Kim *et al.*, 1997; Carver and Stubbs, 1997). Comparative sequencing of related mouse and human chromosome 19 regions is now under way, and should provide new insights regarding the evolution of tandemly clustered gene families and the role they have played in genome evolution.

MATERIALS AND METHODS

Molecular probes

All markers analyzed in this study and the variant fragments used to assign them to mouse chromosome maps are described in Table 1. *Notch3* was obtained from Dr. M. Lardelli (Karolinska Institute; Stockholm, Sweden). The templates consisted of human cosmid DNA (100ng/100µl reaction) prepared by Qiagen columns (Qiagen, Valencia, CA) as described by the manufacturer. All PCR reactions were conducted in 100 µl volumes, using 0.2 U Taq polymerase (Perkin-Elmer/Cetus, Norwalk, CT) in standard buffers supplied by the manufacturer, using the following program in a GeneAmp 9600 thermocycler (Perkin-Elmer/Cetus, Norwalk, CT): 95°C, 4 min.; (94°C, 1 min.; 55 °C, 1 min.; 72°C, 2 min.) X 35 cycles; 72°C, 4 min.

Isolation of olfactory receptor-specific probes

Human cosmid clones were used as templates to isolate specific olfactory receptor gene segments. A pair of degenerate oligonucleotide primers corresponding to the protein sequence PMY(L/F)FL (primer NL61) and TC(A/G)SHL (primer NL63), which are the olfactory receptor transmembrane domains 2 and 6 (Buck and Axel, 1991), were used for PCR analysis. The sequences utilized for PCR included; primer NL61: 5'CGGAATTCCC(GATC)ATGTA(CT)(CT)T(GATC)TT(CT)(CT) and primer NL63: 5'ATAAGCTTAG(GA)TG(GATC)(GC)(TA)(GATC)(GC)C(GA)CA(GATC)GT (Issel-Tarver and Rine, 1996), which together would PCR amplify a 560-bp product from human genomic DNA.

Table one. Description of probes and variant fragments.

For cloned probes, the type of clone used (c=cDNA, g=genomic) and species of origin are listed (H=human; M=mouse). The sizes of informative variant fragments used to follow the segregation of each marker in *Mus musculus* (M) or *Mus spretus* (S) DNA, and the restriction enzymes used to generate those fragments, are also shown. References describing the probes used and their sources are listed at right (ATCC=American Type Culture Collection, Rockville, MD).

TABLE ONE

Chr.	Locus	Type/ Species	M(kb)	S(kb)	Enz.	Reference
17	*D17Leh180*	M/g	9.8	4.3	*Bam*HI	Bucan et al., 1987
17	*Crya1*	M/c	1.6	2.2	*Pvu*II	King et al., 1982
17	*31431-Olf*	H/c	10	15	*Eco*R1	This paper
17	*15232-Olf*	H/c	9.8	6.2	*Eco*RI	This paper
17	*Notch3*	M/c	2.8	3.3	*Bam*HI	Lardelli et al., 1994
17	*Col11a2*	M/c	7.7	10	*Eco*RI	Stubbs et al., 1993
17	*C3*	M/c	20	5.8	*Hinc*II	Doyle et al., 1996
10	*SCL1A6*	H/c	17	23	*Eco*RI	EST #52422
10	*14281-Olf*	H/c	2.1	2.3	*Eco*RV	This paper
10	*Myb*	M/c	3.1	4.4	*Taq*I	Mucenski et al., 1991
10	*CfOlf4rs2*	D/c	3.5	3.2	*Eco*RV	Issel-Tarver and Rine, 1996
10	*PFKL*	H/c	2.6	4.2	*Eco*RI	ATCC#61648 Levanon et al., 1989
10	*Igtb2*	M/c	9.8	17	*Eco*RV	ATCC#63056 Wilson et al., 1989

Mouse genetic mapping

To localize the mouse orthologs of the human genes, we followed the segregation of variant restriction fragments detected by probes in the progeny of a *Mus musculus* X *Mus spretus* interspecific backcross (IB; [(C3Hf/Rl-*Mgf* ^{Sl2ENURg}/+ X *M. spretus*) X C3Hf/Rl], Doyle *et al.*, 1996a; Stubbs *et al.*, 1996), which has been typed for multiple markers mapping to all mouse chromosomes. Radiolabled probes were prepared and hybridized as previously described (Stubbs *et al.*, 1990) to Southern blots carrying restriction digests of DNA samples prepared from each of 118 backcross progeny. The mouse probes were hybridized at higher stringencies (hybridized in buffer containing 50% formamide; washed at 0.2 X SSC at 65°C), and human markers were hybridized under conditions of lower stringency (hybridized in buffer containing 45% formamide at 42°C; washed at 1 X SSC at 65°C). Mapping data were interpreted using Map Manager data analysis program (Manly, 1993), with standard errors derived using standard statistical methods (Silver, 1985).

DNA sequence analysis

The PCR products were subcloned into the TA cloning vector (Invitrogen, Inc), per manufacturer's protocols using fluorescence-based DNA cycle-sequencing kits (Dye terminator and Dye primer kits, Perkin Elmer-Cetus), and reactions analyzed on an ABI373 automated sequencer. For the dye terminator sequencing reaction, 5 ul of each reaction mixture was used for the sequencing reaction. For the dye primer sequencing reaction, 2 ul of PCR products was directly used for sequencing without any purification. Database search and sequence alignments were analyzed using GCG software, Version 8.1 (Genetics Computer Group, Inc).

RESULTS

Mapping of NOTCH3 and surrounding olfactory receptors in mouse

To better define homology relationships within the 19p13.1 region, we followed the segregation of variant restriction fragments detected by specific probes in the progeny of a *Mus musculus* X *Mus spretus* interspecific backcross. Results of this analysis are summarized in Figure 1. The previously reported location of mouse *Notch3* provided the first clue to the existence of a new region of syntenic homology between human 19p13.1 and mouse chromosome 17 (Gao *et al.*, 1998). To define the extent of this homology segment, we isolated fragments of two *Olfr* gene sequences, 31431-Olf and 15232-Olf, from two cosmids located between *NOTCH3* and *MEL* in the human 19p13.1 physical map (Ashworth *et al.*, 1995; Figure 2). These two Olfr sequences were mapped to mouse 17, co-segregating with the murine *Notch3* locus (Figure 1). Olfr-PCR fragments isolated from a second set of cosmids, located between *NOTCH3* and *CD97*, were also sequenced and mapped using the IB system. These probes detected a much more extensive subfamily of Olfr gene sequences, and most of the fragments were not polymorphic or were difficult to distinguish from other similarly-sized fragments (data not shown). As previously reported, *CfOLF4* (*Canis familiaris* olfactory receptor gene 4) detected two independent, related gene clusters when hybridized to gridded human chromosome 19 cosmid arrays (Carver *et al.*, 1998). The Olfr cluster located between *NOTCH3* and *CD97* was one of these clusters, and one of the positively-hybridizing cosmids was clone 14281. PCR using cosmid 14281 as a template yielded a single Olfr sequence, 14281-Olf; this probe was mapped to the region of mouse chromosome 10 where one variant set of CfOlf4-related gene sequences (*CfOlf4-rs2*) had been mapped in the earlier study (Carver *et al.*, 1998; Figure 1). These mapping data indicated the presence of a new homology group associated with genes within this region of human chromosome 19p13.1 and mouse chromosome 10.

Figure 1. Segregation patterns of *Notch3* and surrounding olfactory receptor genes mapping to mouse chromosome 17 and 10.

Sequences were analyzed in 118 progeny of an IB system (IB; [(C3Hf/Rl-*Mgf* ^{Sl2ENURg}/+ X *Mus spretus*) X C3Hf/Rl], Doyle *et al.*, 1996; Stubbs *et al.*, 1996) typed for markers derived from all mouse chromosomes. Probes and variant fragments used to follow each locus are described in Table 1. The partial chromosome linkage maps showing the location of each *Olfr* gene in relation to linked genes appear above each chromosome segregation pattern. The positions of the human homologs of each locus are given, when known, at the left of the chromosome. Each column represents a type of parental or recombinant chromosome; the number of backcross animals inheriting a particular type of chromosome is listed at the bottom of each column. White boxes represent the inheritance of a *M. spretus* allele for a particular set of loci, while black boxes denote the presence of the C3Hf/Rl allele for the same gene. Recombination distances between loci are shown to the right of the columns in centimorgans. **A)** Segregation of *Notch3* and olfactory receptors which map to mouse chromosome 17. **B)** Segregation of *Slc1a6* and olfactory receptors which map to mouse chromosome 10.

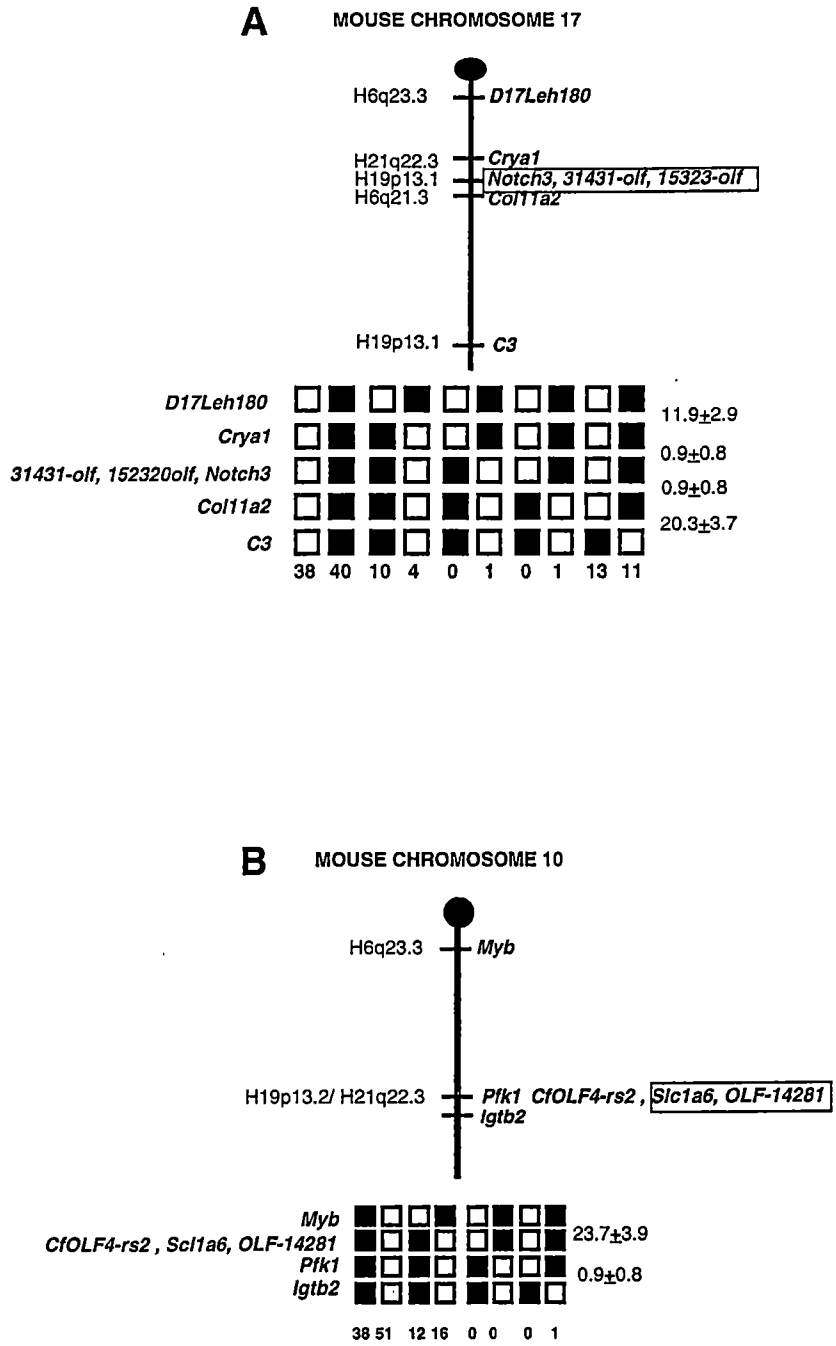

A
MOUSE CHROMOSOME 17
H6q23.3
D17Leh180
H21q22.3
Crya1
H19p13.1
Notch3, 31431-olf, 15323-olf
H6q21.3
Col11a2
H19p13.1
C3
D17Leh180
Crya1
31431-olf, 152320olf, Notch3
Col11a2
C3
11.9±2.9
0.9±0.8
0.9±0.8
20.3±3.7
38 40 10 4 0 1 0 1 13 11
B
MOUSE CHROMOSOME 10
H6q23.3
Myb
H19p13.2/ H21q22.3
Pfk1 CfOLF4-rs2 , Slc1a6, OLF-14281
Igtb2
Myb
CfOLF4-rs2 , Scl1a6, OLF-14281
Pfk1
Igtb2
23.7±3.9
0.9±0.8
38 51 12 16 0 0 0 1

Figure 2. Location of genes in the human metric map, and their locations in the mouse genome.

The human metric map and distances from the p-telomere are indicated at the bottom of the figure. Cosmid and BAC clone designations are indicated underneath the markers that hybridize positively for them. These markers were used to anchor and extend the mouse homology groups that appear at the top of the figure. The homology group sizes are illustrated with respect to their known distances in the human genome.

19pcen
19ptel
p13.1
MY09B
MEL
CYPs
OLFR
OLFR
OLFR
OLFR
NOTCH3
OLFR
SLC1A6
OLFR
OLFR
ZNF
EMR1 like
OLFR
CD97
L82 cDNA
RFX1
CACNL1A4
15734
15232
31431
12909
BC629401
14281
29368
27752
11664
23890
28441
MMU 8
~3Mb
Unknown
~100kb
MMU17
~700kb
Unknown
~100kb
MMU10
~200kb
Unknown
~250kb
MMU 8
~2Mb
19pcen
19ptel
17
16
15
Mb

To verify the existence of this homology segment and to define its physical extent, we searched the genomic sequence of the human 19p13.1 region (http://www-bio.llnl.gov/genome/genome.html) for additional, non-*Olfr* genes. Sequences from the region near 14281-Olf (accession #004659; Lamerdin *et al.*, 1998) contained a significant match to a human EST clone, #52422, which was obtained from the I.M.A.G.E. consortium (Lennon *et al.*, 1996). The sequence of this EST and the human genomic sequence identified a gene as a member of the Solute carrier family, subsequently designated as *SLC1A6*. IB analysis of the SLC1A6 EST probe localized the related gene to mouse chromosome 10, co-segregating with the 14281-Olf locus (Figure 1). These data confirmed the existence of the new homology group between human 19p13.1 and mouse chromosome 10.

Sequence and classification of 19p13.1 *Olfr* genes

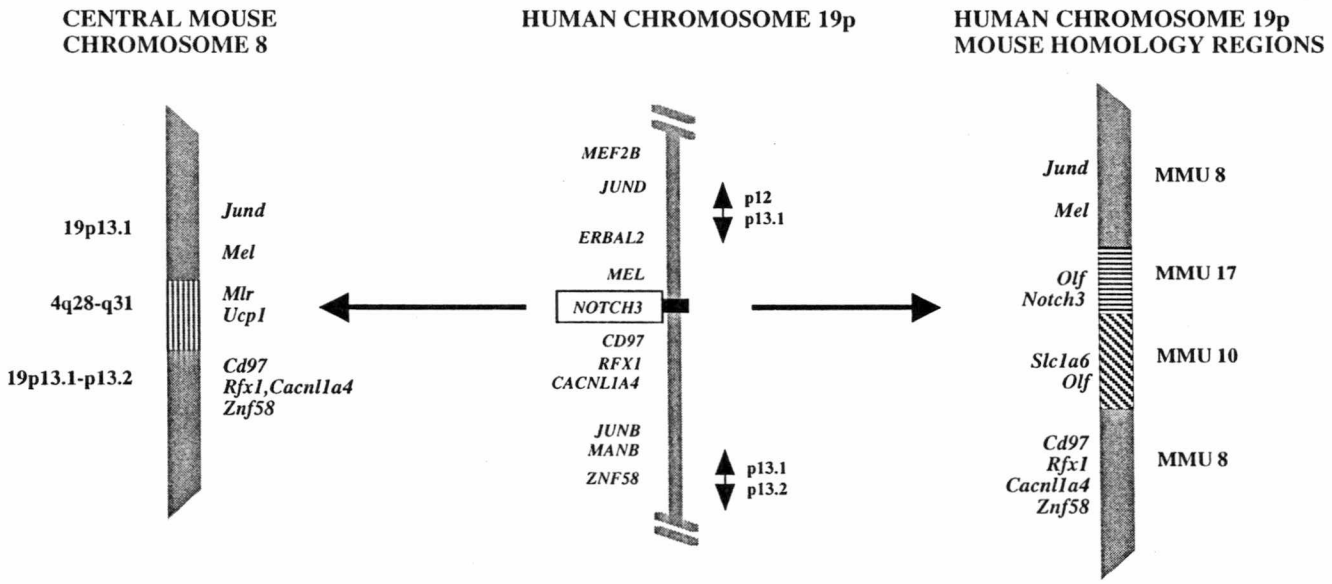
According to one classification system, olfactory genes that share at least 40% amino acid identity belong to the same family of genes, and those that share at least 60% amino acid identity belong to the same subfamily of genes (Ben-Arie *et al.*, 1994). Analysis of the 14281-Olf showed it to be 75% identical to the canine olfactory gene CfOLF4 at the DNA level (data not shown). By contrast, 14281-Olf has 32% and 33% amino acid sequence identity with the 15232-Olf and 31431-Olf olfactory receptor genes, respectively. These data demonstrate, as previously suggested (Carver et al., 1998), that *Olfr* genes located proximally (e.g. 14281-Olf) and distally (e.g. 15232-Olf and 31431-Olf) to *NOTCH3* in human 19p13.1 are of separate families of olfactory receptor genes. 15232-Olf and 31431-Olf olfactory genes are located very close to each other and share 71% amino acid sequence identity (Figure 2; data not shown) demonstrating that these neighboring genes belong to the same subfamily.

Discussion

Previous studies have located the borders of homology between human chromosome 19 and related regions of the mouse genome to intervals that are occupied by clustered gene family members (Stubbs *et al.*, 1996; Kim *et al.*, 1997; and Carver and Stubbs, 1997). The detailed comparative mapping of 19p13.1 confirms this general observation. Borders defining the four segments of human-mouse homology within this 6 Mb human segment contain Znf, Olfr, and/or Cyp family gene members (Figure 2). These regions have been, or are in the process of being, sequenced in humans (http://www-bio.llnl.gov/genome/genome.html), and related regions are scheduled for sequencing in the mouse. The region located between *MEL* and *CD97* in both humans and mice appears to have been particularly prone to rearrangements (Figure 3). Genes located proximal to *MEL* and distal to *CD97* in human chromosome 19, respectively, are conserved in order and organization in human 19p13 and mouse chromosome 8 (Figure 2 and 3, Doyle *et al.*, 1996, 1997). The distance between *CD97* and *MEL* is ~1.9 Mb in human, whereas the orthologs of these genes are separated by ~5 cM in the mouse genome. However, these two conserved segments are not contiguous in either human or mouse; in both species, the two intervals are interrupted by insertion of unrelated genetic material. Genes whose homologs are located in human chromosome 4q28-q31, such as *Mlr* and *Ucp1*, are inserted between *Mel* and *Cd97* on mouse chromosome 8 (Figure 3). In contrast, *MEL* and *CD97* on human chromosome 19 show the insertion of two separate homology groups containing genes related to mouse chromosome 10 and mouse chromosome 17. Therefore, rearrangements appear to have occurred within this interval independently in both lineages. The human 19p13.1 region sequence will be a critical tool to further characterize the homology borders in the mouse. However, a real understanding of the evolutionary history of this region will require in depth study of the linkage of resident genes in other mammals and more distantly related vertebrate species.

Figure 3. Location of genomic regions in human and mouse subject to evolutionary rearrangements.

The region encompassing *MEL* and *CD97* shows linkage in both mouse chromosome 8 and human chromosome 19, although the interval between the two markers is not conserved between the two species. Regions conserved between Mmu 8 and human chromosome 19p13 are illustrated as gray bars. Regions that are not conserved are shown as different hatched boxes, corresponding to their individual homology groups. The left side of the figure illustrates the mouse chromosome 8 region with human homology regions shown on the left and individual genes on the right. As indicated by the vertical hatched box, there is an insertion of material that has homology to human chromosome 4q in the interval between *Mel* and *Cd97*. The center section shows human chromosome 19. This region serves as an anchor to compare the homology groups. On the left side of the figure the conserved gene markers are shown. The *NOTCH3* gene is boxed, to highlight the region that is not conserved between human chromosome 19p13.1 and Mmu 8. The right portion of the figure illustrates the 19p13 human chromosome homology regions in the mouse. The horizontal and angle-hatched boxes show the insertion of two different homology groups from mouse chromosome 17 and 10, respectively. The gene probes are shown on the left of the figure, while the mouse homology regions are indicated on the right.


CENTRAL MOUSE CHROMOSOME 8
HUMAN CHROMOSOME 19p
HUMAN CHROMOSOME 19p MOUSE HOMOLOGY REGIONS
19p13.1
4q28-q31
19p13.1-p13.2
Jund
Mel
Mlr
Ucp1
Cd97
Rfx1,Cacnl1a4
Znf58
MEF2B
JUND
ERBAL2
MEL
NOTCH3
CD97
RFX1
CACNL1A4
JUNB
MANB
ZNF58
p12
p13.1
p13.1
p13.2
Jund
Mel
Olf
Notch3
Slc1a6
Olf
Cd97
Rfx1
Cacnl1a4
Znf58
MMU 8
MMU 17
MMU 10
MMU 8

As the borders of regions of homology in the mouse and human genomes are better defined, it is clear that a significant portion of borders seem to lie within regions occupied by locally reiterated sequences (e.g. Saleh *et al.*, 1992; van Deutekom *et al.*, 1996; Dutly and Schnitzel, 1996). These repeated sequences might facilitate genomic rearrangements by providing substrates for illegitimate inter- and intrachromosomal recombination events (e.g. Kulozik *et al.*, 1992; Dutly and Scnitzel, 1996; van Deutekom *et al.*, 1996; Kehrer-Sawatski *et al.*, 1997; Rabbitts, 1994). As demonstrated by these studies, human 19p13.1 contains members of several different tandemly clustered gene families. Our own observations have shown that similar features characterize the bulk of homology borders defined on human chromosome 19 (Stubbs *et al.*, 1996; Kim *et al.*, 1997; Carver and Stubbs, 1997; and unpublished results). A discussion of the next stages of this research, and ways that the hypotheses put forth in this chapter can be tested, is presented in Part IV.

ACKNOWLEDGMENTS

We would like to express our thanks to Gene Rinchik and Dabney Johnson for establishing and conducting initial analyses of the mouse IB mapping system used in this study. Also, we extend our gratitude to Loren Hauser and Melissa York for assistance in the sequencing of the olfactory receptor gene fragments. This work was supported by grants from the U.S. DOE under contract nos. DE-AC05-96ORO22464 (Oak Ridge National Laboratory) and W-7405-ENG-48 (Lawrence Livermore National Laboratory).

REFERENCES

Ashworth, L.K., Batzer, M.A., Brandriff, B., Branscomb, E., deJong, P., Garcia, E., Garnes, J., Gordon, L., Lamerdin, J.E., Lennon, G., Mohrenweiser, H., Olsen, A., Slezak, T., and Carrano, A.V. (1995). A metric physical map of human chromosome 19. *Nature Genet* **11**:422-427.

Ben-Arie, N., Lancet, D., Taylor, C., Khen, M., Walker, N., Ledbetter, D.H., Carrozzo, R., Patel, K., Sheer, D., Lehrach, H., and North, M.A. (1994). Olfactory receptor gene cluster on human chromosome 17: possible duplication of an ancestral receptor repertoire. *Hum Mol Genet* **3**:229-235.

Bucan, M., Herrmann, B.G., Frischauf, A.M., Bautch, V.L., Bode, V., Silver, L.M., Martin, G.R., and Lehrach, H. (1987). Deletion and duplication of DNA sequences is associated with the embryonic lethal phenotype of the t9 complementation group of the mouse t complex. *Genes Dev* **1**:376-385.

Buck, L. and Axel, R. (1991). A novel multigene family may encode odorant receptors: a molecular basis for odor recognition. *Cell* **65**:175-187.

Carver, E.A., Hamann, J., Olsen, A.S., and Stubbs, L. (1999). Physical mapping of *EMR1* and *CD97*, two distantly related genes, to human chromosome 19 and assignment of *Cd97* to mouse chromosome 8. *Mammal Genome* Submitted.

Carver, E.A., Issel-Tarver, L., Rine, J., Olsen, A., and Stubbs, L. (1998). Location of mouse and human genes corresponding to conserved canine olfactory receptor gene subfamilies. *Mammal Genome* **9**:349-354.

Carver, E.A., and L. Stubbs. (1997). Zooming in on the human-mouse comparative map: genome conservation re-examined on a high-resolution scale. *Genome Res* **7**:1123-1137.

Doyle J., Hoffman S., Ucla C., Reith W., Mach B., and Stubbs L. (1996). Locations of human and mouse genes encoding the RFX1 and RFX2 transcription factor proteins. *Genomics* **35**:227-230.

Doyle, J., Hellevuo, K., and Stubbs, L. (1996). The gene encoding adenylyl cyclase VII is located in central mouse chromosome 8. *Mammal Genome* **7**:320-321.

Doyle, J., Ren, X., Lennon, G., and Stubbs, L. (1997). Mutations in the *Cacnl1a4* calcium channel gene are associated with seizures, cerebellar degeneration, and ataxia in tottering and leaner mutant mice. *Mammal Genome* **8**:113-120.

Dutly, F., and Schinzel, A. (1996). Unequal interchromosomal rearrangements may result in elastin gene deletions causing the Williams-Beuren syndrome. *Hum Mol Genet* **5**:1893-1898.

Gao, X., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., and Gridley, T. (1998). Assignment of the murine Notch2 and Notch3 genes to chromosomes 3 and 17. *Genomics* **49**:160-161.

Issel-Tarver, L. and Rine, J. (1996). Organization and expression of canine olfactory receptor genes. *Proc Natl Acad Sci* **93**:10897-10902.

Joutel, A., Corpechot, C., Ducros, A., Vahedi, K., Chabriat, H., Mouton, P., Alamowitch, S., Domenga, V., Cecillion, M., Marechal, E., Maciazek, J., Vayssiere, C., Cruaud, C., Cabanis, E-A., Ruchoux, M., Weissenbach, J., Bach, J., Bousser M., and Tourneir-Lasserve, E. (1996). Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke and dementia. *Nature* **383**:707-710.

Kehrer-Sawatzki, H., Haussler, J., Krone, W., Bode, H., Jenne, D.E., Mehnert, K.U., Tummers, U., and Assum, G. (1997). The second case of a t(17;22) in a family with neurofibromatosis type 1: sequence analysis of the breakpoint regions. *Hum Genet* **99**:237-247.

Kim, J., Ashworth, L., Branscomb, E., and Stubbs, L. (1997). The human homolog of a mouse-imprinted gene, *Peg3,* maps to a zinc finger gene-rich region of human chromosome 19q13.4. *Genome Res* **7**:532-540.

King, C.R., Shinohara, T., and Piatigorsky, J. (1982). alpha A-crystallin messenger RNA of the mouse lens: more noncoding than coding sequences. *Science* **215**:985-987.

Kulozik, A.E., Bellan-Koch, A., Kohne, E., and Kleihauer, E. (1992). A deletion/inversion rearrangement of the beta-globin gene cluster in a Turkish family with delta beta zero-thalassemia intermedia. *Blood* **79**:2455-2459.

Lamerdin, J.E., McCready, P.M., Skowronski, E., Adamson, A.W., Burkhart-Schultz, K., Gordon, L., Kyle, A., Ramirez, M., Stilwagen, S., Phan, H., Velasco, N., Garnes, J., Danganan, L., Poundstone, P., Christensen, M., Georgescu, A., Avila, J., Liu, S., Attix, C., Andreise, T., Trankheim, M., Amico-Keller, G., Coefield, J., Duarte, S., Lucas, S., Bruce, R., Thomas, P., Quan, G., Kronmiller, B., Arellano, A., Montgomery, M., Ow, D., Nolan, M., Trong, S., Kobayashi, A., Olsen, A.S. and Carrano, A.V. (1998). Sequence analysis of an ~1.5 Mb olfactory receptor (OLFR) cluster. Unpublished.

Lardelli, M., Dahlstrand, J., and Lendahl, U. (1994). The novel Notch homologue mouse Notch 3 lacks specific epidermal growth factor-repeats and is expressed in proliferating neuroepithelium. *Mech Dev* **46**:123-136.

Lennon, G., Auffray, C., Polymeropoulos, M., and Soares, M.B. (1996). The I.M.A.G.E. Consortium: an integrated molecular analysis of genomes and their expression. *Genomics* **33**:151-152.

Levanon, D., Danciger, E., Dafni, N., Bernstein, Y., Elson, A., Moens, W., Brandeis, M., and Groner, Y. (1989). The primary structure of human liver type phosphofructokinase and its comparison with other types of PFK. *DNA* **8**:733-743

Manly, K.F. (1993). A MacIntosh program for the storage and analysis of experimental genetic mapping data. *Mammal Genome* **4**:303-313.

Mucenski, M.L., McLain, K., Kier, A.B., Swerdlow, S.H., Schreiner, C.M., Miller, T.A., Pietryga, D.W., Scott, W.J. Jr., and Potter, S.S. (1991). A functional c-myb gene is required for normal murine fetal hepatic hematopoiesis. *Cell* **65**:677-689.

Rabbitts, T H. (1994). Chromosomal translocation in human cancer. *Nature* **372**:143-149.

Saleh, M., Selleri, L., Little, P.F., and Evans, G.A. (1992). Isolation and expression of linked zinc finger gene clusters on human chromosome 11q. *Genomics* **14**:970-978.

Saunders, A.M., and Seldin, M.F. (1990). A molecular genetic linkage map of mouse chromosome 7. *Genomics* **8**:525-35.

Silver, J. (1985). Confidence limits for estimates of gene linkage analysis of recombinant inbred strains. *J Heredity* **76**:436-440.

Stubbs, L., Carver, E., Ashworth, L., Lopez-Molina, L. (1996). Location of the DBP transcription factor gene in man and mouse. *Mammal Genome* **7**:65-67.

Stubbs, L., Carver, E.A., Shannon, M.E., Kim, J., Geisler, J., Generoso, E.E., Stanford, B.G., Dunn, W.C., Mohrenweiser, H., Zimmermann, W., Watt, S.M., and Ashworth L.K. (1996). Detailed comparative map of human chromosome 19q and related regions of the mouse genome. *Genomics* **35**:499-508.

Stubbs, L., Huxley, C., Hogan, B., Evans, T., Fried, M., Duboule, D., and Lehrach, H. (1990). The HOX-5 and surfeit gene clusters are linked in the proximal portion of mouse chromosome 2. *Genomics* **6**:645-650.

Stubbs, L., Lui, V.C., Ng, L.J., and Cheah, K.S. (1993). The alpha 2(XI) collagen gene lies within 8 kb of Pb in the proximal portion of the murine major histocompatibility complex. *Mammal Genome* **4**:95-103.

Sullivan, S.L., Adamson, M.C., Ressler, K.J., Kozak, C.A., and Buck, L.B. (1996). The chromosomal distribution of mouse odorant receptor genes. *Proc Natl Acad Sci* **93**:884-888.

van Deutekom, J.C., Bakker, E., Lemmers, R.J., van der Wielen, M.J., Bik, E., Hofker, M.H., Padberg, G.W., and Frants, R.R. (1996). Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counseling and etiology of FSHD1. *Hum Mol Genet* **5**:1997-2003.

Wilson, R.W., O'Brien, W.E., and Beaudet, A.L. (1989). Nucleotide sequence of the cDNA from the mouse leukocyte adhesion protein CD18. *Nucleic Acids Res* **17**:5397.

PART III

LOCALIZATION AND CLONING OF THE MMU2 TRANSLOCATION BREAKPOINT IN THE 1GSO MUTANT MOUSE

CHAPTER 1.

INTRODUCTION

Part III describes a mapping and positional cloning study that was undertaken to localize breakpoints of the reciprocal translocation, 1Gso, associated with the *diver* phenotype in mice (Rutledge *et al.*, 1986). This work represents the first step in the search for a gene(s) responsible for the neurological and embryonic lethal phenotypes exhibited by these animals.

The 1Gso mutation was generated during mutagenesis studies conducted by Dr. Walderico Generoso at Oak Ridge National Laboratory. The original founder male was obtained from the offspring of a triethylenemelamine-treated (SEC X C57BL) F_1 male mated to a (SEC X C57BL) F_1 female. The founder male was semisterile, which is indicative of a balanced translocation (Generoso *et al.*, 1981). 1Gso animals were subsequently found to carry a reciprocal translocation involving chromosomes 2, band E4, and 14, band E3 (Figure 1; Rutledge *et al.*, 1986). Homozygous 1Gso mice have never been documented and presumably die early in embryogenesis, whereas 1Gso heterozygotes exhibit a subtle, complex neurological phenotype. In heterozygous animals, no significant difference in weight is observed between newborn translocation carriers and the normal littermates, but older 1Gso heterozygous animals are weigh less than their wild-type littermates (Rutledge *et al.*, 1986). The neurological phenotype is seen at birth, and is neither debilitating nor progressive. The heterozygous neonates can be recognized by their decreased ability to right themselves as compared to normal littermates (Rutledge *et al.*, 1986). 1Gso mice are distinguished by the inability to swim (Figure 2), hence the name "diver" mice. When challenged or startled, the mutant mice will move backwards first, instead of turning and running like a normal mouse (Figure 3). Often, the animals move backwards in circles. Additional details regarding the 1Gso phenotype are discussed in Section 4.

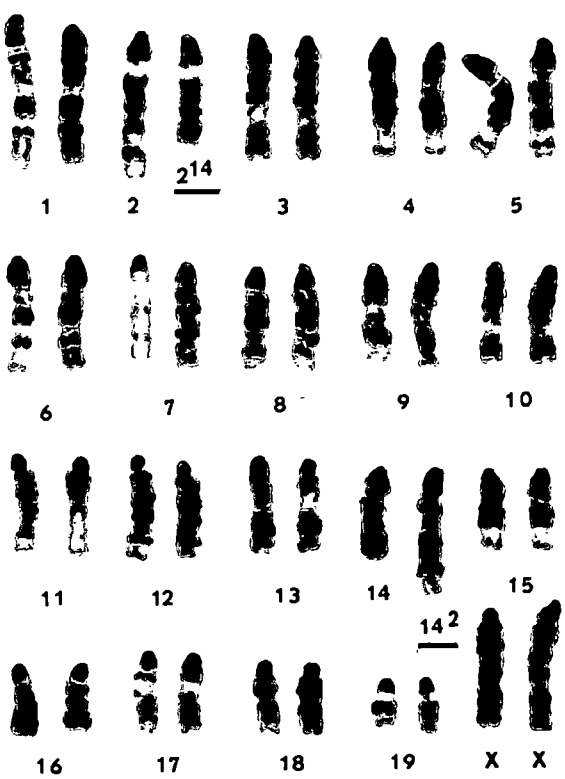
Figure 1. Giemsa-banded karyotype of a 1Gso female mouse.

A Giemsa-band stain and karyotype of the 1Gso translocation heterozygote (Rutledge *et al.*, 1986). This karyotype illustrates that the translocation occurs between chromosome 2E4 and chromosome 14E3. On the upper right, the ideogram of chromosomes 2 and 14 are shown, with approximate breakpoints illustrated with arrows. The karyotype of the entire genome of the 1Gso mutant is shown with chromosome designation below the chromosomes. The translocated chromosomes are underlined. One product of the translocation is formed by the centromere and proximal region of mouse chromosome 2, with the addition of a small distal end of mouse chromosome 14, 2^{14}. The other product of the translocation is formed by the centromere and major portion of mouse chromosome 14, with the addition of the distal end of mouse chromosome 2, 14^{2}. Also, at the gross level, the G-bands do not indicate a large deletion associated with the translocation.



Figure 2. 1Gso "diver" phenotype.

This illustration shows the "diver" phenotype of the 1Gso mice. A normal mouse, shown on the left of the figure, has the inherent ability to swim. 1Gso animals when introduced to the water either "freeze" and do not move at all, or swim in an abnormal manner. Upon attempting to swim, it will paddle in circles, and when submerged, will lose orientation and dive under the water, as seen in the mouse in the right of the figure. Even if it resurfaces, it seems incapable of proper orientation, and will continue back underwater. If the mouse is not physically rescued from the water, it will drown.


2^{14}
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
14^{2}
16
17
18
19
X
X

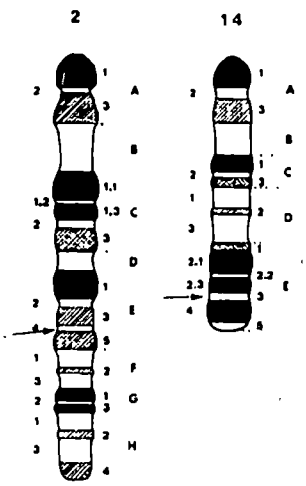
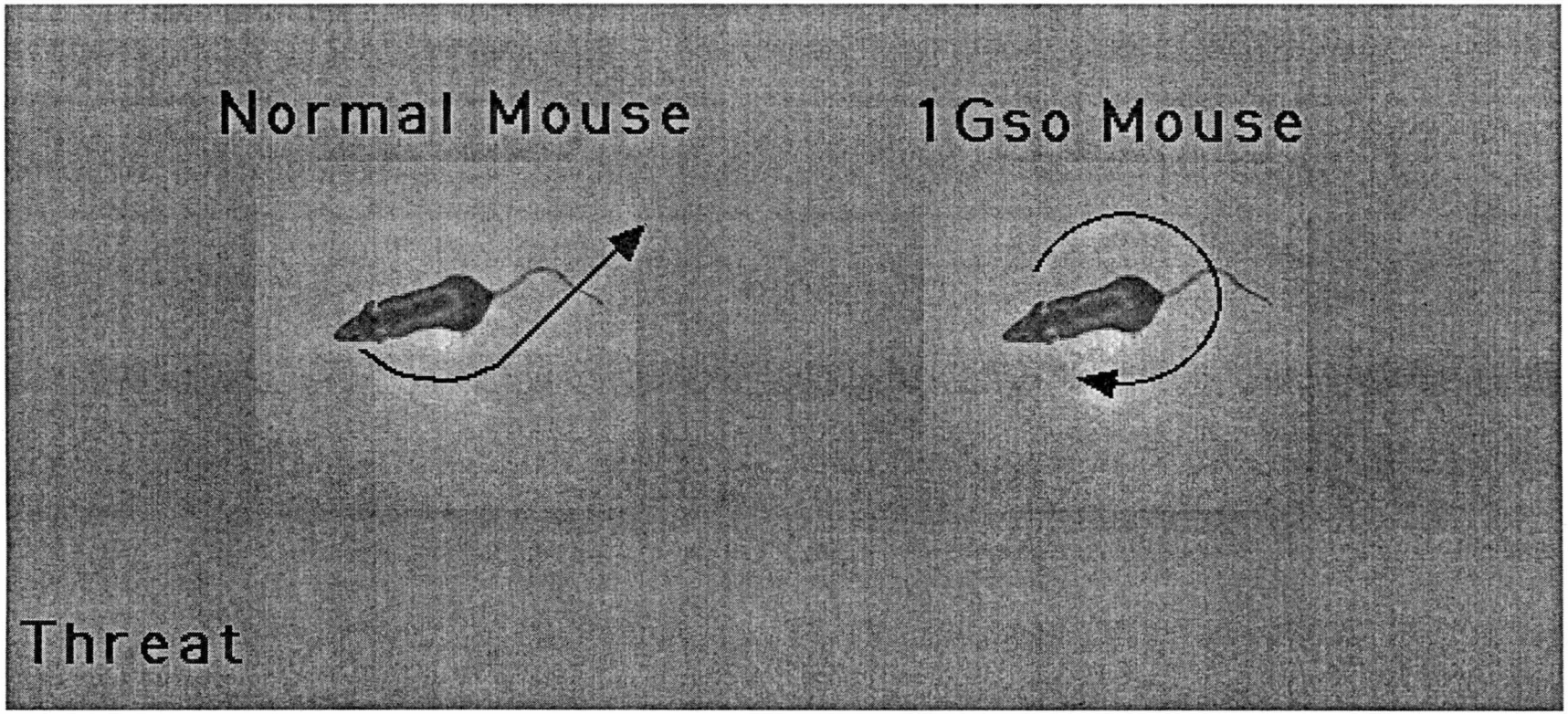

2
14
A
B
C
D
E
F
G
H

Figure 3. 1Gso escape response.

The 1Gso mouse has an abnormal startle response. When normal mice are challenged with a threat, either auditory or visual, they quickly turn around and run away (animal on left). When the 1Gso animal is threatened, it will plant a hindfoot, and push backwards with its front limds away from the threat (animal on the right). This circling behavior is primarily in response to a perceived threat; usually it walks in a normal manner in its cage.


Normal Mouse
1Gso Mouse
Threat

These phenotypes exhibited by 1Gso mice indicate that the rearrangement disrupts gene(s) that are critical to the development and function of the nervous system. Existing knowledge regarding regulation and function of genes critical to the developing nervous system is scanty, despite the fact that numerous genes associated with neurological phenotypes have been identified in mice (Burgess *et al.*, 1995; Burgess *et al.*, 1997; Doyle *et al.*, 1997; Letts *et al.*, 1998; Patil *et al.*, 1995; and others). Moreover, the complexity of neuronal function, and the roles of genes that are active in embryonic neurogenesis, makes the processes difficult to study in humans. Therefore, the study of mutant mice that exhibit neurological phenotypes provides an important venue for scientists to gain access to novel genes associated with the development and function of the mammalian nervous system.

Translocations can disrupt the structure and/or expression of critical genes

A balanced reciprocal translocation results from breaks in two nonhomologous chromosomes followed by a symmetrical exchange of chromatin. This breakage and subsequent exchange can alter the environment of the DNA closely associated with the translocation breakpoints, and give rise to an altered phenotype in the organism. Translocations can cause a phenotype in many different ways. The chromosomal rearrangements can occur inside a gene, thereby disrupting the coding sequence and disrupting the gene's function. The breakage and fusion can introduce stop codons and cause RNA instability and/or premature termination of a protein; translocations can also result in fusion of separate genes to form an entirely different set of proteins, disrupting the functions of each individual protein profoundly (Weterman *et al.*, 1996). Similarly, a translocation can couple regulatory elements from a gene on one chromosome with intact coding sequences located on another chromosome (Morrison *et al.*, 1998). In this case,

both genes would be expected to show aberrant expression, including alterations in the time, amount, and tissue specificity of both proteins (Morrison *et al.*, 1998). Aberrant expression of genes located near, but not physically disrupted by, the translocation has been observed, presumably due to the new heterochromatin environment (Fantes *et al.*, 1995; Surralles *et al.*, 1998).

Because translocations always involve breakage and reunion of sequences at *two* chromosomal sites, the possibility that phenotypes associated with these mutations might represent the effects of disrupting more than one gene must always be considered. In addition, although most of the translocations that have been studied have proved to result from relatively simple breakage-reunion events (Mercer *et al.*, 1993; Corral *et al.*, 1993; Super *et al.*, 1997; Kurahashi *et al.*, 1998), translocations can also be associated with deletions, insertions and other rearrangements at the translocation site (Woychik *et al.*, 1990; Bernasconi, *et al.*, 1998; Edelhoof *et al.*, 1998). In these cases, a mutant phenotype can reflect the disruption of several genes and not a single sequence located immediately adjacent to the breakpoint site. Therefore, it is important to keep in mind the different types of disruptions that can be caused when searching for gene(s) associated with a translocation. These different possibilities and their relevance to the 1Gso translocation are discussed in more detail in Part IV.

DNA rearrangements provide visible access to mutant genes

Chromosomal rearrangements have historically served as physical landmarks for regions containing genes responsible for many human disorders, such as Duchenne muscular dystrophy (de Martinville *et al.*, 1985), neurofibromatosis type 1 (Wallace *et al.*, 1990), aniridia (Gessler *et al.*, 1989), Menkes disease (Mercer *et al.*, 1993), and cancer (reviewed by Rowley, 1998). DNA altered by chromosomal rearrangements has

also served as a guide to localize and identify genes associated with a variety of phenotypes in mice (e.g. Bultman *et al.*, 1992; Herrmann *et al.*, 1990).

Translocations have been particularly useful in positional cloning studies because they disrupt the structure of chromosomes profoundly and in a manner that can easily be seen, often without removing extensive amounts of genetic material. There are four main stages associated with a positional cloning project that are critical to its success. First, the mutation must be localized to a specific chromosomal region. The second stage involves localizing the mutation within a small physical distance that can be approached through conventional positional cloning methods. The third stage involves identifying genes within the targeted regions, and then establishing molecular proof that links a particular candidate gene to a specific mutation completes the positional cloning process.

Translocations provide advantages in each of these different steps. Karyotyping of mouse metaphase cells by G-banding techniques allows each individual chromosome to be identified by the unique pattern of light and dark bands which constitute the chromosome (Committee on Standardized Genetic Nomenclature for Mice, 1972). If a large disruption of the genome has occurred, such as a translocation, then the banding pattern of the affected chromosome will change accordingly (Wurster, 1972). By studying the karyotype of the mutant animal, most chromosomal variations can be quickly identified, as is seen with the 1Gso translocation (Figure 2; Cacheiro *et al.*, 1994; Rutledge *et al.*, 1986). Visualization of chromosomes via Giemsa-staining techniques can be accomplished within a few days. The fact that a translocation can be mapped by examination of metaphase chromosomes from one or a few mutant animals completely avoids a long, expensive and tedious series of breeding and genotyping experiments that normally must be designed to localize a mouse mutation. To better define the position of breakpoint sites, a strategy based on the use of fluorescence *in situ* hybridization (FISH) can be employed (Mercer *et al.*, 1993). This strategy, which will be discussed in detail in

section 2, permits mapping to proceed from chromosome band to a single clone without the need for large breeding or genotyping experiments (reviewed by Stubbs *et al.*, 1997).

Once a translocation has been localized to chromosomal regions through G-band analysis, the surrounding interval can be searched for mapped markers and other features that can aid in positional cloning. The greater the density of mapped markers in the interval flanking the breakpoint, the easier it is to bracket the translocation between two markers that are closely linked. Therefore, the choice of which chromosome to concentrate positional cloning efforts becomes dependent upon reagents that are available for physical map construction.

As has been mentioned, since translocations affect two distinct chromosomes, disruption of gene(s) located on either or both chromosomes might contribute to the phenotype expressed by the mutant animals. Therefore, in the absence of definitive complementation data that link the mutant phenotype to a specific genetic interval, both translocation breakpoint regions must be considered in the search for candidate genes. However, sequences isolated from one chromosomal breakpoint can be used as reagents to isolate sequences from the other side of the rearrangement (Bultman *et al.*, 1992; Menkes *et al.*, 1993; Weterman *et al.*, 1996). Therefore, initial mapping and positional cloning steps are concentrated on the most accessible chromosomal region and do not depend upon prior knowledge of where critical genes reside. The fact that one can "walk" between affected chromosomes of a translocation negates the need to simultaneously map and clone both breakpoint regions, and allows initial research to be concentrated on the chromosome with the highest density of available markers. The studies reported here were concentrated on analysis of mouse chromosome 2 because the greater marker density on that chromosome provided an advantage in the isolation of the translocation breakpoint. Once the translocation breakpoint is isolated, access to the other chromosome is greatly simplified by use of standard molecular biology techniques

similar to those used to isolate the breakpoints of the Is1Gso mutation (Bultman *et al.*, 1992).

Translocations not only are visible on the cytogenetic level, but also provide a concrete means to identify the mutation on a physical level (Gessler *et al.*, 1989; Bultman *et al.*, 1992; Mercer *et al.*, 1993). Any gene disrupted by the translocation can be expected to display differences in size, levels of expression, or sites of transcription in the mutant mice due to the physical rearrangement of DNA. For this reason, testing of candidate genes is also greatly simplified. These points will be addressed in more detail in the ensuing sections that describe the positional cloning of the breakpoint of 1Gso mutant mice.

CHAPTER 2.

MATERIALS AND METHODS

Interspecific backcross mapping

To localize sequences on mouse chromosomes 2 and 14, we followed the segregation of restriction fragment length variants (RFLVs) detected by probes in a subset of the progeny of a previously-described *Mus musculus* X *Mus spretus* interspecific backcross ([IB; (C3Hf/Rl-*Mgf* ^{Sl2ENURg}/+ X *M. spretus*) X C3Hf/Rl], Doyle *et al.*, 1996; Stubbs *et al.*, 1996). Probes were ordered from commercial sources, obtained directly from individual researchers, or generated from sequence by PCR. Southern blots carrying restriction digests of DNA samples prepared from each of 118 backcross progeny were hybridized with radiolabeled probes as previously described (Stubbs *et al.*, 1990). Map Manager data analysis program (Manly, 1993) was used to store and analyze mapping data.

Genomic DNA isolation

Genomic DNA clones were obtained by Southern hybridizations to mouse libraries (BACs, Research Genetics; Huntsville, AL; PACs, Roswell Park Cancer Institute, Buffalo, NY) using informative markers from either mouse chromosome 2 or 14. DNAs were prepared from 100 ml cultures using Qiagen tip-100 column purification according to manufacturer protocols (Qiagen, Palo Alto, CA). 150-200ng of DNA from each individual clone was used for subsequent FISH analysis.

Mouse metaphase chromosome preparation

Mouse metaphase chromosomes were derived from spleen cells of animals heterozygous for the 1Gso translocation. Splenocytes were cultured in media and cells

were dispensed on slides according to protocols previously established (Cacheiro and Russell, 1975).

Fluorescence *in situ* hybridization

Large genomic clones were isolated for FISH experiments to increase the size and intensity of visualized signals during FISH analysis. These DNAs were random-prime labeled using Fluor-12-dUTP, to produce directly fluoresceinated DNA probes according to manufacturer's instructions (Prime-It Fluor Fluorescence labeling kit; Stratagene, La Jolla, CA). Fluorescence *in situ* hybridization was performed with these probes onto previously prepared mouse metaphase chromosome spreads using the Chromosomal *in situ* Hybridization kit with modifications (Stratagene, La Jolla, CA, a detailed protocol to perform the Fluorescence *in situ* hybridization technique is given in the appendix).

Single primer PCR to generate endclones

To generate endclones from BAC clones in an efficient manner, a PCR based method was designed and tested for generating endclones in collaboration with Dr. Joomyeong Kim (Kim *et al.*, 1997). This technique utilized a single primer PCR method (Parker *et al.*, 1991, Parks, *et al.*, 1991, and Wang, *et al.*, 1991) to amplify different ends of the BAC, and was demonstrated to consistently derive true endclones (Kim *et al.*, 1997).

This method employs the M13 reverse and T7 sites present in the BAC vector, as an anchor location. Utilizing a low annealing temperature, PCR conditions allow nonspecific binding of the same primer to occur with insert DNA. If the two binding sites were close enough, and in opposite orientation, logarithmic amplification can occur. The resulting products would contain inserts from the appropriate ends of the BACs.

PAC endclones are generated in a similar manner, with T7 and Sp6 sites already

present in the pPAC4 vector. Sp6 generated the same size band in all the PACs by PCR, suggesting that it was amplifying vector sequence. The pPAC vector sequence (accession #U75992; Frengen *et al.*, 1996) was examined to design an oligonucleotide on the same side as Sp6, designated PAC41PA (5'AATTGACTAGTGGGTAGGCC3').

100 ng of BAC or PAC clone DNA is used as a template for PCR, along with 400 ng of primer DNA. PCR reactions used 0.2 U Taq polymerase (Perkin-Elmer/Cetus) in standard, supplied buffers, and were amplified with the following program in a GeneAmp 9600 thermocycler (Perkin-Elmer/Cetus): 95°C, 4 min.; (94°C, 1 min.; 50°C 1 min.; 72°C, 2 min.) X 45 cycles; 72°C, 5 min.

Yeast Artificial Chromosome experiments

A YAC clone, ICFRy903D0214DD2 (D0214D, Dr. Hans Lehrach, Imperial Cancer Research Fund, England), was isolated and found to contain the mouse *Pax6*, but not the *Wt1* gene. Yeast cells containing the YAC D0214D clone were grown and DNA was isolated into agarose plugs as described (Lai *et al.*, 1994). Partial digests of the YAC plugs, using rare-cutting enzymes, were performed for different time periods. These digested fragments were separated by size using pulsed-field gel electrophoresis (PFGE) and then Southern blotted (Chaplin and Brownstein, 1993). YAC endclones were generated using a modified ligated-linker method (Mueller and Wold, 1989).

Transformation associated recombination (TAR) cloning

In collaboration with Vladimir Larionov (National Institute for Environment Health and Sciences, Research Triangle Park, NC), a construct was made to incorporate a portion of the *Rcn* gene using oligonucleotides Rcn3in2 (5'gttcaaactcctctgggtgc3') and Rcn5in2Sal1 (5'ctcgtcgacgaaggcatcttaggtcaggg3') corresponding to portions of *Rcn* genomic sequence in exons two and three (accession #D43954, Ozawa, 1995). PCR, on a BAC clone 115-

E-3 that contained the *Rcn* gene, isolated a 917 bp fragment that incorporated a *Sal*I restriction enzyme site on one end. The fragment was ligated into a plasmid, using the TA Cloning Kit (Invitrogen, Carlsbad, CA), and transfected according to manufacturer's protocols. An internal *Eco*RV site within the genomic fragment allowed the plasmid to be digested with *Eco*RV and *Sal*I restriction enzymes to isolate a 425 bp band that was directionally ligated into the corresponding sites of the pTAR4-B1 vector. This TAR clone was given to Vladimir Noskov as well as a liver sample of the Sey^{Dey}/1Gso compound mutant animal for use as the target genomic DNA. Two primers, G832 and G833 (Kent *et al.*, 1997), were used for the negative control. A second set of primers from the *Rcn* sequence, ConF (5'CTAACCTAGGGGACCACGCC3') and ConR (5'CGTACTGGAAGCTCTGGTTGTCC'), were derived from the promoter sequence and exon one (accession #D43952, Ozawa, 1995), and used as a positive control for the TAR cloning.

CHAPTER 3.

THE 1GSO PROJECT

SECTION 1.

CONSTRUCTION OF A BACKCROSS MAP OF MMU2 AND 14

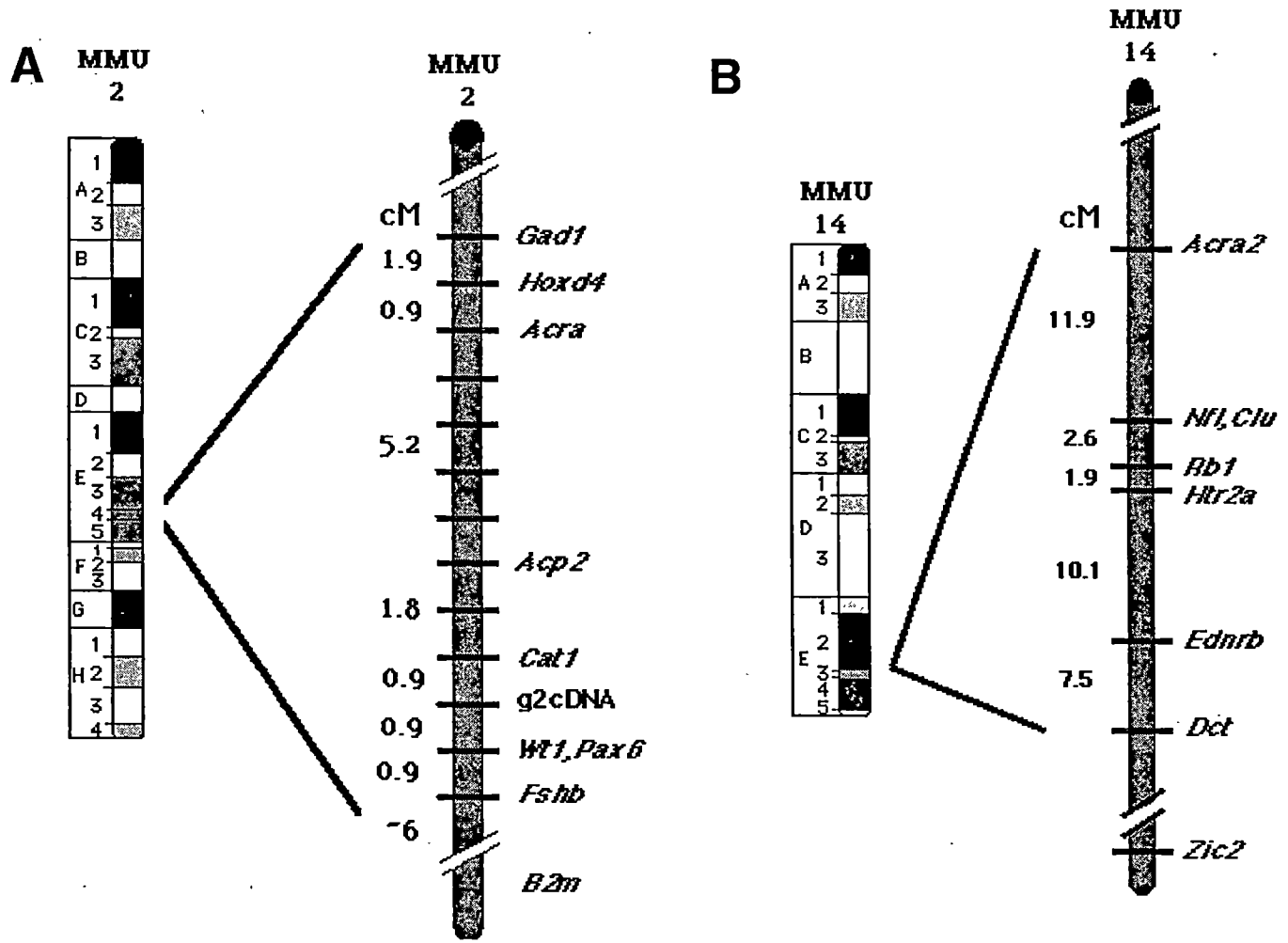
As mentioned in the Introduction, the first stage associated with a positional cloning project is the localization of the mutation on a chromosome. Cytological studies showed that the 1Gso translocation occurred between chromosome 2 (band E4) and chromosome 14 (band E3) (Figure 1) (Rutledge *et al.*, 1986). This cytogenetic mapping localized the breakpoints to individual chromosome regions, and allowed the two regions to be compared with established murine genetic maps to begin to choose markers as potential candidate genes and mapping reagents.

At the outset of the mapping effort, we had no idea whether gene(s) associated with the diver neurological or embryo-lethal phenotypes would be located on chromosomes 2, 14, or both. However, chromosome 2 was chosen for preliminary mapping and breakpoint cloning, because it had, and still has, a greater density of known molecular markers relative to chromosome 14 (Figure 4). Cytogenetic mapping data allowed us to concentrate our marker search to a small interval. This mouse region is related to two human regions (HSA) 11p13 or 2q24-3. Since the homology regions were established, we were able to use the human maps to gain access to numerous additional mapped gene markers that had not previously been located in mouse.

On chromosome 14, the marker density is very low, and related regions of human 13q contained few markers to help us build an informative genetic map in the mouse. However, a few potential candidate genes were mapped using our IB panel to establish linkage and order on Mmu 14 (Figure 4).

Figure 4. Backcross Map of mouse chromosome 2 and 14.

A summary backcross map encompassing the regions of mouse chromosome 2 band E4 and mouse chromosome 14 E3 is shown. To the left of each map is an ideogram of the chromosome, highlighting the regions where the translocation breakpoint occurs. A) A backcross map of mouse chromosome 2 surrounding the region of band E4, encompassing over 26 cM in the mouse genome. B) A backcross map of mouse chromosome 14 surrounding the region of band E3 including 34 cM of chromosome 14.


A
MMU
2
MMU
2
cM
1.9
0.9
5.2
1.8
0.9
0.9
~6
Gad1
Hoxd4
Acra
Acp2
Cat1
g2cDNA
Wt1,Pax6
Fshb
B2m
B
MMU
14
MMU
14
cM
11.9
2.6
1.9
10.1
7.5
Acra2
Nf1,Clu
Rb1
Htr2a
Ednrb
Dct
Zic2

To locate new conserved gene markers with the Mmu2 and Mmu14 breakpoint regions, we used the same *Mus musculus* X *Mus spretus* interspecific backcross that generated the comparative map of human chromosome 19 and related regions of the mouse genome (other than mouse chromosome 7, see Part II) to map markers for mouse chromosome 2 and 14.

A compilation of information regarding probes used in the construction of the backcross maps of Mmu2 and Mmu14 is given in Tables 1 and 2. These probes were used in backcross analysis to determine marker location and order relative to genes in the region (Avner *et al.* 1988, and Copeland *et al.*, 1993). Although some of the markers we tested were already on published maps, we repeated IB mapping analysis of all loci before FISH analysis because of known discrepancies within the published consensus maps. The Stubbs' laboratory has identified several such discrepancies, thus, establishing the precise gene order, even of previously mapped markers, is critical to breakpoint mapping studies. For example, one probe corresponding to the cDNA for *Crabp2* (cellular retinoic acid-binding protein type II) did not map to the region of interest on chromosome 2, as had been published (MacGregor *et al.*, 1992). *Crabp2* mapped to mouse chromosome 3, within a homologous segment in agreement with its human location on chromosome 1q21 (Flagiello *et al.*, 1997). Several other instances of erroneous gene mapping were encountered in the course of our overall comparative mapping studies, and for this reason no markers were used for FISH analysis before map position was verified through IB mapping.

Table one. Clones and probes used to construct a backcross map on mouse chromosome 2.

Cloned or PCR-generated probes that were used to map the loci on mouse chromosome 2 are described.

For cloned probes, the type of clone used and its species of origin are listed (H=human; M=mouse; G=genomic, C=cDNA). The sizes of variant fragments used to follow the segregation of each marker in *Mus musculus* (M) or *Mus spretus* (S) DNA, and the restriction enzymes used to generate those fragments are also shown. If the probe was used in FISH analysis, the type of genomic clone is indicated. References describing the probes used and source are listed at right (ATCC= American Type Culture Collection, Rockville, MD). At right the primary references for gene sequences from which oligonucleotide primers were designed are listed. Southern blotting, probe preparation, and hybridizations were carried out as previously described (Stubbs *et al.*, 1990).

Mouse Chromosome 2 probes

Locus	Species/ DNA type	Enzyme	Size kb M	S	Origin	FISH clone	Reference
Acp2	M/C	*Taq*I	2	1.8	ATCC #59900	cosmid	Pohlmann et al., 1988
Acra	H/C	*Eco*RI	13	6.7	ATCC #63103	cosmid	Heidmann et al., 1986
B2m	M/G	*Taq*I	7.5	6	Dr. Stubbs	cosmid	Stubbs et al., 1990
Cat1	H/C	*Taq*I	1.3	1.4	ATCC #57354	cosmid	Quan et al., 1986
Cd59	H/C				ATCC #65965	- - - - -	Philbrick et al., 1990
Cf2	M/C	*Taq*I	1.9	3.5	Dr. Degen	cosmid	Friezner-Degen et al., 1992
Cyct	M/C	*Taq*I	5.4	3.5	Dr. Hecht	- - - - -	Hake et al., 1990
D2H11s149	M/G				GDB#197530	- - - - -	Rochelle et al., 1992
Evx2	M/G				Dr. Gruss	- - - - -	Bastian et al., 1992
f3	H/G				Dr. Glaser	cosmid	Glaser et al., 1989
f8	H/G				Dr. Glaser	- - - - -	Glaser et al., 1989
g2	M/C	*Bam*HI	9.4	11	Dr. Glaser	cosmid	Glaser et al., 1989
Hoxd4	M/G	*Eco*RI	2.8	2	ATCC#63103	BAC/cosmid	Featherstone et al., 1988
Kcna4	M/C	*Taq*I	5.4	7	Dr. Tamkun	- - - - -	Klock et al., 199
Pax6	M/C	*Taq*I	3.2	1.9	ATCC#63180	YAC/BAC/cosmid	Ton et al., 1991
Rag1	H/C	*Taq*I	1.3	3.4	Dr. Baltimore	- - - - -	Schatz et al., 1989
Rag2	H/C	*Taq*I	2.4	4.4	Dr. Baltimore	- - - - -	Oettinger et al., 1992
Rcn	M/G				GDB#D13003	YAC/BAC/PAC	Kent et al., 1997
Scn2a	H/G	*Pvu*II	13	9.4	ATCC #59742	- - - - -	Litt et al., 1989
Spi1	M/G	*Bam*HI	10	17	GDB#X52056	cosmid	Smith et al., 1993
Vim	M/C				ATCC #63023	- - - - -	Wood et al., 1989
Wt1	M/C	*Eco*RV	7	9.8		YAC/BAC/cosmid	

Table two. Clones and probes used to construct a backcross map on mouse chromosome 14.

Cloned or PCR-generated probes that were used to map the loci on mouse chromosome 14 are described.

For cloned probes, the type of clone used and its species of origin are listed (H=human; M=mouse; G=genomic, C=cDNA). The sizes of variant fragments used to follow the segregation of each marker in *Mus musculus* (M) or *Mus spretus* (S) DNA, and the restriction enzymes used to generate those fragments, are also shown. If the probe was used in FISH analysis, the type of genomic clone is indicated. References describing the probes used and source are listed at right (ATCC= American Type Culture Collection, Rockville, MD). At right, the primary references for gene sequences from which oligonucleotide primers were designed are listed. Southern blotting, probe preparation, and hybridizations were carried out as previously described (Stubbs *et al.*, 1990).

Mouse Chromosome 14 probes

Locus	Species/ DNA type	Enzyme	Size kb M	S	Origin	FISH clone	Reference
Clu	H/C	*Bam*HI	16	4.3	EST#1091828	BAC	Adams et al., 1991
Rb1	H/C	*Eco*RI	23	10	EST#876598	BAC	Huang et al., 1988
NEFL	H/C	*Hinc*II	8	7	ATCC #57464	BAC	Myers et al., 1987
ESD	H/C	*Eco*RV	9.4	3.8	ATCC #57608	BAC	Squire et al., 1986
Dct	M/C	*Bgl*II	8.5	4.1	EST#418154	BAC	Marra et al., 1996
Htr2A	H/C	*Bgl*II	20	10	EST#40580	-----	Hillier et al., 1995

SECTION 2.

USING FLUORESCENCE *IN SITU* HYBRIDIZATION TO LOCALIZE THE TRANSLOCATION TO A SMALL INTERVAL ON MOUSE CHROMOSOME 2.

Use of FISH to narrow down the location of the translocation breakpoint site

Once a breakpoint site is mapped by G-banding methods, alignment of cytogenetic data with DNA marker maps permits selection of probes distributed throughout the breakpoint region.

The second stage associated with a positional cloning project involves identifying a small, defined "critical" region associated with the mutation that can be approached through physical mapping techniques. To refine the position of the chromosomal breakpoints and to correlate cytogenetic position with mapped DNA markers we applied a fluorescent *in situ* hybridization (FISH)-based strategy using markers from the Mmu2 breakpoint region. Briefly, FISH is a method that allows visualization of marker loci on metaphase chromosomes (Lawrence, 1990; and Trask, 1991). DNA is labeled with a fluorescent "tag" and subsequently denatured and hybridized to chromosome preparations from cells arrested at metaphase. The fluoresceinated DNA probes will anneal to their homologous sequence on the chromosome, and are visualized by microscopy with a fluorescent light source. Translocation breakpoints can be mapped on the molecular level by using DNA probes that flank the breakpoint (Arai *et. al.*, 1993). Similar approaches have also been used successfully in a number of human gene cloning experiments (e.g. Mercer *et al.*, 1993).

Establishment of the FISH methods that are now routinely used in the laboratory was an important aim of this positional cloning project. The FISH technique was initially

learned with the help of Dr. Ed Michaud, and facility with this technique was further developed while a student in a Cold Spring Harbor Laboratory summer course.

FISH strategy to map translocation breakpoints

To ensure that the breakpoint would fall between two selected markers, probes were chosen from either side of band E4 on chromosome 2. Sequential application of the FISH mapping strategy eventually permitted the narrowing of the breakpoint into a small physical interval (probes used and genomic clones isolated for FISH analysis are summarized in table 3). The FISH strategy is outlined in Figure 5, and illustrated with data derived from the mapping of the 1Gso Mmu2 breakpoint. For example, in mapping 1Gso, markers *Hoxd4* and *B2m* were first used to flank the Mmu2E4 region (Fig 5). Since these probes are both derived from Mmu2, FISH analysis of normal metaphase spreads should show the probes *Hoxd4* and *B2m* to be linked together on one pair of chromosomes (Figure 5, Case A). However, if the translocation is located between these markers, 1Gso mice will carry copies of *Hoxd4* and *B2m* on separate translocation chromosomes (Figure 5). If both probes are located distal (Figure 5, Case B) or proximal (Figure 5, Case C), to the breakpoint, the markers will remain linked together on a single chromosome pair. This simple, visual test allows for the localization of the translocation between any two genetically mapped markers. FISH analysis was performed on metaphase spreads of spleen cells taken from animals heterozygous for the 1Gso translocation, so that the normal chromosome 2 carried by these mice provided an internal control for FISH experiments.

TABLE 3. Probes and genomic clones used for FISH analysis.

Probes and genomic clones that were used in FISH analysis of the 1Gso translocation breakpoint are summarized.

The chromosome 2 probes and FISH clones are illustrated on the left of the table, while chromosome 14 probes and FISH clones are illustrated to the right of the table. A list of genomic YAC, BAC, PAC, and cosmid clones is given to the right of each probe, with the clone designation in parentheses.

Mouse Chromosome 2 FISH clones		Mouse Chromosome 14 FISH clones	
Locus	FISH clone	Locus	FISH clone
Acp2	cosmid (cAcp2-1)	*Clu*	BAC (86-N-11)
Acra	cosmid (cAcra-1)	*Dct*	BAC (267-K-9)
B2m	cosmid (cB2m-1)	*Rb1*	BAC (51-M-20)
Cat1	cosmid (cCat1-1)		
Cf2	cosmid (cCf2-1)		
f3	cosmid (cf3-1)		
g2	cosmid (cg2-1)		
Hoxd4	cosmid (cHoxd-1)		
Pax6	YAC (D0214D)		
	BAC (279-O-16)		
	BAC (202-A-6)		
	BAC (311-O-6)		
	PAC (437-G-5)		
	Cosmid (cPax6-1)		
Rcn	YAC (D0214D)		
	BAC (596-D-16)		
	BAC (541-C-3)		
	BAC (287-B-20)		
	BAC (44-N-23)		
	PAC (460-J-8)		
Spi1	cosmid (c-Sfpi1-1)		
Wt1	BAC (587-C-27)		
	Cosmid (cWt1-1)		
DO214D#4	BAC (224-L-6)		
44-N-23T7	PAC (54-J-2)		
311-O-6	PAC (54-J-2)		

Figure 5. FISH analysis of translocation mutations.

In normal mice, linked probes on mouse chromosome 2 will be visualized as two sets of markers linked together as seen on the top left of the figure.

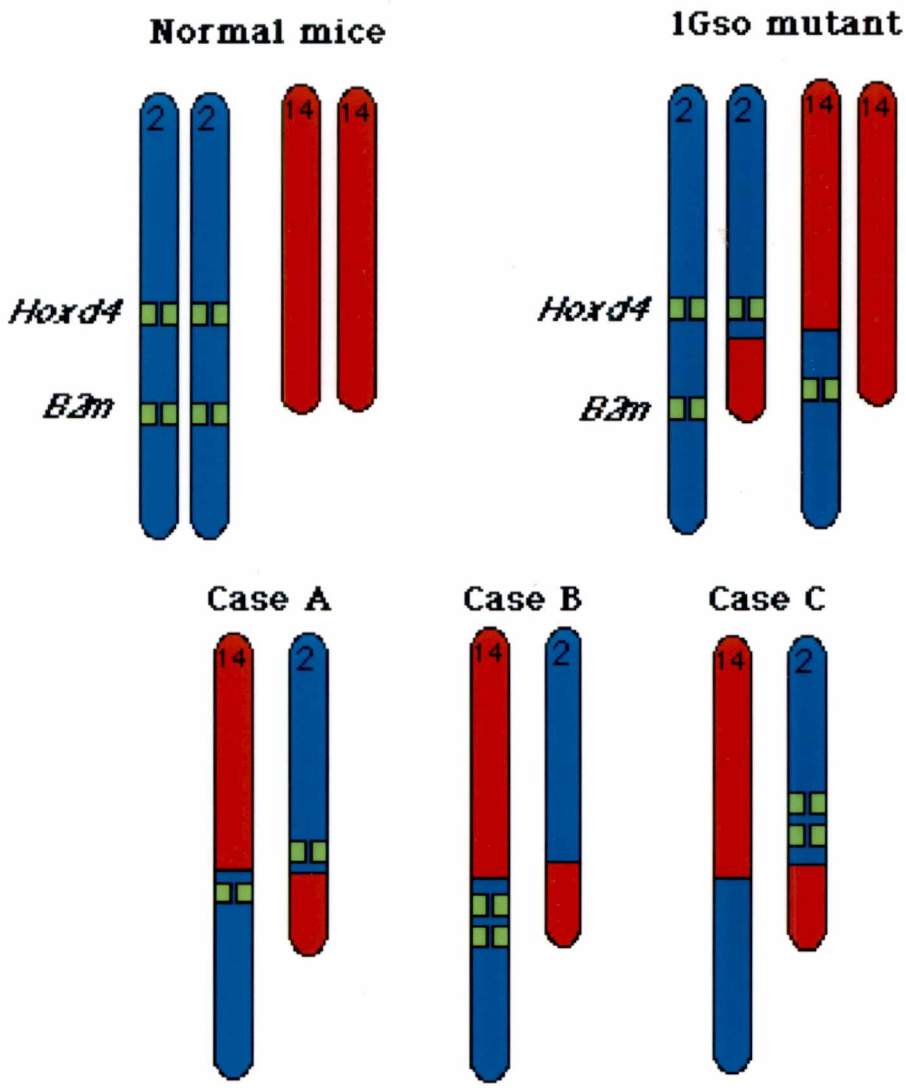
In translocation heterozygotes, markers flanking the translocation will be seen as syntenic on the normal chromosome 2 and will be juxtaposed on the translocation containing chromosomes, if the translocation occurs between the two markers, as seen in the top right of the figure.

Case A: Flanking linked probes are separated by the translocation.

Case B: Distal linked probes remain syntenic.

Case C: Proximal linked probes remain syntenic also.

Using FISH to Map Translocations



FISH mapping of the translocation breakpoint

Preliminary applications of FISH verified the predicted map positions of *Hoxd4* and *B2m*, by showing that the cosmids containing those genes flanked the 1Gso breakpoint. *Hoxd4* and *B2m* were known to be ~16 cM apart on chromosome 2, but were seen to be on different translocated chromosomes when used on 1Gso/+ metaphase spreads (data not shown). The next step was to select another probe detecting a locus positioned between the two previously mapped markers. A *Pax6* (paired box-6) probe was used to isolate a cosmid clone for FISH analysis. Two possible combinations were utilized with the previous markers; *Hoxd4* and *Pax6*, and *Pax6* and *B2m*. FISH results showed that the translocation breakpoint is located between *Hoxd4* and *Pax6* (Figure 6). The other set of cosmids, *Pax6* and *B2m*, remain linked in the translocation mutant mice. Further FISH analysis with internal markers refined the map of the breakpoint region. *Acp2* (acid phosphatase 2) was used to divide the interval between *Hoxd4* and *Pax6* into two, approximately equal-sized regions. A cosmid clone was isolated for *Acp2* and used in FISH with either the *Hoxd4* or the *Pax6* cosmid. The results (Figure 7a) indicated that the breakpoint resides in an ~ 3.6 cM interval between *Acp2* and *Pax6*. *Cat1* (catalase1) was another marker used to further delineate the remaining 3.6 cM region. FISH techniques were used with a *Cat1* cosmid clone with either the *Hoxd4* or the *Pax6* cosmid. The region was then shown to be in the ~1.8 cM region between *Cat1* and *Pax6* (Figure 7b). *Cat1* is one of the markers used to define homology between Mmu2 and a region on human chromosome 11p13, containing genes responsible for the WAGR syndrome (Wilms tumor, aniridia, genitourinary malformations, and mental retardation (Gessler *et al.*, 1989). The well-defined homologous region is within the interval bracketed by the breakpoint-spanning probes, indicating that the neurological defect in 1Gso mice might possibly be analogous to the mental retardation seen in the WAGR syndrome patients.

Figure 6. *Hoxd4* and *Pax6* FISH analysis.

On the left is an ideogram of Mmu2 showing the interval between *Hoxd4* and *Pax6*. On right is the results of FISH with cosmid clones containing *Hoxd4* and *Pax6* hybridized to heterozygous 1Gso metaphase cells. Fluorescent signals are linked on the normal chromosome 2 and separated in the translocated chromosomes, as indicated by arrows. These data anchored the translocation breakpoint relative to the genetic map on chromosome 2 and served as a starting point for subsequent mapping and FISH experiments.

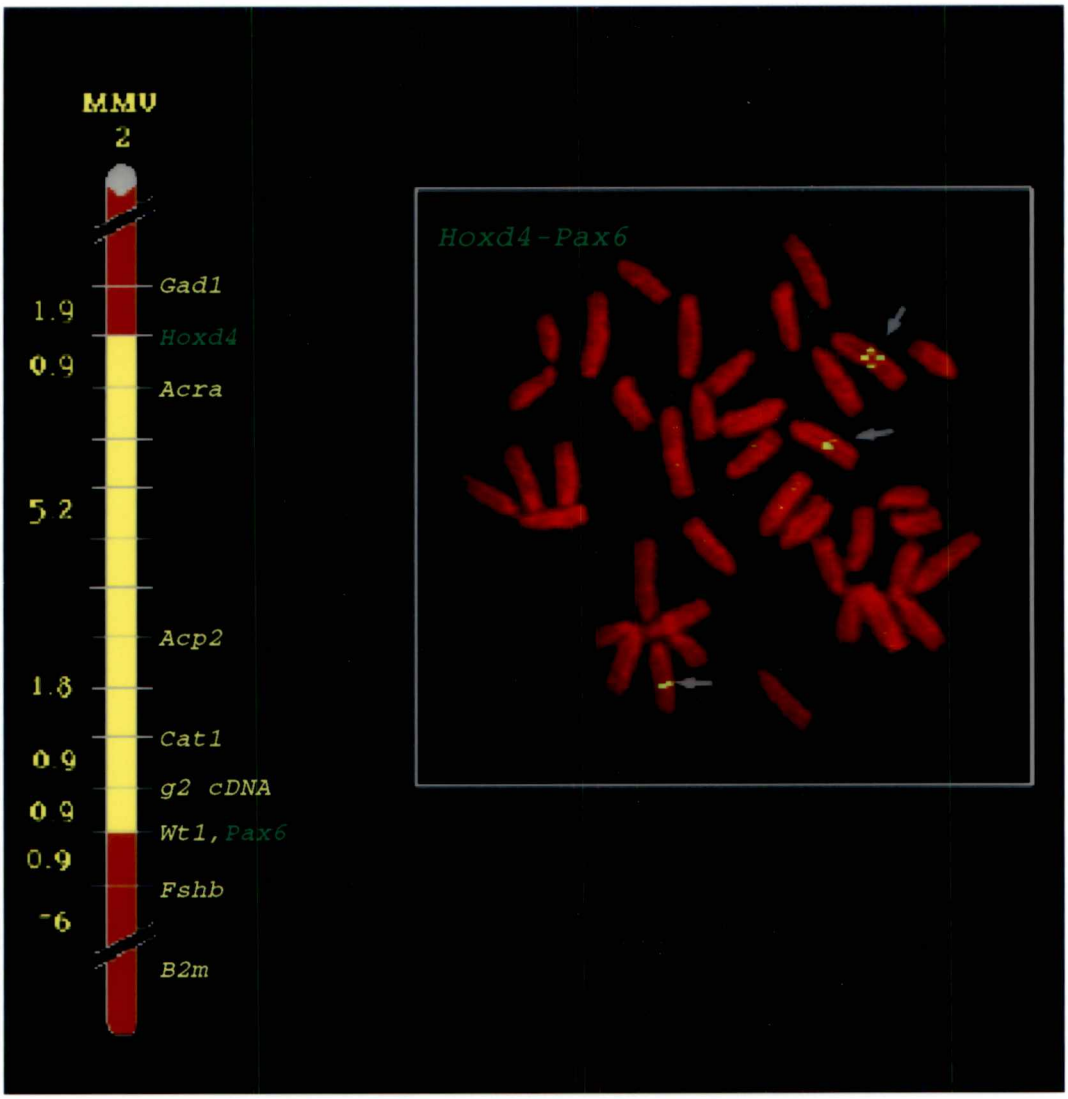

MMU
2
Gad1
1.9
Hoxd4
0.9
Acra
5.2
Acp2
1.8
Cat1
0.9
g2 cDNA
0.9
Wt1, Pax6
0.9
Fshb
6
B2m
Hoxd4-Pax6

Figure 7. Summary of FISH steps.

This figure shows a summary of successive FISH experiments undertaken to refine the mapping of the translocation breakpoint of 1Gso on mouse chromosome 2. A backcross map of chromosome 2 is depicted in each figure, with probes used during each FISH step highlighted in blue.

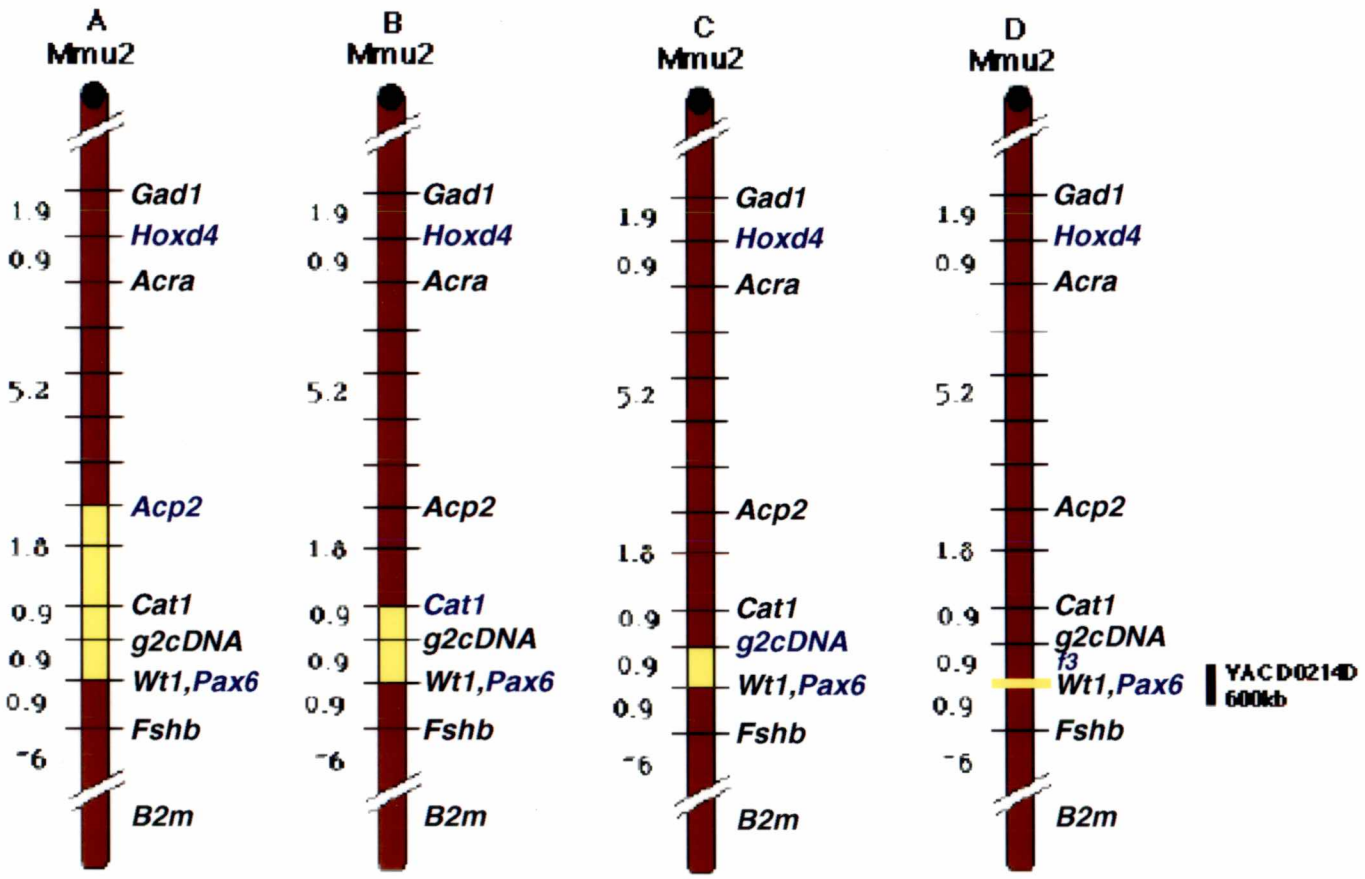
A) FISH between the markers *Hoxd4* and *Acp2*, and *Acp2* and *Pax6*. The yellow indicates the localization of the translocation breakpoint between *Acp2* and *Pax6*.

B) FISH between the markers *Hoxd4* and *Cat1*, and *Cat1* and *Pax6*. The yellow indicates the localization of the translocation breakpoint between *Cat1* and *Pax6*.

C) FISH between the markers *Hoxd4* and *g2*, and *g2* and *Pax6*. The yellow indicates the localization of the translocation breakpoint between *g2* and *Pax6*.

D) FISH between the markers *Hoxd4* and *f3*, and*f3* and *Pax6*. The yellow indicates the localization of the translocation breakpoint between *f3* and *Pax6*.

This left a small region containing *Wt1* as the only other known probe to use to narrow down the region further.


A
Mmu2
B
Mmu2
C
Mmu2
D
Mmu2
Gad1
Hoxd4
Acra
Acp2
Cat1
g2cDNA
Wt1,Pax6
Fshb
B2m
1.9
0.9
5.2
1.8
0.9
0.9
0.9
~6
f3
YAC D0214D
600kb

Assignment of the 1Gso breakpoint within this well studied region also provided access to new markers for comparative mapping purposes, enabling further delineation of the breakpoint interval (Figure 8; Glaser *et al.*, 1990).

A number of conserved markers had already been mapped to the human 11p13 region, and a physical map had already been generated (Rose *et al.*, 1990). Linkage of the markers in this region was known to be conserved in mouse (Glaser *et al.*, 1990). Two markers, *g2* and *f3*, generously donated by Dr. Tom Glaser, allowed for further refinement of this region through physical map construction and FISH analysis.

A cosmid clone containing *g2* was used in FISH experiments in combination with cosmids for both *Hoxd4* and *Pax6*, respectively (Figure 7c). Another series of FISH experiments used an *f3* cosmid as a probe in conjunction with *Hoxd4* and *Pax6* cosmids. The FISH results placed the breakpoint between *f3* and *Pax6*, with *Wt1* as the only other marker available for further analysis (Figure 7d). FISH analysis using a *Wt1* cosmid in combination with *Hoxd4* and *Pax6* clones indicated that the translocation was between *Wt1* and *Pax6* (Figure 9). These two markers are closely linked, with only an ~700 kb interval between them in humans and mice (Rose *et al.*, 1990; Glaser *et al.*, 1990). On the backcross panel these two markers cosegregated, suggesting that in the mouse they were tightly linked as well. This distance was small enough to be amenable to positional cloning techniques, fulfilling the requirements for the second stage of a positional cloning strategy. The third stage of a positional cloning project was then undertaken. A physical map and a contig of overlapping genomic DNA clones for the region were assembled, with the ultimate goal of isolating the translocation breakpoint and identifying candidate genes.

Figure 8. Mouse and human homology surrounding the WAGR region.

The mouse chromosome 2 and homologous 11p13 region in humans corresponding to the WAGR region allowed for the identification of new markers. Probes known to map in both species are indicated. The vertical lines are spaced 200 kb apart to represent physical distances. This illustrates the linkage synteny that is preserved in this region between the two species.

Homology between the mouse and the human WAGR region

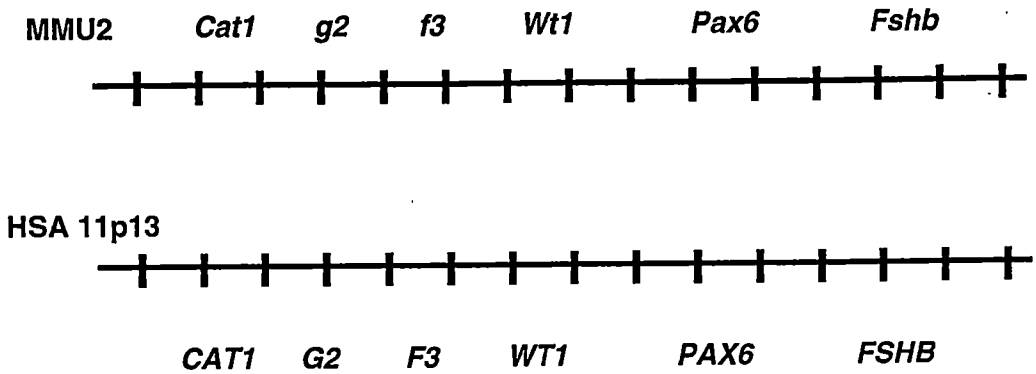
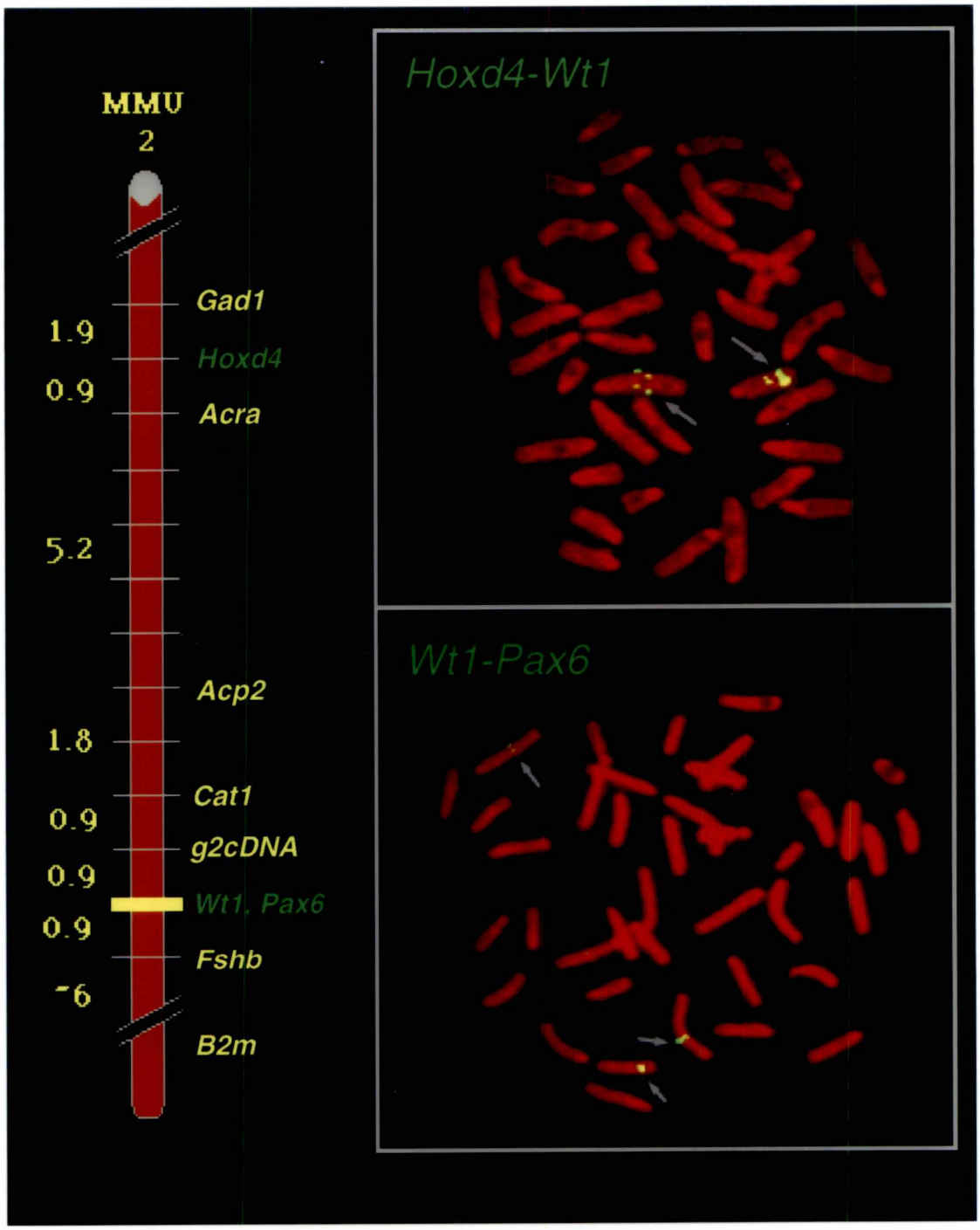


Figure 9. FISH analysis between *Hoxd4-Wt1* and *Wt1-Pax6*.

On the left is an ideogram of Mmu2 showing the interval between *Hoxd4* and *Pax6*. On the right are results of FISH with cosmid clones containing *Hoxd4* and *Wt1* (Top box) or *Wt1* and *Pax6* (bottom box) hybridized to heterozygous 1Gso metaphase cells. The interval between *Hoxd4-Wt1* does not contain the breakpoint, as seen in the top FISH figure, which shows the probes remaining linked as indicated by arrows. In the bottom FISH figure, the position of the breakpoint was narrowed to between *Wt1* and *Pax6* , as indicated by arrows, a region less than 1 cM in length on chromosome 2.


MMU
2
Gad1
1.9
Hoxd4
0.9
Acra
5.2
Acp2
1.8
Cat1
0.9
g2cDNA
0.9
Wt1, Pax6
0.9
Fshb
~6
B2m
Hoxd4-Wt1
Wt1-Pax6

SECTION 3.

PHYSICAL LOCALIZATON OF THE BREAKPOINT ON MMU2

In the final stages of the positional cloning strategy, we initiated a chromosomal walk (Ballabio, 1992; Collins, 1992; reviewed by Stubbs, 1992) to narrow the breakpoint down from the 700 kb interval between *Wt1* and *Pax6*, to a smaller region contained within a single BAC or PAC clone. Screening a YAC library with both *Pax6* and *Wt1* probes initially isolated a YAC (DO214D) which contained *Pax6*.

To assay whether D0214D contained the breakpoint, the entire YAC was used as a probe for FISH. Hybridization of the labeled YAC to 1Gso/+ metaphase spreads showed a clear "split" signal, verifying that the YAC contained the 1Gso Mmu2 translocation breakpoint (Figure 10).

The YAC was end-cloned in the laboratory using established protocols (Mueller and Wold, 1989). The YAC is 600 kb in length according to pulsed-field gel electrophoresis (data not shown). Endclone #4 was used as a marker for the left end of the YAC and endclone #2 was used as a marker for the right end (Figure 11). A rare-cutting restriction enzyme map was generated for YAC D0214D (Figure 11). Also, the endclones from the YAC were mapped to the backcross panel on mouse chromosome 2, and were shown to cosegregate with both *Wt1* and *Pax6*. This showed that the YAC was not chimeric. This YAC has provided probes for further isolation of clones within this interval, and as a control mapping reagent to verify overlap and position of other genomic clones isolated within this region (see below).

Figure 10. YAC D0214D FISH analysis.

On the left is an ideogram of Mmu2 showing the interval between *Hoxd4* and *Pax6*. On right shows FISH analysis of YAC D0214D hybridized to heterozygous 1Gso metaphase cells. The YAC is present in a single copy on the normal mouse chromosome 2, but is split into two signals across the translocation breakpoint on two different chromosomes. The split signals occur in the genomic location expected in the translocation, visually verifying its authenticity.

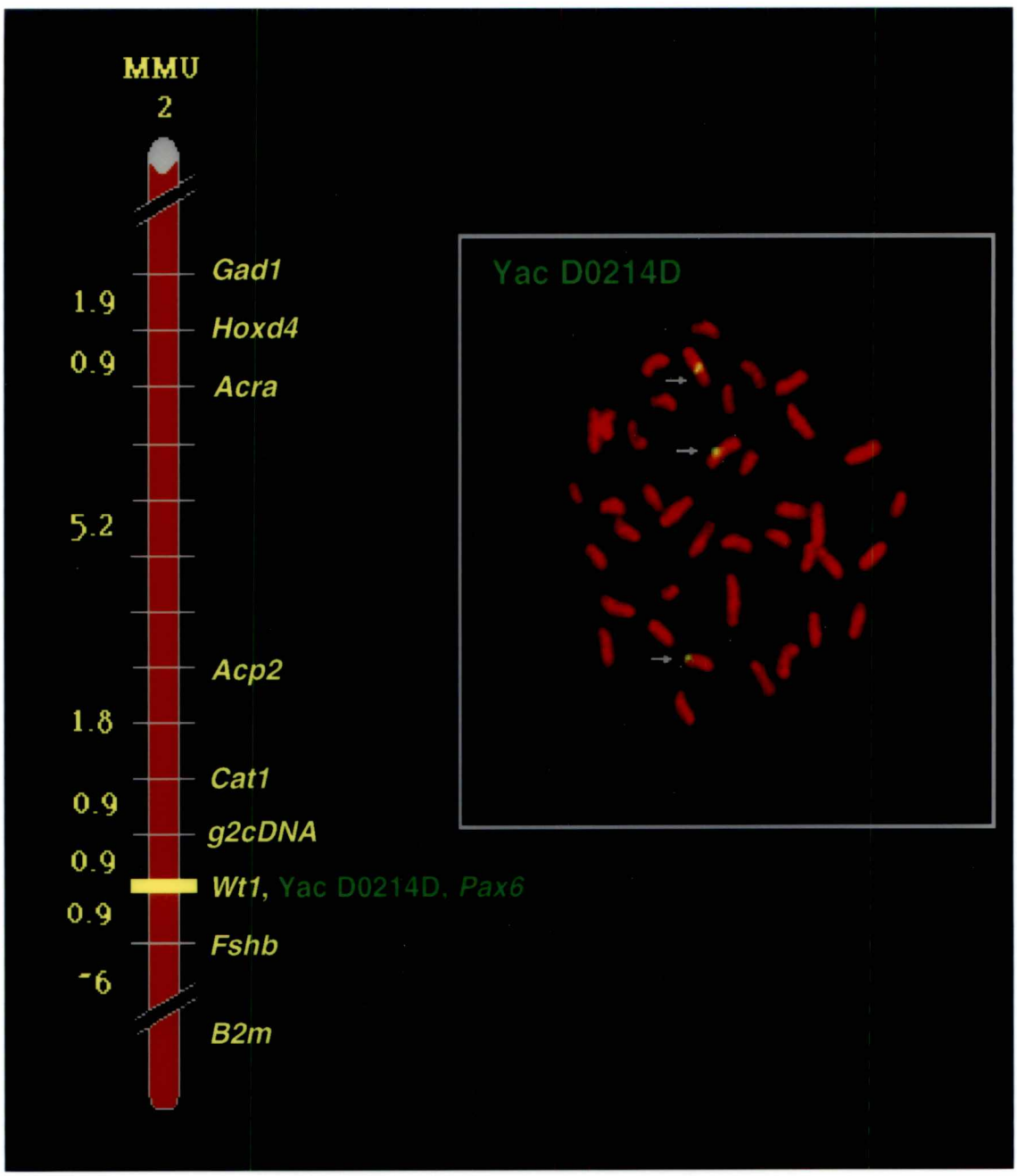
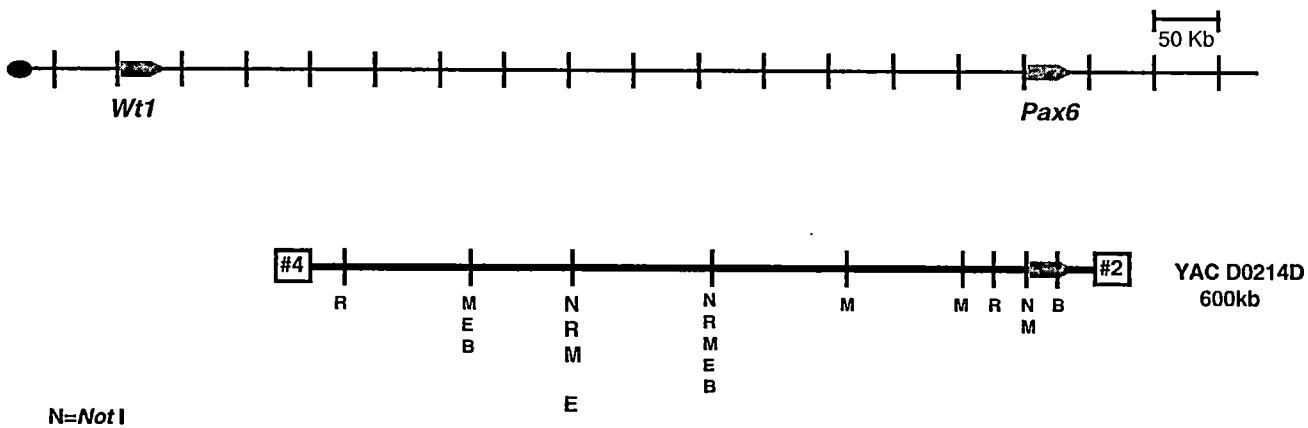

MMU
2
Gad1
1.9
Hoxd4
0.9
Acra
5.2
Acp2
1.8
Cat1
0.9
g2cDNA
0.9
Wt1, Yac D0214D, Pax6
0.9
Fshb
6
B2m
Yac D0214D

Figure 11. Restriction map of YAC D0214D.

A restriction map of YAC D0214D was generated. The orientation of the YAC in relationship to mouse chromosome 2 is indicated in the upper measure. The left of the measure is toward the centromere, while the right is toward the telomere of the chromosome. Distances are indicated with vertical lines spaced every 50 kb. The genes for *Wt1* and *Pax6*, are also illustrated by boxes. The YAC is approximately 600 kb in length. Vertical lines show the restriction sites, and approximate distances on the YAC, with the corresponding restriction enzyme site noted underneath. The YAC contains *Pax6*, as seen on the distal end to the YAC.



N=*Not* I
R=*Nru* I
M=*Mlu* I
E=*Eag* I
B=*BssH*

Isolation of smaller clones to further narrow the breakpoint region

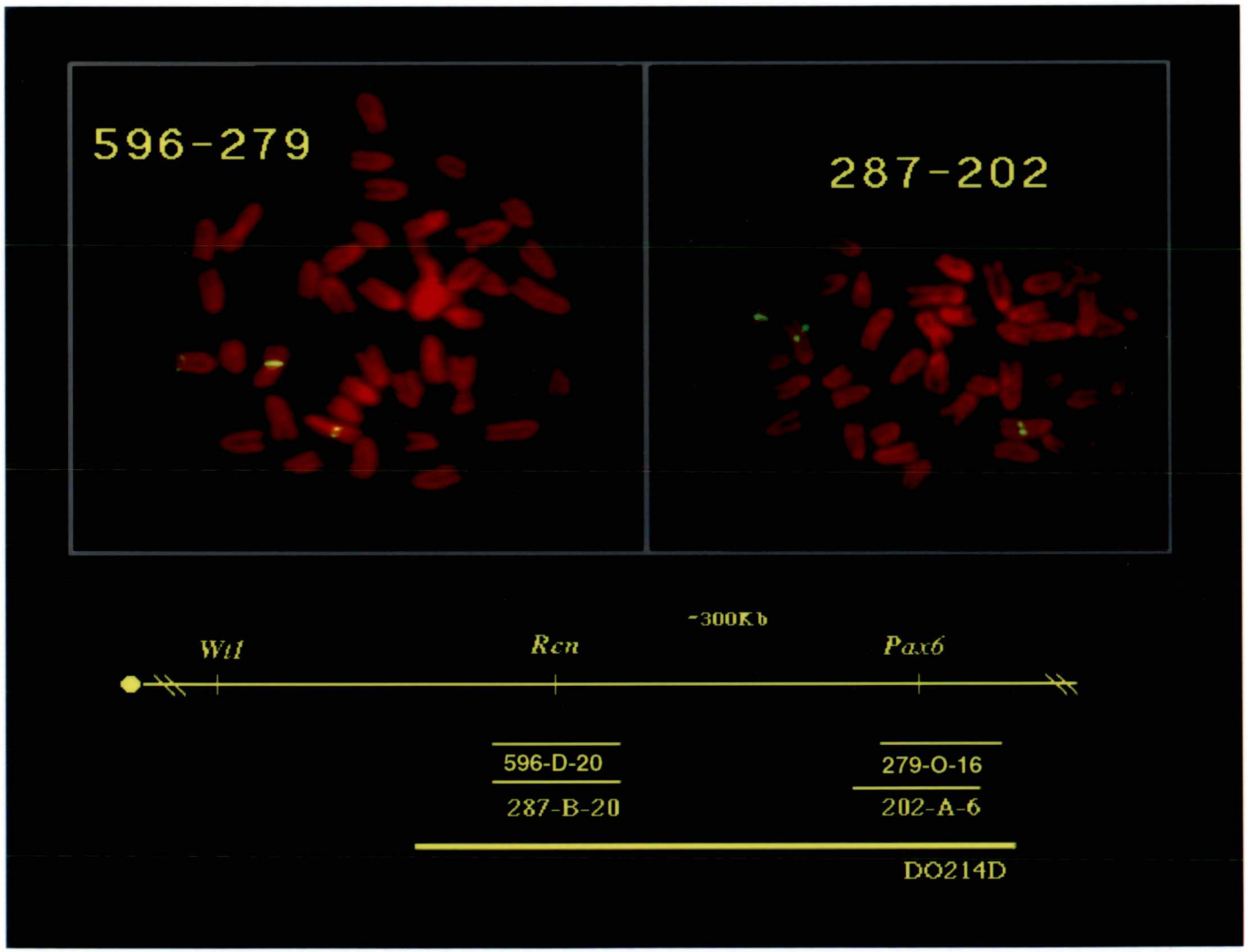
Once the YAC that spanned the breakpoint was discovered, internal markers and endclones were used to screen a BAC library (Research Genetics, and LLNL) to isolate BACs, and further narrow down the breakpoint-containing region. One endclone, endclone #2, was particularly informative, because it hybridized to BACs previously isolated and containing *Pax6*. The other end of the YAC, endclone #4, was also used to isolate clones. A BAC clone, BAC 224-L-6, was isolated with the YAC endclone #4, and was used as a FISH probe on 1Gso/+ metaphase spreads, but it did not span the breakpoint. However, BAC 224-L-6 was used as a FISH probe in conjunction with BACs for *Pax6* and *Hoxd4*, respectively (data not shown). The results confirmed that the 1Gso breakpoint was contained in the region spanned by the YAC, as was seen in earlier YAC FISH experiments. The breakpoint was not between *Hoxd4* and YAC-end #4, but was included in the YAC-end #4 to *Pax6* interval. These experiments narrowed down the region containing the breakpoint to less than 600 kb, and oriented the YAC (Figure 11).

Before the walking steps were initiated, a mouse gene, Reticulocalbin (*Rcn*) was mapped to the human interval between *WT1* and *PAX6*, and to the Small eye Harwell deletion (Sey^H) (Kent *et al.*, 1997). The deletion of *Rcn* in this mutant localized the gene to the corresponding homology region on mouse chromosome 2.

The *Rcn* marker was very informative for further mapping of the 1Gso translocation breakpoint. The interval between *Pax6* and *Wt1* was effectively divided into two equal lengths. When BACs isolated with the *Rcn* cDNA probe were used in FISH analysis, the breakpoint was seen to map between *Rcn* and *Pax6* (Figure 12), a distance of only ~ 300 kb. *Rcn* BAC endclones were used to continue the chromosomal walking.

Figure 12. FISH analysis of *Rcn* and *Pax6*.

FISH analysis of 1Gso using internal markers for *Rcn* and *Pax6*. The horizontal line illustrates the region between *Wt1*, *Rcn* and *Pax6*, with gene order indicated by vertical lines. The small horizontal lines represent BAC clones, with BAC designations underneath. A YAC clone is shown by the larger, thicker horizontal line with the YAC name underneath. Two separate BACs were used for each marker to demonstrate that the translocation occurs between *Rcn* and *Pax6*. The left FISH figure shows BAC 596-D-20 and BAC 279-O-16 flank the translocation breakpoint. The right FISH figure illustrates BAC 287-B-20 and BAC 202-A-6 also flank the translocation breakpoint. These BACs were also used independently to verify that they did not harbor the breakpoint.


596-279
287-202
~300Kb
Wt1
Rcn
Pax6
596-D-20
287-B-20
279-O-16
202-A-6
DO214D

By comparing rare-cutting restriction site maps of the BAC clones with the known restriciton map of the region (Figure 11), the orientation of the new clones was determined and verified.

Endclones for the new BACs were made and used to rescreen the BAC library. The ends of the new BACs that would extend the walk into the uncloned, breakpoint-containing region were isolated and used to screen the BAC libraries again.

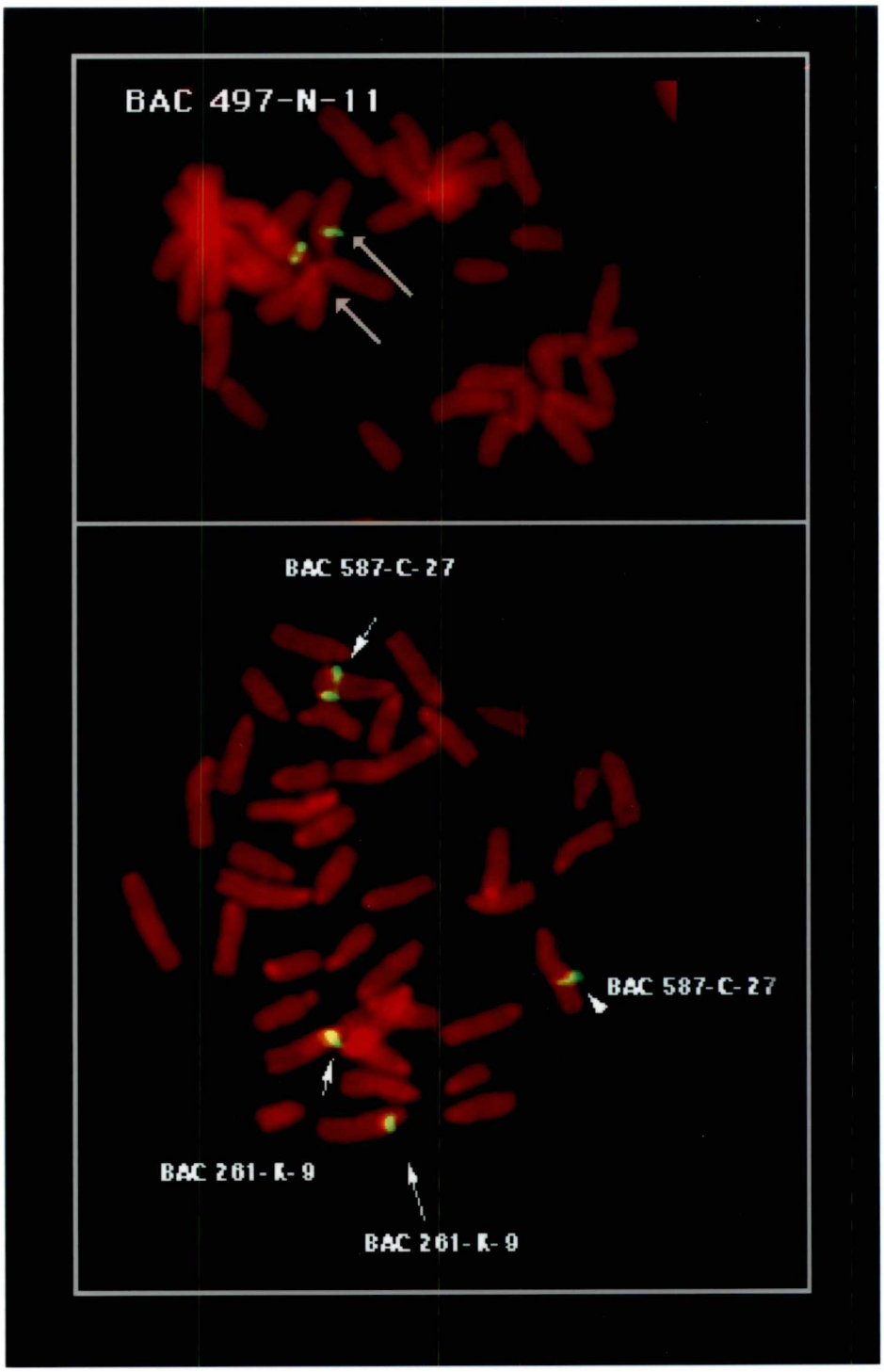
Chromosome walking steps are not always as straightforward as they may seem. BAC 261-K-9, which contained the *Rcn* gene, was of particular interest. One PCR generated endclone 261-K-9-M13R was anchored by hybridization to the existing contig and the other endclone 261-K-9-T7 did not map into the established BAC contig. This made endclone 261-K-9-T7 the best candidate to fill the remaining gap. New BACs were isolated with endclone 261-K-9-T7. The newly isolated BAC clones did not seem to cover the gap, and no other probe in the contig could be localized to these BACs.

To determine the location of BAC 261-K-9 relative to other markers on mouse chromosome 2, BAC 261-K-9 was used in FISH with BAC 587-C-27, which has previously been mapped to mouse chromosome 2 (Figure 13). Also, BAC 497-N-11, one of the new BACs that was isolated with endclone 261-K-9-T7 was used in FISH experiments alone (Figure 13).

These results indicated that these clones were not located on mouse chromosome 2. Therefore, BAC 261-K-9 was chimeric, and its endclone isolated BACs which were FISH mapped onto another chromosome. BAC 261-K-9 was the only BAC clone we encountered that was rearranged in this manner, and illustrates that during walking steps different levels of clone verification must be used to insure that subsequent research is indeed progressing in a correct orientation and manner. After multiple screening of the BAC library with a variety of different endclones, no new BACs could be found, suggesting that the breakpoint region is not represented in this BAC library.

Figure 13. Chromosome walk onto wrong chromosome.

The top figure shows the localization of BAC 497-N-11, isolated by the 261-K-9-T7 endclone, to the centromeric region of an unknown chromosome, indicating that this BAC is not in the central region of mouse chromosome 2 where the breakpoint occurs. The bottom FISH figure shows BAC 261-K-9 with BAC 587-C-27 (*Wt1* positive clone). FISH analysis indicated that BAC 261-K-9 is located on a chromosome other than chromosome 2, in the centromeric region, similar to BAC 497-N-11.


BAC 497-N-11
BAC 587-C-27
BAC 587-C-27
BAC 261-K-9
BAC 261-K-9

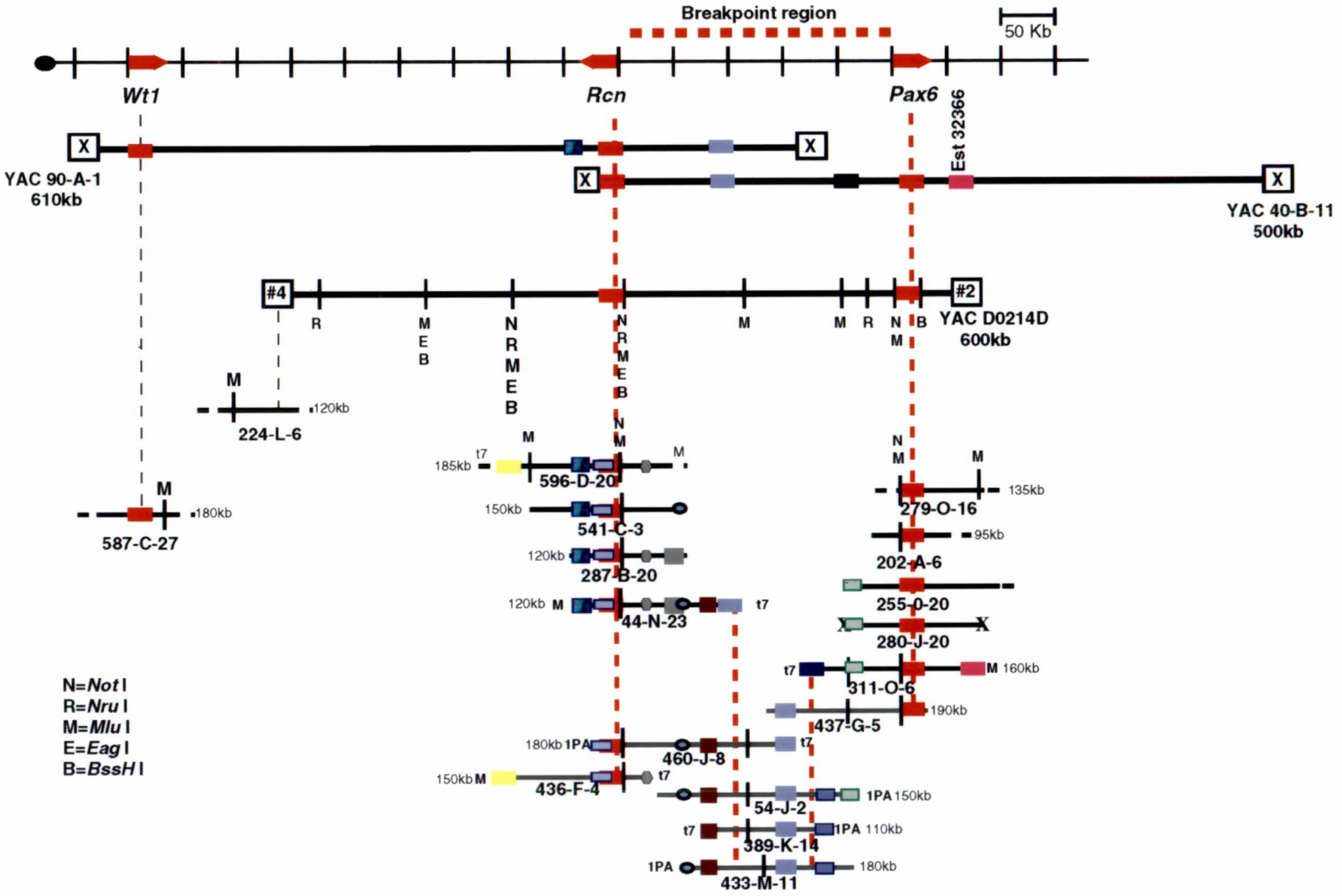
To attempt to close the gap in the contig, the laboratory acquired a new, commercially available large-insert mouse genomic library. The PAC (P1-derived artificial chromosome) library contains inserts that are as large or larger than the BAC library, but were cloned into a different vector with a different restriction enzyme than the BAC library. A map of this region shows the contig generated, as well as the endclone markers used for gap closure (Figure 14). The library was hybridized using a pair of endclones (BAC 44-N-23-T7 end, and BAC 311-O-16-T7 end). PAC clones were isolated, and by subsequent hybridizations, three PAC clones (54-J-2, 389-K-14, and 433-M-11) were discovered to contain both endprobes. The three PAC clones seemed to overlap with both flanking BAC clones, and gave the first evidence of gap closure, suggesting that the region between *Rcn* and *Pax6* has been completely cloned. The generation of endclones from these PACs and subsequent hybridization of these endclones to the existing BAC contig verified this overlap. Endclones generated by the PAC41PA primer verified that the contig has been completed, and therefore the region containing the breakpoint has been cloned in wild type DNA (Figure 14).

Isolation of breakpoint containing clones

The new PAC clones spanning the BAC contig gap have been used in FISH analysis. Similar to the YAC clone D0214D, any individual clone that spans the breakpoint should be visualized as a spilt signal on 1Gso metaphase spreads. However, if the breakpoint is close to one end of the clone, visualizing the breakpoint with FISH can be difficult, because the signal would proportionally be very small. By the use of clones, BAC 44-N-23, PAC 54-J-2, and BAC 311-O-6, which overlap each other (Figure 14), this problem was circumvented.

Figure 14. Physical Map of the 1Gso breakpoint region on mouse chromosome 2.

The portion of distal mouse chromosome 2 surrounding the breakpoint is illustrated. The top line represents the position and transcription orientation of the genes known to be located within this interval. Vertical lines delineate 50 kb intervals, and these distances are applied throughout the entire figure. YACs are shown below the top line, with their locations and sizes indicated to one side of the clone. YAC D0214D was studied in more detail, and shows a restriction map of the region, as indicated by the vertical lines. Letters below the vertical lines represent restriction enzymes, with a key for them in the lower left-hand corner of the figure. Below the YAC clones, each BAC (black line) or PAC (gray line) is depicted with individual number designations located underneath. On the left or right of the clone is a number indicating its size in kb. *Not*1 and *Mlu*1 sites are shown on the BAC/PAC contig by solid vertical lines. The different boxes present at the ends of the BAC or PAC clones denote the presence of an endclone. During map construction, if the endclone hybridized to other BACs, PACs, or YACs, an identical box was placed on those clones consistent with the location of the original endclone. The red dashed lines denote hybridization to the BAC and PAC clones with the critical probes (*Rcn*, 44-N-23-T7, 311-O-6-T7, and *Pax6*) that illustrate contig closure.


Breakpoint region
50 Kb
Wt1
Rcn
Pax6
Est 32366
YAC 90-A-1
610kb
YAC 40-B-11
500kb
YAC D0214D
600kb
#4
#2
224-L-6
120kb
587-C-27
180kb
596-D-20
185kb
541-C-3
150kb
287-B-20
120kb
44-N-23
120kb
279-O-16
135kb
202-A-6
95kb
255-O-20
280-J-20
311-O-6
160kb
437-G-5
190kb
460-J-8
180kb
436-F-4
150kb
54-J-2
150kb
389-K-14
110kb
433-M-11
180kb
N=NotI
R=NruI
M=MluI
E=EagI
B=BssHI

The translocation should be seen within two overlapping clones, even if the breakpoint occurs near the end of a clone. The breakpoint was seen to map between *Rcn* and *Pax6* (Figure 13 and 15a). BAC 44-N-23 and PAC 54-J-2 overlap (Figure 14), and when used together in FISH, show a split signal, one of which is very small and not always seen, suggesting that the breakpoint may occur at the end of PAC 54-J-2 (Figure 15b). The proximal end of PAC 54-J-2, where the breakpoint is thought to occur, overlaps with BAC 311-O-6 (Figure 16). A FISH experiment using these two clones shows that the breakpoint occurs between the two, with strong signals for both clones (Figure 15c). This experiment supports the idea that the breakpoint is located on the end of PAC 54-J-2 that overlaps with BAC 311-O-6, because the signal strength is not diminished as seen with FISH hybridizations using BAC 44-N-23 and PAC 54-J-2. FISH analysis using BAC 311-O-6 alone does not show the breakpoint, but the breakpoint may lie very close the end of the BAC. These results confirm that the 1Gso translocation breakpoint is located within a small, well-defined interval spanned by a series of BAC and PAC clones on mouse chromosome 2.

A minimal tiling path of BAC and PAC clones from this interval is now being sequenced by the D.O.E. Joint Genome Institute sequencing teams. The sequence will permit searches for new candidate genes, and generation of new markers that can help isolate and study the 1Gso breakpoint site. This sets the stage for experiments that will permit us to gain access to the 1Gso chromosome 14 breakpoint region and to begin the search for candidate genes on that chromosome.

FISH analysis on chromosome 14

FISH experiments were also performed with markers from mouse chromosome 14 to define the region harboring the breakpoint and investigate some potential candidate genes that were on chromosome 14 (Figure 5). The FISH experiments narrowed down the

interval on mouse chromosome 14 to about 15 cM (Figure 17), which has homology with human chromosome 13. No obvious disease phenotypes on human chromosome 13 can be associated with the 1Gso mouse. However, some potential candidate genes were excluded due to the fact that they do not fall into the region where the breakpoint occurs.

Genes excluded by the FISH analysis on both MMU2 and MMU14

FISH analysis was undertaken on mouse chromosome 14 to begin to narrow down the region of the translocation breakpoint. These experiments were done to initially anchor the breakpoint in association with a set of markers on chromosome 14, in case any potential candidate genes were discovered.

During this series of experiments, a few candidate genes were excluded from the interval. Neurofilament genes, including *Nfl*, are clustered together on mouse chromosome 14 (Julien *et al.*, 1986). Neurofilament genes are differentially expressed in the brain during development and their functions are not well known. A zinc-finger gene was also excluded from candidacy. Zinc finger protein of the cerebellum 2, *Zic2*, is expressed in the adult cerebellum, the area of the brain that coordinates fine motor movement (Aruga *et al.*, 1996).

The brain-derived neurotrophic factor (*Bdnf*) locus in mice was an excellent candidate based upon the phenotype of a targeted deletion of this gene in mice (Ernfors *et al.*, 1994), but is located distal to the *Pax6* locus on mouse chromosome 2, and therefore is not a candidate gene for 1Gso. However, the gene belongs to an interesting gene family that will be discussed in more detail in Part IV.

Figure 15. Overlapping clones used in FISH.

Chromosome 2 FISH analysis. These experiments show that the translocation breakpoint on mouse chromosome 2 occurs between clones which overlap. A) BAC 44-N-23 and clone BAC 311-O-6 are shown to be split with FISH analysis. B) BAC 44-N-23 and overlapping clone PAC 54-J-2 are shown to split with FISH, the signal for the region which spans the breakpoint is very faint, as indicated by the asterisk. Also, this small signal is not always seen with this combination of probes, indicating that the breakpoint is close to one end of the PAC 54-J-2, and is therefore more difficult to visualize. C) PAC 54-J-2 and BAC 311-O-6 overlap, and show strong split signals with FISH, and verifies the location of the translocation breakpoint to the end of PAC 54-J-2 which overlaps with BAC 311-O-6.

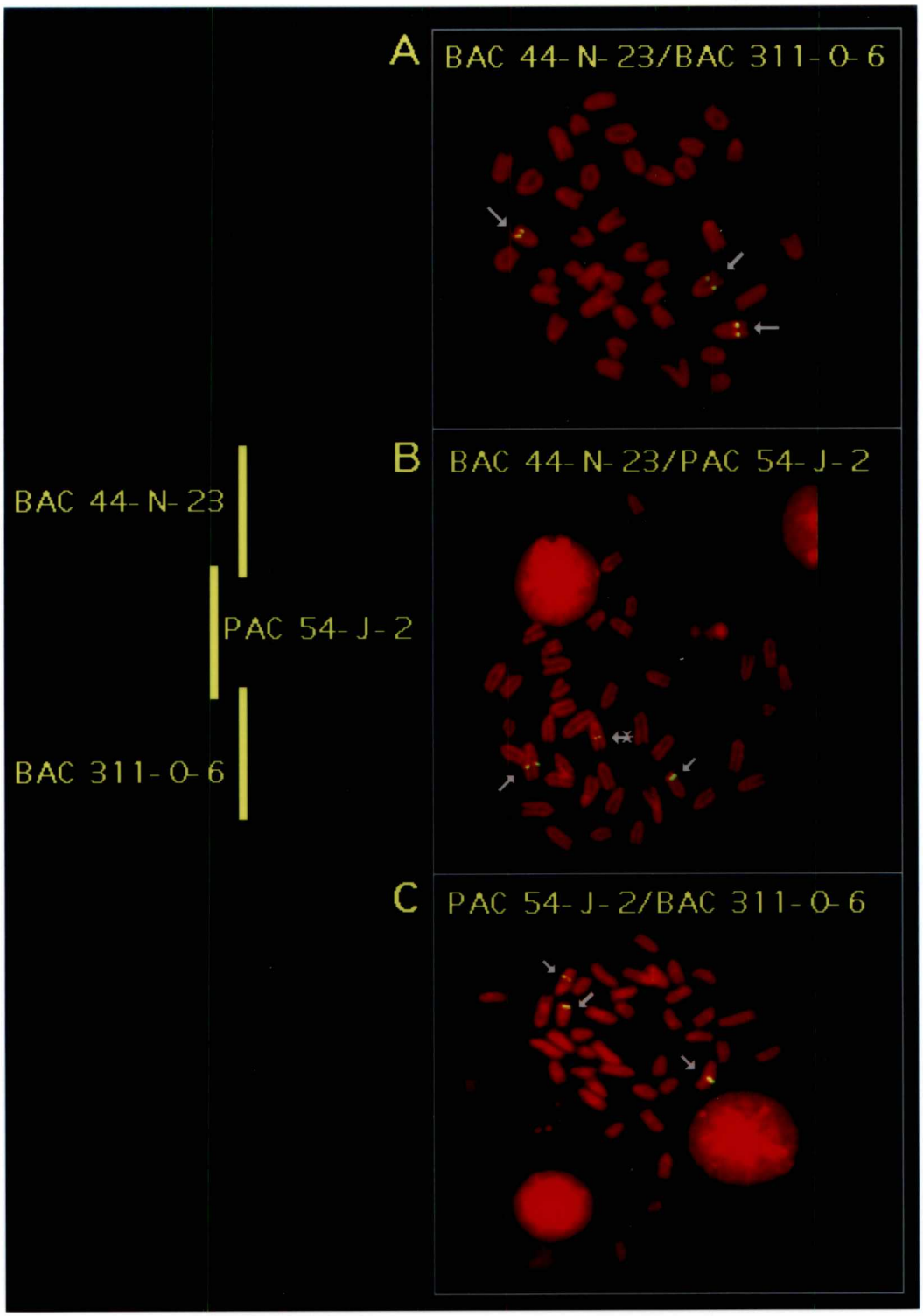

A
BAC 44-N-23/BAC 311-O-6
B
BAC 44-N-23/PAC 54-J-2
BAC 44-N-23
PAC 54-J-2
BAC 311-O-6
C
PAC 54-J-2/BAC 311-O-6

Figure 16. The region of the translocation breakpoint in relation to the physical map.

The three genomic clones used for FISH are illustrated in red. PAC 54-J-2 and BAC 311-O-6 overlap, and show strong split signals with FISH. The region of the translocation breakpoint in the end of PAC 54-J-2 that overlaps with BAC 311-O-6 is highlighted through the whole contig.

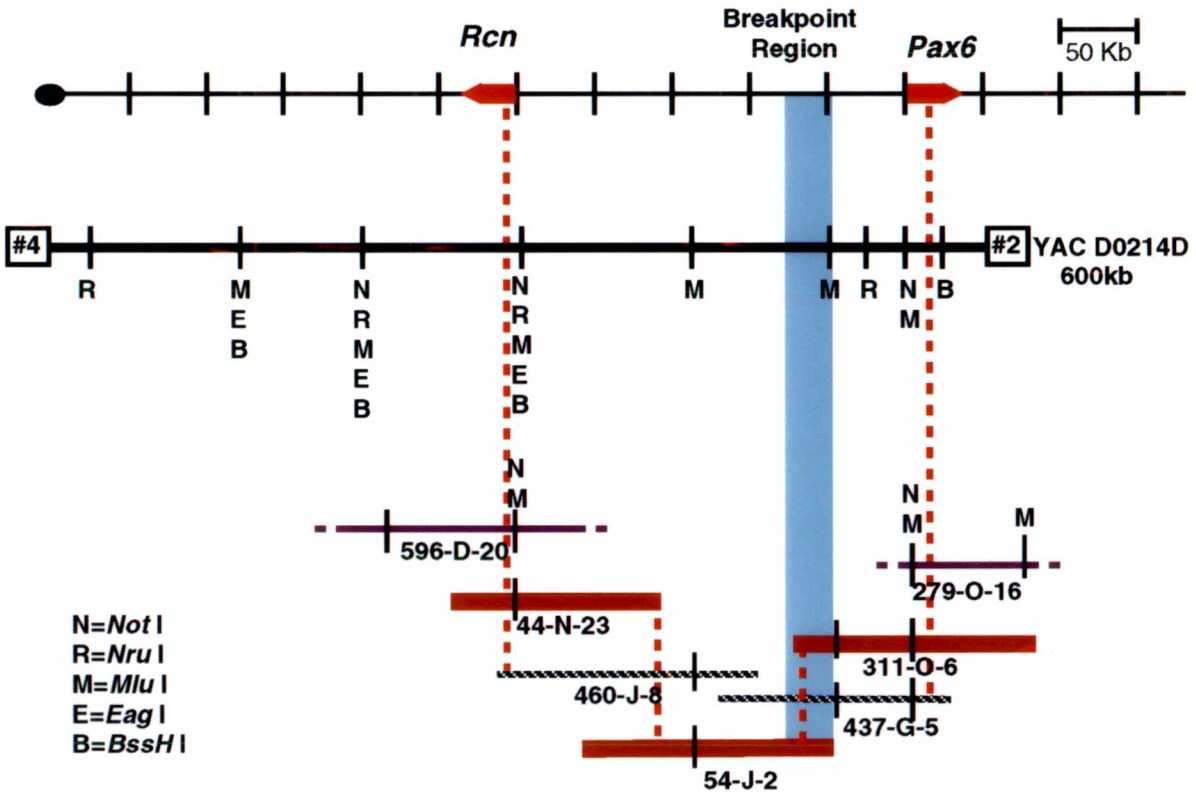
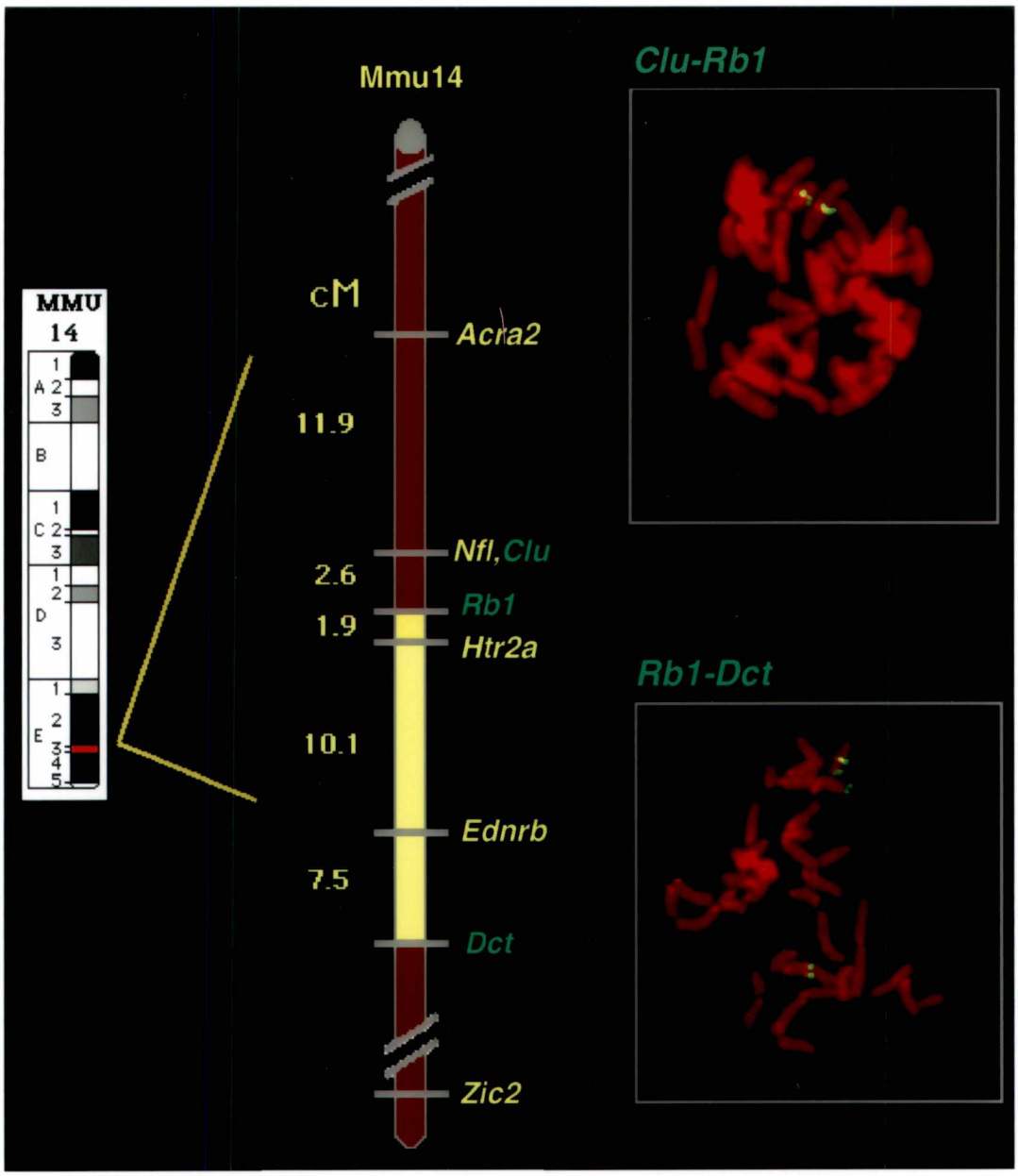

Rcn
Breakpoint Region
Pax6
50 Kb
#4
#2
YAC D0214D
600kb
R
M E B
N R M E B
N R M E B
M
M
R
N M
B
N M
596-D-20
N M
M
279-O-16
44-N-23
311-O-6
460-J-8
437-G-5
54-J-2
N=Not I
R=Nru I
M=Mlu I
E=Eag I
B=BssH I

Figure 17. Chromosome 14 FISH analysis.

On the left is an ideogram of Mmu14 showing the interval between *Acra2* and *Dct*. On right is the results of FISH with BAC clones containing *Clu* and *Rb1* (top box) and *Rb1* and *Dct* (bottom box) hybridized to heterozygous 1Gso metaphase cells. *Clu* and *Rb1* remain linked together during FISH analysis to 1Gso translocation heterozygotes (top box), while *Rb1* and *Dct* are linked on the normal chromosome 14 and separated in the translocated chromosomes (bottom box).


Mmu14
cM
Acra2
11.9
Nfl,Clu
2.6
Rb1
1.9
Htr2a
10.1
Ednrb
7.5
Dct
Zic2
MMU
14
A
B
C
D
E
Clu-Rb1
Rb1-Dct

SECTION 4.

OBSERVATIONS ON THE 1GSO MUTANT PHENOTYPE AND GENETIC STUDIES

Detailed studies of the phenotype of 1Gso animals

In addition to the previously described abnormal escape responses, inability to swim, and delayed ability to right itself, (Figure 2 and 3; Rutledge et al., 1986), 1Gso mutants do not appear to sense environmental cues in a normal manner. For example, the mice do not seem to sense the edges of a table, and will run off the table freely. In contrast, wild-type mice will run away from a threat, but not run off the table unless under great duress. 1Gso animals tilt their head to one side or the other instead of keeping it centered when they walk. The animals tend to curl up with their hindlimbs clasped when held by the tail. Overall, the mutant mice do not appear to have good spatial orientation, which could contribute to the inability to swim. The mice do not exhibit excessive nervousness, which is generally associated with neurological problems. All breeding and cytological data indicate a complete concordance between the inheritance of heterozygosity for the 1Gso translocation and the dominant neurological phenotypes associated with the 1Gso mice. Breeding data have correlated semisterility (indicating the presence of the heterozygous translocation) and the neurological phenotype of the 1Gso animals (Rutledge *et al.*, 1986). Also, cytological studies of twenty 1Gso animals displaying the neurological phenotype indicated that all were translocation heterozygotes as condirmed by either G-banding or FISH techniques (N. Cacheiro, K.T. Cain and E. Carver, unpublished).

Mice homozygous for the translocation have never been reported to survive until birth, and the mutation is therefore presumed to be associated with recessive embryonic lethality (Rutledge *et al.*, 1986). In studies conducted over the past several years, the

number of reabsorption moles relative to healthy fetuses that have been observed in 1Gso/+ females crossed to 1Gso carrier males is not significantly elevated compared to the numbers seen when carrier males are crossed to normal females; the proportion of dead implants is also not higher than percentages observed in crosses involving translocations that are clearly not associated with embryonic lethality (W. Generoso and K.T. Cain, personal communication). Unusual rates of late fetal or neonatal death have not been observed in association with 1Gso carrier crosses in over 5 years of breeding experiments (K.T. Cain, E. Carver, unpublished observations). In 28 late fetuses obtained in crosses between 1Gso heterozygotes, no specific deformities of neurological tissues were observed (Rutledge *et al.*, 1986). Similar results were obtained in a second experiment where late fetuses from four 1Gso translocation female mice were examined (Carver *et al.*, unpublished).

Necropsies were performed on 1Gso/+ mice and wild-type mice, focusing on examinations of basic histopathology and the central nervous system (Rutledge *et al.*, 1986). During these necropsies, no obvious morphological defects of the central nervous system or brain were found in the mutant mice. Brains of ten 1Gso/+ animals were compared to ten wild-type mice by staining and light microscopy for obvious anatomical and histological defects, but no gross abnormalities have been noted (Rutledge *et al.*, 1986). The 1Gso/+ mice do not appear to have defects in vestibular innervation or obvious malformations of the inner ear (Rutledge *et al.*, 1986), as seen in mice with balance defects such as *shaker* and *waltzer* (Lyon and Searle, 1989). Most mutant mice that have an inability to swim, or have a circling phenotype, exhibit inner ear defects (Osako and Hilding, 1971; Cordes and Barsh, 1994; Lyon, M.F. 1953; and others). There are however, some mutant mice with phenotypic characteristics similar to the 1Gso animals, and some of these will be discussed in Part IV.

Complementation studies involving 1Gso mice and animals carrying *Sey*^{Dey}, a deletion spanning the breakpoint region on Mmu2

The region containing *Wt1* and *Pax6* in mice has been extensively investigated , and several strains of mice have been discovered or generated which carry mutations that correlate with homologous human diseases. Several deletions were discovered that completely remove the region containing *Wt1* and *Pax6*, including, *Sey*^{Dey}, which has been well-characterized (Glaser *et al.*, 1990). It is a deletion of ~3000 kb that encompasses *Wt1, f3,* and *Pax6* (Figure 18; Glaser *et al.*, 1990). Animals homozygous for the *Sey*^{Dey} deletion die as embryos, as do *Sey* point mutants (Hill *et al.*, 1991), while heterozygotes for the deletion have eye abnormalities (Theiler *et al.*, 1978). *Sey*^{Dey} was mapped on mouse chromosome 2 (Theiler *et al.*, 1978), and an allele, *Sey*^{MH}, was localized to the human WAGR homologous region specifically (van der Meer-de Jong, 1990). Different *Sey* mutants, including animals heterozygous for deletions, do not exhibit a 1Gso-like neurological phenotype. These observations demonstrate that the *diver* neurological phenotype cannot result from haploinsufficiency of any gene within the *Wt1-Pax6* interval. However, it is possible that genes within the *Sey*^{Dey} deletion might contribute to the neurological or embryonic lethal phenotypes observed in 1Gso mice.

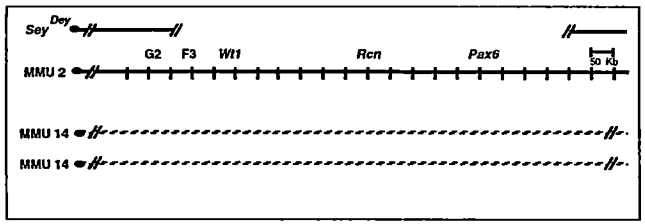
To test for complementation between the 1Gso translocation and genes within the *Sey*^{Dey} deletion, we mated heterozygotes for each mutation to generate compound heterozygous mutants (Figure 18). If any genes contributing to 1Gso neurological defects or lethality were present in the deleted interval, their effects might be uncovered in the compound mutant animals. The offspring of matings between the 1Gso translocation heterozygote and the *Sey*^{Dey} heterozygote were, therefore, examined.

Figure 18. Compound mutant, 1Gso and *Sey*^{Dey}, and the corresponding region in the mouse genome.

The potential chromosome constitutions of offspring generated by matings of 1Gso and *Sey*^{Dey} heterozygotes are depicted, illustrating the region uncovered by the deletion and its association with the 1Gso translocation. Chromosomes are shown as horizontal lines, with the chromosome 2 depicted as a solid line and chromosome 14 a hatched line. Vertical lines on the chromosome delineate a distance of 50 kb, and informative probes are listed above chromosome 2. Illustrated in the box on the top left is the region of mouse chromosome 2 surrounding the deletion in *Sey*^{Dey} heterozygotes and the unaffected chromosome 14. The regions of mouse chromosome 2 and 14 associated with a 1Gso translocation heterozygote are seen in the top right box. Matings between these animals form an F_1 generation with the Mmu 2 and and Mmu 14 genotypes as shown in the lower boxes:

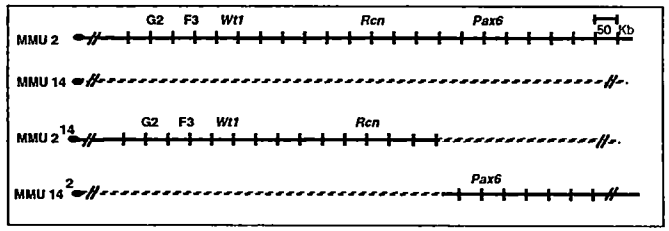
1. A *Sey*^{Dey} phenotype alone.
2. A normal phenotype.
3. A 1Gso phenotype alone.
4. A compound phenotype involving both the 1Gso and the *Sey*^{Dey} phenotype together.

Dickie's Small eye (Sey^{Dey}) Deletion



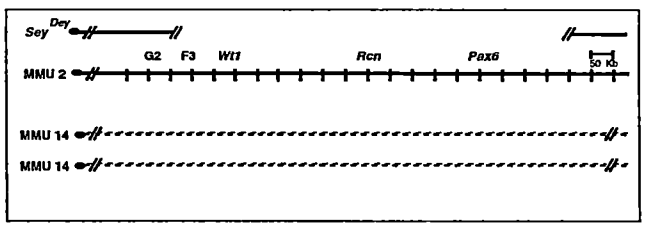
X

1Gso translocation

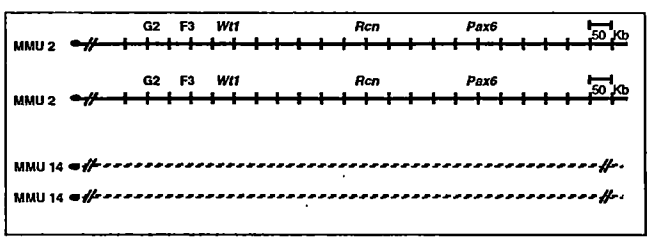


F_1 GENERATION

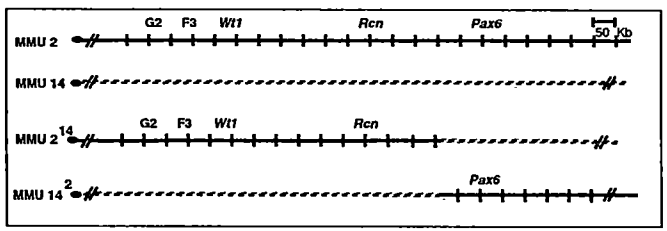
1 Dickie's Small eye (Sey^{Dey}) Deletion



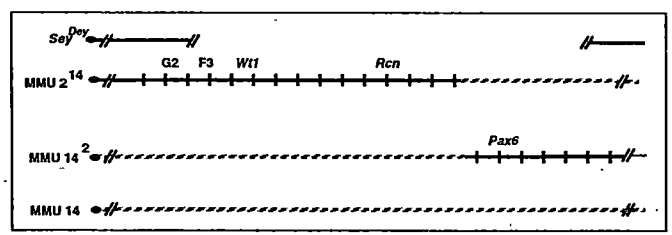
2 Normal animal



3 1Gso translocation



4 Compound Mutant



Five *Sey*^{Dey} male mice were mated to five 1Gso/+ females and to eight different wild-type females. The *Sey*^{Dey} male mice failed to produce offspring in any matings, suggesting that they were sterile. Ten *Sey*^{Dey} female mice were then mated to ten 1Gso/+ males. The phenotypes of the expected offspring were of four types: normal, 1Gso type alone, *Sey*^{Dey} type alone, or a combination of 1Gso and *Sey*^{Dey} (Figure 18), in a Mendelian ratio of 1:1:1:1.

Six litters have been born with an average of three pups per litter. All possible phenotypes were seen, with four animals displaying the compound 1Gso and *Sey*^{De} phenotypes. The number of compound mutants born (4 compound mutants out of 18 total offspring) correlates roughly with the 25% ratio expected from a cross of this nature. The compound phenotype was slightly more delayed, by an average of 3-4 days, in the ability to right themselves than animals carrying either *Sey*^{De} or 1Gso alone, but at weaning age these mice were no better or worse than normal-eyed *diver* littermates. These data indicate that neither the dominant neurological defect nor the recessive lethality associated with the 1Gso mutation can be attributed to null mutations of genes located on chromosome 2. If null mutations of genes on chromosome 2 did contribute significantly to either phenotype, we would expect a more severe expression of the neurological defect or a lethality that would block survival of the compound mutants. Other possible explanations for the basis of the 1Gso phenotype are discussed in the Part IV.

SECTION 5.

CONSTRUCTION OF A COSMID LIBRARY OF 1GSO/+ DNA AND TAR CLONING TO GAIN ACCESS TO THE BREAKPOINT

Since specific segments of chromosomes 2 and 14 are fused together in 1Gso DNA, genomic clones containing the translocation breakpoint will contain DNA sequences derived from both affected chromosomes. This will enable a "walk" across the breakpoint and the beginning of research that will help define the contribution of genes on Mmu 14 to the 1Gso phenotype. Two different experiments have been initiated which will allow access to the breakpoint region of Mmu 14.

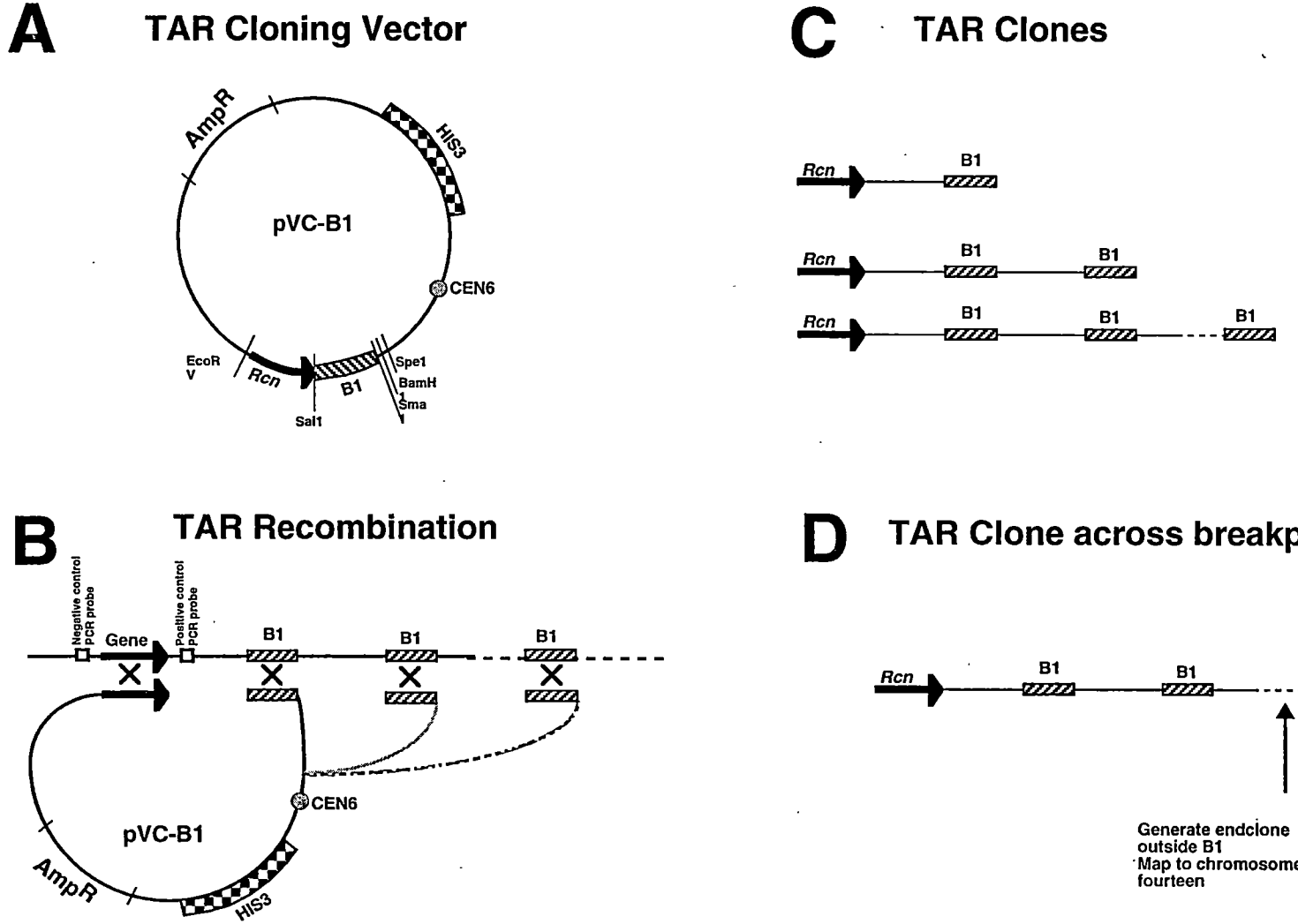
The first approach involves construction of a 1Gso/+ cosmid library. Cosmid clones that are near the translocation should carry the translocated pieces of DNA together on one clone. Therefore, this library will be screened with the closest markers known to flank the translocation, to isolate a clone that harbors the translocation breakpoint. The ends of any clone discovered by this method will be sequenced, and if both ends match sequence previously generated by the Mmu 2 contig, then they would not harbor the breakpoint. However, if one end matched Mmu 2 sequences, while the other does not, the clone could contain the translocation. The end would be mapped, and if found to localize to Mmu 14, then this clone would contain one of the 1Gso translocation breakpoints. This clone would show the location of the rearrangement on Mmu 14, and provide a reagent for sequencing and gene discovery. The sequence of this clone would pinpoint the rearrangement, and start a gene search on Mmu 14. Also, it would serve as a template to isolate markers that in turn could be used to isolate clones on mouse chromosome 14, to extend the region and search for more candidate genes. The construction of a cosmid library of 1Gso/+ DNA has been initiated. The vector, pCos2, has been grown, purified, and modified for ligation according to laboratory protocols.

The 1Gso/+ DNA has been purified, digested, and is ready for ligation. A series of tests have been performed, and the library is ready for large-scale construction. However, because of recent developments we have decided to construct a new library using DNA from a compound mutant *Sey*^{Dev}/1Gso. This mutant carries the Mmu2 translocation breakpoint in hemizygous form, since the corresponding *Wt1* to *Pax6* region on mouse chromosome 2 is deleted in *Sey*^{Dev} mice (Figure 18). Use of DNA from this compound mutant for the library means that the library will contain only the region containing the translocation breakpoint and no clones derived from a normal chromosome 2.

As a second approach, a new technique, transformation-associated recombination (TAR) cloning (Larionov *et al.*, 1996), will be used to cross over the disrupted region on Mmu2 to Mmu14. TAR cloning in yeast has been used to selectively isolate regions of interest with minimal sequence data (Larionov *et al.*, 1996, Kouprina *et al.*, 1998, others). This technique has been used to isolate large genes, such as the human BRAC2, which is over 90 kb in length (Larionov *et al.*, 1997). By using homologous recombination that occurs during yeast transformation, one can isolate specific DNA sequences and clone them into the TAR vector without construction of large genomic libraries, which is tedious and time consuming. Also, the construction of a library for each individual positional cloning project involving numerous different translocation mutant mice would be very difficult, time consuming, and ultimately wasteful, since the library would be of limited use outside the original project. Therefore, the TAR cloning method has been applied to the region of mouse chromosome 2 associated with the translocation breakpoint to see if it is a viable option for gaining molecular access to mouse chromosome 14.

Figure 19. TAR cloning vector and strategy.

This figure illustrates the TAR cloning vector and strategy to capture the 1Gso breakpoint in compound mutant *Sey*^{Dey}/1Gso DNA. A) The TAR vector is shown with the *Rcn* segment of DNA inserted into the vector. B) Double homologous recombination occurs between Rcn and B1 sequences in the TAR cloning vector and *Sey*^{Dey}/1Gso genomic DNA. This recombination event "captures" the sequences contained between *Rcn* and the B1 repeat in *Sey*^{Dey}/1Gso genomic DNA. C) TAR clones isolated from this region will vary in length depending on the position of B1 repeat elements present in the region. The longer the clone, the more likely it crosses over the translocation breakpoint. D) The translocation breakpoint-containing TAR clone can be discovered by isolating clones from the B1-end and hybridizing them to the interspecific backcross to determine if one localizes to mouse chromosome 14.


A
TAR Cloning Vector
AmpR
HIS3
pVC-B1
CEN6
EcoR V
Rcn
B1
Spe1
BamH 1
Sma
Sal1
B
TAR Recombination
Negative control PCR probe
Gene
Positive control PCR probe
B1
B1
B1
CEN6
pVC-B1
AmpR
HIS3
C
TAR Clones
Rcn
B1
Rcn
B1
B1
Rcn
B1
B1
B1
D
TAR Clone across breakpoint
Rcn
B1
B1
B1
Generate endclone outside B1
Map to chromosome fourteen

A TAR vector was used to incorporate the *Rcn* gene (Figure 19a). The mouse *Rcn* gene was used as a directional anchor, while a mouse B1 repeat sequence was used as a "hook" to isolate different portions of the corresponding mouse genomic region in a site specific manner (Figure 19b). The directionality of the clone is critical. If the orientation of the clone were incorrect, the clone would undergo TAR cloning in an improper direction, and not be useful. PCR amplification from different areas surrounding the anchor is used for a positive and negative control. The negative control would exclude recombination that occurs in the wrong orientation, and therefore would not extend toward the breakpoint region. Positive controls are used to verify that recombination has occurred, and that the TAR clone is extending toward the translocation breakpoint (Figure 19c-d). The mouse genomic DNA isolated from Sey^{Dey}/1Gso compound mutant mice will be used as the template for the TAR clones for the same reasons it is being used in construction of the cosmid library. The TAR clones circumvent the need for screening a 1Gso mutant cosmid library and the length of time needed to analyze and verify a clone that contains the breakpoint. One of these clones could be long enough to cross the breakpoint (Figure 19d). This clone can be verified by end-cloning the B1-end and hybridizing the endclone to a backcross panel to see if it maps to chromosome 14. If it does, then the breakpoint is contained within the TAR clone. The completed vector and DNA have been sent to our collaborators, V. Larionov, V. Noskov, and N. Kouprina, for further analysis.

SUMMARY

This project has shown the feasibility of using FISH analysis to narrow down translocation breakpoints in mice to a defined interval that is amendable to positional cloning strategies. The region on mouse chromosome 14 that carries the breakpoint was narrowed down the region to an ~17 cM interval where there are few known genes, but all have homology to human chromosome 13. The breakpoint on mouse chromosome 2 has been localized to a 150 kb PAC clone in wild type DNA. Two different strategies for the positional cloning of the breakpoint are being pursued, and were discussed. An in-depth summary of the research, future directions, and consequences of rearrangements caused by translocations are investigated in PART IV.

REFERENCES

Adams, M.D., Kelley, J.M., Gocayne, J.D., Dubnick, M., Polymeropoulos, M.H., Xiao, H., Merril, C.R., Wu, A., Olde, B., Moreno, R.F., Kerlavage, A.R., McCombie, W.R., and Venter, J.C. (1991). Complementary DNA sequencing: expressed sequence tags and human genome project. *Science* **252**:1651-1656.

Arai, E., Tokino, T., Imai, T., Inazawa, J., Ikeuchi, T., Tonomura, A., and Nakamura, Y. (1993). Mapping the breakpoint of a constitutional translocation on chromosome 22 in a patient with NF2. *Genes Chrom & Cancer* **6**:235-238.

Aruga, J., Nagai, T., Tokuyama, T., Hayashizaki, Y., Okazaki, Y., Chapman, V.M., and Mikoshiba, K. (1996). The mouse zic gene family. Homologues of the Drosophila pair-rule gene odd-paired. *J Biol Chem* **271**:1043-1047.

Avner, P., Amar, L., Dandolo, L., and Guénet, J. L. (1988). Genetic analysis of the mouse using interspecific backcross analysis. *Trends Genet* **4**:18-23.

Ballabio, A. (1991). The rise and fall of positional cloning? *Nature Genet* **3**:277-279.

Bastian, H., Gruss, P, Duboule, D., and Izpisua-Belmonte, J.C. (1992). The murine even-skipped-like gene Evx-2 is closely linked to the Hox-4 complex, but is transcribed in the opposite direction. *Mammal Genome* **3**:241-243.

Bernasconi, P., Cavigliano, P.M., Genini, E., Castagnola, C., Malcovati, L., Calatroni, S., Caresana, M., Boni, M., Alessandrino, E.P., Astori, C., Corso, A., and Bernasconi, C.

(1998). A complex translocation (5;7) in a patient with acute nonlymphocytic leukemia evolved from a myelodysplastic syndrome. *Cancer Genet Cytogenet* **105**:182-186.

Burgess, D.L., Jones, J.M., Meisler, M.H., and Noebels, J.L. (1997). Mutation of the Ca2+ channel beta subunit gene Cchb4 is associated with ataxia and seizures in the lethargic (lh) mouse. *Cell* **88**:385-392.

Burgess, D.L., Kohrman, D.C., Galt, J., Plummer, N.W., Jones, J.M., Spear, B., and Meisler, M.H. (1995). Mutation of a new sodium channel gene, Scn8a, in the mouse mutant 'motor endplate disease'. *Nature Genet* **10**:461-465.

Bultman S.J., Russell, L.B., Gutierrez-Espeleta, G.A., and Woychik, R.P. (1991). Molecular characterization of a region of DNA associated with mutations at the agouti locus in the mouse. *Proc Natl Acad Sci* **88**:8062-8066.

Bultman, S.J., Michaud E.J., and Woychik, R.P. (1992). Molecular characterization of the mouse agouti locus. *Cell* **71**:1195-1204.

Burke, D. T., Carle, G.F., and Olson, M. V. (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* **236**:806-812.

Cacheiro, N.L., and Russell, L.B. (1975). Evidence that linkage group IV as well as linkage group x of the mouse are in chromosome 10. *Genet Res* **25**:193-195.

Cacheiro, N.L., Rutledge, J.C., Cain, K.T., Cornett, C.V., and Generoso, W. M. (1994). Cytogenetic analysis of malformed mouse fetuses derived from balanced translocation heterozygotes. *Cytogenet Cell Genet* **66**:139-148.

Chaplin, D.D., and Brownstein, B.H. (1993). Yeast artificial chromosome libraries. In Current Protocols in Molecular Biology" (F.M. Ausubel, R., Brent, R.E. Kingston, D.D. Moore,J.G. Seidman, J.A. Smith, K. Struhl, L.M. Albright, D.M. Coen and A. Varki, Eds) Greene Pub. Assoc. & John Wiley & Sons, New York. Vol. 1: pp 610-19.

Collins F. S. (1992). Positional cloning: Let's not call it reverse anymore. *Nature Genet* **1**:3-6.

Committee on Standardized Genetic Nomenclature for Mice. (1972). Standard karyotype of the mouse, *Mus Musculus. J Hered* **63**:69-72.

Cooke, H. (1987). Cloning in yeast: an appropriate scale for mammalian genomes. *Trends Genet* **3**:173-174.

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

Cordes, S.P., and Barsh, G.S. (1994). The mouse segmentation gene kr encodes a novel basic domain-leucine zipper transcription factor. *Cell* **79**:1025-34.

de Martinville, B., Kunkel, L.M., Bruns, G., Morle, F., Koenig, M., Mandel, J.L., Horwich, A., Latt, S.A., Gusella, J.F., Housman, D., and Francke, U. (1985). Localization of DNA sequences in region Xp21 of the human X chromosome: search for molecular markers close to the Duchenne muscular dystrophy locus. *Am J Hum Genet* **37**:235-249.

Doyle, J., Ren, X., Lennon, G., and Stubbs, L. (1997). Mutations in the Cacnl1a4 calcium channel gene are associated with seizures, cerebellar degeneration, and ataxia in tottering and leaner mutant mice. *Mammal Genome* **8**:113-120.

Edelhoff, S., Grond-Ginsbach, C., von Eggeling, F., Korholz, D., Gobel, U., and Senger, G. (1998). Complex chromosomal rearrangements associated with congenital erythrophagocytotic histiocytosis. *Clin Genet* **53**:298-302.

Ernfors, P., Lee, K.F., and Jaenisch, R. (1994). Mice lacking brain-derived neurotrophic factor develop with sensory deficits. *Nature* **368**:147-150.

Fantes, J., Redeker, B., Breen, M., Boyle, S., Brown, J., Fletcher, J., Jones, S., Bickmore, W., Fukushima, Y., Mannens, M., Danes, S., van Heyningen, V., and Hanson, I. (1995). Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum Mol Genet* **4**:415-422.

Flagiello, D., Apiou, F., Gibaud, A., Poupon, M. F., Dutrillaux, B., and Malfoy, B. (1997). Assignment of the genes for cellular retinoic acid binding protein 1 (CRABP1) and 2 (CRABP2) to human chromosome band 15q24 and 1q21.3, respectively, by *in situ* hybridization. *Cytogenet Cell Genet* **76**:17-18.

Frengen, E., Zhao, B., Weichenhan, D., Gjernes, E., Jessee, J., Prydz, H., and de Jong,P.J. (1996). Modular large insert cloning vectors for propagation in bacterial and mammalian cells. Unpublished.

Friezner-Degen, S.J., Schaefer, L.A., Jamison, C.S., Grant, S.G., Fitzgibbon, J.J., Pai, J.A., Chapman, V.M., and Elliott, R.W. (1990). Characterization of the cDNA coding for mouse prothrombin and localization of the gene on mouse chromosome 2. *DNA Cell Biol* **9**:487-498.

Generoso, W. M., Krishna, M., Cain, K.T., and Shue, C.W. (1981). Comparison of two methods for detecting translocation heterozygotes in mice. *Mutat Res* **81**:177-186.

Gessler, M., Simola, K.O.J., and Bruns, G.A.P. (1989). Cloning of breakpoints of a translocation identifies the AN2 locus. *Science* **244**:1575-1578.

Glaser, T., Jones, C.A., Lewis, W.H., Call, K., Rose, E.A., Buckler, A., Ito, C., And Housman, D.E. (1989). The ultrafine structure of the WAGR gene complex. *Cytogenet Cell Genet* **51**:1005.

Glaser, T., Lane, J., and Housman, D. (1990). A mouse model of the aniridia-Wilms tumor deletion syndrome. *Science* **250**:823-827.

Hake, L.E., Alcivar, A.A., Hecht, N.B. (1991). Changes in mRNA length accompany translational regulation of the somatic and testis-specific cytochrome c genes during spermatogenesis in the mouse. *Development* **110**:249-257.

Heidmann, O., Buonanno, A., Geoffroy, B., Robert, B., Guenet, J.L., Merlie, J.P., and Changeux, J.P. (1986). Chromosomal localization of muscle nicotinic acetylcholine receptor genes in the mouse. *Science* **234**:866-868.

Herrmann, B.G., Labeit, S., Poustka, A,, King, T.R., and Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. *Nature* **343**:617-622.

Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R. (1995). The WashU-Merck EST Project. Unpublished.

Hill, R.E., Favor, J., Hogan, B.L., Ton, C.C., Saunders, G.F., Hanson, I.M., Prosser, J., Jordan, T., Hastie, N.D., and van Heyningen, V. (1991). Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* **354**:522-525.

Huang, H.J., Yee, J.K., Shew, J.Y., Chen, P.L., Bookstein, R., Friedmann, T., Lee, E.Y., and Lee, W.H. (1988). Suppression of the neoplastic phenotype by replacement of the RB gene in human cancer cells. *Science* **242**:1563-1566.

Ioannou, P.A., Amemiya, C.T., Garnes, J., Kroisel, P.M., Shizuya, H., Chen, C., Batzer, M.A., and de Jong, P.J. (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet* **1**:84-89.

Julien, J.P., Meyer, D., Flavell, D., Hurst, J., and Grosveld, F. (1986). Cloning and developmental expression of the murine neurofilament gene family. *Brain Res* **387**:243-250.

Kent, J., Lee, M., Schedl, A., Boyle, S., Fantes, J., Powell, M., Rushmere, N., Abbott, C., van Heyningen, V., and Bickmore, W.A. (1997). The reticulocalbin gene maps to the WAGR region in human and to the Small eye Harwell deletion in mouse. *Genomics* **42**:260-267.

Kim, J., Carver, E.A., and Stubbs, L. (1997). Amplification and sequencing of end fragments from bacterial artificial chromosome clones by single-primer polymerase chain reaction. *Anal Biochem* **253**:272-275.

Kouprina, N., Graves, J., Cancilla, M.R., Resnick, M.A., and Larionov, V. (1997). Specific isolation of human rDNA genes by TAR cloning. *Gene* **197**:269-276.

Lai, F., Stubbs, L., Lehrach, H., Huang, Y., Yeom, Y., and Artzt, K. (1994). Genomic organization and expressed sequences of the mouse extended H-2K region. *Genomics* **23**:338-343.

Lamhonwah, A.M., Barankiewicz, T.J., Willard, H.F., Mahuran, D.J., Quan, F., and Gravel, R.A. (1986). Isolation of cDNA clones coding for the alpha and beta chains of human propionyl-CoA carboxylase: chromosomal assignments and DNA polymorphisms associated with PCCA and PCCB genes. *Proc Natl Acad Sci* **83**:4864-4868.

Larionov, V., Kouprina, N., Graves, J., Chen, X.N., Korenberg, J.R., and Resnick, M.A. (1996). Specific cloning of human DNA as yeast artificial chromosomes by transformation-associated recombination. *Proc Natl Acad Sci* **93**:491-496.

Larionov, V., Kouprina, N., Graves, J., and Resnick, M.A. (1996). Highly selective isolation of human DNAs from rodent-human hybrid cells as circular yeast artificial chromosomes by transformation-associated recombination cloning. *Proc Natl Acad Sci* **93**:13925-13930.

Lawrence, J. B. (1990). A fluorescence *in situ* hybridization approach for gene mapping and the study of nuclear organization. In Davies, K.E. and Tilghman, S.M. (eds), Genome Analysis Volume 1; Genetic and Physical Mapping. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 1-38.

Lehrach, H. (1990). Hybridization fingerprinting in genome mapping and sequencing. In Davies, K.E. and Tilghman, S.M. (eds), Genome Analysis Volume 1; Genetic and Physical Mapping. Cold Spring Harbor Laboratory, Cold Spring Harbor, pp. 39-81.

Letts, V.A., Felix, R., Biddlecome, G.H., Arikkath, J., Mahaffey, C.L., Valenzuela, A., Bartlett, F.S. 2nd, Mori, Y., Campbell, K.P., and Frankel, W.N. (1998). The mouse stargazer gene encodes a neuronal Ca2+-channel gamma subunit. *Nature Genet* **19**:340-7.

Litt, M., Luty, J., Kwak, M., Allen, L., Magenis, R.E., and Mandel, G. (1989). Localization of a human brain sodium channel gene (SCN2A) to chromosome 2. *Genomics* **5**:204-208.

Lyon, M. F. (1953). Absence of otoliths in the mouse: An effect of the pallid mutant. *J Genet* **51**:638-650.

Lyon, M.F., and Searle, A.G. (eds) (1989). Genetic Variant and Strains of the Laboratory Mouse. 2nd edn. Oxford Univ. Press

MacGregor, T.M., Copeland, N.G., Jenkins, N.A., and Gigiere, V. (1992). The murine gene for cellular retinoic acid-binding protein type II: genomic organization, chromosomal localization, and post-transcriptional regulation by retinoic acid. *J Biol Chem* **267**:7777-7783.

Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisel, S., Kucaba, T., Lacy, M., Le, M., Martin, J ., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Wilson, R., and Waterston, R. (1996). The WashU-HHMI Mouse EST Project. Unpublished.

Mercer, J.F.B., J. Livingston, B. Hall, J.A. Paynter, C. Begy, S. Chandrasekharappa, P. Lockhart, A. Grimes, M. Bhave, D. Siemieniak, and T.W. Glover (1993). Isolation of a partial candidate gene for Menkes disease by positional cloning. *Nature Genet* **3**:20-25.

Morrison, A.M., Jager, U., Chott, A., Schebesta, M., Haas, O.A., and Busslinger. M. (1998). Deregulated PAX-5 transcription from a translocated IgH promoter in marginal zone lymphoma. *Blood* **92**:3865-3878.

Mueller, P.R., and Wold, B. (1989). *In vivo* footprinting of a muscle specific enhancer by ligation mediated PCR. *Science* **246**:780-786.

Myers, M.W., Lazzarini, R.A., Lee, V.M., Schlaepfer, W.W., and Nelson, D.L. (1987). The human mid-size neurofilament subunit: a repeated protein sequence and the relationship of its gene to the intermediate filament gene family. *EMBO J* **6**:1617-1626.

Oettinger, M.A., Stanger, B., Schatz, D.G., Glaser, T., Call, K., Housman, D., and Baltimore, D. (1992). The recombination activating genes, RAG 1 and RAG 2, are on chromosome 11p in humans and chromosome 2p in mice. *Immunogenetics* **35**:97-101.

Osako, S., Hilding, D.A. (1971). Electron microscopic studies of capillary permeability in normal and ames waltzer deaf mice. *Acta Otolaryngol* **71**:365-376.

Ozawa, M. (1995). Cloning of a human homologue of mouse reticulocalbin reveals conservation of structural domains in the novel endoplasmic reticulum resident Ca(2+)-binding protein with multiple EF-hand motifs. *J Biochem* **117**:1113-1119.

Parker, J. D., Rabinovitch, P. S., and Burmer, G. C. (1991). Targeted gene walking polymerase chain reaction. *Nucleic Acids Res* **19**:3055-3060.

Parks, C. L., Chang, L.-S., and Shenk, T. (1991). A polymerase chain reaction mediated by a single primer: cloning of genomic sequences adjacent to a serotonin receptor protein coding region. *Nucleic Acids Res* **19**:7155-7160.

Patil, N., Cox, D.R., Bhat, D., Faham, M., Myers, R.M., and Peterson, A.S. (1995). A potassium channel mutation in weaver mice implicates membrane excitability in granule cell differentiation. *Nature Genet* **11**:126-129.

Philbrick, W.M., Palfree, R.G., Maher, S.E., Bridgett, M.M., Sirlin, S., and Bothwell, A.L. (1990). The CD59 antigen is a structural homologue of murine Ly-6 antigens but lacks interferon inducibility. *Eur J Immunol* **20**:87-92.

Pohlmann, R., Krentler, C., Schmidt, B., Schroder, W., Lorkowski, G., Culley, J., Mersmann, G., Geier, C., Waheed, A., Gottschalk, S., Grzeschik, K-H., Hasilik, A., and von Figura, K. (1988). Human lysosomal acid phosphatase: cloning, expression and chromosomal assignment. *EMBO J* **7**:2343-2350.

Quan, F., Korneluk, R.G., Tropak, M.B., and Gravel, R.A. (1986). Isolation and characterization of the human catalase gene. *Nucleic Acids Res* **14**:5321-5335.

Rochelle, J.M., Watson, M.L., Oakey, R.J., and Seldin, M.F. (1992). A linkage map of mouse chromosome 19: definition of comparative mapping relationships with human chromosomes 10 and 11 including the MEN1 locus. *Genomics* **14**:26-31.

Rose, E.A., Glaser, T., Jones, C., Smith, C.L., Lewis, W.H., Call, K.M., Minden, M., Champagne, E., Bonetta, L., Yeger, H., and Housman, D.E. (1990). Complete physical map of the WAGR region of 11p13 localizes a candidate Wilms' tumor gene. *Cell* **60**:495-508.

Rowley, J.D. (1998). The critical role of chromosome translocations in human leukemias. *Annu Rev Genet* **32**:495-519.

Rutledge, J. C., Cain, K.T., Cacheiro, N. L. A., Cornett, C. V., Wright, C. G., and Generoso, W. M. (1986). A balanced translocation in mice with a neurological defect. *Science* **231**:395-397.

Schatz, D.G., Oettinger, M.A., and Baltimore, D. (1989). The V(D)J recombination activating gene, RAG-1. *Cell* **59**:1035-1048.

Schlessinger, D. (1990). Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Genet* **6**:251-258.

Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. *Proc Natl Acad Sci* **89**:8794-8797.

Smith, M.W., Clark S.P., Hutchinson, J.S., Wei, Y.H., Churukian, A.C., Daniels, L.B., Diggle, K.L., Gen, M.W., Romo, A.J., Lin, Y., Selleri, L., McElligott, D.L., and Evans, G.A. (1993). A sequence-tagged site map of human chromosome 11. *Genomics* **17**:699-725.

Squire, J., Dryja, T.P., Dunn, J., Goddard, A., Hofmann, T., Musarella, M., Willard, H.F., Becker, A.J., Gallie, B.L., and Phillips, R.A. (1986). Cloning of the esterase D gene: a polymorphic gene probe closely linked to the retinoblastoma locus on chromosome 13. *Proc Natl Acad Sci* **83**:6573-6577.

Stubbs, L. (1992). Long-range walking techniques and positional cloning strategies. *Mammal Genome* **3**:127-142.

Stubbs, L., Carver, E.A., Cacheiro, N.L.A., Shelby, M., and Generoso, W. (1997). Generation and characterization of heritable reciprocal translocations in mice. *Methods: A Companion to Methods in Enzymology* **13**:397-408.

Stubbs, L., Huxley, C., Hogan, B., Evans, T., Fried, M., Duboule, D., and Lehrach. H. (1990). The Hox-5 and surfeit gene clusters are linked in the proximal portion of the mouse chromosome 2. *Genomics* **6**:645-650.

Surralles, J., and Natarajan, A.T. (1998). Position effect of translocations involving the inactive X chromosome: physical linkage to XIC/XIST does not lead to long-range *de novo* inactivation in human differentiated cells. *Cytogenet Cell Genet* **82**:58-66.

Theiler, K., Varnum, D.S., and Stevens, L.C. (1978). Development of Dickie's small eye, a mutation in the house mouse. *Anat Embryol* **155**:81-86.

Ton, C.C., Hirvonen, H., Miwa, H., Weil, M.M., Monaghan, P., Jordan, T., van Heyningen, V., Hastie, N.D., Meijers-Heijboer, H., Drechsler, M., Royer-Pokora, B., Collins, F., Swaroop, A., Strong, L.C., and Gaunders, G. (1991). Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell* **67**:1059-1074.

Trask B. J. (1991). Fluorescence *in situ* hybridization: applications in cytogenetics and gene mapping. *Trends Genet* **7**:149-151.

van der Meer-de Jong, R., Dickinson, M. E., Woychik, R. P., Stubbs, L., Hetherington, C., and Hogan, B. L. M. (1990). Location of the gene involving the small eye mutation on mouse chromosome 2 suggests homology with human aniridia 2 (AN2). *Genomics* **7**:270-275.

Wallace, M.R., Marchuk, D.A., Anderson, L.B., Letcher, R., Odeh, H.M., Saulino, A.M., Fountain, J.W., Brereton, A., Nicholson, J., Mitchell, A.M., Brownstein, B.H., and Collins, F.S. (1990). Type 1 neurofibromatosis gene: Identification of a large transcript disrupted in three NF1 patients. *Science* **249**:181-186.

Wang W.-P., Myers, R. L., and Chiu, I.-M. (1991). Single primer-mediated polymerase chain reaction: application in cloning of two different 5'-untranslated sequences of acidic fibroblast growth factor mRNA. *DNA and Cell Biol* **10**:771-777.

Weterman, M.A., Wilbrink, M., and Geurts van Kessel, A. (1996). Fusion of the transcription factor TFE3 gene to a novel gene, PRCC, in t(X;1)(p11;q21)-positive papillary renal cell carcinomas. *Proc Natl Acad Sci* **93**:15294-15298.

Wood, L., Theriault, N., and Vogeli, G. (1989). Vimentin cDNA clones covering the complete intermediate-filament protein are found in an EHS tumor cDNA library. *Gene* **76**:171-175.

Woychik, R.P., Generoso, W.M., Russell, L.B., Cain, K.T., Cacheiro, N.L., Bultman, S.J., Selby, P.B., Dickinson, M.E., Hogan, B.L., and Rutledge, J.C. (1990). Molecular and genetic characterization of a radiation-induced structural rearrangement in mouse chromosome 2 causing mutations at the limb deformity and agouti loci. *Proc Natl Acad Sci* **87**:2588-2592.

Woychik, R.P., Stewart, T.A., Davis, L.G., D'Eustachio, P., and Leder, P. (1985). An inherited limb deformity created by insertional mutagenesis in a transgenic mouse. *Nature* **318**:36-40.

Wurster, D.H. (1972). Mouse chromosomes identified by trypsin-Giemsa (T-G) banding. *Cytogenetics* **11**:379-387.

PART IV

CONCLUSIONS

Chapter 1.

INTRODUCTION

The research projects detailed in Parts II and III of this dissertation utilized comparative mapping of mouse and human homology regions to further our understanding of genome evolution and applied this genomic homology data to aid in the isolation of a translocation breakpoint in the mouse.

As described in Part II, comparative maps of human chromosome 19 and related regions of the mouse genome were established. The analysis of these genomic region relationships revealed both similarities that have been maintained and differences that have arisen in chromosome structure during mammalian evolution. In the course of these studies, a fine-structure comparative map of two small mouse homology segments was generated, including one that had not been previously described, carrying murine genes related to those located in human 19p13.1 region. Physical maps of these regions are now being constructed and will provide interesting substrates for comparative sequence analysis, providing data to aid in understanding the structure of evolutionarily unstable regions and the mechanisms associated with their creation.

The positional cloning study, Part III, centered on the 1Gso translocation mutant and relied upon the establishment of an interspecific backcross map of a 15 cM region of central mouse chromosome 2. The map, in turn, was used to guide a fluorescence *in situ* hybridization (FISH) strategy that permitted the localization of the 1Gso Mmu 2 translocation breakpoint to a small interval. A chromosome walk established a YAC, BAC, and PAC clone contig across the breakpoint interval, and FISH was finally used to bracket the breakpoint within a single PAC clone of 150 kb. This research provides critical reagents for future studies aimed at isolating the gene(s) associated with the 1Gso

phenotype. The next section fully summarizes the research and discusses future directions for each individual project.

SECTION 1.

THE COMPARATIVE MAP OF CHROMOSOME 19p AND RELATED REGIONS OF THE MOUSE GENOME

Summary of research

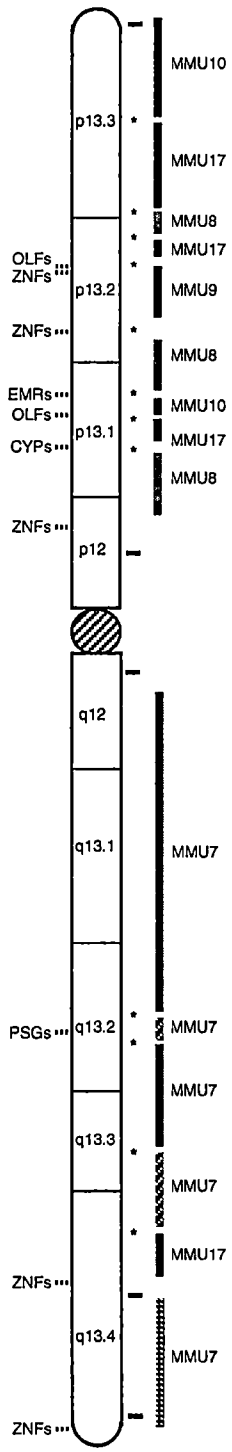
Comparative mapping studies focused on human chromosome 19 and the related regions of the mouse genome have been very informative for many aspects of research including comparative physical map construction, evolutionary studies (Lamerdin *et al.*, 1996), disease gene cloning in humans (Ophoff *et al.*, 1996) and identification of candidate genes for mouse mutations (Doyle *et al.*, 1997). In fact, the utility of these mapping data have just begun to be realized; the comparative sequence analysis of homologous human and mouse regions, the real endpoint of these studies, has just started. The biological information that will be gleaned from this "ultimate comparative map" will have an enormous, wide-ranging impact.

The overall comparative map of human 19 and the related regions in the mouse genome illustrates some general principles. The human chromosome shares homology with at least five mouse chromosomes: 7, 8, 9, 10, and 17 (Figure 1). Sequences located on the p-arm of human chromosome 19 are rearranged to a much greater extent than 19q in mouse, carrying homology with at least nine distinct segments over four mouse chromosomes, compared to the larger q-arm which carries homology to only two known regions in the mouse (Figure 1). There are twelve regions of homology between the human and mouse maps that have been spanned by cosmid, BAC, and/or YAC contigs in humans to date and some are currently being sequenced (Figure 1 and http://www-bio.llnl.gov/genome/genome.html). The sequence derived from these regions will help elaborate the homology relationships and provide more insight into mechanisms of genome diversity.

Figure 1. Human chromosome 19 and homology regions in the mouse genome.

The ideogram of chromosome 19 is shown in the center. Located to the right are the corresponding mouse homology regions. In between the two as indicated by an asterisk are homology region borders that are currently being sequenced, or have completed contigs that span them. The horizontal lines indicate gaps in the human contigs. To the left of the human chromosome are gene family member localizations on the chromosome. These are: OLFs, olfactory receptor genes; ZNFs, zinc-finger genes; EMRs, EGF-like module containing, mucin-like, hormone receptor-like sequences; and CYPs, cytochrome P450, gene families. This figure shows the correlation with the homology regions, gene families, and the homology borders that are spanned by completed contigs.

CHROMOSOSOME 19 AND MOUSE HOMOLOGY REGION BORDERS



The research presented in Part II provides the basis for construction of comparative physical maps in several interesting regions and illustrates the need for further, more detailed characterization of homology regions in the human and mouse genomes. Studies focusing on human chromosome 19p13.1 have revealed the existence of a previously unknown segment of homology between human chromosome 19 and central mouse chromosome 10. This research has permitted further documentation of the observation that appears to hold true for most or all of human chromosome 19, that is, that the borders of segments that are rearranged in mouse relative to human are populated by tandemly clustered gene family members.

Conclusions and future directions

Are certain regions of the mouse and human genomes especially prone to chromosome translocations, inversions, or other rearrangements that give rise to the diverse genomic organization of these two species?

Genomic rearrangements and expansions appear to have played a critical role in the early divergence of many species. One thing that has become apparent is that even if gene order has been preserved, the distances between the genes can vary, and the repeats that fall between genes may be different (J. Kim and M. Shannon, personal communication). Duplication and differential expansion of gene families, including ZNF, OLF, and HOX (Homeo box) genes, have been seen in several cases (Glusman *et al.*, 1996; Gao *et al.*, 1998; Shannon *et al.*, 1996 and personal communication; Rouquier *et al.*, 1998; Amores *et al.*, 1998; Trask *et al.*, 1998; Brand-Apron *et al.*, 1999). By studying individual genes, one can determine if the duplicated gene(s) have also diverged in function. Only with the detailed comparisons of the various gene family clusters from different genomes will the evolutionary impact of gene duplication become clear (Shannon *et al.*, 1996; Holland and Garcia-Fernadez. 1997; Amores *et al.*, 1998).

Inter- and intrachromosomal recombination have been suggested as common mechanisms of genomic rearrangement, and there is evidence that repeated sequences may help facilitate these events within the genome (Dutly and Schnitzel, 1996; van Deutekom *et al.*, 1996; Kehrer-Sawatski *et al.*, 1997; Lupski, 1998). Recombination between misaligned, related sequences seems to occur within regions of the greatest identity. For example, *GH1* leads to the absence of growth hormone in certain families exhibiting growth retardation or short stature (reviewed by Procter *et al.*, 1998). A deletion of growth hormone 1, *GH1*, in patients with short stature was characterized in detail and found to result from illegitimate recombination between flanking segments of 594 bp, which are 99% identical (Vnencak-Jones and Phillips, 1990). Also, genes associated with Smith-Magenis syndrome (SMS), one of the more commonly observed microdeletion syndromes, are flanked by repeated gene clusters between which homologous recombination can occur, this resulting in the deletion of internal segments. The SMS region contains three different repeat gene clusters and more than 90% of deletion patients surveyed were shown to have deletions associated with recombination in one region between the outer two gene clusters (Chen *et al.*, 1997).

Evolutionary rearrangements might also be associated with repeated sequences. For example, the breakpoint of a translocation t(X;1) (p11; q21) associated with papillary renal cell carcinoma (RCC) was mapped to a small region of H1q21 between *SPTA1* and a clustered gene family, including *CD1C, CD1B, CD1A, CD1D, CACY*, and at least four other members (Weterman *et al.*, 1996). Interestingly, the boundary between the two segments of H1q21 that are related to Mmu1 and Mmu3, respectively, is located between *SPTA1* and *CD1C*, a region of less than 200 kb (Oakey *et al.*, 1992). The *S100* genes in this region are clustered in mouse chromosome 3 and human chromosome 1q21 (Ridinger *et al.*, 1998). This cluster of genes is rearranged in organization between

human and mouse (Ridinger *et al.*, 1998), and the entire region of 1q21 is prone to rearrangements including translocations, deletions and duplications.

Comparisons between related regions of Mmu7 and H19q permitted the borders of several evolutionary rearrangements to be localized physically on the human map (Stubbs *et al.*, 1996). One rearrangement involves the transposition of clustered pregnancy specific glycoprotein (PSG) gene family members to a more proximal position within Mmu7, while an adjacent gene family, carcinoembryonic antigen (CEA), has remained in the same relative position in human and mouse (Stubbs *et al.*, 1996). The borders of some homology region segments are surrounded by tandemly clustered zinc-finger (ZNF) genes (Stubbs *et al.*, 1996; Kim *et al.*, 1997; Carver and Stubbs *et al.*, 1997), which are particularly prevalent on human chromosome 19 (Ashworth *et al.*, 1995). Recent studies within several of the ZNF clusters have demonstrated that these regions have changed rapidly during recent evolutionary time, and have been subject to differential rearrangements and expansions since the divergence of human and murine species (M. Shannon, J. Kim, unpublished data). Clustered ZNF genes are frequently involved in translocations associated with human cancer (Rabbitts *et al.*1994, and references therein), confirming the notion that these tandemly duplicated sequences are particularly prone to rearrangement.

Sequence comparison across evolutionary breakpoints and cytogenetic band boundaries has the potential to discover and open up whole new arenas for future research that cannot be approached today. Large-scale sequence comparison has just begun and limited data are available at the present time. Some of the regions that have been examined include the T-cell receptor complex (Koop and Hood, 1994), the BTK region of the X-chromosome (Oeltjen *et al.*, 1997), the ERCC2 region (Lamerdin *et al.*, 1996) and the XRCC1 region (Lamerdin *et al.*, 1995). The sequence analysis of a human olfactory gene cluster on chromosome 17 indicates that, indeed, the olfactory genes can

serve as a site for nonhomologous recombination, as seen in the fusion of two adjacent olfactory receptor genes (Glusman *et al.*, 1996).

The structure of the region surrounding *MEL* and *CD97* on human chromosome 19 does seem to suggest that this region has been prone to rearrangements in both species (Part II. Chapter 2 section 3, Carver *et al.*, in preparation). This region is occupied by members of several different gene families, and the presence of these locally repeated sequences could cause the region to be less stable by providing a template for recombination between gene family members. The sequencing of homology region borders in the human and mouse genomes will further reveal the structural similarities that might have provided substrates for recombination and rearrangements during genome evolution. If our hypothesis is correct, sequences located on "opposite sides" of a homology region "breakpoint" should be closely related to one another. The rearrangements should appear as a translocation, inversion or transposition of sequences from one homology region to another through recombination within closely related sequences.

State of sequence generated and analyzed

The sequencing of other model organisms will highlight differences in genomes that will begin to elucidate the evolution of animals into different species. The large scale comparisons of genomic regions in humans and other species will provide insights into gene structure, function, and regulatory regions, as well as chromosomal evolution which would not easily be seen in the analysis of any one species alone. The genomes of other model organisms have been completely sequenced. The bakers' yeast, *Saccharomyces cerevisiae*, was the first eukaryote whose genome is entirely sequenced (Goffeau *et al.*, 1996). *E. Coli* (Blattner *et al.*, 1997), and *C. elegans* (The *C. elegans* Sequencing Consortium, 1998) have also been completely sequenced. These genomes are providing

researchers with a first look into the types of genes that are in each genome, and are giving an idea as to which genes are critical to the life of a cell. As more organisms' sequence becomes available, a whole new area of informatics-based research is becoming popular (Bailey *et al.,* 1998; Seldin, 1997).

The Human Genome project now has goals to complete the entire human sequence by 2003 (Collins *et al.,* 1998). With significant portions of the human sequence becoming available, large sequencing efforts are focusing on the mouse genome. Mouse sequencing goals (Collins *et al.,* 1998) include the development of physical and genetic mapping resources, and additional cDNA resources, with the complete sequence of the mouse genome by 2005.

Once the sequence of the mouse is completed, the science of positional cloning will forever change. Before, regional sequencing was done as a means to clone a gene, or as a by-product of the gene search itself. Before long, once the sequence is done, it will be only a matter of localizing a mutation to a certain region, and acquiring all the candidate genes for testing. There will begin to be a shift from gene discovery to analysis of gene function, and the mouse will be the model organism in which to do this type of research for many years. Already, there are methodologies to generate new mutations in the mouse genome to take advantage of the sequencing efforts (Justice *et al.,* 1998; You *et al.,* 1997).

SECTION 2. CONCLUSIONS OF THE 1GSO POSITIONAL CLONING PROJECT AND FUTURE DIRECTIONS

Summary of research

The positional cloning project that served as a second major aim of this dissertation project focused upon a translocation mutant, 1Gso, and the development and implementation of techniques required to map and clone regions surrounding the translocation breakpoints. This project began with generation of an interspecific backcross map of the region surrounding the translocation breakpoint located in central chromosome 2. Construction of the map, in turn, was facilitated greatly by comparative studies that permitted the derivation of new markers based upon the better-developed human genetic and physical maps; the ability to use mapped markers from one species to increase the resolution of maps in the other species was a critical factor in the success of this work.

This project applied fluorescence *in situ* hybridization techniques to localize 1Gso Mmu 2 and Mmu 14 translocation breakpoints, and to define the Mmu 2 breakpoint within an interval that was small enough to be accessible by standard positional cloning techniques. Multiple applications of the FISH strategy permitted 1Gso breakpoint to be narrowed down from and initial region of ~15 cM to a single PAC clone, 54-J-2, which is ~150 kb in length. This effort provides data and critical reagents that will advance the final search for gene(s) associated with the 1Gso phenotype. The ~450 kb BAC and PAC clone contig that was constructed in the breakpoint region has now been submitted to the Joint Genome Institute for complete sequencing. This sequence will be used to further define the breakpoint site.

Conclusions and future directions

Sequencing across the breakpoint region in mouse chromosome 2.

The minimal overlapping set of clones being sequenced (BAC 596-D-20, PAC 460-J-8, PAC 437-G-5, and BAC 279-O-16 (Chapter III, Figure 14)) is ~450 kb in length and contains the region on mouse chromosome 2 harboring the breakpoint. Since the homologous *RCN-PAX6* region has been sequenced in *Fugu rubripes* (Miles *et al.,* 1998) and is in the process of being sequenced in humans (Niederfuhr *et al.,* 1998), the mouse sequence that will be generated in this region will permit a study of genome evolution of unusual depth. This region will be the first, of any substantial size, to be sequenced in three vertebrate species. The sequence will be a good template for comparing the structure of the region among three diverse species, and will give researchers access to multiple-species genomic DNA scans that are just now beginning to be explored. These data will also provide useful reagents to researchers focused on analysis of WAGR mutations in human patients.

No new genes were shown to be present between *Rcn* and *Pax6* in *Fugu rubripes* (Miles *et al.,* 1998), and no other genes have been discovered in the corresponding human homology region (Kent *et al.,* 1997; Niederfuhr *et al.,* 1998). These data suggest that additional genes are also not likely to be found within this region in mouse. Even so, during evolution the addition of gene sequences can occur through retrotransposition, duplication, or other means (Carlton *et al.,* 1995; Holland and Garcia-Fernandez, 1996). Therefore, the sequencing of this region will be important to search for new genes; any discovered would be a potential candidate gene, and would be investigated further.

The mouse genomic sequence will serve as an anchor to compare sequence data derived from the cosmid and TAR cloning efforts, and provide a wild-type sequence against which the 1Gso breakpoint region can be aligned. Comparison of sequences of breakpoint spanning cosmid or TAR clones with that of the existing, wild-type contig

sequence will allow us to pinpoint the breakpoint to a specific sequence and identify small deletions or other rearrangements that might be associated with the translocation at that site.

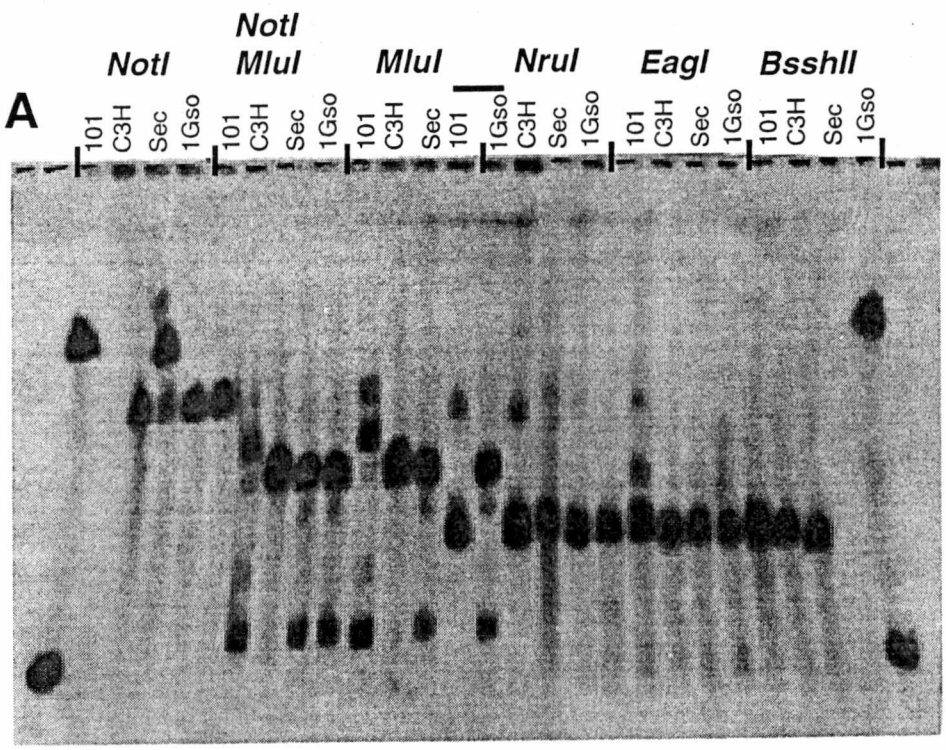
Once the Mmu 2 breakpoint site is defined, it can be used to isolate chromosome 14-containing clones so that that breakpoint can be explored. Any gene discovered by sequencing will be isolated and characterized to see if it has been disrupted by the 1Gso translocation. Size variations between wild-type and 1Gso genomic DNA can be seen if the gene is located close to or is disrupted by the translocation, however, genes near the translocation on Mmu 2 have not shown any size variations (Figure 2).

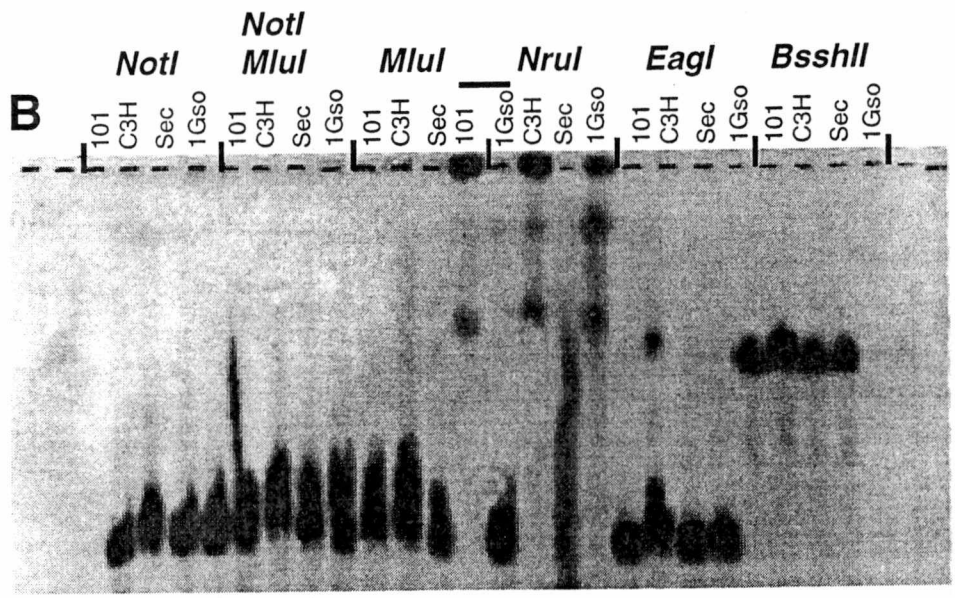
A suitable candidate gene should be expressed at the correct time and in the right tissues during development. Any candidate gene involved in the dominant neurological phenotype could be expressed in developing neural tissues, or tissues which can effect the developing central nervous system. However, as described below, it is also possible that the gene associated with the dominant effects is normally expressed elsewhere, but turned on inappropriately in the brains of mutant mice.

In the case of translocations, screening for aberrant expression through analysis of simple Northern blots, RNase protection, and/or *in situ* hybridization, is a good test to establish if a candidate gene is involved in the mutant phenotype. A gene that has been disrupted by rearrangements can give rise to different expression, aberrantly-size transcripts, or no expression at all (Bultman *et al.*, 1992; Culiat *et al.*, 1993; Mercer *et al.*, 1993). These possibilities are outlined in detail in the following section.

Figure 2. Analysis of genomic DNA of 1Gso and background animals.

DNA isolated from 1Gso and inbred strains of mice from which the 1Gso stock was derived were digested with rare cutter enzymes and size separated by pulsed-field gel electrophoresis. These digests were then Southern blotted and hybridized with probes surrounding the translocation breakpoint. The restriction enzyme used to cleave the DNA is illustrated at the top of each figure. Underneath the restriction enzymes are the strains of animals used as sources of DNA. The horizontal line indicates where two samples were inadvertently switched during the preparation of the pulsed-field gel. The clones used detected no rearrangements between the 1Gso and background animals at the genomic level. A) No rearrangements between the 1Gso and other animals were visualized with a Wilms tumor probe. B) No rearrangements between the 1Gso and other animals were visualized with a *Pax6* probe.


A
NotI
NotI
MluI
MluI
NruI
EagI
BsshII
101
C3H
Sec
1Gso


B
NotI
NotI
MluI
MluI
NruI
EagI
BsshII
101
C3H
Sec
1Gso

Types of mutations associated with translocations and their possible relevance to the *diver* phenotype

Translocations have been associated with many types of disease-related phenotypes, ranging from anirida to cancer (Gessler *et al.*, 1989; Mercer *et al.*, 1993; reviewed by Rabbitts, 1994; Rowley, 1998). Translocations cause the DNA environment to be altered greatly, and can affect a gene(s) in a variety of different ways. Therefore, it is important to keep in mind the different types of disruptions that can be caused when searching for gene(s) associated with a translocation. These possibilities must be considered in the context of the fact that 1Gso exhibits both a dominant neurological and a recessive lethal phenotype. This phenotype could represent the effects of disrupting one or more genes on either or both of the affected chromosomes. These possibilities, and their potential relevance to the 1Gso mutant given data presented in this thesis, are outlined below.

Types of mutations associated with translocations

1. Null mutations

The chromosomal rearrangement could occur inside a gene or within the regulatory elements controlling the gene's expression, knocking out the gene and obliterating its function. If the mutation occurs inside the coding sequence, a translocation may, for example, introduce premature stop codons that would lead to instability of the message and lack of translation. Such a rearrangement would lead to a null mutation. A null mutation could easily explain the recessive lethality, but could produce the dominant neurological phenotype only through haploinsufficiency of a gene whose function is dosage-dependent. Some translocations have been seen to cause haploinsufficiency in humans (Gessler *et al.*, 1989; Kurahashi *et al.*, 1998). Haploinsufficiency of the *Pax6* gene is associated with small eyes in mice and aniridia in humans (Glaser *et al.*, 1996).

However, if a gene on mouse chromosome 2 is disturbed by the 1Gso mutation, the disruption cannot be associated with haploinsufficiency, based upon the complementation experiments with the *Sey*^{Dey} mutation presented in chapter III. First, *Sey*^{Dey} heterozygotes carry a deletion of sequences surrounding the 1Gso breakpoint but do not exhibit a diver-like neurological disorder. Second, the compound mutant animals (*Sey*^{Dey}/1Gso) survived and exhibited a simple, compound phenotype typical of each of the two dominant mutations. These experiments showed that the 1Gso phenotype - either the dominant neurological defects or the recessive lethality - cannot be due to knockout mutations of gene(s) contained within the *Sey*^{Dey} deletion.

Null mutations of genes located at or near the breakpoint site should be detected by hybridization of candidate gene probes to Northern blots containing brain RNA isolated from *Sey*^{Dey}/1Gso animals (if the gene is on Mmu 2), or through RNase protection of RNA isolated from 1Gso/+ mice (if the gene is located on Mmu 14). The hybridization of a candidate gene probe to Northern blots containing RNA from tissues isolated from Sey^{Dey}/1Gso and wild-type control animals should show a complete lack of a transcript in the Sey^{Dey}/1Gso tissue compared to wild-type, if the candidate gene is associated with a null mutation on Mmu 2. If the translocation generates a null mutation of a gene from Mmu 14, on similar Northern blots the transcript should be differentially expressed, indicating a lower expression level. However, to accurately quantify mRNA levels, RNase protection assays would be employed to measure the differences between 1Gso and wild-type gene expression. Definitive association of a null mutation with the 1Gso phenotype could be accomplished with transgenic experiments described later.

2. Production of truncated proteins

Translocations can also cause truncation of a gene so that only part of the protein is produced. This type of mutation could result in a simple loss-of-function mutation, or it

could produce a protein with altered functions. Several truncation mutations in *PAX6* show an increased DNA-binding capacity, but are nonfunctional since they lack the transactivation activity from the C-terminal portion of the normal protein (Singh *et al.*, 1998). Competition of the truncated protein with the normal PAX6 protein for DNA-binding sites causes the mutations to function as dominant-negative repressors (Singh *et al.*, 1998). This type of mutation could explain the 1Gso phenotype if the gene is located on chromosome 14. Because 1Gso does not break within the body of either *Rcn* or *Pax6*, production of truncated proteins involving these genes can be ruled out. However, another, yet to be identified gene on chromosome 2 may still be involved. A mutation that results in expression of a truncated but stable protein with altered activities would not be expected to complement the Sey^{Dey} deletion of the normal gene.

Northern blot analysis of RNA from tissues of 1Gso/+ and wild-type animals can be utilized to visualize the presence of a truncated gene from either chromosome 2 or 14. RNA species produced from a candidate gene that is truly truncated by a translocation should be discernable as a smaller-than-normal sized transcript in 1Gso/+ RNA; this transcript should be clearly visible in addition to the normal-sized transcript present in RNA from both animals. Isolation of cDNA from the mutant animal and subsequent sequencing can give an indication of this rearrangement by the absence of sequence corresponding to the entire, intact cDNA.

3. Overexpression or expression of genes at ectopic sites

A translocation could couple regulatory elements from a gene on one chromosome with a gene on another chromosome (Morrison *et al.*, 1998a; Busslinger *et al.*, 1996). Because one gene has lost its regulatory element and the other has gained a new regulatory element, both genes would have aberrant expression, which could alter the time, amount, and tissue specificity of both proteins. This scenario is plausible for the

1Gso translocation. For instance, *Pax6*, since it is close to the translocation and its 5' end is oriented toward the translocation breakpoint (Part III, Figure 14), could be under the control of regulatory elements from an unknown gene from chromosome 14. The new regulation could cause a gain-of-function mutation of *Pax6* that causes the 1Gso phenotype due to overexpression of the gene combined with ectopic expression. Still, the gene on chromosome 14, which lost regulatory elements, would have to be investigated to determine if it contributed to the phenotype. The opposite exchange of regulatory elements could be true as well. There are no known regulatory elements located more than 8 kb 5' of the *Pax6* gene (Okladnova *et al.*, 1998). However, an unknown regulatory element farther upstream from *Pax6* could be altering expression of a gene from chromosome 14 and disrupting *Pax6* expression simultaneously.

Altered expression patterns should be seen by hybridization of Northern blots with a candidate gene probe, or RNase protection assays that compare 1Gso/+ and wild-type RNA. By disturbing endogenous regulatory sequences or by fusing unlinked regulatory and coding sequences, translocations can be associated with expression of a gene in the wrong tissues, at different times in development, or at different concentrations in tissues where the gene in normally expressed.

4. Association of the translocation with other, complex rearrangements

Translocations might also be associated with deletions, insertions, or inversions of sequences at the translocation site. In this case, the phenotype expressed could represent a complex one; comparable to those associated with syndromic phenotypes in humans (e.g. DiGeorge syndrome, Wilson *et al.*, 1991; WAGR, Riccardi *et al.*, 1978). However, FISH mapping and contig construction have not detected a deletion on mouse chromosome 2, and our present mapping data show that if a region of chromosome 2 is

deleted or otherwise rearranged, the affected region should be very small. This does not rule out the possibility that a deletion is associated with the chromosome 14 breakpoint, however. Evaluation of this possibility must await cloning and molecular analysis of the breakpoint region of chromosome 14.

Once the region containing the breakpoint has been cloned from mutant genomic DNA, sequences from the chromosome 14 breakpoint region will be used as probes to isolate surrounding DNA from wild-type genomic libraries. The mapping, contig construction, and sequencing of this region between wild-type DNA and 1Gso DNA will reveal the presence or absence of a deletion associated with the Mmu 14 translocation. If a deletion does occur on Mmu 14, the subsequent cloning and verification of candidate genes will become problematic only if the deletion is very large. Otherwise, the sequencing of the region in wild-type DNA will identify candidate genes, and the association of the candidate gene to the 1Gso phenotype can still be proven by techniques discussed below.

5. Disruptions of genes on both chromosome 2 and 14

In summary, our data provide strong evidence that null mutations of any gene located on chromosome 2 cannot explain the 1Gso phenotype. However, it is possible that both chromosomes involved with the translocation could contain genes that contribute to the 1Gso phenotype. Translocations are also frequently associated with fusion of transcription units, and production of fusion proteins with novel activities (Correl, 1993; Reviewed by Rabbitts, 1994; Rowley, 1998). These types of mutations can also produce dominant phenotypes through gain of novel functions that are not seen in normal animals. Two previously separate genes could be combined to form an entirely different set of proteins, disrupting the functions of each individual protein profoundly. Many human cancers have been associated with this type of gene fusion (Corral *et al.*, 1993;

Megonigal *et al.*, 1998; Super *et al.*, 1997 Kurokoawa *et al.*, 1998). The fusion products can lead to "gain- or loss-of-function" mutations in either or both genes (Morrison *et al.*, 1998b; Bennicelli *et al.*, 1996). The formation of a chimeric protein can link two different proteins together to so that the function of each individual protein is lost, and the new protein functions in a completely novel manner. The normal function of downstream target proteins can also be disturbed by the fusion protein (Kurokawa *et al.*, 1998). A chimeric protein can also produce a dominant-negative effect by competing with the normal protein products of the translocated genes for binding sites.

Because of the location of the 1Gso Mmu 2 breakpoint, fusion proteins involving *Rcn* or *Pax6* can be ruled out in this mutation. A fusion protein can only be made as a result of the 1Gso translocation if another gene or large open reading frame is found at the Mmu 2 breakpoint site.

The fusion of two separate genes can be visualized on Northern blots containing RNA from the 1Gso/+ and wild-type animals. The fusion products can be expressed at the wrong time or in the wrong tissue, and be seen as a size variant, or combinations of these patterns. By isolation of cDNA from the mutant animal, subsequent sequencing can give an indication of this type of rearrangement by containing novel sequence attached somewhere to the mutant cDNA, as compared to normal cDNA.

It is, of course possible that genes on Mmu 2 are overexpressed or ectopically expressed, and at the same time a gene or genes on Mmu 14 are also disrupted in 1Gso mice. In this case, the phenotype would represent composite effects of these independent sets of mutations.

Dissecting out the contributions of candidate genes

Definitive proof that a gene is responsible for certain aspects of the mutant phenotype may ultimately need to be accomplished by generation of transgenic mice. For example,

in the case that a deletion or null mutation is discovered, mice could be designed to carry a targeted disruption of the candidate gene; alternatively, a BAC or other clone could be used to determine whether expression of genes contained on those clones can rescue the phenotype in mutant animals (Culiat *et al.*, 1995; Antoch *et al.*, 1997). Animals carrying the extra gene could be bred with the 1Gso translocation to generate a compound mutant. If these animals were normal, and did not exhibit the phenotype, the candidate gene would have proof of association to the phenotype.

These scenarios would seem to work only if the 1Gso mutation creates a nonfunctional allele of the gene. If the gene is missexpressed due to being linked to a different promoter, or forms a fusion protein with another gene, then straightforward transgenic approaches would not be as valid an option to formally prove gene candidacy. Still, by recreating the mutation - be it a fusion protein, or by the addition of regulatory elements to the candidate gene - and introducing it into mice, the phenotype might be recapitulated.

Potential types of genes involved in the 1Gso phenotype, as seen by other murine phenotypes.

The unique series of defects of the 1Gso heterozygotes are not seen in many neurological mutations. However, some mouse mutations, including both spontaneous mutations and targeted deletions of single genes, have some similarities with the 1Gso phenotype. Other mutant animals can provide clues as to the type of gene involved in the 1Gso phenotype.

1. *Pallid*

One recessive mutation, pallid, occurs in a different region of mouse chromosome 2. These animals have an inability to swim and some have abnormal behavior including head tilting, but a defined inner ear defect accounts for the phenotype (Lyon, M.F. 1953;

Trune and Lim, 1983). Even though this is an obvious inner ear defect, the nervous system pathway between the inner ear and brain could be affected in 1Gso animals.

2. Targeted deletions of dopamine receptors

Dopamine is a neurotransmitter in different brain pathways that are involved in emotional, neuroendocrine, cognitive and locomotor functions (Lindvall and Bjorland, 1983). Dopamine is also synthesized in neurons which project from the substantia nigra pars compacta to the striatum, a structure associated with initiation and coordination of voluntary movement (Graybiel, 1990). The dopamine receptors have been associated with sensorimotor behavior and somatosensory orientation in mice (Weihmuller *et al.*, 1990). Different dopamine receptors show a variety of phenotypes. For example, animals with a targeted disruption of the dopamine beta-hydroxylase gene were poor swimmers, and some would drown unless rescued (Thomas and Palmiter, 1997). Mice with disruptions of the dopamine D4 receptor are less active in open field tests, but display increased coordination of fine motor movements as seen on a rotarod test (Rubinstein *et al.*, 1997). In contrast, targeted deletions of the dopamine D1 and D3 receptor genes both show an increase in locomotor activity in open field tests (Xu *et al.*, 1994; Accili *et al.*, 1996). Most of the targeted deletions did not produce any major changes in the anatomy of the brain (Thomas and Palmiter, 1997; Rubinstein *et al.*, 1997; Xu *et al.*, 1994; Accili *et al.*, 1996). The different receptor pathways could be explored in collaboration with other researchers. These collaborations could give functional information about the 1Gso phenotype, and provide insight into the biological consequences associated with the translocation.

3. Neurotrophic factors and targeted deletion of the *Bdnf* gene in mice

A null mutation was generated for the brain-derived neurotrophic factor (*Bdnf*) locus in mice (Ernfors *et al.*, 1994). Initially, this was a potential candidate gene for the 1Gso phenotype, due to its localization to the breakpoint region on mouse chromosome 2. Mice lacking *Bdnf* showed defective coordination of movements and balance, as well as head bobbing and tilting, circular movements, and a reduced size when compared to normal littermates (Ernfors *et al.*, 1994). Some aspects of this phenotype are similar to the 1Gso animal. The coordination and balance defects, and reduced body size are the most apparent. However, unlike the 1Gso animals, these mutants had observable defects in the peripheral sensory ganglia. After FISH experiments were begun, this gene was excluded from the breakpoint interval, and therefore was not a direct candidate for the 1Gso phenotype. However, *Bdnf* belongs to a family of genes, neurotrophins, which are located in various positions within the mouse genome. These genes are implicated in neuronal development in mouse (Conover *et al.*, 1995; Klein *et al.*, 1994). If a gene is disrupted that affects the signaling between peripheral nerves, or in the processing of the sensory input itself, the resulting phenotype could be very close to the 1Gso phenotype, as illustrated in the *Bdnf* mutant animals.

A hypothesis regarding the molecular basis of the 1Gso mutation and correlation to its phenotype

Within Parts III and IV of this dissertation, the 1Gso phenotype and consequences of the translocation have been described. The fact that no obvious defects in the inner ear, or gross hearing loss has been observed, compounded with the sensory defects and lack of any gross difference in the brain size, or number of cell types in the brain, implicates a gene related to a neurotrophin or a neurotransmitter as the best candidate gene for the 1Gso phenotype.

Evidence already indicates that Mmu 14 is more likely than Mmu 2 to harbor a gene that is disrupted by the 1Gso translocation. The translocation could easily disrupt a single neurotransmitter on Mmu 14 to produce a truncated protein that is nonfunctional. The nonfunctional copy would lead to haploinsufficency in translocation heterozygotes, which survive embryogenesis and have an observable neurological defect. If both copies of the gene are nonfunctional, as in the translocation homozygotes, the embryo is not viable, and dies before birth. The ways to isolate, and test for candidacy of genes has been previously discussed. However, once a gene(s) has been found that is associated with the 1Gso translocation, many questions can begin to be answered. What type of gene(s) is actually involved in the phenotype? How exactly is it altered in the mutant animals? Is the gene responsible for both the embryonic lethality and the neurological defects? What developmental pathways are disrupted, and how are they related to the generation of the 1Gso phenotype? What specific cell types are being affected? Is the gene a secreted, membrane bound or cytoplasmic factor? Does it belong to a known family of genes, such as the dopamine receptor or neurotrophin family? These are the next research questions that will be addressed to begin to understand the function of the gene(s) associated with 1Gso translocation mutant mice.

Future applications of translocation cloning in the context of the completed Genome project

The type of research that was initiated during the course of the translocation studies will be greatly facilitated by the sequencing of the human and mouse genomes. The region of a translocation can be isolated much more quickly with the resources that will be available in the near future. Once a translocation has been localized by microscopy to a set of chromosomes, databases will be used to locate a series of BAC clones that are separated by 5 Mb surrounding the region of interest. These BACs can then be used in

FISH to bracket the translocation to a defined interval. Then, another series of internal BACs can be chosen to delineate the breakpoint to less that 1 Mb and so on. Once the region is small, within 100-500 kb, the sequence can be accessed to look for candidate genes. These genes can be chosen from the EST databases, and screened to see if they are rearranged on genomic DNA or if they are incorrectly expressed. This could isolate a candidate gene very quickly. Also, once a small region has been shown to be disrupted, locating the translocation is simplified as well. The genomic sequence can be exploited to generate a series of PCR primers, which would amplify the region in both wild type animals and the mutant animals. Any series of primers that produce a correct product in the wild type, but an incorrect or no product in the mutant animal would locate the translocation to a few hundred base pairs. This would then be used to sequence across the breakpoint in the mutant animals. Therefore, the next few years of research within the human genome project will greatly facilitate this type of research. It will make cloning of a mutation more a beginning to understanding the biology of the gene, than a means to isolate the gene itself.

Final summation

These research projects utilized comparative mapping of mouse and human homology regions to both further our understanding of genome evolution and apply these relationships to isolate a translocation breakpoint in the mouse. The establishment of the comparative maps of human chromosome 19 and related regions of the mouse genome revealed the similarities and differences that have arisen in chromosome structure during mammalian evolution. One region in particular has been shown to be prone to evolutionary rearrangements, and its homology borders are associated with clustered gene families. It will be of great interest to further elucidate the nature of these rearrangements and to see if the clustered gene families cause the region to be

destabilized and are a reagent for non-homologous inter- and intrachromosomal recombination to occur.

The positional cloning study of the 1Gso translocation mutant relied upon a backcross map on mouse chromosome 2 as a guide for a fluorescence *in situ* hybridization (FISH) strategy to localize the breakpoint to a small interval. A contig was built across the small breakpoint interval, and FISH was then employed to bracket the breakpoint within a single clone. This clone will form the basis for further characterization of the breakpoint, with the ultimate goal being the isolation of a gene(s) responsible for the 1Gso phenotype.

REFERENCES

Accili, D., Fishburn, C.S., Drago, J., Steiner, H., Lachowicz, J.E., Park, B.H., Gauda, E.B., Lee, E.J., Cool, M.H., Sibley, D.R., Gerfen, C.R., Westphal, H., and Fuchs, S. (1996). A targeted mutation of the D3 dopamine receptor gene is associated with hyperactivity in mice. *Proc Natl Acad Sci* **93**:1945-1949.

Amores, A., Force, A., Yan, Y.L., Joly, L., Amemiya, C., Fritz, A., Ho, R.K., Langeland, J., Prince, V., Wang, Y.L., Westerfield, M., Ekker, M., and Postlethwait, J.H. (1998). Zebrafish hox clusters and vertebrate genome evolution. *Science* **282**:1711-1714.

Antoch, M.P., Song, E.J., Chang, A.M., Vitaterna, M.H., Zhao, Y., Wilsbacher, L.D., Sangoram, A.M., King, D.P., Pinto, L.H., and Takahashi, J.S. (1997). Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. *Cell* **89**:655-667.

Ashworth, L.K., Batzer, M.A., Brandriff, B., Branscomb, E., deJong, P., Garcia, E., Garnes, J., Gordon, L., Lamerdin, J.E., Lennon, G., Mohrenweiser, H., Olsen, A., Slezak, T., and Carrano, A.V. (1995). A metric physical map of human chromosome 19. *Nature Genet* **11**:422-427.

Bailey, L.C. Jr, Fischer, S., Schug, J., Crabtree, J., Gibson, M., and Overton, G.C. (1998). GAIA: framework annotation of genomic sequence. *Genome Res* **8**:234-250.

Bennicelli, J.L., Edwards, R.H., and Barr, F.G. (1996). Mechanism for transcriptional gain of function resulting from chromosomal translocation in alveolar rhabdomyosarcoma. *Proc Natl Acad Sci* **93**:5455-5459.

Blattner, F.R., Plunkett, G. 3rd., Bloch, C.A., Perna, N.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, H.A., Goeden, M.A., Rose, D.J., Mau, B., and Shao, Y. (1997). The complete genome sequence of Escherichia coli K-12. *Science* **277**:1453-1474.

Brand-Arpon, V., Rouquier, S., Massa, H., de Jong, P.J., Ferraz, C., Ioannou, P.A., Demaille, J.G., Trask, B.J., and Giorgi, D. (1999). A genomic region encompassing a cluster of olfactory receptor genes and a myosin light chain kinase (MYLK) gene is duplicated on human chromosome regions 3q13-q21 and 3p13. *Genomics* **56**:98-110.

Bultman, S.J., Michaud, E.J., and Woychik, R.P. (1992). Molecular characterization of the mouse agouti locus. *Cell* **71**:1195-204.

Busslinger, M., Klix, N., Pfeffer, P., Graninger, P.G., and Kozmik, Z. (1996). Deregulation of PAX-5 by translocation of the Emu enhancer of the IgH locus adjacent to two alternative PAX-5 promoters in a diffuse large-cell lymphoma. *Proc Natl Acad Sci* **93**:6129-6134.

Carlton, M.B, Colledge, W.H., and Evans, M.J. (1995). Generation of a pseudogene during retroviral infection. *Mammal Genome* **6**:90-95.

Chen, K.S., Manian, P., Koeuth, T., Potocki, L., Zhao, Q., Chinault, A.C., Lee, C.C., and Lupski, J.R. (1997). Homologous recombination of a flanking repeat gene cluster is a mechanism for a common contiguous gene deletion syndrome. *Nature Genet* **17**:154-163.

Chen, W., Iida, S., Louie, D.C., Dalla-Favera, R., and Chaganti, R.S. (1998). Heterologous promoters fused to BCL6 by chromosomal translocations affecting band 3q27 cause its deregulated expression during B-cell differentiation. *Blood* **91**:603-607.

Collins, F.S., Patrinos, A., Jordan, E., Chakravarti, A., Gesteland, R., Walters, L. and the members of the DOE and NIH planning groups. (1998) New goals for the U.S. Human Genome Project: 1998-2003. *Science* **282**:682-689.

Conover, J.C., Erickson, J.T., Katz, D.M., Bianchi, L.M., Poueymirou, W.T., McClain, J., Pan, L., Helgren, M., Ip, N.Y., Boland, P., Friedman, B., Wiegand, S., Vejsada, R., Kato, A.C., DeChiara, T.M., and Yancopoulos, G.D. (1995). Neuronal deficits, not involving motor neurons, in mice lacking BDNF and/or NT4. *Nature* **375**:235-238.

Corral, J., Forster, A., Thompson, S., Lampert, F., Kaneko, Y., Slater, R., Kroes, W.G., van der Schoot, C.E., Ludwig, W.D., Karpas, A., Pocock, C., Cotter, F., and Rabbitts, T.H. (1993). Acute leukemias of different lineages have similar MLL gene fusions encoding related chimeric proteins resulting from chromosomal translocation. *Proc Natl Acad Sci* **90**:8538-8542

Culiat, C.T., Stubbs, L., Nicholls, R.D., Montgomery, C.S., Russell, L.B., Johnson, D.K., and Rinchik, E.M. (1993). Concordance between isolated cleft palate in mice and

alterations within a region including the gene encoding the beta 3 subunit of the type A gamma-aminobutyric acid receptor. *Proc Natl Acad Sci* **90**:5105-5109.

Culiat, C.T., Stubbs, L.J., Woychik, R.P., Russell, L.B., Johnson, D.K., and Rinchik, E.M. (1995). Deficiency of the beta 3 subunit of the type A gamma-aminobutyric acid receptor causes cleft palate in mice. *Nature Genet* **11**:344-346.

Doyle, J., Ren, X., Lennon, G., and Stubbs, L. (1997). Mutations in the *Cacnl1a4* calcium channel gene are associated with seizures, cerebellar degeneration, and ataxia in tottering and leaner mutant mice. *Mammal Genome* **8**:113-120.

Dutly, F., and Schinzel, A. (1996). Unequal interchromosomal rearrangements may result in elastin gene deletions causing the Williams-Beuren syndrome. *Hum Mol Genet* **5**: 1893-1898.

Ernfors, P., Lee, K.F., and Jaenisch, R. (1994). Mice lacking brain-derived neurotrophic factor develop with sensory deficits. *Nature* **368**:147-150.

Gao, J.L., Chen, H., Filie, J.D., Kozak, C.A., and Murphy, P.M. (1998). Differential expansion of the N-formylpeptide receptor gene cluster in human and mouse. *Genomics* **51**:270-276.

Gessler, M., Simola, K.O.J., and Bruns, G.A.P. (1989). Cloning of breakpoints of a translocation identifies the AN2 locus. *Science* **244**:1575-1578.

Glusman, G., Clifton, S., Roe, B., and Lancet, D. (1996). Sequence analysis in the olfactory receptor gene cluster on human chromosome 17: recombinatorial events affecting receptor diversity. *Genomics* **37**:147-160.

Goffeau, A., Barrell, B.G., Bussey, H., Davis, R.W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J.D., Jacq, C., Johnston, M., Louis, E.J., Mewes, H.W., Murakami, Y., Philippsen, P., Tettelin, H., and Oliver, S.G. (1996). Life with 6000 genes. *Science* **274**:563-567

Glaser, T., Jepeal, L., Edwards, J.G., Young, S.R., Favor, J., and Maas, R.L. (1994). PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nature Genet* **7**:463-471.

Graybiel, A.M. (1990). Neurotransmitters and neuromodulators in the basal ganglia. *Trends Neurosci* **13**:244-254.

Higgs, D.R., Old, J.M., Pressley, L., Clegg, J.B., and Weatherall, D.J. (1980). A novel alpha-globin gene arrangement in man. *Nature* **284**:632-635.

Holland, P.W., and Garcia-Fernandez, J. (1996). Hox genes and chordate evolution. *Dev Biol* **173**:382-395.

Justice, M.J., Zheng, B., Woychik, R.P., and Bradley, A. (1997). Using targeted large deletions and high-efficiency N-ethyl-N-nitrosourea mutagenesis for functional analyses of the mammalian genome. *Methods: A Companion to Methods in Enzymology* **13**:423-436.

Kehrer-Sawatzki, H., Haussler, J., Krone, W., Bode, H., Jenne, D.E., Mehnert, K.U., Tummers, U., and Assum, G. (1997). The second case of a t(17;22) in a family with neurofibromatosis type 1: sequence analysis of the breakpoint regions. *Hum Genet* **99**:237-247.

Kent, J., Lee, M., Schedl, A., Boyle, S., Fantes, J., Powell, M., Rushmere, N., Abbott, C., van Heyningen, V., and Bickmore, W.A. (1997). The reticulocalbin gene maps to the WAGR region in human and to the Small eye Harwell deletion in mouse. *Genomics* **42**:260-267.

Klein, R. (1994). Role of neurotrophins in mouse neuronal development. *FASEB J* **8**:738-744.

Koop, B.F., and Hood, L. (1994). Striking sequence similarity over almost 100 kilobases of human and mouse T-cell receptor DNA. *Nature Genet* **7**:48-53.

Kurahashi, H., Sakamoto, M., Ono, J., Honda, A., Okada, S., and Nakamura, Y. (1998). Molecular cloning of the chromosomal breakpoint in the LIS1 gene of a patient with isolated lissencephaly and balanced t(8;17). *Hum Genet* **103**:189-192.

Lamerdin, J.E., Montgomery, M.A., Stilwagen, S.A., Scheidecker, L.K., Tebbs, R.S., Brookman, K.W., Thompson, L.H., and Carrano, A.V. (1995). Genomic sequence comparison of the human and mouse XRCC1 DNA repair gene regions. *Genomics* **25**:547-554.

Lamerdin, J.E., Stilwagen, S.A., Ramirez, M.H., Stubbs, L., and Carrano, A.V. (1996). Sequence analysis of the ERCC2 gene regions in human, mouse, and hamster reveals three linked genes. *Genomics* **34**:399-409.

Lindvall, O., and Bjorland, A. (1983). Dopamine and norepinephrine-containing neuron systems: their anatomy in the rat brain. In Chemical Neuroanatomy, P.C. Emson, ed. New York; Raven Press. pp.229-255.

Lupski, J.R. (1998). Genomic disorders: structural features of the genome can lead to DNA rearrangements and human disease traits. *Trends Genet* **14**:417-422.

Lyon, M. F., (1953). Absence of otoliths in the mouse: An effect of the pallid mutant. *J Genet* **51**:638-650.

Megonigal, M.D., Rappaport, E.F., Jones, D.H., Williams, T.M., Lovett, B.D., Kelly, K.M., Lerou, P.H., Moulton, T., Budarf, M.L., and Felix, C.A. (1998). t(11;22)(q23;q11.2) in acute myeloid leukemia of infant twins fuses MLL with hCDCrel, a cell division cycle gene in the genomic region of deletion in DiGeorge and velocardiofacial syndromes. *Proc Natl Acad Sci* **95**:6413-6418.

Mercer, J.F.B., J. Livingston, B. Hall, J.A. Paynter, C. Begy, S. Chandrasekharappa, P. Lockhart, A. Grimes, M. Bhave, D. Siemieniak, and T.W. Glover (1993). Isolation of a partial candidate gene for Menkes disease by positional cloning. *Nature Genet* **3**:20-25.

Miles, C., Elgar, G., Coles, E., Kleinjan, D.J., van Heyningen, V., and Hastie, N. (1998). Complete sequencing of the Fugu WAGR region from WT1 to PAX6: dramatic

compaction and conservation of synteny with human chromosome 11p13. *Proc Natl Acad Sci* **95**:13068-13072.

Morrison, A.M., Jager, U., Chott, A., Schebesta, M., Haas, O.A., and Busslinger. M. (1998a). Deregulated PAX-5 transcription from a translocated IgH promoter in marginal zone lymphoma. *Blood* **92**:3865-3878.

Morrison, A.M., Nutt, S.L., Thevenin, C., Rolink, A., and Busslinger, M. (1998b). Loss- and gain-of-function mutations reveal an important role of BSAP (Pax-5) at the start and end of B cell differentiation. *Semin Immunol* **10**:133-142.

Niederfuhr, A., Hummerich, H., Gawin, B., Boyle, S., Little, P.F., and Gessler, M. (1998). A sequence-ready 3-Mb PAC contig covering 16 breakpoints of the Wilms tumor/anirida region of human chromosome 11p13. *Genomics* **53**:155-163.

Oakey, R.J., Watson, M.L., and Seldin, M.F. (1992). Construction of a physical map on mouse and human chromosome 1: comparison of 13 Mb of mouse and 11 Mb of human DNA. *Hum Mol Genet* **1**:613-620.

Oeltjen, J.C., Malley, T.M., Muzny, D.M., Miller, W., Gibbs, R.A., and Belmont, J.W. (1997) Large-scale comparative sequence analysis of the human and murine Bruton's tyrosine kinase loci reveals conserved regulatory domains. *Genome Res* **7**:315-329.

Okladnova, O., Syagailo, Y.V., Mossner, R., Riederer, P., and Lesch, K.P. (1998). Regulation of PAX-6 gene transcription: alternate promoter usage in human brain. *Brain Res Mol Brain Res* **60**:177-192.

Ophoff, R.A., Terwindt, G.M., Vergouwe, M.N., van Eijk, R., Oefner, P.J., Hoffman, S.M., Lamerdin, J.E., Mohrenweiser, H.W., Bulman, D.E., Ferrari, M., Haan, J., Lindhout, D., van Ommen, G.J., Hofker, M.H., Ferrari, M.D., and Frants, R.R. (1996). Familial hemiplegic migraine and episodic ataxia type-2 are caused by mutations in the Ca2+ channel gene CACNL1A4. *Cell* **87**:543-552.

Procter, A.M., Phillips, J.A. 3rd, and Cooper, D.N. (1998). The molecular genetics of growth hormone deficiency. *Hum Genet* **103**:255-272.

Rabbitts, T H. (1994). Chromosomal translocation in human cancer. *Nature* **372**:143-149.

Riccardi, V.M., Sujansky, E., Smith, A.C., and Francke, U. (1978). Chromosomal imbalance in the Aniridia-Wilms' tumor association: 11p interstitial deletion. *Pediatrics* **61**:604-610.

Ridinger, K., Ilg, E.C., Niggli, F.K., Heizmann, C.W., and Schafer, B.W. (1998). Clustered organization of S100 genes in human and mouse. *Biochim Biophys Acta* **1448**:254-263.

Rouquier, S., Taviaux, S., Trask, B.J., Brand-Arpon, V., van den Engh, G., Demaille, J., and Giorgi, D. (1998). Distribution of olfactory receptor genes in the human genome. *Nature Genet* **18**:243-250.

Rowley, J.D. (1998). The critical role of chromosome translocations in human leukemias. *Annu Rev Genet* **32**:495-519.

Rubinstein, M., Phillips, T.J., Bunzow, J.R., Falzone, T.L., Dziewczapolski, G., Zhang, G., Fang, Y., Larson, J.L., McDougall, J.A., Chester, J.A., Saez, C., Pugsley, T.A., Gershanik, O., Low, M.J., and Grandy, D.K. (1997). Mice lacking dopamine D4 receptors are supersensitive to ethanol, cocaine, and methamphetamine. *Cell* **90**:991-1001.

Seldin, M.F. (1997). Genome surfing: using internet-based informatic tools toward functional genetic studies in mouse and humans. *Methods: A Companion to Methods in Enzymology* **13**:445-457.

Shannon, M., Ashworth, L.K., Mucenski, M.L., Lamerdin, J.E., Branscomb, E., and Stubbs, L. (1996). Comparative analysis of a conserved zinc finger gene cluster on human chromosome 19q and mouse chromosome 7. *Genomics* **33**:112-120.

Singh, S., Tang, H.K., Lee, J.Y., and Saunders, G.F. (1998). Truncation mutations in the transactivation region of PAX6 result in dominant-negative mutants. *J Biol Chem* **273**:21531-21541.

Stubbs, L., Carver, E.A., Shannon, M.E., Kim, J., Geisler, J., Generoso, E.E., Stanford, B.G., Dunn, W.C., Mohrenweiser, H., Zimmermann, W., Watt, S.M., and Ashworth L.K. (1996). Detailed comparative map of human chromosome 19q and related regions of the mouse genome. *Genomics* **35**:499-508.

Super, H.G., Strissel, P.L., Sobulo, O.M., Burian, D., Reshmi, S.C., Roe, B., Zeleznik-Le, N.J., Diaz, M.O., and Rowley, J.D. (1997). Identification of complex genomic breakpoint junctions in the t(9;11) MLL-AF9 fusion gene in acute leukemia. *Genes Chrom Cancer* **20**:185-195.

The C. elegans Sequencing Consortium. (1998) Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282**:2012-2018.

Thomas, S.A., and Palmiter, R.D. (1997). Disruption of the dopamine beta-hydroxylase gene in mice suggests roles for norepinephrine in motor function, learning, and memory. *Behav Neurosci* **111**:579-589.

Trask, B.J., Massa, H., Brand-Arpon, V., Chan, K., Friedman, C., Nguyen, O.T., Eichler, E., van den Engh, G., Rouquier, S., Shizuya, H., and Giorgi, D. (1998). Large multi-chromosomal duplications encompass many members of the olfactory receptor gene family in the human genome. *Hum Mol Genet* **7**:2007-2020.

Trune, D.R., and Lim, D.J. (1983). The behavior and vestibular nuclear morphology of otoconia-deficient pallid mutant mice. *J Neurogenet* **1**:53-69.

van Deutekom, J.C., Bakker, E., Lemmers, R.J., van der Wielen, M.J., Bik, E., Hofker, M.H., Padberg, G.W., and Frants, R.R. (1996). Evidence for subtelomeric exchange of 3.3 kb tandemly repeated units between chromosomes 4q35 and 10q26: implications for genetic counseling and etiology of FSHD1. *Hum Mol Genet* **5**:1997-2003.

Vnencak-Jones, C.L., and Phillips, J.A. 3d. (1990). Hot spots for growth hormone gene deletions in homologous regions outside of Alu repeats. *Science* **250**:1745-1748.

Weihmuller, F.B., Hadjiconstantinou, M., and Bruno, J.P. (1990). Dopamine receptors and sensorimotor behavior in MPTP-treated mice. *Behav Brain Res* **38**:263-73.

Weterman, M.A., Wilbrink, M., Dijkhuizen, T., van den Berg, E., and Geurts van Kessel, A. (1996). Fine mapping of the 1q21 breakpoint of the papillary renal cell carcinoma-associated (X;1) translocation. *Hum Genet* **98**:16-21.

Weterman, M.A., Wilbrink, M., and Geurts van Kessel, A. (1996). Fusion of the transcription factor TFE3 gene to a novel gene, PRCC, in t(X;1)(p11;q21)-positive papillary renal cell carcinomas. *Proc Natl Acad Sci* **93**:15294-15298.

Wilson, D.I., Cross, I.E., Goodship, J.A., Coulthard, S., Carey, A.H., Scambler, P.J., Bain, H.H., Hunter, A.S., Carter, P.E., and Burn, J. (1991). DiGeorge syndrome with isolated aortic coarctation and isolated ventricular septal defect in three sibs with a 22q11 deletion of maternal origin. *Br Heart J* **66**:308-312.

Xu, M., Moratalla, R., Gold, L.H., Hiroi, N., Koob, G.F., Graybiel, A.M., and Tonegawa, S. (1994). Dopamine D1 receptor mutant mice are deficient in striatal expression of dynorphin and in dopamine-mediated behavioral responses. *Cell* **79**:729-42.

You, Y., Browning, V.L., and Schimenti, J.C. (1997). Generation of radiation-induced deletion complexes in the mouse genome using embryonic stem cells. *Methods: A Companion to Methods in Enzymology* **13**:409-421.

Appendix

FISH Protocol

(Combined and revised by EAC from Stratagene Prime-It Fluor Fluorescence Labeling Kit and Chromosomal *In Situ* Hybridization Kit)

DAY ONE

Place slides in 65°C oven four hours to overnight prior to first day of FISH.
(This bakes material to slide, preventing cell loss during denaturation and washing steps.)

Turn on waterbaths to 37°C, 42°C and 70°C .
Place petri dish with moist towels into 37°C incubator.

Will have already checked slides for number and quality of metaphase cells, and scored a good region of slide. (*Can start to make probes here*)
Make 50 mls of 2xSSC by adding 45 mls of dH20 and 5mls of 20xSSC into a Coplin jar.
Prepare 2 ml of 2xSSCbuffer and add 100ul of RNA solution to the buffer.
Spot each slide with 150ul of RNase in 2x SSC, cover with a large coverslip and incubate at 37°C for 1 hour.
Rinse the slides in 2xSSC and place directly into 80%, 90%, and then 100% EtOH, for 5 minutes each and air-dry slides. The slides are now ready to denature.

Prepare probe: (Dim/dark lighting as possible)

100-200ng of probe per slide, can be YACs, BACs, PACs, P1s, or cosmids.

Prepare the reaction buffer by mixing 8ul of fluor-12-dUTP with 92ul of 5x nucleotide buffer in a sterile microcentrifuge tube. This mixture is stable for up to one month at -20°C.

Mix the probe up as follows:

100-200 ng (1-29ul) of DNA template
0-28ul of distilled water
10ul of random 9-mer primers
to a total reaction volume of 39 ul.

Heat the sample to 95-100°C for five minutes in a heating block.
Briefly centrifuge the samples at room temperature and place reaction tubes on ice.

Add the following reagents to the sample tubes:
10ul of reaction buffer
1ul (5U) of exonuclease-free Klenow
mix with the pipet tip.
Spin down samples

Incubate the sample tubes at 37°C for 30 minutes.

Add 2ul of the Stop Mix to each reaction tube, and store in 4°C in the dark.

During probe preparation you can:
Set up a Coplin jar for chromosome denaturation:

35ml deionized formamide
5ml 20X SSC
10ml ddH_2O
Place in 70°C waterbath

Place 3 Coplin jars with EtOH into an ice bucket

80% EtOH
90% EtOH
100% EtOH
100% EtOH at room temperature

Turn on 42°C slide warmer

Precipitation of probes.

Precipitate probes by addition of:

5ul of 5M NaOAc
5ul of mouse Cot-1 DNA
5ul of sheared salmon sperm DNA
120ul of 100% EtOH

Combine probes that will be hybridized on the same slide.
Incubate on ice for 30 minutes
Spin down in microcentrifuge for 30 minutes
Decant supernate and lightly dry.
Add 30ul of hybridization buffer and resuspend gently.

Denature chromosomes:

Place slides in denaturation solution at 70°C* for 3 min.
Remove and place directly into ice cold EtOH
80% 2 minutes , 90% 2 minutes , 100% 1 minute , then room temperature 100% EtOH, 1 minute .
Air-dry slides.

*Note-Be sure denaturation solution(not just waterbath) is at 70°C. This step is critical and the length of time must be determined empirically for various samples and methods of slide preparation. For our method of preparing mitotic metaphase cell from spleen, 3 min. works well.

Turn 70°C waterbath up to 75°C

Denature probe:

Incubate probes 75°C for 15 min.
Incubate probes 42°C for 30 min.

Hybridization of probes to slides
Place air-dried slides on 42°C slide warmer.
Add all of probe.
Cover with 22x22mm coverslip eliminate as many air bubbles as possible.
Seal edges with rubber cement (syringe with large gauge needle).
Place slides into covered petri dish with moist towels in 37°C overnight.

DAY TWO

Conduct remainder of protocol in the DARK/dim lighting as much as possible

Turn on 45° and 60°C waterbaths

Make wash solutions.

Wash solution A
50 mls of formamide
10 mls of 20xSSC
40 mls of distilled water

Wash solution B
0.5 mls of 20xSSC
95.5 mls of distilled water

Wash solution C
0.5 mls of Tween 20
10 mls of 20xSSC
39.5 mls of distilled water

Fill 3 Coplin jars with Wash A, 3 for Wash B, 1 for Wash C

Place jars with Wash 1 into 45°C waterbath
Place jars with Wash 2 into 60°C waterbath

After washes have equilibrated, take slides from incubator and remove rubber cement without disturbing coverslip

VERY important to never let the slides dry out.

WASH A

Place in first jar of Wash 1 for 5 min. with moderate agitation. Coverslips will slide off in first jar of Wash A.
Transfer to consecutive Wash 1s for 5 min. each.

WASH B

After last Wash 1. Transfer to Wash 2 in 60°C waterbath.
3X 5 min. each.

WASH C

Rinse each slide with slight agitation.

Drain slides well(do not let dry!)
Add 30ul Antifade with PI, do not leave open longer than necessary, sensitive to oxidation and to light.
Cover with 22x40mm coverslip.

Place in light-proof box.
View results with epifluorescence microscope.
Store slides at 4°C in the dark (several months).
Or in -20°C or -80°C for longer periods (wrap boxes and store with desiccant).

Vita

Ethan Allan Carver, the youngest of two sons of Audrey and Glenn Carver, was born on September 6, 1968 in Bryson City, North Carolina. He attended public schools in the Rhea County, Tennessee system, and graduated from Rhea County High School in June of 1986. He enrolled at The University of Tennessee at Chattanooga, Chattanooga, Tennessee in 1986, and was partially supported for three years by a music performance scholarship. In May, 1991, he obtained a Bachelor of Science degree in Biology. In August, 1992, he began pursuing a Doctor of Philosophy degree at The University of Tennessee Graduate School of Biomedical Sciences at Oak Ridge. His research, under the mentorship of Dr. Lisa Stubbs began at the Oak Ridge National Laboratory. In the fall of 1997, Dr. Stubbs moved to Lawrence Livermore National Laboratory, and he completed his research there in the spring of 1999. He was awarded an NCI training grant and will begin postdoctoral research work focused on mouse genes related to the *Drosophila* snail genes, and their role in development with Dr. Thomas Gridley at The Jackson Laboratory in Bar Harbor, Maine.