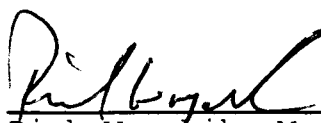


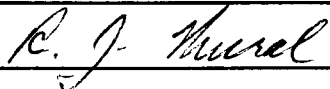
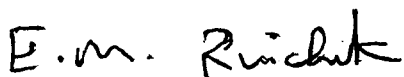
To the Graduate Council:

I am submitting herewith a dissertation written by Scott J. Bultman entitled "MOLECULAR CHARACTERIZATION OF THE MOUSE AGOUTI LOCUS." 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.



Rick Woychik, Major Professor

We have read this dissertation
and recommend its acceptance:



Accepted for the Council:



Associate Vice Chancellor
and Dean of the Graduate School

MOLECULAR CHARACTERIZATION OF THE MOUSE AGOUTI LOCUS

A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Scott Justin Bultman

May 1994

Acknowledgements

I would like to acknowledge a number of people who helped with my graduate school training. I am indebted to my mentor, Dr. Rick Woychik, for teaching me a variety of molecular biology techniques, and more importantly, how to design well controlled experiments that address complex biological questions. Rick's guidance and insight certainly played an important role in the success of this work. I am also grateful to the other members of my committee, Drs. Richard Mural, Gene Rinchik, and Lee Russell, for their advice and constructive criticism.

I thank my colleagues in the Woychik lab, particularly Dr. Ed Michaud, for numerous discussions and helpful suggestions. The fact that Ed and I have been able to work synergistically was paramount to the success of the agouti-locus project. I would also like to recognize another member of the Woychik lab, Jeff Schrick, for teaching me several computer- and photographic-related techniques.

Much of the work presented in this dissertation would have not been possible without the foresight of Drs. Bill and Lee Russell who established the Mammalian Genetics section at the Oak Ridge National Laboratory fifty years ago. Bill and Lee along with their past and present colleagues, particularly Pat Hunsicker and Clyde Montgomery, generated many important agouti-locus mutants and made the mice available to me for molecular analyses. I would also like to acknowledge Dr. Waldy Generoso for providing perhaps the most important agouti-locus mutation, Is1Gso, which enabled the agouti gene to be cloned.

It is also a pleasure to thank my wife, Laura, and my parents, J.B. and Judy, for their strong support and understanding during my Ph.D. training.

Abstract

The agouti locus in mouse chromosome 2 regulates the differential production of eumelanin (black or brown) and phaeomelanin (yellow) pigment granules. Agouti coloration, which is the true wild-type coat color of mice, arises because the individual hair shafts are black or brown with a subapical yellow band. Over 20 distinct alleles have been identified that form an intricate dominance hierarchy and account for an array of different phenotypes. Unlike most coat-color genes, which act in a cell-autonomous manner within the melanocyte, the agouti locus functions from within the microenvironment of the hair follicle, suggesting that the agouti gene product participates in an intercellular signaling pathway. The objective of this dissertation was to learn more about agouti-locus function by analyzing it at the molecular level.

I utilized a radiation-induced mutation to gain molecular access to the mouse agouti locus. Characterization of genomic and cDNA clones revealed that agouti is a large, complex locus that consists of a single gene that spans over 100 kb of DNA and encodes a novel 131 amino acid protein. The agouti gene is normally expressed in skin, and more importantly, its structure and expression are disrupted in a number of mutant alleles. Furthermore, this gene produces several alternately processed transcripts that differ in their 5' untranslated regions but splice to the same coding sequence. These different forms of mRNA are differentially expressed within the skin, particularly over the dorsoventral surface of the animal. A model is presented that describes how the agouti protein, which appears to be secreted, might signal melanocytes to produce eumelanin or phaeomelanin. Additionally, ectopic overexpression of the agouti gene in several dominant mutants is responsible for a pleiotropic syndrome that includes an all-yellow coat, diabetes, obesity, and an increased susceptibility to tumors. Finally, using a cross-species hybridization approach, I was able to isolate the human agouti gene. It shares considerable sequence homology with the mouse gene and maps to chromosome 20q11; its potential role in human hair pigmentation and diabetes is discussed.

TABLE OF CONTENTS

SECTION	PAGE
PART I: INTRODUCTION	
BACKGROUND AND SIGNIFICANCE.....	2
OBJECTIVES AND OVERVIEW.....	6
REFERENCES.....	8
PART II: MOLECULAR AND GENETIC CHARACTERIZATION OF A RADIATION-INDUCED STRUCTURAL REARRANGEMENT IN MOUSE CHROMOSOME 2 CAUSING MUTATIONS AT THE <i>LIMB</i> <i>DEFORMITY</i> AND <i>AGOUTI</i> LOCI	
ABSTRACT.....	19
INTRODUCTION.....	19
MATERIALS AND METHODS.....	21
Mutagenesis and Genetic and Cytogenetic Analysis.....	21
Northern and Southern Blot Analysis.....	21
Cloning the DNA Breakpoint Within the <i>ld</i> Locus.....	21
RESULTS.....	22
Mutant Recovery and Genetic Analysis.....	22
Cytogenetic Analysis Reveals a Chromosomal Rearrangement.....	23
Altered Gene Expression at the <i>ld</i> Locus.....	24
DNA Breakpoint in the <i>ld</i> Locus.....	25
DISCUSSION.....	26

SECTION	PAGE
REFERENCES.....	28
APPENDIXES.....	32

PART III: MOLECULAR CHARACTERIZATION OF A REGION OF DNA ASSOCIATED WITH MUTATIONS AT THE AGOUTI LOCUS IN THE MOUSE

ABSTRACT.....	45
INTRODUCTION.....	45
MATERIALS AND METHODS.....	47
Animals.....	47
Southern Blotting.....	47
Cloning of a 5.5-kb Wild-Type Fragment from a Region Corresponding to the Distal Inversion Breakpoint of Is1Gso.....	47
Cloning a Rearranged Fragment Associated with the α^{5MNU} Mutation.....	47
Cloning a Large Region of DNA at or Near the Agouti Locus.....	48
RESULTS.....	48
Identification of a DNA Probe Deleted in Several Radiation- or Chemical-Induced α -Locus Mutations.....	48
Characterization of the Structure of the Distal Inversion Breakpoint in Is1Gso.....	50
Characterization of a Structural Alteration in the α^{5MNU} Viable α -Locus mutation.....	51
DISCUSSION.....	52

SECTION	PAGE
REFERENCES.....	55
APPENDIXES.....	61

PART IV: MOLECULAR CHARACTERIZATION OF THE MOUSE AGOUTI LOCUS

ABSTRACT.....	76
INTRODUCTION.....	76
MATERIALS AND METHODS.....	78
Animals.....	78
DNA Isolation and Southern Blot Analysis.....	78
RNA Preparation and Northern Blot Analysis.....	78
Isolation of cDNA.....	79
Isolation of Genomic Clones.....	79
DNA Sequencing.....	79
RESULTS.....	80
Isolation of a cDNA Clone and Genomic Structure of the Agouti Locus.....	80
Wild-Type Pattern of Expression.....	81
Changes in Gene Structure and Expression in Several Agouti Locus Mutations.....	82
Deregulated Expression in Lethal Yellow Heterozygotes.....	83
DISCUSSION.....	84
Characterization of the Agouti Locus Gene.....	84
Hypothetical Interaction with the Extension Locus.....	86
Deregulated Expression in Lethal Yellow Heterozygotes.....	87
Agouti is a Single Gene Locus.....	89

SECTION	PAGE
REFERENCES.....	91
APPENDIX.....	100

**PART V: MOLECULAR ANALYSIS OF REVERSE MUTATIONS
FROM NONAGOUTI (α) TO BLACK-AND-TAN (α^t) AND WHITE-
BELLIED AGOUTI (A^w) REVEALS ALTERNATIVE FORMS OF
AGOUTI TRANSCRIPTS**

ABSTRACT.....	120
INTRODUCTION.....	121
MATERIALS AND METHODS.....	122
Animals.....	122
Cloning Overlapping Genomic DNA Fragments that Span the Insertions in α , α^t , and A^w	123
DNA Sequencing.....	123
Southern Blot Analysis.....	124
RT-PCR Analysis.....	124
RESULTS.....	124
Structural DNA Rearrangements Associated with the α , α^t , and A^w alleles.....	124
Molecular Analysis of Several New α^t and A^w Reverse Mutations.....	126
Additional Forms of Agouti Transcripts.....	128
DISCUSSION.....	129
DNA Insertions in α , α^t -2Gso, and A^w -38J.....	129

SECTION	PAGE
Reverse Mutations From a to a^t and A^w are Mediated by Homologous Recombination.....	130
Multiple Forms of Agouti mRNA.....	131
The Insertions in a and a^t Modulate the Expression of Form I and Form II Agouti Transcripts.....	132
A^w is Wild Type.....	133
REFERENCES.....	135
APPENDIXES.....	142

**PART VI: THE EMBRYONIC LETHALITY OF HOMOZYGOUS
LETHAL YELLOW MICE (A^y/A^y) IS ASSOCIATED WITH THE
DISRUPTION OF A NOVEL RNA-BINDING PROTEIN**

ABSTRACT.....	159
INTRODUCTION.....	159
MATERIALS AND METHODS.....	162
Animals.....	162
RNA Analysis.....	162
Isolation of cDNA Clones.....	163
Isolation of Genomic Clones.....	163
DNA Sequencing.....	163
RESULTS.....	164
The A^y Transcript is a Chimera Composed of Agouti and a Second Gene.....	164

SECTION	PAGE
Characterization of <i>Raly</i> cDNA Clones.....	164
The <i>Raly</i> Transcription Unit is Alternately Spliced at the 5' End.....	165
<i>Raly</i> and Agouti are Tightly Linked in Chromosome 2.....	166
<i>Raly</i> is Expressed in a Widespread Temporal and Spatial Manner.....	167
The <i>A^y</i> Mutation Disrupts the Expression of the <i>Raly</i> Gene.....	168
The <i>A^y</i> Transcription Unit Forms Unique Transcripts that are Alternately Spliced at the 5' End.....	169
The Embryonic Lethality of <i>A^y/A^y</i> Mice May be the Result of the Disruption of <i>Raly</i> , and Not the Ectopic Overexpression of the Agouti Protein.....	170
DISCUSSION.....	170
REFERENCES.....	176
APPENDIX.....	184

PART VII: A MOLECULAR MODEL FOR THE GENETIC AND PHENOTYPIC CHARACTERISTICS OF THE MOUSE LETHAL YELLOW (*A^y*) MUTATION

ABSTRACT.....	202
INTRODUCTION.....	202
MATERIALS AND METHODS.....	204
Pulsed-Field Gel Analysis.....	204
Southern Blot Analysis.....	204
Probes.....	205

SECTION	PAGE
RESULTS.....	205
Long-Range Physical Map of <i>Raly</i> and Agouti.....	205
Localization of the <i>A^y</i> Deletion Breakpoints.....	206
DISCUSSION.....	208
REFERENCES.....	212
APPENDIXES.....	217

PART VIII: MOLECULAR STRUCTURE AND CHROMOSOMAL MAPPING OF THE HUMAN HOMOLOG OF THE AGOUTI GENE

ABSTRACT.....	230
INTRODUCTION.....	230
MATERIALS AND METHODS.....	232
Southern Blotting.....	232
Isolation of a Human Genomic Clone.....	232
DNA Sequencing.....	232
Chromosomal Assignment.....	233
Mapping by Fluorescence <i>In Situ</i> Hybridization (FISH).....	233
RNA Isolation and RT-PCR.....	234
RESULTS.....	234
Isolation and Characterization of a Human Clone Homologous to the Mouse Agouti Gene.....	234
Mapping and Regional Assignment of the Human Agouti Gene.....	236
Expression of the Agouti Gene in Human Adipose Tissue.....	237

SECTION	PAGE
DISCUSSION.....	237
REFERENCES.....	241
APPENDIX.....	246

PART IX: SUMMARY

CONCLUSIONS.....	257
Gaining Molecular Access to a Region of DNA at the Agouti Locus.....	257
Cloning the Agouti Gene.....	258
Expression of the Agouti Gene.....	259
Agouti is a Single Gene Locus.....	260
The Agouti Gene Produces Multiple Forms of mRNA.....	260
Proposed Function of the Agouti Protein.....	261
Molecular Basis of <i>A^y</i> the Phenotype.....	262
Is Agouti a Human Disease Gene?.....	263
REFERENCES.....	265
VITA.....	271

LIST OF TABLES

TABLE	PAGE
PART II: MOLECULAR AND GENETIC CHARACTERIZATION OF A RADIATION-INDUCED STRUCTURAL REARRANGEMENT IN MOUSE CHROMOSOME 2 CAUSING MUTATIONS AT THE <i>LIMB DEFORMITY</i> AND AGOUTI LOCI	
1. Manifestation of allelism between R and ld^{OR} as expressed by limb morphology of progeny.....	34
PART III: MOLECULAR CHARACTERIZATION OF A REGION OF DNA ASSOCIATED WITH MUTATIONS AT THE AGOUTI LOCUS IN THE MOUSE	
1. Radiation- and chemical-induced agouti-locus mutations that exhibit detectable structural alterations.....	63
PART V: MOLECULAR ANALYSIS OF REVERSE MUTATIONS FROM NONAGOUTI (a) TO BLACK-AND-TAN (a^t) AND WHITE-BELLIED AGOUTI (A^w) REVEALS ALTERNATIVE FORMS OF AGOUTI TRANSCRIPTS	
1. New A^w and a^t mutations analyzed by Southern analysis.....	143
2. Dorsal (D) and ventral (V) distribution of form I and form II transcripts in neonatal skin from mice with different agouti genotypes.....	144
PART VII: A MOLECULAR MODEL FOR THE GENETIC AND PHENOTYPIC CHARACTERISTICS OF THE MOUSE LETHAL YELLOW (A^y) MUTATION	
1. Agouti exceptions from crosses involving lethal yellow (A^y) mice at the Oak Ridge National Laboratory.....	219

LIST OF FIGURES

FIGURE	PAGE
PART II: MOLECULAR AND GENETIC CHARACTERIZATION OF A RADIATION-INDUCED STRUCTURAL REARRANGEMENT IN MOUSE CHROMOSOME 2 CAUSING MUTATIONS AT THE <i>LIMB</i> <i>DEFORMITY</i> AND <i>AGOUTI</i> LOCI	
1. Structure of the rearranged chromosomes 2 and 17 in the radiation mutant.....	36
2. RNA expression from a complex gene located at the <i>ld</i> locus.....	38
3. Identification of rearranged DNA fragments at the <i>ld</i> locus of the radiation-induced mutation.....	40
4. Genetic mapping of the Is1Gso breakpoint sequence.....	42
PART III: MOLECULAR CHARACTERIZATION OF A REGION OF DNA ASSOCIATED WITH MUTATIONS AT THE <i>AGOUTI</i> LOCUS IN THE MOUSE	
1 Southern blot analysis of radiation- or chemical-induced extreme nonagouti mutations.....	65
2. Southern blot of wild-type, heterozygous, or homozygous adult genomic DNA from the Is1Gso mutant line.....	69
3. Southern blot of C3H, 101, or homozygous α^{5MNU} adult genomic DNA digested with either <i>EcoRI</i> , <i>PstI</i> , or <i>BamHI</i>	71
4. Fine structure analysis of the DNA lesion in the α^{5MNU} mutation.....	73

PART IV: MOLECULAR CHARACTERIZATION OF THE MOUSE AGOUTI LOCUS

1. Identification of an exon near the distal inversion breakpoint of Is1Gso...	101
2. Nucleotide and predicted amino acid sequence of the cDNA.....	103
3. Intron/exon structure of the wild-type locus and two extreme nonagouti mutations.....	105
4. Northern blot analysis of wild-type (<i>A/A</i>) neonatal skin.....	107
5. Northern blot analysis of adult tissues from wild-type (<i>A/A</i>) mice.....	109
6. Identification of DNA structural alterations associated with the <i>a</i> and <i>a^t</i> alleles.....	111
7. Northern blot analysis of neonatal skin from several agouti locus mutants.....	113
8. Northern blot analysis of adult tissues from wild-type and lethal yellow heterozygotes.....	115
9. Molecular analysis of the size-altered <i>A^y</i> transcript.....	117

PART VI: MOLECULAR ANALYSIS OF REVERSE MUTATIONS FROM NONAGOUTI (*a*) TO BLACK-AND-TAN (*a^t*) AND WHITE- BELLIED AGOUTI (*A^w*) REVEALS ALTERNATIVE FORMS OF AGOUTI TRANSCRIPTS

1. Identification of DNA insertions in the first intron of the agouti gene in the <i>a</i> , <i>a^t</i> , and <i>A^w</i> alleles.....	146
2. Nucleotide sequence of a region of DNA spanning the insertion site in <i>A</i> , <i>a</i> , <i>a^t-2Gso</i> , and <i>A^{w-38J}</i>	148

3. Nucleotide sequence comparison between the 526-bp direct terminal repeats in the α and α^t insertions and part of an intragenic regulatory region between two TCR γ variable gene segments.....150
4. Southern blot analysis of reverse mutations from α to α^t 152
5. Southern blot analysis of reverse mutations from α to A^w 154
6. Identification of two forms of agouti transcripts by RT-PCR.....156

**PART VI: THE EMBRYONIC LETHALITY OF HOMOZYGOUS
LETHAL YELLOW MICE (A^y/A^y) IS ASSOCIATED WITH THE
DISRUPTION OF A NOVEL RNA-BINDING PROTEIN**

1. The A^y allele encodes a *Raly/a* fusion transcript.....185
2. Nucleotide and predicted amino acid sequence of *Raly*.....187
3. *Raly* is a member of the heterogeneous nuclear ribonucleoprotein particle (hnRNP) protein gene family.....189
4. Southern blot analyses reveal that the A^y mutation deletes the 3' end of *Raly* but does not alter the 5' end of agouti.....191
5. Expression of the wild-type *Raly* gene.....193
6. The expression of *Raly* is disrupted from the A^y allele.....195
7. The A^y transcription unit is alternately processed at the 5' end, giving rise to three separate transcripts.....197
8. Northern blot analysis of agouti locus expression in tissues of adult viable yellow (A^{vy}) mice.....199

FIGURE

PAGE

**PART VII: A MOLECULAR MODEL FOR THE GENETIC AND
PHENOTYPIC CHARACTERISTICS OF THE MOUSE LETHAL
YELLOW (A^y) MUTATION**

1. Pulsed-field gel analysis of the chromosome 2 region containing the *Raly* and agouti loci in wild-type and lethal yellow DNAs.....221
2. Detection of the A^y deletion breakpoints by Southern blot analyses.....223
3. A simple model for recombination between A^y and other agouti-locus alleles.....225
4. A model for the production of chimeric *Raly*/agouti transcripts from the A^y allele.....227

**PART VIII: MOLECULAR STRUCTURE AND CHROMOSOMAL
MAPPING OF THE HUMAN HOMOLOG OF THE AGOUTI GENE**

1. Genomic structure and nucleotide sequence of the human agouti gene.....247
2. Amino acid sequence comparison between the mouse and human agouti genes.....249
3. Physical mapping of the human agouti gene to chromosome 20q11.2.....251
4. Tissue distribution of the human agouti mRNA.....254

PART I
INTRODUCTION

BACKGROUND AND SIGNIFICANCE

Analysis of coat color mutations in the mouse has provided one of the premier systems for studying mammalian gene action and interaction (reviewed in Silvers, 1979). Because coat-color mutants are readily visible, they are particularly amenable to genetic and phenotypic analyses. As a consequence, genetic analyses of coat-color mutants have been conducted for over 100 years and have led to the identification of over 70 loci that control coat pigmentation. Phenotypic analyses of coat-color mutants, including histological examination of wild-type and mutant tissues and numerous *in vitro* and *in vivo* experiments, have complemented the genetic studies by providing information on the cell biology and biochemistry of pigment-producing melanocytes. Additionally, genetic and phenotypic analyses have shown that a number of genes that affect coat color also play crucial roles in developmental processes such as neural development (Epstein et al., 1991; Mercer et al., 1991; Hirata et al., 1993), hematopoiesis (Chabot et al., Geissler et al., 1988; Witte, 1990, and references therein), gametogenesis (Chabot et al., Geissler et al., 1988; Witte, 1990, and references therein), bone resorption (Hodgkinson et al., 1993), eye differentiation (Hodgkinson et al., 1993), and hearing (Hodgkinson et al., 1993).

Interest in mouse coat-color genetics has recently been bolstered by the cloning of a number of coat-color genes. A variety of strategies, including positional cloning (Brilliant et al., 1991; Gardner et al., 1992), examination of previously cloned genes as candidate genes (Chabot et al., 1988; Geissler et al., 1988; Epstein et al., 1991; Rinchik et al., 1993; Robbins et al., 1993), differential hybridization (Shibahara et al., 1986), insertional mutagenesis (Keller et al., 1990; Mercer et al., 1991; Hodgkinson et al., 1993), immunobiological screening (Kwon et al., 1987; Jackson et al., 1988; Kwon et al., 1991; Jackson et al., 1992), exploitation of protein-protein interactions (Flanagan and Leder, 1990), and traditional protein purification procedures (Anderson et al., 1990; Huang et al., 1990; Martin et al., 1990; Williams et al., 1990; Zsebo et al., 1990), have been employed to accomplish this feat. Molecular characterization of these genes revealed that they encode proteins such as membrane-bound and soluble ligands (Witte, 1990, and references

therein), a tyrosine-kinase (Chabot et al., 1988; Geissler et al., 1988) and other types of transmembrane receptors (Gardner et al., 1992; Rinchik et al., 1993; Robbins et al., 1993), intracellular enzymes (Kwon et al., 1987; Jackson, 1988; Jackson et al., 1992; Tsukamoto et al., 1992), and helix-loop-helix-leucine zipper (Hodgkinson et al., 1993) and paired box (Epstein et al., 1991) transcription factors. These types of proteins are of special interest because they play important roles in sundry signal transduction pathways. The papers referenced above, combined with subsequent reports, have provided considerable insight into the molecular basis of coat pigmentation and several other developmentally important processes. Furthermore, most of the coat-color loci that have been identified in the mouse are conserved throughout many mammalian species, including humans (Searle, 1968). As a result, knowledge gained from studying these genes in the mouse can be directly applied to corresponding human conditions. In fact, mutations of several human genes that are homologous to mouse coat-color genes are responsible for inherited diseases in humans that closely resemble the phenotypes observed in mice (Giebel and Spritz, 1991; Giebel et al., 1991; King et al., 1991; Baldwin et al., 1992; Fleischman, 1992; Tassabehji et al., 1992; Oetting et al., 1993; Rinchik et al., 1993; Spritz et al., 1993).

The principal cell type involved in coat pigmentation is the neural crest-derived melanocyte. During embryonic development, melanocyte precursors, called melanoblasts, migrate from the neural tube to the base of developing hair follicles (Billingham and Silvers, 1960; Searle, 1968). The melanoblasts differentiate into melanocytes and essentially function as unicellular glands by secreting melanin granules into actively growing hair shafts (Billingham and Silvers, 1960; Searle, 1968). Importantly, melanocytes produce two distinct types of melanin: eumelanin (black or brown, depending on alleles present at the brown locus and other coat-color loci) and phaeomelanin (yellow).

The agouti locus in chromosome 2 of the mouse regulates the differential production of eumelanin and phaeomelanin. Agouti pigmentation, which is the true wild-type coat color of mice, arises because melanocytes initially produce eumelanin, then switch to phaeomelanin synthesis, and subsequently return to eumelanin production before the hair

growth process is complete (Galbraith, 1964; Sakurai et al., 1975). As a result, the individual hairs of agouti mice are black or brown with a subapical yellow band and give the animal an overall grizzled appearance.

At least 20 dominant and recessive agouti alleles form a complex dominance hierarchy. Alleles associated with increased phaeomelanin production are generally dominant over alleles associated with agouti pigmentation, which in turn, are dominant over alleles primarily associated with eumelanin production. However, the degree of dominance of yellow pigmentation varies in different regions of the pelage. For example the null allele, extreme nonagouti (α^e), confers a completely black phenotype. Proceeding up the dominance hierarchy, phaeomelanin first appears near the extremities in nonagouti (α) while the rest of the pelage remains black. Phaeomelanin is found in the ventrum and then the dorsum in the intermediate alleles, black-and-tan (α^t), wild-type agouti (A), and white-bellied agouti (A^w), resulting in mice that have a black dorsum with yellow ventrum, agouti dorsum with agouti ventrum, and agouti dorsum with yellow ventrum, respectively. All-yellow dorsal hairs are associated with only the most dominant agouti alleles (e.g., lethal yellow, A^y ; viable yellow, A^{vy}), which usually confer a completely yellow phenotype with little or no eumelanin. In other words, the production of phaeomelanin is graded in such a manner that hair near the extremities (e.g., the pinnae, mammae, and perineum) has the greatest tendency of being pigmented yellow, followed by hair across the ventral surface of the animal, and finally by hair across the dorsal surface of the animal (reviewed in Silvers, 1979).

Genetic data can be interpreted as evidence for the agouti locus being comprised of either a single gene, which is regulated differently in each body region, or of several closely linked genes that work in concert (reviewed in Silvers, 1979). The observation that A^y can recombine with several other agouti-locus alleles, namely α^x (Russell et al., 1963), α (Siracusa et al., 1987a), and α^t (Siracusa et al., 1987a), certainly supports the multiple gene hypothesis. The recombination frequency in each case was shown to be on the order of 0.1%, with A^y lying proximal to α^x (Russell et al., 1963; Siracusa et al., 1987a). Additionally, the finding that reverse mutations from α to α^t or A^w are relatively common (Schlager and Dickie, 1967; Dickie, 1969;

Silvers, 1979), whereas all other spontaneous agouti-locus mutations are quite rare, also suggests that certain agouti-locus alleles may involve a complex of genetic sites, some of which, as is the case for A^y , are separable by recombination.

Perhaps the most interesting aspect of the agouti-locus function is its mode of action. Elegant skin-grafting experiments performed in the 1950s demonstrated that when melanocytes of one agouti genotype migrate into the follicular environment of another agouti genotype, they produce pigment typical of the follicular environment into which they have migrated (Silvers and Russell, 1955; Silvers, 1958a, 1958b, 1961, 1979). This indicated that agouti acts from within the microenvironment of the hair follicle, and not from within the melanocytes themselves. Most other coat-color genes, in contrast, act in a cell-autonomous manner within the melanocyte. Therefore, agouti must be regulating coat pigmentation by some direct or indirect form of cell-cell communication between the follicular environment and the melanocyte.

The agouti locus is also significant because several dominant mutations of the agouti gene are associated with embryonic lethality, obesity, diabetes, or an increased susceptibility to a variety of tumors. Lethal yellow has the most remarkable phenotype and, as a result, has been the most intensively studied. A^y was the first embryonic lethal mutation to ever be reported in a mammal (Cuenot, 1905, 1908). The lethality, which was revealed as a deviation in normal Mendelian inheritance (Cuenot, 1908; Castle and Little, 1910; Ibsen and Steigleder, 1917; Kirkham, 1917, 1919), occurs before implantation is complete (Robertson, 1942; Eaton and Green, 1963; Pederson, 1974; Calarco and Pederson, 1976; Pederson and Spindle, 1976; Granholm and Johnson, 1978) and is associated with defects in both the trophectoderm and the inner cell mass of blastula-stage embryos (Papaioannou and Gardner, 1979; Papaioannou, 1988; Barsh et al., 1990). In the heterozygous condition (where A^y is balanced with any other agouti allele), A^y confers an all-yellow coat color, pronounced obesity (Dickerson and Gowen, 1947; Fenton and Chase, 1951; Carpenter and Mayer, 1958; Plocher and Powley, 1976; Friedman and Leibel, 1992), a non-insulin-dependent diabetic-like condition (Hellerstrom and Hellman, 1963), and a propensity

to acquire spontaneous and induced tumors (reviewed in Silvers, 1979; Wolff et al., 1986; Wolff, 1987). Several other dominant agouti-locus alleles, including A^{vy} (viable yellow), A^{iy} (intermediate yellow), and A^{sy} (sienna yellow), confer the same suite of dominant pleiotropic effects as A^y , but with reduced penetrance (Silvers, 1979). A^{vy} , A^{iy} , and A^{sy} differ from A^y , however, because they are homozygous viable. Lethal nonagouti (α^x) is a recessive, radiation-induced allele, which, like A^y , is lethal when homozygous. The lethal effect of α^x is distinct from that of A^y , however, since embryos do not die until shortly after implantation (Russell et al., 1963; Papaioannou and Mardon, 1983) and, more importantly, because A^y/α^x embryos are perfectly viable and fertile (Silvers, 1979; Papaioannou and Mardon, 1983).

The agouti locus must be studied at the molecular level in order to determine whether it is comprised of a single gene or multiple genes and to elucidate its role in coat pigmentation as well as embryonic development, diabetes, and tumorigenesis. The endogenous ecotropic murine leukemia virus *Emv-15* has been localized very close to A^y (Copeland et al., 1983), and probes flanking *Emv-15* have been used to initiate a variety of physical mapping experiments and positional cloning strategies (Lovett and Epstein, 1987; Lovett et al., 1987; Siracusa et al., 1987a; Siracusa et al., 1987b; Barsh and Epstein, 1989a; Barsh and Epstein, 1989b; Siracusa et al., 1989). However, the results of certain presumed spontaneous crossover events indicate that *Emv-15* is not responsible for the A^y phenotype and that it does not lie between A^y and α (Lovett et al., 1989; Russell, 1989). As a result, despite considerable effort over the past few years, the molecular structure of the agouti locus has remained uncharacterized.

OBJECTIVES AND OVERVIEW

The objective of this dissertation was to characterize the mouse agouti locus at the molecular level in order to understand how it normally regulates coat pigmentation as well as how it contributes to embryonic lethality, diabetes, obesity, and tumorigenesis when mutated. Each of the specific aims listed below are addressed in the parts shown in parentheses:

- A. To gain molecular access to a region of DNA at the agouti locus (Parts II and III).
- B. To isolate and characterize cDNA and genomic clones representing the agouti-locus gene(s) (Part IV).
- C. To predict how the agouti protein functions based on its deduced amino acid sequence(s) (Part IV).
- D. To elucidate the temporal and spatial patterns of agouti gene expression (Parts IV and V).
- E. To ascertain whether agouti is a single gene or multiple gene locus (Part IV).
- F. To characterize the molecular structure of the α , α^t , and A^w alleles in an attempt to gain insight into each of their unique coat color phenotypes and identify the mechanism responsible for reverse mutations from α to α^t or A^w (Part V).
- G. To identify and characterize the DNA lesion in A^y in order to understand the molecular basis for the embryonic lethality of A^y homozygotes and the dominant pleiotropic effects of A^y heterozygotes (Parts VI and VII).
- H. To determine whether the agouti gene has been conserved in humans and evaluate its potential function (Part VIII).

REFERENCES

- Anderson, D. M., Lyman, S. D., Baird, A., Wignall, J. M., Eisenman, J., Rauch, C., March, C. J., Boswell, H. S., Gimpel, S. D., Cosman, D., and Williams, D. E. (1990). Molecular cloning of mast cell growth factor, a hematopoietin that is active in both membrane bound and soluble forms. *Cell* **63**: 235-243.
- Baldwin, C. T., Hoth, C. F., Amos, J. A., da-Silava, E. O., and Milunsky, A. (1992). An exonic mutation in the HuP2 paired domain gene causes Waardenburg's syndrome. *Nature* **355**: 637-638.
- Barsh, G. S., and Epstein, C. (1989a). Physical and genetic characterization of a 75-kilobase deletion associated with a^l , a recessive lethal allele at the mouse *agouti* locus. *Genetics* **121**: 811-818.
- Barsh, G. S. and Epstein, C. (1989b). The long-range restriction map surrounding the mouse *agouti* locus reveals a disparity between physical and genetic distances. *Genomics* **5**: 9-18.
- Barsh, G. S., Lovett, M., and Epstein, C. J. (1990). Effects of the lethal yellow (A^y) mutation in mouse aggregation chimeras. *Development* **109**: 683-690.
- Billingham, R. E. and Silvers, W. K. (1960). The melanocytes of mammals. *Quart. Rev. Biol.* **35**: 1-40.
- Brilliant, M. H., Gondo, Y., and Eicher, E. M. (1991). Direct molecular identification of the mouse pink-eyed unstable mutation by genome scanning. *Science* **252**: 566-569.
- Calarco, P. G., and Pedersen, R. A. (1976). Ultrastructural observations of lethal yellow (A^y/A^y) mouse embryos. *J. Embryol. Exp. Morphol.* **35**: 73-80.

- Carpenter, K. J., and Mayer, J. (1958). Physiologic observations on yellow obesity in the mouse. *Am. J. Physiol.* **193**: 499-504.
- Castle, W. E., and Little, C. C. (1910). On a modified Mendelian ratio among yellow mice. *Science* **32**: 868-870.
- Chabot, B., Stephenson, D. A., Chapman, V. M., Besmer, P., and Bernstein, A. (1988). The proto-oncogene *c-kit* encoding a transmembrane tyrosine kinase receptor maps to the mouse *W* locus. *Nature* **335**: 88-89.
- Copeland, N. G., Jenkins, N. A., and Lee, B. K. (1983). Association of the lethal yellow *A^y* coat color mutation with an ecotropic murine leukemia virus genome. *Proc. Natl. Acad. Sci. USA* **80**: 247-249.
- Cuenot, L. (1905). Les races pures et leurs combinaisons chez les souris. *Arch. Zool. Exp. Gen.* **3**: 123-132.
- Cuenot, L. (1908). Sur quelques anomalies apparentes des proportions Mendeliennes. *Arch. Zool. Exp. Gen.* **9**: 7-15.
- Dickerson, G. E., and Gowen, J. W. (1947). Hereditary obesity and efficient food utilization in mice. *Science* **105**: 496-498.
- Dickie, M. M. 1969. Mutations at the agouti locus in the mouse. *J. Hered.* **60**: 20-25.
- Eaton, G. J., and Green, M. M. (1963). Giant cell differentiation and lethality of homozygous yellow mouse embryos. *Genetica* **34**: 155-161.
- Epstein, D. J., Vekemans M., and Gros, P. (1991). Splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell* **67**: 767-774.

- Fenton, P. F., and Chase, H. B. (1951). Effect of diet on obesity of yellow mice in inbred lines. *Proc. Soc. Exper. Biol.* **77**: 420-422.
- Flanagan, J. G., and Leder, P. (1990). The *kit* ligand: A cell surface molecule altered in steel mutant fibroblasts. *Cell* **63**: 185-194.
- Fleischman, R. A. (1992). Human piebald trait resulting from a dominant negative allele of the c-kit membrane receptor gene. *J. Clin. Invest.* **89**: 1713-1717.
- Friedman, J. M., and Leibel, R. L. (1992). Tackling a weighty problem. *Cell* **69**: 217-220.
- Galbraith, D. B. (1964). The agouti pigment pattern of the mouse: A quantitative and experimental study. *J. Exp. Zool.* **155**: 71-90.
- Gardner, J. M., Nakatsu, Y., Gondo, Y., Lee, S., Lyon, M. F., King, R. A., and Brilliant, M. H. (1992). The mouse pink-eyed dilution gene; association with human Prader-Willi and Angelman syndromes. *Science* **257**: 1121-1124.
- Granholt, N.H., and Johnson, P.M. (1978). Enhanced identification of lethal yellow (A^y/A^y) mouse embryos by means of delayed development of four-cell stages. *J. Exp. Zool.* **205**: 327-333.
- Green, M. C. (1989). Catalogue of mutant genes and polymorphic loci. In *Genetic variants and strains of the laboratory mouse*, (eds., M. F. Lyon, and A. G. Searle), pp. 17-20. Oxford, New York: Oxford University Press, Oxford and New York.
- 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. USA.* **88**: 8896-8899.

- Hellerstrom, C., and Hellman, B. (1963). The islets of Langerhans in yellow obese mice. *Metabolism* **12**: 527-536.
- Hodgkinson, C. A., Moore, K. J., Narayama, A., Steingrimsson, E., Copeland, N. G., Jenkins, N. A., and Arnheiter, H. (1993). Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. *Cell* **63**: 395-404.
- Hirata, T., Morli, E., Morimoto, M., Kasugai, T., Tsujimura, T., Hirota, S., Kanakura, Y., Nomura, S., and Kitamura, Y. (1993). Stem cell factor induces outgrowth of c-kit-positive neurites and supports the survival of c-kit-positive neurons in dorsal root ganglia of mouse embryos. *Development* **119**: 49-56.
- Huang, E., Nocka, K., Beier, D. R., Chu, T.-Y., Buck, J., Lahm, H.-W., Wellner, D., Leder, P., and Besmer, P. (1990). The hematopoietic growth factor KL is encoded by the *Sl* locus and is the ligand of the c-kit receptor, the gene product of the *W* locus. *Cell* **63**: 225-233.
- Ibsen, H. L., and Steigleder, E. (1917). Evidence for the death *in utero* of the homozygous yellow mouse. *Am. Nat.* **51**: 740-752.
- Jackson, I. J. (1988). A cDNA encoding tyrosinase-related protein maps to the brown locus in mice. *Proc. Natl. Acad. Sci. USA*. **85**: 4392-4396.
- Jackson, I. J., Chambers, D. M., Tsukamoto, K., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Hearing, V. J. (1992). A second tyrosinase-related protein, TRP-2, maps to and is mutated at the mouse slaty locus. *EMBO J.* **11**: 527-535.
- Kirkham, W. B. (1917). Embryology of the yellow mouse. *Anat. Rec.* **11**: 480-481.
- . (1919). The fate of homozygous yellow mice. *J. Exp. Zool.* **28**: 125-135.

- Kwon, B. S., Haq, A. K., Pomerantz, S. H., and Halaban, R. (1987). Isolation and sequence of a cDNA locus for human tyrosinase that maps at the mouse c-albino locus. *Proc. Natl. Acad. Sci. USA*. **84**: 7473-7477.
- Kwon, B. S., Chintamaneni, C. D., Kozak, C. A., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Barton, D. E. (1991). Pmel-17, maps near the silver coat color locus on mouse chromosome 10 and is a syntenic region on human chromosome 12. *Proc. Natl. Acad. Sci. USA*. **88**: 9228-9232.
- Keller, S. A., Liptay, S., Hajra, A., and Meisler, M. H. (1990). Transgene-induced mutation of the murine steel locus. *Proc. Natl. Acad. Sci. USA*. **87**: 10019-10022.
- Lovett, M., and Epstein, C. J. (1987). The lethal yellow allele associated provirus results in the production of chimeric viral-host RNAs. *Proc. Natl. Acad. Sci. USA*. **84**: 2853-2857.
- Lovett, M., Cheng, Z. Y., Lamela, E. M., Yokoi, T., and Epstein, C. (1987). Molecular markers for the *agouti* coat color locus of the mouse. *Genetics* **115**: 747-754.
- Martin, F. H., Suggs, S. V., Langley, K. E., Lu, H. S., Ting, J., Okino, K. H., Morris, C. F., McNiece, I. K., Jacobsen, F. W., Mendiaz, E. A., Birkett, N. C., Smith, K. A., Johnson, M. J., Parker, V. P., Flores, J. C., Patel, A. C., Fisher, E. F., Erjavec, H. O., Herrera, C. J., Wypych, J., Sachdev, R. K., Pope, J. A., Leslie, I., Wen, D., Lin, C.-H., Cupples, R. L., and Zsebo, K. M. (1990). Primary structure and functional expression of rat and human stem cell factor DNAs. *Cell* **63**: 203-211.
- Mercer, J. A., Seperack, P. K., Strobel, M. C., Copeland, N. G., and Jenkins, N. A. (1991). Novel myosin heavy chain encoded by murine dilute coat color locus. *Nature* **349**: 709-713.

- Papaioannou, V. E. (1988). Investigation of the tissue specificity of the lethal yellow (A^y) gene in mouse embryos. *Devl. Genet.* **9**: 155-165.
- Papaioannou, V., and Gardner, R. L. (1979). Investigations of the lethal yellow A^y/A^y embryo using mouse chimaeras. *J. Embryol. Exp. Morph.* **52**: 153-163.
- Papioannou, V. E. and Mardon, H. (1983). Lethal nonagouti (a^x): Description of a second embryonic lethal at the agouti locus. *Dev. Genet.* **4**: 21-29.
- Pedersen, R. A. (1974). Development of lethal yellow (A^y/A^y) mouse embryos *in vitro*. *J. Exp. Zool.* **188**: 307-319.
- Pedersen, R. A., and Spindle, A.I. (1976). Genetic effects on mammalian development during and after implantation. In *Embryogenesis in mammals, CIBA Foundation Symp. 40* (eds. K. Elliot and M. O'Connor), pp. 133-154. Elsevier, New York.
- Plocher, T.A., and Powley, T.L. (1976). Effect of hypophysectomy on weight gain and body composition in the genetically obese yellow (A^y/a) mouse. *Metabolism* **25**: 593-602.
- Rinchik, E. M., Bultman, S. J., Horsthemke, B., Lee, S.-T., Strunk, K. M., Spritz, R. A., Avidano, K. M., Jong, M. T. C. and Nicholls, R. D. (1993). A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. *Nature* **361**: 72-76.
- Robbins, L. S., Nadeau, J. H., Johnson, K. R., Kelly, M. A., Roselli-Rehfuss, L., Baack, E., Mountjoy, K. G., and Cone, R. D. (1993) Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* **72**: 827-834.

- Robertson, G. G. (1942). An analysis of the development of homozygous yellow mouse embryos. *J. Exp. Zool.* **89**: 197-231.
- Russell, L.B. (1989). Functional and structural analyses of mouse genomic regions screened by the morphological specific-locus test. *Mutation Research* **212**: 23-32.
- Russell, L. B., McDaniel, N. C., and Woodiel, F. N. (1963). Crossing-over within the a "locus" of the mouse. *Genetics* **48**: 907a.
- Sakurai, T., H. Ochia, and T. Takeuchi. 1975. Ultrastructural change of melanosomes associated with agouti pattern formation in mouse hair. *Dev. Biol.* **47**: 466-471.
- Schlager, G. and M.M. Dickie. 1967. Spontaneous mutations and mutation rates in the house mouse. *Genetics* **57**: 319-330.
- Searle, A. G. (1968). *Comparative genetics of coat color in mammals*. Academic press, New York.
- Shibahara, A., Tomita, Y., Sakakura, T., Nager, C., Chaudhuri, B., and Muller, R. (1986). Cloning and expression of cDNA encoding mouse tyrosinase. *Nucl. Acids Res.* **14**: 2413-2427.
- Silvers, W. K. (1958a). An experimental approach to action of genes at the agouti locus in the mouse. II. Transplants of newborn aa ventral skin to $a^t a$, $A^w a$, and aa hosts. *J. Exp. Zool.* **137**: 181-188.
- Silvers, W. K. (1958b). An experimental approach to action of genes at the agouti locus in the mouse. III. Transplants of newborn A^w -, A -, a^t -skin to A^y -, A^w -, A - and aa hosts. *J. Exp. Zool.* **137**: 189-196.
- Silvers, W. K. (1961). Genes and pigment cells of mammals. *Science* **134**: 368-373.

- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6-44. Springer-Verlag, New York.
- Silvers, W. K., and Russell, E. S. (1955). An experimental approach to action of genes at the agouti locus in the mouse. *J. Exp. Zool.* **130**: 199-220.
- Siracusa, L. D., Russell, L. B., Jenkins, N. A., and Copeland, N. G. (1987a). Allelic variation within the *Emv-15* locus defines genomic sequences closely linked to the *agouti* locus on mouse chromosome 2. *Genetics* **117**: 85-92.
- Siracusa, L. D., Russell, L. B., Eicher, E. M., Corrow, D. J., Copeland, N. G., and Jenkins, N. A. (1987b). Genetic organization of the *agouti* region of the mouse. *Genetics* **117**: 93-100.
- Siracusa, L. D., Buchberg, A. M., Copeland, N. G., and Jenkins, N. A. (1989). Recombinant inbred strain and interspecific backcross analysis of molecular markers flanking the murine *agouti* coat color locus. *Genetics* **122**, 669-679.
- Spritz, R. A., Holmes, S. A., Berg, S. Z., Nordlund, J. J., and Fukai, K. (1993). A recurrent deletion in the KIT (mast/stem cell growth factor receptor) proto-oncogene is a frequent cause of human piebaldism. *Hum. Mol. Genet.* **2**: 1499-1500.
- 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.

- Tsukamoto, K., Jackson, I. J., Urabe, K., Montague, P. M., and Hearing, V. J. (1992). A second tyrosinase-related protein, TRP-2, is a melanogenic enzyme termed DOPAchrome tautomerase. *EMBO J.* **11**: 519-526.
- Witte, O. N. (1990). Steel locus defines new multipotent growth factor. *Cell* **63**: 5-6.
- Wolff, G. L. (1987). Body weight and cancer. *Am. J. Clin. Nutr.* **45**: 168-180.
- Wolff, G. L., Roberts, D. W., and Galbraith, D. B. (1986). Prenatal determination of obesity, tumor susceptibility, and coat color pattern in viable yellow (A^{vy}/a) mice: The yellow mouse syndrome. *J. Hered.* **77**: 151-158.
- Zsebo, K. M., Williams, D. A., Geissler, E. N., Broudy, V. C., Martin, F. H., Atkins, H. L., Hsu, R.-Y., Birkett, N. C., Okino, K. H., Murdock, D. C., Jacobsen, F. W., Langley, K. E., Smith, K. A., Takeishi, T., Cattanaach, B. M., Galli, S. J., and Suggs, S. V. (1990). Stem cell factor is encoded at the *Sl* locus of the mouse and is the ligand for the c-kit tyrosine kinase receptor. *Cell* **63**: 213-224.

PART II

MOLECULAR AND GENETIC CHARACTERIZATION OF A RADIATION-INDUCED STRUCTURAL REARRANGEMENT IN MOUSE CHROMOSOME 2 CAUSING MUTATIONS AT THE LIMB DEFORMITY AND AGOUTI LOCI

ABSTRACT

Molecular characterization of mutations in the mouse, particularly those involving agent-induced major structural alterations, is proving to be extremely useful for correlating the structure and expression of individual genes with their function in the whole organism. One radiation-induced mutation simultaneously generated new alleles of the limb deformity (*ld*) and agouti (*a*) loci, two regions of chromosome 2 normally separated by 20 cM. Cytogenetic analysis provides evidence that an interstitial segment of chromosome 17 (17A2-17B) has been translocated into the distal end of chromosome 2, resulting in a smaller-than-normal Chromosome 17 (17^d) and a longer form of chromosome 2 (2¹⁷). Additionally, an interstitial segment of chromosome 2 (2E4-2H2), immediately proximal to the insertion, does not match the corresponding normal segment. Molecular probes from an insertional mutation at the *ld* locus in a transgenic mouse allowed us to clone a DNA structural rearrangement associated with this radiation-induced mutation. Analysis of this rearrangement indicated that a region at or near the *a* locus has become juxtaposed with a portion of the *ld* locus on the mutant chromosome. Our data suggest that the two new mutations arose by an inversion of a section of Chromosome 2 with DNA breakpoints and sequence disruptions in the *ld* and *a* loci.

INTRODUCTION

Molecular characterization of mutations is a powerful way of correlating the structure and expression of genes with their function in the organism. In mice, mutations produced in chemical- and radiation-mutagenesis programs are important tools for the characterization of the mammalian genome. Many of these mutations are major structural chromosomal alterations, namely, deletions, inversions and translocations, and are particularly useful for functional studies of the mouse genome. For example, deletion mutations, arising mainly from radiation-exposure experiments, are being used as tools in complementation (Russell, 1989) and saturation mutagenesis (Rinchik et al., 1989) experiments to define and

study the function of individual genes located in several different regions of the genome. DNA probes mapping to the deleted regions (Johnson et al., 1989) are being used to derive physical maps of deleted genome sections and to clone and characterize the genes associated with these regions (Johnson et al., 1989).

Here we describe the genetic, cytogenetic and molecular characterization of a complex, radiation-induced chromosomal rearrangement in the mouse associated with two phenotypes controlled by chromosome 2 loci. These loci, *ld* (limb deformity) and *a* (agouti), normally are separated by 20 centimorgans (cM). The *ld* locus was originally defined over 30 years ago by two independently arising spontaneous mutations, one at the Oak Ridge National Laboratory and the other at the Jackson Laboratory (Cupp, 1962; Green, 1962; Green, 1968). This locus, which became accessible at the molecular level with probes derived from an insertional mutation in a transgenic mouse (Woychik et al., 1985), is associated most notably with a skeletal abnormality affecting both the fore- and hindlimbs of the animal and with kidney defects (Cupp, 1962; Green, 1962; Green, 1968; Kleinebrecht et al., 1982; Woychik et al., 1985; Zeller et al., 1989). The *a* locus regulates the distribution of the eumelanin and pheomelanin pigments in individual hairs as well as over the entire coat (Silvers, 1979; Green, 1981). Additionally, analysis of radiation-induced mutations has revealed the presence of genes within or near the *a* locus that regulate embryonic development (Russell, 1989; Silvers, 1979; Green, 1981; Siracusa et al., 1987a). Despite a considerable effort, using genetic mapping and molecular "walking" techniques, the gene associated with the nonagouti phenotype has not been cloned (Siracusa et al., 1987b; Lovett et al., 1987; Siracusa et al., 1989; Barsh and Epstein, 1989). The radiation-induced mutation described here involves an inversion of chromosome 2 that has allowed us to use probes from the *ld* locus to "jump" to a region tightly linked to the *a* locus. Consequently, we have been able to derive DNA probes within or very close to the genes associated with this developmentally interesting region of the genome.

MATERIALS AND METHODS

Mutagenesis and Genetic and Cytogenetic Analysis

A ^{137}Cs source was used to generate gamma radiation, delivered at a dose rate of 0.7 R/min (1R=0.258mC/kg). (C3H/Rl X 101/Rl)F₁ males, of agouti-locus genotype A/A^w , were irradiated with 300 R and subsequently mated to (SEC/Rl X C57BL/E) F₁ females of genotype a/a ; $+/b$; $+/c^{ch}$; $+ +/d^{se}$. The mutant described in this report was conceived 41 weeks after irradiation and was therefore derived from an irradiated spermatogonial stem cell. All stocks of mice employed in the radiation experiment, as well as those used in subsequent crosses of the mutant and his descendants, were bred at the Oak Ridge National Laboratory. All methods for cytogenetic analysis were as described previously (Pathek et al., 1975). The *M. spretus*/*M. musculus* backcross experiments were conducted as described (van der Meer-de Jong et al., 1990).

Northern and Southern Blot Analysis

For Southern blots, 10 μg of genomic liver DNA was digested with restriction enzymes, electrophoresed through agarose gels and blotted and UV cross-linked to Genescreen (Dupont) using standard procedures (Ausubel et al., 1988). For Northern blots, poly(A)⁺ RNA, prepared from 500 μg of total RNA, was electrophoresed in formaldehyde gels and blotted and UV cross-linked to Genescreen (Dupont) using standard procedures (Ausubel et al., 1988; Maniatis et al., 1982). All blots were probed with random-hexamer labeled double-stranded restriction fragments (Feinberg and Vogelstein, 1984).

Cloning the DNA Breakpoint Within the *ld* Locus

A total of 250 μg of genomic liver DNA from a male heterozygous for the rearrangement was digested to completion with EcoRI and size fractionated on a 10-40% sucrose gradient (Ausubel et al., 1988; Maniatis et al., 1982). An EMBL4 phage library, prepared with fractions containing 6.5-7.5-kb fragments, was screened using standard procedures using a region of the *ld* locus as a probe (Ausubel et al., 1988; Maniatis et al., 1982). The fragments of interest were subcloned into the pGEM4 vector (Promega), restriction mapped, and partially

sequenced using standard procedures (Ausubel et al., 1988; Maniatis et al., 1982).

RESULTS

Mutant Recovery and Genetic Analysis

Among 2870 male offspring of irradiated (C3H/Rl X 101/Rl) F_1 males (A/A^w) mated to (SEC/Rl x C57BL/E) F_1 females (a/a), there was one nonagouti male, presumed to be a/a^* , where a^* represents a mutation involving the a locus on either the A or A^w chromosome. In crosses to a/a mice, this mutant male produced 52 nonagouti offspring, and no other types. In crosses to A/A stocks, 17 offspring were obtained, and all were agouti. Also, in crosses to A/a mice, 6 agouti and 10 nonagouti offspring were obtained. These results indicate that a^* is recessive to A , and is either equivalent or recessive to a .

The mutant was revealed to be "partially sterile" due to embryonic death of some of his progeny. In matings with 12 normal females, 29% of the conceptuses were found to die at around the time of implantation. Among surviving sons of the mutant, 19 were also "partially sterile" (with, again, 29% of their conceptuses dying at around implantation), whereas two were completely sterile, and 22 had normal fertility.

Homozygotes were obtained from matings of partially sterile males to daughters of partially sterile males. Such homozygotes of either sex were found to be fully fertile, allowing establishment of a homozygous stock. The phenotype of the homozygous animals resembles aj/aj "jet" (probably equivalent to a^e/a^e , extreme nonagouti), which is recessive to a . Crosses of homozygotes to aj/aj produced only jet offspring.

Unexpectedly, these homozygous a^*/a^* jet-colored animals also had severely deformed fore- and hindlimbs. Close examination of alizarin-red-S stained skeletons from 19 homozygous, young adult mice revealed an abnormality where, in all cases, there was an absence of at least two digits on each extremity, radioulnar synostosis, complete fibular aplasia, and aplasia or hypoplasia of some carpals and many tarsals. Some variable features were also observed, which included boney syndactyly and

brachydactyly, fusions of various carpals and tarsals, malpositioning of the patella, and hindfeet that do not point in the correct direction. In extreme cases, the hindfeet contained only a single digit, or the foot articulated with a projection from the side of the tibia instead of from the distal end of the tibia. The defects were bilateral but generally asymmetric. The radioulnar synostosis could often be interpreted as also involving radial hypoplasia except near the elbow, and, in some mice, there appeared to be no radius.

In addition to the skeletal malformations, kidney defects were also observed in homozygous progeny from crosses involving heterozygotes and homozygotes in various combinations. The kidneys of eleven post-weaning mice, nine of which were homozygotes and expressed the limb defect, were analyzed. Six of the 9 animals with the limb defect had renal abnormalities consisting of various degrees of hydronephrosis and hydroureter, and 1 animal had unilateral aplasia and hypoplasia. None of the animals with normal limbs had kidney defects. Additionally, kidneys were examined in 133 fetuses at day 17, of which 80 clearly expressed the limb defect. In this case, 17 of the 80 fetuses with the limb defect had urinary-tract defects, including dilation of one or both ureters (ten fetuses), renal hypoplasia (five fetuses), and renal agenesis (two fetuses). None of the fetuses with normal limbs showed any kidney defects.

Since the limb and kidney defects resembled the phenotype of the *ld* mutation, various crosses were made between mice carrying the radiation-induced mutation described above and mice of a stock carrying the allele *ld^{OR}* (Cupp, 1962). The results of these crosses clearly reveal that the new mutation is allelic to *ld* (Table 1). Therefore, the new mutation involves alterations at two different loci, *a* and *ld*, which are normally separated by 20 cM on the standard genetic map (Davisson and Roderick, 1978).

Cytogenetic Analysis Reveals A Chromosomal Rearrangement

The partial sterility of the original mutant and some of his descendents suggested that the radiation-induced mutation might have resulted from a chromosomal rearrangement. Therefore, karyotypes were prepared of Giemsa-banded metaphase chromosomes from four phenotypically normal heterozygous adults (2 males, 2 females), and from

two adult homozygous males expressing both the nonagouti and limb-deformity conditions (Fig. 1A). Although high-resolution Giemsa-banding analysis has thus far failed to provide a definitive cytogenetic interpretation of this mutation, it is clear that all karyotypes show a longer-than-normal chromosome 2 (designated 2^{17}) and a shortened chromosome 17 (designated 17^{del}) (Fig. 1A). The banding pattern suggests that an interstitial segment of chromosome 17, involving bands 17B-17C, integrated into the subdistal region of chromosome 2 proximal to band 2H2 (Fig. 1B). (An alternative interpretation is that bands 17A2-17B are inserted in the same location but in an inverted fashion.) In addition, a segment of the 2^{17} chromosome immediately proximal to the insertion does not match the 2E4-2H1 segment of the normal chromosome 2 and is presumably rearranged (Fig. 1).

Altered Gene Expression at the *ld* Locus

We were interested in analyzing how this new mutation affects the expression of a gene at the *ld* locus whose structure and patterns of expression will be described in detail elsewhere (R.P.W., R. Maas, R. Zeller, T. Vogt, and P. Leder, unpublished results). In brief, this gene is structurally very complex and is capable of producing a family of mRNAs by alternate tissue-specific splicing and the utilization of multiple polyadenylation sites. Expression of this gene is normally detectable in numerous sections of the developing fetus as well as in several tissues of the adult (Zeller et al., 1989; R.P.W. et al., unpublished results). The composite structure of some RNA transcripts produced from this gene was derived from overlapping cDNA clones and is shown in Figure 2A.

By using a small 0.7-kb section of cDNA as a probe (Fig. 2A), the analysis of RNA from the adult brain, kidney, testes, salivary gland, and small intestine revealed that the expression of this gene is markedly altered in animals homozygous for the radiation-induced mutation (Fig. 2B). With the exception of a very small transcript in the testis, none of the normal transcripts, covering a range of sizes up to approximately 13 kb, were detectable in any of the mutant samples (Fig. 2B). However, aberrant-sized transcripts were observed in the brain and testes of the mutant animals.

DNA Breakpoint in the *ld* Locus

To test the hypothesis that the altered expression of the gene at the *ld* locus was caused by a radiation-induced DNA structural rearrangement, we searched for an altered-size DNA fragment unique to the mutant allele and not detectable in the DNA from either the C3H or the 101 inbred parents. When the same region of cDNA that had been used for the Northern blot in figure 2B was employed as a probe, an altered-size DNA fragment was observed with each of several different restriction enzymes, including *Eco*RI, *Xba* I, and *Hind*III (Fig. 3A). The observation that multiple enzymes generated a DNA fragment unique to the mutant allele suggested that this region of the *ld* locus had undergone a structural rearrangement.

To characterize the structural rearrangement, the normal 6.5-kb and the corresponding mutant 7.5-kb *Eco*RI fragments (Fig. 3A), were cloned and analyzed (Fig. 3B). Hybridization and nucleotide-sequencing experiments revealed that the two clones diverged in sequence at a position approximately 0.9-Kb from the 3' end of the normal 6.5-kb fragment (Figure 3B; S.B., C. Tancongo, R.P.W., unpublished results). Therefore, the 7.5-kb fragment arose from a structural alteration of the *ld* locus.

Mouse-hamster somatic cell hybrids were used to determine whether this structural alteration represented a DNA breakpoint region in the 2¹⁷ chromosome associated with the inserted sequences from chromosome 17. A DNA probe, corresponding to a 0.2-kb restriction fragment unique to the region 3' to the breakpoint on the mutant 7.5-kb fragment (Fig. 3B), was hybridized to DNA from mouse-hamster somatic cell hybrids containing, along with other mouse chromosomes, either both chromosomes 2 and 17 (hybrid line BEM 1-4) or just mouse chromosome 17 (hybrid line R44-1) (D'Eustachio et al., 1987). Results showed that the probes hybridized to chromosome 2 but not chromosome 17 (data now shown). Therefore, the alteration in the *ld* locus represented by the cloned 7.5-kb fragment did not involve the DNA breakpoint associated with the insertion of the chromosome 17 sequences.

In additional experiments, this same probe, derived from the region 3' to the breakpoint on the 7.5-kb fragment, was hybridized to DNA from 91 different *M.spretus*/*M.musculus* backcross animals (van der Meer-de Jong et

al., 1990). *Msp*I digestion of the DNA allowed us to differentiate the *M. spretus* from the *M. musculus* hybridizing alleles and to determine the genetic distance of the sequence 3' to the breakpoint on the 7.5-kb fragment relative to the normal *ld* and α loci, which had been mapped previously in this manner (R. van der Meer-de Jong et al., unpublished results). As shown in Fig. 4, a total of 13 backcross animals (14%, with 95% confidence limits 7.7 and 22.7) were recombinants between the breakpoint sequence on the 7.5-kb fragment and the *ld* locus. This is comparable to the established 20 cM distance between *ld* and α loci (22). However, there were no recombinants between this sequence and the α locus (upper 95% confidence limit of zero recombination frequency, 4%), indicating that the sequence 3' to the breakpoint on the mutant 7.5-kb fragment is normally tightly linked to the α locus. Therefore, the *ld* locus is juxtaposed with a region at or near the α locus to form the structurally altered 7.5-kb *Eco*RI fragment on the mutant allele.

DISCUSSION

A mouse mutation that is allelic to both *ld* and α arose in a radiation-mutagenesis experiment. Cytogenetic analysis revealed that a short interstitial segment of chromosome 17 (17B-17C; or, possibly, inverted 17A2-17B) was inserted into the distal region of chromosome 2, proximal to band 2H2 and that a large segment in the 2¹⁷ chromosome, immediately proximal to the insertion site, did not match the 2E4-2H1 region of the normal chromosome 2 (Fig. 1). Molecular analysis revealed that a region at or near the α locus has been joined to the *ld* locus in the mutant allele. Based on the genetic, cytogenetic and molecular findings presented here, we are proposing that the nonmatching interstitial segment, probably containing bands 2E4-2H1, was inverted on the 2¹⁷ chromosome, with DNA breaks and sequence disruptions in the *ld* and α loci. Also, based on the cytogenetic analysis, we are speculating that the 17B-17C (or possibly, 17A2-17B) interstitial segment on chromosome 17 had integrated at a point immediately adjacent and distal to the 2E4-2H1 inverted region of the 2¹⁷ chromosome. On the basis of the available evidence concerning this

structural rearrangement, we have designated the mutation *Is(17;In2)ld,a1Gso*. For convenience, the name of the rearrangement will be abbreviated to *Is1Gso*, and the genotype at the mutant loci will be designated as *ld^{In2}a^{jIn2}*.

The DNA breakpoint within *ld^{In2}* occurs at a position that is approximately 8 kb on the 3' side of the site where the transgene integrated in the insertional allele, *ld^{Hd}* (Fig. 3B) (Woychik et al., 1985). The close proximity of the structural alteration in the *ld^{Hd}* and the *ld^{In2}* mutations not unexpectedly leads to a similar disruption in the expression of the gene at the *ld* locus. In experiments to be described elsewhere (D. Maas et al., unpublished results), the analysis of RNA from several adult tissues indicated that the alteration in the expression of this gene in *ld^{In2}* is similar to that observed with *ld^{Hd}*. The fact that the expression of this gene is disrupted by both of these alleles of *ld* provides compelling evidence that this particular transcription unit represents the gene that is directly related to the mutant phenotype.

The inversion in the structurally altered chromosome 2 in the *Is1Gso* rearrangement has allowed us to jump directly from the cloned region of the *ld* locus to a region within or very close to the *a* locus. The proximity of this cloned region to the *a* locus, in addition to its association with the *a* locus phenotype of the mutation described here, leads us to hypothesize that the breakpoint region on the 7.5-kb mutant fragment (Fig. 3) is very close to, or possibly within, the gene directly associated with the extreme-nonagouti ("jet") phenotype. If correct, this is of considerable significance in that it may now provide us with molecular probes to initiate a study of the molecular biology of the genes situated at the *a* locus. However, before any detailed analysis can be performed, it will first of all be necessary to map this putative agouti-region probe relative to the mutations at this locus. A number of agent-induced mutations that have been generated over the past several decades should be particularly useful for this purpose.

REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1988). *Current Protocols in Molecular Biology*. John Wiley & Sons/Greene, New York.
- Barsh, G. S., and Epstein, C. (1989a). Physical and genetic characterization of a 75-kilobase deletion associated with α^l , a recessive lethal allele at the mouse *agouti* locus. *Genetics* **121**: 811-818
- Cupp, M. B. (1962). *Mouse News Lett* **26**: 51.
- Davisson, M. T., and Roderick, T. H. (1978). Status of the linkage map of the mouse. *Cytogenet. Cell Genet.* **22**: 552-557.
- D'Eustachio, P., Jadidi, S., Fuhlbrigge, R. C., Gray, P. W., and Chaplin, D. D. (1987). Interleukin-1 α and β genes: linkage on chromosome 2 in the mouse. *Immunogenetics* **26**: 339-343
- Feinberg, A. P., and Vogelstein, B. (1984). Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**: 266-267.
- Granholm, N. H., Jeppesen, K. W., and Japs, R. A. (1986). Progressive infertility in female lethal yellow mice. *J. Reprod. Fertil.* **76**: 279-287.
- Green, M. C. (1962). *Mouse News Lett.* **26**: 34.
- Green, M. C. (1966). Mutant genes and linkages. In *Biology of the Laboratory Mouse*, 2ND ed (ed. E. L. Green), pp.87-150. McGraw-Hill, New York.
- Johnson, D. K., Hand, R. E., and Rinchik, E. M. (1989). Molecular mapping within the mouse albino-deletion complex. *Proc. Natl. Acad. Sci. USA* **86**: 8862-8866.

- Kleinebrecht, J., Selow, J., and Winkler, W. (1982). The mouse mutant limb-deformity (*ld*). *Anat. Anz. Jena* **152**: 313-321.
- Lovett, M., Cheng, Z. Y., Lamela, E. M., Yokoi, T., and Epstein, C. (1987). Molecular markers for the *agouti* coat color locus of the mouse. *Genetics* **115**: 747-754.
- Maniatis, T., Fritsch, E. F., Sambrook, J. (1982). *Molecular cloning: A laboratory manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
- Pathak, S., Stock, A. D., and Lusby, A. (1975). A combination of sister chromatid differential staining and giemsa banding. *Experimentia* **31**: 916-918.
- Rinchik, E. M., Carpenter, D. A., and Selby, P. B. (1990). A strategy for fine-structure analysis of a 6- to 11-cM region of mouse chromosome 7 by high efficiency mutagenesis. *Proc. Natl. Acad., Sci USA*. **87**: 896-900.
- Russell, L.B. (1989). Functional and structural analyses of mouse genomic regions screened by the morphological specific-locus test. *Mutation Research* **212**: 23-32.
- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6-44. Springer-Verlag, New York.
- Siracusa, L. D., Russell, L. B., Jenkins, N. A., and Copeland, N. G. (1987a). Allelic variation within the *Emv-15* locus defines genomic sequences closely linked to the *agouti* locus on mouse chromosome 2. *Genetics* **117**: 85-92.

- Siracusa, L. D., Russell, L. B., Eicher, E. M., Corrow, D. J., Copeland, N. G., and Jenkins, N. A. (1987b). Genetic organization of the *agouti* region of the mouse. *Genetics* **117**: 93-100.
- 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.
- 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.
- Zeller, R., Jackson-Grusby, L., and Leder, P. (1989). The limb deformity gene is required for apical ectodermal ridge differentiation and anteroposterior limb pattern formation. *Genes & Dev.* **3**: 1481-1492.

APPENDIXES

APPENDIX A

TABLES

Table 1. Manifestation of allelism between R^* and ld^{OR} as expressed by limb morphology of progeny.

Cross			<u>Mice, no.</u>	
			Abnormal	Normal
R/R	x	ld^{OR}/ld^{OR}	48	0
R/ ld^{OR}	x	R/ ld^{OR}	73	0
ld^{OR}/ld^{OR}	x	R/ ld^{OR}	78	0
$ld^{OR}/+$	x	R/R	58	49

The R^* designation refers to the radiation-induced mutation Is1Gso.

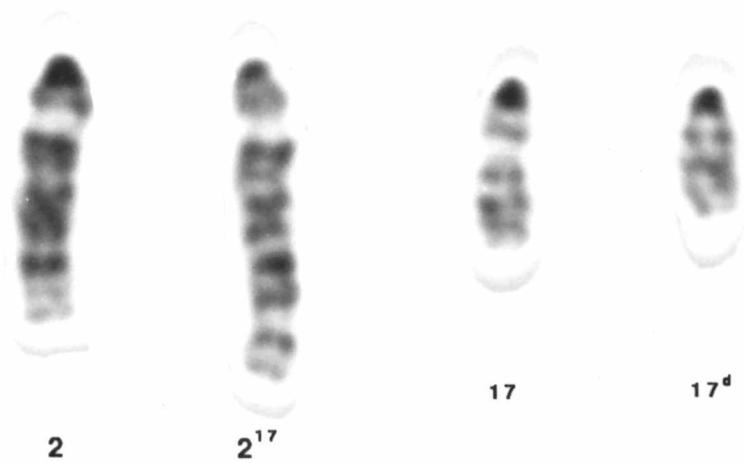
APPENDIX B

FIGURES

Figure 1. Structure of the rearranged chromosomes 2 and 17 in the radiation mutant.

(A) Chromosomes from Giemsa-banded mitotic metaphase karyotypes. From left to right, normal chromosome 2, rearranged chromosome 2 (designated 2^{17}), normal chromosome 17, and deleted chromosome 17 (designated 17^{del}). (B) G-banding diagrams illustrating the proposed structure for the radiation-induced chromosomal rearrangement involving chromosomes 2 and 17 (see text). Broken, curved arrows show the position of the inverted 2E4-2H2 section of chromosome 2. Solid linear arrows between chromosomes 2 and 2^{17} show the position of the interstitial segment of chromosome 17 that has integrated into the distal portion of chromosome 2, immediately adjacent and proximal to the 2E4-2H2 inversion. The probable positions of the *a* and *ld* loci relative to the banded map are shown to the left of the diagram for the normal chromosome 2.

A.



B.

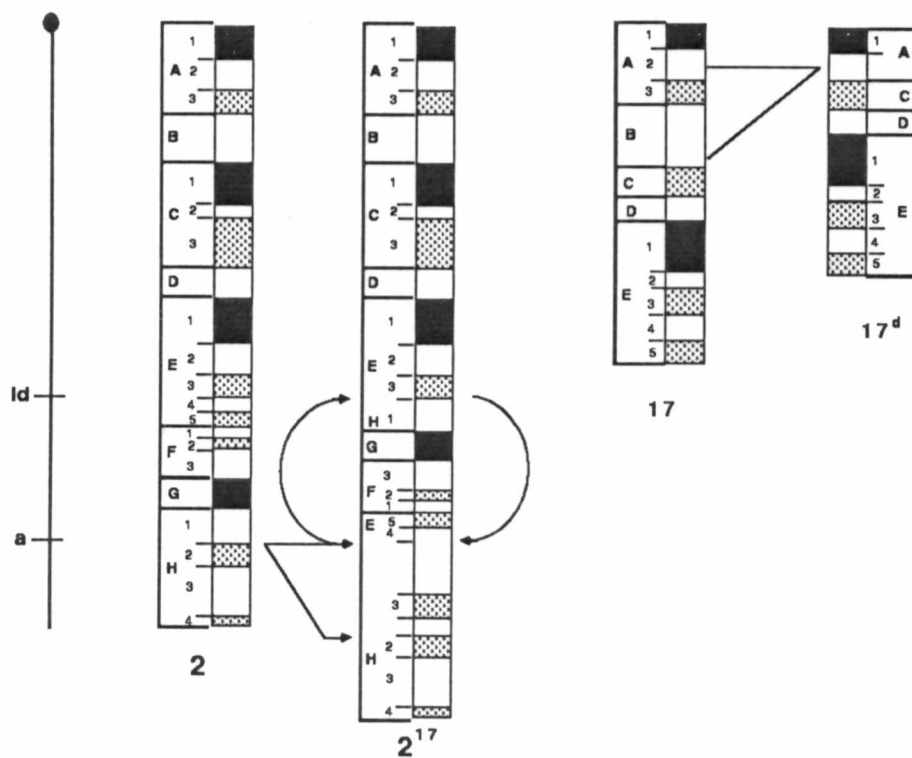


Figure 2. RNA expression from a complex gene located at the *ld* locus.

(A) Structure of mRNAs expressed from this gene as derived from overlapping cDNA clones (R.P.W. et al., unpublished results). Solid box represents a putative protein coding region; narrow lines represent noncoding regions. The designation A_n refers to the sites of polyadenylation. (B) Northern blot of Poly(A)⁺ RNA from the adult brain (B), kidney (K), testes (T), salivary gland (S), and small intestine (I) of the radiation-induced homozygous mutant (left) and wild-type (right) animals. A total of 500 μ g of total RNA was oligo-dT selected and loaded into each lane of the gel. The probe was a 0.7-kb fragment of cDNA corresponding to the sequence immediately 5' to the first polyadenylation site (see A). Numerals refer to RNA size markers in kb.

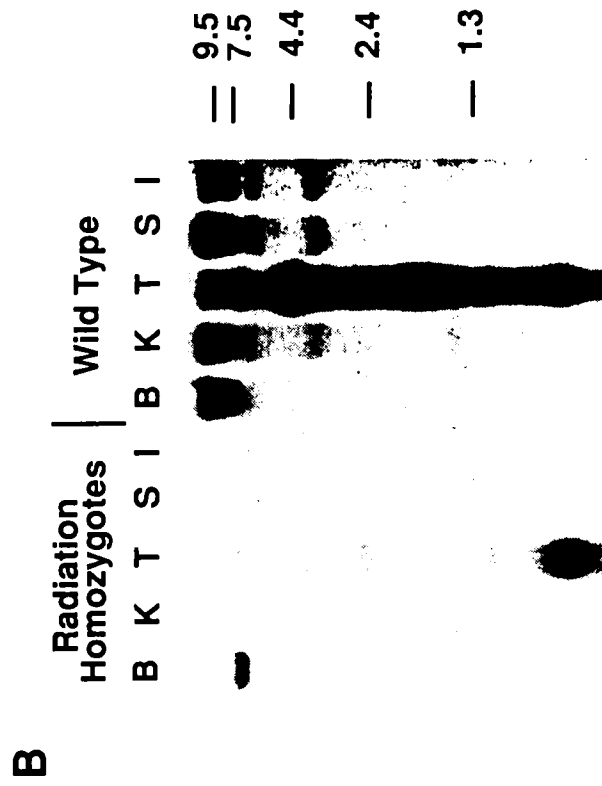
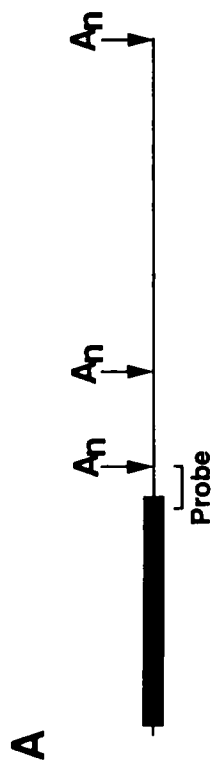


Figure 3. Identification of rearranged DNA fragments at the *ld* locus of the radiation-induced mutation.

(A) Genomic DNA, prepared from the liver of mice homozygous (lane 1) or heterozygous (lane 2) for the radiation-induced mutation, was digested with *Xba*I, *Hind*III, or *Eco*RI, transferred to a nylon membrane, and probed with the 0.7-kb fragment shown in Fig. 2A. Several genomic fragments were detected with this cDNA probe; arrows identify aberrant sized DNA fragments, unique to the mutant locus. Numerals refer to DNA size markers in kb. (B) Structure of the cloned 7.5-kb mutant, and corresponding normal 6.5-kb, *Eco*RI fragments. The boxed region at the 3' end of the 7.5-kb clone represents a segment of the rearranged portion of the *ld* locus in the mutant allele. The position of the 6.5-kb normal fragment on a larger region of genomic DNA is also shown along with the position of the radiation-induced breakpoint in the mutant allele and the transgene integration site in the insertional mutant, *ld*^{Hd} (Woychik et al., 1985). Small solid boxes represent two exons of the gene on the genomic DNA (R.P.W. et al., unpublished results).

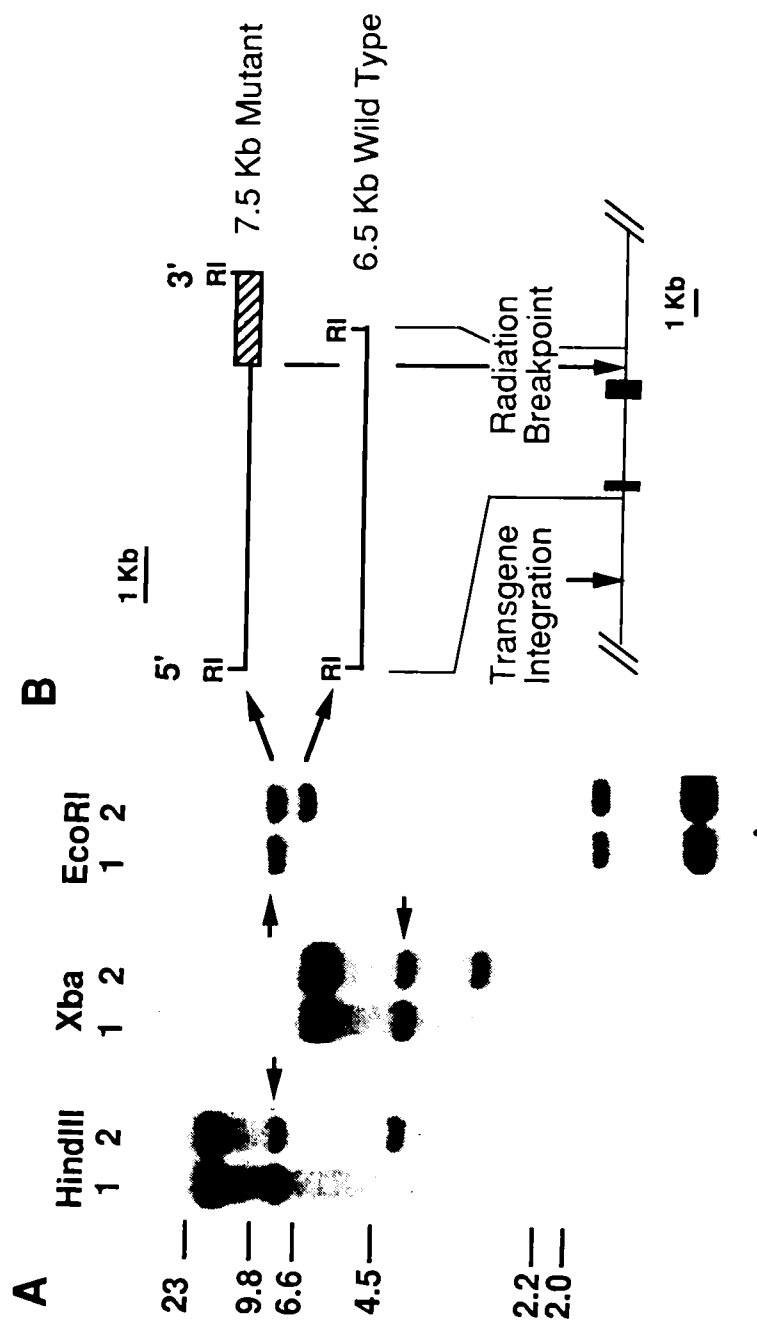


Figure 4. Genetic mapping of the *Is1Gso* breakpoint sequence.

Analysis of 93 N2 progeny derived from a (C57BL/10Nimr-+/Sey^{MH} X *Mus spretus*)F1 X C57BL/10Nimr backcross described (van der Meer-de Jong et al., 1990; R. van der Meer-de Jong et al., unpublished results). Genotypes at the *agouti* locus were scored by examination of coat color (agouti versus black [nonagouti]). For the genotype analysis of the limb deformity locus, genomic DNA samples were digested with *HindIII* and probed with a 1.5-kb *Sau3A-PvuII* fragment located approximately 7-kb 3' to the transgene integration site in the *ld^{Hd}* allele (Woychik et al., 1985). For genotyping of the breakpoint sequence, samples were digested with *MspI* and probed with a 0.2-kb *PvuII* fragment (called "breakpoint sequence" in the figure) derived from the region 3' to the breakpoint on the cloned 7.5-kb mutant *EcoRI* shown in Fig. 3B. "B" indicates the C57BL/10 allele, and "S" indicates the *Mus spretus* allele. Each column represents the genetic composition of the chromosome inherited from the F₁ parent for the three loci. The numeral above each column represents the number of progeny that inherited a given chromosome.

No. of Backcross Animals

48	32	8	5	0	0
----	----	---	---	---	---

limb deformity

S	B	B	S	S	B
---	---	---	---	---	---

agouti

S	B	S	B	S	B
---	---	---	---	---	---

breakpoint sequence

S	B	S	B	B	S
---	---	---	---	---	---

PART III

MOLECULAR CHARACTERIZATION OF A REGION OF DNA ASSOCIATED WITH MUTATIONS AT THE AGOUTI LOCUS IN THE MOUSE

ABSTRACT

Molecular characterization of a radiation-induced agouti (α) locus mutation allowed us to isolate a segment of DNA that maps at or near the α locus on chromosome 2 in the mouse. This region of DNA is deleted in several radiation- and chemical-induced homozygous-lethal α -locus mutations and is associated with specific DNA structural alterations in two viable α -locus mutations. We propose that DNA probes from this region of chromosome 2 will be useful for ultimately characterizing the individual gene or genes associated with α -locus function.

INTRODUCTION

The agouti (α) locus in chromosome 2 of the mouse controls the production of two distinct types of pigment, phaeomelanin (yellow) and eumelanin (black or brown, depending on alleles present at other loci), by the melanocytes within the hair follicle. Under control of the α locus, individual hairs may contain only eumelanin, only phaeomelanin, or a subapical band of phaeomelanin in an otherwise eumelanotic hair (agouti type). The α locus also regulates the distribution of phaeomelanotic, eumelanotic, or agouti hair over the surface of the animal; certain genotypes show a marked variation in dorso-ventral pigmentation, whereas others exhibit differential pigment patterns in various regions, including the perineum, pinnae, nipples, flanks, and tail. Interestingly, the regulatory signal produced by the α locus to control pigmentation does not originate within the melanocytes themselves but is transmitted from the follicular environment within which the melanocytes reside (Silvers, 1958a; 1958b; 1961; 1979; Silvers and Russell, 1955).

To date, 18 dominant and recessive alleles and pseudoalleles have been described for the α locus (Green, 1989), and numerous other alleles, some associated with chromosome rearrangements, have been recovered in mutagenesis studies (Woychik et al., 1990; Russell et al., 1990; L.B.R., unpublished). Genetic data can be interpreted as evidence for a complex, multiallelic gene that is regulated differently in each body region or as

evidence for the existence of several closely linked genes that work in concert (Silvers, 1979). A^y , lethal-yellow, described in 1905 (and the first lethal gene ever reported in a mammal) (Cuenot, 1905), has been shown to recombine with a number of α -locus alleles-namely α^x (lethal nonagouti) (Russell et al., 1963), α^t (black-and-tan) (Siracusa et al., 1987a), and α (nonagouti) (Siracusa et al., 1987a)-the recombination frequency in each case being of the order of 0.1%, with A^y lying proximal to α^x (Russell et al., 1963; Siracusa et al., 1987a). One crossover product in each case (the presumed double wild type) was indistinguishable from A ; the other product (presumed double mutant) has not been identified. The finding that among spontaneous α -locus mutations, certain ones (specifically, α to α^t and α to A^w) are common, whereas others are quite rare, also suggests that " α -locus alleles" may involve a complex of genetic sites, some of which, as is the case for A^y , are separable by recombination.

Although considerable effort has been expended in investigating the genetics and biology of the α -locus region, the structure and expression of the gene(s) remain unknown. The endogenous ecotropic murine leukemia virus *Emv-15* has been localized very close to A^y , and probes flanking *Emv-15* have been used for physical mapping experiments (Copeland et al., 1983; Lovett and Epstein, 1987; Lovett et al., 1987; Siracusa et al., 1987a; Siracusa et al., 1987b; Barsh and Epstein, 1989a; Barsh and Epstein, 1989b; Siracusa et al., 1989). However, the results of certain presumed spontaneous crossover events indicate that *Emv-15* does not lie between A^y and α (Russell, 1989).

We have earlier reported the derivation of a DNA probe associated with a radiation-induced α -locus mutation (Woychik et al., 1990). Here we describe how we have used this DNA probe to detect genomic alterations in chromosomes carrying mutations that express the α^j phenotype (α^j , jet-black, is probably equivalent to α^e , extreme nonagouti). The data we present support the hypothesis that this region of DNA is directly associated with the α locus, and that cloned segments of DNA from this region will be useful for the ultimate characterization of the individual gene or genes directly associated with α -locus function.

MATERIALS AND METHODS

Animals

All mice were bred at our laboratory in stocks propagating induced mutations. Five of the mutations were induced at Oak Ridge (Woychik et al., 1985; L.B.R., W.L. Russell, and P.A. Hunsicker, unpublished); one was induced at Harwell (Lyon et al., 1985).

Southern Blotting

Ten μ g of genomic adult liver or tail DNA was digested with restriction enzymes, electrophoresed through agarose gels, and blotted to GeneScreen (Dupont) with standard procedures (Ausubel et al., 1988; Maniatis et al., 1982). Radiolabeled hybridization probes were prepared with the random-hexamer labeling technique (Feinberg and Vogelstein, 1984).

Cloning of a 5.5-kb Wild-type Fragment From a Region Corresponding to the Distal Inversion Breakpoint of *Is1Gso*

A total of 250 μ g of genomic adult-liver DNA from an *Is1Gso* heterozygote (Woychik et al., 1990) was digested to completion with *EcoRI* and size-fractionated on a 10-40% sucrose gradient (Maniatis et al., 1982). An EMBL4 phage library, prepared with fractions containing 5- to 6-kb fragments, was screened with standard procedures utilizing a 0.2-kb *PvuII* fragment from the cross-hatched region of the 7.5-kb *EcoRI* rearranged fragment corresponding to probe A in Fig. 1A (Maniatis et al., 1982; Ausubel et al., 1988; Woychik et al., 1990). The 5.5-kb *EcoRI* fragment was then subcloned into the pGEM4 vector (Promega), restriction mapped, and partially sequenced as described (Maniatis et al., 1982; Ausubel et al., 1988).

Cloning a Rearranged Fragment Associated with the α^{5MNU} Mutation

A total of 500 μ g of genomic tail DNA from an $\alpha^{5MNU}/\alpha^{5MNU}$ homozygote was partially digested with *Sau3A* and size-fractionated on a 10-40% sucrose gradient (Maniatis et al., 1982). An EMBL3 phage library

was prepared with fractions containing 15- to 20-kb fragments and was screened by standard procedures (Maniatis et al., 1982; Ausubel et al., 1988) utilizing a 1.0-kb *EcoRI-XbaI* fragment from the distal end of the 5.5-kb *EcoRI* fragment as a probe (see probe B, Fig. 1A). A rearranged 3.5-kb *EcoRI* fragment was subcloned into pGEM4 (Promega), restriction mapped, and partially sequenced as described (Maniatis et al., 1982; Ausubel et al., 1988). Oligonucleotide primers were synthesized to facilitate the sequencing of the deletion breakpoint region.

Cloning a Large Region of DNA at or Near the Agouti Locus

A total of 250 µg of genomic liver DNA was partially digested with *Sau3A* and size-fractionated on a 10-40% sucrose gradient (Maniatis et al., 1982). Fractions containing 35- to 45-kb fragments were ligated into the cosmid vector c2RB (Bates and Swift, 1983), packaged *in vitro*, and screened by standard procedures using a 0.2-kb *PvuII* fragment from the hatched region of the 7.5-kb *EcoRI* rearranged fragment corresponding to probe A in Fig. 1A.

RESULTS

Identification of a DNA Probe Deleted in Several Radiation- or Chemical-Induced α -Locus Mutations

We previously described the genetic, cytogenetic, and molecular characterization of a new radiation-induced mutation that simultaneously gave rise to new alleles at the limb deformity (*ld*) and α loci, two regions normally separated by 20 cM on mouse chromosome 2 (Woychik et al., 1990). This mutation, Is(17;In2)*ld, α* 1Gso (abbreviated Is1Gso), contains an inversion of chromosome 2 that involves a segment lying between the limb deformity (*ld*) locus and a site that maps at or near the α locus (Woychik et al., 1990). At the proximal end of the inversion, the breakpoint in the DNA caused a physical disruption and alteration in the expression of the gene giving rise to the formin family of mRNAs. The disruption of this gene appears to be associated with the limb-deformity defect in this and the *ld^{Hd}* mutant (Maas et al., 1990; Woychik et al., 1990). Since the extreme

nonagouti phenotype arose simultaneously with the limb-defect phenotype in Is1Gso, we reasoned that it was possible that the DNA breakpoint at the distal end of the inversion had caused a physical disruption and alteration in the expression of a gene directly associated with the nonagouti trait in this mutant.

To begin to test the possibility that the distal inversion breakpoint in Is1Gso might be associated with a gene at the α locus, we utilized a DNA probe (probe A in Fig. 1A), consisting of a sequence adjacent to the proximal DNA breakpoint (hatched segment of the 7.5-kb fragment in Fig. 1A), to determine whether this region of the genome is altered in four radiation- or chemical-induced agouti mutations. These four mutations, α^{jl10} , α^{jl41} , α^{jl85} and α^{9H} , all express the α^j (α^j , jet-black; probably equivalent to α^e , extreme nonagouti) phenotype and are lethal in the homozygous condition (Table 1). Viable compound heterozygotes were produced for this experiment by balancing each of these alleles with Is1Gso. One DNA breakpoint of Is1Gso is near the segment that was used as a probe, so the Is1Gso allele provided us with an *Eco*RI restriction fragment length variant (RFLV) for this region of DNA; the Is1Gso allele is associated with a 7.5-kb *Eco*RI fragment, while the wild-type allele appears as a 5.5-kb *Eco*RI fragment. Therefore, this RFLV allowed us readily to determine whether the region identified by the probe was deleted in these agent-induced mutations.

Analysis of *Eco*RI-digested DNA from the above Is1Gso/ α^{jl} compound heterozygotes, as expected, revealed that each contained the 7.5-kb Is1Gso-specific fragment; none of the samples yielded a wild-type size 5.5-kb fragment, or any other detectable fragment, associated with the agent-induced allele (Fig. 1B, lanes A-D). To rule out the unlikely possibility that the induced allele in each of the four mutations might appear as a 7.5-kb fragment that co-migrates with the Is1Gso-specific fragment, we analyzed DNA from heterozygotes containing a balancing chromosome with a normal-sized 5.5-kb fragment (Fig. 1B, lanes H-L). No fragments of a length other than 5.5 kb were detectable. These data clearly indicate that all four of these α -locus mutations contain a deletion in the region corresponding to the distal inversion breakpoint in Is1Gso and supports the hypothesis that this section of chromosome 2 contains at least one gene associated the α locus.

Characterization of the Structure of the Distal Inversion Breakpoint in Is1Gso

The above data indicate that large radiation- and chemical-induced deletion mutations are useful for initially mapping DNA probes to the general vicinity of a particular gene on the genome. However, such mutations are often associated with the deletion of hundreds, thousands, or perhaps even millions of base-pairs of DNA sequence and, therefore, may be deficient in multiple genes, not all of which are directly related to the phenotype of interest (Rinchik and Russell, 1990). Viable mutations, on the other hand, are potentially more useful than large deletions for locating specific genes in the DNA because any structural alteration is more likely to affect only the expression of the gene directly associated with the phenotype under consideration. For this reason, we subsequently focused our attention on the analysis of viable nonagouti mutations in an attempt to identify a gene associated with the function of the α locus.

Since the Is1Gso mutation is a viable allele at the α locus, we continued our analysis of this mutation, concentrating on the DNA structural alteration at the distal inversion DNA breakpoint. As described above, the initial hypothesis was that the distal inversion breakpoint itself caused a disruption of the expression of a gene directly associated with the function of the α locus. While this may have been the case, it was also possible that during the process of generating the chromosomal structural alterations in Is1Gso, a modest-sized segment of DNA had been deleted from the distal breakpoint region on chromosome 2. (It was unlikely that a large deletion had occurred since the mutation is viable.)

To begin to test this possibility, we cloned and characterized the wild-type 5.5-kb *Eco*RI fragment that hybridized with the probe used in the experiment described in Fig. 1. One end of this wild-type 5.5-kb *Eco*RI fragment contains all of the 1.9 kb of sequence located adjacent to the proximal inversion breakpoint in Is1Gso (corresponding to the hatched box at the end of the 7.5-kb rearranged fragment in Fig. 1A), whereas the other end of the fragment contains the remaining 3.6 kb of sequence. A probe prepared from the distal end of the remaining 3.6 kb of sequence (probe B,

Fig. 1A) detected mutant allele-specific RFLVs in *EcoRI*- and *PstI*- but not *BamHI*- digested genomic DNA of wild-type, heterozygous and homozygous animals from the Is1Gso line (Fig. 2). The Is1Gso RFLV detected with this probe occurs at a position within a small interval between a *PstI* and *BamHI* site on the 5.5-kb fragment (indicated with asterisks in Fig. 1A). Since this small interval occurs at approximately the same position as the structural alteration site detected on the 7.5-kb fragment-i.e., 1.9 kb from the proximal end of the 5.5-kb fragment-it appears that no significant deletion had occurred at the distal inversion breakpoint site of Is1Gso. In fact, in additional experiments to be described elsewhere, DNA sequence analysis of cloned fragments from this region revealed that only 29 base pairs (bp) of sequence were deleted from the distal inversion breakpoint site (S.J.B., R.P.W., unpublished results). No additional structural alterations were detected in this region of the mutant allele. This result supports our prediction that a gene associated with the function of the α locus has been disrupted by the breakage of the DNA at the distal inversion site of Is1Gso.

Characterization of a Structural Alteration in the α^{5MNU} Viable α -Locus Mutation

Major structural alterations like Is1Gso could exert position effects on genes that are not located at the actual DNA structural alteration site. Therefore, although we have demonstrated that no additional major structural alterations occurred in the region near the distal DNA inversion breakpoint of Is1Gso, it was possible that the expression of the gene or genes at the α locus might be affected from a distance, and that the α locus might not map within the immediate vicinity of the breakpoint region. For this reason, we examined other viable α -locus mutations with the DNA probes derived from the Is1Gso distal breakpoint region.

One viable mutation, α^{5MNU} , arose in a specific-locus experiment (W.L. Russell, unpublished results; Russell et al., 1990) six weeks after methylnitrosourea exposure of differentiating spermatogonia (Table 1). DNA from an $\alpha^{5MNU}/\alpha^{5MNU}$ homozygote, along with DNA from the C3H and 101 parental types, was digested with *EcoRI*, *PstI*, and *BamHI*. The samples were then hybridized with probe A or probe B, derived from either

the proximal or distal ends of the normal 5.5-kb fragment (Fig. 1A). Although the distal probe B hybridized to an identical *Bam*HI fragment in mutant and wild-type DNA, nonparental fragments were detected in the *Eco*RI and *Pst*I-digested mutant samples (Fig. 3). Also, a deletion of the region of DNA corresponding to the proximal probe (probe A) had occurred in this mutant (Fig. 3). This result indicated that, in a DNA structural rearrangement, which included a deletion of the proximal end of the normal 5.5-kb *Eco*RI fragment, had occurred between the *Pst*I and distal *Bam*HI sites on the 5.5-kb fragment in α^{5MNU} (sites shown with asterisk in Fig. 1A).

To characterize the nature of the structural alteration in the α^{5MNU} allele, the mutant 3.5-kb *Eco*RI fragment detected in the above analysis was cloned and compared with a large section of DNA derived from overlapping cosmid genomic clones from this region of the genome (Fig. 4A). DNA sequence analysis of the structurally altered region demonstrated that a small segment of DNA, estimated to be 2.8 kb in length, had been deleted from the α^{5MNU} mutant allele (Fig. 4B). No other major structural alterations were detected in the α^{5MNU} mutant allele within several kilobases of this deletion (S.J.B., R.P.W., unpublished results). Interestingly, the distal deletion breakpoint site of the α^{5MNU} occurs only 129 bp away from the site of the distal inversion breakpoint on the *Is1Gso* mutation (S.J.B., R.P.W., unpublished results).

DISCUSSION

Characterization of the molecular structure of the α locus has been the focus of a considerable amount of research by many investigators over the past several years (Barsh and Epstein, 1989a; Barsh and Epstein, 1989b; Copeland et al., 1983; Lovett et al., 1987; Lovett and Epstein, 1987; Russell, 1989; Siracusa et al., 1987a; Siracusa et al., 1987b; Siracusa et al., 1989). We have utilized a DNA probe derived from a radiation-induced extreme nonagouti mutation to characterize a region of DNA that has undergone structural changes in several α -locus mutations. Although several of these mutations are homozygous lethals and are deletions that

may be very large, we did identify specific DNA sequence alterations in two viable mutations, Is1Gso and α^{5MNU} . We hypothesize that the sequence alterations in these viable mutations are likely to define the position of at least one gene associated with α -locus function.

Interestingly, the Is1Gso and α^{5MNU} mutations contain a structural alteration at sites very close to each other in the DNA: the distal inversion breakpoint in Is1Gso is located within the region of DNA deleted in the α^{5MNU} allele, 129 bp from the distal α^{5MNU} deletion breakpoint. The fact that this region is deleted in several lethal mutations and has undergone a structural alteration at almost the same site in two viable mutations provides compelling evidence that this region contains at least a portion of the α -locus gene(s).

Molecular characterization of the agouti region should ultimately reveal whether the function of the α locus is associated with a single, large gene that undergoes complex regulation, or whether there are multiple linked genes that undergo a coordinated pattern of expression. Additionally, cloning of cDNAs and the analysis of RNA transcripts arising from the gene(s) at the α locus, coupled with *in situ* hybridization techniques, may also aid in understanding the molecular basis for the complex biology associated with the production of the banded agouti pattern of coat pigmentation.

The agouti region has also been of interest because some of the alleles or pseudoalleles are associated with an increased susceptibility to the development of a variety of spontaneous and induced tumors (Heston and Vlahakis, 1961; Silvers, 1979; Wolff et al., 1986), a stimulation in normal body growth (Heston and Vlahakis, 1961), obesity (Dickerson and Gowen, 1947; Heston and Vlahakis, 1961; Wolff et al., 1986), a diabetes-like syndrome (Coleman, 1982), reduced fertility (Granholt et al., 1986), and embryonic death (Robertson, 1942; Calarco and Pederson, 1976; Papioannou and Mardon, 1983). The lethality associated with A^y has been investigated in some detail; embryos homozygous for this mutation display characteristic abnormalities at the morula or blastocyst stage and die during implantation (Robertson, 1942; Calarco and Pederson, 1976; Papioannou and Mardon, 1983). Therefore, in addition to understanding the molecular

basis for the agouti coat color, the characterization of the agouti locus at the molecular level should also provide useful information for understanding the molecular basis of many or all of these additional phenotypes.

REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1988). *Current Protocols in Molecular Biology*. John Wiley & Sons/Greene, New York.
- Barsh, G. S., and Epstein, C. (1989a). Physical and genetic characterization of a 75-kilobase deletion associated with a^l , a recessive lethal allele at the mouse *agouti* locus. *Genetics* **121**: 811-818.
- Barsh, G. S. and Epstein, C. (1989b). The long-range restriction map surrounding the mouse *agouti* locus reveals a disparity between physical and genetic distances. *Genomics* **5**: 9-18.
- Bates, P. F., and Swift, R. A. (1983). Double cos site vectors: simplified cosmid cloning. *Gene* **26**: 137-146.
- Calarco, P. G., and Pedersen, R. A. (1976). Ultrastructural observations of lethal yellow (A^y/A^y) mouse embryos. *J. Embryol. Exp. Morphol.* **35**: 73-80.
- Cattanach, B. M., Lyon, M. F., Peters, J., and Searle, A. G. (1987). *Mouse News Lett.* **77**: 123-125.
- Coleman, D. L. (1982). Diabetes-obesity syndromes. In *The mouse in biomedical research IV* (eds. H. L. Foster, J. D., Small, and J. G., Fox), pp. 125-132. Academic Press, New York.
- Copeland, N. G., Jenkins, N. A., and Lee, B. K. (1983). Association of the lethal yellow A^y coat color mutation with an ecotropic murine leukemia virus genome. *Proc. Natl. Acad. Sci. USA* **80**: 247-249.
- Cuenot, L. (1905). Les races pures et leurs combinaisons chez les souris. *Arch. Zool. Exp. Gen.* **3**: 123-132.

- Dickerson, G. E., and Gowen, J. W. (1947). Hereditary obesity and efficient food utilization in mice. *Science* **105**: 496-498.
- Feinberg, A. P., and Vogelstein, B. (1984). Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**: 266-267.
- Granholm, N. H., Jeppesen, K. W., and Japs, R. A. (1986). Progressive infertility in female lethal yellow mice. *J. Reprod. Fertil.* **76**: 279-287.
- Green, M. C. (1966). in *Biology of the Laboratory Mouse*, 2ND ed. (Green, E.L. ed.), (McGraw-Hill, New York), pp 87-150.
- Green, M. C. (1989). Catalogue of mutant genes and polymorphic loci. In *Genetic variants and strains of the laboratory mouse*, (eds., M. F. Lyon, and A. G. Searle), pp. 17-20. Oxford, New York: Oxford University Press, Oxford and New York.
- Heston, W. E., and Vlahakis, G. (1961). Influence of the *A^y* gene on mammary-gland tumors, hepatomas, and normal growth in mice. *J. Natl. Cancer Inst.* **26**: 969-983.
- Lovett, M., Cheng, Z. Y., Lamela, E. M., Yokoi, T., and Epstein, C. (1987). Molecular markers for the *agouti* coat color locus of the mouse. *Genetics* **115**: 747-754.
- Lyon, M. F., Fisher, G., and Glenister, P. H. (1985). A recessive allele of the mouse *agouti* locus showing lethality with yellow, *A^y*. *Genet. Res. Camb.* **46**: 95-99.
- Maas, R. L., Zeller, R., Woychik, R. P., Vogt, T. F., and Leder, P. (1990). Disruption of formin-encoding transcripts in two mutant *limb deformity* alleles. *Nature* **346**: 853-855.

- Maniatis, T., Fritsch, E. F., Sambrook, J. (1982). *Molecular cloning: A laboratory manual*. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York).
- Papioannou, V. E. and Mardon, H. (1983). Lethal nonagouti (a^x): Description of a second embryonic lethal at the agouti locus. *Dev. Genet.* **4**: 21-29.
- Rinchik, E.M. and Russell, L.B. (1990) Germ-line deletion mutations in the mouse: Tools for intensive functional and physical mapping of regions of the mammalian genome. In *Genome analysis, vol. 1*, (eds. K. Davies, S. Tilghman), pp. 121-158. Cold Spring Harbor Press, Cold Spring Harbor, New York.
- Robertson, G. G. (1942). An analysis of the development of homozygous yellow mouse embryos. *J. Exp. Zool.* **89**: 197-231.
- Russell, L.B. (1989). Functional and structural analyses of mouse genomic regions screened by the morphological specific-locus test. *Mutation Research* **212**: 23-32.
- Russell, L. B., McDaniel, N. C., and Woodiel, F. N. (1963). Crossing-over within the a "locus" of the mouse. *Genetics* **48**: 907a.
- Russell, L. B., Russell, W. L., Rinchik, E. M., and Hunsicker, P.R. (1990) Factors affecting the nature of induced mutations. In *Biology of mammalian germ cell mutagenesis, Banbury report 34*, (eds. J. W. Allen, B. A. Bridges, M. F. Lyon, M. J. Moses, and L. B. Russell), pp. 271-285. Cold Spring Harbor Lab., Cold Spring Harbor, New York.
- Silvers, W. K. (1958a). An experimental approach to action of genes at the agouti locus in the mouse. II. Transplants of newborn aa ventral skin to $a^t a$, $A^w a$, and aa hosts. *J. Exp. Zool.* **137**: 181-188.

- Silvers, W. K. (1958b). An experimental approach to action of genes at the agouti locus in the mouse. III. Transplants of newborn A^w -, A -, a^t -skin to A^y -, A^w -, A - and aa hosts. *J. Exp. Zool.* **137**: 189-196.
- Silvers, W. K. (1961). Genes and pigment cells of mammals. *Science* **134**: 368-373.
- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6-44. Springer-Verlag, New York.
- Silvers, W. K., and Russell, E. S. (1955). An experimental approach to action of genes at the agouti locus in the mouse. *J. Exp. Zool.* **130**: 199-220.
- Siracusa, L. D., Russell, L. B., Jenkins, N. A., and Copeland, N. G. (1987a). Allelic variation within the *Emv-15* locus defines genomic sequences closely linked to the *agouti* locus on mouse chromosome 2. *Genetics* **117**: 85-92.
- Siracusa, L. D., Russell, L. B., Eicher, E. M., Corrow, D. J., Copeland, N. G., and Jenkins, N. A. (1987b). Genetic organization of the *agouti* region of the mouse. *Genetics* **117**: 93-100.
- Siracusa, L. D., Buchberg, A. M., Copeland, N. G., and Jenkins, N. A. (1989). Recombinant inbred strain and interspecific backcross analysis of molecular markers flanking the murine *agouti* coat color locus. *Genetics* **122**, 669-679.
- Wolff, G. L., Roberts, D. W., and Galbraith, D. B. (1986). Prenatal determination of obesity, tumor susceptibility, and coat color pattern in viable yellow (A^{vy}/a) mice. *J. Hered.* **77**: 151-158.

- Woychik, R. P., Generoso, W. M., Russell, L. B., Cain, K. T., Cacheiro, N. L. A., Bultman, S. J., Selby, P. B., Dickinson, M. E., Hogan, B. L. M., 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. USA* **87**: 2588-2592.

APPENDIXES

APPENDIX A**TABLES**

Table 1. Radiation- and chemical-induced agouti-locus mutations that exhibit detectable structural alterations.

Mutation (<i>a</i> *)	Mutagen	Germ-Cell Type†	Phenotype of		Structural Alteration	Ref.
			<i>a</i> */ <i>a</i> *	<i>a</i> */ <i>a</i> j##		
<i>a</i> ^l l0	X rays	Oocyte	lethal	jet	Deletion	§
<i>a</i> ^l 41	Ethylmethanesulfonate	Spermatozoa	lethal	jet	Deletion	§
<i>a</i> ^l 85	Gamma rays	Oocyte	lethal	jet	Deletion	§
<i>a</i> ⁹ HΩ	X rays	Oocyte	lethal	jet	Deletion	20
Is1Gso	Gamma rays	Spermatogonia	jet	jet	Inversion	7
<i>a</i> ⁵ MNU	Methylnitrosourea	Spermatogonia	jet	jet	Deletion	ψ

*a** Radiation- or chemical-induced mutant agouti allele.

† Indicates the specific germ-cell stage exposed to mutagen.

a^j, jet black, is the most recessive *a*-locus allele and is probably equivalent to *a*^e, extreme nonagouti.

20 Russell, 1989

7 Woychik et al., 1990

§ L. B. Russell, unpublished results.

Ω *a*⁹H is the Harwell designation for *a*^l (Cattanach et al., 1987)

ψ W. L. Russell, and P. R. Hunsicker, unpublished results.

APPENDIX B

FIGURES

Figure 1. Southern blot analysis of radiation- or chemical-induced extreme nonagouti mutations.

(A) Structure of a 7.5-kb rearranged fragment from the Is1Gso radiation-induced mutation containing the proximal inversion breakpoint, and the structure of a 5.5-kb wild-type fragment mapping at or near the α locus.

The 7.5-kb fragment contains a portion of the *ld* locus (filled-box) juxtaposed with a section of DNA (cross-hatched) that we previously mapped to a region at or near the α locus (Woychik et al., 1990). The 5.5-kb wild-type fragment was cloned utilizing a probe from the cross-hatched region of the 7.5-kb fragment (region corresponding to probe A). Comparison of the two cloned fragments revealed that the inversion DNA breakpoint on the 7.5-kb fragment was approximately 1.9 kb from the proximal end of the 5.5-kb fragment. B, *Bam*HI; P, *Pst*I; X, *Xba*I; R, *Eco*RI; H, *Hind*III. Probes A and B correspond to proximal and distal regions of the 5.5-kb fragment and are represented by cross-hatched boxes.

Figure 1. Southern blot analysis of radiation- or chemical-induced extreme nonagouti mutations.

(B) Southern blot of *Eco*RI-digested genomic DNA from several radiation- or chemical-induced extreme nonagouti mutations hybridized with probe A from the 5.5-kb fragment (see A).

Individual mutations are described in Table 1. Lanes A and H, α^{l10} ; lanes B and I, α^{l41} ; lanes C and J, α^{l85} ; lanes D and K, α^{9H} . The four individual radiation- or chemical-induced mutations were analyzed as compound heterozygotes balanced with Is1Gso (lanes A-D) or heterozygotes balanced with either a wild-type chromosome (lanes H-J) or the α^{16H} allele (Lyon et al., 1985), which contains a normal-size 5.5-kb fragment in this region of DNA (lane K). (The α^{9H} mutation was balanced with α^{16H} in lane L to verify that the α^{16H} mutation contains a 5.5-kb fragment in this region.) Lanes E-G, homozygous, heterozygous, and wild-type DNA, respectively, from the Is1Gso mutant line. Hybridizing fragments corresponding to the Is1Gso-specific fragment, or to a wild-type-size 5.5-kb specific fragment, are indicated.

B.

A B C D E F G H I J K L

7.5 kbs1Gso→



5.5 kb Normal→



Figure 2. Southern blot of wild-type (lane A), heterozygous (lane B), or homozygous (lane C) adult genomic DNA from the Is1Gso mutant line.

Each sample was digested with either *Eco*RI, *Pst*I, or *Bam*HI and then hybridized with probe B (Fig. 1A). Numbers refer to sizes of DNA fragments in kb; asterisks denote rearranged Is1Gso fragments.

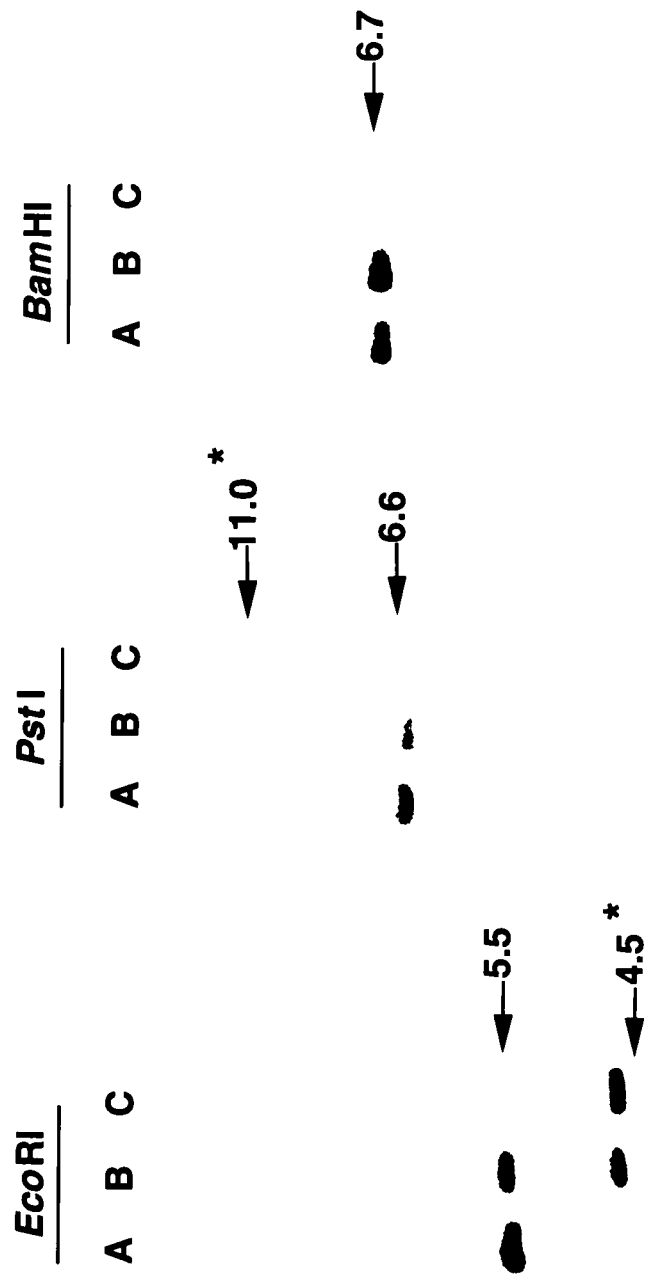


Figure 3. Southern blot of C3H (lane A), 101 (lane B) or homozygous α^{5MNU} (lane C) adult genomic DNA digested with *EcoRI*, *PstI*, or *BamHI*.

Each sample was hybridized with either probe A or probe B from the proximal and distal ends of the cloned 5.5-kb normal fragment (Fig. 1A). Numerals refer to the sizes of the DNA fragments, and asterisks denote rearranged α^{5MNU} fragments identified with the distal probe.

Distal Probe B

EcoRI

A B C

PstI

A B C

BamHI

A B C

6.7*
6.6

6.7

←5.5

←3.5*

Proximal Probe A

EcoRI

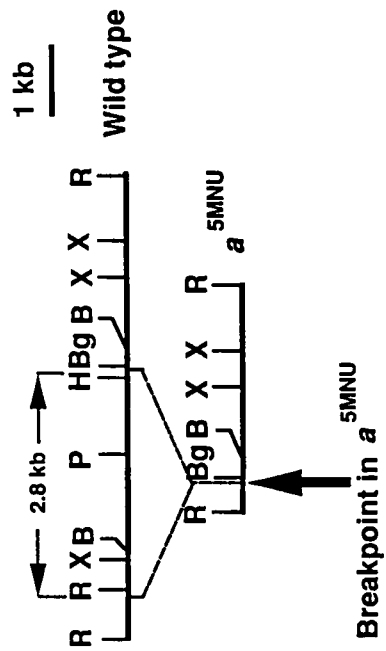
A B C

←5.5

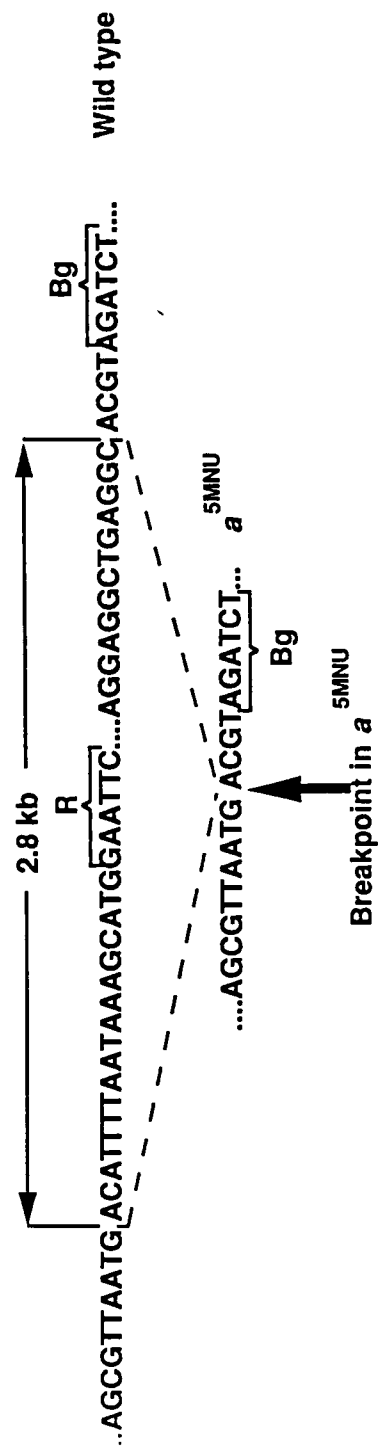
Figure 4. Fine structure analysis of the DNA lesion in the α^{5MNU} mutation.

(A). Structure of the 3.5 kb *Eco*RI fragment cloned from the α^{5MNU} mutation compared with the corresponding structure of a cloned segment of wild-type DNA. The presence of a 2.8 kb deletion in the 3.5 kb 5MNU fragment is indicated with dashed lines. B, *Bam*HI; P, *Pst*I; X, *Xba*I; R, *Eco*RI; Bg, *Bgl*II; H, *Hind*III. (B) Nucleotide sequence analysis of the α^{5MNU} breakpoint region compared with the sequence of wild-type DNA.

A.



B.



PART IV

**MOLECULAR CHARACTERIZATION OF THE MOUSE AGOUTI
LOCUS**

ABSTRACT

The agouti locus acts within the microenvironment of the hair follicle to regulate coat color pigmentation in the mouse. We have characterized a gene encoding a novel 131 amino acid protein which we propose is the one gene associated with the agouti locus. This gene is normally expressed in a manner consistent with α -locus function and, most importantly, its structure and expression is affected by a number of representative alleles in the agouti dominance hierarchy. Additionally, we found that the pleiotropic effects associated with the lethal yellow (A^Y) mutation, which include pronounced obesity, diabetes, and the development of neoplasms, are accompanied by deregulated overexpression of the agouti gene in numerous tissues of the adult animal.

INTRODUCTION

The agouti (α) locus in chromosome 2 regulates the differential production of black and yellow pigment granules that give rise to the agouti coat color of the mouse. Agouti coloration, which is the true wild-type coat color of mice, is unusual in that it arises not from a homogeneous pigmentation of the pelage, but rather from a banded coloration pattern in which each hair is black with a subapical band of yellow. One of the most interesting aspects of the agouti locus is that it functions within the microenvironment of the hair follicle (Silvers, 1958a, 1958b, 1961, 1979; Silvers and Russell, 1955), unlike many other coat color genes that act in a cell-autonomous manner within the melanocytes. Therefore, agouti must be regulating coat pigmentation by some direct or indirect form of intercellular signaling within the follicular environment.

Like many other genes that play a role in the regulation of coat pigmentation in the mouse, the agouti locus contributes to essential developmental processes unrelated to pigmentation (Geissler et al., 1988; Witte, 1990 and references therein; Epstein et al., 1991; Mercer et al., 1991). For example, some of the individual alleles at the agouti locus are associated with embryonic lethality, obesity, diabetes, and the development

of tumors in a wide variety of tissues. In fact, the lethal yellow (A^y) mutation at agouti was the first embryonic lethal mutation to be characterized in the mouse (Cuenot, 1905). Embryos homozygous for A^y die very early in development, around the time of implantation, possibly due to a defect in trophoblast differentiation (Eaton and Green, 1963; Calarco and Pederson, 1976; Papaioannou et al., 1979).

Genetic analyses of numerous α -locus mutants have been ongoing for nearly a century and have led to the identification of at least eighteen dominant and recessive alleles and pseudoalleles of agouti (Sillers, 1979; Green, 1989). Different combinations of alleles account for an array of different phenotypes, ranging from subtle differences in coat color compared to wild type, to drastic changes in the distribution of pigmentation in different regions of the animal, particularly across the dorso-ventral surface. An intricate dominance hierarchy exists in which alleles associated with pheomelanin (yellow) production are generally dominant over alleles associated with eumelanin (black or brown depending on alleles at other loci) production (Sillers, 1979). This relationship is exemplified by several alleles that date back to the mouse fancy: lethal yellow (A^y), which confers an all-yellow phenotype in the heterozygous condition, black-and-tan (a^t), which gives rise to an all-black dorsum and an all-yellow ventrum (Dunn, 1928), nonagouti (a), which gives rise to a predominantly black phenotype except for small amounts of pheomelanin around the pinnae, nipples, and perineum (Sillers, 1979), and extreme nonagouti (a^e), which confers a completely black phenotype (Hollander and Gowen, 1956).

Genetic data can be interpreted as evidence for the agouti locus being comprised of either a single gene, which is regulated differently in each body region, or several closely linked genes which work in concert (reviewed in Silvers, 1979). In order to distinguish between these alternate hypotheses and to ultimately elucidate its mode of action, the agouti locus must be studied at the molecular level. Although the agouti alleles have been extensively characterized with classical genetic techniques, the structure of the gene(s) responsible for α -locus function have not yet been determined.

We previously described a radiation-induced inversion mutation, called $Is(17;In2)ld,a^j1Gso$ (abbreviated $Is1Gso$), which contains DNA breakpoints

in the limb deformity (*ld*) and agouti loci, two regions that are normally separated by 22 cM on chromosome 2 (Woychik et al., 1990a; Bultman et al., 1991). Utilizing a DNA probe from the *ld^{Hd}* insertional mutant (Woychik et al., 1985), we were able to jump 22 cM with the inversion and clone a region of DNA that maps to the agouti locus (Woychik et al., 1990a). Moreover, this cloned region also hybridizes to sequences that are rearranged in several agent-induced α -locus mutations (Bultman et al., 1991). Here, we present the molecular characterization of a gene from this region of DNA that is structurally altered in several α -locus mutations and is expressed in a manner appropriate for agouti function. We propose that this gene is, in fact, the one gene associated with the agouti locus.

MATERIALS AND METHODS

Animals

All mice were maintained at the Oak Ridge National Laboratory (ORNL) including the SB B+T stock, derived from a mutation that arose spontaneously at ORNL in an SEC/E x C57BL/E mating.

DNA Isolation and Southern Blot Analysis

Ten micrograms of genomic DNA was digested with restriction enzymes, electrophoresed through agarose gels, and blotted to GeneScreen (DuPont) utilizing standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). Radiolabeled hybridization probes were prepared with the random-hexamer labeling technique (Feinberg and Vogelstein, 1984). Prior to hybridization, probe 1.5 was reassociated with sheared, genomic mouse DNA to prevent the hybridization of repetitive sequence (Sealey et al., 1985). Post-hybridization filter washing was conducted under high stringency (0.2X SSC, 0.1% SDS at 68°C) or reduced stringency (0.2X SSC, 0.1% SDS at 50°C) conditions for membranes containing DNA from mice or other mammalian species, respectively.

RNA Preparation and Northern Blot Analysis

Total RNA was prepared using the guanidine isothiocyanate procedure (Ausubel et al., 1988), enriched for poly(A)⁺ RNA using an oligo(dT)-cellulose column (Aviv and Leder, 1972), electrophoresed through formaldehyde gels, and blotted to GeneScreen (DuPont) utilizing standard procedures (Ausubel et al., 1988). The membrane was hybridized and washed under high stringency conditions as described above for Southern blot analysis.

Isolation of cDNA

Poly(A)⁺ RNA was prepared from day five neonatal skin (C3H strain) or adult *A^y/Is1Gso* kidney or testis as described above, and double-stranded cDNA was subsequently prepared with reverse transcriptase using standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). After the addition of *Eco*RI linkers, the cDNA was ligated into the λ gt10 vector (Stratagene), packaged in vitro, and screened with Probe B (Figure 1A) for the neonatal skin cDNA library or the wild-type cDNA clone (for the *A^y/Is1Gso* cDNA libraries) using standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). Positive clones were purified using conventional methods and subcloned into pGEM (Promega) or pBluescript (Stratagene) for further analysis.

Isolation of Genomic Clones

Genomic spleen DNA from the strain 129/R1 was partially digested with *Sau*3A and size-fractionated on a 10-40% sucrose gradient (Sambrook et al., 1989). Fractions containing 35- to 45-kb fragments were ligated into the cosmid vector c2RB (Bates and Swift, 1983), packaged in vitro, and screened using standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). Utilizing the cDNA clone as a probe, cosmid subfragments were isolated and subcloned into pGEM (Promega) or pBluescript (Stratagene) by standard procedures (Sambrook et al., 1989).

DNA Sequencing

Genomic and cDNA clones were sequenced by the Sanger dideoxynucleotide method (Sanger et al., 1977) using T7 DNA polymerase

(USB) (Tabor and Richardson, 1987). Analysis of the DNA sequence was performed using the University of Wisconsin Genetics Computing Group (GCG) sequence analysis programs (Devereux et al., 1984).

RESULTS

Isolation of a cDNA Clone and Genomic Structure of the Agouti Locus

Having established from previous work that the distal inversion breakpoint of Is1Gso likely lies within the agouti locus (Bultman et al., 1991), a search was initiated to find an α -locus exon. Interspecific hybridizations (i.e. "zoo blots") identified a region of DNA near the distal inversion breakpoint of Is1Gso (Fig. 1A) that is conserved in a number of mammalian species (Fig. 1B). This evolutionarily conserved region was subsequently shown to be expressed based on its ability to hybridize to a 0.8-kb transcript in RNA prepared from neonatal skin of wild-type animals (Fig. 1C). Neonatal skin cDNA libraries were subsequently prepared and screened with probe B (Fig. 1A), resulting in the identification of several clones, one of which contained the entire coding region and is nearly full-length. The total size of this cDNA clone is 800 bp, which is comparable to the size of the mRNA seen on Northern blots. The complete nucleotide sequence is 692 bp, excluding the poly(A) tract (Fig. 2). The cDNA contains an open reading frame (ORF) extending from nucleotide 83 through 478, beginning with an ATG codon flanked by sequence that is in agreement with the consensus sequence for translation initiation (Kozak, 1987). The hexanucleotide AATAAT is present within the 3' untranslated sequence immediately preceding the poly(A) tract and probably represents the polyadenylation signal of the gene (Sheets et al., 1990; Durkop et al., 1992).

The translation product deduced from the ORF is 131 amino acids in length with a molecular weight estimated to be 15,000 daltons (Fig. 2). Searches of the NBRF and SWISSPROT databases using the algorithm FASTA (Pearson and Lipman, 1988) failed to identify any proteins with significant sequence homology. The N-terminus may comprise a signal peptide since it is hydrophobic (Kyte and Doolittle, 1982) and includes a possible cleavage site after residue 22 (von Heijne, 1986). A highly basic

domain containing 16 lysine or arginine residues in a stretch of 29 amino acids is present in the center of the predicted protein and is followed by a poly-proline stretch. At the carboxy-terminus, ten of the final forty amino acids are cysteines, and in four instances, cysteine residues are spaced six amino acids apart. Additionally, one putative N-linked glycosylation site and nine potential serine/threonine kinase phosphorylation sites are present within the predicted protein, five of which lie within the highly basic domain.

To elucidate the intron/exon composition of the gene corresponding to the cDNA clone presented in Fig. 2, overlapping genomic clones were isolated and characterized utilizing the cDNA clone as a probe. Comparing the structure of the genomic and cDNA clones revealed that the gene contains four exons (Fig. 3A). The first exon of the cDNA is 72 bp in length and composed entirely of 5' untranslated sequence. The first intron is approximately 11.5 kb in length and is followed by the second (170 bp), third (65 bp), and fourth (385 bp) exons which are separated by introns of 2.5 kb and 2.8 kb, respectively (Fig. 3A). The gene spans an overall distance of 18 kb.

Wild-Type Pattern of Expression

Utilizing the cDNA clone as a probe, poly(A)⁺ RNA from the skin of newborn wild-type mice was examined at daily intervals and shown to express the 0.8-kb transcript throughout early postnatal development; the level of expression appeared to be greatest, however, during days two through seven (Fig. 4). To determine whether the 0.8-kb transcript is expressed exclusively by melanocytes within the hair follicle, *W/W^v* mice were examined because they lack hair-bulb melanocytes due to a defect in the *c-kit* gene (Geissler et al., 1988). The 0.8-kb transcript was shown to be present in neonatal skin from *W/W^v* mice, indicating that it is not expressed in melanocytes but, presumably, within the follicular environment (Fig. 4).

Poly(A)⁺ RNA was also analyzed from a variety of wild-type adult tissues (Fig. 5) and whole embryos from day 10 postcoitum through term (data not shown). The only sample showing any expression was testis, which produced multiple transcripts ranging in size from 7.5 to 1.0 kb (Fig. 5). Based on Northern blot analysis and RNase protection assays, it was

determined that none of the testis-specific transcripts were expressed in neonatal skin (data not shown). Therefore, it is unlikely that the testis-specific transcripts are playing any role in α -locus function.

Changes in Gene Structure and Expression in Several Agouti Locus Mutations

In an attempt to provide evidence that the gene we have characterized is in fact the agouti gene, we tested whether it is structurally altered in agent-induced or spontaneous α -locus mutations. First of all, utilizing two probes from within the 5.5-kb wild-type *Eco*RI fragment at the 3' end of the gene, mutant locus-specific restriction fragment length variants (RFLVs) were detected of 7.5-kb and 3.5-kb for the agent-induced extreme nonagouti mutations, Is1Gso (Woychik et al., 1990a; Bultman et al., 1991) and α^{5MNU} (Bultman et al., 1991) (Fig. 3B), respectively. Additional experiments revealed that the 7.5-kb RFLV arises from a DNA structural alteration at the distal inversion breakpoint of Is1Gso which causes the 5' half of the gene to be juxtaposed with a portion of the limb deformity gene (Maas et al., 1990; Woychik et al., 1990b) in the opposite transcriptional orientation (Fig. 3A, B). For α^{5MNU} , the 3.5-kb RFLV is due to an intragenic deletion encompassing 2.8 kb of genomic DNA, which includes the third exon. (Fig. 3A, B). The distal inversion breakpoint of Is1Gso and the deletion breakpoints of α^{5MNU} have also been characterized by nucleotide sequence analysis (Fig. 3C).

In additional experiments we also analyzed the molecular structure of the original nonagouti mutation (α) and a black-and-tan (α^t) allele called SB B+T that arose spontaneously in a cross of the SEC/E and C57BL/E inbred lines at the Oak Ridge National Laboratory (W. Generoso, unpublished results). For these experiments, probes containing either the first exon (probe 1.5) or a portion of the first intron (probe 0.6) each detected a distinct mutant locus-specific RFLV associated with the α and α^t mutations (Fig. 6A, B). Based on this analysis, it appears that each of these mutant alleles contain a structural alteration caused by the presence of extra DNA (at least 11-kb or 5-kb for the α or α^t mutations, respectively) within a region

corresponding to a 700 bp *Bgl*III-*Eco*RI fragment within the first intron of the wild-type gene (Fig. 6B).

As expected, the DNA structural alterations in *Is1Gso*, α^{5MNU} , α , and α^t cause detectable changes in the expression of the gene. Production of the 0.8-kb transcript is absent in neonatal skin from *Is1Gso* homozygotes and is replaced by a low-abundance 8.0-kb transcript that likely arises from cryptic elements on the rearranged gene segments (Fig. 7, lane 1). In neonatal skin from α^{5MNU} homozygotes, a transcript is present at wild-type levels but is decreased in size to an extent that is consistent with a deletion of the third exon of the gene (Fig. 7, compare lanes 2 and 3). Deletion of the third exon results in the removal of 21 codons and the introduction of a frame-shift within the last exon (Fig. 3A). The original nonagouti mutation in C57BL mice does not express the 0.8-kb transcript (Fig. 7, lanes 8 and 9). In black-and-tan neonates, the 0.8-kb transcript is absent in skin derived from the black, dorsal surface but is overexpressed in skin derived from the yellow, ventral surface (Fig. 7, lanes 4 and 5). Control samples from wild-type (*A/A*) mice showed similar levels of expression of the 0.8-kb transcript in the dorsal and ventral surfaces. (Figure 7, lanes 10 and 11). Notably, *A^y* is associated with a marked increase in expression of a larger-than-normal transcript (Fig. 7, lanes 6 and 7).

Deregulated Expression in Lethal Yellow Heterozygotes

Lethal yellow (*A^y*) heterozygotes display a number of pleiotropic effects including pronounced obesity, a diabetes-like condition, and a propensity to develop a variety of spontaneous and induced neoplasms in the adult (Silvers, 1979). In order to examine the molecular defect associated with the *A^y* allele more precisely, a number of adult tissues from lethal yellow heterozygotes were analyzed for expression of the gene we have characterized. In striking contrast to wild-type mice where expression is restricted to testis and neonatal skin, *A^y* animals overexpressed a size-altered transcript in every tissue examined (Fig. 8).

To characterize the nature of the increased size of the *A^y* transcript, an adult kidney and testis cDNA library was prepared from lethal yellow heterozygotes and screened using the wild-type cDNA clone as a probe.

Analysis of several A^y cDNA clones indicated that the first exon of the agouti gene has been replaced by novel sequence on the A^y transcript, whereas the second, third, and fourth agouti exons, which contain the entire coding region, remain intact (Fig. 9A). Moreover, characterization of the genomic region flanking the novel sequence revealed a potential intron/exon consensus sequence precisely at point of divergence with the cDNA clone (Fig. 9B). This result strongly suggests that the altered size A^y transcript arises through a mechanism that involves splicing of the region of novel sequence to the second exon of the gene.

DISCUSSION

Characterization of the Agouti Locus Gene

We previously described the isolation of a region of DNA that maps to the agouti locus and is directly associated with DNA structural rearrangements that give rise to the extreme nonagouti phenotype (Woychik et al. 1990a; Bultman et al. 1991). Here we present the cloning and characterization of a gene contained within this region of DNA that is approximately 18 kb in length. This gene is structurally altered in several different α -locus mutations. For example, the distal inversion breakpoint associated with the recessive mutation, $Is1Gso$, breaks this gene within its third intron, and the 2.8 kb genomic deletion associated with α^{5MNU} is an intragenic mutation that specifically deletes the third exon and causes a frame-shift in the last exon of this gene (Fig. 3). We have also identified DNA structural alterations of this gene that are associated with the α and α^t alleles (Fig. 6). Additionally, the mRNA expressed from the A^y allele is longer than normal and arises from a replacement of the first exon with a region of novel sequence added to the 5' end of the transcript (Fig. 9). Further evidence that this gene is associated with the agouti locus is based on our finding that the original recessive nonagouti allele (α) associated with the C57BL strain fails to express the normal 0.8-kb transcript in neonatal skin (Fig. 7, lanes 8 and 9). The transcripts produced from the α^t and A^y alleles are also expressed in a deregulated manner (Fig. 7, lanes 4-7), with A^y being ectopically expressed in many adult tissues (Fig. 8). The fact that

this gene is directly associated with these mutations strongly suggests that it is, in fact, the one gene associated with α -locus function.

The normal pattern of expression of this gene is also consistent with it being responsible for the functions associated with the agouti locus. The production and deposition of phaeomelanin in the growing agouti hair shaft occurs during days three through seven of neonatal development, coincident with the appearance of the subapical yellow band (Sakurai et al., 1975). The observation that expression of this gene is maximal during days two through seven correlates with phaeomelanin production. Expression of this gene also correlates with the production of phaeomelanin in the a^t , A^y , and a mutants, particularly in a^t where expression is present at high levels in the yellow colored hair of the ventrum, but absent in the black colored hair of the dorsum (Fig. 7, lanes 4-9). Most significantly, we found that this gene is expressed in neonatal skin from W/W^v animals (which lack hair-bulb melanocytes), consistent with early skin grafting experiments by Silvers and Russell indicating that the agouti gene is expressed in the follicular environment and not within the melanocytes (Silvers and Russell, 1955; Silvers, 1958a, 1958b, 1961, 1979). However, in situ hybridization analysis of neonatal skin will be necessary to identify the cell type(s) expressing the agouti gene.

Expression from this region of DNA in wild-type, adult tissues is limited to testis (Fig. 5). The functional significance of the testis-specific expression is presently unclear, particularly since it appears that the unique-sized RNA transcripts produced in testis are not expressed in neonatal skin, elsewhere during development, or in other adult tissues (Fig. 5). Other characterized developmental genes that do not have any obvious function in testis, like limb deformity (Woychik et al., 1990b), also express testis-specific transcripts. It may turn out that either there is no functional significance to the expression of these genes in the testis whatsoever, or that there is a presently unknown testis-specific function associated with this expression.

Based on the fact that the structure and expression of this gene is directly affected by the most dominant agouti locus mutation (A^y), mutations in the middle of the dominance hierarchy (a^t , a), and the most

recessive member of the agouti allelic series (a^e), and the fact that the pattern of expression of this gene is compatible with normal α -locus function, we propose that the gene we have characterized is, in fact, the agouti gene.

Hypothetical Interaction With the Extension (e) Locus

Analysis of the sequence of the cDNA clone predicts a small polypeptide 131 amino acids in length and approximately 15,000 daltons in molecular weight (Fig. 2). This protein is probably secreted since it contains a putative signal peptide at its N-terminus (Fig. 2). Additionally, the molecule has several potential phosphorylation sites, a highly basic domain in the middle of the protein, and a cysteine-rich region near its C-terminus. However, neither of these domains resemble any characterized motif and cannot yet be ascribed a particular function.

Since the agouti locus, like steel (McCulloch et al., 1965; Mayer and Green, 1968; Dexter and Moore, 1977), does not act in a cell-autonomous manner, it is attractive to speculate that the agouti gene product, like the steel factor, functions as a ligand for some as yet undetermined receptor on the melanocyte. Interestingly, the extension (e) locus in chromosome 8 of the mouse acts in a cell-autonomous manner and produces pigmentation phenotypes similar to agouti (Silvers, 1979). However, dominant e -locus mutations cause an all-black phenotype similar to that observed in recessive α -locus mutations, and recessive e -locus mutations cause an all-yellow phenotype similar to that observed in dominant α -locus mutations. Making the assumption that dominant mutations at these two loci are associated with gain-of-function and recessive mutations are associated with loss-of-function, then a straightforward model (like that for the steel and dominant white spotting loci) in which the e -locus encodes a melanocyte-specific receptor that is directly activated by a ligand produced by the agouti locus is unlikely.

Alternatively, Takeuchi et al. (1989) described a model for a potential interaction between the α and e loci based on their experiments with skin explants from A^y/a or e/e mice. They proposed that α -melanocyte stimulating hormone (α -MSH) binds to its receptor on the surface of the melanocyte,

triggering a cascade of biochemical events culminating in the elevation of cAMP levels, which acts as a second messenger to produce eumelanin. The *e* locus hypothetically encodes a protein that modulates the production of cAMP via interaction with adenylate cyclase in the membrane of melanocytes. In this model, the agouti protein would function as a ligand to compete with α -MSH for binding of its receptor, preventing an increase of cAMP in the melanocyte, resulting in the production of phaeomelanin. As discussed above, our finding that the product of the agouti locus may be a small secreted protein is compatible with this model. The agouti gene product shares no sequence homology with α -MSH and, if this model is correct, it is unclear whether the agouti gene product would bind directly to α -MSH and inactivate its binding capacity, or would bind directly to the α -MSH receptor and competitively interfere with the binding of α -MSH. We expect that the cloned sequences described in this report, in combination with the recent cloning of the α -MSH receptor (Mountjoy et al., 1992), will generate the molecular reagents necessary to test this model.

Deregulated Expression in Lethal Yellow Heterozygotes

Lethal yellow (*A^y*) is an allele of particular interest because it is a recessive lethal and is associated with a number of dominant pleiotropic effects in the heterozygous condition. In 1905, *A^y* was the first lethal mutation to be documented in a mammal (Cuenot, 1905; Castle and Little, 1910). Lethal yellow homozygotes display characteristic abnormalities at the morula and blastula stages of development (Calarco and Pedersen, 1976) and die before implantation is complete (Robertson, 1942). The cause of death is presumably due to a delay in giant cell differentiation, thereby preventing attachment of the embryo to the endometrial cells lining the uterus (Eaton and Green, 1963). In addition, lethal yellow heterozygotes exhibit a systemic stimulation in body growth, reduced fertility in females (Granholm et al., 1986), a diabetes-like syndrome (Hellerstrom and Hellman, 1963), pronounced obesity (Dickerson and Gowen, 1947; Fenton and Chase, 1951; Carpenter and Mayer, 1958; Plocher and Powley, 1976; Friedman and Leibel, 1992), and a propensity to develop a variety of

spontaneous and induced neoplasms (Heston, 1942; Heston and Deringer, 1947; Heston and Vlahakis, 1961, 1962, 1963; Deringer, 1970).

In an attempt to characterize the nature of the molecular defect associated with the A^y allele, mRNA was prepared from a variety of adult tissues from A^y heterozygotes (Fig. 8). Whereas the agouti gene is normally expressed only in testis and neonatal skin, the A^y allele is associated with overproduction of agouti RNA transcripts in both neonatal skin and in every adult tissue that we have analyzed to date (Fig. 8). Based on these findings, our working hypothesis is that deregulated overexpression of the α -locus gene is directly responsible for the yellow coat color and, most importantly, for the dominant pleiotropic effects observed in lethal yellow heterozygotes. It is unlikely that the dominant pleiotropic effects associated with A^y are due to the production of an altered agouti gene product because the coding exons of the gene are unaltered, with the larger-than-normal transcript resulting from a 5' extension of the 0.8-kb mRNA (Fig. 9). Also, our preliminary analysis of a new allele with a phenotype identical to viable yellow (A^{vy}) (for a review of A^{vy} , see Wolff et al., 1986), revealed that the agouti gene is overexpressed in the same deregulated manner but produces a normal-sized mRNA (E. M., S. B., and R. W., unpublished data). To test our hypothesis, we are currently in the process of generating lines of transgenic mice that each overexpress the normal agouti gene product in several different tissues of the adult animal.

The molecular nature of the embryonic lethality associated with A^y homozygotes is presently unclear. Based on our preliminary molecular analysis of the DNA structure associated with the A^y allele and the fact that A^y produces an altered-sized transcript with an extension at its 5' end, we currently hypothesize that the A^y allele arose from a genomic rearrangement that led to a replacement of the normal promoter and first exon of the agouti gene with the promoter and 5' end of another gene located within the vicinity of the agouti locus. As a consequence of this structural alteration, we predict that the expression of a second gene was disrupted, and that the expression of this second gene, and not the agouti gene per se, is directly responsible for the embryonic lethality phenotype of A^y . Work currently in progress to

completely characterize the structure of the A^y DNA will allow us to test this hypothesis directly.

Agouti is a Single Gene Locus

Genetic experiments at the Oak Ridge National Laboratory involving several thousand offspring of α -locus mutants revealed three unequivocal and two probable recombination events between A^y and α^x (Russell et al., 1963), as well as one and two possible recombinations between A^y and α^t , and A^y and α , respectively (Siracusa et al., 1987). In each case, the recombination frequency was on the order of 0.1%, and flanking markers in the crosses showed A^y to lie proximal to α^x (Russell et al., 1963). Based on this result, it was hypothesized that the agouti gene was either very large, or that A^y was associated with a separate gene that is pseudoallelic with agouti. Our results help to resolve this issue and point to a single gene hypothesis. In each case, with the A^y , α^t and α alleles (we have not yet characterized the molecular structure of α^x), we see unique alterations in the structure and expression of the single 18-kb gene that we have characterized. If A^y were associated with a different gene, then we would not have expected to see structural and expression changes in the same gene for all three mutations. While this resolves the single versus multiple gene hypothesis, it is still difficult to reconcile the fact that recombination does occur between A^y and one to three alleles in a gene that is only 18 kb in length. The 0.1% recombination frequency would, on average, predict about a 100-kb distance between A^y and the other alleles. Since we know that this is not the case, we are led to hypothesize that there may be a highly localized recombination hotspot within the agouti gene, similar to that which occurs within the murine major histocompatibility complex (Kobori et al., 1986; Steinmetz et al., 1986) and human insulin and β -globin loci (Lebo et al., 1983; Treco et al., 1985).

The large number of alleles and the wide range of phenotypes associated with the agouti locus have been used as evidence by some investigators to propose that the agouti locus is comprised of multiple "mini-loci" and not a single gene (reviewed in Silvers, 1979). According to this hypothesis, each gene of the mini-locus plays a role in regulating

pigmentation in different parts of the body, particularly over the dorsal and ventral surfaces, and around the pinnae, nipples, and perineum. Support for this assertion stems from the finding that changes from yellow to black pigmentation proceed from the dorsal to the ventral regions as one progresses from the most dominant to the most recessive mutation of the agouti allelic series. For example, phaeomelanin progressively disappears from the mid-dorsum with A^i/a (A^i , intermediate yellow), from the lateral dorsum with a^t/a^t , from the ventral surface with a/a , and from the pinnae, nipples, and perineum with a^e/a^e (Silvers, 1979). With the mini-locus hypothesis, different genes should be affected by mutations associated with the individual alleles in the hierarchy. Our finding that the structure and expression of the same gene is affected with mutations at the top (A^y), middle (a^t and a), and the bottom (a^e) of the allelic series clearly demonstrates that the mini-locus hypothesis is not correct. On the other hand, a single gene hypothesis, which has also been proposed for the agouti locus (Silvers, 1979), is supported by our data and by many genetic observations, the most compelling of which is the fact that the mini-locus hypothesis would be expected to produce some phenotypes with darker hair on the ventrum than on the dorsum, or with solid black or yellow hair on the dorsum and agouti hair on the ventral surface, and these phenotypes have never been observed (Silvers, 1979). The data reported here clearly support the single gene hypothesis for the agouti locus.

REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1988). *Current Protocols in Molecular Biology*. John Wiley & Sons/Greene, New York.
- Aviv, H., and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. USA* **69**: 1408-1412.
- Bates, P. F., and Swift, R. A. (1983). Double cos site vectors: simplified cosmid cloning. *Gene* **26**: 137-146.
- Bultman, S. J., Russell, L. B., Gutierrez-Espeleta, G., 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. USA* **88**: 8062-8066.
- Calarco, P. G., and Pedersen, R. A. (1976). Ultrastructural observations of lethal yellow (A^y/A^y) mouse embryos. *J. Embryol. Exp. Morphol.* **35**: 73-80.
- Carpenter, K. J., and Mayer, J. (1958). Physiologic observations on yellow obesity in the mouse. *Am. J. Physiol.* **193**: 499-504.
- Castle, W. E., and Little, C. C. (1910). On a modified Mendelian ratio among yellow mice. *Science* **32**: 868-870.
- Cuenot, L. (1905). Les races pures et leurs combinaisons chez les souris. *Arch. Zool. Exp. Gen.* **3**: 123-132.
- Deringer, M. K. (1970). Influence of the lethal yellow (A^y) gene on development of reticular neoplasms. *J. Natl. Cancer Inst.* **45**: 1205-1210.

- Dexter, T. M., and Moore, M. A. S. (1977). In vitro duplication and cure of hematopoietic defects in genetically anaemic mice. *Nature* **269**: 412-414.
- Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**: 387-395.
- Dickerson, G. E., and Gowen, J. W. (1947). Hereditary obesity and efficient food utilization in mice. *Science* **105**: 496-498.
- Dunn, L. C. (1928). A fifth allelomorph in the agouti series of the house mouse. *Proc. Natl. Acad. Sci. USA* **14**: 816-819.
- Durkop, H., Latza, U., Hummel, M., Eitelbach, F., Seed, B., and Stein, H. (1992). Molecular cloning and expression of a new member of the nerve growth factor receptor family that is characteristic for Hodgkin's disease. *Cell* **68**: 421-427.
- Eaton, G. J., and Green, M. M. (1963). Giant cell differentiation and lethality of homozygous yellow mouse embryos. *Genetica* **34**: 155-161.
- Epstein, D. J., Vekemans, M., and Gros, P. (1991). Splotch, a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell* **67**: 767-774.
- Feinberg, A. P., and Vogelstein, B. (1984). Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**: 266-267.
- Fenton, P. F., and Chase, H. B. (1951). Effect of diet on obesity of yellow mice in inbred lines. *Proc. Soc. Exper. Biol.* **77**: 420-422.
- Friedman, J. M., and Leibel, R. L. (1992). Tackling a weighty problem. *Cell* **69**: 217-220.

- Geissler, E. N., Ryan, M. A., and Housman, D. E. (1988). The dominant-white spotting (*W*) locus of the mouse encodes the *c-kit* proto-oncogene. *Cell* **55**: 185-192.
- Granholm, N. H., Jeppesen, K. W., and Japs, R. A. (1986). Progressive infertility in female lethal yellow mice. *J. Reprod. Fertil.* **76**: 279-287.
- Green, M. C. (1989). Catalogue of mutant genes and polymorphic loci. In *Genetic Variants and Strains of the Laboratory Mouse*, (eds. M. F. Lyon, and A. G. Searle) pp. 17-20. Oxford University Press, Oxford and New York.
- Hellerstrom, C., and Hellman, B. (1963). The islets of Langerhans in yellow obese mice. *Metabolism* **12**: 527-536.
- Heston, W. E. (1942). Relationship between the lethal yellow (*A^y*) gene of the mouse and susceptibility to induced pulmonary tumors. *J. Cancer Inst.* **3**: 303-308.
- Heston, W. E., and Deringer, M. K. (1947). Relationship between the lethal yellow (*A^y*) gene of the mouse and susceptibility to spontaneous pulmonary tumors. *J. Cancer Inst.* **7**: 463-465.
- Heston, W. E., and Vlahakis, G. (1961). Influence of the *A^y* gene on mammary-gland tumors, hepatomas, and normal growth in mice. *J. Natl. Cancer Inst.* **26**: 969-983.
- Heston, W. E., and Vlahakis, G. (1962). Genetic obesity and neoplasia. *J. Natl. Cancer Inst.* **29**: 197-209.
- Heston, W. E., and Vlahakis, G. (1963) Increase of induced skin tumors in the mouse by the lethal yellow gene (*A^y*). *J. Natl. Cancer Inst.* **31**: 189-195.

- Hollander, W. F. and Gowen, J. W. (1956). An extreme nonagouti mutant in the mouse. *J. Hered.* **47**: 221-224.
- Kobori, J. A., Strauss, E., Minard, K., and Hood, L. (1986). Molecular analysis of the hotspot of recombination in the murine major histocompatibility complex. *Science* **234**: 173-179.
- Kozak, M. (1987). An analysis of 5'-noncoding sequence from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* **15**: 8125-8148.
- Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* **157**: 105-132.
- Lebo, R. V., Chakravarti, A., Buetow, K. H., Cheung, M. C., Cann, H., Cordell, B., and Goodman, H. (1983). Recombination within and between the human insulin and β -globin gene loci. *Proc. Natl. Acad. Sci. USA* **80**: 4808-4812.
- Maas, R. L., Zeller, R., Woychik, R. P., Vogt, T. F., and Leder, P. (1990). Disruption of formin-encoding transcripts in two mutant *limb deformity* alleles. *Nature* **346**: 853-855.
- Mayer, T. C., and Green, M. C. (1968). An experimental analysis of the pigment defect caused by mutations at the *W* and *Sl* loci in mice. *Dev. Biol.* **18**: 62-75.
- McCulloch, E. A., Siminovitch, L., Till, J. E., Russell, E. S., and Bernstein, S. E. (1965). The cellular basis of the genetically determined hemopoietic defect in anemic mice of genotype *Sl/Sl^d*. *Blood* **26**: 399-410.
- Mercer, J. A., Seperack, P. K., Strobel, M. C., Copeland, N. G., and Jenkins, N. A. (1991). Novel myosin heavy chain encoded by murine *dilute* coat color locus. *Nature* **349**: 709-713.

- Mount, S. M. (1982). A catalogue of splice junction sequences. *Nucl. Acids Res.* **10**: 459-472.
- Mountjoy, K. G., Robbins, L. S., Mortrud, M. T., and Cone, R. D. (1992). The cloning of a family of genes that encode the melanocortin receptors. *Science* **257**: 1248-1251.
- Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**: 2444-2448.
- Papayioannou, V. E., and Gardner, R. L. (1979). Investigation of the lethal yellow A^y/A^y embryo using mouse chimeras. *J. Embryol. Exp. Morphol.* **52**: 153-163.
- Plocher, T. A. and Powley, T. L. (1976). Effect of hypophysectomy on weight gain and body composition in the genetically obese yellow (A^y/a) mouse. *Metabolism* **25**: 593-602.
- Robertson, G. G. (1942). An analysis of the development of homozygous yellow mouse embryos. *J. Exp. Zool.* **89**: 197-231.
- Russell, L. B., McDaniel, N. C., and Woodiel, F. N. (1963). Crossing-over within the a "locus" of the mouse. *Genetics* **48**: 907a.
- Sakurai, T., Ochiai, H., and Takeuchi, T. (1975). Ultrastructural change of melanosomes associated with agouti pattern formation in mouse hair. *Develop. Biol.* **47**: 466-471.
- Sambrook, J., Fritsch, E.F., Maniatis, T., (1989) *Molecular cloning: A laboratory manual, second edition* . Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.
- Sealey, P. G., Whittaker, P. A., and Southern, E. M. (1985). Removal of DNA sequences from hybridisation probes. *Nucl. Acids Res.* **13**: 1905-1922.
- Sheets, M. D., Ogg, S. C., and Wickens, M. P. (1990). Point mutations in AAUAAA and the poly(A) addition site: effects on the accuracy and efficiency of cleavage and polyadenylation in vitro. *Nucl. Acids Res.* **18**: 5799-5805.
- Silvers, W. K. (1958a). An experimental approach to action of genes at the agouti locus in the mouse. II. Transplants of newborn *aa* ventral skin to *a^ta*, *A^wa*, and *aa* hosts. *J. Exp. Zool.* **137**: 181-188.
- Silvers, W. K. (1958b). An experimental approach to action of genes at the agouti locus in the mouse. III. Transplants of newborn *A^w-*, *A-*, *a^t-*skin to *A^y-*, *A^w-*, *A-* and *aa* hosts. *J. Exp. Zool.* **137**: 189-196.
- Silvers, W. K. (1961). Genes and pigment cells of mammals. *Science* **134**: 368-373.
- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6-44. Springer-Verlag, New York.
- Silvers, W. K., and Russell, E. S. (1955). An experimental approach to action of genes at the agouti locus in the mouse. *J. Exp. Zool.* **130**: 199-220.
- Siracusa, L. D., Russell, L. B., Eicher, E. M., Corrow, D. J., Copeland, N. G., and Jenkins, N. A. (1987). Genetic organization of the *agouti* region of the mouse. *Genetics* **117**: 93-100.

- Steinmetz, M., Stephan, D., and Lindahl, K. F. (1986). Gene organization and recombinational hotspots in the murine major histocompatibility complex. *Cell* 44: 895-904.
- Tabor, S., and Richardson, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* 84: 4767-4771.
- Takeuchi, T., Kobunai, T., and Yamamoto, H. (1989). Genetic control of signal transduction in mouse melanocytes. *J. of Invest. Dermatol.* 92: 239S-242S.
- Treco, D., Thomas, B., and Arnheim, N. (1985). Recombination hot spot in the human β -globin gene cluster: meiotic recombination of human DNA fragments in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 5: 2029-2038.
- von Heijne, G. (1986). A new method for predicting signal cleavage sites. *Nucl. Acids Res.* 14: 4683-4690.
- Witte, O. N. (1990). Steel locus defines new multipotent growth factor. *Cell* 63: 5-6.
- Wolff, G. L., Roberts, D. W., and Galbraith, D. B. (1986). Prenatal determination of obesity, tumor susceptibility, and coat color pattern in viable yellow (A^{vy}/a) mice. *J. Hered.* 77: 151-158.
- Woychik, R. P., Generoso, W. M., Russell, L. B., Cain, K. T., Cacheiro, N. L. A., Bultman, S. J., Selby, P. B., Dickinson, M. E., Hogan, B. L. M., and Rutledge, J. C. (1990a). 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. USA* 87: 2588-2592.

- Woychik, R. P., Maas, R. L., Zeller, R., Vogt, T. F., and Leder, P. (1990b). 'Formins': proteins deduced from the alternative transcripts of the *limb deformity* gene. *Nature* **346**: 850-853.
- 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.

APPENDIX

FIGURES

Figure 1. Identification of an exon near the distal inversion breakpoint of Is1Gso.

(A) Restriction map of a region of DNA that maps to the agouti locus and is associated with structural alterations in the extreme nonagouti mutations, Is1Gso and α^{5MNU} . The positions of two DNA probes that lie in close proximity to the distal inversion breakpoint in Is1Gso and the 2.8-kb deletion in α^{5MNU} are indicated. R, *Eco*RI; B, *Bam*HI; X, *Xba*I.

(B) Identification of a segment of DNA near the distal inversion breakpoint of Is1Gso that is conserved in genomic DNA from several mammalian species. Probe A in Figure 1A was ^{32}P -labeled and hybridized to *Eco*RI- (R) or *Pst*I- (P) digested DNA from hamster, cat, and dog. The extensive smearing in hamster DNA is due to the presence of a rodent-specific repetitive element in probe A. The λ *Hind*III molecular weight standard is shown on the left in kb.

(C) Identification of coding sequence near the distal inversion breakpoint of Is1Gso by Northern blot analysis. Probe B in Figure 1A was ^{32}P -labeled and hybridized to poly(A)⁺ RNA from wild-type skin of a day 4 neonate. The positions of the 28S and 18S rRNA subunits are indicated on the left. Higher molecular weight smearing was due to the presence of moderately repetitive sequences in the probe.

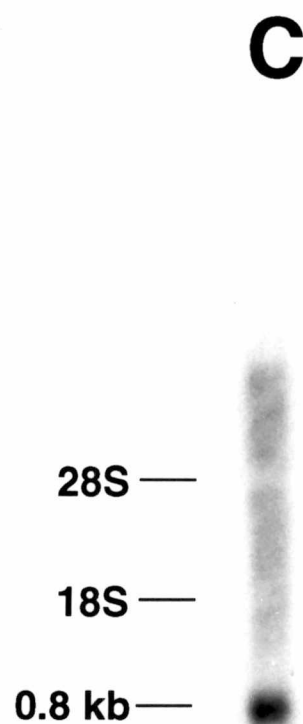
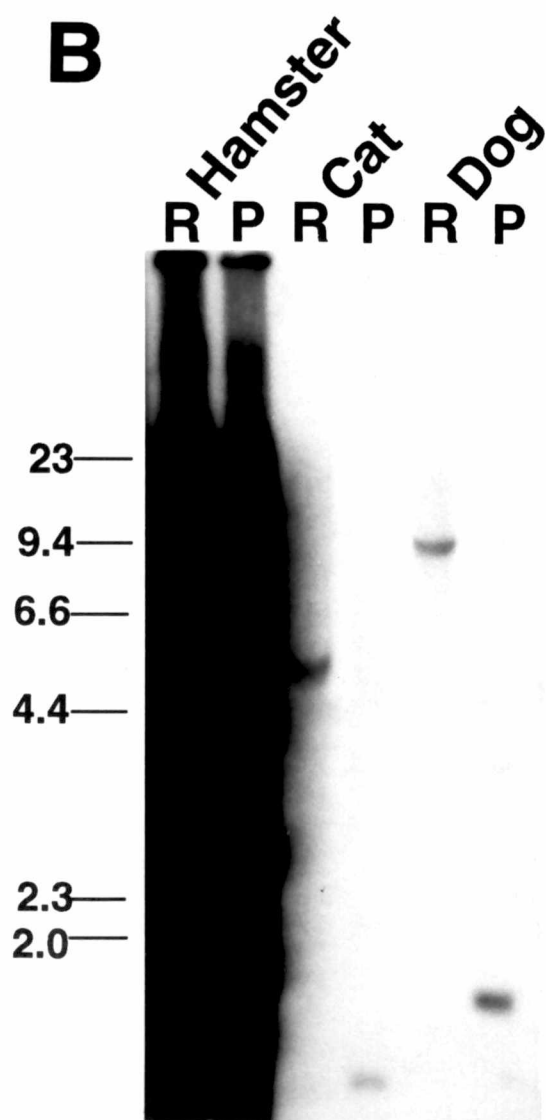
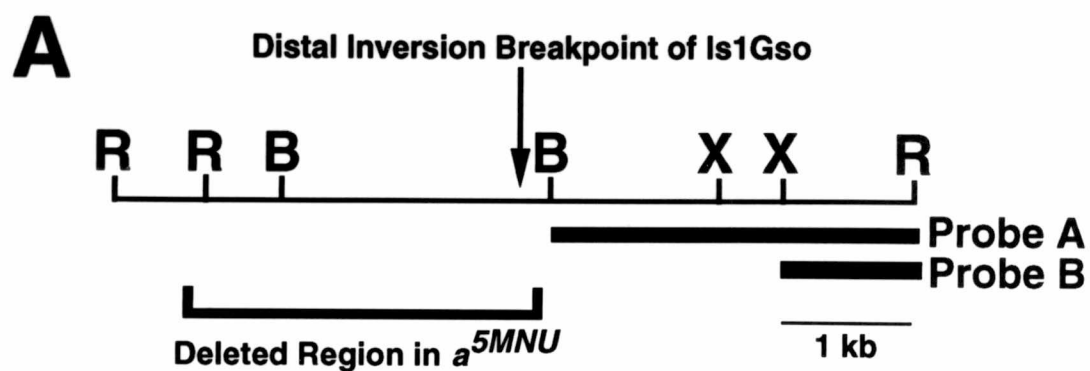


Figure 2. Nucleotide and predicted amino acid sequence of the cDNA.

The putative signal peptide sequence is underscored with a double line and the polyadenylation signal with a single line. The boxed region represents a highly basic domain that has the potential to be phosphorylated at a number of sites designated by PO₄. An asterisk denotes a potential N-linked glycosylation site, and cysteine residues are highlighted by solid circles. Arrowheads delimit the boundaries of the four individual exons.

TTCAAGGACAGGAAAGACATTCTGGCCTGGCTTCCCTTAGGGGAGCTGATGCGGAATAGAGTC 63

ACTTGTGCTGCTTCTCAGG ATG GAT GTC ACC CGC CTA CTC CTG GCC ACC CTA 115
Met Asp Val Thr Arg Leu Leu Leu Ala Thr Leu 11

GTG AGC TTC CTG TGC TTC TTC ACC GTC CAC AGC CAC CTG GCA CTC GAG 163
Val Ser Phe Leu Cys Phe Phe Thr Val His Ser His Leu Ala Leu Glu 27

PO₄ PO₄ *
GAG ACG CTT GGA GAT GAC AGG AGT CTG CGG AGT AAC TCC TCC ATG AAC 211
Glu Thr Leu Gly Asp Asp Arg Ser Leu Arg Ser Asn Ser Ser Met Asn 43

TCG CTG GAT TTC TCC TCT GTT TCT ATC GTG GCA CTG AAC AAG AAA TCC 259
Ser Leu Asp Phe Ser Ser Val Ser Ile Val Ala Leu Asn Lys Lys Ser 59

PO₄ PO₄ PO₄
AAG AAG ATC AGC AGA AAA GAA GCC CAG AAG CGG AAG AGG TCT TCC AAG 307
Lys Lys Ile Ser Arg Lys Glu Ala Glu Lys Arg Lys Arg Ser Ser Lys 75

PO₄
AAA AAG GCT TCG ATG AAG AAG GTG GCA AGG CCC CCG CCA CCT TCG CCC 355
Lys Lys Ala Ser Met Lys Lys Val Ala Arg Pro Pro Pro Pro Ser Pro 91

PO₄
TGC GTG GCC ACC CGC GAC AGC TGC AAG CCA CCC GCA CCC GCC TGC TGC 403
Cys Val Ala Thr Arg Asp Ser Cys Lys Pro Pro Ala Pro Ala Cys Cys 107

PO₄
GAC CCG TGC GCC TCC TGC CAG TGC CGT TTC TTC GGC AGC GCC TGC ACC 451
Asp Pro Cys Ala Ser Cys Gln Cys Arg Phe Phe Gly Ser Ala Cys Thr 123

TGT CGA GTA CTC AAC CCC AAC TGC TGA CGCAGCTTCTTCGCTGCGCGCAGCT 505
Cys Arg Val Leu Asn Pro Asn Cys End 131

TCGGGAACGGGTGATTGGGCGGGGCTTCAGGGTCCCGCGCTTCTAGGCTGAGGGGCGGGTCTC 568

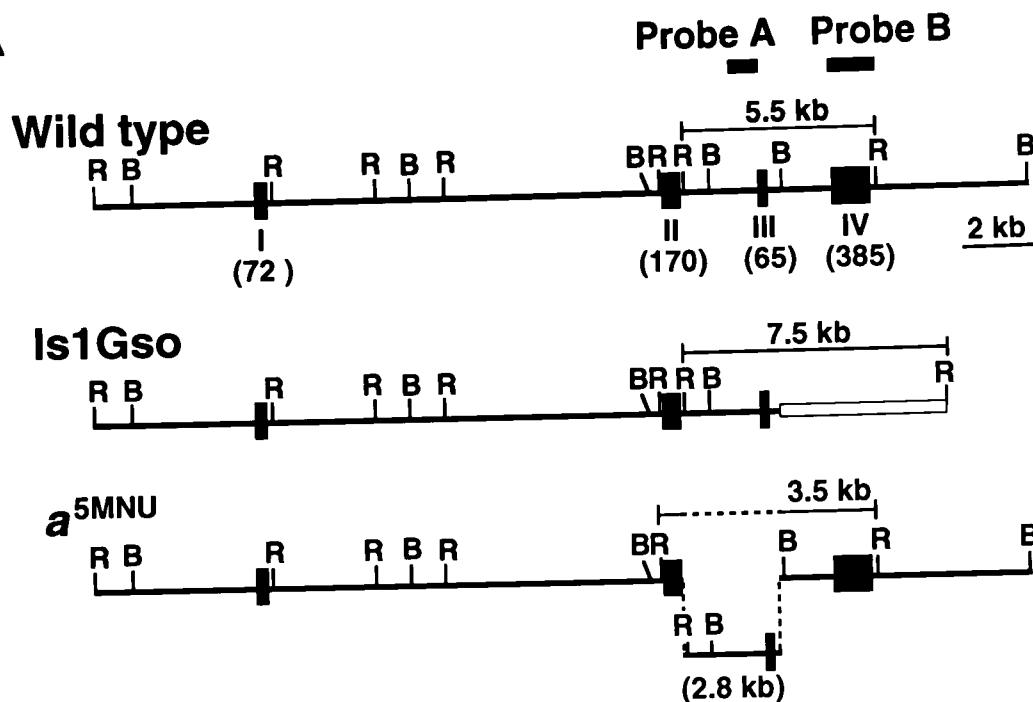
TGTGGGTGGGGCTTGTGGGTGGGCGTGGTCACTGGTTGTGACTTGTGGGCGCTTTCAAAAAAC 631

CGGTTTCTAGGAAACCTAGTGGAGCTAAAAATCAGAATACAATAATATTTTAGGCTGCC(A) 692

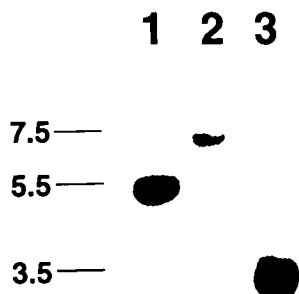
Figure 3. Intron/exon structure of the wild-type locus and two extreme nonagouti mutations.

(A) Schematic representation of the genomic structure of the agouti gene in DNA from wild-type, Is1Gso, and α^{5MNU} mice. The four exons are depicted as solid boxes, and the introns and flanking sequences are shown as a solid line. The first exon, which is 72 bp long in the cDNA clone, is 8-10 bp longer at its 5' end based on RNase protection experiments (data not shown). The 3' junction of the last exon corresponds to the 3' end of the cDNA clone, immediately upstream from the poly(A) tract. The precise location of each intron/exon junction was ascertained by identifying where genomic DNA sequences diverged from the cDNA sequence. At each splice junction, the genomic DNA sequence matched the canonical sequence for 5' splice donor and 3' splice acceptor sites (Mount, 1982) (data not shown). The open box in the Is1Gso schematic represents genomic DNA from the limb deformity (*ld*) gene in the opposite transcriptional orientation relative to agouti. The 2.8 kb intragenic deletion in the α^{5MNU} mutation is depicted by a horizontal dashed line above the mutant locus, with the deleted region shown below the mutant locus. The positions of two probes (Probes A and B) and the sizes of the *Eco*RI fragments they identify are shown. R, *Eco*RI; B, *Bam*HI. (B) Identification of RFLVs specifically associated with the Is1Gso and α^{5MNU} mutations. Wild-type (lane 1, C3H strain), Is1Gso homozygous (lane 2), and α^{5MNU} homozygous (lane 3) genomic DNA was digested with *Eco*RI, blotted, and hybridized with a ^{32} P-labeled fragment of DNA corresponding to probe A (lanes 1 and 2) or probe B (lane 3) in Figure 3A. The sizes of the DNA fragments detected by probes A and B are shown on the left in kb. (C) Nucleotide sequence of the distal inversion breakpoint of Is1Gso and the 5' and 3' deletion breakpoints of α^{5MNU} compared to wild type. The vertical bars indicate the positions of the DNA breakpoints. The lower case nucleotides in Is1Gso represent sequence of the *ld* gene, and the horizontal lines in α^{5MNU} represent deleted nucleotides.

A



B



C

Wild type CATTGAGGACAACGTCCCTAGGCTGTGGGAGTGTGTCTGTATGTAGCGTTT

Is1Gso CATTGAGGACAACGTCCgcgagtaaatacgaacccggctacacttttatgt

Wild type AAGTCAAGATGCTTGGTGGACTTGGTTTTCTTTAGCGTTAATGACATTTTAA

5' *a*^{5MNU} AAGTCAAGATGCTTGGTGGACTTGGTTTTCTTTAGCGTTAATG

Wild type AGGAGGCTGAGGCACGTAGATCTGAGTTTGAGGCCAGCCTGGTCTACAGAGT

3' *a*^{5MNU} ACGTAGATCTGAGTTTGAGGCCAGCCTGGTCTACAGAGT

Figure 4. Northern blot analysis of wild-type (*A/A*) neonatal skin.

The full-length cDNA clone (Figure 2) was ^{32}P -labeled and hybridized to poly(A)⁺ RNA (2.5 μg per lane) from skin of day one through day eight wild-type neonates (1-8) and day 6 *W/W^v* neonates (W). RNA molecular weight standards are shown on the left in kb. The filter was subsequently hybridized with a tubulin probe as a control to analyze the quantity and quality of the RNA in each lane.

1 2 3 4 5 6 7 8 W

9.5—
7.5—

4.4—

2.4—

1.35—

0.33—



—0.8 kb



Tubulin

Figure 5. Northern blot analysis of adult tissues from wild-type (*A/A*) mice.

The full-length cDNA clone was ^{32}P -labeled and hybridized to a variety of poly(A)⁺ RNAs (2.5 μg per lane). RNA molecular weight standards are shown on the left in kb. The filter was subsequently hybridized with a tubulin probe as a control to analyze the quantity and quality of the RNA in each lane.

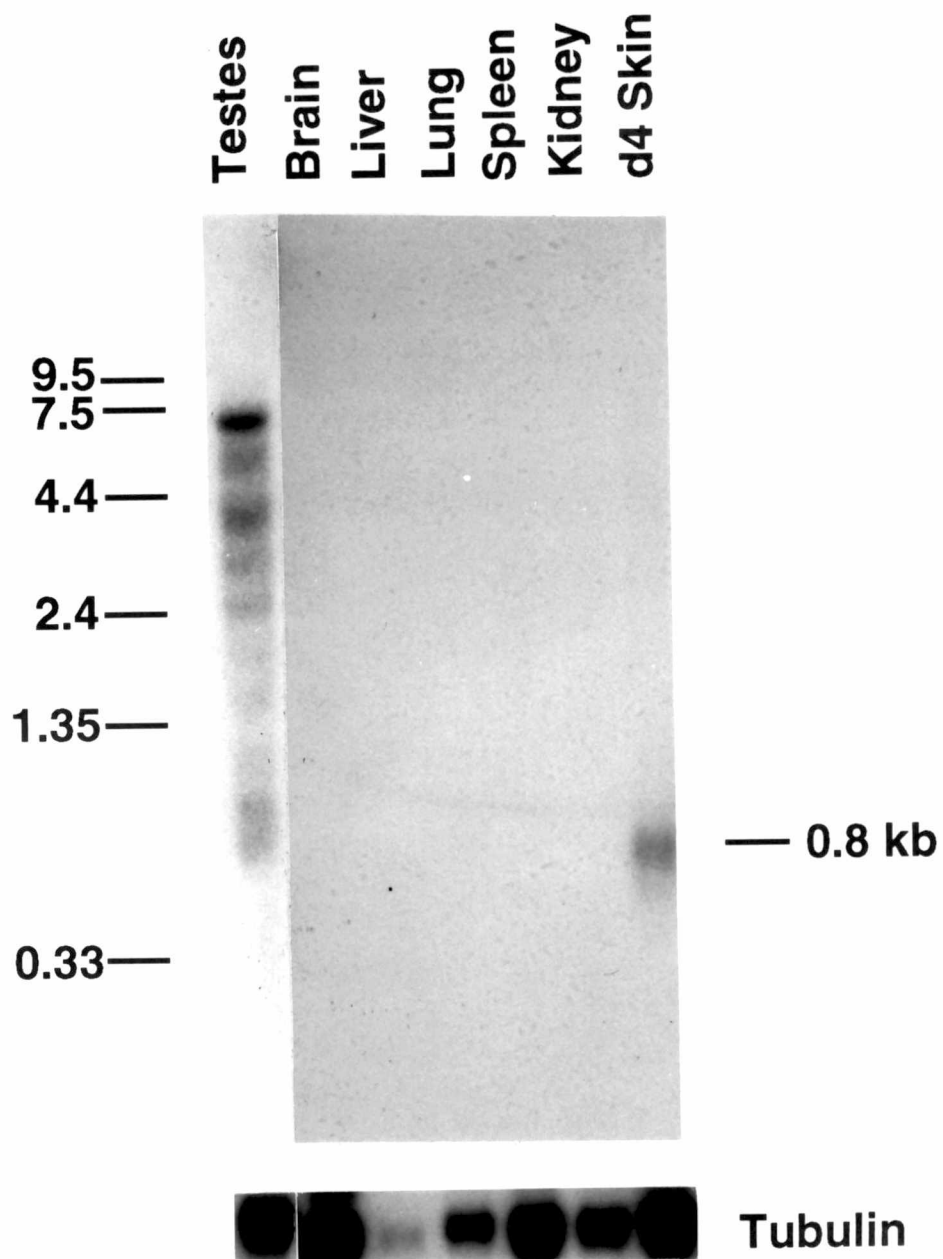


Figure 6. Identification of DNA structural alterations associated with the α and α^t alleles.

(A) Wild-type (A, C3H strain), nonagouti (α , C57BL/E strain), and black-and-tan (α^t , SB B+T stock) genomic DNA was digested with *Bam*HI or *Bgl*II, blotted, and hybridized with a ^{32} P-labeled fragment of DNA corresponding to probe 1.5 (*Bam*HI digest) or probe 0.6 (*Bgl*II digest), which are shown as thick horizontal lines in the illustration of the wild-type allele in B. The sizes of the DNA fragments detected by the two probes are shown on the left margin in kb. The SB B+T mutation arose in a cross between the strains SEC/E and C57BL/E, both of which are nonagouti (α/α) homozygotes and display a 16.0-kb *Bam*HI fragment with probe 1.5 and a 6.0-kb *Bgl*II fragment with probe 0.6, although only the C57BL/E result is shown. The parental strain from which the original nonagouti α mutation arose is not known, however, two wild-type inbred strains, FVB/N and 101, were analyzed in addition to C3H, and all displayed the wild-type pattern illustrated for C3H. (B) Schematic representation of the 5' end of the gene shown in Figure 3 from wild-type (A), nonagouti (α), and SB B+T mutant (α^t) mice. The first exon is shown as a solid box, and the bold horizontal line represents the 5' flanking sequence and portion of the first intron. Probe 1.5 detects a 8.0-kb fragment in wild type, and RFLV fragments of 16.0-kb and 10.0-kb in the α and α^t mutant alleles, respectively. Probe 0.6 detects a 2.0-kb wild-type fragment, and RFLV mutant allele-specific fragments of 6.0-kb in α and 8.0-kb in α^t . These RFLV's are due to the presence of at least 11-kb (α) or 5.5-kb (α^t) of additional sequence in the mutant alleles within a highly localized 700-bp region on the wild-type DNA between *Bgl*II and *Eco*RI sites. The insertions for each allele have not yet been cloned and mapped, and, therefore, are depicted by vertical bars with an open parentheses at the end. B, *Bam*HI; R, *Eco*RI; Bg, *Bgl*II.

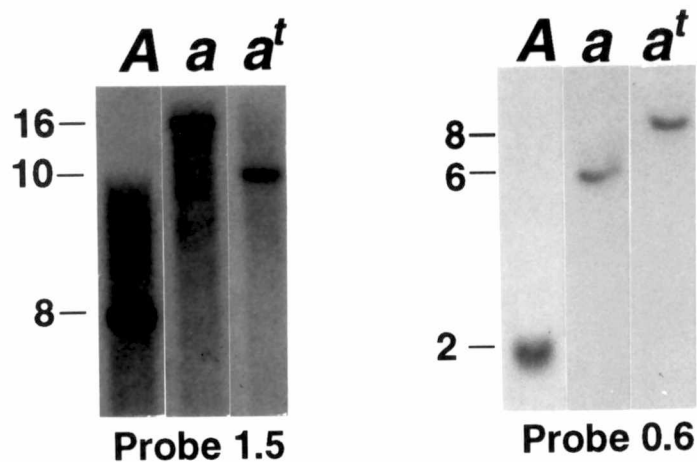
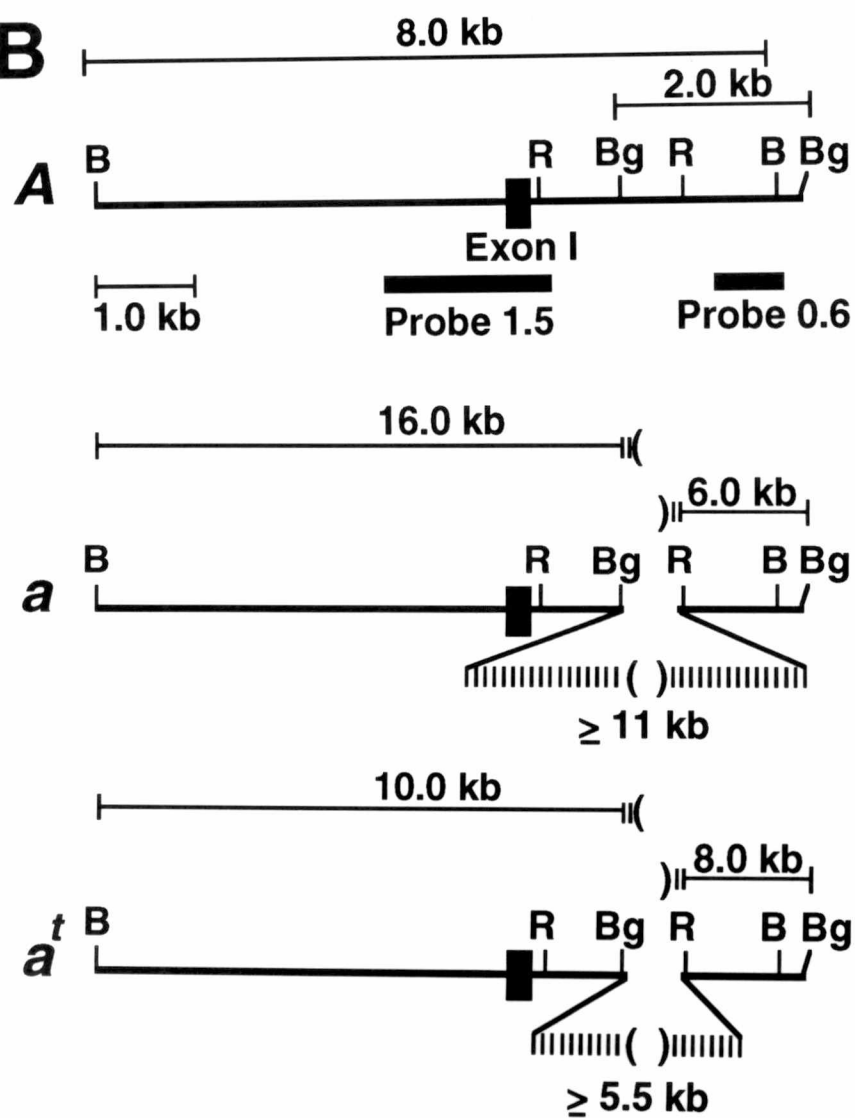
A**B**

Figure 7. Northern blot analysis of neonatal skin from several agouti locus mutants.

The full-length cDNA clone (Figure 2) was ^{32}P -labeled and hybridized to poly(A)⁺ RNA (2.5 μg per lane) from neonatal skin of spontaneous or agent-induced α -locus mutations. RNA molecular weight standards (kb) and the positions of the 28S and 18S rRNA subunits are shown. The filter was subsequently hybridized with a tubulin probe as a control to analyze the quantity and quality of the RNA in each lane. D, skin derived from the dorsal surface of neonates; V, skin derived from the ventral surface of neonates; Is1Gso, day four Is1Gso homozygote; 5MNU, day five $\alpha^{5\text{MNU}}$ homozygote; A/A , day five wild type; α^t/α^t , day five black-and-tan; A^y/α , day six lethal yellow heterozygote; α/α , day six nonagouti (C57BL/10).

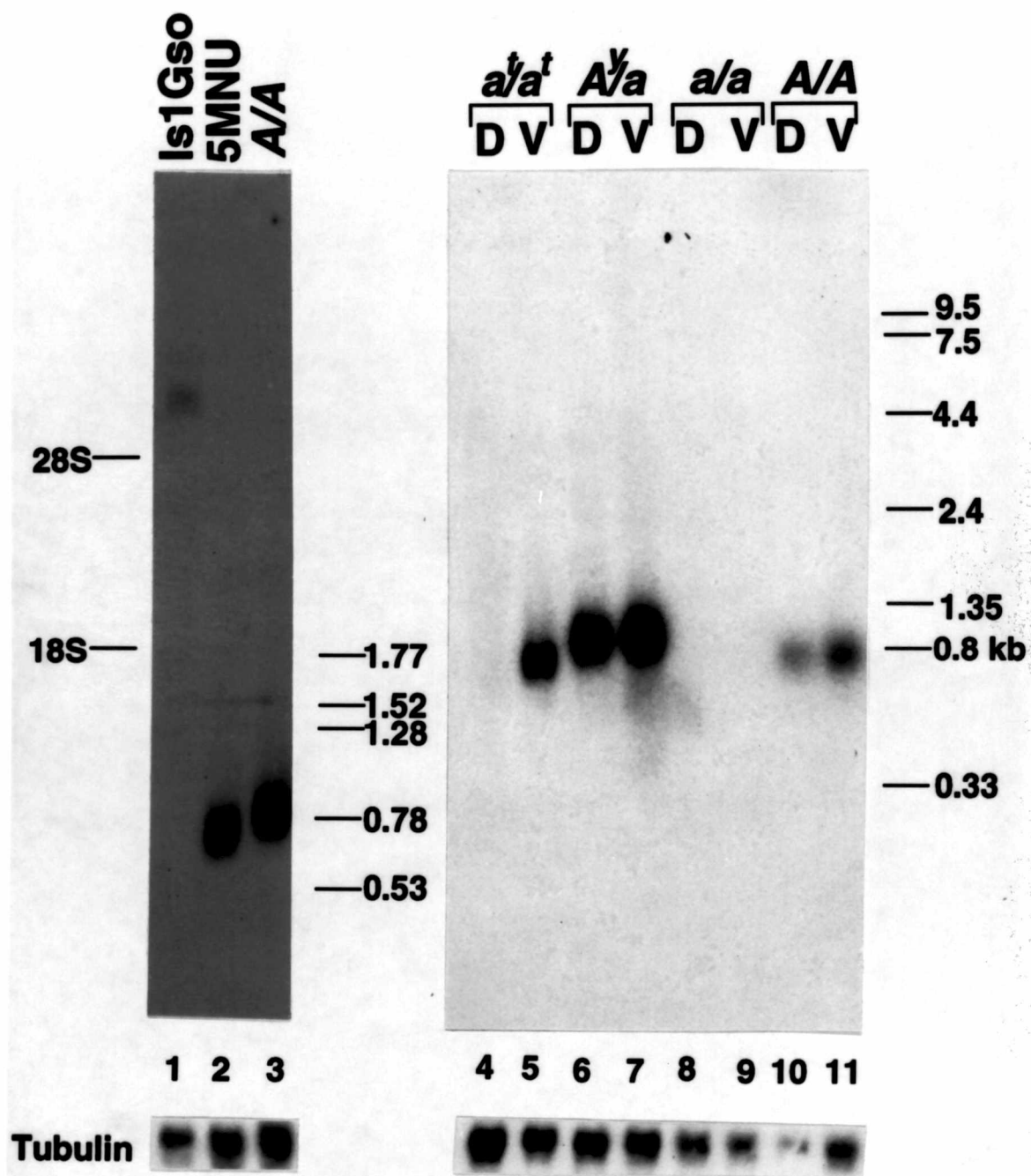


Figure 8. Northern blot analysis of adult tissues from wild-type and lethal-yellow heterozygotes.

The full-length cDNA clone (Figure 2) was ^{32}P -labeled and hybridized to poly(A)⁺ RNA (2.5 μg per lane) from adult tissues or neonatal skin of wild-type or $A^y/_-$ mutant animals. RNA molecular weight standards are shown on the left in kb, and the positions of the 28S and 18S rRNA subunits are indicated on the right. The filter was subsequently hybridized with a tubulin probe as a control to analyze the quantity and quality of the RNA in each lane. $A^y/_-$, lethal yellow heterozygotes with genotypes of A^y/a and A^y/a^e were used in this analysis; d4 skin, day four postnatal skin; d6 skin, day six postnatal skin.

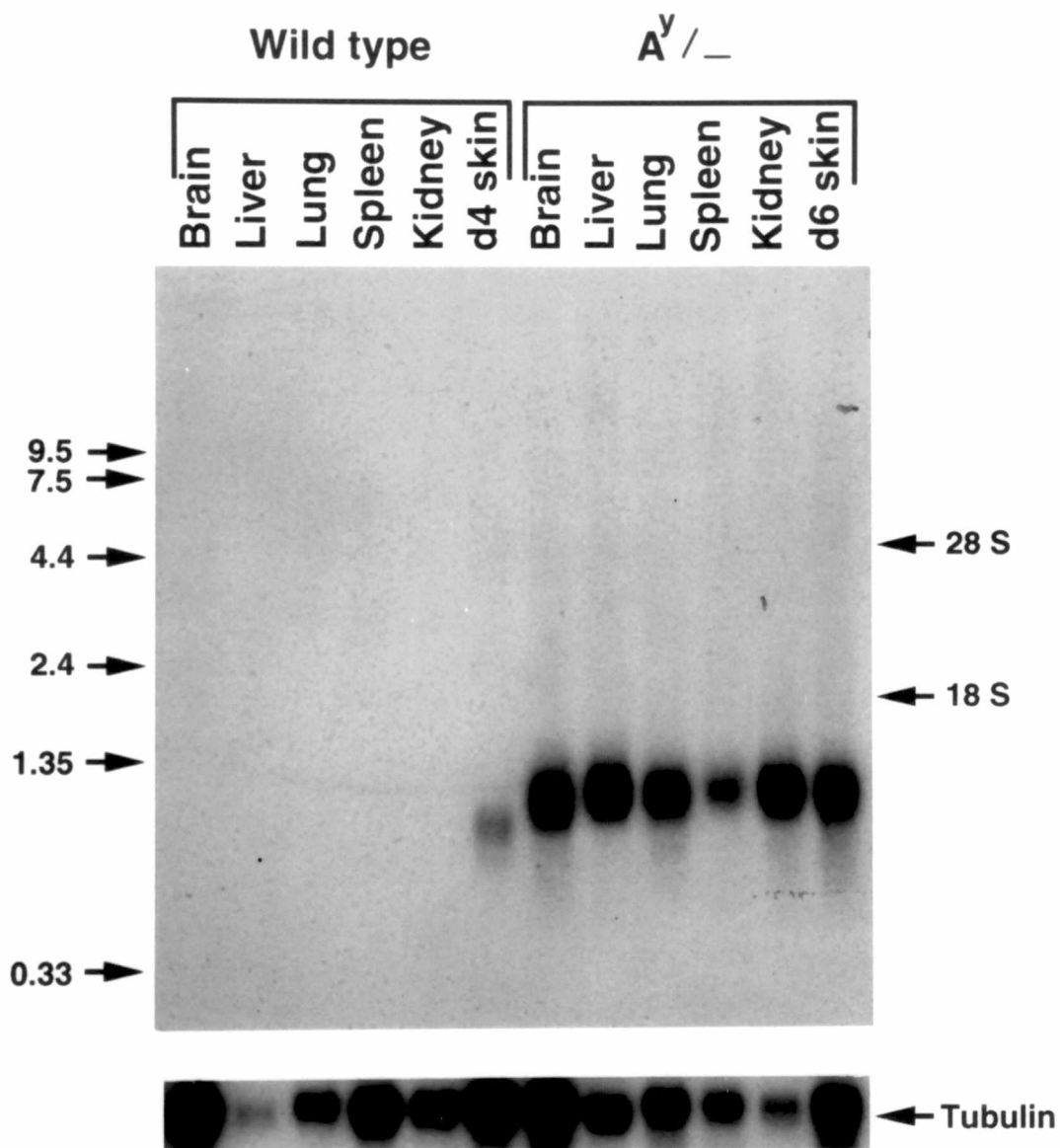
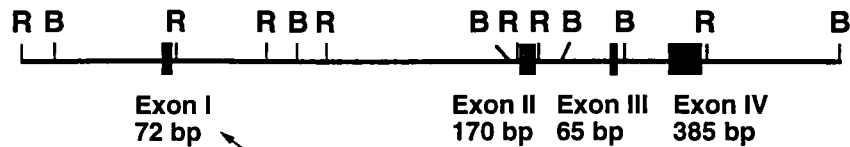


Figure 9. Molecular analysis of the size-altered A^y transcript.

(A) Schematic representation of the wild-type and lethal yellow (A^y) transcripts produced by the agouti locus. Based on the analysis of cDNA clones, the A^y transcript is identical to the wild-type transcript for the sequence derived from the second, third, and fourth exons. However, the 5' end of the second exon on the A^y transcript is flanked by sequence that is different from the region corresponding to the first exon on the wild-type transcript. Moreover, our preliminary evidence suggests that this sequence unique to A^y corresponds to the first noncoding exon (approximately 300 nucleotides in length) of another gene (E.M., S.B, R.W., unpublished results). The genomic organization of the wild-type gene is shown above with the four exons depicted as solid boxes, and the introns and flanking sequences shown as a solid line. B, *Bam*HI; R, *Eco*RI. (B) Nucleotide sequence comparison of the wild-type and A^y cDNA clones, and comparison of the 5' end of the A^y cDNA with its corresponding genomic region. The sequence divergence point is positioned with a vertical line and corresponds to the junction between the first and second exons of the wild-type cDNA. Also, at this same junction point, the genomic sequence diverges from the A^y cDNA and is followed by a canonical 5' splice donor (Mount, 1982).

A

Wild-type DNA



Wild-type transcript



A^y transcript



Unique to A^y transcript

B

	EXON I	EXON II
Wild-type cDNA	 t t g t g c t	g c t t c t
A^y cDNA	 <u>g g c g c g g</u>	g c t t c t
Unique genomic	 g g c g c g g	g t g a g t
	EXON	INTRON

PART V

MOLECULAR ANALYSIS OF REVERSE MUTATIONS FROM NONAGOUTI (α) TO BLACK-AND-TAN (α^t) AND WHITE-BELLIED AGOUTI (A^w) REVEALS ALTERNATE FORMS OF AGOUTI TRANSCRIPTS

ABSTRACT

The agouti gene regulates the differential production of eumelanin (black or brown) and phaeomelanin (yellow) pigment granules by melanocytes in the hair follicles of mice. The original nonagouti (a) allele, which confers a predominantly black coat color, has been shown to revert to two other more dominant agouti alleles, black-and-tan (a^t) and white-bellied agouti (A^w), with an exceptionally high frequency. The a^t and A^w alleles confer phenotypes in which the pigmentation is not uniformly distributed over the dorsal and ventral surfaces of the animal; in both cases the ventral surface of the animal is markedly lighter than the dorsal surface due to an increase in phaeomelanin production. In order to understand the unusually high reversion rate of a to a^t or A^w , and to decipher the molecular events associated with the different pigmentation patterns associated with these three agouti alleles, we have characterized a , a^t and A^w at the molecular level. Here we report that insertions of 11 kb, 6 kb, and 0.6 kb are present at precisely the same position in the first intron of the agouti gene in a , a^t , and A^w , respectively. The a insertion consists of a 5.5-kb VL30 element that has incorporated 5.5 kb of additional sequence internally; this internal sequence is flanked by 526-bp direct repeats. The a^t allele contains only the VL30 element and a single, internal 526-bp repeat. The A^w allele has only a solo VL30 LTR. Based on the comparison of the structure of the a^t and A^w insertions, we propose that reverse mutations occur by excision of inserted sequences in a through homologous recombination, utilizing either the 526-bp direct repeats to generate a^t or the VL30 LTRs to generate A^w . Moreover, the analysis of these three alleles has allowed us to identify additional exons of the agouti gene that give rise to alternatively processed forms of agouti mRNA. We demonstrate that the distinct insertions in a , a^t and A^w cause pigmentation differences by selectively inactivating the expression of different forms of agouti transcripts.

INTRODUCTION

The agouti locus in chromosome 2 of the mouse controls the differential production of eumelanin (black or brown depending on alleles present at other loci) and phaeomelanin (yellow) pigment granules by the melanocytes within the hair follicle. Agouti pigmentation, which is the true wild-type coat color of mice, arises from the fact that the individual hair shafts contain a subapical yellow band on an otherwise black or brown background. At least nineteen different alleles have been identified at the agouti locus that form a complex dominance hierarchy in which alleles associated with phaeomelanin production are generally dominant over alleles associated with eumelanin production (Silers, 1979; Green, 1989). Moreover, a number of these alleles regulate the amount and distribution of phaeomelanin and eumelanin differently in different regions of the pelage, particularly across the dorsal and ventral surfaces (Silers, 1979; Green, 1989).

The regional pigmentation differences observed with different agouti alleles are exemplified by the original nonagouti (a), black-and-tan (a^t), and white-bellied agouti (A^w) alleles. The a/a genotype gives rise to a predominantly black coat color, with relatively small amounts of yellow pigmentation restricted to the pinnae, mammae, and perineum. Nonagouti is recessive to the A^w , wild-type (A), and a^t alleles. The a^t/a^t genotype causes a unique pattern of pigmentation in which the dorsal surface is entirely black, the ventral surface and pinnae are completely yellow, and a sharp lateral line of demarcation separates the dorsal and ventral surfaces. Black-and-tan is dominant to a but recessive to A^w . It is dominant to A across the ventral surface of the animal consistent with yellow being dominant to agouti but is recessive to A across the dorsal surface of the animal, which is compatible with the observation that agouti is dominant to black. White-bellied agouti, which confers an agouti dorsal surface and a cream color over the ventral surface, is dominant to the A , a^t , and a alleles.

The a , a^t , and A^w alleles are also of interest because reverse mutations from a to a^t or A^w are more common than any other spontaneous agouti-locus mutation (Dickie, 1969). In fact, the incidence of 4.7×10^{-6} per

generation (3.29 for a^t and 1.41 for A^w) represents one of the highest reverse mutation frequencies documented in the mouse (Schlager and Dickie, 1967). This high reversion frequency implies that there may be some unusual structural features associated with the a mutation.

We and others previously cloned the agouti gene and determined that it contains four exons, the last three of which are coding (Bultman et al., 1992; Miller et al., 1993). The gene produces a 0.8-kb mRNA in A/A mice that encodes a novel 131-amino-acid protein that includes a putative signal peptide. The A protein has been hypothesized to be a secreted signaling molecule that functions as an antagonist to the α -melanocyte stimulating hormone receptor on the surface of melanocytes (Bultman et al., 1992; Jackson, 1993). Because the structure and expression of the agouti gene are altered by the most dominant agouti-locus mutation (lethal yellow, A^y), mutations in the middle of the dominance hierarchy (a and a^t), and the most recessive member of the agouti allelic series (extreme nonagouti, a^e), we proposed that there is only one gene associated with agouti-locus function (Bultman et al., 1992).

Here we report that the a , a^t , and A^w alleles are associated with DNA insertions at precisely the same position in the first intron of the agouti gene. A portion of each inserted fragment corresponds to a member of the VL30 family of retrovirus-like transposable elements in the mouse. We describe how homologous recombination events within the integrated sequences of the a allele are associated with the reversions from a to a^t or A^w . Additionally, we have discovered that the agouti gene produces alternatively processed transcripts that are expressed in the a^t and A^w alleles but not in the A allele. We propose that the insertions in a , a^t , and A^w confer their unique coat-color phenotypes by selectively inactivating the expression of the different processed forms of agouti mRNA.

MATERIALS AND METHODS

Animals

All of the new a^t and A^w mutants reported here arose spontaneously in the production colony at The Jackson Laboratory (see Table 1), except for

a^t -2 G^{so} (formerly referred to as SB B+T), which arose spontaneously at the Oak Ridge National Laboratory in a cross between the strains SEC/E (a/a , b/b , c^{ch}/c^{ch}) and C57BL/E (a/a). All animals were subsequently maintained at the Oak Ridge National Laboratory.

Cloning Overlapping Genomic DNA Fragments that Span the Insertions in a , a^t , and A^w

A total of 250 μ g of genomic liver DNA from wild-type A/A (C3H), a/a (C57BL/E), or a^t/a^t (a^t -2 G^{so} stock) mice were digested with either *Bam*HI or *Bgl*II, and were size fractionated on a 10%-40% sucrose gradient (Sambrook et al., 1989). Fractions containing appropriately sized fragments (see below) were ligated into the EMBL3 vector, packaged *in vitro*, and screened using standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). Utilizing a fragment that includes the first agouti exon plus some 5' flanking sequence and part of the first intron (probe 1.5, Fig. 1), *Bam*HI fragments of 8 kb, 16 kb and 11 kb were cloned from A/A , a/a , and a^t/a^t DNAs, respectively. Utilizing a 0.6-kb *Bam*HI-*Pst*I fragment in the first intron of agouti (probe 0.6, Fig. 1), *Bgl*II fragments of 2 kb, 6 kb, and 7 kb were cloned from A/A , a/a , and a^t/a^t DNAs, respectively. These fragments were subsequently subcloned into the pGEM (Promega) or pBluescript (Stratagene) vectors and purified as described (Ausubel et al., 1988; Sambrook et al., 1989). For A^w , a total of 250 μ g of genomic liver DNA was prepared from the original A^w -38J mutant, partially digested with *Sau*3A, and size fractionated on a 10%-40% sucrose gradient (Sambrook et al., 1989). Fractions containing 15-20 kb fragments were ligated into the EMBL3 lambda vector, packaged *in vitro*, and screened using standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). Clones that hybridized to probes 1.5 and 0.6 (see Fig. 1) were purified, and subfragments were subsequently isolated and subcloned into pGem or pBluescript (Stratagene) for further analysis.

DNA Sequencing

Selected portions of genomic clones were sequenced by the dideoxynucleotide method (Sanger et al., 1977) using T7 DNA polymerase (US Biochemical) (Tabor and Richardson, 1987). Analysis of the DNA

sequence was conducted using the University of Wisconsin Genetics Computing Group sequence analysis programs (Devereux et al., 1984), and GenBank database searches were performed using the BLAST algorithm (Altschul et al., 1990).

Southern Blot Analysis

Ten μg of genomic adult liver or tail DNA were digested with restriction enzymes, electrophoresed through 0.8% agarose gels, and blotted to GeneScreen (DuPont) with standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). Radiolabeled hybridization probes were prepared with the random-hexamer labeling technique (Feinberg and Vogelstein, 1984), and posthybridization filter washing was conducted under high stringency conditions (0.2X SSC, 0.1% SDS, at 68°C).

RT-PCR Analysis

Ten μg of total RNA, prepared by the guanidine isothiocyanate procedure (Ausubel et al., 1988), were reverse transcribed (Kawasaki, 1990), ethanol precipitated, and resuspended in 20 μl of H_2O . PCR was performed using 4 μl of the sample as described previously (Pieretti et al., 1991). The sequences of the oligonucleotide primers in Fig. 6 are as follows: Primer 1, 5'-G A A C T G G C A T C A A A G T A C C A; primer 2, 5'-CAATGCTCCTGCCTCTGCCA; primer 3, 5'-TTCCGCTTCTCGGCTTCTTT. The β -actin primers were 5'-ATGGGTCAGAAGGACTCCTA (forward) and 5'-CAACATAGCACAGCTTCTCT (reverse). PCR products were electrophoresed through 3.0% agarose gels, transferred to Gene Screen Plus (DuPont), and hybridized to an agouti cDNA probe as described above.

RESULTS

Structural DNA Rearrangements Associated with the α , α^t , and A^w Alleles

As part of our recent molecular characterization of the agouti gene, we identified DNA insertions within a 700-bp interval in the first intron both in α and in an α^t mutation called $\alpha^t\text{-}2Gso$ (formerly referred to as SB B+T,

Bultman et al., 1992). This 700-bp region is defined by *EcoRI* and *BglII* restriction sites and is illustrated in Fig. 1A. More recently, we determined that a new *Aw* mutation called *Aw-38J* also contains a DNA insertion in this same 700-bp interval (see below). To elucidate the molecular nature of the inserted sequences, overlapping DNA fragments spanning the insertions were cloned from the α , α^{t-2Gso} , and *Aw-38J* alleles. In all cases, the structure of the mutant fragments was compared to the corresponding cloned wild-type (*A*) region (Fig. 1). These analyses demonstrated that insertions of approximately 11 kb, 6 kb, and 0.6 kb are present in the first intron of the α , α^{t-2Gso} , and *Aw-38J* alleles, respectively (Fig. 1).

DNA sequence analysis of the cloned fragments revealed that the ends of the inserted DNA are identical in α , α^{t-2Gso} , and *Aw-38J* and correspond to the 5' and 3' regions of the long terminal repeat (LTR) of a retrovirus-like VL30 element (Figs. 1 and 2). In fact, restriction mapping and DNA sequencing analyses of the inserted regions indicated that α and α^t each contain a complete 5.5-kb VL30 element, nearly identical to the element already reported (Adams et al., 1988; Hodgson et al., 1990) (Fig. 1). The VL30 sequences within both α and α^{t-2Gso} each contain an additional DNA segment that is unique to each allele (see below), while in *Aw-38J* only a single LTR is present (Fig. 1). The VL30 sequences in all three alleles are integrated in the opposite transcriptional orientation relative to agouti (Figs. 1 and 2).

For both the α and α^{t-2Gso} mutations, sequence analyses indicated that the additional segment of DNA within the VL30 element interrupts the VL30 sequence between nucleotides 2473 and 2474 of the published sequence (Hodgson et al., 1990) (Fig. 1). The size of this additional sequence was determined to be 5.5 kb for the α allele and 526 bp for the α^{t-2Gso} allele. The complete nucleotide sequence of the 526-bp insert within the VL30 element of α^{t-2Gso} was determined and compared to the sequence of selected regions from the α allele. This analysis revealed that the extra 5.5-kb insert within the VL30 sequence in α contains direct terminal repeats of 526 bp in length, each of which corresponds to the single 526-bp insert in the α^{t-2Gso} allele (Fig. 1). Comparison of the sequence of this 526-bp direct repeat with GenBank indicated that it is 90% identical to a region between two T-cell

receptor (TCR) γ variable region gene segments, V10.8A and V10.8B, which lie 3.5 kb apart in an opposite transcriptional orientation (Hayday et al., 1985; Garman et al., 1986) (Fig. 3). The functional significance of this homology is presently unclear because the 526-bp region does not correspond to a defined regulatory element or cis-acting sequence responsible for somatic recombination events that produce functional TCRs. Furthermore, hybridization of the 526-bp region to total mouse DNA revealed that it is repetitive in the mouse genome, suggesting that the 526-bp region is a member of an, as yet, uncharacterized family of DNA repetitive elements (data not shown). The nature of the other sequences within the inserted 5.5-kb of DNA in the α allele is currently under investigation.

Characterization of the sequence flanking the inserted VL30 sequences for the α , α^t -2Gso, and A^w -38J alleles showed that the integrated DNA occurs at precisely the same position in each case (Fig. 2). Furthermore, a duplication of four nucleotides of host DNA at the integration site, and the fact that two bp are missing from the end of each LTR, suggests that the integration of VL30 into the agouti gene occurred through the same mechanism that is normally associated with the integration of retroviral sequences (Dhar et al., 1980; Shimotohno et al., 1980; Van Buren et al., 1980; Varmus, 1982). This finding is in accord with VL30 integrations that have been characterized previously (Hodgson et al., 1983; Itin and Keshet, 1983).

Molecular Analysis of Several New α^t and A^w Reverse Mutations

Based on the analysis of the α^t -2Gso and A^w -38J mutations, we hypothesize that each arose by excision of inserted DNA from the α allele through homologous recombination. In the case of α^t -2Gso, we predict that homologous recombination between the 526-bp direct repeats occurred to remove most of the additional 5.5 kb of inserted sequence, and in A^w -38J we predict that homologous recombination between the VL30 LTRs removed most of the 11 kb of inserted sequence present in the α allele (Fig. 1). If homologous recombination is indeed playing a role in the reversion from α to α^t or A^w , then one would expect several independent α^t and A^w mutations that arise from α to exhibit the same structural alterations as those

observed for α^{t-2Gso} and A^{w-38J} , respectively. To test this possibility, we analyzed DNA from twenty-one new mutant animals that arose in the C57BL/6J (α/α) colony at The Jackson Laboratory over the past two years. Eleven of these new mutants exhibited a phenotype identical to α^t , and ten others displayed a phenotype identical to A^w (Table 1). All of these mutations pass through the germline and presumably represent spontaneous reverse mutations from α to α^t or A^w .

Genomic DNA from each of the original mutant animals listed in Table 1 was analyzed utilizing probe 0.6, which lies just 3' to the insertion site in the α , α^{t-2Gso} , and A^{w-38J} mutations (see Fig. 1). This probe hybridizes to *Bgl*II fragments of 6 kb, 7 kb, and 2.6 kb in DNA from the α , α^{t-2Gso} , and A^{w-38J} mutations, respectively (Figs. 4 and 5). As expected, DNA from each of the eleven new α^t mutant animals contained a *Bgl*II fragment of 7 kb, which is identical in size to that of α^{t-2Gso} (Fig. 4). Similarly, DNA from each of the 10 new A^w mutant animals contained a *Bgl*II fragment of 2.6 kb, which is identical in size to that of A^{w-38J} (Fig. 5). The agouti alleles in the inbred strains 101 (A^w/A^w) and C3H (A/A), in contrast, are associated with a 2-kb *Bgl*II fragment, indicating the absence of any inserted sequence (Fig. 5). Further analysis of genomic DNA from each of the twenty-one new reversion mutants indicated that they also have the same size-altered *Bam*HI fragments as those observed for the α^{t-2Gso} and A^{w-38J} mutations when hybridized with a probe that lies just 5' to the DNA insertions (see probe 1.5 in Fig. 1; data not shown). Based on these data, we propose that all of the new reverse mutations contain the same DNA structural alterations that are present in the α^{t-2Gso} and A^{w-38J} mutations. Furthermore, because we analyzed each of the "original" mutants we were able to determine that the α^t reversions arose through intrachromosomal recombination. If the reversions had occurred through interchromosomal recombination, then we would have detected a size-altered restriction fragment associated with the reciprocal recombination product on the other chromosome, and we did not (data not shown). Due to the relatively large size of the DNA segment that is excised from α in A^w revertants, we have been unable to identify a diagnostic restriction enzyme necessary to perform a similar analysis for A^w .

Additional Forms of Agouti Transcripts

As part of our recent molecular analysis of the lethal-yellow (A^y) mutation, we determined that A^y produces chimeric transcripts in which the first noncoding exon of agouti is replaced by the noncoding first exon of a closely linked gene called *Raly* (hnRNP associated with lethal yellow) (Michaud et al., 1993). During the course of this analysis, a number of A^y cDNA clones were isolated that contained differentially spliced segments, 111 bp and 46 bp in length, integrated between the *Raly* and agouti exons (Michaud et al., 1993). These segments have been mapped to a region immediately 3' to a deletion breakpoint associated with the A^y mutation, which positions them approximately 110 kb upstream from the first coding exon of agouti (Michaud et al., 1994). We initially attributed these segments to cryptic splicing events associated with the A^y allele but could not rule out the possibility that the 111-bp and 46-bp segments actually represent additional 5' exons of the agouti gene.

Here we demonstrate that the 111-bp and 46-bp segments are differentially incorporated into agouti-specific transcripts that are structurally different from the agouti mRNA originally described (Bultman et al., 1992; Miller et al., 1993). The RNAs that incorporate the 111-bp and/or 46-bp segments, here referred to as form II transcripts, were detected with a reverse transcriptase-polymerase chain reaction (RT-PCR) assay using an oligonucleotide primer from the 111-bp region (primer 1) coupled with a primer from the third agouti exon (primer 3) (Fig. 6A). These primers produced PCR products of 380 bp and 334 bp in size, depending on the presence or absence of the 46-bp exon in skin from the ventrum, but not the dorsum, of four and five day old A^w/a , a^t/a^t , and A^w/A mice (Fig. 6B). Both of the form II transcripts are also present in A^y/a neonatal skin (and in adult tissues, see Michaud et al., 1993a), which was included as a positive control. In contrast, neither of the form II transcripts could be detected as ethidium-stained RT-PCR fragments in the ventrum or the dorsum of age-matched neonatal A/A or a/a mice (Fig. 6B), which explains why these forms of agouti mRNA were not detected originally (Bultman et al. 1992).

Additionally, a similar RT-PCR assay was utilized to detect the presence of agouti transcripts that incorporate the 5' noncoding exon that we originally identified as the first agouti exon (Bultman et al. 1992). This form of agouti mRNA, here referred to as the form I transcript, was detected with a primer from the originally defined noncoding first exon (primer 2) coupled with same third exon primer as above (primer 3). With these primers, form I mRNA produced a single 315-bp PCR product of the expected size (Fig. 6A). Form I transcripts were detected in both the dorsum and the ventrum of four and five day old *A/A*, *A^w/a*, and *A^w/A* mice, but not in the dorsum or the ventrum of age-matched *a/a* or *a^t/a^t* mice (Fig. 6C).

Overall, these RT-PCR analyses demonstrate that *A/A* mice produce form I transcripts in both their dorsal and ventral skin but do not produce form II transcripts. Conversely, *a^t/a^t* animals do not express form I transcripts and express form II transcripts but only in ventral skin and not in dorsal skin. Animals carrying the *A^w* allele (derived either from *A^{w-42J}* or 101) express both the form I and the ventral-specific form II transcripts, and *a/a* animals do not express detectable levels of either the form I or the form II transcripts (Table 2).

DISCUSSION

DNA Insertions in *a*, *a^{t-2Gso}*, and *A^{w-38J}*

As part of our recent molecular characterization of the agouti gene, we determined that the *a* and *a^{t-2Gso}* mutations each contain an extra fragment of DNA within an intervening sequence of the gene (Bultman et al., 1992). More recently, we determined that a new *A^w* mutation, called *A^{w-38J}*, also contains an extra fragment of DNA within this same region. Based on the analysis of genomic clones that span these inserted sequences, we now report that the insertions are 11 kb, 6 kb, and 0.6 kb in *a*, *a^{t-2Gso}*, and *A^{w-38J}*, respectively, and that they are all integrated at precisely the same position. The 11-kb insertion in *a* includes an entire 5.5-kb VL30 element plus an internal 5.5 kb of additional sequence flanked by direct terminal repeats of 526 bp. The 6-kb insertion in *a^{t-2Gso}* also includes an entire VL30

element but contains only one 526-bp direct repeat internally. The 0.6-kb insertion in *A^{w-38J}* corresponds to a single, intact VL30 LTR.

VL30 elements are retrovirus-like structures, approximately 5.5 kb in size, that are dispersed throughout rodent genomes at about 100 to 200 copies (Besmer et al., 1979; Keshet et al., 1980; Keshet and Itin, 1982). Although VL30 elements have not been shown to encode any polypeptides and are replication-defective, RNA copies of their genome are capable of being co-packaged by endogenous murine leukemia viruses and retrotransposed when the pseudovirion infects new cells and the VL30 RNA undergoes reverse transcription and integrates into the host cell genome (Scolnick et al., 1979; Carter et al., 1986). Because retroviral-like elements, particularly intracisternal A particles, have been shown to affect the expression of numerous genes, even when integrated in the opposite transcriptional orientation (Kuff and Lueders, 1988), we propose that the inserted VL30 elements in the *a* and *a^t* insertions are responsible for the alterations in agouti gene expression (see below).

Reverse Mutations From *a* to *a^t* and *A^w* are Mediated by Homologous Recombination

Comparison of the structure of the *a*, *a^t*, and *A^w* insertions suggests a model that explains the high rate of reverse mutations from *a* to *a^t* or *A^w*. Based on the finding that the VL30 element in *a^{t-2Gso}* contains a single copy of the 526-bp direct repeat associated with *a*, we propose that the reversion from *a* to *a^t* occurs by excision of 5 kb of DNA (the sequence integrated in the VL30) from the *a* allele, brought about by homologous recombination between the 526-bp direct terminal repeats. Likewise, since the insertion in *A^{w-38J}* is comprised of only a single VL30 LTR, we propose that the reversion from *a* to *A^w* occurs by excision of the entire insertion in *a* except for one VL30 LTR, brought about by homologous recombination between the VL30 LTRs. This model is supported by the fact that each of eleven new *a^t* mutations and ten new *A^w* mutations appear to have the same structural rearrangements as those found in the *a^{t-2Gso}* or *A^{w-38J}* mutations. Additional support for a homologous recombination mechanism comes from the fact that the reversion of *a* to *a^t* and *A^w* occurs with a frequency of $4.7 \times$

10^{-6} , which is comparable to the reversion frequency of 3.3×10^{-6} for the original dilute (*d*) coat-color mutation (Schlager and Dickie, 1967). The *d* mutation was shown to be caused by a retroviral insertion, and reversion to wild type occurs by homologous recombination between the LTRs (Copeland et al., 1983). Although we expect that this model will explain the genetics of *a*, *a^t*, and *A^w*, we cannot definitively rule out the possibility that other structural alterations might also be associated with these alleles.

Multiple Forms of Agouti mRNA

We previously demonstrated that the agouti gene produces a 0.8-kb mRNA in neonatal skin and that the abundance of this mRNA correlates well with the presence of phaeomelanin (Bultman et al., 1992; Michaud et al., 1993). However, RT-PCR experiments presented here clearly demonstrate that, instead of producing a single species of mRNA, the agouti gene produces a number of similarly-sized transcripts that are different at their 5' ends. One class of transcript, referred to here as form I, consists of sequence derived from the first, second, third, and fourth agouti exons as originally described (Bultman et al., 1992; Miller et al., 1993). This form I transcript is expressed in skin from both the ventrum and the dorsum of four and five day old *A/A*, *A^w/a*, and *A^w/A* mice, but neither in the ventrum nor the dorsum of age-matched *a/a* or *a^t/a^t* mice (Table 2). Furthermore, expression of the form I transcript is regulated during early postnatal development such that it is produced while melanocytes are actively depositing phaeomelanin into agouti hair follicles between days three and seven (Galbraith, 1964; Sakurai et al., 1975), but is absent both before and after this interval (data not shown). A second class of transcripts detected by RT-PCR, referred to here as form II, consists of differentially incorporated 111-bp and 46-bp segments spliced to the agouti coding exons. Both of these transcripts are expressed in skin from the ventrum, but not the dorsum, of four and five day old *A^w/a*, *a^t/a^t*, and *A^w/A* mice. In contrast, form II transcripts are not detected in the ventrum or the dorsum of age-matched *A/A* or *a/a* mice (Table 2); this is the reason why these different forms of agouti mRNA were not originally identified in the analysis of the *A* allele, which was presumed to be wild type (Bultman et al. 1992). Furthermore, the form II transcripts

are expressed throughout the hair-growth process in the ventrum of A^w/A mice (data not shown). Taken together, these findings provide compelling evidence that the form I transcript is associated with the production of the subapical band of phaeomelanin in agouti hair throughout the pelage and that the form II transcripts are responsible for the predominating yellow pigmentation over the ventral surface in the A^w and a^t alleles.

The fact that A/a^t mice have a phenotype that is identical to that of A^w (Dunn, 1928; Silvers, 1979) is consistent with the expression data presented here. A^w produces both form I and form II transcripts, whereas A produces only the form I transcript and a^t produces only form II transcripts. A/a^t mice would be expected to produce both forms of transcripts as a result and exhibit a phenotype identical to that of A^w .

From our analyses, form I and II agouti transcripts differ only in their 5' noncoding regions and apparently encode the same protein. Additional analysis of the 111-bp segment has revealed that it may be comprised of two separate exons (Michaud et al., 1994), neither of which contains a termination codon (Michaud et al., 1993). For this reason, we cannot exclude the possibility that the 334-bp form II mRNA that contains only the 111-bp segment actually has a 5' extension of the agouti open reading frame. Also, because the 111-bp and 46-bp segments each contains a functional splice acceptor at its 5' junction (Michaud et al., 1993; unpublished data), we predict that there must be at least one more exon upstream of the 111-bp segment which is normally at the 5' end of form II transcripts. These matters are currently under investigation.

The Insertions in a and a^t Modulate the Expression of Form I and Form II Agouti Transcripts

Comparison of the structure of the a , a^t , and A^w alleles suggests a model that explains how each insertion alters agouti gene expression. We propose that the VL30 element, along with the 5.5 kb of additional sequence, essentially shuts off agouti gene expression in the a mutation by negatively regulating the production of both form I and form II mRNA. This is presumably the case throughout the pelage, except for very restricted areas near the pinnae, mammae, and perineum, where phaeomelanin

deposition still occurs. It is likely that any agouti transcripts produced from the mammae in the samples we analyzed (skin samples containing the pinnae and perineum were not used in our experiments) were present at a sufficiently low level as to be undetectable with our molecular techniques. Additionally, we propose that in the α^t mutation, the inserted segment of DNA (VL30 containing a single copy of the 526-bp direct repeat) downregulates the production of form I transcripts but does not significantly affect the production of form II transcripts. Perhaps the inserted element in α^t is acting in a localized manner such that it does not influence the expression of the form II-specific exons, which are located approximately 100 kb upstream of the inserted VL30 sequences. The mechanism through which the inserted elements in α and α^t alter agouti expression is currently under investigation.

A^w is Wild Type

As demonstrated here, all of the A^w alleles that arise through reversion from α contain a solo VL30 LTR in the first intron of the agouti gene. It is possible that this solo LTR is redirecting agouti gene expression in a manner that is consistent with the production of form I and form II mRNAs in the A^w allele. Alternatively, based on the analysis of wild-type reversion alleles of the dilute mutation (Copeland et al., 1983), it is possible that the remaining LTR has no functional significance and does not interfere with the normal expression of the agouti gene. We favor the latter possibility based on our analysis of the A^w allele in the 101 inbred strain; this A^w allele does not contain a solo VL30 LTR in the same site as does A^{w-38J} or any of the other reversion mutants that we have thus far analyzed (S.J.B., H.O.S., M.T.D., and R.P.W., unpublished results). This implies that the original α mutation arose on a white-bellied agouti background. Furthermore, since form II transcripts are expressed in A^w but not in A , we propose that A^w is the true wild-type allele and that A actually represents a loss-of-function mutation in which form II transcripts are not expressed. This is of particular interest because both A and A^w have traditionally been referred to as wild-type agouti alleles. The proposal that A^w is the true wild-type agouti allele is not only compatible with the molecular data

presented here but is also consistent with reports that most populations of wild mice have a white-bellied agouti phenotype (Silvers, 1979).

REFERENCES

- Adams, S. E., Rathjen, P. D., Stanway, C. A., Fulton, S. M., Malim, M. H., Wilson, W., Ogden, J., King, L., Kingman, S. M., and Kingsman, A. J. (1988). Complete nucleotide sequence of a mouse VL30 retro-element. *Mol. Cell Biol.* **8**: 2989-2998.
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215**: 403-410.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1988). *Current protocols in molecular biology*. John Wiley & Sons/Greene, New York.
- Besmer, P., Olshevsky, U., Baltimore, D., Dolberg, D., and Fan, H. (1979). Virus-like 30S RNA in mouse cells. *J. Virology* **29**: 1168-1176.
- Bultman, S. J., Michaud, E. J., and Woychik, R. P. (1992). Molecular characterization of the mouse agouti locus. *Cell* **71**: 1195-1204.
- Carter, A. T., Norton, J. D., Gibson, Y., and Avery, R. J. (1986). Expression and transmission of a rodent retrovirus-like VL30 gene family. *J. Mol. Biol.* **188**: 105-108.
- Copeland, N. G., Hutchison, K. W., and Jenkins, N. A. (1983). Excision of the DBA ecotropic provirus in dilute coat-color revertants of mice occurs by homologous recombination involving the viral LTRs. *Cell* **33**: 379-387.
- Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**: 387-395.
- Dhar, R., McClements, W. L., Enquist, L. W., and Vande Woude, G. F. (1980). Nucleotide sequences of integrated Moloney sarcoma provirus long terminal repeats and their host and viral junctions. *Proc. Natl. Acad. Sci.* **77**: 3937-3941.

- Dickie, M. M. (1969). Mutations at the agouti locus in the mouse. *J. Hered.* **60**: 20-25.
- Dunn, L. C. (1928). A fifth allelomorph in the agouti series of the house mouse. *Proc. Natl. Acad. Sci.* **14**: 816-819.
- Feinberg, A. P., and Vogelstein, B. (1984). Addendum: A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**: 266-267.
- Galbraith, D. B. (1964). The agouti pigment pattern of the mouse: A quantitative and experimental study. *J. Exp. Zool.* **155**: 71-90.
- Garman, R. D., Doherty, P. J., and Raulet, D. H. (1986). Diversity rearrangement, and expression of murine T cell gamma genes. *Cell* **45**: 733-742.
- Green, M. C. (1989). Catalogue of mutant genes and polymorphic loci. In *Genetic variants and strains of the laboratory mouse*, (eds. M.F. Lyon, and A.G. Searle), pp. 17-20. Oxford Univ. Press, Oxford and New York.
- Hayday, A. C., Saito, H., Gillies, S., Kranz, D. M., Tanigawa, G., Eisen, H. N., and Tonegawa, S. (1985). Structure, organization, and somatic rearrangement of T cell gamma genes. *Cell* **40**: 259-269.
- Hodgson, C. P., Elder, P. K., Ono, T., Foster, D. N., and Getz, M. J. (1983). Structure and expression of mouse VL30 genes. *Mol. Cell Biol.* **3**: 2221-2231.
- Hodgson, C. P., Fisk, R. Z., Arora, P., and Chotani, M. (1990). Nucleotide sequence of mouse virus-like (VL30) retrotransposon BVL-1. *Nucl. Acids Res.* **18**: 673.

- Itin, A., and Keshet, E. (1983). Nucleotide sequence analysis of the long terminal repeat of murine virus-like DNA (VL30) and its adjacent sequences: resemblance to retrovirus proviruses. *J. Virology* **47**: 656-659.
- Jackson, I. J. (1993). More to colour than meets the eye. *Current Biology* **3**: 518-521.
- Kawasaki, E. S. (1990). Amplification of RNA. In *PCR protocols: A guide to methods and applications* (e.d. M.A. Innis, D.H., Gelfand, J.J. Sninsky, and T.J. White), pp. 21-27. Academic Press, San Diego, CA.
- Keshet, E., Shaul, Y., Kaminchik, J., and Aviv, H. (1980). Heterogeneity of "virus-like" genes encoding retrovirus-associated 30S RNA and their organization with the mouse genome. *Cell* **20**: 431-439.
- Keshet, E., and Itin, A. (1982). Patterns of genomic distribution and sequence heterogeneity of a murine "retrovirus-like" multigene family. *J. Virology* **43**: 50-58.
- Kuff, E. L., and Lueders, K. K. (1988). The intracisternal A-particle gene family: structure and functional aspects. *Adv. Cancer Res.* **51**: 183-276.
- Miller, M. M., Duhl, D. M. J., Vrieling, H., Cordes, S. P., Ollmann, M. M., Winkes, B. M., and Barsh, G. S. (1993). Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation. *Genes & Dev.* **7**: 454-467.
- Michaud, E. J., Bultman, S. J., Stubbs, L. J., and Woychik, R. P. (1993). The embryonic lethality of homozygous lethal yellow mice (A^y/A^y) is associated with the disruption of a novel RNA-binding protein. *Genes & Dev.* **7**: 1203-1213.

- Michaud, E. J., Bultman, S. J., Klebig, M. L., van Vugt, M. J., Stubbs, L. J., Russell, L. B., and Woychik, R. P. (1994). A molecular model for the genetic and phenotypic characteristics of the mouse lethal yellow (A^y) mutation. *Proc. Natl. Acad. Sci.*, in press.
- Pieretti, M., Zhang, F., Fu, Y.-H., Warren, S. T., Oostra, B. A., Caskey, C. T., and Nelson, D. L. (1991). Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* **66**: 817-822.
- Sakurai, T., Ochia, H., and Takeuchi, T. (1975). Ultrastructural change of melanosomes associated with agouti pattern formation in mouse hair. *Dev. Biol.* **47**: 466-471.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: A laboratory manual, Second Edition*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74**: 5463-5467.
- Schlager, G., and Dickie, M.M. (1967). Spontaneous mutations and mutation rates in the house mouse. *Genetics* **57**: 319-330.
- Scolnick, E. M., Vass, W. C., Howk, R. S., and Duesberg, P. H. (1979). Defective retrovirus-like 30S RNA species of rat and mouse cells are infectious if packaged by type C helper virus. *J. Virol.* **29**: 964-972.
- Shimotohno, K., Mizutani, S., and Temin, H. M. (1980). Sequence of retrovirus provirus resembles that of bacterial transposable elements. *Nature* **285**: 550-554.
- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The Coat Colors of Mice: A Model for Mammalian Gene Action and Interaction*, pp. 6-44. Springer-Verlag, New York.

- Tabor, S. and Richardson, C. C. (1987). DNA sequence analysis with a modified form of T7 DNA polymerase. *Proc. Natl. Acad. Sci.* **84**: 4767-4771.
- Van Buren, C., Goddard, J. G., Berns, A., Verma, I. M. (1980). Structure of Moloney murine leukemia viral DNA: nucleotide sequence of the 5' long terminal repeats and adjacent cellular sequences. *Proc. Natl. Acad. Sci.* **77**: 3307-3311.
- Varmus, H. E. (1982). Form and function of retroviral proviruses. *Science* **216**: 812-820.

APPENDIXES

APPENDIX A**TABLES**

Table 1. New A^W and a^t mutations analyzed by Southern analysis.

<u>Remutation Number</u>	<u>Strain of Origin</u>
<i>A^W-34J</i>	C57BL/6J
<i>A^W-35J</i>	B10.BL/SgSnJ
<i>A^W-36J</i>	C57BL/6J
<i>A^W-37J</i>	C57BL/6J
<i>A^W-38J</i>	C57BL/6J
<i>A^W-39J</i>	C57BL/6JN-Sp
<i>A^W-40J</i>	C57BL/6J
<i>A^W-41J</i>	C57BL/KsJ- <i>dbm</i>
<i>A^W-42J</i>	C57BL/6J
<i>A^W-43J</i>	C57BL/6J
<i>A^W-44J</i>	C57BL/6J
<i>a^t-46J</i>	C57BL/6J
<i>a^t-47J</i>	C57BL/KsJ- <i>dbm</i>
<i>a^t-48J</i>	C57BL/6J
<i>a^t-49J</i>	C57BL/6J
<i>a^t-50J</i>	C57BL/6J
<i>a^t-51J</i>	C57BL/6J
<i>a^t-52J</i>	C57BL/6J
<i>a^t-53J</i>	C57BL/6J
<i>a^t-54J</i>	C57BL/6J
<i>a^t-55J</i>	C57BL/KsJ- <i>dbm</i>
<i>a^t-56J</i>	C57BL/6J x DBA/2J

Table 2. Dorsal (D) and ventral (V) distribution of form I and form II transcripts in neonatal skin from mice with different agouti genotypes.

Genotype	Form I	Form II
<i>A/A</i>	D & V	-
<i>A^w/A</i>	D & V	V
<i>A^w/a</i>	D & V	V
<i>a^t/a^t</i>	-	V
<i>a/a</i>	-	-

APPENDIX B

FIGURES

Figure 1. Identification of DNA insertions in the first intron of the agouti gene in the α , α^t , and A^w alleles.

(A) Schematic representation of the genomic structure of the wild-type agouti gene with the 5' end enlarged immediately below. The four exons are depicted as solid boxes, and introns and flanking sequences are drawn as a solid, horizontal line. The insertion sites in α , α^t , and A^w are highlighted by a vertical arrow. The positions of two probes, 1.5 and 0.6, are indicated. B, *Bam*HI; R, *Eco*RI; Bg, *Bgl*II. (B) Schematic representation of the 5' end of the agouti gene in the α and α^t -2*Gso* mutations. Insertions of 11 kb in α and 6 kb in α^t are represented by rectangles. The wide portion of the rectangles represents VL30 sequence with cross-hatched regions corresponding to the LTRs. Numerals refer to nucleotides within VL30 in the 5' to 3' orientation and are based on the published sequence (Hodgson et al., 1990). The narrow portion of the rectangles represents additional sequence integrated between positions 2473 and 2474 of VL30. This additional sequence is approximately 5.5 kb in α and is bound by 526-bp direct repeats (stippled region with arrows above). All of this additional sequence is absent in the α^t insertion, except for one of the 526-bp direct repeats (indicated by dashed lines). (C) Schematic representation of the 5' end of the agouti gene in the α and A^w -38*J* alleles. All of the 11-kb insertion in α is absent in the A^w allele, except for one VL30 LTR (indicated by dashed lines).

Figure 2. Nucleotide sequence of a region of DNA spanning the insertion site in *A*, *a*, *a^{t-2Gso}*, and *A^{w-38J}*.

Wild-type agouti intron sequence is shown 5' to 3' in upper case letters. The sequences at the ends of the insertions in *a*, *a^t*, and *A^w* are identical and correspond to the LTRs of a VL30 provirus, as shown below in lower case letters. The stippled box highlights a 4-bp duplication of agouti sequence at the VL30 integration site in *a*, *a^t*, and *A^w*.

5'

3'

A

...TGTCGATAGGGTACTTGCCT AACCTGTGTGTAACCCCTGGG...

a, a^t A^w

...TGTCGATAGGGTACTTGCCT AACCTGTGTGTAACCCCTGGG...

igaagatttt...tcttcagcct

3' LTR

5' LTR

Figure 3. Nucleotide sequence comparison between the 526-bp direct terminal repeats in the α and α^t insertions (stippled region of the narrow rectangles in α and α^t in Figure 1) (top sequence) and part of an intergenic regulatory region between two TCR γ variable gene segments (bottom sequence).

The sequence is presented colinear with the arrows shown above the 526-bp repeats in Fig. 1B. The three capitalized nucleotides at the 5' end of the 526-bp repeat are present in α^t and in the repeat apposed with nucleotide 2474 of VL30 in α but are absent in the repeat apposed with nucleotide 2473 of VL30 in α .

```

1  agaCTCtgtcagagaccatcccgtgagaaagtgagcccttcacaatctcc 50
   | | |||||
1682 GAATTGTGTCAGAGACCATCCGGTGAGAAAGCGAGCCCTTCACAATCTCC 1731
      .
51  ctagcaggccaaatgg.cttggtactgagagctggttgtcacccccctttc 99
   ||| | |||||
1732 CTAACCAGCCAAATGGCCTTGTGCTAGTAGCTGGTTATTACTCCC...TC 1778
      .
100 ctccctattcctttcctggcacctgaggctgtaaaagctgaattatagtc 149
   ||| | |||||
1779 CTCCGTATTCCCTTCTTGGCACCTGAGGCTGTAAAAGCTGAATTATAGTC 1828
      .
150 ccctcttccctatctcttccctgagttcccatgccatccaaggacatgagt 199
   |||||
1829 CCCTCTTCCCTATCTCTTCTTAACTCCCATGCCATCCAAGGACATGAGT 1878
      .
200 tacgcctgagcctagcctgaccccccaaggctgtcaaggaggatcgatggt 249
   | |||||
1879 TTCGCCTGAGCCCAGCCTGACACCCAAGGCTGTCAAAGAGGATCTATGTT 1928
      .
250 ccagagataagatccagagtgcccgctgcctggcgctgacttcggcccc 299
   |||||
1929 CCAGAGATAAGATGCAGAGTGCCCGCTGCCTGGCGCCTGACTTTGGCCCC 1978
      .
300 catgtcagcagatgcccacttctttgttctttgtataattctccctcgac 349
   |||||
1979 CATGTCAGCAGATGCCCGCTTCTTTGTTCTTTGTAAAATTCTCCCTCAAC 2028
      .
350 ccctcccatattccccgcgatgtatgctttaaaaagaaggcacctcagcc 399
   |||||
2029 CCCTCCCTATT.CCCGGGATGTATGCTTTAAAAAGAAGGCACCTCAGCA 2077
      .
400 taataaacgagaccttgataggttcaatctacttggctcctcgactcttct 449
   |||||
2078 TAATAAACGAGACCTTGATAGGT..ACCCTTCCTGGTCTCCTGCCTCTCC 2125
      .
450 ttcttttactcccatttcccttccaggtttgcggtccccctcgcacccacg 499
   ||| | |||||
2126 TTCCCTTCCTCCCATCTCCTTCCAGGTTTGCGGTCCCCCTCTCACCCACG 2175
      .
500 aataactgagtcccgcgggacgggata 526
   |||||
2176 AATAACTGAATCCCGCGGGACGGGATA 2202

```

Figure 4. Southern blot analysis of reverse mutations of a to a^t .

Genomic DNA prepared from the a^t -2 G_{so} stock (a^t/a^t), nine of the new a^t mutants summarized in Table 1 (a^t/a), and the C57BL/6J genetic background on which eight of these mutations occurred on (a/a) were digested with *Bgl*III, blotted, and hybridized with a ^{32}P -labeled fragment of DNA corresponding to probe 0.6 (see Fig. 1). Sizes of the DNA fragments detected by probe 0.6 are shown on the right. The a^t -46 J and a^t -47 J mutants display the same 7-kb *Bgl*III fragment as the other a^t mutants (data not shown). The signal corresponding to the 7-kb a^t fragment is faint relative to the 6-kb a fragment in a^t -48 J due to mosaicism.

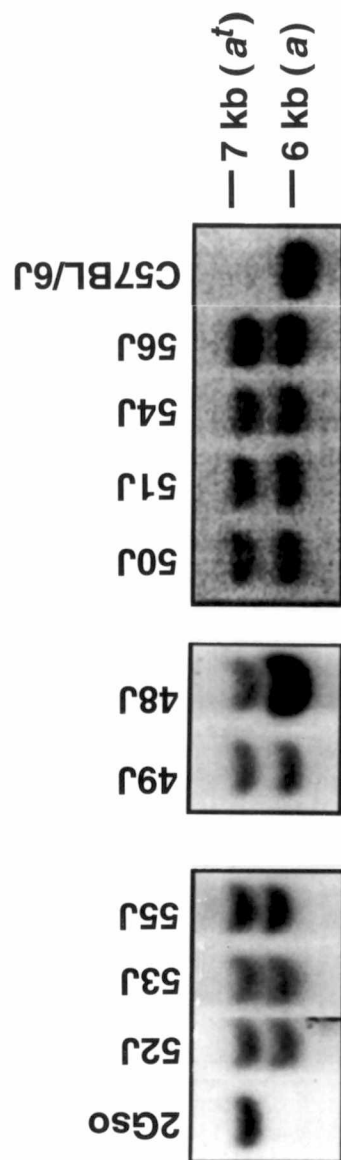


Figure 5. Southern blot analysis of reverse mutations of a to A^w .

Genomic DNA prepared from the eleven new A^w mutants described in Table 1 (A^w/a), the C57BL/6J genetic background on which nine of these mutations occurred on (a/a), and the inbred strains 101 (A^w/A^w) and C3H (A/A) were digested with *Bgl*II, blotted, and hybridized with a ^{32} P-labeled fragment of DNA corresponding to probe 0.6 (see Figure 1). Sizes of the DNA fragments detected by probe 0.6 are shown on the right. The signal corresponding to the 2.6-kb A^w fragment is faint relative to the 6-kb a fragment in A^{w-35J} , A^{w-39J} , and A^{w-41J} due to mosaicism. $A^{w-41J}.1$ (A^w/a) is a non-mosaic offspring of the original A^{w-41J} mutant that was included because the 2.6-kb A^w fragment is more apparent.

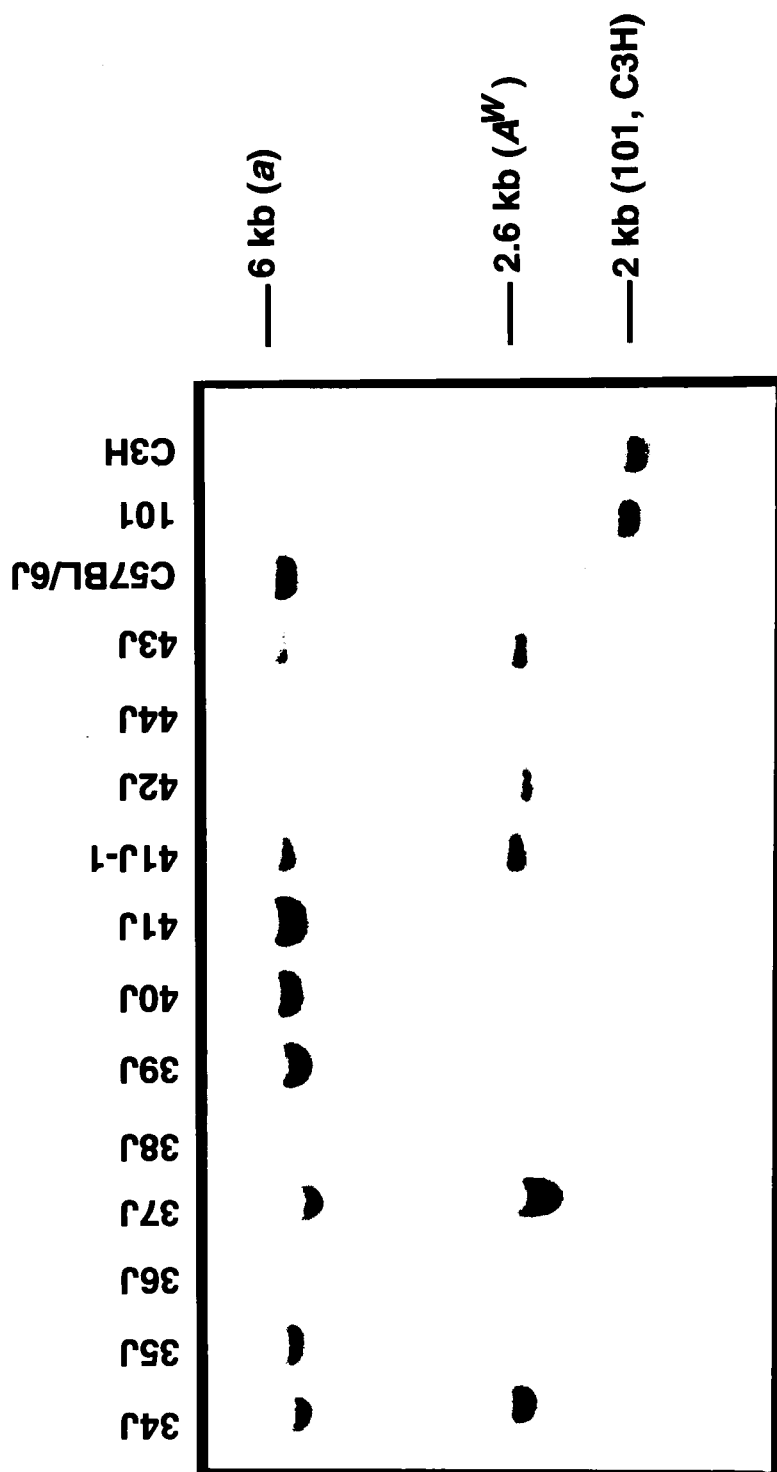
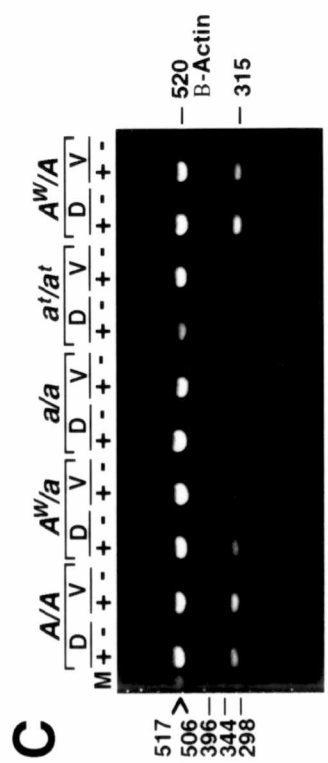
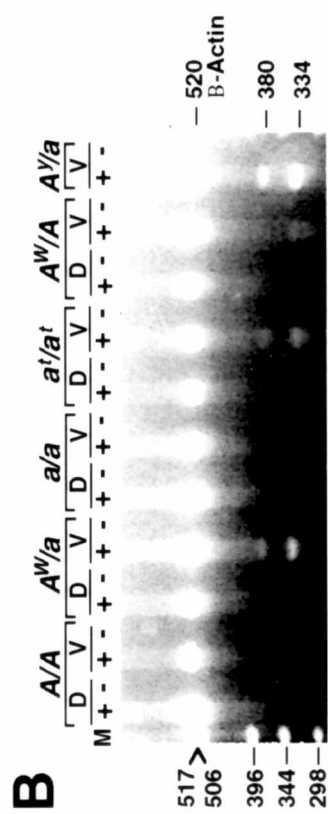
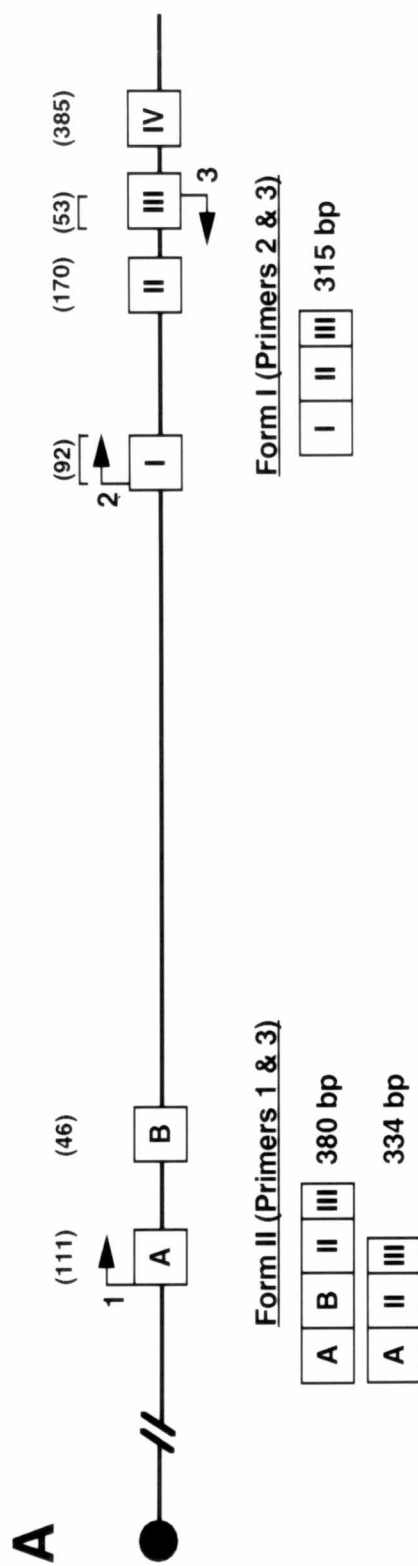


Figure 6. Identification of two forms of agouti transcripts by RT-PCR.

(A) A schematic map of the agouti gene showing the approximate locations of exons (boxes labeled A, B, I, II, III, and IV) and oligonucleotide primers (arrows labelled 1, 2, and 3) used in the PCR analysis. The map is shown 5' to 3' (left to right), and the position of the chromosome 2 centromere is indicated to the left (Michaud et al. 1994). The size of each exon or portion of exon amplified with the indicated primers is given in bp enclosed in parentheses above. Each of the boxes labelled A and B may consist of one or more exons (see Discussion). The RT-PCR products detected are shown below the map. Primers 1 and 3 amplify 380-bp and 334-bp portions of form II transcripts that consist of the exons indicated, whereas primers 2 and 3 amplify a 315-bp segment of the form I transcript comprised of the indicated exons. (B and C) RT-PCR products amplified with primers 1 and 3 (B) or primers 2 and 3 (C) from total RNA treated with (+) or without (-) reverse transcriptase. All of the RNA samples were derived from the dorsal (D) or ventral (V) skin of 4 to 5 day old neonatal mice of the following genotypes and strains: A/A (C3H_f/Rl), A^w/a (A^{W-42J} stock), a/a (C57BL), a^t/a^t (original), A^w/A (C3H x 101), and A^y/a . Shown are photographs of RT-PCR products size-fractionated in agarose gels and stained with ethidium bromide. Numbers to the left of each panel indicate the sizes in bp of the 1 kb ladder molecular weight marker (lane M); numbers to the right of the panels indicate the sizes of the RT-PCR products. The 520-bp fragment was amplified using oligonucleotide primers specific for the β -actin mRNA and serves as a positive control. Although the 380-bp fragment of the form II transcript is not readily visible in the ventral skin of A^w/A (panel B), it and the other agouti-specific fragments hybridized to an agouti cDNA probe after the gel was Southern blotted to a nylon membrane (data not shown).



PART VI

**THE EMBRYONIC LETHALITY OF HOMOZYGOUS LETHAL
YELLOW MICE (A^y/A^y) IS ASSOCIATED WITH THE DISRUPTION
OF A NOVEL RNA-BINDING PROTEIN**

ABSTRACT

Lethal yellow (A^y) is a mutation at the mouse agouti locus that is associated with an all yellow coat color, obesity, diabetes, and tumors in heterozygotes, and preimplantation embryonic lethality in homozygotes. We previously cloned and characterized the wild-type agouti gene and demonstrated that it expresses a 0.8-kb mRNA in neonatal skin. In contrast, lethal yellow expresses a 1.1-kb transcript that is ectopically overexpressed in all tissues examined. The A^y mRNA is identical to the wild-type agouti transcript for the entire coding region, but the 5' untranslated sequence of the agouti gene has been removed and replaced by novel sequence. Here we demonstrate that the novel 5' sequence in the A^y mRNA corresponds to 5' untranslated sequence of another gene that is normally tightly linked to agouti in mouse chromosome 2. This other gene (*Raly*) has the potential to encode a novel RNA-binding protein that is normally expressed in the preimplantation embryo, throughout development, and in all adult tissues examined. Importantly, the A^y mutation disrupts the expression of the *Raly* gene. The data suggest that the A^y mutation arose from a DNA structural alteration that affects the expression of both agouti and *Raly*. We propose that the dominant pleiotropic effects associated with A^y may result from the ectopic overexpression of the wild-type agouti gene product under the control of the *Raly* promoter, and that the recessive embryonic lethality is likely due to the lack of *Raly* gene expression in the early embryo.

INTRODUCTION

The agouti (*a*) locus in mouse chromosome 2 is recognized most widely for its role in the regulation of coat pigmentation (reviewed by Silvers, 1979). There are, however, several agouti locus mutations that also cause embryonic lethality in homozygotes, and these mutations have served as useful genetic reagents for the analysis of early mouse development (Papaioannou and Mardon, 1983; Lyon et al., 1985; Papaioannou and Gardner 1992). The most notable of these mutations is lethal yellow (A^y),

which dates back to the mouse fancy and derives its name from the fact that homozygotes die prenatally and heterozygotes have a completely yellow pelage. The A^y mutation causes a number of dominant pleiotropic effects in addition to the yellow pelage, including obesity (Dickerson and Gowen, 1947; Fenton and Chase, 1951; Carpenter and Mayer, 1958; Plocher and Powley, 1976; Friedman and Leibel, 1992), non-insulin dependent diabetes (Hellerstrom and Hellman, 1963), and increased tumor susceptibility (reviewed by Wolff et al., 1986; Wolff, 1987).

A^y was the first mammalian mutation to be found associated with embryonic lethality, which was revealed by a modified Mendelian ratio among offspring of phenotypically yellow parents (Cuenot, 1908; Castle and Little, 1910; Ibsen and Steigleder, 1917; Kirkham 1917, 1919). Early histological studies placed the time of death of A^y/A^y mice between 5.5 and 6.5 days *post coitum* (Robertson 1942; Eaton and Green, 1963), but subsequent experiments with preimplantation embryos in culture indicated that deleterious effects of the A^y mutation may be manifested in early cleavage (Pedersen, 1974; Pedersen and Spindle, 1976; Granholm and Johnson, 1978). Overall, these experiments suggested that the A^y mutation causes developmental defects over a range of time between early cleavage and implantation, but no tissue or stage-specific effect of the A^y mutation was identified.

Papaioannou and Gardner (1979) sought to determine whether A^y affects either the trophectoderm or the inner cell mass (ICM) of the blastocyst exclusively, or if A^y acts as a general cell lethal. To address these issues, they constructed chimeras by transferring the ICM of an A^y/A^y embryo into the blastocoel cavity of a normal embryo, and vice versa. The key results from these analyses were that homozygous lethal ICMs survived and differentiated in normal blastocysts up to at least 10.5 days of gestation, but that normal ICMs could not rescue the preimplantation lethality of A^y/A^y blastocysts. Since neither the genotype nor phenotype of the A^y/A^y embryos could be unequivocally ascertained in these experiments, the authors analyzed their data statistically and concluded that the A^y mutation primarily affects the trophectoderm and does not cause a generalized cell lethality.

More recently, however, Papaioannou (1988) further investigated the tissue specificity of the lethal yellow mutation by culturing ICM and trophoctoderm that had been separated from individual blastocysts derived from ($A^y/a^e \times A^y/a^e$) and control ($a^e/a^e \times A^y/a^e$) matings. Based on these experiments it was concluded that homozygosity of A^y exerts a detrimental effect in both the ICM and the trophoctoderm. This *in vitro* result is consistent with the *in vivo* experiments later reported by Barsh et al. (1990) in which aggregation chimeras prepared between A^y/A^y and normal blastocysts were analyzed for the presence of A^y with a molecular probe at day 9.5 of gestation or at birth. These experiments suggested that the lethal effects of A^y/A^y cannot be rescued in a chimeric environment.

The molecular nature of the gene(s) responsible for the embryonic lethality and the dominant pleiotropic effects of A^y is currently unknown. However, as part of our recently reported work involving the molecular characterization of the wild-type agouti gene, we made the observation that the agouti transcript expressed from the A^y allele is slightly larger than the wild-type transcript. The entire coding region of the agouti gene is intact in the larger-than-normal A^y transcript, but the first, noncoding exon of the agouti gene has been removed and replaced by novel sequence, partly through an aberrant splicing event (Bultman et al., 1992). Based on these observations, and the fact that the A^y transcript is ectopically overexpressed, we previously hypothesized that the A^y allele arose through a structural alteration that caused the disruption of both the agouti gene and a second gene, resulting in the replacement of the normal promoter and first exon of agouti with the promoter and 5' end of the second gene (Bultman et al., 1992). We further proposed that disruption of the agouti gene and its concomitant ectopic overexpression is associated with the yellow coat pigmentation and dominant pleiotropic effects of A^y , and that the recessive embryonic lethality is not related to the agouti gene *per se*, but is instead due to the disruption of the function of the second gene.

Here we present a more detailed molecular analysis of the A^y allele, including the cloning and characterization of that second gene, which we propose is directly responsible for the embryonic lethality of A^y . This gene maps very close to the agouti locus, and has a broad temporal and spatial

pattern of expression, which includes the blastocyst. Most importantly, the expression of this gene is disrupted in the *A^y* allele. Sequence analysis of this gene indicates that it is a member of a family of RNA-binding proteins that have been implicated in pre-mRNA processing and developmental regulation.

MATERIALS AND METHODS

Animals

All mice were maintained at the Oak Ridge National Laboratory. The *A^{vy}* mice represent a new spontaneous mutation from The Jackson Laboratory that we determined is homozygous viable, and appears to be genetically identical to the classical *A^{vy}* allele (E.J.M., S.J.B., Muriel T. Davisson, and R.P.W., manuscript in preparation).

RNA Analysis

Total cellular RNA from all tissues (except blastocysts) was extracted using the guanidine isothiocyanate procedure (Ausubel et al., 1988), enriched for poly(A)⁺ RNA using an oligo(dT)-cellulose column (Aviv and Leder, 1972), electrophoresed through formaldehyde gels, and blotted to GeneScreen (DuPont) using standard procedures (Ausubel et al., 1988). Radiolabeled hybridization probes were prepared with the random hexamer labeling technique (Feinberg and Vogelstein, 1984). Posthybridization filter washes were conducted at high stringency (0.2X SSC, 0.1% SDS, 68° C) unless otherwise noted. Reduced stringency washes were conducted at 0.2X SSC, 0.1% SDS, 50° C.

Total cellular RNA from blastocysts was prepared as described by Rothstein et al. (1992).

For RT-PCR analyses, eight µg of total RNA was reverse transcribed (Kawasaki, 1990), ethanol precipitated, resuspended in 20 µl of H₂O, and PCR analysis was performed with 2 µl of the sample as previously described (Pieretti et al., 1991).

Isolation of cDNA Clones

Total RNA was prepared from *Ay/a^e* adult kidney, and double-stranded cDNA was prepared by standard procedures (Sambrook et al., 1989). After the addition of *Eco*RI linkers, the cDNA was ligated into the λ gt10 vector (Stratagene), packaged in vitro, and screened with a ³²P-labeled wild-type agouti cDNA probe (Fig. 2 in Bultman et al. 1992) using standard procedures (Sambrook et al., 1989). Positive clones were purified and subcloned into pBluescript (Stratagene) for further analysis. *Raly* cDNA clones were isolated from a C57BL/6 day 8.5 whole embryo library (Fahrner et al., 1987) by screening with ³²P-labeled probe A (Fig. 1 and 2), and positive clones were purified and subcloned into pBluescript (Stratagene).

Isolation of Genomic Clones

Genomic liver DNA from the strain FVB/N was partially digested with *Sau*3A and size fractionated on a 10-40% sucrose gradient (Sambrook et al., 1989). Fractions containing 18 to 23-kb fragments were ligated into the lambda vector, EMBL3 (Stratagene), packaged in vitro, and screened with ³²P-labeled probe A (Fig. 1 and 2) using standard procedures (Sambrook et al., 1989). Positive clones were purified and subcloned into pBluescript (Stratagene).

DNA Sequencing

Genomic and cDNA clones were sequenced by the Sanger dideoxynucleotide method (Sanger et al., 1977) using T7 DNA polymerase (US Biochemical) (Tabor and Richardson, 1987). For direct sequencing of RT-PCR amplified cDNA from the *M. musculus*/*M. spretus* hybrids, the Sanger dideoxynucleotide method was used along with a modified PCR-template protocol as described (Winship, 1989). Analysis of DNA sequence was performed using the University of Wisconsin Genetics Computing Group sequence analysis programs (Devereux et al., 1984).

RESULTS

The A^y Transcript is a Chimera Composed of Agouti and a Second Gene

We previously demonstrated that the agouti gene consists of four exons and is normally expressed as a 0.8-kb transcript in neonatal skin, whereas the A^y allele is expressed as an approximately 1.1-kb transcript in neonatal skin and in all adult tissues examined to date (Bultman et al., 1992). An adult kidney cDNA library was prepared from a lethal yellow heterozygote (A^y/a^e) and was screened with the wild-type agouti cDNA clone (Bultman et al., 1992). Sequence analysis of a partial A^y cDNA clone revealed that the entire coding region and 3' untranslated sequence of the agouti gene is identical in the larger-than-normal A^y transcript, but the first, noncoding exon of the agouti gene has been replaced by novel sequence (Fig. 1). This segment of novel 5' sequence in the A^y cDNA was subsequently used as a probe (probe A in Fig. 1) to obtain wild-type genomic and cDNA clones. Sequence analyses of these clones revealed that the novel sequence in the A^y transcript corresponds to a 5' noncoding exon of a previously uncharacterized gene (Fig. 1), which we have named *Raly* because it has the potential to encode an hnRNP protein that is associated with lethal yellow (see below).

Characterization of *Raly* cDNA Clones

To characterize the structure of the wild-type *Raly* gene, we screened approximately 1.2×10^6 plaque forming units of a day 8.5 total embryo cDNA library (Fahrner et al., 1987) with probe A (Fig. 1), resulting in the identification of six clones, three of which were characterized further. Each of these three clones contains the entire coding region and appears to be full-length or nearly full-length. The total size of the cDNA clone presented in Figure 2 is 1517 bp, and with the addition of a poly(A) tract is similar in size to the 1.6-kb mRNA observed on Northern blots (see below). The cDNA contains an open reading frame (ORF) extending from nucleotide 262 through 1152, beginning with an ATG codon flanked by sequence that conforms well to the consensus for translation initiation (Kozak, 1987).

The translation product deduced from the ORF is 296 amino acids in length, with an estimated molecular weight of 31.2 kD and an isoelectric point of 10.17 (Fig. 2). Searches of the NBRF and SWISSPROT data bases using the algorithm FASTA (Pearson and Lipman, 1988) identified both human (Swanson et al., 1987; Burd et al., 1989) and *Xenopus* (Preugschat and Wold, 1988) heterogeneous nuclear ribonucleoprotein particle C (hnRNP C) proteins as having striking sequence similarity to *Raly*.

The hnRNP proteins, a family of proteins involved in pre-mRNA packaging and processing, belong to a larger family of RNA-binding proteins that are all characterized by a generally conserved domain consisting of approximately 90 amino acids (the RNA-binding domain, RBD) that occurs at least once in these proteins (Bandziulis et al., 1989). The most highly conserved segments in the RBD are an octapeptide (referred to as RNP 1) and a hexapeptide (RNP 2); these residues are indicated in the *Raly* sequence in Fig. 2. A comparison of the putative 99 amino acid RBD of *Raly* to the 94 amino acid RBD of human hnRNP C indicates that, over a 93 residue region, these two proteins are 77% identical and 89% similar, with no gaps introduced into the alignment (Fig. 3). However, when the remaining C-terminal portions of these proteins are compared, an optimal alignment results in 37% identity and 58% similarity after introducing 7 gaps (data not shown). Therefore, although the RBDs of *Raly* and hnRNP C share striking sequence similarity and likely have a related function within the cell, the remainder of *Raly* is sufficiently unique to indicate that it is not simply the mouse homologue of hnRNP C. Rather, it appears that *Raly* represents a novel member of the hnRNP family of proteins.

The *Raly* Transcription Unit is Alternately Spliced at the 5' End

Based on the analysis of *Raly* cDNA clones, we discovered that there are two forms of the mature *Raly* transcript present in the day 8.5 total embryo. These two forms differ only by the presence or absence of an 83-bp segment in the 5' untranslated region (indicated by capital letters in Fig. 2). The form depicted in Fig. 2, containing the 83-bp segment, has in-frame termination codons upstream of the translation initiation codon. The alternate form of the *Raly* transcript lacks the 83-bp segment (the form

illustrated in Fig. 1), which results in the removal of the upstream termination codons and the extension of the ORF from the ATG to the 5' end of the cDNA clone. Whether this extension of the ORF is functionally significant is presently unclear, but it is unlikely since there are no ATG codons present within this region.

A reverse transcriptase-polymerase chain reaction (RT-PCR) assay was used to determine if both forms of the *Raly* mRNA are present in a variety of tissues, in addition to the day 8.5 total embryo. Oligonucleotide primers were prepared which correspond to base pairs 108-127 (5' to the 83-bp segment) and to the reverse complement of base pairs 457-438 (3' to the 83-bp segment) in the *Raly* cDNA (Fig. 2). If both forms of the *Raly* mRNA are present in a given tissue, this assay is expected to result in RT-PCR products of 350 bp for the transcript containing the 83-bp segment, and 267 bp for the alternate transcript that lacks this sequence. To date, the assay has been used to examine the blastocyst, neonate skin, and adult brain. The two expected size fragments were produced in each tissue, and both fragments were observed at approximately equal intensities (data not shown). The results of this assay indicate that both forms of the *Raly* transcript are expressed at approximately equal levels in each of these tissues.

***Raly* and *Agouti* are Tightly Linked in Chromosome 2**

The observation that the *A^y* allele expresses a chimeric transcript containing both *Raly* and *agouti* sequences, and that the *A^y* allele has no detectable rearrangement at the cytogenetic level (N.L.A. Cacheiro, personal communication), suggested that these two genes may normally be tightly linked on chromosome 2. To address this issue, genomic probes from both the *Raly* and *agouti* loci were hybridized to *TaqI* digested DNA from 80 *Mus musculus*/*M. spretus* F₁ animals (Johnson et al., 1989). The *Raly* probe identified fragments of 2.2, 1.8, and 0.9 kb in *M. musculus*, and fragments of 3.5, 1.8, and 0.9 kb in *M. spretus*. The *agouti* probe identified fragments of 3.5, 1.8, 1.7, and 0.5 kb in *M. musculus*, and a single fragment of 1.9 kb in *M. spretus*. The unique sized restriction fragments allowed us to differentiate the *M. musculus* and *M. spretus* alleles. This analysis revealed no

recombinants between the *Raly* and agouti sequences, indicating that these genes are tightly linked on chromosome 2 (95% confidence limits of 0-4.1 cM, data not shown). Moreover, preliminary pulsed-field data suggest that *Raly* and agouti may be no farther than 300 kb apart (data not shown).

In an attempt to begin to analyze the structure of the DNA associated with the *A^y* allele, we probed genomic blots with probes specific to the 5' end of agouti and to the 3' end of *Raly* (Fig. 4). For the 5' end of agouti, we were able to differentiate the *A^y* allele from the balancing nonagouti allele by exploiting a sequence polymorphism. As shown in Figure 4A, we clearly demonstrated that the 5' end of agouti in the *A^y* allele is not deleted or rearranged. Likewise, to analyze the 3' end of *Raly*, we utilized a DNA polymorphism to differentiate the *M. musculus* *A^y* allele from the *M. spretus* wild-type agouti allele in an F₁ hybrid. In this case, we demonstrated that the 3' end of the *Raly* gene is deleted in the *A^y* allele (Fig. 4B). This information, coupled with the fact that the promoter and first exon of *Raly* are not deleted or rearranged in the *A^y* allele (E.J. Michaud, S.J. Bultman, L.J. Stubbs, and R.P. Woychik, unpublished), leads us to conclude that a deletion is associated with the *A^y* allele. The deletion removes at least a portion of the coding sequences at the 3' end of *Raly* and occurs within the region between the *a* gene and the 5' end of *Raly*.

***Raly* is Expressed in a Widespread Temporal and Spatial Manner**

Given that the 5' end of *Raly* has been spliced to the coding region of agouti in the *A^y* transcript, and that *A^y* is ectopically expressed (presumably from the *Raly* promoter) in a widespread manner (Bultman et al., 1992), it seemed likely that wild-type *Raly* would also be expressed in the same broad temporal and spatial pattern as *A^y*. A cDNA probe was used to demonstrate that *Raly* is, in fact, expressed as a 1.6-kb transcript in the developing embryo and fetus, in neonatal skin, and in a variety of adult tissues (Fig. 5A, B). Both forms of the *Raly* transcript apparently comigrate on Northern blots because they differ in size by only 83 bases. When the Northern blot filters were washed at high stringency, only the 1.6-kb transcript was present in each tissue. Interestingly, when the same filter was initially washed at a reduced stringency, hybridizing fragments of

approximately 1.0, 2.3, and 2.7 kb were also present in neonatal skin (Fig. 5A). Under the same reduced stringency conditions, these additional transcripts are also apparent in advanced stage whole fetuses, possibly due to the increased percentage of skin in the total fetal RNA (Fig. 5B). Whether these additional transcripts represent *Raly* family members that are uniquely expressed in the skin, or are simply unrelated skin-specific transcripts that cross-hybridize to *Raly* remains to be determined. Importantly, *Raly* is also expressed in the blastocyst (Fig. 5C), the stage in which the recessive lethality of the A^y mutation is manifested.

The A^y Mutation Disrupts the Expression of the *Raly* Gene

The molecular structure of the A^y genomic DNA indicated that the A^y mutation contains a deletion that removes at least a portion of the *Raly*-coding sequences near the 3' end of the gene (Fig. 4B). may disrupt wild-type *Raly* expression. To provide additional supporting evidence that *Raly* is not expressed from the A^y allele, we tested for the expression of one type of *Raly* transcript in the A^y allele using a method that exploited a sequence polymorphism between *Mus musculus* and *Mus spretus*. Sequence analysis of a portion of the 5' untranslated region of the wild-type *Raly* gene from both species revealed that *M. musculus* has a "g" at nucleotide 175 (complement of sequence in Fig. 2), which is replaced by a "t" in *M. spretus* (first two panels in Fig. 6). For this purpose, *M. musculus* females (A^y/a) were mated to *M. spretus* males (A/A) to produce the F_1 hybrids A^y/A (distinguished from a/A littermates by their solid yellow pelage). Total RNA from both the A^y/A and a/A F_1 hybrids was prepared from day 10 neonate skin, reverse transcribed, and oligonucleotide primers corresponding to base pairs 108-127 and the reverse complement of base pairs 457-438 (Fig. 2) were used to PCR amplify the 5' portion of the *Raly* mRNA containing the sequence polymorphism at position 175. The PCR-amplified fragments were then gel isolated and directly sequenced. The A^y/A F_1 hybrid has a "t" at position 175, indicating that *Raly* is expressed only from the *M. spretus* allele (A), and not from the *M. musculus* allele (A^y) (Fig. 6, third panel). The a/A F_1 hybrid, which should express the wild-type *Raly* of both species, served as a control for this experiment and, as expected, has both a "g" and a "t" comigrating at

position 175. The analysis of the α/A control indicated that *Raly* is capable of being expressed from both the *M. musculus* (α) and *M. spretus* (A) alleles (Fig. 6, last panel), and that our assay can detect the simultaneous expression of these two forms of the gene.

The A^y Transcription Unit Forms Unique Transcripts that are Alternately Spliced at the 5' End

Because *Raly* is alternately processed at its 5' end, we also tested for alternative processing in the A^y transcription unit with an RT-PCR assay similar to the one used above for *Raly*. In this case, however, we used an oligonucleotide primer from the first exon of *Raly* coupled with a primer from the third exon of agouti (Fig. 7). Results indicated that three different sized PCR fragments of 284, 395, and 441 bp were produced (Fig. 7), all of which were of the same approximate abundance in adult spleen and brain (the tissues examined to date in this manner, data not shown). These three PCR fragments were subcloned and sequenced, revealing that each differs at its 5' end, and that the alternate processing of the A^y transcription unit occurs in a unique manner compared to *Raly* (Fig. 7). The 5' end of the first form contains sequence derived from a region we are referring to as *Raly* exon I, the second form contains *Raly* exon I and an additional 111-base segment, and the third form contains *Raly* exon I, the 111-base segment, and an additional 46-base segment. Each of these different 5' ends is spliced to sequence derived from the second, third, and fourth agouti exons, which contain the entire coding region of the agouti gene (Fig. 7).

A feature in common to all the A^y and *Raly* transcripts we have identified is the presence of sequence derived from *Raly* exon I. This *Raly* sequence was identified as an exon by comparing the wild-type cDNA sequence to genomic clones, and is considered to be the first *Raly* exon because it is present at the 5' end of all cDNA clones characterized thus far. Interestingly, the 83-base segment that is alternately processed in the *Raly* transcription unit has not been found in any of the A^y transcripts, and the 111-base and 46-base segments that are alternately processed in the A^y transcription unit have not been found in either form of the mature *Raly* transcript. It has not yet been determined whether the 111- and 46-base

segments in the A^y transcription unit are, in fact, *Raly* exons, or whether these sequences may arise from cryptic splicing events in a unique pre-mRNA transcript resulting from the molecular nature of the A^y mutation.

The largest of the A^y transcripts, containing *Raly* exon I, and the 111- and 46-base segments, has four in-frame termination codons upstream of the agouti translation initiation codon. The two smaller A^y transcripts have an in-frame ORF that continues from the agouti translation initiation codon to the 5' end of the transcript, however, no ATG codons are present in this upstream sequence.

The Embryonic Lethality of A^y/A^y Mice May be the Result of the Disruption of *Raly*, and not the Ectopic Overexpression of the Agouti Protein

Viable yellow (A^{vy}) is another dominant agouti locus mutation that confers the same suite of pleiotropic effects as A^y (reviewed by Wolff et al., 1986; Wolff, 1987), however, as the name suggests, A^{vy}/A^{vy} mice are viable. As a first step in analyzing the molecular nature of the A^{vy} mutation, we tested for the expression of the agouti mRNA in several adult tissues of an A^{vy} heterozygote. We found that A^{vy} mice, like A^y mice, ectopically overexpress agouti. However, unlike A^y , the transcript size of A^{vy} is the same as that from the wild-type allele (Fig. 8).

DISCUSSION

The lethal yellow (A^y) mutation at the agouti locus in mouse chromosome 2 confers a number of dominant pleiotropic effects in heterozygotes and results in the preimplantation lethality of homozygous embryos. The dominant pleiotropic effects and recessive embryonic lethality associated with A^y have been intensively studied for several decades, but the molecular nature of the underlying defects has remained elusive. As part of our recent molecular characterization of the agouti gene, we identified a size-altered agouti mRNA associated with the A^y allele (Bultman et al., 1992). This larger-than-normal A^y transcript contains the entire coding region of the agouti gene, but the untranslated first exon has been replaced by novel

sequence (Bultman et al., 1992). Here we demonstrate that at least a portion of this novel sequence in the *A^y* transcript corresponds to the 5' region of *Raly*, another gene that is tightly linked to agouti. Moreover, we have determined that *Raly* is normally expressed in a broad temporal and spatial manner, which includes the blastocyst, and that it normally produces two alternatively spliced mRNAs that are approximately 1.6 kb in length. *Raly* has the potential to encode a protein that has striking amino acid sequence similarity to the RNA-binding domain of vertebrate hnRNP C proteins (Swanson et al., 1987; Preugschat and Wold, 1988; Burd et al., 1989), which avidly bind to RNA and are probably involved in the processing of hnRNA within the nucleus of the cell. Most importantly, we have determined that *Raly* is not expressed from the *A^y* allele, and we propose that this lack of expression is responsible for the recessive embryonic lethality of *A^y*.

The precise nature of the DNA structural alteration that generated the *A^y* allele is not yet clear. We do know that the rearrangement involves two closely linked genes, *Raly* and agouti, and that a deletion has removed at least a portion of the 3' end of *Raly* but has not altered the genomic structure of the *a* locus. Our current working hypothesis for the *A^y* allele is that a deletion has removed the coding exons of the *Raly* gene but has not affected the promoter and 5'-noncoding sequences. In this scenario, transcription would initiate normally at the *Raly* promoter, but since the 3' end of the *Raly* gene is deleted, transcription would proceed into the intergenic region and through the downstream agouti gene. The resulting large primary transcript would be spliced in such a manner that the intergenic DNA and the first exon of the agouti gene would be treated as an "intron", because the next available splice acceptor for the splice donor at the 5' end of *Raly* would be provided by the 3' end of the first intron of the agouti gene. The processed transcript would be expected to contain the 5' noncoding exon of *Raly* and the second, third, and fourth exons of agouti, which is precisely the nature of one of the *A^y*-specific transcripts described in this paper. Physical mapping analyses of the *Raly* and agouti loci in the *A^y* allele are currently in progress and should provide direct evidence to support or refute this working model.

Previously we reported that the wild-type agouti 0.8-kb mRNA is expressed specifically within the skin of the neonate, whereas the A^y allele ectopically overexpresses a 1.1-kb agouti mRNA in a variety of tissues of the adult animal (Bultman et al., 1992). We further demonstrated that the increased size of the A^y transcript is due to a rearrangement upstream of the agouti translation initiation codon, and that A^y retains the potential to encode the wild-type agouti protein. Based on these observations, and the likelihood that the agouti protein is a secreted ligand, we proposed that the A^y allele ectopically overexpresses the wild-type agouti protein, leading to the dominant pleiotropic effects associated with A^y (Bultman et al., 1992). However, after sequencing the heterogeneous 5' ends of several A^y cDNA clones (Fig. 7), we discovered that two of the three different forms of the A^y transcript have in-frame ORFs that continue from the agouti translation initiation codon to the 5' end of the cloned portion of the transcript. Therefore, we must entertain the possibility that a fusion protein is generated from at least some forms of the A^y transcript. Since no ATG codons are present in the upstream sequence that we have characterized thus far, we continue to favor the hypothesis that the A^y transcription unit produces only a wild-type agouti protein. Antibodies to the agouti protein and further analysis of the 5' end of the A^y transcript to search for additional ATG codons should facilitate the resolution of this uncertainty.

The characterization of the A^{vy} allele lends further support to the idea that the pleiotropic effects associated with A^y result from the ectopic overexpression of the wild-type agouti protein. The A^{vy} allele causes the same dominant pleiotropic effects as A^y , and also ectopically overexpresses the agouti mRNA. Unlike A^y , however, the A^{vy} transcript is the same size as the wild-type agouti transcript, suggesting that A^{vy} expresses the wild-type agouti protein from this normal sized transcript. Experiments are currently under way to characterize the molecular nature of the A^{vy} mutation.

Whereas the ectopic overexpression of the agouti protein is likely to be responsible for the dominant pleiotropic effects associated with A^y and A^{vy} , this is unlikely to be the cause of the recessive embryonic lethality associated with A^y . Not unexpectedly, the A^y allele expresses the agouti mRNA in the same broad temporal and spatial pattern as observed for the

wild-type *Raly* gene product because the agouti gene is under the control of the *Raly* promoter in the A^y allele. Based on our finding that *Raly* is expressed in the blastocyst, it is reasonable to deduce that the agouti mRNA is expressed from the A^y allele in the blastocyst of A^y heterozygotes. Following this line of reasoning, it is unlikely that the ectopic overexpression of the agouti gene product is responsible for the embryonic lethality of A^y . If this were the case, then embryos heterozygous for A^y would also be expected to show some early developmental phenotype, and they do not. A dosage effect argument could be invoked in which two agouti expressing A^y alleles, not just one, are necessary to induce the recessive lethality of A^y . However, since *Raly* is associated with the A^y mutation, and is inactivated in the A^y allele, we favor the alternative possibility that A^y homozygotes fail to develop beyond the blastocyst stage due to their inability to produce the *Raly* gene product. This hypothesis is consistent with the recessive nature of the embryonic lethality. Because we have not yet fully characterized the structural alteration associated with the A^y allele, we cannot rule out the possibility that a gene(s) other than *Raly* has also been affected by the A^y mutation and contributes to the lethality phenotype. The definitive experiment to determine if the loss of *Raly* expression alone is directly responsible for the prenatal lethality of A^y/A^y mice is to rescue the lethal phenotype with the wild-type *Raly* gene in transgenic mice, and this work is currently in progress.

The observation that the *Raly* gene encodes a protein whose amino terminus has 77% sequence identity to the RNA-binding domain of the hnRNP C proteins provides a framework with which to speculate about its biochemical function in the early embryo. There are at least 20 abundant hnRNP proteins present in the vertebrate cell nucleus that are physically associated with nascent hnRNA transcripts to form hnRNP complexes (Dreyfuss, 1986; Pinol-Roma et al., 1988; Bennett et al., 1992). The posttranscriptional processing of hnRNA into mRNA is believed to occur within these hnRNP complexes. The hnRNP proteins have, in fact, been shown to bind to pre-mRNAs in a sequence-dependent manner. Notably, the hnRNP C proteins bind to the 3' end of introns (Swanson and Dreyfuss, 1988) and to sequences downstream of polyadenylation cleavage sites

(Wilusz et al., 1988), suggesting that they may have a role in pre-mRNA processing. However, relatively little is known about the precise physiological functions of hnRNP proteins.

The recent cloning and sequencing of cDNAs for a number of hnRNP proteins revealed that many of them contain one or two putative RNA-binding domains (RBD) of approximately 90 amino acids. In addition to the hnRNP proteins, a whole host of nuclear and cytoplasmic RNA-binding proteins contain one or more of these 90 residue RBDs that are evolutionarily conserved from yeast to man. Comparisons of these various proteins revealed that the most highly conserved portions of the RBD are an octapeptide (RNP 1) and a hexapeptide (RNP 2) (reviewed by Dreyfuss et al. 1988, Bandziulis et al., 1989).

The ribonucleoprotein consensus sequences (RNP 1 and RNP 2) have also been found in the deduced proteins of several different developmental regulatory genes of *Drosophila*. For example, the genes *sex-lethal* (Bell et al., 1988) and *transformer-2* (Amrein et al., 1988) regulate somatic sexual development, and the genes *elav* (embryonic lethal, abnormal visual system) (Robinow et al., 1988) and *couch potato* (Bellen et al., 1992) are involved in the normal development and maintenance of the nervous system. Based on the presence of RNP consensus motifs in each of these gene products, and the observation that these genes most likely exert their regulatory effects at the posttranscriptional level, it has been suggested that these genes regulate development by controlling the RNA-processing of other genes that directly participate in establishing the somatic sexual and neuronal developmental pathways (Bandziulis et al., 1989).

The striking sequence similarity between the *Raly* and hnRNP C proteins suggests that *Raly* may be functioning in the processing of mRNA in the preimplantation embryo. Based on the observation that the embryonic lethality of *A^y* may occur as late as implantation, we propose that *Raly* is not involved in the generalized RNA processing of most or all of the genes produced in the early embryo. If this were the case, embryonic lethality would probably occur much earlier, since *de novo* transcription begins in the two-cell embryo of the mouse, and most maternal mRNA in the embryo is degraded at this time (reviewed by Schultz, 1986). Therefore, it is

conceivable that *Raly* may have a specific regulatory function later in development that somehow relates to the maturation of the blastocyst. on the other hand, it may turn out that *Raly* is essential for generalized processing of transcripts in the early embryo and that enough *Raly* protein from maternal mRNA persists in the early embryo to allow development to the blastocyst stage. Experiments are currently under way to address this issue.

REFERENCES

- Amrein, H., Gorman, M., and Nothiger, R.. (1988). The sex-determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell* **55**: 1025-1035.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1988). *Current protocols in molecular biology*. John Wiley & Sons/Greene, New York.
- Aviv, H., and Leder, P. (1972). Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci USA* **69**: 1408-1412.
- Bandziulis, R. J., Swanson, M. S., and Dreyfuss, G. (1989). RNA-binding proteins as developmental regulators. *Genes & Dev.* **3**: 431-437.
- Barsh, G. S., Lovett, M., and Epstein, C. J. 1990. Effects of the lethal yellow (*A^y*) mutation in mouse aggregation chimeras. *Development* **109**: 683-690.
- Bell, L. R., Maine, E. M., Schedl, P., and Cline, T. C. (1988). *Sex-lethal*, a *Drosophila* sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell* **55**: 1037-1046.
- Bellen, H. J., Kooyer, S., D'Evelyn, D., and Pearlman, J. (1992). The *Drosophila* Couch potato protein is expressed in nuclei of peripheral neuronal precursors and shows homology to RNA-binding proteins. *Genes & Dev.* **6**: 2125-2136.
- Bennett, M., Pinol-Roma, S., Staknis, D., Dreyfuss, G., and Reed, R. (1992). Differential binding of heterogeneous nuclear ribonucleoproteins to mRNA precursors prior to spliceosome assembly in vitro. *Mol. Cell. Biol.* **12**: 3165-3175.

- Bultman, S. J., Michaud, E. J., and Woychik, R. P.. (1992). Molecular characterization of the mouse agouti locus. *Cell* **71**: 1195-1204.
- Burd, C. G., Swanson, M. S., Gorlach, M., and Dreyfuss G. (1989). Primary structures of the heterogeneous nuclear ribonucleoprotein A2, B1, and C2 proteins: A diversity of RNA binding proteins is generated by small peptide inserts. *Proc. Natl. Acad. Sci. USA* **86**: 9788-9792.
- Carpenter, K. J., and Mayer, J. (1958). Physiologic observations on yellow obesity in the mouse. *Am. J. Physiol.* **193**: 499-504.
- Castle, W. E., and Little, C. C. (1910). On a modified Mendelian ratio among yellow mice. *Science* **32**: 868-870.
- Cuenot, L. (1908). Sur quelques anomalies apparentes des proportions Mendeliennes. *Arch. Zool. Exp. Gen.* **9**: 7-15.
- Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**: 387-395.
- Dickerson, G. E., and Gowen, J.W. (1947). Hereditary obesity and efficient food utilization in mice. *Science* **105**: 496-498.
- Dreyfuss, G. (1986). Structure and function of nuclear and cytoplasmic ribonucleoprotein particles *Annu. Rev. Cell Biol.* **2**: 459-498.
- Dreyfuss, G., Swanson, M. S., and Pinol-Roma, S. (1988). Heterogeneous nuclear ribonucleoprotein particles and the pathway of mRNA formation. *Trends Biochem. Sci.* **13**: 86-91.
- Eaton, G. J., and Green, M.M. (1963). Giant cell differentiation and lethality of homozygous yellow mouse embryos. *Genetica* **34**: 155-161.

- Fahrner, K., Hogan, B. L. M., and Flavell, R. A. (1987). Transcription of *H-2* and *Qa* genes in embryonic and adult mice. *EMBO J.* **6**: 1265-1271.
- Feinberg, A. P., and Vogelstein, B. (1984). Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**: 266-267.
- Fenton, P. F., and Chase, H. B. (1951). Effect of diet on obesity of yellow mice in inbred lines. *Proc. Soc. Exper. Biol.* **77**: 420-422.
- Friedman, J. M., and Leibel, R. L. (1992). Tackling a weighty problem. *Cell* **69**: 217-220.
- Granholm, N.H., and Johnson, P.M. (1978). Enhanced identification of lethal yellow (*A^y/A^y*) mouse embryos by means of delayed development of four-cell stages. *J. Exp. Zool.* **205**: 327-333.
- Hellerstrom, C., and Hellman, B. (1963). The islets of Langerhans in yellow obese mice. *Metabolism* **12**: 527-536.
- Ibsen, H. L., and Steigleder, E. (1917). Evidence for the death *in utero* of the homozygous yellow mouse. *Am. Nat.* **51**: 740-752.
- Johnson, D. K., Hand, R. E., and Rinchik, E. M. (1989). Molecular mapping within the mouse albino-deletion complex. *Proc. Natl. Acad. Sci USA* **86**: 8862-8866.
- Kawasaki, E. S. (1990). Amplification of RNA. In *PCR protocols: A guide to methods and applications* (eds. M.A. Innis, D.H. Gelfand, J.J. Sninsky, and T.J. White), pp. 21-27. Academic Press, San Diego, CA.
- Kirkham, W. B. (1917). Embryology of the yellow mouse. *Anat. Rec.* **11**: 480-481.

- . (1919). The fate of homozygous yellow mice. *J. Exp. Zool.* **28**: 125-135.
- Kozak, M. (1987). An analysis of 5'-noncoding sequence from 699 vertebrate messenger RNAs. *Nucl. Acids Res.* **15**: 8125-8148.
- Lyon, M. F., Fisher, G., and Glenister, P. H. (1985). A recessive allele of the mouse agouti locus showing lethality with yellow, A^y . *Genet. Res.* **46**: 95-99.
- Papaiouannou, V. E. (1988). Investigation of the tissue specificity of the lethal yellow (A^y) gene in mouse embryos. *Devl. Genet.* **9**: 155-165.
- Papaiouannou, V., and Gardner, R. L. (1979). Investigations of the lethal yellow A^y/A^y embryo using mouse chimaeras. *J. Embryol. Exp. Morph.* **52**: 153-163.
- . (1992). Effects of diapause on lethal yellow (A^y/A^y) mouse embryos. *J. Exp. Zool.* **263**: 309-315.
- Papaiouannou, V. E., and Mardon, H. (1983). Lethal nonagouti (α^x): Description of a second embryonic lethal at the agouti locus. *Dev. Genetics* **4**: 21-29.
- Pearson, W. R., and Lipman, D. J. (1988). Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. USA* **85**: 2444-2448.
- Pedersen, R. A. (1974). Development of lethal yellow (A^y/A^y) mouse embryos *in vitro*. *J. Exp. Zool.* **188**: 307-319.
- Pedersen, R. A., and Spindle, A.I. (1976). Genetic effects on mammalian development during and after implantation. In *Embryogenesis in mammals, CIBA Foundation Symp. 40* (eds. K. Elliot and M. O'Connor), pp. 133-154. Elsevier, New York.

- Pieretti, M., F., Zhang, F., Fu, Y. -H., Warren, S. T., Oostra, B. A., Caskey, C. T., and Nelson, D. L. (1991). Absence of expression of the FMR-1 gene in fragile X syndrome. *Cell* **66**: 817-822.
- Pinol-Roma, S., Choi, Y. D., Matunis, M. J., and Dreyfuss, G. (1988). Immunopurification of heterogeneous nuclear ribonucleoprotein particles reveals an assortment of RNA-binding proteins. *Genes & Dev.* **2**: 215-227.
- Plocher, T.A., and Powley, T.L. (1976). Effect of hypophysectomy on weight gain and body composition in the genetically obese yellow (*A^y/a*) mouse. *Metabolism* **25**: 593-602.
- Preugschat, F., and Wold, B. (1988). Isolation and characterization of a *Xenopus laevis* C protein cDNA: Structure and expression of a heterogeneous nuclear ribonucleoprotein core protein. *Proc. Natl. Acad. Sci. USA* **85**: 9669-9673.
- Robertson, G. G. (1942). An analysis of the development of homozygous yellow mouse embryos. *J. Exp. Zool.* **89**: 197- 231.
- Robinow, S., Campos, A. R., Yao, K.- M., and White, K.. (1988). The *elav* gene product of *Drosophila*, required in neurons, has three RNP consensus motifs. *Science* **242**:1570-1572.
- Rothstein, J. L., Johnson, D., DeLoia, J. A., Skowronski, J., Solter, D., and Knowles, B. (1992). Gene expression during preimplantation mouse development. *Genes & Dev.* **6**: 1190-1201.
- Russell, L. B., McDaniel, M. N. C., and Woodiel, F. N. (1963). Crossing-over within the *a* "locus" of the mouse. *Genetics* **48**: 907a.

- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: A laboratory manual, second edition*. Cold Spring Harbor Press, Cold Spring Harbor, New York.
- Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.
- Schultz, G. A. (1986). Utilization of genetic information in the preimplantation mouse embryo. In *Experimental approaches to mammalian embryonic development* (eds. J. Rossant and R. A. Pedersen), pp. 239-265. Cambridge University Press, New York.
- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6-44. Springer-Verlag, New York.
- Siracusa, L. D., Russell, L. B., Eicher, E. M., Corrow, D. J., Copeland, N. G., and Jenkins, N. A. (1987). Genetic organization of the *agouti* region of the mouse. *Genetics* **117**: 93-100.
- Swanson, M. S., and Dreyfuss, G. (1988). RNA binding specificity of hnRNP proteins: a subset bind to the 3' end of introns. *EMBO J.* **7**: 3519-3529.
- Swanson, M. S., Nakagawa, T. Y., LeVan, K., and Dreyfuss, G. (1987). Primary structure of human nuclear ribonucleoprotein particle C proteins: Conservation of sequence and domain structures in heterogeneous nuclear RNA, mRNA, and pre-rRNA-binding proteins. *Mol. Cell. Biol.* **7**: 1731-1739.
- Tabor, S., and Richardson, C. C.. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84**: 4767-4771.

- Wilusz, J., Feig, D. I., and Shenk, T. (1988). The C proteins of heterogeneous nuclear ribonucleoprotein complexes interact with RNA sequences downstream of polyadenylation cleavage sites. *Mol. Cell. Biol.* **8**: 4477-4483.
- Winship, P. R. (1989). An improved method for directly sequencing PCR amplified material using dimethyl sulphoxide. *Nucleic Acids Res.* **17**: 1266.
- Wittekind, M., Gorlach, M., Friedrichs, M., Dreyfuss, G., and Mueller, L. (1992). ^1H , ^{13}C , and ^{15}N NMR assignments and global folding pattern of the RNA-binding domain of the human hnRNP C proteins. *Biochemistry* **31**: 6254-6265.
- Wolff, G. L. (1987). Body weight and cancer. *Am. J. Clin. Nutr.* **45**: 168-180.
- Wolff, G. L., Roberts, D. W., and Galbraith, D. B. (1986). Prenatal determination of obesity, tumor susceptibility, and coat color pattern in viable yellow (A^{vy}/a) mice: The yellow mouse syndrome. *J. Hered.* **77**: 151-158.

APPENDIX

FIGURES

Figure 1. The *A^y* allele encodes a *Raly/agouti* fusion transcript.

The wild-type agouti transcript is composed of sequence derived from four exons (numbered 1-4, with their respective sizes indicated below). The *A^y* transcript is identical to the wild-type agouti transcript for the sequence derived from the second, third and fourth exons (coding region), but the noncoding first exon of agouti has been removed and replaced by the noncoding first exon of *Raly* (represented by the shaded region, 169 bases in length). Probe A corresponds to sequence that was initially identified as being unique to the 5' end of the *A^y* transcript, and subsequently used to isolate wild-type *Raly* cDNA and genomic clones. The vertical line indicates that the replacement of the first agouti exon with the first *Raly* exon in the *A^y* transcript occurs precisely at an exon-intron splice junction in both wild-type genes. The proposed sites for translation initiation in all three transcripts are indicated by AUGs.

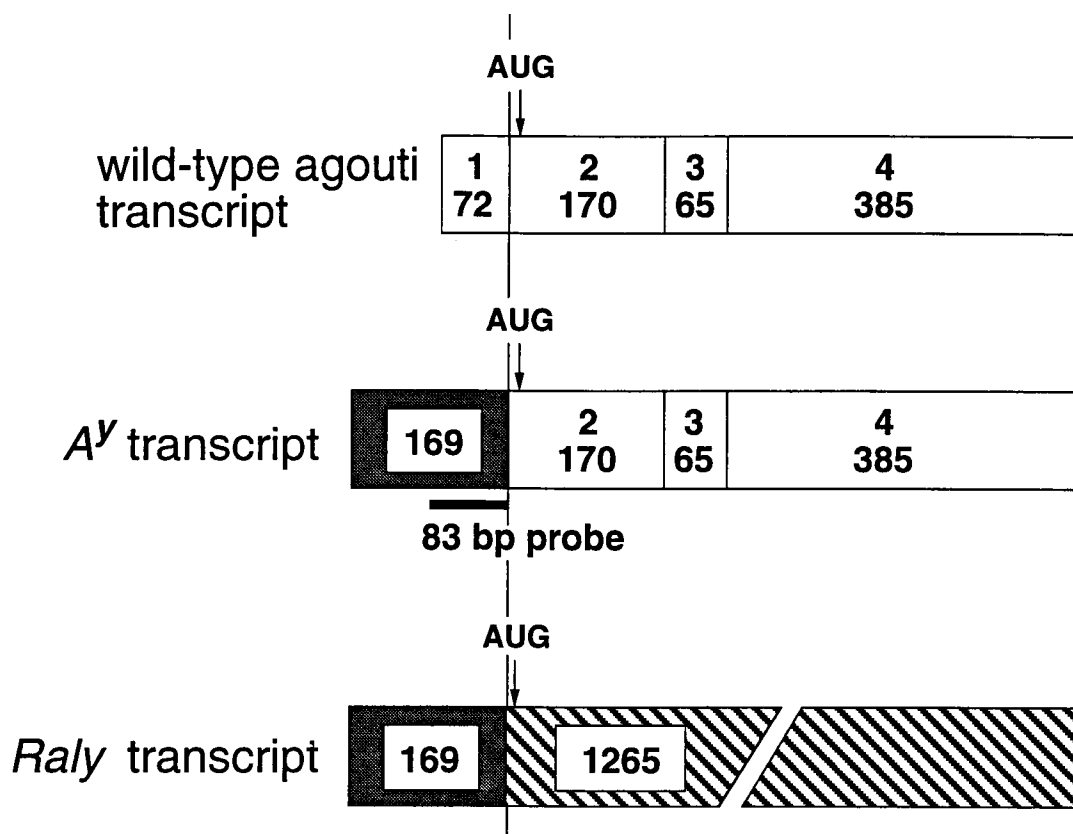


Figure 2. Nucleotide and predicted amino acid sequence of *Raly*.

The sequence of probe A (Fig. 1) is underscored by a dashed line. The 83-bp exon that is alternately processed in the mature transcription unit is indicated by capital letters. The "aataaa" polyadenylation signal is underscored by a solid line. A putative 99 amino acid RNA-binding domain (RBD) is located at the amino terminus of the 296 amino acid translation product and is enclosed by a box, with the highly conserved hexapeptide and octapeptide within the RBD underscored by double and single lines, respectively. The middle of the predicted protein contains a region (amino acids 120-148) where 19 of 29 residues are hydrophobic, and 16 of these residues are proline or valine. Enclosed in brackets near the carboxyl terminus of the predicted protein is a hydrophilic domain where 30 of the 33 residues are glycine or serine, and the domain contains two potential glycosaminoglycan attachment sites (indicated by filled boxes). The translation product contains one potential asparagine glycosylation site (open circle), 12 potential serine/threonine kinase phosphorylation sites (open boxes), and is acidic, especially over its carboxyl terminus where 17 of the final 47 residues are aspartic or glutamic acid. The cDNA sequence presented is a composite of two cDNA clones; cDNA 1 provided all of the sequence except for the first 59 bp which came from cDNA 4.

cggggtgaggagccgaggggaagccgagggggcggaagcggtcgcgactctcgcgctgtg 60
 ctcgggctcctcacgcgggccagggcgccctcttccctcccgccctccgagagcagac 120
 gcgcgctcgcccttcgggtgcgcgcgggcttccctccagacctcggcgggGTGAGCCCTAT 180
 TTCTAGAGACAGCTGCTGCTGACCCTGTAACCTCAAAGGACAACTAGCTGGCTAAACTCA 240

TTCTTGGTACTGgtgaacacc atgtccttgaagattcagaccagcaatgtaaccaacaag 300
 M S L K I Q T S N V T N K 13

aatgaccctaagtccatcaactctcgggcttctcatcggaatctaaacacagctgtgggtg 360
 N D P K S I N S R V F I G N L N T A V V 33

aagaagtcagatgtggagaccatcttttccaagtaaggccgagtggtggttgcctgtg 420
 K K S D V E T I F S K Y G R V A G C S V 53

cacaaaggctatgcttttgtccagtatgccaatgagcgccatgcccgggcagctgtgctg 480
 H K G Y A F V Q Y A N E R H A R A A V L 73

ggagagaatgggcggtgctggctggacagaccctggacatcaacatggctggagagccc 540
 G E N G R V L A G Q T L D I N M A G E P 93

aagcctaataagacccaag gggctaaagagagcagcaactgccatctacaggctgtttgat 600
 K P N R P K G L K R A A T A I Y R L F D 113

tatcaggccgctttct (ccagtgcctgtgccagggcagttccgggtgaagcgaccccggt 660
 Y R G R L S P V P V P R A V P V K R P R 133

gttacagtccctttggttcgctgtcctcaggtacgatacctgtc) aagctctttgcccgc 720
 V T V P L V R R V K T T I P V K L F A R 153

tccacagctgtcactactggctcagccaaaatcaagttaaagagcagtgagctacagacc 780
 S T A V T T G S A K I K L K S S E L Q T 173

atcaaaacagagctgacacagatcaagtccaacatcgatgccctgttggtcgcttgga 840
 I K T E L T Q I K S N I D A L L G R L E 193

cagattgctgaggaacagaaggccaaccagatggcaagaagaagggtgac [agcagcagt 900
 Q I A E E Q K A N P D G K K K G D S S S 213

ggtggaggaggaggcagcagtggtggaggcgccagtagcaatgttggtggtggcagcagc 960
 G G G G G S S G G G G S S N V G G G S S 233

ggcggcagcgaggagctgcagcagcagc] cggctaccagcgccccaagaagacacggct 1020
 G G S G S C S S S R L P A P Q E D T A 253

tctgaggcaggcacaccccaaggagaagtccaaactcgagatgatggtgatgaggagga 1080
 S E A G T P Q G E V Q T R D D G D E E G 273

ctgctaacacatagcgaggaggagctggagcacagccaggacacagatgcagaagatgga 1140
 L L T H S E E E L E H S Q D T D A E D G 293

gccttgacagtaagcagcttaacaggagcattggccaccagcagaaggccatcactgtctc 1200
 A L Q * 296

aggcctcaagccaggcacccatctctggtatgccagtctatagcggtaccagaggaaagc 1260
 tggcagcagtaactctctcccatgcacctagccagtgagtgctacatcctttgcaagt 1320
 ggagttactggcctacccttaccctatgcattcttctgtctgcactgcctgggccaagg 1380
 ggcagaaacactctgtcttcttccccaggacattccccaggttgggtttttctatagg 1440
 ttgaaagtaaaaggggggtgggaaggggtgggaggaacctgacaataaagagattgg 1500
 atccaaataaaaaaaaaa 1517

Figure 3. *Raly* is a member of the heterogeneous nuclear ribonucleoprotein particle (hnRNP) protein gene family.

Shown is the amino acid sequence comparison of the putative RNA-binding domains from the mouse *Raly* and human hnRNP C (Swanson et al., 1987; Wittekind et al., 1992) proteins. Residue identities are depicted by vertical lines, and moderate and highly conserved substitutions are denoted by one or two dots, respectively. The highly conserved signature sites of the RNA-binding domain, RNP 1 and RNP 2, are indicated by solid lines.

RNP 2

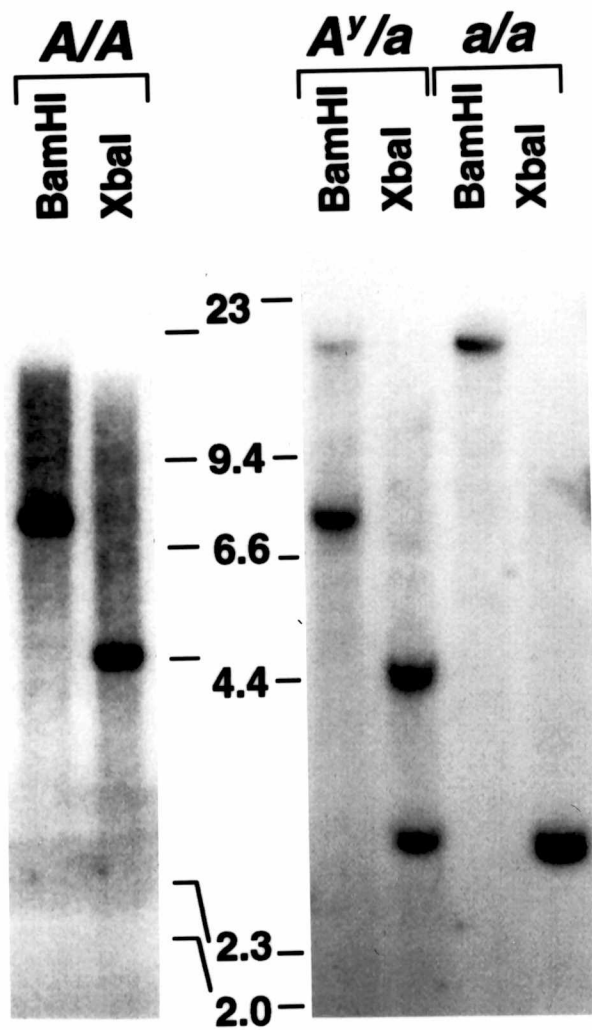
mouse <i>Raly</i>	1	MSLKIQT	SNVTN	KN	DPKS	INSR	VF	IGNL	NTA	VV	KKSD	VE	TIF	SKYGR	VAG	50																																
human hnRNP C	1	M		AS	NVTN	KTD	PR	SM	NSR	VF	IGNL	NTL	VV	KKSD	VEA	IFSKYGI	VG	45																														
mouse <i>Raly</i>	51	CS	VHK	G	YAF	VQ	YAN	ER	H	A	AA	V	L	G	EN	GR	V	LAG	QT	L	D	I	N	M	A	G	E	P	K	P	N	R	P	K	99													
human hnRNP C	46	CS	VHK	G	F	A	F	V	Q	Y	V	N	E	R	N	A	R	A	A	V	A	G	E	D	G	R	M	I	A	G	Q	V	L	D	I	N	L	A	A	E	P	K	V	N	R	G	K	94

RNP 1

Figure 4. Southern blot analyses reveal that the A^y mutation deletes the 3' end of *Raly* but does not alter the 5' end of *agouti*.

(A) Wild-type (A/A , C3H strain), lethal yellow (A^y/a), and nonagouti (a/a) genomic DNA was digested with *Bam*HI or *Xba*I, blotted, and hybridized with a 1.5-kb ^{32} P-labeled fragment of DNA containing the first exon of *agouti* and flanking sequences (probe 1.5 in Bultman et al., 1992). Probe 1.5 detects an 8.0-kb *Bam*HI fragment and a 4.6-kb *Xba*I fragment in both the wild-type *agouti* and A^y alleles and 16.0-kb *Bam*HI and 3.0-kb *Xba*I fragments in the a allele. (B) *M. musculus* (A/A , FVB/N strain), *M. spretus* (A/A), and $A^y/spretus$ [an F_1 hybrid from the cross *M. musculus* (A^y/a) x *M. spretus* (A/A)] genomic DNA was digested with *Bgl*II, blotted, and hybridized with a 255-bp ^{32} P-labeled fragment of DNA that was PCR amplified from the *Raly* cDNA clone with oligonucleotide primers corresponding to base pairs 1080-1099 and the reverse complement of base pairs 1334-1315 (Fig. 2). The 255-bp probe corresponding to the 3' end of *Raly* detects a 5.5-kb fragment in *M. musculus*, an 8.0-kb fragment in *M. spretus*, and only the 8.0-kb *M. spretus*-specific fragment in $A^y/spretus$. DNA molecular size standards are shown in kb.

A



B

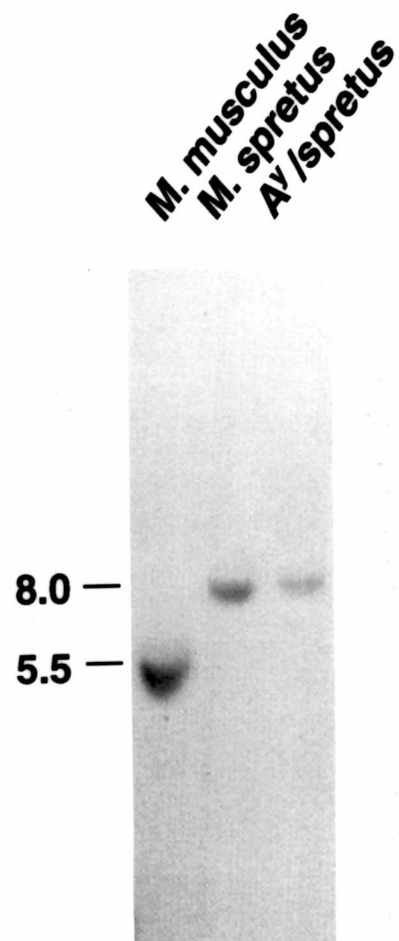


Figure 5. Expression of the wild-type *Raly* gene.

(A and B) Northern analysis. The *Raly* cDNA clone was ^{32}P -labeled and hybridized to poly(A)⁺ RNA (~2.5 µg per lane) from (A) a variety of adult mouse tissues and day 4 (d4) neonate skin, and to (B) whole mouse embryos ranging from day 10 through day 18 of gestation (E10 to E18). The filter in (A) was washed under high stringency (0.2X SSC, 0.1% SDS at 68° C) or reduced stringency (0.2X SSC, 0.1% SDS at 50° C) conditions as indicated. The filter in (B) was washed under reduced stringency conditions. RNA molecular size standards are shown on the left in kilobases. (C) RT-PCR analysis. Eighteen C57BL/6J blastocysts were collected at day 3.5 *post coitum*, total RNA was prepared, reverse transcribed, and the entire coding region of the *Raly* mRNA was amplified by PCR with oligonucleotide primers corresponding to base pairs 108-127 (5' to the translation initiation codon) and the reverse complement of base pairs 1334-1315 (3' to the termination codon) in the *Raly* cDNA (Fig. 2). The PCR product was electrophoresed through an agarose gel, transferred to GeneScreen Plus (DuPont), and hybridized to a ^{32}P -labeled *Raly* cDNA probe. A non-distinct fragment of approximately 1.2 kb is present in the blastocyst sample (lane 1), which is consistent with a comigration of the two expected fragments of 1144 and 1227 bp. Lane 2 is the control sample consisting of *E. coli* tRNA instead of blastocyst RNA. DNA molecular size standards are shown on the left in base pairs.

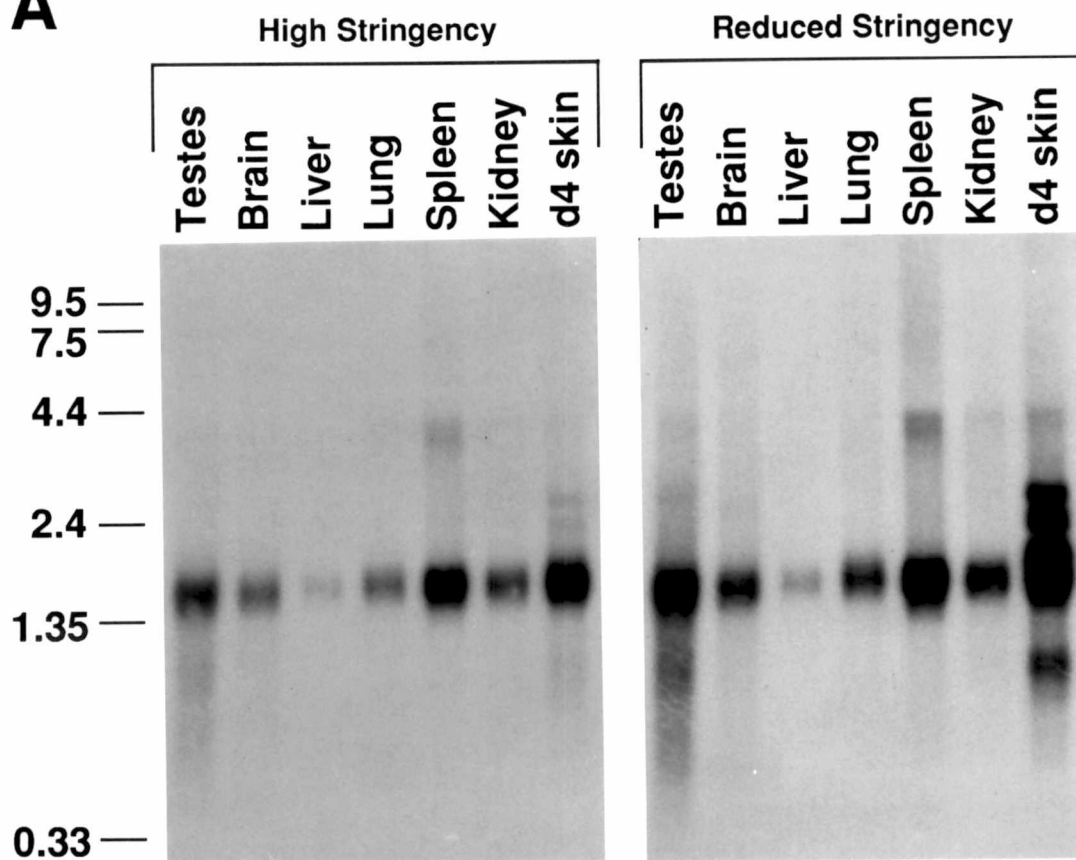
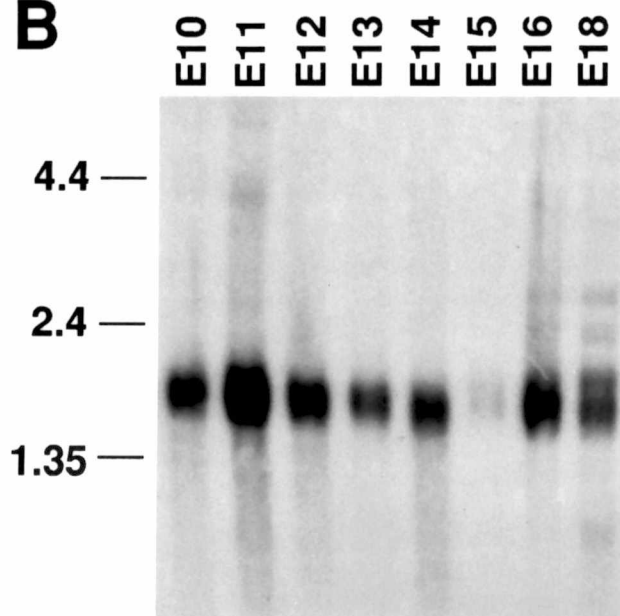
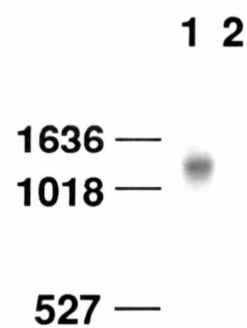
A**B****C**

Figure 6. The expression of *Raly* is disrupted from the *A^y* allele.

Total RNA from the adult liver of wild-type *Mus musculus* and *M. spretus*, and from day 10 neonate skin of two *M. musculus*/*M. spretus* F₁ hybrids (*A^y/spretus* and *a/spretus*) was reverse transcribed, and a portion of the 5' end of the *Raly* cDNA was amplified by PCR and sequenced. Shown are bases 178-171 (bottom to top) of the *Raly* cDNA from these wild-type mice and F₁ hybrids. *Mus musculus* has a "g" at position 175 (first panel), whereas a species specific sequence polymorphism at nucleotide 175 results in a "g" to "t" transversion in the *Raly* cDNA of *M. spretus* (second panel). The *A^y/spretus* hybrid (third panel) has a "t" at position 175, indicating that *Raly* is expressed from the *M. spretus* allele (+), but is not expressed from the *M. musculus* allele (*A^y*). The *a/spretus* hybrid has a "g" and "t" comigrating at position 175 (fourth panel), indicating that the *Raly* gene is expressed from both the *M. spretus* (+) and *M. musculus* alleles (*a*).

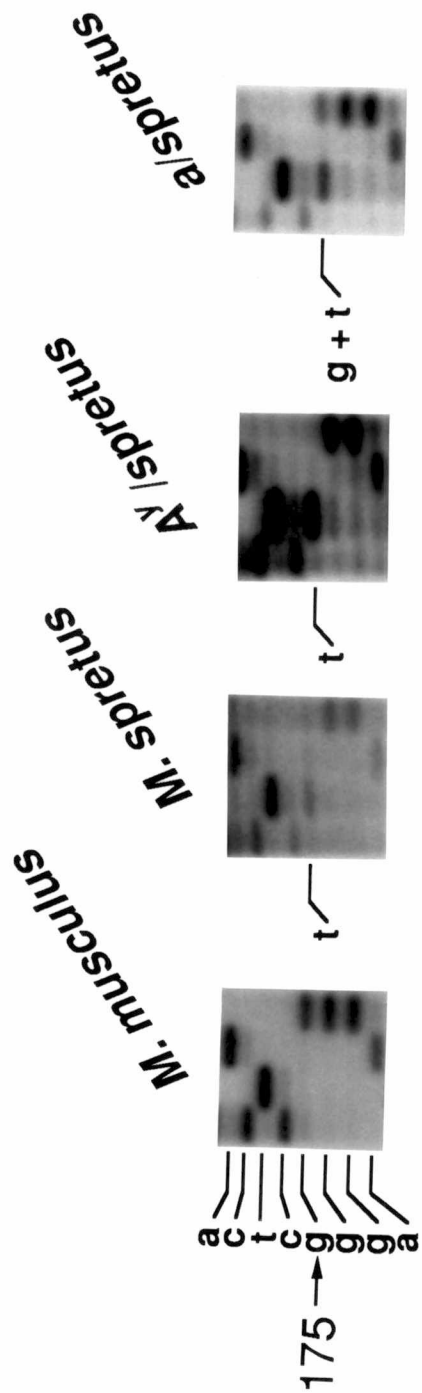
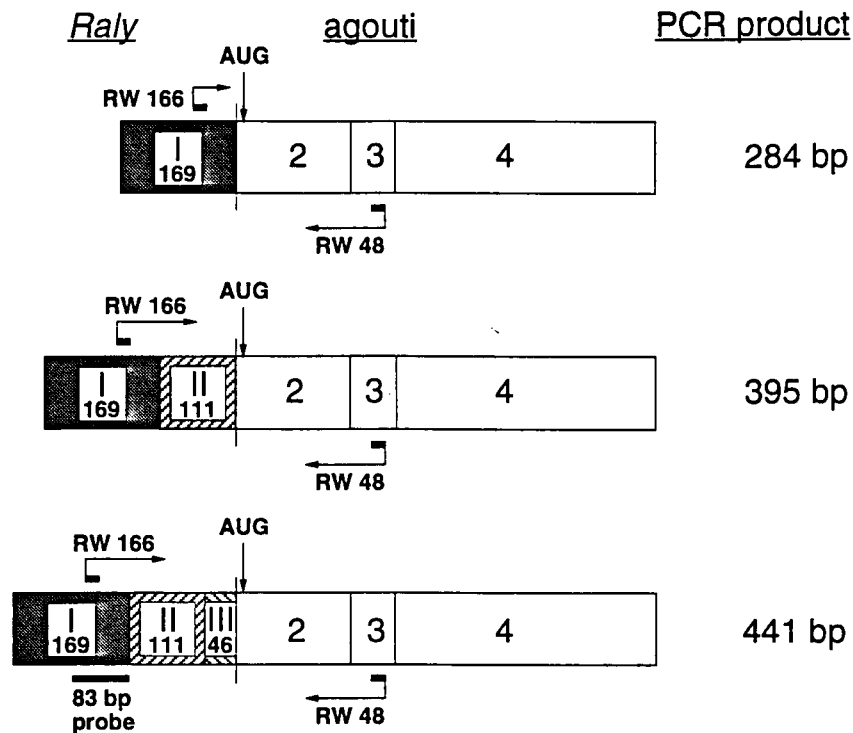


Figure 7. The *A γ* transcription unit is alternately processed at the 5' end, giving rise to three separate transcripts.

Total RNA was prepared from adult *A γ* spleen, reverse transcribed, and PCR amplified across the *Raly/agouti* chimeric junction with oligonucleotide primers corresponding to base pairs 108-127 of the wild-type *Raly* gene [first 20 bases in (B)] and the reverse complement of base pairs 294-275 of the wild-type agouti gene [reverse complement of the last 20 bases in (B)]. As shown in (A), this analysis gave rise to fragments of 284, 395, and 441 bp. Each of these fragments were subcloned into the pCRTM II vector (Invitrogen) and sequenced. Based on sequence analysis of these PCR fragments, we deduced that there are three different *A γ* transcripts whose 5' ends are comprised of the first exon of *Raly* (Roman numeral one, 169 bases, stippled box), and variously comprised of 111- and 46-base segments (cross-hatched and black box, respectively). The sequence derived from the agouti exons is common to all three *A γ* transcripts and is shown in the open boxes numbered 2-4. The positions of the oligonucleotide primers and translation initiation codons are indicated. (B) Combined sequence of the PCR fragments in (A). The boxed regions represent the sequences that are differentially incorporated in the three different *A γ* transcripts. The sequence of *Raly* exon I, and the 111- and 46-base segments are shown in lower case letters, and the agouti sequence is presented in capital letters. The boundary between agouti exons 2 and 3 is delimited by a vertical line, and the predicted amino acid sequence is indicated. The numbers of nucleotides and amino acids are given on the right.

A

Three alternately spliced lethal yellow transcripts

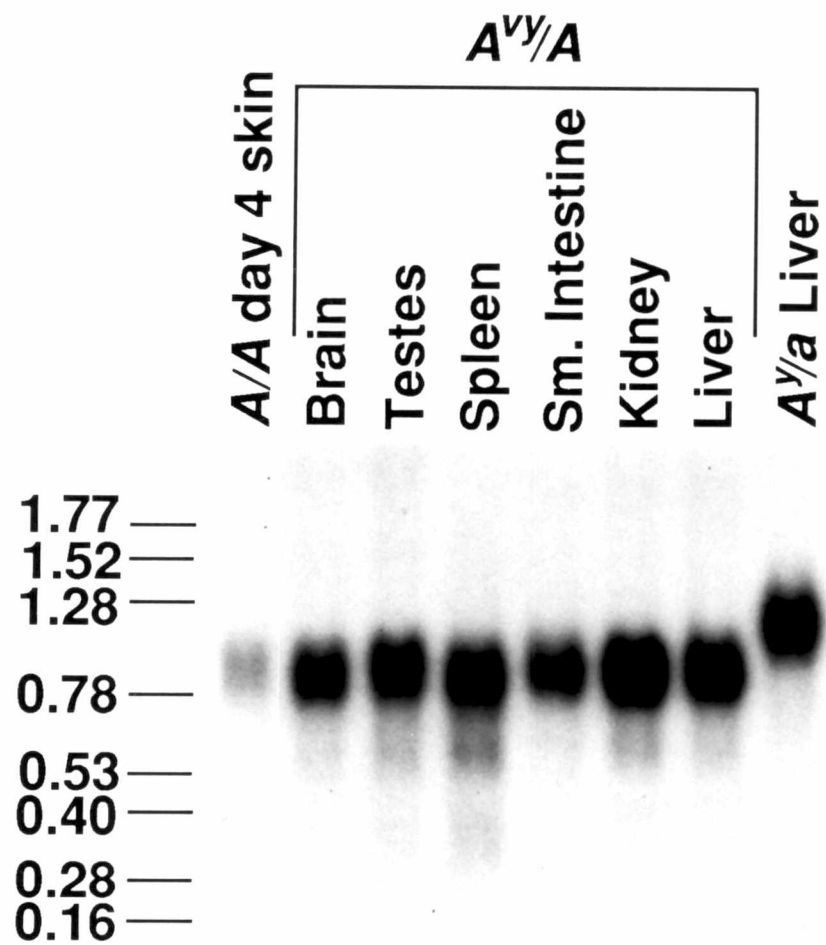


B

tccgagagcagacgcgcgccttcggtgccgcggccttccctccagacctcggcgc	60	Exon I
gggaactggcatcaaagtaccatttcccaccagctctgagtccttgagcctctgcagcctc	120	Exon II
agaagaggggagtcacagctgaaacctccaggaaccaccgggggtcccagaagagagcga	180	Exon III
ttattgatgaatcctggaacatgccctgtactgtaagagGCTTCTCAGGATGGATGTCAC	240	
	M D V T 4	
CCGCCTACTCCTGGCCACCCTAGTGAGCTTCTGTGCTTCTCACCGTCCACAGCCACCT	300	
R L L L A T L V S F L C F F T V H S H L	24	Exon 2
GGCACTCGAGGAGACGCTTGGAGATGACAGGAGTCTGCGGAGTAACCTCCATGAATC	360	
A L E E T L G D D R S L R S N S S M N S	44	
GCTGGATTTCTCCTCTGTTTCTATCGTGGCACTGAACAAGAAATCCAAGAAGATCAGCAG	420	
L D F S S V S I V A L N K K S K K I S R	64	Exon 3
AAAAGAAGCCGAGAAGCGGAA	441	
K E A E K R	70	

Figure 8. Northern blot analysis of agouti locus expression in tissues of adult viable yellow (A^{vy}) mice.

The wild-type agouti cDNA clone was ^{32}P -labeled and hybridized to poly(A)⁺ RNA (2.5 μg per lane) from a variety of adult A^{vy} tissues, and to wild-type agouti (day 4 neonate skin) and lethal yellow (adult liver) controls. RNA molecular size standards are shown on the left in kilobases.



PART VII**A MOLECULAR MODEL FOR GENETIC AND PHENOTYPIC
CHARACTERISTICS OF THE MOUSE LETHAL YELLOW (A^y)
MUTATION**

ABSTRACT

Lethal yellow (A^y) is a mutation at the mouse agouti locus that causes a number of dominant pleiotropic effects, including a completely yellow coat color, obesity, an insulin-resistant type II diabetic condition, and an increased propensity to develop a variety of spontaneous and induced tumors. Additionally, homozygosity of A^y results in preimplantation lethality, which terminates development by the blastocyst stage. Previously, we demonstrated that A^y is associated with a deletion of at least part of a gene called *Raly*, which maps close to agouti in chromosome 2. Interestingly, we determined that this deletion does not affect the DNA structure of the agouti locus. Here we demonstrate that the 5' end of *Raly* lies 280 kb proximal to the 3' end of agouti, and that *Raly* is transcribed in the same orientation as agouti. We also show that the A^y mutation is the result of a 170-kb deletion that removes the entire *Raly* gene, except for the promoter and non-coding first exon. Furthermore, we present a model for the A^y allele that can explain the recessive lethality and the dominant pleiotropic effects associated with this mutation.

INTRODUCTION

Lethal yellow (A^y) is a dominant mutation at the agouti locus that dates back to the mouse fancy and has been intensively studied for decades. Lethal yellow heterozygotes develop a uniform yellow coat color over the entire body, instead of the wild-type agouti pigmentation in which each hair shaft is black with only a subapical band of yellow. In addition to its effect on pigmentation, the A^y allele also causes a number of dominant pleiotropic effects which include a non-insulin dependent diabetic-like condition (Hellerstrom and Hellman, 1963), marked obesity (Dickerson and Gowen, 1947; Fenton and Chase, 1951; Carpenter and Mayer, 1958; Plocher and Powley, 1976; Friedman and Leibel, 1992), and an increased propensity to develop a variety of spontaneous and induced tumors (reviewed in Wolff, 1987; Wolff et al., 1986). When homozygous, the A^y allele also causes a preimplantation lethality, which was first revealed at the turn of the century

as an alteration in normal Mendelian inheritance (Cuenot, 1908; Castle and Little, 1910; Ibsen and Steinleder, 1917; Kirkham, 1917; Kirkham, 1919). The lethality has subsequently been shown to occur prior to implantation (Robertson, 1942; Eaton and Green, 1963; Pederson, 1974; Pederson and Spindle, 1976; Granholm and Johnson, 1978) and to be associated with abnormalities in both the trophectoderm and the inner cell mass (Papaioannou and Gardner, 1979; Papaioannou, 1988; Barsh et al., 1990).

Extensive genetic experiments conducted at the Oak Ridge National Laboratory demonstrated that recombination can occur between A^y and other alleles at the agouti locus (Russell et al., 1963; Siracusa et al., 1987). This unusual finding was first revealed in a cross involving A^y and the recessive-lethal nonagouti (a^x) allele. In this case, wild-type agouti offspring arose in crosses between A^y/a^x compound heterozygotes, and were unequivocally shown to result from recombination between these two alleles (Russell et al., 1963). Recombination also appears to have occurred in crosses involving A^y and the black-and-tan mutation (a^t), and in crosses between A^y and the nonagouti (a) allele (Siracusa et al., 1987). Collectively, the results from these recombination events place A^y 0.1 cM proximal to agouti, which led to the suggestion that A^y is pseudoallelic with agouti (Russell et al., 1963; Siracusa et al., 1987).

The structure of the agouti gene was recently characterized by us (Bultman et al., 1992) and by others (Miller et al., 1993), and shown to produce a mRNA of approximately 0.8 kb. We also demonstrated that A^y is an allele of agouti and it gives rise to three size-altered mRNAs, each approximately 1.1 kb in length, that are ectopically overexpressed in every tissue examined to date (Bultman et al., 1992; Michaud et al., 1993). The A^y -specific transcripts consist of the normal coding and 3' untranslated regions of the agouti transcription unit, coupled with novel sequences at their 5' ends (Michaud et al., 1993). Further characterization revealed that the 5'-most portion of the novel sequence in these A^y transcripts corresponds to the non-coding first exon of a novel gene. This gene, *Raly* (hnRNP associated with lethal yellow), belongs to a family of RNA-binding protein genes implicated in pre-mRNA processing and developmental regulation. *Raly* is normally expressed in a ubiquitous manner, including the

preimplantation embryo, and is tightly linked to the agouti locus in chromosome 2 (Michaud et al., 1993). In the A^y allele, it appears that the coding region of the agouti gene is under the transcriptional control of the ubiquitous *Raly* promoter, leading to the proposal that the ectopic overexpression of the wild-type agouti gene product is responsible for the suite of dominant pleiotropic effects in A^y heterozygotes (Michaud et al., 1993). Moreover, because the *Raly* gene is not expressed from the A^y allele, we hypothesized that the lack of the *Raly* gene product in the preimplantation embryo is associated with the recessive lethality of homozygous A^y mice (Michaud et al., 1993).

In an attempt to understand further the molecular basis of the dominant pleiotropic effects, the recessive lethality, and the unusual recombination events associated with the lethal yellow mutation, we have undertaken a more thorough characterization of the A^y allele. Here we demonstrate that the 5' end of the *Raly* gene normally lies 280 kb proximal to the 3' end of the agouti gene, and that there is a 170-kb deletion associated with the A^y mutation that removes the entire *Raly* gene except for the promoter and non-coding first exon. Additionally, we present a model that explains how A^y is able to recombine with other agouti locus alleles, and also how A^y produces novel 1.1-kb transcripts.

MATERIALS AND METHODS

Pulsed-Field Gel Analysis

Pulsed-field gel electrophoresis (PFGE) analyses were conducted essentially as described (Klebig et al., 1992). The digested DNAs were electrophoresed in the CHEF-DR™ II Pulsed Field Electrophoresis System (BioRad) at 200 V, 12°C, 10-40 second ramp for 25 hr.

Southern Blot Analysis

Genomic DNA (10 µg) was digested with restriction enzymes, electrophoresed through agarose gels, and blotted to GeneScreen (DuPont) using standard procedures (Ausubel et al., 1988; Sambrook et al., 1989). Radiolabeled hybridization probes were prepared with the random-hexamer

labeling technique (Feinberg and Vogelstein, 1984). Posthybridization filter washes were conducted at high stringency (0.2X SSC, 0.1% SDS, 68° C), except in the case of the 83-bp *Raly* second exon and 111-bp agouti probes which were washed at reduced stringency (0.2X SSC, 0.1% SDS, 50° C).

Probes

Isolation of *Raly* genomic clones was as described (Michaud et al., 1993). Probe A is a 2.0-kb *Xba*I fragment that was isolated from a *Raly* genomic clone and subcloned into pGEM4 (Promega). Probes B and C were amplified by PCR from plasmid templates and their nucleotide sequences have been published (Michaud et al., 1993: uppercase letters in Fig. 2.; C, 111-bp boxed region in Fig. 7B).

RESULTS

Long-Range Physical Map of *Raly* and Agouti

Previously, we demonstrated that *Raly* and agouti are tightly linked in mouse chromosome 2, and that the *A^y* mutation deletes at least the 3' end of the *Raly* gene but does not affect the genomic DNA structure of the agouti locus (Michaud et al., 1993). We have now compared the DNA structure of the *A^y* and wild-type alleles using PFGE. DNA samples were digested with the enzymes *Bss*HII, *Eag*I, and *Ksp*I, subjected to PFGE, blotted, and hybridized consecutively with probes corresponding to a 5' segment of the *Raly* first intron (Fig. 1A) and the wild-type agouti cDNA (Fig. 1B). In wild-type DNA, both the *Raly* and agouti probes hybridized to a 280-kb *Bss*HII fragment. Because the agouti gene contains a *Bss*HII site within its first intron and its last exon, the agouti cDNA probe also hybridized to an internal 17-kb *Bss*HII fragment, and to a 75-kb *Bss*HII fragment that contains the 3' end of the agouti gene and 3' flanking sequences. In *A^y* DNA, both probes detected a smaller *Bss*HII fragment of 110 kb. A similar result was obtained with the enzyme *Ksp*I, in which both the *Raly* and agouti probes hybridized to a 300-kb fragment in wild-type DNA, and a smaller 130-kb fragment in *A^y* DNA. These data are consistent with the interpretation that both the *Raly* and agouti probes lie on single *Bss*HII and

*Ksp*I fragments of 280 kb and 300 kb, and that the *A^y* mutation is associated with a 170-kb deletion.

The indications with *Bss*HII and *Sma*I that the *Raly* and agouti genes lie on the same 280 kb of DNA, and that the *A^y* mutation is due to a 170-kb deletion, was substantiated by *Eag*I digestions. We determined that there are *Eag*I sites flanking *Raly* (in the CpG island at the 5' end of the gene, Fig. 1C) and agouti (immediately 3' of the gene, Fig. 1C) that cut to completion in genomic DNA (data not shown). In wild-type DNA, the probe from the first intron of the *Raly* gene hybridized with a 110-kb *Eag*I fragment, and the agouti cDNA probe hybridized with a 170-kb fragment. This result indicates that there is at least one *Eag*I site between the two genes. In the *A^y* allele, both the *Raly* and agouti probes hybridized to a 110-kb *Eag*I fragment. These data strongly suggest that the *Eag*I site between these loci is encompassed by an *A^y* deletion of approximately 170 kb in length. There is an additional *Eag*I site in the 5' end of the *Raly* gene (not shown in Fig. 1C) that is also encompassed by the *A^y* deletion, but this site does not appear to cut in wild-type genomic DNA based on the *Eag*I digestions shown in Fig. 1A. To verify this observation, the *Raly* cDNA was hybridized with the same PFGE blot shown in Fig. 1A and 1B. Like the *Raly* first intron probe, the cDNA probe hybridized with a single 110-kb *Eag*I fragment (data not shown), again suggesting that the *Eag*I site in the 5' end of the *Raly* cDNA may be methylated in wild-type genomic DNA.

Based on the results from *Bss*HII, *Ksp*I, and *Eag*I, and the additional enzymes *Mlu*I, *Sma*I, and *Not*I (data not shown), a long-range restriction map of wild-type and *A^y* DNA surrounding the *Raly* and agouti loci was constructed (Fig. 1C). Taken together, these data indicate that the 5' end of *Raly* lies about 280 kb proximal to agouti and that the *A^y* mutation is the result of a 170-kb deletion at the *Raly* locus.

Localization of the *A^y* Deletion Breakpoints

We previously determined that, in the *A^y* allele, the first exon of *Raly* is present, the 3' end of *Raly* is deleted, and none of the four exons of the agouti gene are deleted or structurally altered (Michaud et al., 1993). To define the location of the *A^y* 5' deletion breakpoint in the *Raly* gene more

precisely, Southern blots of genomic DNA were hybridized with probes specific to various portions of the 5' end of the *Raly* gene (Fig. 2). To facilitate these analyses, we utilized DNA polymorphisms to differentiate the *Mus musculus* wild-type agouti and A^y alleles from the *M. spretus* wild-type allele in an F_1 hybrid ($A^y/M.s.$).

Given that the non-coding first exon of *Raly* is present in the A^y allele (Michaud et al., 1993), a portion of the first intron of *Raly* was used as a probe (probe A in Fig. 2E). The *Raly* first intron probe detected a 4.5-kb *M. musculus*-specific *EcoRI* fragment, a 7.0-kb *M. spretus*-specific fragment, and both the 4.5- and 7.0-kb fragments in $A^y/M.s.$, indicating that this portion of the *Raly* first intron is present in the A^y allele (Fig. 2A). However, a probe specific to the 83-bp second exon of *Raly* (probe B in Fig. 2E), which detected a 12.0-kb *M. musculus*-specific *BglII* fragment and a 7.8-kb *M. spretus*-specific fragment, hybridized only to the 7.8-kb *M. spretus*-specific fragment in $A^y/M.s.$, indicating that the second exon of *Raly* is deleted in the A^y allele (Fig. 2B). In addition to the *Raly* second-exon probe, probes from the first and last coding exons of *Raly* demonstrated that these regions of the *Raly* gene are also deleted in the A^y allele (data not shown, and Michaud et al., 1993, respectively). The observation that the *Raly* cDNA probe detected only a 110-kb *EagI* fragment on the PFGE blot (see above), indicates that the *Raly* gene does not extend in the 3' direction past the last *EagI* site that is encompassed by the A^y deletion. Collectively, these data provide compelling evidence that the entire *Raly* gene, except for the promoter and non-coding first exon, is deleted in A^y mice.

In order to localize further the 5' deletion breakpoint between the first and second *Raly* exons, the *Raly* first intron probe (probe A) was additionally hybridized to DNAs digested with *BamHI* and *SacI*. A size-altered fragment unique to the A^y allele was detected with each of these enzymes (Fig. 2C). To test whether these size-altered A^y -specific fragments correspond to the deletion breakpoint region, this same blot was also hybridized with a probe mapping 3' to the deletion breakpoint. This probe corresponds to a 111-bp region (Probe C in Fig. 2E) that is differentially incorporated into A^y -specific transcripts (Michaud et al., 1993). This 111-bp probe is present in A^y DNA and normally hybridizes to a 170-kb *EagI*

fragment in wild-type DNA(data not shown), as does the agouti cDNA probe, which unequivocally places it 3' to the deleted region in the A^y allele (see above and Fig. 1C). As shown in Fig. 2D, the same size-altered A^y -specific *Bam*HI and *Sac*I fragments that were detected with the *Raly* first intron probe (Probe A) also hybridized with the 111-bp probe (Probe C). As expected, probes A and C each detected different wild-type *Bam*HI and *Sac*I fragments because these probes lie on opposite sides of the deleted region in the A^y allele. Taken together, these data demonstrate that the 5' A^y deletion breakpoint occurs approximately 12 kb downstream from the *Raly* first exon (Fig. 2E), and that the 3' deletion breakpoint occurs 170-kb downstream from this region, at a position approximately 105 kb upstream from the first exon of the agouti gene (Fig. 1C).

DISCUSSION

As part of our effort to clone and characterize the agouti gene, we previously determined that the lethal yellow mutation expresses three size-altered 1.1-kb mRNAs that each contain the first exon of another gene (*Raly*) that is closely linked to agouti in mouse chromosome 2. We also previously determined that at least a portion of the *Raly* gene has been removed from the A^y allele, which indicated that the A^y mutation is associated with a deletion (Michaud et al., 1993). Here we have utilized PFGE to demonstrate that the 5' end of the *Raly* gene lies 280 kb proximal to the 3' end of the agouti gene (see below), and that the deletion associated with the A^y mutation is 170 kb in length. The deletion encompasses a region that starts at a site about 12 kb downstream from the non-coding first exon of *Raly*, extends through the remainder of the *Raly* gene, and terminates at a position estimated to lie 105 kb 5' of the first exon of agouti.

The 111-bp probe that was used for characterizing the 3' A^y deletion breakpoint was originally identified by the fact that is differentially incorporated into three alternately processed A^y -specific transcripts (Michaud et al., 1993). All three A^y transcripts contain the non-coding first exon of *Raly* at their 5' termini. Two of the three A^y transcripts contain differentially spliced regions, 111 bp and 46 bp in length (labeled A and B,

respectively, in Figs. 3 and 4) derived from DNA located between the *Raly* first exon and the agouti coding exons. We originally proposed that these 111- and 46-bp regions might be sequences that arose from cryptic splicing events in a unique A^y pre-mRNA. More recently, however, we have determined that these A^y sequences are also found in transcripts derived from other alleles, including A^w and a^t . Therefore, the 111-bp and 46-bp sequences actually represent alternatively processed exons of the agouti gene (Bultman et al., 1993). The fact that the 111-bp agouti exon detects the 3' A^y deletion breakpoint indicates that the 3' breakpoint occurs very close to, or possibly within, the 5' end of the agouti gene, and that this 111-bp agouti exon maps to a site located approximately 105 kb upstream of what was previously identified as the first exon of the wild-type agouti gene. Based on the observation that the 111-bp exon does not contain *Bam*HI or *Sac*I enzyme recognition sites, but detects two wild-type fragments with each enzyme (Fig. 2D), the 111-bp exon may actually represent two agouti exons. The nature and functional significance of the agouti transcripts that incorporate the 111- and 46-bp regions will be reported elsewhere (Bultman et al., 1993).

Our interpretation of these results is based on the assumption that the 170-kb deletion in A^y is contiguous. We have not, however, mapped the 111-bp probe with respect to the remainder of the agouti gene. For this reason, it is possible that the 170 kb of deleted DNA is actually comprised of more than one deletion, with a single deletion of less than 170 kb occurring between the *Raly* first intron and the downstream 111-bp region, and an additional deletion occurring between the 111-bp region and the remainder of the agouti gene. Whether or not the 170-kb A^y deletion is contiguous, these data allow us to propose a model which can explain the recombination between A^y and other agouti alleles, and how the novel 1.1-kb transcripts arise from the A^y allele.

The A^y allele can recombine with a^x (Russell et al., 1963) and probably also recombines with a and a^t (Siracusa et al., 1987). Utilizing the confirmed and probable recombination events, the recombination frequency between A^y and these other agouti locus alleles is estimated to be 0.1% (Table 1), with A^y being proximal to a^x . The fact that A^y can recombine with these other agouti alleles led us to propose that the agouti gene is either

very large or that A^y is associated with a separate gene that is pseudoallelic with agouti. We ruled out the latter possibility as a result of our recent cloning and characterization of the agouti gene. To explain the 0.1% recombination frequency between A^y and the other agouti alleles, we hypothesized that a recombination hot-spot exists within the first intron of agouti (Bultman et al., 1992). However, based on the unique structure of the A^y allele, the recombination between A^y and other agouti alleles can be reconciled with a model that incorporates a conventional recombination mechanism without the need to invoke a recombination hot-spot. This model is presented in Figure 3 and is based on the following observations: the 5' end of the *Raly* gene lies about 280 kb proximal to the 3' end of the agouti gene; the A^y deletion starts within the first *Raly* intron and continues for 170 kb through the remainder of the *Raly* gene; two of the mutations that A^y likely recombines with, a and a^t , are each associated with the insertion of novel DNA sequences at the same position, located in the intron upstream of the first coding exon (Bultman et al., 1992); and the deletion in A^y , whether it is contiguous or not, would leave approximately 105 kb of DNA that normally occurs between the 3' end of the remaining portion of *Raly* and the mutated region associated with the a and a^t mutations. Recombination anywhere within this 105-kb region would give rise to a wild-type agouti allele. The 0.1% recombination frequency observed between A^y and other agouti locus alleles is compatible with the 105-kb physical distance between the deletion in A^y and the mutated region associated with the a and a^t mutations, using the estimate that 1% recombination frequency corresponds to 1000 kb.

The data regarding the structure of the A^y allele described in this report also suggests a model to account for the production of the different-sized *Raly*/agouti chimeric transcripts that are expressed from the A^y allele (Fig. 4). We propose that transcription initiates normally at the *Raly* promoter and proceeds through the first *Raly* exon; however, because the remainder of the gene, which likely includes any transcription termination signal, is deleted, transcription proceeds into the remaining 105 kb of intergenic DNA and through the downstream exons of the agouti gene. The resulting novel primary transcript could be spliced in a manner whereby the

splice donor associated with the first exon of *Raly* connects to the available splice acceptors in the downstream region. Because all of the normal potential splice acceptors of *Raly* are deleted, splicing would occur to the available splice acceptors of the downstream exons of the agouti gene. Alternative splicing could also occur by differentially incorporating exons from the 111-bp and 46-bp regions that are positioned just 3' to the *A^y* deletion breakpoint. Overall, this model predicts that the *Raly* gene would not be functional in the *A^y* allele, and that transcripts with the coding potential of the wild-type agouti gene would be overexpressed in a ubiquitous manner under the influence of the *Raly* promoter. In accordance with this model, we propose that deletion of the *Raly* gene is responsible for the recessive embryonic lethality of *A^y*, and that the ubiquitous overexpression of the agouti gene, driven by the *Raly* promoter, is directly associated with the dominant pleiotropic effects of *A^y*.

REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1988). *Current protocols in molecular biology*. John Wiley & Sons, New York.
- Barsh, G. S., Lovett, M., and Epstein, C. J. 1990. Effects of the lethal yellow (A^y) mutation in mouse aggregation chimeras. *Development* **109**: 683-690.
- Bultman, S. J., Michaud, E. J., and Woychik, R. P.. (1992). Molecular characterization of the mouse agouti locus. *Cell* **71**: 1195-1204.
- Carpenter, K. J., and Mayer, J. (1958). Physiologic observations on yellow obesity in the mouse. *Am. J. Physiol.* **193**: 499-504.
- Castle, W. E., and Little, C. C. (1910). On a modified Mendelian ratio among yellow mice. *Science* **32**: 868-870.
- Cuenot, L. (1908). Sur quelques anomalies apparentes des proportions Mendeliennes. *Arch. Zool. Exp. Gen.* **9**: 7-15.
- Dickerson, G. E., and Gowen, J.W. (1947). Hereditary obesity and efficient food utilization in mice. *Science* **105**: 496-498.
- Eaton, G. J., and Green, M.M. (1963). Giant cell differentiation and lethality of homozygous yellow mouse embryos. *Genetica* **34**: 155-161.
- Feinberg, A. P., and Vogelstein, B. (1984). Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**: 266-267.
- Fenton, P. F., and Chase, H. B. (1951). Effect of diet on obesity of yellow mice in inbred lines. *Proc. Soc. Exper. Biol.* **77**: 420-422.

- Friedman, J. M., and Leibel, R. L. (1992). Tackling a weighty problem. *Cell* **69**: 217-220.
- Granholm, N.H., and Johnson, P.M. (1978). Enhanced identification of lethal yellow (A^y/A^y) mouse embryos by means of delayed development of four-cell stages. *J. Exp. Zool.* **205**: 327-333.
- Hellerstrom, C., and Hellman, B. (1963). The islets of Langerhans in yellow obese mice. *Metabolism* **12**: 527-536.
- Ibsen, H. L., and Steigleder, E. (1917). Evidence for the death *in utero* of the homozygous yellow mouse. *Am. Nat.* **51**: 740-752.
- Kirkham, W. B. (1917). Embryology of the yellow mouse. *Anat. Rec.* **11**: 480-481.
- . (1919). The fate of homozygous yellow mice. *J. Exp. Zool.* **28**: 125-135.
- Klebig, M. L., Kwon, B. S. and Rinchik, E. M. (1992). *Mammalian Genome* **2**, 51-63.
- Miller, M. M., Duhl, D. M. J., Vrieling, H., Cordes, S. P., Ollmann, M. M., Winkes, B. M., and Barsh, G. S. (1993). Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation. *Genes & Dev.* **7**: 454-467.
- Michaud, E. J., Bultman, S. J., Stubbs, L. J., and Woychik, R. P. (1993). The embryonic lethality of homozygous lethal yellow mice (A^y/A^y) is associated with the disruption of a novel RNA-binding protein. *Genes & Dev.* **7**: 1203-1213.
- Papaioannou, V. E. (1988). Investigation of the tissue specificity of the lethal yellow (A^y) gene in mouse embryos. *Devl. Genet.* **9**: 155-165.

- Papaioannou, V., and Gardner, R. L. (1979). Investigations of the lethal yellow A^y/A^y embryo using mouse chimaeras. *J. Embryol. Exp. Morph.* **52**: 153-163.
- Pedersen, R. A. (1974). Development of lethal yellow (A^y/A^y) mouse embryos *in vitro*. *J. Exp. Zool.* **188**: 307-319.
- Pedersen, R. A., and Spindle, A.I. (1976). Genetic effects on mammalian development during and after implantation. In *Embryogenesis in Mammals, CIBA Foundation Symposium 40* (eds. K. Elliot and M. O'Connor), pp. 133-154. Elsevier, New York.
- Plocher, T.A., and Powley, T.L. (1976). Effect of hypophysectomy on weight gain and body composition in the genetically obese yellow (A^y/a) mouse. *Metabolism* **25**: 593-602.
- Robertson, G. G. (1942). An analysis of the development of homozygous yellow mouse embryos. *J. Exp. Zool.* **89**: 197- 231.
- Russell, L. B., McDaniel, M. N. C., and Woodiel, F. N. (1963). Crossing-over within the α "locus" of the mouse. *Genetics* **48**: 907a.
- Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). *Molecular cloning: A laboratory manual, second edition*. Cold Spring Harbor Press, Cold Spring Harbor, New York.
- Siracusa, L. D., Russell, L. B., Eicher, E. M., Corrow, D. J., Copeland, N. G., and Jenkins, N. A. (1987). Genetic organization of the *agouti* region of the mouse. *Genetics* **117**: 93-100.
- Wolff, G. L. (1987). Body weight and cancer. *Am. J. Clin. Nutr.* **45**: 168-180.
- Wolff, G. L., Roberts, D. W., and Galbraith, D. B. (1986). Prenatal determination of obesity, tumor susceptibility, and coat color pattern in

viable yellow (A^{vy}/a) mice: The yellow mouse syndrome. *J. Hered.* **77**: 151-158.

APPENDIXES

APPENDIX A

TABLES

Table 1. Agouti exceptions from crosses involving lethal yellow (A^y) mice at the Oak Ridge National Laboratory*.

<u>Cross</u>			N	Agouti exceptions	M [†]	Recombination frequency(cM) [‡]
Females		Males				
A^y/a^t	x	a^t/a	2160	1	2	0.09 (0.005-0.5)
A^y/a	x	a/a	3631 [§]	2	2	0.11 (0.02-0.4)
A^y/a^x	x	A^y/a^x	4276	5 [¶]	1	0.12 (0.05-0.3)

*Modified from reference Siracusa et al., 1987.

[†]Multiplier (M). In all crosses, a presumably equivalent number of recombinants would be yellow, and thus nondetectable. In the third cross, however, recombinants can be generated by both parents.

[‡]Assuming exceptions were the result of recombination rather than mutation. Proof for recombination exists for the third cross where flanking markers were present. Parentheses indicate 95% confidence limits.

[§]Almost all offspring had a female A^y/a parent. In crosses predominantly a/a females x A^y/a males, made at other laboratories, no agouti exceptions were reported among 6632 offspring (Siracusa et al., 1987).

[¶]Four agouti exceptions were observed among 3784 offspring (the remainder were yellow, a^x/a^x being lethal); one A^y/A was identified among 492 tested yellow breeders.

APPENDIX B

FIGURES

Figure 1. Pulsed-field gel analysis of the chromosome 2 region containing the *Raly* and agouti loci in wild-type and lethal yellow DNAs.

(A) A PFGE blot of single and double digests of wild-type (A/A, FVB/N strain, denoted by A) and lethal yellow (A^y/A , denoted by Y) genomic DNAs was hybridized with a ^{32}P -labeled fragment of wild-type DNA (probe A in Fig. 1C and Fig. 2E) corresponding to a 5' segment of the *Raly* first intron. *Bss*HII (B), *Eag*I (E), *Ksp*I (K). DNA molecular size standards are shown at left in kb and LM indicates limiting mobility DNA. (B) The same filter was stripped and rehybridized with a ^{32}P -labeled wild-type agouti cDNA probe (see Fig. 2 in Bultman et al., 1992). (C) Long-range restriction map of the *Raly*-agouti region in wild-type and lethal yellow alleles. Shown below the restriction map are expanded views of the 5' end of the *Raly* gene (the solid box denotes the promoter and non-coding first exon) and the 3' end of the agouti gene (numbered solid boxes indicate exons). Separate scale bars are shown for the long-range restriction map and for each of the expanded regions. The expanded agouti region is shown to scale, except for the four exons, which are enlarged for clarity. The 2.4-kb *Eco*RI (R) interval shown at the 5' end of *Raly* is a CpG island containing one or more of each of the following enzyme recognition sites: *Eag*I (E), *Sma*I (S), *Ksp*I (K), *Bss*HII (B), and *Mlu*I (M). All of these enzyme sites are indicated on the long-range map, but only the *Bss*HII and *Mlu*I sites are shown in the expanded region due to space considerations. The probes used to detect the A^y deletion breakpoints (Probes A and C) are also shown. The 111-bp probe C is shown larger than scale. The *Eag*I and *Sma*I sites denoted by asterisks (E* and S*) are tentatively positioned on the map based on their occurrence in the 5' end of the *Raly* cDNA. N, *Not*I.

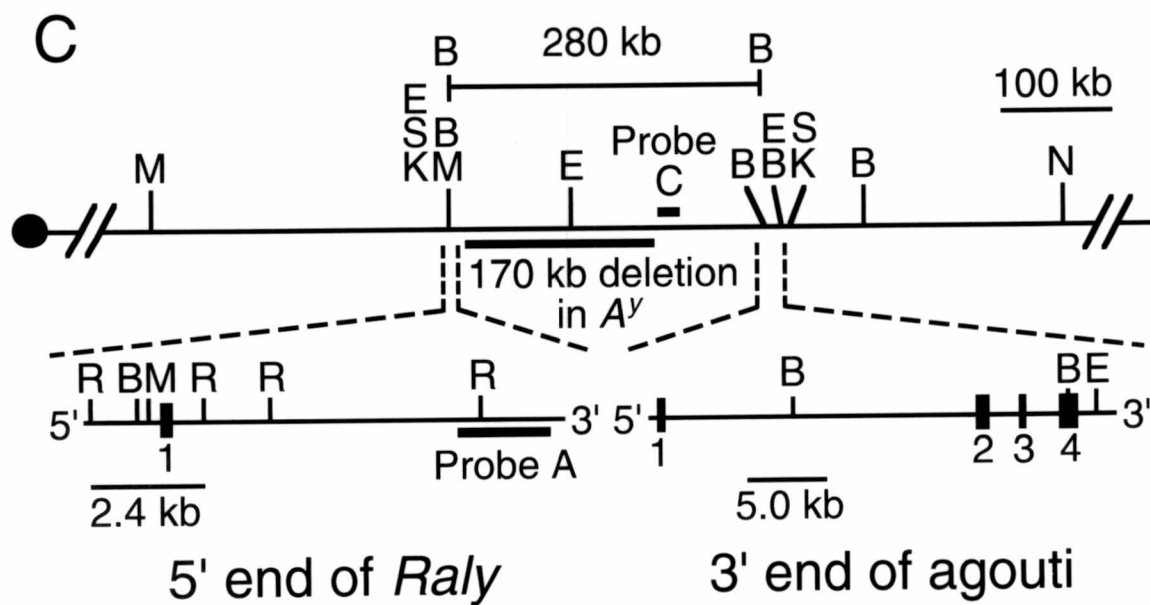
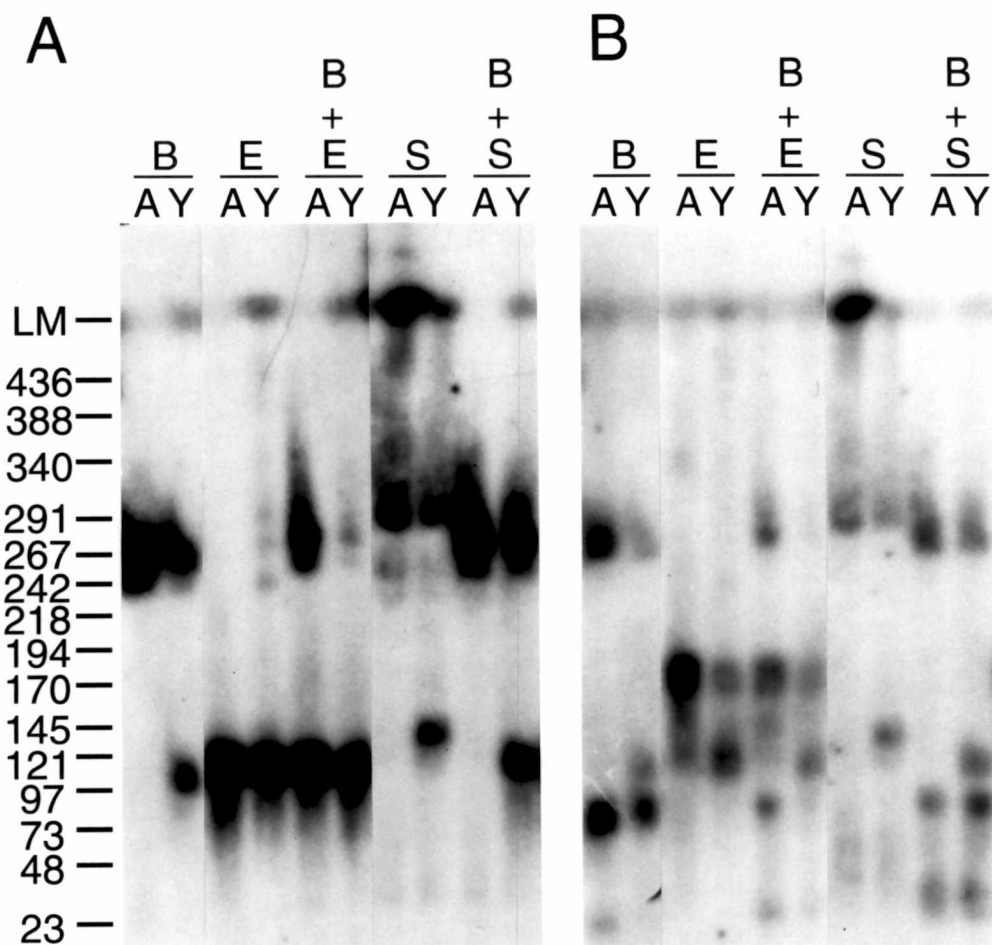


Figure 2. Detection of the A^y deletion breakpoints by Southern blot analyses.

(A) *Mus musculus* (A/A, FVB/N strain), *M. spretus* (A/A), and $A^y/M.s.$ [an F₁ hybrid from the cross *M. musculus* (A^y/a) x *M. spretus* (A/A)] genomic DNAs were digested with *EcoRI*, blotted, and hybridized with a ³²P-labeled fragment of DNA (Probe A in Fig. 2E) corresponding to a 5' segment of the *Raly* first intron. (B) The same genomic DNAs were digested with *Bgl*III, blotted, and hybridized with an 83-bp ³²P-labeled fragment of DNA corresponding to the non-coding second exon of *Raly* (Probe B in Fig. 2E). (C) The same genomic DNAs were digested with *Bam*HI or *Sac*I, blotted, and hybridized with probe A. Probe A detects A^y -specific breakpoint fragments of 14.0-kb with *Bam*HI and 13.5-kb with *Sac*I (highlighted by the boxed region). (D) The same filter used in part C was stripped and rehybridized with the 111-bp probe (Probe C in Fig. 2E). Probe C detects the same 14.0-kb *Bam*HI and 13.5-kb *Sac*I A^y -specific breakpoint fragments as did probe A, but each of these probes detects unique wild-type *M. musculus* and *M. spretus* fragments. (E) Genomic restriction map of the regions surrounding the 170-kb of DNA that is deleted in the A^y allele. The 5' deletion breakpoint in the A^y allele occurs in the first intron of *Raly*, within the 3 kb *Eco*RI-*Bam*HI interval indicated. The *Raly* second exon (and the remainder of the *Raly* gene) lies within the 170-kb region that is deleted in the A^y allele. The 111-bp segment maps to the 3' end of the 170-kb region and identifies the A^y 3' deletion breakpoint. The restriction map is shown to scale except for *Raly* exons one and two, which are enlarged for clarity. B, *Bam*HI; E, *Eco*RI; X, *Xba*I.

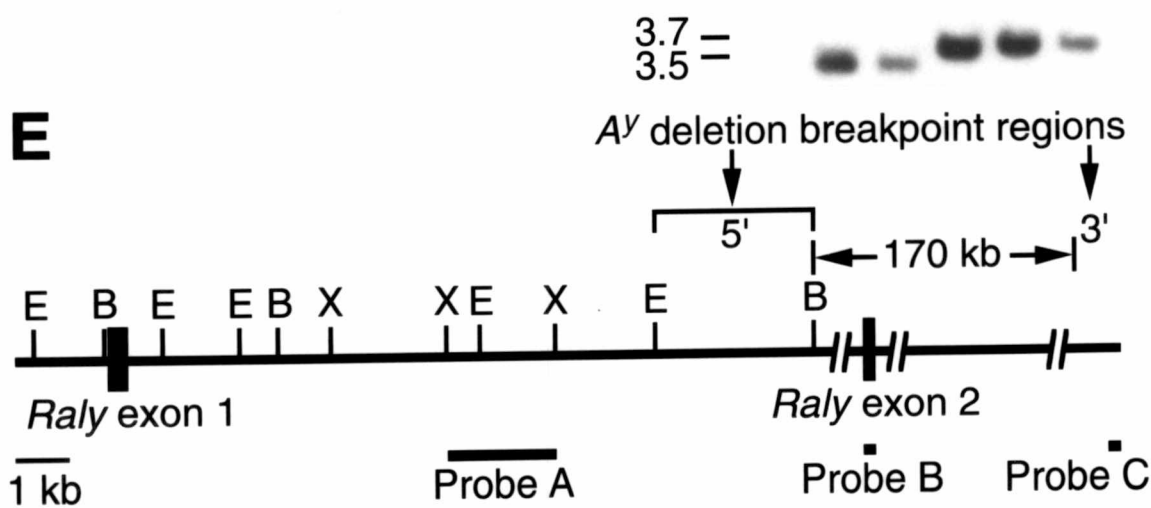
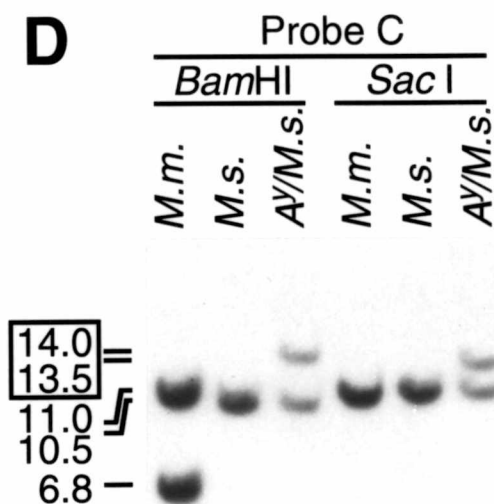
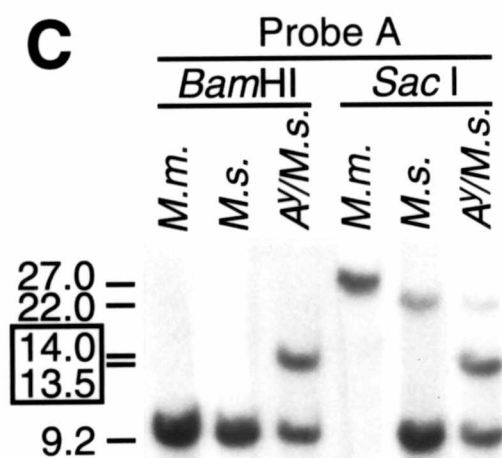
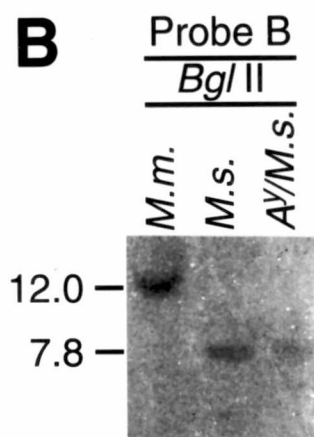
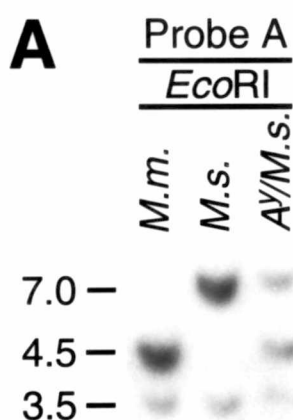
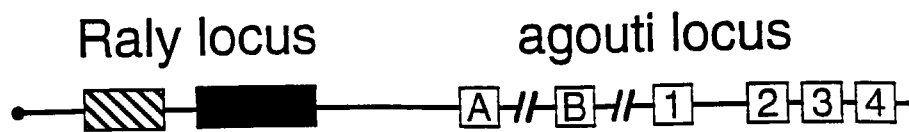


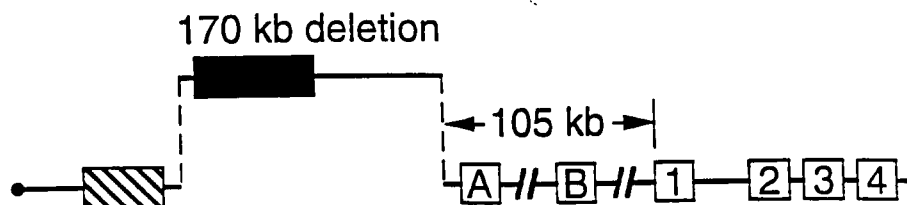
Figure 3. A simple model for recombination between A^y and other agouti-locus alleles.

The 5' end of the wild-type *Raly* gene is located about 280 kb proximal to the 3' end of the wild-type agouti gene and is transcribed in the same orientation (top panel). The wild-type *Raly* gene is shown as a cross-hatched box (indicating the promoter and untranslated first exon) and a solid box (depicting the remainder of the gene). The four exons of the wild-type agouti gene (as the gene was originally described) are shown as numbered boxes, and the recently identified upstream agouti exons are indicated by the boxes labeled A (corresponding to the 111-bp probe used to identify the 3' A^y deletion breakpoint) and B (46 bp). The A^y mutation is the result of a 170-kb deletion at the *Raly* locus, with deletion breakpoints occurring in the first intron of *Raly* and immediately upstream of agouti exon A. Approximately 105 kb of DNA remains between the 3' deletion breakpoint and the last 5' non-coding exon of the agouti gene (labeled exon 1). The agouti-locus alleles α and α^t each arose by the insertion of novel DNA sequences (shown by the vertical lines; Bultman et al., 1992) in the intron that immediately precedes the coding region of the gene. Crossing over within this 105-kb region of DNA between A^y and either α or α^t would result in the generation of a wild-type agouti allele (wild type recombination product) and the regeneration of an A^y allele (not shown) at an expected frequency of about 0.05%. This model also predicts the recombination observed between A^y and α^x , as α^x occurs in a separate complementation group from A^y , suggesting that the DNA lesion in α^x may occur 3' to the agouti gene.

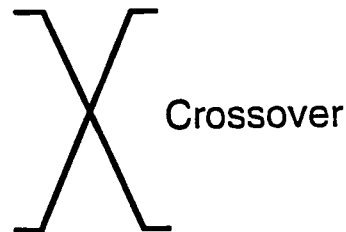
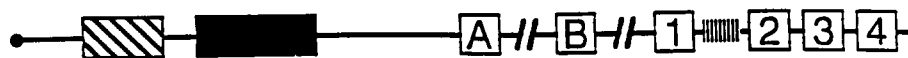
wild type



A^y allele



a or a^t allele



Result of crossover

wild type recombination product

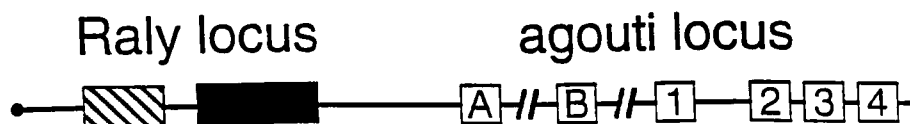
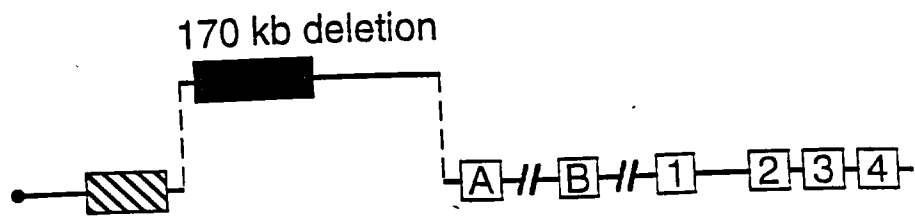


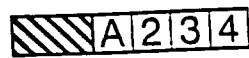
Figure 4. A model for the production of chimeric *Raly/agouti* transcripts from the *A^y* allele.

Due to the 170-kb deletion in the *A^y* allele, transcriptional initiation at the *Raly* promoter is proposed to result in the transcription of the 5' untranslated *Raly* sequence, the intergenic sequence, and the downstream agouti gene. The processing of this novel primary transcript could result in the remaining intergenic DNA being spliced out as if it were an intron, because the next available splice acceptor sites for the only remaining *Raly* splice donor would be provided by the downstream agouti gene. The result would be *A^y* transcripts consisting of the non-coding first exon of *Raly* and the three coding exons of the agouti gene (exons two, three, and four), with alternative splicing of agouti exons A and B, as previously described (Michaud et al., 1993).

A^y allele



Three A^y transcripts



PART VIII

**MOLECULAR STRUCTURE AND CHROMOSOMAL MAPPING OF
THE HUMAN HOMOLOG OF THE AGOUTI GENE**

ABSTRACT

The agouti (*a*) locus in mouse chromosome 2 normally regulates coat-color pigmentation. The agouti gene was recently cloned and shown to encode a novel 131 amino acid protein with a consensus signal peptide. Here we describe the cloning of the human homolog of the mouse agouti gene utilizing an interspecies hybridization approach. Sequence analysis revealed that the coding region of the human agouti gene is highly homologous to the mouse gene and has the potential to encode a protein of 132 amino acids that has all of the features of the mouse protein. Chromosomal assignment utilizing somatic cell hybrid mapping panels and fluorescence *in situ* hybridization (FISH) demonstrated that the human agouti gene maps to 20q11.2. This was expected, since the position on human chromosome 20 exhibits syntenic conservation with the agouti region of mouse chromosome 2. The mapping experiments revealed that human agouti is closely linked to a locus called MODY1 (Maturity Onset Diabetes of the Young). Initial expression studies with RNA from several adult human tissues showed that the agouti gene is expressed in adipose tissue and testis.

INTRODUCTION

The agouti (*a*) locus in mouse chromosome 2 normally regulates the differential production of eumelanin (black or brown) and phaeomelanin (yellow) pigment granules by the melanocytes within the hair follicle. Unlike the homogeneous black or yellow coat coloration associated with certain mouse mutants, agouti pigmentation arises from individual hair shafts containing a subapical yellow band on an otherwise black or brown background (Silvers, 1979) and is conserved in many mammalian species (Searle, 1968). We and others recently cloned the agouti gene and found that it is normally expressed in neonatal skin and encodes a small protein containing a consensus signal peptide (Bultman et al., 1992; Miller et al., 1993), indicating that the agouti protein is probably secreted.

Several dominant agouti alleles (e.g., A^y , A^{vy} , A^{sy} , A^{ty}) cause a predominantly yellow pelage (Silvers, 1979) as well as a number of additional effects which include a non-insulin-dependent diabetic-like condition, pronounced obesity, and a marked increase in the incidence of spontaneous and induced tumors (Dickerson and Gowen, 1947; Heston and Vlahakis, 1961; Silvers, 1979; Coleman, 1982; Wolff et al., 1986). Assessment of the physiology of these mutants revealed that their diabetic/obese condition has a notable similarity to non-insulin-dependent diabetes in humans (Hellerstrom and Hellman, 1963; Coleman, 1982). We recently demonstrated that several of these mutant alleles exhibit deregulated agouti gene expression caused by structural alterations within the noncoding regions of the gene (Bultman et al., 1992; Michaud et al., 1993). Based on this finding, we hypothesized that ectopic overexpression of the agouti protein in certain tissues of the animal is responsible for the dominant pleiotropic effects associated with these mutants.

The similarities between the phenotype in dominant agouti mutants and non-insulin-dependent diabetes in humans raised the question of whether a corresponding human gene could be involved in certain forms of diabetes/obesity in humans. Additionally, the association between the ectopic expression of agouti and the development of tumors in mice led to the question of whether human tumor incidence may be associated with agouti expression. To address these issues, we cloned and mapped the human agouti gene. The coding region of the human agouti gene is highly homologous to the mouse gene and maps to a region of chromosome 20 that exhibits synteny conservation with the corresponding region of mouse chromosome 2 (Siracusa and Abbott, 1991; Loffler et al., 1993). This portion of human chromosome 20 is closely linked to several loci, including a deletion complex involved in myeloid leukemia (Roulston et al., 1993) and a region associated with a locus (MODY) that has been implicated in the development of non-insulin-dependent diabetes mellitus (NIDDM) (Rothschild et al., 1993).

MATERIALS AND METHODS

Southern Blotting

Ten micrograms of genomic DNA was digested with restriction enzymes, electrophoresed through agarose gels, and blotted to Gene-Screen (DuPont) using standard procedures (Ausubel, 1988; Sambrook et al., 1989). Radiolabeled hybridization probes were prepared with the random hexamer-labeling technique (Feinberg and Vogelstein, 1984). Repetitive sequences were reassociated with sheared genomic DNA before hybridization. Posthybridization filter washes were conducted under reduced stringency conditions (0.2 x SSC, 0.1% SDS at 46°C) for experiments involving human DNA.

Isolation of a Human Genomic Clone

A *Sau3A* partial-digested human genomic lambda library in EMBL3 (Stratagene) was prepared and screened utilizing a ³²P-labeled mouse agouti full-length cDNA that had been reassociated with human sheared genomic DNA. One positive clone was purified and several restriction fragments were subcloned into the pGEM4 vector (Promega) for sequencing. Specifically, a 1.2-kilobase (kb) *Pst*I fragment, which hybridized to the mouse cDNA probe but not total mouse DNA, was subcloned to generate the sequence homologous to the first coding exon. A 4.5-kb *Eco*RI fragment, which hybridized to two overlapping 42 nucleotides (nt) oligonucleotides (5'-cactgaacaagaaatccaagaagatcagcagaaaagaagccg-3') (5'-ttggaagacctcttcgcttctcggttctttctgctgac-3') complementary to 5' end, or 3' end of mouse exon III, respectively, with 25 bp overlapping in the middle, was subcloned and used to generate the sequence homologous to the second coding exon. A 6.2-kb *Eco*RI fragment, which hybridized to probe Xbr 1.2 containing mouse exon IV, was subcloned and used to generate the sequence of the region homologous to the last agouti exon.

DNA Sequencing

Genomic subclones were sequenced by the Sanger dideoxynucleotide method (Sanger et al., 1977) using T7 DNA polymerase (U.S. Biochemical)

(Tabor and Richardson, 1987). Analysis of DNA sequence was performed using the University of Wisconsin Genetics Computing Group sequence analysis programs (Devereux et al., 1984).

Chromosomal Assignment

For chromosomal assignment of the human agouti gene, polymerase chain reactions (PCR) with 24 DNA samples from the NIGMS human rodent somatic cell hybrid mapping panel 2 (Drwinga et al., 1993), and 3 samples from the human (NAIMR91), mouse (NA05862), and chinese hamster (NA106588) parental cell line, respectively, were performed. The oligonucleotide (5'-CCTCTTACCATTACCCCTGA-3'), which corresponds to flanking intron sequence 35 bp upstream of the 5' end of the putative exon II and the oligonucleotide (5'-CTAGGTGACTTACCCACAAT-3'), which corresponds to the 3' end of the putative exon II were used as primers A and B, respectively. PCR was carried out in 50 µl volumes containing 100 ng template DNA, 15 pM of primer A and B, 50 mM KCl, 10 mM Tris-Cl, 15 mM MgCl₂, 0.01% (w/v) gelatin, 250 mM of each dNTP and 1.4 U Taq polymerase. Initial denaturation was at 95°C for 4 min, followed by 32 cycles, each consisting of 1 min at 94°C, 1 min at 55°C, and 2 min at 72°C. Ethidium bromide stained PCR fragments were visualized after separation on a nondenaturing PAGE.

Mapping by Fluorescence *In Situ* Hybridization (FISH)

Prometaphase chromosomes for FISH analysis were obtained from cultured lymphocytes following a slightly modified version of the method described by Rybak *et al.* (1982). For regional mapping of the human agouti gene, the 17-kb genomic probe h20B1 was labeled with biotin-16-dUTP (Boehringer, Mannheim) by nick translation. Labeling, hybridization and signal detection was carried out as described (Lichter et al., 1988; Rao et al., 1992). The signal was amplified twice using biotinylated goat-anti-avidin immunoglobulins (Vector). Slides were counterstained with ethidium bromide or DAPI for band identification and were mounted in anti-fade solution containing 90% glycerol, 10% PBS, and 1 mg/ml phenylene diamine. A Zeiss axiophot fluorescence microscope (63X, filter

LP420 and LP520) was used for analysis, and representative metaphases were photographed on Scotch T640 film. The map position of the agouti locus was determined by the assignment of FISH signals to G-bands and also by length measurements and its expression in terms of fractional length from pter values (FL-pter) according to Lichter *et al.* (1990).

RNA Isolation and RT-PCR

Poly (A)⁺ RNA from human adipose tissue was isolated using the FastTrack mRNA isolation kit with guanidinium-based lysis buffer (Invitrogen). Poly (A)⁺ RNA for the remaining tissues was purchased from Clontech. Two hundred ng of Poly (A)⁺ RNA was reverse transcribed to first strand cDNA using random-hexamer primers and avian myeloblastosis virus reverse transcriptase (Invitrogen). One twentieth of the reaction mixture was then subjected to specific PCR amplification as follows. PCR was carried out for 35 cycles (95°C, 1 min; 65°C, 1 min; 72°C, 2min) in 20% glycerol with Taq I polymerase and 20 pM 5' primer (5'-ATGGATGTCACCCGCTTACTCCTGGCC-3') and 3' primer (5'-GCGCTCAGCAGTTGAGGCTGAGCACGC-3'), respectively. Amplified PCR products were subjected to electrophoresis on 1% agarose gel.

RESULTS

Isolation and Characterization of a Human Clone Homologous to the Mouse Agouti Gene

To determine whether the agouti gene has been conserved in the human genome, the mouse cDNA clone was hybridized to Southern blots containing human genomic DNA. Reduced stringency hybridization conditions detected cross-hybridizing DNA fragments with a variety of restriction enzymes (data not shown). Based on this result, the same mouse agouti cDNA probe was used to screen a human genomic library. This led to the isolation of a human genomic clone called h20B1. Restriction mapping of this clone revealed that it contains an insert of about 17 kb.

Characterization of h20B1 revealed the presence of three distinct segments that each cross-hybridize with one of the three coding exons of the

mouse agouti gene (Fig. 1). These cross-hybridizing regions were subcloned and subjected to nucleotide sequence analysis utilizing primers derived from the sequence of the individual mouse exons. This analysis revealed that each cross-hybridizing region of h20B1 contains sequence that is highly homologous with one of the three coding exons of the mouse agouti gene. Additionally, each conserved region is flanked by consensus splice donor and acceptor sites, except for the last region which, as expected, was flanked only by a splice acceptor at its 5' end. A polypyrimidine tract and putative branch point signal 10-50 bp upstream of each conserved 3' splice acceptor site was also identified (data not shown).

Most notably, the open reading frame (ORF) defined by the mouse agouti gene is retained within the individual blocks of conserved sequences in h20B1 (Fig. 1). In fact, juxtaposition of these conserved regions through the putative splice donor and acceptor sequences gives rise to an ORF of 396 bp. This ORF has the potential to encode a protein of 132 amino acids, which is very similar to the predicted 131 amino acid protein for the mouse. The ORF begins with a translation initiation codon (ATG), and ends with a stop codon (TGA) at a position nearly identical to that of the mouse (Fig. 1). In the region corresponding to exon IV, a single polyadenylation signal (AATATA) is located at a position similar to where the AATAAT found in the mouse.

Comparison of the mouse and human regions indicates that the ORF sequences are 85% identical at the nucleotide sequence level (Fig. 1). In contrast, the regions corresponding to the 5' and 3' untranslated portions of have diverged considerably (Fig. 1). The human sequence has a 3-bp deletion within the region corresponding to exon III, and a 6-bp insertion in the coding region of sequences corresponding to exon IV. Both of the sequence differences are in increments of 3 bp, which preserves the evolutionarily conserved ORF. Overall, comparison of the deduced amino acid sequences shows that the predicted human protein is 80% identical to that in the mouse (85% similarity-conserved changes) (Fig. 2). Like the mouse protein, the amino terminus of the putative human protein contains all of the features of a signal peptide, including a consensus cleavage site.

The highly basic domain and the cysteine-rich domain near the carboxy-terminus are also conserved.

Mapping and Regional Assignment of the Human Agouti Gene

Several genes linked to agouti in mouse chromosome 2 map to the long arm of human chromosome 20 (Siracusa et al., 1991; Loffler et al., 1993). In an attempt to determine whether h20B1 also maps to human chromosome 20, PCR was performed on 24 DNA samples from the NIGMS human-rodent somatic cell hybrid panel 2 (Drwinga et al., 1993). Primers for this experiment were from within, and flanked the putative exon II, and as expected, gave rise to a 238-bp amplified fragment utilizing, as a template, human genomic DNA or DNA from the parental human cell line NAIMR91 (data not shown). Analysis of the somatic cell hybrid panel 2 revealed that the agouti-specific 238-bp fragment was detected only in cell line NA10478 (data not shown), which is known to carry only human chromosome 20 in 76% of its cells and only chromosome 4 in another 12% of its cells (Drwinga et al., 1993). No specific products were amplified from the parental mouse line (NA05862), from the chinese hamster cell line (NA10658), or from any of the other cell hybrids, particularly hybrid NA 10115, which carries only human chromosome 4. These data unequivocally map h20B1 to human chromosome 20.

For regional localization of human agouti, FISH was performed using h20B1 as a probe. Hybridization was seen in the majority of metaphases in the proximal long arm of chromosome 20 (Fig. 3). Specific FISH signals were not observed on any other chromosome. According to the banding pattern of DAPI-stained metaphase chromosomes, the FISH signals for the agouti gene could be assigned to band 20q11.2. A few individual signals were detected on chromatids at other chromosomal sites in some metaphases, although their pattern was inconsistent; therefore, we attributed these signals to nonspecific background hybridization.

The position of h20B1 on the physical FISH map of chromosome 20 (Fig. 3) was determined by length measurements from 13 selected chromosomes 20, and the results from this analysis was expressed in terms of fractional-length-pter (FLpter). The mean FLpter-value was determined

to be 0.55 ± 0.05 , which corroborates our localization to band 20q11.2 (Fig. 3A). This assignment was also recently confirmed by FISH mapping of a 450-kb YAC carrying the sequences hybridizing to h20B1. For this purpose, length measurements of 11 selected chromosomes 20 hybridized with this YAC revealed an FLpter-value of 0.55 ± 0.04 , which is identical to that obtained with the h20B1 clone (data not shown).

Expression of the Agouti Gene in Human Adipose Tissue

In an initial attempt to evaluate expression of the human homolog of the agouti gene, we utilized RT-PCR (Reverse Transcriptase-Polymerase Chain Reaction) to assay for the presence of agouti transcripts in a variety of adult human tissues, including three tissues that have been implicated in NIDDM (muscle, liver and fat). As shown in Fig. 4, the expected 403-bp agouti-specific fragment was detected only in adipose tissue and the testis. To confirm that this 403-bp fragment actually arose from the corresponding human mRNA, the PCR fragment was purified, subjected to nucleotide sequence analysis and found to be identical to the mouse-human homologous regions in the h20B1 clone (data not shown). This indicated that the homologous sequences on h20B1 clone actually correspond to exons of a functional gene that is expressed in selected human tissues.

DISCUSSION

Utilizing a cross-species hybridization approach, we have isolated a human genomic clone that contains the three coding exons of the agouti gene. The human gene produces a mRNA that is predicted to encode a protein that is highly homologous and nearly identical in size to that of the mouse gene. Based on these findings, coupled with the mapping data, we propose that the sequences in h20B1 correspond to the coding exons of a human homolog of the agouti gene.

FISH mapping experiments described in this report indicate that human agouti is closely linked to the MODY1 locus (Fig. 3). MODY1 is a form of NIDDM that has an early age of onset and an autosomal dominant mode of inheritance (Fajans et al., 1992). Based on the fact that the

dominant yellow agouti mutations in the mouse exhibit a non-insulin dependent diabetic-like condition similar to that observed in MODY1 patients, we evaluated whether agouti was a candidate gene for MODY1 based on the close linkage between the two. Analysis of our comprehensive FISH data revealed that agouti maps to band 20q11.2, which is the region where the loci HCK1 and RPN2 have been mapped (Fig. 3). All three of these loci, HCK1-agouti-RPN2, are located slightly more proximal to the 13 cM sex-averaged interval encompassing ADA, D20S17, PPGB, D20S16 and D20S75, which is the region where MODY1 has been mapped (Rothschild et al., 1993). Therefore, taking into account only the genetic linkage data of MODY1 and our physical assignment of agouti to a slightly more proximal 20q segment (Fig. 3B), we would conclude that it appears that agouti is unrelated to MODY1. However, since the MODY1 locus mapping to chromosome 20 is based solely on a single family called RW (Rothschild et al., 1993), it is possible that a unique chromosomal structural alteration, such as an inversion, occurred in this family, and that this feature of the mutation affects genetic recombination in a manner that would interfere with its accurate mapping on a physical map of the region. This being the case, it is still possible that human agouti is associated with MODY1 in the RW family. Further experiments utilizing microsatellite repeats, genetic mapping of the two, and mutation analyses will help to resolve this issue.

With respect to whether the human agouti gene is involved in tumor formation in humans, Roulston et al. (1993) recently reported that a region of human chromosome 20q is commonly deleted in patients with myeloid leukemia. By cytogenetic and FISH analysis, a commonly deleted interstitial segment (SRO) mapping to 20q11.2-q12 was identified which contains the ADA and SRC loci and is flanked proximally by the ribophorin 2 gene locus (RPN2) and distally by D20S17 (Roulston et al., 1993). Our FISH results place the agouti gene close to the proximal boundary of this commonly deleted segment (Roulston et al., 1993). Whether agouti maps within the deleted region is presently unclear, although, in the event that agouti occurs adjacent to the deleted region, then a similar situation would exist to that observed in the dominant mouse mutation, lethal yellow (Bultman et al., 1992; Michaud et al., 1993). Lethal yellow animals have an

increased susceptibility to the development of both spontaneous and induced tumors. We recently determined that a deletion in the region upstream of the agouti coding exons causes an activation of the gene and is responsible for the pleiotropic effects (Michaud et al., 1993). Therefore, it could be that the human agouti gene is activated by deletion of sequences adjacent to the coding region of the gene, and that this may be related to the development of human myeloid leukemia.

Although the agouti pattern of pigmentation is evolutionarily conserved throughout many mammalian species, there has never been a published report describing human hair with banded, agouti-like pigmentation. Therefore, the question arises as to what role the agouti gene may have on pigmentation in humans. The very strong sequence conservation within the coding exons of the gene would argue that the agouti gene has some functional role in humans. One possibility is that agouti has an identical function in humans and mice, and that this includes signaling the melanocyte to produce phaeomelanosomes. In this case, the lack of banded agouti hair in humans could be related to the fact that humans have a continuous hair growth cycle, which is different from the discontinuous molting hair growth cycles exhibited by various mammalian species that display agouti pigmentation. Alternatively, it is possible that the human and mouse agouti genes have evolved different regulatory mechanisms for expression. Experiments currently under way to identify and characterize the promoter and untranslated regions of the human gene will address this matter.

Finally, based on our limited expression analysis of the human agouti gene, the production of agouti transcripts in adipose tissue from presumably normal individuals is provocative. In the mouse, it appears that agouti and other pigmentation genes exert their effects by regulating the level of cAMP within the melanocyte (Takeuchi et al., 1989). If agouti regulates the cAMP level in the adipocyte in a similar manner, then the net effect of adipose-specific expression of agouti would be decreased cAMP levels within the adipocyte. In this case, lower levels of cAMP would lead to decreased lipolysis due to inactivation of the hormone sensitive lipase, which would contribute to adipocyte hypertrophy and obesity (Stralfors et al., 1984). The

resulting obesity may lead to insulin resistance (Bonadonna et al., 1992), and possibly non-insulin dependent diabetes (Reaven, 1988). Therefore, in humans, it may be that agouti has some functional role in the regulation of lipid metabolism within the adipocyte. Further experiments are currently underway to study the expression of the agouti gene in adipose tissue and to define its role, if any, in the development of obesity and non-insulin-dependent diabetes in humans.

REFERENCES

- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (1988). *Current Protocols in Molecular Biology*. John Wiley & Sons/Greene, New York.
- Bonadonna, R. C. and DeFronzo, R. A. (1992). in *Obesity*. (eds. Björntorp, P., and Brodoff, B. N.), J. B. Lippincott, Philadelphia.
- Bultman, S. J., Michaud, E. J., and Woychik, R. P.. (1992). Molecular characterization of the mouse agouti locus. *Cell* **71**: 1195-1204.
- Coleman, D. L. (1982). Diabetes-obesity syndromes. In *The mouse in biomedical research IV* (eds. H. L. Foster, J. D., Small, and J. G., Fox), pp. 125-132. Academic Press, New York.
- Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**: 387-395.
- Dickerson, G. E., and Gowen, J. W. (1947). Hereditary obesity and efficient food utilization in mice. *Science* **105**: 496-498.
- Drwinga, H. L., Toji, L. M., Kim, C. H., Greene, A. E. and Mulivor, R. A. (1993). NIGMS human/rodent somatic cell hybrid mapping panels 1 and 2. *Genomics* **16**: 311-314.
- Fajans, S. S., Bell, G. I. and Bowden, D. W. (1992). *J. Lab. Clin. Med.* **119**: 206-210
- Feinberg, A. P., and Vogelstein, B. (1984). Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**: 266-267.
- Hellerstrom, C., and Hellman, B. (1963). The islets of Langerhans in yellow obese mice. *Metabolism* **12**: 527-536.

- Heston, W. E., and Vlahakis, G. (1961). Influence of the A^y gene on mammary-gland tumors, hepatomas, and normal growth in mice. *J. Natl. Cancer Inst.* **26**: 969-983.
- Lichter, P., Cremer, T., Borden, J., Manuelidis, L., and Ward, D. C. (1988). Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. *Hum. Genet.* **80**: 224-234.
- Lichter, P., Tang, C. C., Call, K., Hermanson, G., Evans, A., Housman, D and Ward, D. C. (1990). *Science* **247**: 64-69.
- Löffler, C., Rao, V. V. N. G., Schnittger, S., Pastuszak, D., Schubert, S., and Hansmann, I. (1993). A first physical FISH map of human chromosome 20 disclose complete homology with mouse chromosome 2 and assigns candidatemutations. *Am. J. Hum. Genet.* **53S**: 1324.
- Miller, M. M., Duhl, D. M. J., Vrieling, H., Cordes, S. P., Ollmann, M. M., Winkes, B. M., and Barsh, G. S. (1993). Cloning of the mouse agouti gene predicts a secreted protein ubiquitously expressed in mice carrying the lethal yellow mutation. *Genes & Dev.* **7**: 454-467.
- Michaud, E. J., Bultman, S. J., Stubbs, L. J., and Woychik, R. P. (1993). The embryonic lethality of homozygous lethal yellow mice (A^y/A^y) is associated with the disruption of a novel RNA-binding protein. *Genes & Dev.* **7**: 1203-1213.
- Rao, V. V. N. G., Löffler, C., Wozney, J. M. and Hansmann, I. (1992). The gene for bone morphogenetic protein 2A (BMP2A) is localized to human chromosome 20p12 by radioactive and nonradioactive in situ hybridization. *Hum. Genet.* **90**: 299-302.
- Reaven, G. M. (1988). *Diabetes* **37**: 1595-1607

- Rothschild, C. B., Akots, G., Hayworth, R., Pettenati, M. J., Nagesh Rao, P., Wood, P., Stolz, F-M., Hansmann, I., Serino, K., Keith, T. P., Fajans, S. S. and Bowden, D. W. (1993). A genetic map of chromosome 20q12-13.1: Multiple highly polymorphic microsatellite and RFLP markers linked to the maturity-onset diabetes of the young (MODY) locus. *Am. J. Hum. Genet.* **52**: 110-123
- Roulston, D., Espinosa III, R., Stoffel, M., and Le Beau, M. M. (1993). Identification of the commonly deleted segment in myeloid leukemias with a del(20q). *Am. J. Hum. Genet.* **53S**: 352.
- Rybak, J., Tharapel, A., Robinett, S., Garcia, M., Maniken, C. and Freeman, M. (1982). A simple reproducible method for prometaphase chromosome analysis. *Hum. Genet.* **60**: 328-333
- Sambrook, J., Fritsch, E.F., Maniatis, T., (1989) *Molecular cloning: A laboratory manual, second edition* . Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
- Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467.
- Searle, A. G. (1968). *Comparative genetics of coat color in mammals*. Academic press, New York.
- Silvers, W. K. (1958a). An experimental approach to action of genes at the agouti locus in the mouse. II. Transplants of newborn *aa* ventral skin to *a^ta*, *A^wa*, and *aa* hosts. *J. Exp. Zool.* **137**: 181-188.
- Silvers, W. K. (1958b). An experimental approach to action of genes at the agouti locus in the mouse. III. Transplants of newborn *A^w-*, *A-*, *a^t-*skin to *A^y-*, *A^w-*, *A-* and *aa* hosts. *J. Exp. Zool.* **137**: 189-196.

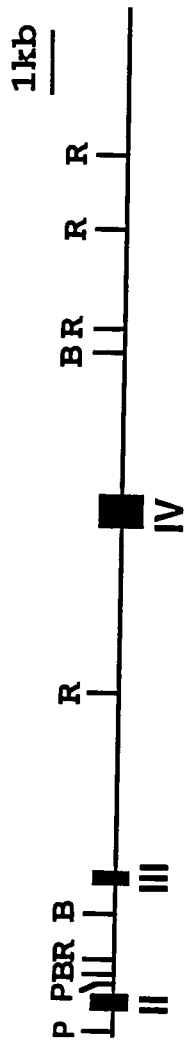
- Silvers, W. K. (1961). Genes and pigment cells of mammals. *Science* **134**: 368-373.
- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6-44. Springer-Verlag, New York.
- Silvers, W. K., and Russell, E. S. (1955). An experimental approach to action of genes at the agouti locus in the mouse. *J. Exp. Zool.* **130**: 199-220.
- Siracusa, L. D., and Abbott, C. M. (1991). Mouse chromosome 2. *Mammalian Genome* **1**: S18-S41.
- Strålfors, P., Björgell, P., and Belfrage, P. (1984). Hormonal regulation of hormone-sensitive lipase in intact adipocytes: Identification of phosphorylated sites and effects on the phosphorylation by lipolytic hormones and insulin. *Proc. Natl. Acad. Sci. USA*. **81**: 3317-3321
- Tabor, S., and Richardson, C. C. (1987). DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. USA* **84**: 4767-4771.
- Takeuchi, T., Kobunai, T., and Yamamoto, H. (1989). Genetic control of signal transduction in mouse melanocytes. *J. of Invest. Dermatol.* **92**: 239S-242S.
- Wolff, G. L., Roberts, D. W., and Galbraith, D. B. (1986). Prenatal determination of obesity, tumor susceptibility, and coat color pattern in viable yellow (A^{vy}/a) mice. *J. Hered.* **77**: 151-158.

APPENDIX**FIGURES**

Figure 1. Genomic structure and nucleotide sequence of the human agouti gene.

(A) Genomic structure of the human agouti gene in clone h20B1. The three coding exons are represented by black boxes and Roman numerals. The restriction enzyme sites utilized to subclone the individual exons are indicated. P, *Pst*I ; Bg, *Bgl*II; R, *Eco*RI; B, *Bam*HI. (B) Nucleotide sequence comparison of the mouse and human exons of the agouti gene. The mouse sequence is shown above, and the human sequence is shown below. Identical nucleotides are noted as dots and differences are indicated. Lack of nucleotides are indicated as dashes. The exon sequence is separated from the 5' splice donors and 3' splice acceptors vertical lines. The proposed translational initiation codons and stop codons are underscored by single lines, and the potential polyadenylation signals are underscored by double lines.

A



B

II

ttag		gcttctcaggatgtcaccgcctactcctctggccaccctagtgagcttcctgtgcttcttcaccgtccacagccacctggca
.c..		..c.ctg.....t.....gt.....c.....t.c.a.....c..
ctcaggagacgcttgagatgacaggagtctcggagtaactcctccatgaactcgtgatttctcctctgttctatcgtgg		gtaa
.ct.....a...cc.....c...a.a.c.....tg.....cta.....g.c.t.....c.....t....		

III

gaag		cactgaacaagaatccaagaagatcagcagaaaaaagcggagaagcggaaagggtcttccaag
....		.g.....ac.....g.....c...a...a.....a.a.....t... ..
		gtaa
		

IV

acag		aaaaaggcttcgatgaagaagggtggcaagccccgccaccttcg-----ccctgcgtggccaccgcgacagctgcaagccac
g...		..gg.....a....tgc.....g.a.c.cctatctgcg.....a.....g.
ccgcacccgcctgctgcgacccgtgcgcctcctcctgcagtgccgttcttcggcagcgccctgcacctgtcgagtactcaaccccaactgctga		.g.....c.....t....c..g...g.t.....
cgcagcttcttcgctgcgcgcagcttcgggaacgggtgattggcgggggttcagggtcccgcgcttcttaggctgagggcggtctctcg		gcgcc.ccac..c.g..cgcgagcaggca...cttc..g.cgc..g.c.ct.ctc...cgggtgatcc...acag.gc..cttccagggc
tgggtggggccttggtggcggtggtcagtggtgtgactgtggcgctttcaaaaaacccggttttctaggaaacctagtggaaagctaaa		..cag.c..gcg.a..t.ccag.a.a.agg.act.cag.gagacct...ttggg.t...t.gaaa.aaaaat.t.taggc..ctc.aaggt
atcagaatacaataataatttttaggtgcc		mouse
g.gc.gc.gttttctg...aagg.cccgaaag		human

Figure 2. Amino acid sequence comparison between the mouse and human agouti genes.

Shown above is the mouse sequence as previously reported (Bultman et al., 1992). Shown below is the human sequence as predicted from the ORF created by the juxtaposition of regions homologous to the individual agouti exons shown in Fig. 1. Identities are represented by dots, and differences are indicated. Lack of amino acids are represented by dashes.

Met Asp Val Thr Arg Leu Leu Leu Ala Thr Leu Val
. Leu

Ser Phe Leu Cys Phe Phe Thr Val His Ser His Leu
Val Ala Asn . . .

Ala Leu Glu Glu Thr Leu Gly Asp Asp Arg Ser Leu
Pro Pro . . Lys . Arg

Arg Ser Asn Ser Ser Met Asn Ser Leu Asp Phe Ser
. Val . Leu . . Val Pro

Ser Val Ser Ile Val Ala Leu Asn Lys Lys Ser Lys
.

Lys Ile Ser Arg Lys Glu Ala Glu Lys Arg Lys Arg
Gln . Gly . . Ala . . . - . .

Ser Ser Lys Lys Lys Ala Ser Met Lys Lys Val Ala
. . . . Glu Val

Arg Pro Pro Pro Pro Ser - - Pro Cys Val Ala
. . Arg Thr . Leu Ser Ala . . .

Thr Arg Asp Ser Cys Lys Pro Pro Ala Pro Ala Cys
. . Asn

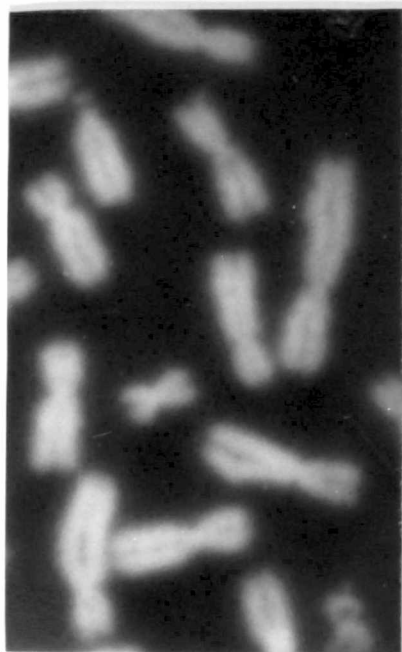
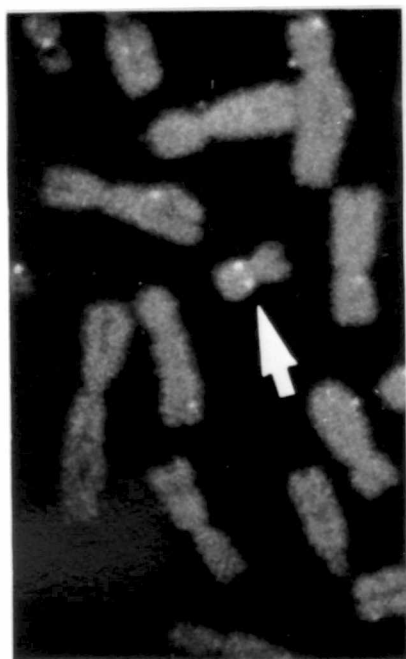
Cys Asp Pro Cys Ala Ser Cys Gln Cys Arg Phe Phe
.

Gly Ser Ala Cys Thr Cys Arg Val Leu Asn Pro Asn
Arg . . . Ser . . . Ser Leu .

Cys mouse
. human

Figure 3. Physical mapping of the human agouti gene to chromosome 20q11.2.

(A) Assignment of the agouti gene locus to human chromosome 20q11.2. Two partial metaphases (top, bottom) showing specific double signals by FISH at 20q11.2 (arrow) after hybridization with h20B1 (left), and the corresponding DAPI-fluorescence pattern of the same metaphases (right). (B) Localization of the agouti gene within the physical FISH map of human chromosome 20 (HSA 20) demonstrating conservation of synteny within the respective segment on the composite linkage map of mouse chromosome 2 (MMU2).



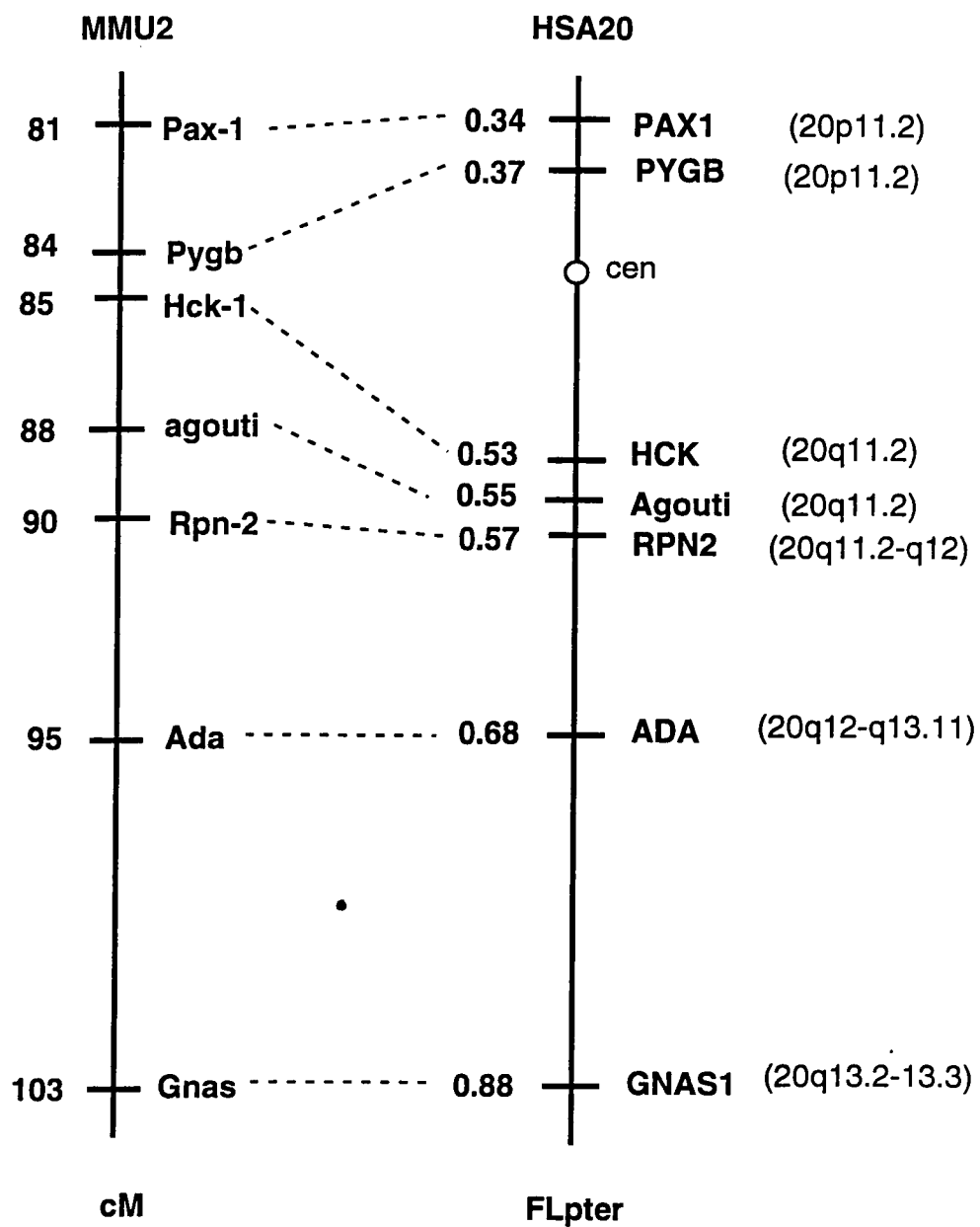
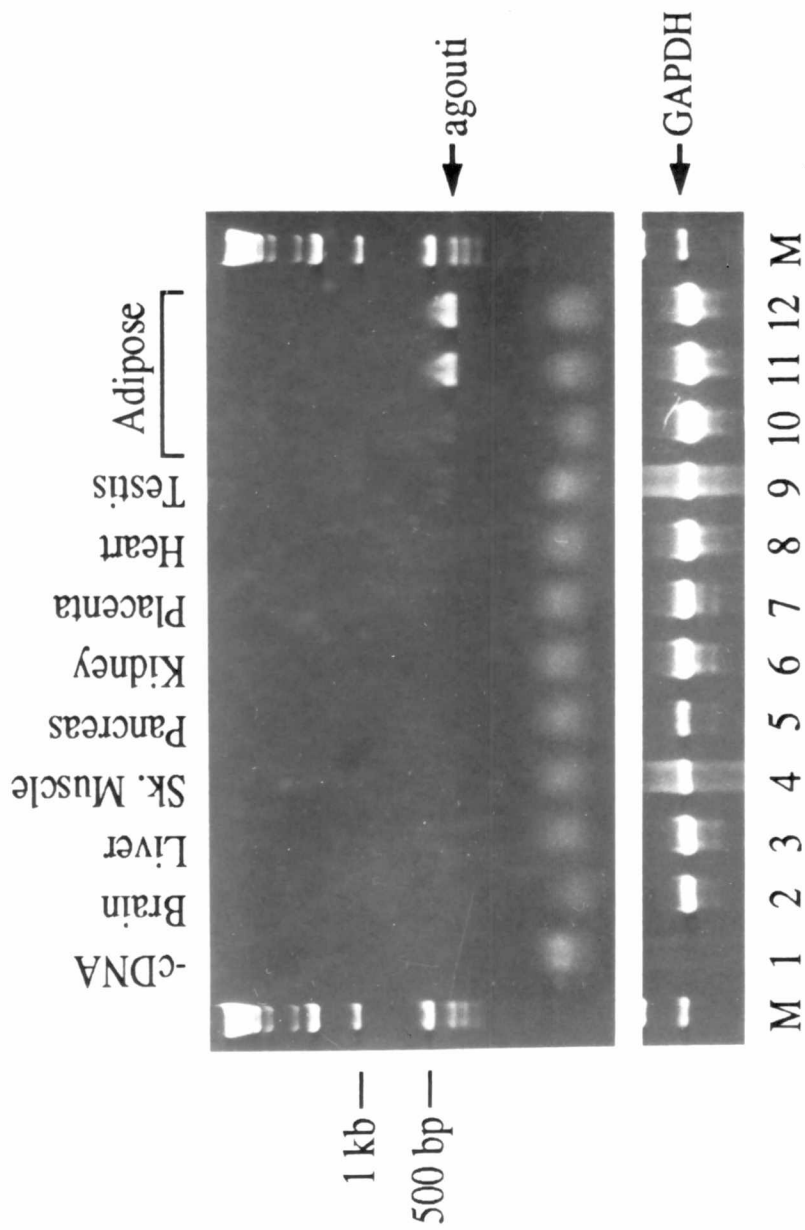


Figure 4. Tissue distribution of the human agouti mRNA.

RT-PCR was used to determine tissue distribution of human agouti RNA. Top panel shows the amplified PCR products using agouti-specific primers and cDNA template reverse transcribed from poly (A)⁺ RNA of various tissues. Lane 1, no cDNA template in PCR reaction; lanes 2-9, RNA from brain, liver, skeletal muscle, pancreas, kidney, placenta, heart, and testis, respectively; lanes 10-12, RNA from adipose tissue of three different individuals. Bottom panel shows the amplified PCR products using human GAPDH (glyceraldehyde-3-phosphate dehydrogenase)-specific primers as an internal control.



PART IX
SUMMARY

CONCLUSIONS

Gaining Molecular Access to a Region of DNA at the Agouti Locus

We exploited a radiation-induced mutation, called Is1Gso, to gain molecular access to a region of DNA at the agouti locus. In Is1Gso, new mutant alleles at both the agouti and limb deformity (*ld*) loci arose simultaneously (see part II; Woychik et al., 1990a). Both of these loci map to mouse chromosome 2 but are separated by 20 centimorgans (cM) (Davisson and Roderick, 1978; Green, 1989; Siracusa and Abbott, 1991). Cytogenetic examination of Is1Gso revealed a complex chromosomal rearrangement involving a large inversion of chromosome 2 with DNA breakpoints at or near the agouti and *ld* loci (see part II; Woychik et al., 1990a). To determine whether the distal inversion breakpoint lies near the agouti locus, we utilized an *ld* DNA probe derived from the *ld^{Hd}* insertional mutant (Woychik et al., 1985; Maas et al., 1990; Woychik et al., 1990b) to detect and clone a size-altered restriction fragment in Is1Gso (see part II; Woychik et al., 1990a). A DNA probe, corresponding to the breakpoint sequence of the size-altered fragment, was subsequently shown to normally map at or near the agouti locus using *Mus. spretus*-*Mus. musculus* backcross analysis (see part II; Woychik et al., 1990a). This indicated that the distal inversion breakpoint is located at or near the agouti locus, and that by cloning a DNA fragment that contains the inversion breakpoint we were able to "jump" 20 cM from *ld* to a region very close to and possibly within the agouti locus.

In an attempt to provide evidence that the region of DNA at the distal inversion breakpoint of Is1Gso is, in fact, associated with a gene at the agouti locus (as opposed to only being near the agouti locus), we analyzed a number of agent-induced mutations for DNA structural alterations. Four homozygous-lethal α^e mutations (*aj^{l10}*, *aj^{l41}*, *aj^{l85}*, *α^{9H}*) and a viable α^e mutation (*α^{5MNU}*) were shown to carry deletions in this region by Southern blot analyses (see part III; Bultman et al., 1991). Although the homozygous-lethal mutations are presumably large deletions that encompass genes unrelated to the agouti locus, the *α^{5MNU}* mutation was cloned and shown to consist of a single 2.8-kb deletion (see part III; Bultman et al., 1991).

Identification of the α^{5MNU} mutation was significant because a relatively subtle DNA lesion of this type is likely to be intragenic and affect only the expression of the gene directly associated with the phenotype of interest.

Cloning the Agouti Gene

Having established that the distal inversion breakpoint of *IslGso* possibly lies within a gene associated with agouti-locus function, a search was initiated to find an exon-containing DNA probe. Interspecies hybridizations (i.e., "zoo blots") identified a probe in the immediate vicinity of the inversion breakpoint that is conserved in genomic DNA from several mammalian species (see part IV; Bultman et al., 1992). Furthermore, this probe was shown to contain an exon based on its ability to hybridize to a 0.8-kb transcript in RNA prepared from neonatal skin (see part IV; Bultman et al., 1992). Screening neonatal skin cDNA libraries with this probe led to the isolation of several clones that contain a complete open reading frame and which were nearly full length (they only lack a short segment of untranslated sequence at the 5' end).

Based on the sequence of the nearly full-length cDNA clones, the agouti gene is predicted to encode a small protein 131 amino acids in length and approximately 15,000 daltons in molecular weight (see part IV; Bultman et al., 1992). Although the predicted protein does not exhibit significant sequence homology with any previously characterized proteins, it does contain several interesting features, including a putative signal peptide at the N-terminus, a highly basic domain in the center of the putative protein (16 lysine or arginine residues in a stretch of 29 amino acids), and a cysteine-rich C-terminus (10 of the final 40 amino acids are cysteines) (see part IV; Bultman et al., 1992).

To elucidate the genomic structure of the gene, overlapping genomic clones were isolated and characterized using the cDNA clone as a probe. Comparison of the structure of genomic and cDNA clones demonstrated that the wild-type (*A*) allele has four exons, the last three of which are coding (see part IV; Bultman et al., 1992). Subsequent molecular analysis of the *A^w* and *a^t* alleles, however, revealed the presence of alternative 5' noncoding exons approximately 100-kb upstream from the rest of the gene (see parts

V and VII; Bultman et al., 1994; Michaud et al., 1994). The agouti gene is therefore rather large and quite complex.

Expression of the Agouti Gene

Consistent with the genetics of the agouti locus (i.e., dominant gain-of-function agouti mutants are yellow and recessive loss-of-function mutants are black), agouti gene expression is positively correlated with the production of phaeomelanin. It is expressed in wild-type (*A/A*) skin throughout early postnatal development, but the level of expression is greatest between days 3 and 7 when the band of phaeomelanin is being deposited into agouti hair shafts (see part IV; Bultman et al., 1992). Expression is even more clearly associated with the production of phaeomelanin in agouti-locus mutants. In black-and-tan mice, for instance, agouti mRNA is absent in skin derived from the black, dorsal surface but overexpressed in skin derived from the yellow, ventral surface of the animal (see parts IV and V; Bultman et al., 1992, 1994).

The agouti gene is also expressed in neonatal skin from *W/W^v* mice (which lack hair-bulb melanocytes) indicating that it is expressed in the microenvironment of the hair follicle and not in the melanocyte (see part IV; Bultman et al., 1992). This is consonant with skin grafting experiments performed in the 1950s which demonstrated that the agouti locus functions from within the follicular environment and not from within the melanocytes (Silvers and Russell, 1955; Silvers, 1958a, 1958b, 1961, 1979). Immunohistochemistry experiments utilizing anti-agouti antibodies are currently being performed to determine precisely where the agouti protein is localized within the follicular environment. Agouti gene expression is restricted to neonatal skin in wild-type mice except for adult testes, which produce a series of transcripts ranging in size from 1.0 kb to 7.5 kb (see part IV; Bultman et al., 1992). None of the testes-specific transcripts are expressed in neonatal skin, however, and it is unlikely that they play any role in agouti-locus function.

Agouti is a Single Gene Locus

Southern blot and Northern blot analyses of existing agouti mutants demonstrated that the DNA structure and pattern of expression of the agouti gene is altered by the most dominant agouti-locus mutation (A^y), alleles in the middle of the dominance hierarchy (A^w , A , a , and a^t), and the most recessive member of the agouti allelic series (a^e) (see part IV; Bultman et al., 1992). This provides compelling evidence that agouti is a single gene locus and that we cloned the one gene associated with agouti-locus function.

The fact that A^y can recombine with several other agouti-locus alleles strongly suggested that agouti is a multiple gene locus (Russell et al., 1963; Siracusa et al., 1987). This finding can be reconciled with the single gene model, however, because the DNA lesion in A^y is a deletion and the recombination likely occurs between agouti and a closely linked gene (see part VII; Michaud et al., 1994). In fact, the recombination frequency of 0.1% is compatible with the 105-kb physical distance between the two genes in the A^y chromosome (see part VII; Michaud et al., 1994). Similarly, the finding that spontaneous reverse mutations from a to a^t or A^w are common, whereas all other spontaneous agouti locus mutations are quite rare, does not contradict the single gene model. It turns out that a is highly unstable due to the insertion of a retrovirus-like element in the agouti gene, and that excision of part or all of the insertion, by homologous recombination, results in the a^t or A^w phenotype (see part V; Bultman et al., 1994).

The Agouti Gene Produces Multiple Forms of mRNA

The finding that the agouti locus is comprised of a single gene suggests that it must be regulated differently in various regions of the pelage, particularly the dorsal and ventral surfaces, in order to produce the phenotypes conferred by alleles like a^t and A^w . RT-PCR experiments demonstrated that, instead of producing a single species of mRNA, the agouti gene produces several similarly-sized transcripts that differ only at their 5' ends (they presumably encode the same protein) (see part V; Bultman et al., 1994). One class of transcript (Form I) appears to be responsible for the subapical band of phaeomelanin in agouti hairs

throughout the animal, whereas a second class of transcript (Form II) appears to be responsible for the production of all-yellow hair on the ventral surface of black-and-tan and white-bellied agouti mice (see part V; Bultman et al., 1994). The different forms of transcripts are selectively inactivated by different agouti-locus mutations and consequently give rise to distinct coat color phenotypes over the dorsal and ventral surfaces of the animal (see part V; Bultman et al., 1994).

Proposed Function of the Agouti Protein

The agouti gene is expressed in the microenvironment of the hair follicle and encodes a novel secreted protein that regulates the differential production of eumelanin and phaeomelanin (see part IV; Bultman et al., 1992). This raises the question as to whether there is an agouti-locus receptor on the surface of the melanocyte and what that receptor might be? The product of the extension (*e*) locus in chromosome 8 of the mouse is a particularly good candidate because *e*-locus mutations give rise to the same all-black and all-yellow phenotypes as observed for agouti mutants. Additionally, the *e* locus acts in a cell-autonomous manner within the melanocyte (Geschwind, 1966; Geschwind et al., 1972; Lamoreux and Mayer, 1975). However, a straightforward receptor-ligand interaction, like that for the *c-kit* receptor and steel ligand (Witte, 1990), seems unlikely because the dominance hierarchy of *e*-locus mutants is reversed relative to that of agouti-locus mutants. For example, dominant agouti mutants are yellow and recessive agouti mutants are black, whereas dominant extension mutants are black and recessive extension mutants are yellow (Silvers, 1979; Green, 1989).

Recently, the *e* locus was shown to encode the α -melanocyte-stimulating hormone receptor (α -MSH-R) (Robbins et al., 1993). The α -MSH-R has seven transmembrane domains spanning the surface of the melanocyte and is coupled to a G-protein which is coupled to adenylate cyclase (Mountjoy et al., 1992). When α -MSH binds to the α -MSH-R, cAMP levels increase inside the melanocyte and trigger a cascade of biochemical events that culminate in eumelanin production overriding the default phaeomelanogenesis pathway (Tamate and Takeuchi, 1981, 1984). Based

on this finding and a number of genetic observations and skin explant experiments, the agouti locus may regulate coat pigmentation by competitively inhibiting the binding of α -MSH to its receptor (Tamate and Takeuchi, 1981, 1984; Takeuchi et al., 1989). In fact, a model has been proposed in which the agouti protein binds directly to the α -MSH-R and acts as an antagonist (see part IV; Tamate and Takeuchi, 1981, 1984; Takeuchi et al., 1989; Bultman et al., 1992; Jackson, 1993a, 1993b). Binding of the agouti protein to the α -MSH-R would prevent α -MSH from binding, thereby preventing eumelanin production and allow the default phaeomelanogenesis pathway to be reestablished. This is only a working hypothesis, however, and the agouti protein could alternatively function by binding to its own, as yet uncharacterized receptor and trigger an independent signal transduction pathway that impinges on, and overrides, the α -MSH eumelanogenesis pathway (Conklin and Bourne, 1993). With the recent cloning of the agouti gene and the α -MSH-R, all of the reagents necessary to test the validity of the antagonist model are available.

Molecular Basis of the A^y Phenotype

The most dominant agouti allele, A^y , is particularly interesting because homozygotes die during embryogenesis before implantation is complete and heterozygous adults are obese, diabetic, and prone to tumor formation (Silvers, 1979). To understand how a coat color gene is involved in such diverse developmental processes, we characterized A^y at the molecular level. Pulsed-field gel electrophoresis (PFGE) analysis of the agouti-locus region revealed that another gene, called *Raly*, lies approximately 280-kb upstream from agouti (see parts VI and VII; Michaud et al., 1993, 1994). PFGE also demonstrated that there is a 170-kb deletion in the A^y chromosome that removes all of *Raly* except for the promoter and first noncoding exon but does not alter the coding exons of the agouti gene (see part VII; Michaud et al., 1994). Because *Raly* encodes a protein putatively involved in pre-mRNA packaging and processing and is expressed in wild-type preimplantation embryos but not A^y/A^y preimplantation embryos, we propose that A^y/A^y embryos die because no *Raly* gene product is produced (see parts VI and VII; Michaud et al., 1993, 1994). Additionally, we propose

that the dominant pleiotropic effects in A^y are due to the ectopic overexpression of the agouti gene because it is under the transcriptional control of the *Raly* promoter (see parts IV, VI, and VII; Bultman et al., 1992; Michaud et al., 1993, 1994). This is supported by the fact that A^{vy}/A^{vy} mice (see part VI; Michaud et al., 1993) and several lines of transgenic mice (M. Klebig and R. Woychik, unpublished data) also ubiquitously overexpress the agouti gene and are perfectly viable and have the same suite of dominant pleiotropic effects. It is our working hypothesis that in these animals, the agouti protein binds to, and interferes with, transmembrane receptor(s) that regulate glucose homeostasis and cell cycle control.

Is Agouti a Human Disease Gene?

Using a cross-species hybridization approach, we cloned the coding exons of the agouti gene from human genomic DNA. Sequence comparison of the mouse and human genes revealed a high degree of similarity (85% identity at nucleotide level and 80% identity at the amino acid level) (see part VIII; Kwon et al., 1994). As expected, based on syntenic conservation between mouse chromosome 2 and human chromosome 20 (Siracusa and Abbott, 1991), the agouti gene was shown to map to human chromosome 20q11-13 using fluorescent *in situ* hybridization (FISH) (see part VIII; Kwon et al., 1994). This finding is significant because maturity onset diabetes of the young (MODY), a non-insulin-dependent form of diabetes mellitus, maps to this same region of chromosome 20q (Bell et al., 1991; Rothschild et al., 1992, 1993).

Importantly, several clinical features of MODY resemble the phenotype of dominant agouti mutants. First, both the human and mouse conditions are inherited in an autosomal dominant manner and are manifest as young adults. Human patients are often diagnosed in their mid-twenties, and the mice usually start displaying the condition shortly after weaning at several months of age. Second, insulin resistance is a fundamental problem affecting both the mouse agouti mutants and the human MODY patients. Insulin is produced in sufficient quantity by the pancreas in insulin-resistant animals, but peripheral tissues like muscle, liver, and fat respond very poorly, resulting in aberrant glucose homeostasis.

Third, the insulin-resistant condition leads to hyperinsulinemia, impaired glucose tolerance, and sometimes hyperglycemia, although the incidence and severity can be quite variable for each of these traits in both species. Fourth, the non-insulin-dependent diabetes in both mouse agouti mutants and human MODY patients is associated with marked obesity. Taken together, the results and observations described above support the notion that the human agouti gene, if mutated in some manner, causes MODY. Furthermore, adipose is one of the few human tissues that has been shown to express the agouti gene (see part VIII; Kwon et al., 1994), and it is normally one of the most insulin-responsive tissues (DeFronzo, 1988).

Despite all of the evidence suggesting that the human agouti gene might be responsible for MODY, contradictory data also exists. Whereas the agouti gene maps proximal to adenosine deaminase (ADA) in both the mouse and human genomes, MODY maps distal to ADA (Bell et al., 1991; Rothschild et al., 1992, 1993). However, it should be noted that the human agouti gene has been physically mapped (using FISH), while MODY has been genetically mapped. Because discrepancies between genetic and physical distances are not uncommon, the mapping of agouti and MODY relative to ADA does not necessarily discount agouti as a candidate gene for MODY. Furthermore, all of the genetics concerning MODY have focused on one relatively small pedigree, the RW family, and it is possible that an inversion or some other type of rearrangement has occurred that suppresses recombination or altered the gene order in this region. Linkage analysis of the human agouti gene relative to MODY using the RW family and other pedigrees, and mutation analysis of MODY patients clearly needs to be performed to resolve the differences described above and determine whether the human agouti gene causes MODY. Finally, even if agouti is not responsible for MODY, it will still play an important role in future diabetes research because it is the first non-insulin-dependent diabetes gene to be cloned that is also associated with an animal model.

REFERENCES

- Bell, G. I., Xiang, K.-S., Newman, M. V., Wu, S.-H., Wright, L. G., Fajans, S. S., Spielman, R. S., and Cox, N. J. (1991). Gene for non-insulin-dependent diabetes mellitus (maturity-onset diabetes of the young subtype) is linked to DNA polymorphism on human chromosome 20q. *Proc. Natl. Acad. Sci. USA* **88**: 1484-1488.
- Bultman, S. J., Klebig, M. L., Michaud, E. J., Sweet, H. O., Davisson, M. T., and Woychik, R. P. (1994). Molecular analysis of reverse mutations from nonagouti (a) to black-and-tan (a^t) and white-bellied agouti (A^w) reveals alternative forms of agouti transcripts. *Genes & Dev.*, in press.
- Bultman, S. J., E. J. Michaud, and R. P. Woychik. (1992). Molecular characterization of the mouse agouti locus. *Cell* **71**: 1195-1204.
- 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. USA* **88**: 8062-8066.
- Conklin, B. R., and Bourne, H. R. (1993). Mouse coat colour reconsidered. *Nature* **364**: 110.
- Davisson, M. T., and Roderick, T. H. (1978). Status of the linkage map of the mouse. *Cytogenet. Cell Genet.* **22**: 552-557.
- DeFronzo, R. A. (1988). The triumvirate: β -cell, muscle, liver. A collusion responsible for NIDDM. *Diabetes* **37**: 667-687.
- Geschwind, I. I., and Huseby, R. A. (1972). Hormonal modification of coat color in the laboratory mouse. In *Pigmentation, Its Genesis and Biologic Control*, (ed. Riley V), pp. 207-214. Appleton-Century-Crofts, New York.

- Geschwind, I. I., Huseby, R. A., and Nishioka, R. (1972). The effect of melanocyte-stimulating hormone on coat color in the mouse. *Rec. Prog. Hormone Res.* **28**: 91-130.
- Green, M. C. (1989). Catalogue of mutant genes and polymorphic loci. In *Genetic variants and strains of the laboratory mouse*, (eds., M. F. Lyon, and A. G. Searle), pp. 17-20. Oxford, New York: Oxford University Press, Oxford and New York.
- Jackson, I. J. (1993a). Colour-coded switches. *Nature* **362**: 587-588.
- Jackson, I. J. (1993b). More to colour than meets the eye. *Current Biology* **3**: 518-521.
- Kwon, H. Y., Bultman, S. J., Loffler, C., Chen, W.-J., Furdon, P. J., Powell, J. G., Usala, A.-L., Wilkinson, W. W., Hansmann, I., and Woychik, R. P. (1994). Molecular structure and chromosomal mapping of the human homolog of the agouti gene. Submitted.
- Lamoreux, M. L., and Mayer, T. C. (1975). Site of gene action in the development of hair pigment in recessive yellow (*e/e*) mice. *Dev. Biol.* **46**: 160-166.
- Michaud, E. J., S. J. Bultman, L. J. Stubbs, and R. P. Woychik. (1993a). The embryonic lethality of homozygous lethal yellow mice (A^y/A^y) is associated with the disruption of a novel RNA-binding protein. *Genes & Dev.* **7**: 1203-1213.
- Michaud, E. J., S. J. Bultman, M. L. Klebig, M. J. van Vugt, L. J. Stubbs, L. B. Russell, and R. P. Woychik. (1994). A molecular model for the genetic and phenotypic characteristics of the mouse lethal yellow (A^y) mutation. *Proc. Natl. Acad. Sci.*, in press.

- Mountjoy, K. G., Robbins, L. S., Mortrud, M. T., and Cone, R. D. (1992). The cloning of a family of genes that encode the melanocortin receptors. *Science* 257: 1248-1251.
- Robbins, L. S., Nadeau, J. H., Johnson, K. R., Kelly, M. A., Roselli-Rehfuß, L., Baack, E., Mountjoy, K. G., and Cone, R. D. (1993) Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* 72: 827-834.
- Rothschild, C. B., Akots, G., Fajans, S. S., and Bowden, D. W. (1992). A microsatellite polymorphism associated with the PLC1 (phospholipase C) locus: Identification, mapping, and linkage to the MODY locus on chromosome 20. *Genomics* 13: 560-564.
- Rothschild, C. B., Akots, G., Hayworth, R., Pettenati, M. J., Rao, N., Wood, P., Stolz, F. M., Hansmann, I., Serino, K., Keith, T. P., Fajans, S. S., and Bowden, D. W. (1993). *Am. J. Hum. Genet.* 52: 110-123.
- Russell, L. B., McDaniel, N. C., and Woodiel, F. N. (1963). Crossing-over within the α "locus" of the mouse. *Genetics* 48: 907a.
- Silvers, W. K. (1958a). An experimental approach to action of genes at the agouti locus in the mouse. II. Transplants of newborn aa ventral skin to $a^t a$, $A^w a$, and aa hosts. *J. Exp. Zool.* 137: 181-188.
- Silvers, W. K. (1958b). An experimental approach to action of genes at the agouti locus in the mouse. III. Transplants of newborn A^w -, A -, a^t -skin to A^y -, A^w -, A - and aa hosts. *J. Exp. Zool.* 137: 189-196.
- Silvers, W. K. (1961). Genes and pigment cells of mammals. *Science* 134: 368-373.

- Silvers, W. K. (1979). The agouti and extension series of alleles, umbrous, and sable. In *The coat colors of mice: A model for mammalian gene action and interaction*, pp. 6-44. Springer-Verlag, New York.
- Silvers, W. K., and Russell, E. S. (1955). An experimental approach to action of genes at the agouti locus in the mouse. *J. Exp. Zool.* **130**: 199-220.
- Siracusa, L. D., Russell, L. B., Jenkins, N. A., and Copeland, N. G. (1987a). Allelic variation within the *Emv-15* locus defines genomic sequences closely linked to the *agouti* locus on mouse chromosome 2. *Genetics* **117**: 85-92.
- Siracusa, L. D., and Abbott, C. M. (1991). Mouse chromosome 2. *Mammalian Genome* **1**: S18-S41.
- Takeuchi, T., Kobunai, T., and Yamamoto, H. (1989). Genetic control of signal transduction in mouse melanocytes. *J. Invest. Dermatol.* **92**: 239S-242S.
- Tamate, H. B., and Takeuchi, T. (1981). Induction of the shift in melanin synthesis in lethal yellow (*A^y/a*) mice in vitro. *Devl. Genet.* **2**: 349-356.
- Tamate, H. B., and Takeuchi, T. (1984). Action of the *e* locus of mice in the response of phaeomelanic hair follicles to α -melanocyte-stimulating hormone in vitro. *Science* **224**: 1241-1242.
- Witte, O. N. (1990). Steel locus defines new multipotent growth factor. *Cell* **63**: 5-6.
- Woychik, R. P., Steart, 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.

Woychik, R. P., Generoso, W. M., Russell, L. B., Cain, K. T., Cacheiro, N. L. A., Bultman, S. J., Selby, P. B., Dickinson, M. E., Hogan, B. L. M., and Rutledge, J. C. (1990a). 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. USA* **87**: 2588-2592.

Woychik, R. P., Maas, R. L., Zeller, R., Vogt, T. F., and Leder, P. (1990b). 'Formins': proteins deduced from the alternative transcripts of the *limb deformity* gene. *Nature* **346**: 850-853.

VITA

Scott Justin Bultman was born and raised in Milwaukee, Wisconsin. After graduating from Wauwatosa East High School in 1983, he attended the University of Wisconsin- Stevens Point. He received a B. S. degree in May of 1988 with a major in Biology and a minor in Chemistry. He entered the University of Tennessee Graduate School of Biomedical Sciences at Oak Ridge in August of 1988 and was awarded the Doctor of Philosophy degree in May of 1994. He will begin postdoctoral study, under the direction of Dr. David J. Anderson, at the California Institute of Technology in July of 1994, and will pursue developmental neurobiology.

VITA

Scott Justin Bultman was born and raised in Milwaukee, Wisconsin. After graduating from Wauwatosa East High School in 1983, he attended the University of Wisconsin- Stevens Point. He received a B. S. degree in May of 1988 with a major in Biology and a minor in Chemistry. He entered the University of Tennessee Graduate School of Biomedical Sciences at Oak Ridge in August of 1988 and was awarded the Doctor of Philosophy degree in May of 1994. He will begin postdoctoral study, under the direction of Dr. David J. Anderson, at the California Institute of Technology in July of 1994, and will pursue developmental neurobiology.