

To the Graduate Council:

I am submitting herewith a dissertation written by Mitchell L. Klebig entitled "Molecular Analysis of Murine Albino (c)-Locus Deletion Mutations Defining the Hepatocyte-Specific Developmental Regulation-1 (*hsdr-1*) and Mesoderm-Deficient (*msd*) Loci Identifies the Fumarylacetoacetate Hydrolase (*Fah*) Gene in the Vicinity of the *hsdr-1* Locus". I have examined the final copy of this dissertation for form and content and recommend that it be accepted for the degree of Ph. D., with a major in Biomedical Sciences.

E. M. Rinchik

Eugene M. Rinchik
Major Professor

We have read this dissertation
and recommend its acceptance:

Paul M. Weyel

Liane B. Russell

Mary Ann Handel

Accepted for the Council:

Cowminkel

Vice Provost
and Dean of The Graduate School

MOLECULAR ANALYSIS OF MURINE ALBINO (*c*)-LOCUS
DELETION MUTATIONS DEFINING THE HEPATOCYTE-
SPECIFIC DEVELOPMENTAL REGULATION-1 (*hsdr-1*)
AND MESODERM-DEFICIENT (*msd*) LOCI
IDENTIFIES THE FUMARYLACETOACETATE
HYDROLASE (*Fah*) GENE IN THE
VICINITY OF THE *hsdr-1* LOCUS.

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

Mitchell Louis Klebig
August 1991

Abstract

Complementation analyses of radiation-induced deletion mutations involving the albino (*c*) locus in chromosome 7 of the mouse have identified several loci, both proximal and distal to *c*, that have important roles in development. Embryos homozygous for a deletion of the most proximal of these loci, "mesoderm-deficient" (*msd*), fail to develop mesoderm and die at Day 11.5 of gestation. Animals homozygous for a deletion of the "hepatocyte-specific developmental regulation-1" (*hsdr-1*) locus, located between the *msd* and *c* loci, develop morphological, ultrastructural, and biochemical abnormalities of the liver and kidney, and die shortly after birth. It has been postulated that the *hsdr-1* locus encodes a *trans*-acting regulatory factor(s) required for normal liver- and kidney-specific gene expression of various enzymes and proteins. Cloning the breakpoints of various albino deletions that define the extents of the *msd* and *hsdr-1* loci is one way of generating molecular probes for studying the particular gene(s) involved in these developmental functions. Mapping of the anonymous locus *D7Was12* with respect to the panel of albino deletions suggested that it, like the previously cloned *Emv-23* locus, might be a useful access point for cloning some of these deletion breakpoints. Using pulsed-field-gel (CHEF) analysis, a long-range map of wild-type DNA spanning about 1.7 megabases surrounding the *c*, *D7Was12*, and *Emv-23* loci was constructed. The molecular identification (with pulsed-field-gel electrophoresis) of structural rearrangements caused by the breakpoints of five albino deletions involved in the *hsdr-1* and *msd* phenotypes, as well as of two viable albino deletions that remove only part of the tyrosinase gene and extend distally, enabled the localization of the deletion breakpoints on the long-range map. This analysis revealed that cloning the breakpoints of the viable deletions can generate additional DNA probes that will, in turn, facilitate the cloning of the breakpoints of some of the deletions defining the *hsdr-1* and *msd* functions.

A deletion-breakpoint-fusion fragment cloned from *c24R75M/c24R75M* DNA with the MTY811 probe proved to be instrumental in cloning the breakpoint of a neonatally lethal deletion, *c14CoS* (complementation designation A1), one of the original mutations used to define the *hsdr-1* locus. DNA probes, isolated from a portion of the A1-breakpoint-fusion fragment located proximal to the breakpoint (i.e., within the *hsdr-1* region), detect transcripts that are most abundant in the liver and kidney. Liver cDNA clones, isolated with one of these probes, were found to encode fumarylacetoacetate hydrolase (FAH). All of the deletions eliciting the *hsdr-1* phenotype prevent expression of this *Fah* gene, and most disrupt the coding segment of the gene. Therefore, the *Fah* gene maps within the *hsdr-1* locus, as defined by deletion breakpoints, and disruption of this gene may be responsible for the *hsdr-1* phenotype. Mutations in the human *FAH* gene are associated with type 1 hereditary tyrosinemia (*HT*), a disorder which, like that observed in mice homozygous for deletions of *hsdr-1*, causes morphological abnormalities and biochemical dysfunctions of the liver and kidney. However, there may be other gene(s) at the *hsdr-1* locus that are associated with the specific biochemical and ultrastructural abnormalities, observed in deletion homozygotes, that have not yet been reported in *HT* patients. Further studies, particularly correction-of-phenotype experiments with transgenic mice, will be required to determine whether or not loss of the *Fah* gene is responsible for the abnormalities present in mice homozygous for several different albino deletions.

TABLE OF CONTENTS

SECTION	PAGE
I. INTRODUCTION	1
Deletion Mutations Are Useful For Identifying Genes Functioning In a Variety of Developmental Processes . . .	2
A Functional Map of a Region of Chromosome 7	2
Surrounding the Albino (<i>c</i>) Locus	
The mesoderm-deficient (<i>msd</i>) Locus and	5
Mesoderm Formation	
The hepatocyte-specific developmental	8
regulation-1 (<i>hsdr-1</i>) Locus	
A Strategy For Gaining Molecular Access To	11
the <i>hsdr-1</i> and <i>msd</i> Loci	
II. MATERIALS AND METHODS	12
Mice	12
Genomic DNA Isolation and Southern-Blot	13
Analysis (Conventional)	
Hybridization Probes	14
Large-Fragment/Pulsed-Field-Gel Electrophoresis Analysis	14
Library Construction	16
RNA Isolation and Northern-Blot Analysis	18

PCR Amplification and Sequence Analysis of cDNA Inserts	18
III. RESULTS	21
The <i>D7Was12</i> and <i>Emv-23</i> Loci Map Distal to Some of the Deletions of the A and Ai Complementation Groups	21
Physical Linkage of the <i>c</i> , <i>D7Was12</i> , and <i>Emv23</i> Loci On a Long-Range Restriction Map of a 1.7-Megabase Region of Chromosome 7	23
Localization of the Breakpoints of Several Deletions (Complementation groups A and Ai) Defining the <i>hsdr-1</i> and <i>msd</i> Loci On the Long-Range Map	28
Molecular Analysis of the Viable Albino Deletions <i>c^{24R75M}</i> and <i>c^{16DFiOD}</i> and Localization of Their Distal Breakpoints in the Vicinity of the Distal Breakpoints of the A1, A2, and Ai1 Deletions	31
Cloning a Deletion-Breakpoint-Fusion Fragment (pRN207) from <i>c^{24R75M}/c^{24R75M}</i> DNA	33
A Distal-End Clone (pRN204) Derived From the Rearranged Fragment RN207 Maps Distal to the A1 Deletion and Hybridizes To Size-Altered Fragments in A1/A1 DNA	34
Cloning a Deletion Breakpoint-Fusion Fragment (pRN226) From <i>c^{14CoS}/c^{14CoS}</i> (i.e., A1/A1) DNA With Probe RN204	36
A Proximal-End Clone (RN247) of the A1-Breakpoint-Fusion Fragment (RN226) Identifies a Locus (<i>D7Rn2</i>) That Maps Proximal to Deletions A1 and A5 But Within the Other Deletions That Define <i>hsdr-1</i>	37

Probes (RN226.1 and RN226.2) Derived From the <i>D7Rn2</i> Locus Hybridize To Abundant Transcripts in Liver and Kidney That Are Absent in A1/A1 and A5/A5 Embryos	39
Liver cDNA Clones Isolated With Probe RN226.2 Have the Same Hybridization Profiles in Wild-Type and Mutant RNAs as Probes RN226.1 and RN226.2	46
The cDNA Clones Are Of Two Classes and Identify a Structural Locus (Or Loci) That Is Deleted Or Disrupted By All the Deletions Defining <i>hsdr-1</i> , With the Exception of the A5 Deletion	47
The Identified cDNA Clones Encode the Enzyme Fumarylacetoacetate Hydrolase (FAH)	48
The <i>Fah</i> (<i>D7Rn2</i>) and <i>D7Rn4</i> Loci Will Contribute To Molecular Analyses of the <i>jdf</i> Locus and Genetic Loci at the Distal Ends of the Di4 and Fp1 Deletions	50
Attempts to Access the <i>msd</i> Locus at the Molecular Level By Cloning Ai-Deletion Breakpoints	54
Mapping the Mouse <i>Fes</i> and <i>D7Rn4</i> Loci By Interspecific Backcross Analysis Indicates That <i>Fes</i> May Facilitate Molecular Access To the <i>msd</i> Locus From the Proximal Side	57
IV. DISCUSSION	61
Gaining Molecular Access to the <i>hsdr-1</i> and <i>msd</i> Loci By Cloning Deletion Breakpoint-Fusion Fragments	61
The Fumarylacetoacetate Hydrolase (<i>Fah</i>) Gene Maps at the Proximal Breakpoint of the <i>c¹⁴CoS</i> (A1) Deletion	64

	vii
Could <i>Fah</i> Deficiency Be Responsible For the <i>hsdr-1</i> Phenotype?	70
The <i>Fah</i> and <i>D7Rn4</i> Loci Will Contribute To Molecular Analyses of Additional Genetic Loci	75
Strategies Used to Facilitate Molecular Access To the <i>msd</i> Locus	78
Potential CpG Islands and Transcription Units Identified by Long-Range Mapping	82
LIST OF REFERENCES	85
APPENDIXES	98
APPENDIX A. TABLES	99
APPENDIX B. FIGURES	108
VITA	152

LIST OF TABLES

TABLE		PAGE
1	Genetic Loci Defined by Complementation Analysis of Albino Deletions	100
2	Hybridization Probes	101
3	Sequences of Primers Used For cDNA Sequencing .	103
4	Wild-type Restriction-fragment Sizes Used to Construct a Long-range Map of the Region of Chromosome 7 Containing the <i>c</i> , <i>D7Was12</i> , and <i>Emv-23</i> Loci.	104
5	Approximate Physical Distances (in kb) of the Distal Breakpoints of the A and Ai Deletions From the <i>c</i> and <i>D7Was12</i> Loci and From the Distal Breakpoints of the <i>c</i> ^{24R75M} and <i>c</i> ^{16DFiOD} Deletions.	105
6	RNAs Analyzed in Fig. 14	106
7	Interspecific Backcross Mapping of <i>D7Rn4</i> and . . . <i>Fes</i> with Respect to Mouse Chromosome-7 Markers <i>Igf1r</i> , <i>Tyr (c)</i> , and <i>Hbb</i>	107

LIST OF FIGURES

FIGURE		PAGE
1	A genetic/functional map of the <i>c</i> -region in chromosome 7 and the associated albino deletion mutations	109
2	A genetic/functional map of a proximal portion of the <i>c</i> -region in chromosome 7 and some of the associated albino deletion mutations	112
3	Deletion-mapping cloned DNA loci within the albino-deletion complex	114
4	Pulsed-field gel analysis of the chromosome-7 region containing the <i>D7Was12</i> , <i>Emv-23</i> , and <i>c</i> loci	116
5	A long-range, wild-type restriction map of the region of chromosome 7 containing the <i>c</i> , <i>D7Was12</i> , and <i>Emv-23</i> loci	118
6	Molecular detection of the distal breakpoints of several viable, neonatally lethal, and prenatally lethal albino-deletion mutations	120
7	The approximate physical locations of the breakpoints of several viable (<i>c</i> ^{24R75M} and <i>c</i> ^{16DFiOD}), group-A, and group-Ai deletions on the long-range restriction map	123
8	Restriction maps of deletion breakpoint-fusion fragments RN207 and RN226 cloned from <i>c</i> ^{24R75M} and <i>c</i> ^{14CoS} (A1) chromosomes	125

		x
9	Deletion-mapping cloned DNA loci within the albino deletion complex	127
10	A proximal portion of the physical/functional map of the <i>c</i> region, as refined by deletion-mapping of the <i>D7Rn1</i> locus	129
11	The physical/functional map as refined by deletion-mapping of the <i>D7Rn2</i> and <i>Fah</i> loci	131
12	Probes RN226.1 and RN226.3 are derived from a portion . of the A1-breakpoint-fusion fragment (RN226) located proximal to the breakpoint	133
13	Probes RN226.1 and RN226.2 identify evolutionarily . . conserved DNA sequences	135
14	Northern-blot analyses of the livers of wild-type . . . and mutant mice, and of other wild-type tissues with genomic-DNA and cDNA probes	137
15	Comparison of the nucleotide sequence of mouse . . . cDNA RN273 with the human <i>FAH</i> gene sequence reported by Phaneuf <i>et al.</i> (1991)	139
16	Comparison of the amino-acid sequence encoded by . . the ORF of mouse cDNA RN273 (see Fig. 15) with the human FAH sequence reported by Phaneuf <i>et al.</i> (1991)	142
17	Possible identification of the A5 breakpoint with a probe . from the <i>D7Rn2</i> locus	144
18	The <i>D7Rn4</i> locus maps proximal to the proximal . . . breakpoints of the Di4 and Fp1 deletions	146

- 19 The physical/functional map as refined by 148
 deletion-mapping of the *D7Rn4* and *D7Rn3* loci
- 20 The *D7Rn3* locus maps distal to the Ai1 and *c^{24R75M}* . . . 150
 breakpoints, but proximal to the A2 breakpoint

LIST OF ABBREVIATIONS

BSA	Bovine-serum albumin
cAMP	Cyclic adenosine monophosphate
cDNA	Complementary DNA derived from reverse transcription of messenger RNA
CHEF	Contour-clamped homogeneous electric field
cM	Centimorgan (1% recombination)
CpG	Dinucleotide of deoxycytidine 5'-monophosphate (5') and deoxyguanosine 5'-monophosphate (3'), which occurs infrequently in mammalian genomic DNA
C.Z.	Compression zone
δ -ALD	δ -aminolevulinic acid dehydratase
EDTA	Disodium ethylenediamine tetraacetate
ENU	N-ethyl-N-nitrosourea
ES cells	Embryonic-stem cells
FAH	Fumarylacetoacetate hydrolase
FGF	Fibroblast growth factor
aFGF	Acidic FGF
bFGF	Basic FGF
GSH	Glutathione (reduced)
GT-B	Glutathione transferase-B
HTF	<i>Hpa</i> II tiny fragments
kb	Kilobase
LMP	Low-melting point
mb	Megabase
MOPS	3-(N-Morpholino)propanesulfonic acid
MW	Molecular weight
oligo-dT	Oligodeoxythymidilic acid
ORF	Open-reading frame
^{32}P -dATP	Deoxyadenosine 5'-[α - ^{32}P] triphosphate
^{32}P -dCTP	Deoxycytidine 5'-[α - ^{32}P] triphosphate
PAGE	Polyacrylamide-gel electrophoresis

PCR	Polymerase chain-reaction
PFG	Pulsed-field gel
PFGE	Pulsed-field-gel electrophoresis
PMSF	Phenylmethanesulfonyl fluoride
poly-(A+) RNA	Eukaryotic RNA after enrichment for messenger RNAs containing long sequences of successive adenylate residues at the 3' end
RFLV	Restriction-fragment-length variant
SAC	Succinylacetone
SSC	Normal saline-citrate
SV40	Simian virus-40
TE	10 mM Tris-Cl, 0.1mM EDTA solution
TGF β	Transforming growth factor β
UDP-GT	UDP glucuronosyltransferase
UV	Ultraviolet
YAC	Yeast-artificial chromosome

I. INTRODUCTION

Deciphering the genetic programs that direct mammalian development will require, at a fundamental level, the continued cloning and characterization of genes identified by mutations that disrupt normal development at different stages of the life cycle (e.g., prenatal lethality, morphological abnormalities, fitness or runting syndromes). These types of genes are often difficult to study molecularly because they are not usually associated with known protein products *a priori*. More traditional cloning approaches that rely on information about the particular proteins affected (e.g., amino-acid sequence, antibodies to the purified protein) are, therefore, not possible with these types of mutations. The techniques of reverse genetics ("positional cloning") are proving to be powerful approaches for identifying mutant loci at the molecular level and for cloning the affected genes, provided appropriately enriched DNA libraries and/or cloned DNA markers closely linked to the mutant locus are available. This is particularly true for loci affected by chromosomal aberrations (translocations, deletions, etc.) that have breakpoints that can identify or delimit a mutant locus on a physical map [e.g., Royer-Pokora *et al.*, 1986 (CGD); Kunkel *et al.*, 1986 and Monaco *et al.*, 1986 (DMD); Viskochil *et al.*, 1990 (NF1); Herrmann *et al.*, 1990 (T)].

Deletion Mutations Are Useful For Identifying Genes Functioning In a Variety of Developmental Processes

A useful characteristic of deletion mutations is that complementation analyses of panels of overlapping deletions, uncovered as mutations at specific loci (Russell, 1951), can be used to construct functional maps of chromosomal regions, several centimorgans (cM) in size, that harbor a variety of genes functioning in various developmental processes (Russell, 1971; Gluecksohn-Waelsch, 1979; Russell *et al.*, 1982; and Rinchik and Russell, 1990). DNA clones and new point mutations can then be rapidly positioned on the functional maps by mapping them with respect to deletion breakpoints (e.g., Disteché and Adler, 1984; Rinchik *et al.*, 1989; Johnson *et al.*, 1989; Rinchik and Russell, 1990; Niswander *et al.*, 1991; and Rinchik, 1991). These DNA probes and point mutations can subsequently be used, in conjunction with the deletions, to facilitate the physical and functional mapping of the regions spanned by the deletions, to help correlate the evolving physical and functional maps, and, eventually, to clone the genes identified by functional analyses (Rinchik and Russell, 1990; Rinchik, 1991).

A Functional Map of a Region of Chromosome 7 Surrounding the Albino (*c*) Locus

Complementation analyses of a panel of radiation-induced pre- and postnatally lethal albino- (*c*-) locus mutations, presumed to be deletions based on genetic studies (Russell *et al.*, 1982; Gluecksohn-

Waelsh, 1979; and references therein), and later histological analyses of embryos homozygous for some of these mutations (Niswander *et al.*, 1988; 1989) have permitted a 6-to-11-cM region of chromosome 7 surrounding the *c* locus to be subdivided into a linear series of at least 9 genetic loci defined by at least 14 complementation groups of deletions [see Rinchik and Russell (1990) and Niswander *et al.* (1991) for the most recent functional maps]. This region includes much of metaphase band 7E (Russell and Cacheiro, 1977) and comprises roughly 10% of chromosome 7, based on a total length of about 90 cM for the chromosome (Doolittle *et al.*, 1990). Several points of molecular access have been achieved within the *c* region by the cloning of the tyrosinase (*c*-locus) gene (Kwon *et al.*, 1988; Ruppert *et al.*, 1988), endogenous proviral integration sites (Rinchik *et al.*, 1989; J. Stoye, personal communication), and numerous anonymous DNA loci (Disteche and Adler, 1984; Johnson *et al.*, 1989; Niswander *et al.*, 1991; and additional references within Rinchik and Russell, 1990). A current genetic/functional map of the *c* region, including the extents of the various albino deletions, is illustrated in Fig. 1*. The positions of standard marker loci and of some of the available cloned loci, with respect to deletion breakpoints, are shown above the chromosome [see Rinchik and Russell (1990) for the locations of all currently mapped DNA loci in this region]. The positions of genetic loci identified solely by complementation analysis of the deletion mutations are shown below the chromosome.

* All figures and tables are in the Appendix.

The function of each of the loci identified by complementation analysis has been defined to date by the phenotype exhibited when the locus is homozygously deleted. These loci are described in Table 1 and have been described in detail previously (Rinchik and Russell, 1990 and references therein). Molecular analyses of these complementation-defined loci and identification of the gene(s) responsible for the function attributed to each locus will contribute significantly to our understanding of the genetic control of development. Currently, the physical locations of these genetic loci can be described only as DNA segments located between deletion breakpoints, none of which have yet been localized on a physical map. The complementation groups that best define the extents of each locus are indicated in Table 1 and illustrated in Fig. 1. These genetic loci were originally referred to as "functional units" (Russell *et al.*, 1982), subsequently called "DNA subregions" (Rinchik and Russell, 1990), and recently renamed according to standard nomenclature for genetic loci in the mouse (Committee for mouse chromosome-7 report, 1991). It should be emphasized that each of these loci may harbor more than one gene, and, in such cases, the phenotype observed from deletion of the locus may result from inactivation of the gene that acts earliest in development or from the additive effects of loss of more than one gene contained within the locus.

This study involved molecular analyses of albino deletions defining two of the complementation-defined loci, hepatocyte-specific developmental regulation-1 (*hsdr-1*; Gluecksohn-Waelsch, 1989) and mesoderm-deficient (*msd*) (see Fig. 1 for the positions of these loci on the map). The goal of the study was to use reverse-genetics strategies to

obtain DNA probes that will facilitate the eventual isolation of the genes contributing to the function of each locus.

The mesoderm-deficient (*msd*) Locus and Mesoderm Formation

The mesoderm-deficient (*msd*) locus, called "implantation survival" ("IS") in previous maps (e.g., Russell *et al.*, 1982; Rinchik and Russell, 1990), is the most proximal locus affected by the albino deletions (Fig. 1). Recent histological analyses have revealed that embryos homozygous for deletions of *msd* (Ai/Ai embryos) are completely devoid of mesoderm and begin to degenerate at about 10.5 days of gestation (B. Holdener-Kenny and T. Magnuson, personal communication). The embryonic and extraembryonic ectoderm of these embryos are present and develop, but the embryonic ectoderm does not differentiate. A primitive streak is also evident in Ai/Ai embryos, but mesodermal precursors do not migrate from the streak. Moreover, the parietal endoderm forms a typical yolk sac and does not collapse, despite the abnormally small size of the embryo inside. Despite these abnormalities, embryonic-stem- (ES-) cell lines have been derived from Ai/Ai embryos (B. Holdener-Kenny and T. Magnuson, personal communication). Molecular analysis of the *msd* locus and isolation of the gene(s) within the locus that is (are) necessary for mesoderm formation will help to unravel the genetic control of this process.

Growth factors from two families, the transforming-growth-factor- β (TGF β) and fibroblast-growth-factor (FGF) families, have been shown to be mesoderm inducers in early *Xenopus* embryogenesis (reviewed in

Harland, 1988 and Slack, 1990; additional references in Thomsen *et al.*, 1990). Animal-pole explants (ectodermal cells) from *Xenopus* blastulae form ectodermal tissue (epidermis) when cultured *in vitro* without inducing agents. In general, when these explants are treated with TGF β -like growth factors (e.g., TGF β 2, activin A, Vg1), they form dorsal-type embryonic mesoderm tissue rather than epidermis. On the other hand, when animal-pole explants are treated with FGF-like factors (e.g., bFGF and aFGF), they form ventrolateral-type mesoderm tissue.

The ability of mammalian TGF β -like (e.g., TGF β 1 and activin A) and FGF-like (e.g., kFGF and INT-2) factors to also induce mesoderm formation in *Xenopus* ectodermal explants (Paterno *et al.*, 1989; Thomsen *et al.*, 1990; and references therein) implies that these factors are probably also functioning as mesoderm-inducers in mammalian systems. The expression of the *int-2* protooncogene in the primitive streak of the 7.5-9.5-day mouse embryo (Wilkinson *et al.*, 1988) is additional evidence that the encoded protein is an inducing factor, since all the germ layers of the embryo (ectoderm, mesoderm, and endoderm) arise from the primitive streak (Paterno *et al.*, 1989). Until the cloning of the Brachyury (*T*) gene of the mouse (Herrmann *et al.*, 1990), however, no other gene products involved in mesoderm formation in mammals had been identified (Wilkinson *et al.*, 1990).

Mouse embryos homozygous for the *T* mutation have a reduced number of mesodermal cells and an increased number of ectodermal cells compared to wild-type (Yanagisawa *et al.*, 1981). In addition, they exhibit severe morphological abnormalities of mesoderm-derived embryonic structures (particularly the notochord), probably due to defects

in morphogenetic movements (Wilkinson *et al.*, 1990 and references therein; Yanagisawa *et al.*, 1981). Using *in situ* hybridization, Wilkinson *et al.* (1990) have found that the *T* gene is expressed in early-stage mesoderm cells and their epithelial precursors, and, later, only in the notochord. They conclude that, because the *T* gene is expressed mainly in the tissues most affected by the *T* mutation, early mesoderm and the notochord, it functions directly in the formation of mesoderm and in the morphogenesis of the notochord. However, the fact that some mesodermal cells are formed in *T/T* embryos (Yanagisawa *et al.*, 1981) implies that the *T* gene functions after mesoderm induction. Wilkinson *et al.* (1990) have, in fact, proposed that a possible relationship between the expression of growth factors in the primitive streak and the expression of the *T* gene should be investigated. It is very likely, however, that other genes are also involved in the induction and initial formation of mesoderm.

The complete absence of the mesoderm in mouse embryos homozygous for deletions of *msd* (B. Holdener-Kenny and T. Magnuson, personal communication) implies that the *msd* gene(s) involved in mesoderm formation functions earlier in the process than the *T* gene. Perhaps the *msd* gene is another growth factor that induces mesoderm. Alternatively, it may have a novel function in the mesoderm-formation pathway, possibly acting between induction and *T*-gene expression.

The hepatocyte-specific developmental regulation-1 (*hsdr-1*) Locus

The *hsdr-1* locus (Gluecksohn-Waelsch, 1989), located proximal to the *c* locus (Fig. 2) and called "neonatal survival" ("NS"; e.g., Russell *et al.*, 1982; Rinchik *et al.*, 1989) in previous *c*-region maps, is the most well-characterized of the genetic functions that have been identified by complementation analyses of lethal albino deletions. Most of the phenotypic analyses have been conducted on mice homozygous for the *c^{14CoS}* (i.e., A1) or *c^{3H}* deletions (Fig. 2). Newborn mice homozygous for deletions of *hsdr-1* exhibit numerous abnormalities, some of which are: (1) morphological abnormalities of the kidney and thymus (Erickson *et al.*, 1968); (2) ultrastructural membrane abnormalities of the endoplasmic reticulum, Golgi apparatus, and nucleus in liver parenchymal cells and certain cells of the proximal convoluted tubules of the kidney (Trigg and Gluecksohn-Waelsch, 1973; Gluecksohn-Waelsch, 1979); (3) deficiencies (often extreme) of activity or quantity of numerous liver- and kidney-specific enzymes and proteins, serum proteins, and hormone receptors (reviewed in Gluecksohn-Waelsch, 1979 and 1987; see also Morris *et al.*, 1988; Baier *et al.*, 1984; Goldfeld *et al.*, 1981; and Goldfeld *et al.*, 1983); (4) the failure of liver- and kidney-specific, glucocorticoid- and cAMP-inducible expression (but not constitutive expression) of the structural genes for many of the affected enzymes and proteins, as evidenced by severely reduced mRNA levels and transcription rates in comparison with normal newborns (reviewed in Gluecksohn-Waelsch, 1987; see also DeFranco *et al.*, 1988; Donner *et al.*, 1988; McKnight *et al.*, 1989; and Ruppert *et al.*, 1990); (5) a >25-fold

increase in the activity of NAD(P)H:oxidoreductase (involved in protection from oxidative stress), as well as a >100-fold increase (in mRNA levels) in the expression of the *Nmo-1* [NAD(P)H:menadione oxidoreductase] gene encoding the enzyme in the liver (Petersen *et al.*, 1989); (6) a twofold increase in the activities in the liver of glutathione transferase B (GT-B; also called ligandin), an enzyme essential for nonoxidative detoxification (Gatmaitan *et al.*, 1977), and of UDP glucuronosyltransferase (UDP-GT) with *p*-nitrophenol as substrate (Thaler *et al.*, 1976); (7) increases in the levels of two proteins identified by 2-dimensional gel electrophoresis (Baier *et al.*, 1984); and ultimately (8) perinatal death of the animal, presumably due to hypoglycemia (Erickson *et al.*, 1968; Russell and DeHamer, 1973).

Based on these and other observations, Gluecksohn-Waelsch (1987) and others (Donner *et al.*, 1988 and Ruppert *et al.*, 1990) have proposed that the *hsdr-1* locus, also called "albino lethal factor" ("*alf*") by Ruppert *et al.* (1990), harbors a gene(s) that encode(s) a *trans*-acting positive regulatory factor(s) involved in the tissue-specific, inducible expression of many structural genes. In addition, due to the pronounced induction of *Nmo-1* gene expression (and also the smaller increases in GT-B and UDP-GT activities) observed in mice homozygous for deletions of *hsdr-1*, Petersen *et al.* (1989) have concluded that the *hsdr-1* locus also harbors a gene important in negatively regulating the expression of the *Nmo-1* gene, and possibly of other genes, in the liver. The relatively small increases in GT-B and UDP-GT activities could also be due to secondary effects (e.g., cellular stress) of the deletions on these enzymes or on their genes (Petersen *et al.*, 1989). Petersen *et al.* (1989) also state that it is

"quite unlikely" that the action of a single gene is responsible for the positive or negative regulation of these distinct types of enzymes.

The above analyses on the expression of the genes affected secondarily by deletion of *hsdr-1* clearly show that the *hsdr-1* gene(s) is (are) acting at the level of transcriptional control. However, no evidence is yet available on whether *hsdr-1* has a direct effect (e.g., as a transcription factor) on these genes, or whether it is acting earlier in a regulatory pathway (Ruppert *et al.*, 1990). Cloning the gene(s) responsible for *hsdr-1* activity may elucidate the nature of its action.

Other studies of liver-specific gene regulation have found that a subset of the genes positively regulated by *hsdr-1* in liver are also the targets of trans-dominant, negative regulation by the tissue-specific extinguisher-1 (*Tse-1*) gene in hepatoma hybrids (Killary and Fournier, 1984; Lem *et al.*, 1988; and Ruppert *et al.*, 1990). Ruppert *et al.* (1990) argue that these findings are evidence for the coordinate control of liver-specific expression of sets of genes by two different *trans*-acting regulatory factors. Additionally, the *Tse-1*-mediated extinction of liver-specific gene expression can be reversed by the administration of glucocorticoids and cAMP *in vitro* (Ruppert *et al.*, 1990; and references therein). This finding suggests that these hormones, which are dependent on *hsdr-1* for the inducibility of the affected liver-specific genes, may act coordinately with *Tse-1* and *hsdr-1* in the expression of these genes *in vivo* (Ruppert *et al.*, 1990; and references therein). Ruppert *et al.* (1990) also point out that *hsdr-1* is probably not a target of *Tse-1* action, since only some of the genes affected by *hsdr-1* are extinguished by *Tse-1*.

A Strategy For Gaining Molecular Access To the *hsdr-1* and *msd* Loci

A current genetic/functional map of a proximal portion of the *c* region, including the extents of some of the deletions affecting this region, is illustrated in Fig. 2. At least a portion of the gene(s) affected by deletion of the *hsdr-1* locus is contained within the smallest region identifiable between the proximal breakpoints of those deletions that delete *hsdr-1*, but not *msd* (e.g., deletions of complementation groups A and E; Fig. 2), and the Bem1 deletion, recently found to have the most proximal breakpoint of all the deletions with breakpoints located between the *c* and *hsdr-1* loci (Niswander *et al.*, 1991 and Rinchik *et al.*, in preparation; see Rinchik and Russell (1990) for a review of the DNA loci recently cloned from this region). In addition, at least part of the essential gene(s) of the *msd* locus is (are) contained within the smallest region identifiable between the proximal breakpoints of those deletions that delete *msd* (e.g., complementation group Ai) and those that leave it intact but delete *hsdr-1* (e.g., groups A and E). Therefore, considering the positions of *msd* and *hsdr-1* on the current functional/complementation map (Fig. 2), it is evident that one approach that could be used to initiate the molecular analysis of these loci is to clone the breakpoints of certain deletions of the A and Ai complementation groups using probes for DNA loci, such as *Emv-23* (Rinchik *et al.*, 1989) and *D7Was12* (Disteche and Adler, 1984), that map distal to the deletions.

II. MATERIALS AND METHODS

Mice

Inbred strains C3Hf/Rl, 101/Rl, and 129/Rl, as well as closed-colony, outbred stocks of 2A (c^{ch}/c^{ch}) mice and of mice carrying various radiation-induced albino deletion mutations, were bred at the Biology Division of the Oak Ridge National Laboratory. Outbred *Mus spretus* mice were obtained from Verne Chapman (Roswell Park Memorial Institute, Buffalo). The albino deletion mutations are referred to by their complementation-group designations (assigned by Russell *et al.*, 1982), with letters representing originally designated complementation groups and numerals specifying independent deletion mutations within each group (many of which have subsequently been demonstrated to identify independent groups). For example, the designation for the c^{14CoS} deletion is "A1", i.e., one of the six independent Oak Ridge deletions belonging to the original A complementation group. Each albino deletion was maintained in the heterozygous state opposite the chinchilla (c^{ch}) allele of the albino (c) locus. The resulting heterozygotes (e.g., c^{ch}/c^{14CoS} ; hereafter designated $c^{ch}/A1$) are a light-chinchilla color and are readily distinguishable from the full-chinchilla color of their c^{ch}/c^{ch} littermates. Mice homozygous for viable (Va) albino deletions (e.g., c^{24R75M}/c^{24R75M}) and for the C-group deletion (C/C) were identified by albino coat color, as were compound heterozygotes carrying lethal deletions from different complementation groups that complement lethality but not albinism (e.g., A/Bem1 and Ai/Bem1). Fetuses

homozygous for neonatally lethal albino deletions (A/A and E/E), or compound heterozygotes carrying different prenatally lethal albino deletions that complement for prenatal lethality but not for neonatal lethality (e.g., Ai/E), were identified by the absence of eye pigment at Day 14.5 of gestation (with day of vaginal plug being Day 0.5).

Genomic DNA Isolation and Southern-Blot Analysis (Conventional)

High-molecular-weight mouse DNA was prepared either from adult spleen and liver, or from entire 14.5-day fetuses, and was analyzed by Southern blotting and hybridization as previously described (Rinchik, *et al.*, 1989), except that: (1) DNA was prepared for transfer by irradiating gels with 254-nm ultraviolet (UV) light either for 1.5 minutes with a DNA Transfer Lamp (Fotodyne) or with 240 millijoules in a UV Stratalinker (Stratagene); (2) some gels were blotted overnight in 1.5 M NaCl, 0.5 M NaOH; (3) DNA was fixed to the nylon membranes either by baking at 80°C with vacuum for at least 2 hours (when blotted in alkali) or by irradiating dry membranes with 254-nm UV light for 15 seconds with the Fotodyne transfer lamp or with 120 millijoules in the Stratalinker; and (4) some blots were washed at lower or higher stringencies after hybridization with some probes by altering the temperature of the wash (described in figure legends).

Hybridization Probes

The previously derived DNA probes, as well as all new DNA clones isolated and used as probes in this study, are described in Table 2. All new DNA clones used as probes were cloned into either the plasmid vector pBSM13+ or pBLUESCRIPT+ (both from Stratagene), or pGEM4 (Promega) by standard procedures (Ausubel *et al.*, 1989). Purification of cloned DNAs, isolation of insert DNA fragments, and radiolabelling of inserts for use as hybridization probes (by random-hexamer priming) were done using standard procedures (Ausubel *et al.*, 1989; Maniatis *et al.*, 1982).

Large-Fragment/Pulsed-Field-Gel Electrophoresis Analysis

Genomic DNA suitable for megabase restriction mapping was prepared from mouse splenic white cells according to the procedures of Herrmann *et al.* (1987) with slight modifications. Single-cell suspensions were embedded in 0.5-1.0% low-melting-point (LMP) agarose (InCert agarose, FMC BioProducts or UltraPure LMP agarose, BRL) at concentrations of either 10^6 viable cells (as determined by Trypan-blue exclusion; approximately 10 μ g DNA) or 2×10^6 cells per block, and were digested with 0.5 ml of the Proteinase-K (Boehringer-Mannheim) solution per block. The digested blocks were rinsed twice in TE [10mM Tris (pH 7.5), 0.1mM EDTA] at room temperature, and then incubated twice for 60 minutes each at 50°C in 1 ml of TE/PMSF [TE containing 40 μ g/ml of phenylmethylsulfonyl flouride (Sigma)] per block.

For restriction-enzyme analysis, blocks that had been stored in 0.5 M EDTA were first dialyzed in TE (25 ml per block) at 50°C for 1 - 2 hours, and then incubated in 1X restriction buffer without enzyme (0.2-0.3 ml/block) for 30 to 60 minutes. For complete digestion, a DNA/agarose block made from 10^6 cells was transferred into 200 μ l of 1X restriction buffer [as recommended by the manufacturer, but supplemented with acetylated BSA (BRL) at 250 μ g/ml] containing 20 - 50 U of the appropriate restriction enzyme (New England Biolabs), and was incubated at the recommended temperature for 6 - 18 hours. Twice the amount of enzyme was used for DNA blocks made from 2×10^6 cells. After restriction digestion, the blocks were incubated and stored in the Proteinase-K solution until ready for use in pulsed-field gel electrophoresis (PFGE), at which time they were loaded directly into the gel.

PFGE was carried out in one of two commercially available systems: the PulsaphorTM Electrophoresis system with a hexagonal electrode array insert (Pharmacia LKB Biotechnology Inc.) or the CHEF-DRTM II Pulsed Field Electrophoresis System (BioRad). Electrophoresis was carried out with the PulsaphorTM system at 170 volts (6.2 V/cm) in 110-150-ml, 1% agarose (BRL Ultra-Pure) gels submerged in 1X TBE buffer [0.1 M Tris base, 0.1 M boric acid, 2 mM EDTA (pH 8)] maintained at a constant temperature of 10-12°C, and with the CHEF-DRTM II system at 170-200 volts (5.1-6 V/cm) in 120-180 ml 1% agarose gels submerged in 0.5X TBE buffer maintained at a constant temperature of ~14°C. The pulse times and total running times, which varied across experiments, are indicated in the respective figure legends. DNA molecular-weight markers used were *Saccharomyces cerevisiae*

chromosomes [either prepared from strain S288C according to the procedures of Carle and Olson (1985), or purchased from BioRad (strain YNN295) or FMC BioProducts (strain YPH80)] and wild-type (cI857) lambda DNA "ladders" [either prepared according to the procedures of Herrmann *et al.* (1987) or purchased from FMC BioProducts] embedded in LMP-agarose blocks. Blotting of PFGE gels and hybridization of filters were done as described above except that the gels were blotted for at least two days in order to get complete transfer. Yeast chromosomes and lambda ladder markers were visualized by including small amounts of total yeast DNA (strain S288C, 4-6 nanograms) and lambda/*Hind*III DNA (about 0.3-0.5 ng) along with the probe in the random-hexamer labelling reaction. The sizes reported for restriction fragments identified with DNA probes have been averaged from two or more pulsed-field-gel blots, except where indicated in Table 1.

Library Construction

A lambda library was constructed from total *Bam*HI-digested *c24R75M/c24R75M* splenic DNA in *Eco*RI-*Bam*HI-digested λ 2001 (American Type Culture Collection) using standard procedures (Frischauf *et al.*, 1983). Phage-insert ligations were packaged *in vitro* (Gigapack Plus; Stratagene) and a library in *Escherichia coli* NM539 was screened with probe MTY811 using standard procedures (Ausubel *et al.*, 1989; Maniatis *et al.*, 1982).

Cosmid libraries were constructed in a *Sma*I/*Bam*HI-digested c2RB vector (Bates and Swift, 1983) or in a *Bam*HI-digested SuperCos

vector (Stratagene) using standard procedures (Ausubel *et al.*, 1989; Maniatis *et al.*, 1982). The breakpoint-fusion-fragment clone pRN226 was derived from a library of *Bam*HI-digested and size-selected (20-30 kb) A1/A1 14.5-day fetal DNA. The pRN224 clone was isolated from a library (provided by D. Johnson) made from a mixture of 129/R1 and C3Hf/R1 splenic DNA partially digested with *Sau*3A and size-selected for 35-45-kb fragments. Both of these libraries were plated in *E. coli* 490A (*recA*⁻) cells. Additional cosmid libraries screened were: (1) a library constructed in the c2RB vector from 129/R1 DNA partially digested with *Sau*3A and size-selected for 35-45-kb fragments, and plated in DH5 α mcr⁻ (BRL); (2) a library constructed in the SuperCos vector from C3Hf/R1 DNA partially digested with *Sau*3A and size-selected for 35-45-kb fragments, and plated in DH5 α mcr⁻ (BRL); and (3) a library constructed in the SuperCos vector from C3Hf/R1 DNA partially digested with *Sau*3A and size-selected for 35-45-kb fragments, and plated in CPLK (mcr⁻; Stratagene). Cosmid-insert ligations were packaged *in vitro* (Gigapack Gold; Stratagene) using standard procedures (Ausubel *et al.*, 1989; Maniatis *et al.*, 1982), and the libraries were screened following the procedures of Herrmann *et al.* (1987).

A cDNA library made from adult BALB/c liver poly-(A⁺) RNA in the lambda gt11 vector (Clonetech; 1.2-kb average insert size; amplified ~500-2500-fold) was screened with probe RN226.2 by standard procedures (Ausubel *et al.*, 1989). About 150-200, of 1.2×10^6 cDNA clones plated, hybridized with the probe, and 23 cDNA clones were plaque-purified.

Plasmid vectors pBSM13+ and pBLUESCRIPT+ (Stratagene), or pGEM4 (Promega) were used for subcloning inserts.

RNA Isolation and Northern-Blot Analysis

Total cellular RNA was isolated from various tissues using the guanidine-isothiocyanate procedure (Ausubel *et al.*, 1989), and poly-(A+) RNA was purified from total RNA by standard oligo-dT selection techniques (Aviv and Leder, 1971; Ausubel *et al.*, 1989). RNA molecules were size-fractionated by electrophoresis through 1% agarose/MOPS gels containing ~0.8 µg/ml ethidium bromide and either 2.2 M or 0.22 M formaldehyde. After washing the gel in 50 mM NaOH for 30 minutes followed by two consecutive 15-min. washes in 0.1 M Tris-Cl, the RNA was blotted to a nylon membrane in 20X SSC. Immobilization of RNA on the membrane and hybridization analysis were performed as described for DNA analysis.

PCR Amplification and Sequence Analysis of cDNA Inserts

The cDNA inserts of 10 of the plaque-purified cDNA clones (ranging in size from 1.1-1.9 kb) were amplified by the polymerase chain-reaction (PCR) according to Friedman *et al.* (1990), using the lambda-gt11 sequencing primers and a DNA Thermal Cycler (Perkin-Elmer Cetus) PCR apparatus. PCR-amplification conditions were: an initial 5 min. denaturation at 94°C; followed by 30 cycles of 72°C for 3 min. (extension), 94°C for 1 min. 20 sec. (denaturation), and 51°C for 2 min.; and a final extension at 72°C for 7 min. Six of the amplified inserts were subcloned into the pGEM4 vector (Promega) to facilitate DNA sequencing analyses.

The inserts of cDNA clones pRN273 and pRN274 were sequenced by the chain-termination method (Sanger *et al.*, 1977). Initially, PCR-amplified and gel-isolated inserts were directly sequenced using 5'-end-labelled primers, according to Kusakawa *et al.* (1990), except that the Sequenase Version 2.0 (US Biochemical) DNA polymerase and chain-termination mixes were used. After subcloning of the PCR-amplified inserts, sequencing reactions were performed as per the Sequenase protocol, except that ^{32}P -dCTP was used in the labelling reaction instead of ^{32}P -dATP. Oligonucleotides used as primers for cDNA sequencing and PCR amplification were synthesized with a PCR-mate DNA Synthesizer (Applied Biosystems). The sequences of the lambda gt11 sequencing primers and the primers for the T7 and SP6 promoters of the pGEM4 vector were obtained from the manufacturer (Promega). The sequences of the primers corresponding to pGEM4 vector sequences located internal to the promoters and of primers corresponding to internal cDNA sequence are shown in Table 3.

The differentially terminated DNA chains were size-fractionated by polyacrylamide-gel electrophoresis (PAGE) in a 100 cm sequencing apparatus (Haake-Buchler), at 50 watts for 8-18 hours. The sequencing gels were dried at 80°C under vacuum, and sequence ladders were visualized by autoradiography. The individual nucleotide sequences obtained from each primer were melded into a single sequence contig, and the assembled cDNA sequence was analyzed for restriction-enzyme sites, open-reading frames, encoded amino-acid sequence, and homologies with DNA sequences contained in the GenBank and EMBL

databases, using the Wisconsin Genetics Computer Group software package (Devereux *et al.*, 1984).

III. RESULTS

The *D7Was12* and *Emv-23* Loci Map Distal to Some of the Deletions of the A and Ai Complementation Groups

It is evident from the functional map derived from complementation analyses (Fig. 2) that the proximal breakpoints of the deletions of the A and Ai complementation groups may occur near or within the hepatocyte-specific developmental regulation-1 (*hsdr-1*) and mesoderm-deficient (*msd*) loci, respectively. Therefore, molecular analyses of these loci can be initiated by identifying and cloning deletion-breakpoint-fusion fragments resulting from the A and Ai deletions. The cloning of the tyrosinase gene at the *c*-locus (Kwon *et al.*, 1987, 1988; Ruppert *et al.*, 1988) and the analysis of the A and Ai mutations (among others) with the tyrosinase cDNA probe has proven, at the molecular level, that the *c* locus is completely deleted by these mutations (Kwon *et al.*, 1987; Rinchik *et al.*, in prep.). The deletion of the *c* locus by all of the A and Ai deletions, as revealed by the results of hybridization of the cDNA probe MTY811 (Table 2) to A/A and Ai/E DNAs digested with *Bam*HI, is demonstrated in Fig. 3A. Because Ai/Ai DNA cannot be readily obtained (the embryos die at Day 11.5) and because the *c* locus was found to be deleted by the neonatally lethal E deletion, deletions of the Ai complementation group were balanced opposite the E deletion. Thus, the *c* locus can be physically mapped with respect to the Ai deletions simply by looking for the presence or absence, in DNAs of 14.5-day E/Ai fetuses, of the wild-type restriction fragments detectable with probe MTY811. In

addition to the *c* locus, some, but not all, of these deletions have been shown to remove the *Emv-23* locus at the molecular level (Rinchik *et al.*, 1989), and the anonymous locus *D7Was12* has been mapped distal to the A1 (*c*^{14CoS}) deletion but within the *c*^{3H} deletion (Disteche and Adler, 1984). [Figs. 1 and 2 summarize the current deletion-mapping profile of these loci.] Therefore, the A and Ai deletions were further analyzed in this study to determine whether probes for the *Emv-23* and *D7Was12* loci, which map distal to some of these deletions, can be used to clone any of their breakpoints.

As the *D7Was12* locus had been mapped with respect to only one of the A deletions and none of the Ai deletions, it was necessary to deletion-map it in more detail. The position of the *D7Was12* locus on the *c*-region functional map was refined by hybridizing the 12A probe (Table 2) to a Southern blot containing *Bam*HI-digested DNAs of the genotypes indicated in Fig. 3A. As with the *c* locus, deletion of the *D7Was12* locus by the neonatally lethal E deletion (Fig. 3A) permits the locus to be mapped with respect to the Ai deletions simply by looking for the presence or absence in E/Ai DNAs of the wild-type restriction fragments detectable with probe 12A. This analysis revealed that the 4.25-kb and 10.8-kb *Bam*HI fragments defining the *D7Was12* locus are not deleted by the A1, A2, A4, Ai1, and Ai2 deletions, but do map within the remaining A and Ai deletions (Fig. 3A). This deletion-mapping profile is identical to that found previously for the *Emv-23* locus (Rinchik *et al.*, 1989).

Physical Linkage of the *c*, *D7Was12*, and *Emv-23* Loci on a Long-Range Restriction Map of a 1.7-Megabase Region of Chromosome 7

Because it was not possible to detect any of the A1, A2, A4, Ai1, and Ai2 deletion breakpoints with the 12A and 23.3 probes using conventional restriction-enzyme analysis and Southern-blotting techniques (data not shown), the segment of wild-type genomic DNA containing the *c*, *D7Was12*, and *Emv-23* loci was analyzed with large-fragment (PFGE) techniques in an attempt to physically link the *c* locus to the *Emv-23* and/or *D7Was12* loci on a long-range map. If successful, this analysis should then permit the physical localization of the deletion breakpoints of interest.

Wild-type [outbred stock 2A (*c^{ch}/c^{ch}*) and inbred strains C3Hf/Rl (+/+) and 101/Rl (+/+)] genomic DNAs (from splenic white cells) digested singly and doubly with the enzymes *Bss*HII, *Sac*II, *Nae*I, *Sma*I, and *Xho*I were subjected to PFGE, blotted to nylon, and hybridized consecutively with the DNA probes MTY811, p12A, and 23.3 (identifying the loci *c*, *D7Was12*, and *Emv-23*, respectively; Table 2). Based on the results of this analysis (Fig. 4 and Table 4), a long-range restriction map of about 1700 kb of wild-type DNA surrounding these loci was constructed (Fig. 5). The enzymes *Sal*I, *Eag*I, *Not*I, *Mlu*I, and *Nru*I were also used in similar analyses of stock 2A DNA with the 12A and MTY811 probes. The DNA fragments generated by these enzymes in this region of the genome were either not detectable under the PFGE conditions used, or were too faint to be detected reproducibly.

The *c*, *D7Was12*, and *Emv-23* loci map to the same 1080-kb *Bss*HII fragment, but to distinct *Sac*II fragments of 820-, 550-, and 160-kb, respectively. Both probes MTY811 and 12A identify the same 920-kb *Nae*I fragment, whereas probe 23.3 detects a distinct *Nae*I fragment of 160 kb. These results strongly suggest that the order of these cloned DNA loci in chromosome 7 is: centromere- *c* - *D7Was12* - *Emv-23* - telomere. Conclusive evidence for this order subsequently came from the detection of structurally altered *Sac*II fragments in animals heterozygous for the A1, A2, and Ai1 deletions with probe 12A but not with probe 23.3 (see below).

All of the *Sac*II and *Nae*I fragments shown in the long-range map (Fig. 5) are placed next to each other, because the sum of the 370-kb *Nae*I-*Sac*II fragment identified with probe MTY811, the 550-kb *Sac*II fragment detected with probe 12A, and the 160-kb *Sac*II fragment (and also *Nae*I fragment) identified with probe 23.3 is 1080 kb, the approximate size of the *Bss*HII fragment recognized by all three probes. The identification of partially digested fragments (data not shown) close to the sizes predicted for these enzymes from the map lends further support to the map shown in Fig. 5. This does not rule out the possibility, however, that two or more *Sac*II and/or *Nae*I sites may exist very close to one another (e.g., as in a CpG island), resulting in additional small fragments that are not detectable and/or partially digested fragments that are impossible to resolve from each other by PFGE. Furthermore, all of these enzymes are methylation-sensitive [i.e., digestion is inhibited by methylation of CpG dinucleotides present in their recognition sequences (Brown and Bird,

1986)]. Therefore, some restriction sites/fragments may be present, but undetectable, in these experiments.

Partially digested fragments are obtained from this genomic region with *Sma*I, *Xho*I, and *Nae*I (Fig. 4 and Table 4) even with excess amounts of enzyme, presumably due to methylation of their restriction sites (Brown and Bird, 1986). According to the manufacturer (New England Biolabs), the partial digestion obtained with *Nae*I may also be due to different site-specific rates of cutting, independent of methylation. Additional and very faint *Nae*I fragments are present, but not consistently detectable, with the probes used; therefore, these *Nae*I sites are not included on the map. The partial *Sma*I and *Xho*I fragments between 350 kb to 755 kb, identified with probe 12A, were not well-resolved in the PFGE blot shown in Fig. 4, but the individual fragments listed in Table 4 have been clearly resolved in other PFGE runs (e.g., see Fig. 6D below). Both probes 12A and 23.3 recognize a 240-kb *Xho*I fragment and the same series of partially digested fragments, indicating that both the *D7Was12* and *Emv-23* loci are contained within the smallest 240-kb *Xho*I fragment obtained. Since these loci are separated by a *Sac*II site, double digestion with *Xho*I and *Sac*II indicates that the *D7Was12* and *Emv-23* loci reside within adjacent 140-kb and 100-kb *Xho*I-*Sac*II fragments, respectively (Table 4 and Figs. 4 and 5). Only the smallest *Xho*I fragment of 240 kb is shown on the map, since this enzyme was not used in later analyses of deletion chromosomes.

The partially cut *Sma*I sites shown in the long-range map (Fig. 5) were placed relative to the distal *Sac*II site of the 550-kb *Sac*II fragment containing the *D7Was12* locus based on the following rationale. The

series of partially digested fragments of 100, 140, 350, 400, and 550 kb in length, identified with probe 12A in DNA doubly digested with *Sma*I and *Sac*II, are also detected in DNA digested with *Sma*I alone (Table 4 and Fig. 4). Because the 12A probe recognizes a completely digested 550-kb *Sac*II fragment, the entire partially digested series of *Sma*I-*Sac*II fragments clearly identified with 12A (100, 140, 350, 400, and 550 kb) must be located within the 550-kb *Sac*II fragment, and are the result of partial *Sma*I digestion within this fragment. The distal *Sma*I site of the smallest *Sma*I fragment (100 kb) detected by probe 12A must be a partially cut site, since some of the fragments identified in *Sma*I-digested DNA (i.e., the 430-, 665-, and 850-kb fragments) with this probe are not present in *Sma*I-*Sac*II-digested DNA, and since probe 23.3 also recognizes fragments of these sizes (Fig. 4 and Table 4). Taken together, this information indicates that a *Sma*I site must be located either very near or 40 kb proximal to the distal *Sac*II site of the 550-kb *Sac*II fragment (the difference between the smallest two *Sma*I fragments of 100 and 140 kb detected with probe 12A is 40 kb). If this *Sma*I site were located 40 kb proximal, then some partially digested *Sma*I fragments other than those actually observed (Table 4) should have been detected with either probe 12A or 23.3. In contrast, the identification, with both probes 23.3 and 12A, of partially digested 430-, 665-, and 850-kb *Sma*I fragments, is consistent with a *Sma*I site coincident with the distal *Sac*II site (i.e., with adjacent 100-kb and 320-kb *Sma*I fragments, the smallest fragments detected with the 12A and 23.3 probes, respectively) as shown in Fig. 5. The only partially digested *Sma*I fragment that was expected to contain both the *D7Was12* and *Emv-23* loci based on the map, but that was

undetectable with either probe, is a 460-kb fragment, comprised of the 140-kb and 320-kb *Sma*I fragments identified by the 12A and 23.3 probes, respectively. This is not surprising, in light of the sensitivity limitations inherent in conventional blotting and hybridization procedures of PFGs, and the differential methylation of *Sma*I sites within this region, as evidenced by the intensity differences observed between different *Sma*I fragments with these probes (Fig. 4; see also Fig. 6D). A *Sma*I fragment of this size was, in fact, weakly detected with both probes in DNAs from short-term cultures of primary embryonic fibroblasts (data not shown).

The identification of this "anchor" *Sma*I/*Sac*II site and mapping of *D7Was12* between this site and the next *Sma*I site located proximal to it, permitted the partially digested *Sma*I sites contained within the 550-kb *Sac*II fragment to be positioned on the map at successively greater distances proximal to the anchor site (Fig. 5), based on the *Sma*I-*Sac*II fragment sizes obtained with probe 12A (Table 4). Finally, the *Sma*I fragments recognized by the tyrosinase cDNA probe MTY811 (Table 4) all fall within the 370-kb *Bss*HII-*Sac*II fragment containing the *c* locus, but are not indicated on the map since it was not possible to orient them with respect to each other with the DNA probes used.

**Localization of the Breakpoints of Several Deletions (Complementation
groups A and Ai) Defining the *hsdr-1* and *msd* Loci
On the Long-Range Map**

The long-range mapping results, along with the fact that the entire *c* locus is deleted in all of the deletions of the A and Ai complementation groups (Figs. 2 and 3), indicate that the A- and Ai-group deletions not deleted for the *D7Was12* and *Emv-23* loci (Figs. 3A and 2) should have distal breakpoints located between the *c* and *D7Was12* loci and, therefore, within the 1080-kb *Bss*HII fragment containing all three loci (Fig. 5). As predicted, PFGE analyses of *Bss*HII-digested *c^{ch}/A1*, *c^{ch}/A2*, *c^{ch}/A4*, *c^{ch}/Ai1*, and *c^{ch}/Ai2* DNAs with both probes 12A and 23.3 did reveal size-altered *Bss*HII fragments (presumed to be deletion-breakpoint-fusion fragments) in all the DNAs (Fig. 6A). In subsequent PFGE analyses of *Sac*II-digested mutant DNAs, probe 12A, but not probe 23.3, identified size-altered *Sac*II fragments in *c^{ch}/A1*, *c^{ch}/A2*, and *c^{ch}/Ai1* DNAs, but not in *c^{ch}/A4* and *c^{ch}/Ai2* DNAs (Fig. 6B and C). These results demonstrate that the distal breakpoints of the A1, A2, and Ai1 deletions lie within the 550-kb wild-type *Sac*II fragment containing the *D7Was12* locus (See Fig. 7). The failure of probe 12A to detect *Sac*II rearrangements in *c^{ch}/A4* and *c^{ch}/Ai2* DNAs indicates that, based on the long-range map, the distal breakpoints of these deletions must be situated somewhere within the 370-kb wild-type *Bss*HII-*Sac*II fragment recognized by probe MTY811, but distal to the *c* locus (Fig. 7). The *Sac*II results, therefore, serve to distinguish among several of the deletions of the A and Ai complementation groups at the

molecular level by identifying differences in the distal extents of the individual deletions.

The A4 breakpoint was localized more precisely, since the size of the rearranged *Bss*HII fragment found in *c^{ch}/A4* DNA with both probes 23.3 and 12A (875 kb; Fig. 6A) is less than the wild-type fragment size (1080 kb). Because the distance from the distal *Bss*HII site of the 1080-kb wild-type fragment to the proximal *Sac*II site of the 550-kb fragment (710 kb) is 165 kb less than the size of the 875-kb rearranged *Bss*HII fragment (Fig. 5), the maximum distance the A4 breakpoint can be from the proximal *Sac*II site of the 550-kb *Sac*II fragment is about 165 kb (Fig. 5). Also, since the *c* locus is comprised of 70 kb (Ruppert *et. al.*, 1988) at an unknown position within the 370-kb *Bss*HII-*Sac*II fragment, the A4 breakpoint may be as far as 300 kb from the *c* locus (Table 2).

The position of the Ai2 breakpoint cannot be refined in this manner, since the size of the rearranged 1440-kb *Bss*HII fragment identified with probe 23.3 in *c^{ch}/Ai2* DNA (Fig. 6A) is larger than the wild-type fragment. Therefore, the Ai2 breakpoint may be located anywhere within 300 kb proximal to the *Sac*II site of the 370-kb *Bss*HII-*Sac*II fragment (Fig. 7), and, like the A4 breakpoint, may be as far as 300 kb from the *c* locus (Table 5).

PFGE analysis of *Sma*I-*Sac*II digested DNAs of the genotypes indicated in Fig. 6D with probe 12A made it possible to map the breakpoints of the A1, A2, and Ai1 deletions into specific *Sma*I fragments located within the 550-kb *Sac*II fragment (Fig. 7). In order to eliminate the wild-type series of partially digested *Sma*I-*Sac*II fragments (i.e., 100, 140, 350, 400, and 550 kb) obtained from the *c^{ch}* chromosome, the A and Ai

deletions were balanced with the Bem1 deletion, which removes both the *c* and *D7Was12* loci (see Fig. 7). As a result, only the *Sma*I-*Sac*II fragments present in the chromosomes carrying the A and Ai deletions are recognized by probe 12A. The point at which the series and intensity of partially digested *Sma*I-*Sac*II fragments detected with probe 12A in mutant DNA deviates from the wild-type pattern identifies the *Sma*I fragment on the long-range map (Fig. 7) that contains the breakpoint of the deletion. The position of the breakpoint can be defined even more precisely, for some of the deletions, by taking into account the size of the structurally altered *Sma*I fragment. For example, in *Sma*I-*Sac*II-digested A1/Bem1 DNA hybridized with probe 12A, the wild-type 100-kb and 140-kb fragments are present; the wild-type 350-, 400-, and 550-kb fragments are absent; and three structurally altered fragments of 255-, 320-, and 435-kb are present (Fig. 6D). These results indicate that the A1 breakpoint occurs within the 210-kb wild-type *Sma*I fragment defined by the proximal *Sma*I sites of the 140-kb and 350-kb *Sma*I-*Sac*II fragments detected with probe 12A (see Fig. 7). Furthermore, the A1 breakpoint can be localized to within a 115-kb segment at the distal end of this 210-kb *Sma*I fragment (Fig. 7) by subtracting the largest wild-type *Sma*I-*Sac*II fragment present in A1/Bem1 DNA (140 kb) from the smallest rearranged fragment detected (255 kb). As a result, the A1 breakpoint can be localized 40-250 kb proximal to the *D7Was12* locus (Table 5). By following the same rationale, the A2- and Ai1-deletion breakpoints can also be localized on the long-range map (Fig. 7).

**Molecular Analysis of the Viable Albino Deletions c^{24R75M} and $c^{16DFiOD}$
and Localization of Their Distal Breakpoints in the Vicinity of
the Distal Breakpoints of the A1, A2, and Ai1 Deletions**

As the distances of the A- and Ai-deletion breakpoints from the *D7Was12* locus were potentially large (Table 5), two viable deletions (c^{24R75M} and $c^{16DFiOD}$; from complementation group Va) were also studied in detail to determine whether they could be used to generate additional probes located near the distal A- and Ai-deletion breakpoints.

The c^{24R75M} and $c^{16DFiOD}$ Deletions Break Within the *c* Locus

The c^{24R75M} and $c^{16DFiOD}$ deletions were recently found to break within the *c* locus, but not to delete the *Emv-23* locus or any of the DNA loci proximal to *c* (Rinchik *et al.*, in preparation). The disruption of the *c* locus by these deletions is illustrated by the results of hybridization of probe MTY811 to *Bam*HI- or *Eco*RI-digested c^{24R75M}/c^{24R75M} and $c^{16DFiOD}/c^{16DFiOD}$ DNAs (Fig. 3; also Rinchik *et al.*, in prep). Deletions and rearrangements of exon-containing fragments are observed with both mutant DNAs. A 16.5-kb wild-type fragment and a 13.7-kb size-altered *Bam*HI fragment are recognized in c^{24R75M}/c^{24R75M} , whereas the wild-type fragments of 18.5 and 26.5 kb are not detected (Fig. 3A). Probe MTY811 detects *Eco*RI fragments (listed in the 5'-3' order determined by Ruppert *et al.*, 1988) of 4.7, 13.8 (10.5 in c^{ch}/c^{ch} DNA), 8.0, 0.7, and 8.0 kb (i.e., two 8-kb fragments are detected) in wild-type DNA (Fig. 3B). However, in $c^{16DFiOD}/c^{16DFiOD}$ and c^{24R75M}/c^{24R75M} DNAs, the 5'-most wild-type fragments of 4.7 and 13.8-kb are missing. Additionally, a size-

altered *Eco*RI fragment of 10 kb is present in *c*^{16DFiOD}/*c*^{16DFiOD} DNA, probably due to rearrangement of the 13.8-kb fragment.

These findings indicate that both the *c*^{24R75M} and *c*^{16DFiOD} deletions have breakpoints occurring within the tyrosinase gene, delete the 5' portion of the gene, and extend into 5' flanking DNA. As a result, molecular "jumps" across the region of DNA spanned by these deletions can be accomplished simply by cloning the above deletion-breakpoint-fusion fragments with the tyrosinase cDNA probe MTY811. However, the available evidence did not indicate whether the *c*^{24R75M} and *c*^{16DFiOD} deletions extend proximally or distally. If they are found to extend proximally, they will not be useful for providing molecular access to *hsdr-1*, as they have been shown not to delete any of the DNA loci already cloned from the *hsdr-1* to *c* interval (Rinchik *et al.*, in preparation). If, on the other hand, either of these viable deletions extends distally into the vicinity of the distal breakpoints of the A and Ai deletions, a breakpoint-fusion-fragment probe from the distal end of the viable deletion will facilitate the cloning of breakpoint-fusion fragments associated with the A and Ai deletions, and, thereby, the molecular analysis of the *hsdr-1* and *msd* loci.

Localization of the Breakpoints of the *c*^{24R75M} and *c*^{16DFiOD} Deletions on the Long-Range Map

In an attempt to determine whether the *c*^{24R75M} or *c*^{16DFiOD} deletions extend distally, they were analyzed by Southern blotting with probe 12A. The results of hybridization of probe 12A to *Bam*HI-digested *c*^{24R75M}/*c*^{24R75M} and *c*^{16DFiOD}/*c*^{16DFiOD} DNAs (Fig. 3A) had revealed that

the *D7Was12* locus is not deleted in these mutations. Therefore, in order to determine whether these deletions have breakpoints located between the *c* and *D7Was12* loci and, if so, to localize their breakpoints on the long-range map, these deletions were analyzed further by PFGE. As expected from the previous long-range mapping analysis, rearranged *Bss*HII fragments were detected in both *c*^{24R75M}/*c*^{24R75M} and *c*^{16DFiOD}/*c*^{16DFiOD} DNAs with probe 12A (Fig. 6A). Fortunately, size-altered *Sac*II fragments were also detected in both DNAs with probe 12A (Fig. 6C). This indicates that both deletions do, in fact, extend distally in the direction of the distal breakpoints of the A1, A2, and Ai1 deletions.

Subsequent hybridization of probe 12A to the *c*^{24R75M}/*c*^{24R75M} and *c*^{16DFiOD}/*c*^{16DFiOD} DNAs, doubly digested with *Sma*I and *Sac*II, identified patterns of wild-type and rearranged fragments (Fig. 6D) indicative of distal breakpoints located as shown in Fig. 7. This analysis revealed that the *c*^{24R75M} breakpoint may be located in the vicinity of the distal breakpoints of the A1 and A2 deletions (0-210 kb), and that the *c*^{16DFiOD} breakpoint may also be near these breakpoints (0-285 kb), as well as the Ai1 breakpoint (0-395 kb; Table 5). Moreover, the *c*^{24R75M} and *c*^{16DFiOD} breakpoints are potentially closer to these A- and Ai-deletion breakpoints than are the *D7Was12* and *c* loci, respectively (Table 5).

Cloning a Deletion-Breakpoint-Fusion Fragment (pRN207) from

***c*^{24R75M}/*c*^{24R75M} DNA**

It is evident, from the results of these large-fragment analyses, that cloning deletion-breakpoint-fusion fragments detected by the

tyrosinase cDNA probe in *c*^{24R75M}/*c*^{24R75M} and *c*^{16DFiOD}/*c*^{16DFiOD} DNAs will yield additional probes that may map closer to the distal breakpoints of some of the A and Ai deletions (Fig. 7). Subsequently, these probes may facilitate the cloning of breakpoint-fusion fragments associated with the A and Ai deletions. This, in turn, would provide DNA probes, derived from portions of the fusion fragments located proximal to the breakpoints, that should be useful for molecular analyses of the *hsdr-1* and *msd* loci.

In light of these possibilities, the 13.7-kb *Bam*HI breakpoint-fusion fragment of *c*^{24R75M}/*c*^{24R75M} DNA was cloned into λ 2001 and subcloned into pBSM13+ (Stratagene), yielding the clone pRN207. A 0.7-kb *Bam*HI-*Hind*III end-fragment (RN207.1) and a 1.25-kb *Bam*HI-*Eco*RI end-fragment clone (pRN204) were generated by recircularization of pRN207 DNA after digestion with *Hind*III and *Eco*RI, respectively. The 13.7-kb RN207 fragment was restriction-mapped by partial digestion with *Eco*RI, *Hind*III, *Sma*I, and *Sfi*I using the end-fragments (RN207.1 and RN204) as hybridization probes. A restriction map of the RN207 fragment, constructed from the partial-digestion results and complete digestion data, is shown in Fig. 8.

**A Distal-End Clone (pRN204) Derived From the Rearranged Fragment
RN207 Maps Distal to the A1 Deletion and Hybridizes To
Size-Altered Fragments in A1/A1 DNA**

To determine which of the two end-clones was the desired distal end-fragment of the 13.7-kb *Bam*HI fragment, both were hybridized to

Southern blots containing *Bam*HI-digested DNAs from A/A, C/C, E/E, and E/Ai fetuses. Based on the long-range-mapping results (above), the distal end-fragment may map distal to one or more of the A and Ai deletions, whereas the proximal end-fragment will map within all these deletions (see Fig. 1). The 0.7-kb end-fragment (probe RN207.1) contains repetitive sequences (indicated by hybridization to DNA throughout the length of each lane) and identifies no distinct genomic DNA fragments (data not shown). The unique-sequence, 1.25-kb end-fragment (probe RN204) fails to hybridize to a *Bam*HI fragment in A3/A3, A5/A5, A6/A6, E/Ai3, C/C, and E/E DNAs, but does detect a wild-type 15.7-kb *Bam*HI fragment in A2/A2, A4/A4, E/Ai1, E/Ai2, and *c*^{16DFiOD}/*c*^{16DFiOD} DNAs (Fig. 9A). RN204 also hybridizes, as expected, to the 13.7-kb rearranged *Bam*HI fragment, from which it was derived, in *c*^{24R75M}/*c*^{24R75M} DNA; and, fortuitously, to a size-altered 21.5-kb *Bam*HI fragment in A1/A1 DNA (suggestive of another deletion breakpoint; Fig. 9A). These results indicate that the locus (designated *D7Rn1*) recognized by probe RN204 maps distal to deletions A1, A2, A4, Ai1, Ai2, and *c*^{16DFiOD} (Fig. 10), which, in turn, indicates that RN204 is the distal end-fragment of the 13.7-kb breakpoint-fusion fragment, RN207. Because RN204 is the distal end of RN207, and because a 0.7-kb *Bam*HI-*Hind*III fragment is located at the other end of RN207 on the restriction map (Fig. 8), the 0.7-kb RN207.1 fragment must be from the proximal end of RN207.

In an attempt to identify other size-altered restriction fragments in A1/A1 DNA with probe RN204, and, thereby, to lend further support to the possibility that the 21.5-kb size-altered *Bam*HI fragment contains the A1 deletion breakpoint, A1/A1 DNA was digested with several other

standard restriction enzymes, blotted, and hybridized with the probe. A2/A2 DNA was also analyzed with several enzymes, in an attempt to detect the A2 breakpoint [which maps within the same 210-kb *Sma*I fragment as do the *c*^{24R75M} and A1 breakpoints (Fig. 7)]. No aberrant restriction fragments could be found in A2/A2 DNA with any of the enzymes *Bam*HI, *Eco*RI, *Kpn*I, *Pst*I, *Hind*III, and *Bst*EII (data not shown). On the other hand, A1/A1 DNA yielded other size-altered fragments, in addition to the 21.5-kb *Bam*HI fragment, with the enzymes *Kpn*I and *Bst*EII (data not shown). This analysis strongly suggests that the RN204 probe is detecting the A1 deletion breakpoint, but cannot rule out the remote possibility that the probe may be detecting *Bam*HI, *Kpn*I and *Bst*EII RFLVs which had arisen in the *c*^{ch}/A1 stock. Ultimate proof that these size-altered restriction fragments are, indeed, deletion-breakpoint fusion fragments caused by the A1 deletion will require the cloning and analysis of one of the fragments.

**Cloning a Deletion Breakpoint-Fusion Fragment (pRN226) From
c^{14CoS}/*c*^{14CoS} (i.e., A1/A1) DNA With Probe RN204**

The discovery of size-altered fragments, with three different enzymes, in A1/A1 DNA using RN204 as a probe is strong evidence that these are A1-breakpoint fusion fragments. It is, therefore, highly likely that cloning one of these fragments will provide DNA probes from a locus situated proximal to the A1 deletion. By definition, from the functional map of the *c* region (see Figs. 1 and 2), this DNA locus will map either within the *hsdr-1* locus or between *hsdr-1* and *msd*.

Therefore, the 21.5-kb size-altered *Bam*HI fragment was cloned from A1/A1 DNA into the c2RB cosmid vector, using RN204 as a probe. The insert of the cosmid clone consists of the 21.5-kb fragment concatamerized to a 15-kb *Bam*HI fragment originating from a different location within the genome. The 21.5-kb fragment was gel-isolated by electroelution and subcloned into the pBSM13+ plasmid, yielding the clone pRN226. End-fragments of the RN226 insert were then subcloned by recircularization of pRN226 after digestion with *Eco*RI or *Hind*III. In addition to a 1.25-kb *Bam*HI-*Eco*RI end-fragment (expected to be obtained as a reisolate of clone pRN204, the distal end of the RN207 insert; Fig. 8), a 160-bp *Bam*HI-*Hind*III fragment (probe pRN247, presumed to be the proximal end-fragment) was derived. The locus identified by probe RN247 is designated *D7Rn2*. The 21.5-kb RN226 fragment was restriction-mapped by partial digestion with *Eco*RI, *Hind*III, *Sma*I, and *Sfi*I using the RN204 and RN247 end-clones as hybridization probes. A restriction map of RN226, constructed from the partial-digestion results and complete digestion data, is shown in Fig. 8.

A Proximal-End Clone (RN247) of the A1-Breakpoint-Fusion Fragment (RN226) Identifies a Locus (*D7Rn2*) That Maps Proximal to Deletions A1 and A5 But Within the Other Deletions That Define *hsdr-1*

Hybridization of the end-probe RN247 to DNAs containing the A, C, E, and Ai deletions revealed that the *D7Rn2* locus identified by the probe does indeed map proximal to the A1 deletion: the probe recognizes the 21.5-kb size-altered *Bam*HI fragment in A1/A1 DNA, as expected, as well

as an 8.9-kb wild-type fragment present in A5/A5 and C/C DNAs but absent from A2/A2, A3/A3, A4/A4, A6/A6, E/E, and each of the Ai/E DNAs (Figs. 9A and 11). This demonstrates conclusively that the 21.5-kb *Bam*HI fragment found in A1/A1 DNA is a deletion breakpoint-fusion fragment, which results from disruption by the A1 deletion of the 15.7-kb and 8.9-kb wild-type *Bam*HI fragments detected by probes RN204 and RN247, respectively. Therefore, the *D7Rn1* locus maps distal to but near the distal breakpoints of the *c^{24R75M}* and A1 deletions, and the *D7Rn2* locus maps proximal to, but near, the proximal A1 breakpoint.

The *D7Rn2* locus was subsequently mapped with respect to the rest of the albino deletions that, like the A deletions, have proximal breakpoints located between the *msd* and *hsdr-1* loci (i.e., deletions of the Di, Dp, and Fp complementation groups; Fig. 1). Hybridization of probe RN247 to *Eco*RI-digested DNAs containing these deletions revealed that all of them delete the *D7Rn2* locus (Fig. 9B). Of all the deletions with proximal breakpoints defining *hsdr-1*, therefore, only the A1 and A5 deletions fail to delete the *D7Rn2* locus, demonstrating that the locus must be distal to the *msd* locus and proximal to or within the *hsdr-1* locus (shown in Fig. 11). Moreover, since the A1 breakpoint is detectable with probe RN247, whereas no size-altered *Bam*HI fragment is found in A5/A5 DNA with the probe, the proximal A5-deletion breakpoint must map distal to the A1 breakpoint.

**Probes (RN226.1 and RN226.2) Derived From the *D7Rn2* Locus Hybridize
To Abundant Transcripts in Liver and Kidney That Are Absent
in A1/A1 and A5/A5 Embryos**

Localization of the A1 breakpoint on the Restriction Map of RN226

It is possible that the A1 breakpoint may occur within a gene at the *hsdr-1* locus, and, therefore, that the RN226 fusion fragment may contain exons located proximal to the breakpoint. Positioning the A1 breakpoint on the restriction map of the RN226 insert (Fig. 8) will identify DNA fragments located proximal to the breakpoint that will be useful for screening for evolutionarily conserved DNA sequences, which are often indicative of coding sequences. In order to localize the A1 breakpoint on the restriction map, the deletion-mapping profiles of several DNA probes isolated from within the RN226 insert were obtained by hybridization to Southern blots containing wild-type and mutant DNAs.

Three DNA probes isolated from the RN226 insert [RN226.1, RN226.2, and RN226.3; see Table 2 and Fig. 8] are expected to hybridize to a 4.6-kb *Eco*RI fragment in A1/A1 DNA, according to the restriction map derived for the RN226 insert (Fig. 8). Indeed, when these probes are hybridized to *Eco*RI-digested DNAs, each recognizes a 6.5-kb *Eco*RI fragment in wild-type DNA and a size-altered 4.6-kb *Eco*RI fragment in A1/A1 DNA (e.g., Figs. 9B and 12). Like probe RN247, these probes also identify the 8.9-kb and 21.5-kb *Bam*HI fragments in wild-type and A1/A1 DNAs, respectively (e.g., probe RN226.1; Fig. 12). Probes 226.1 and 226.3 contain repetitive sequences, but hybridize more stringently to the 6.5-kb and 4.6-kb fragments. Therefore, these fragments can be visualized with

these probes by raising the wash temperature (Fig. 12). Probe RN226.3 also hybridizes stringently to a 1.5-kb *Eco*RI fragment that is unlinked to the deletions. The deletion-mapping results imply that the A1 breakpoint is located within the 4.6-kb *Eco*RI fragment from which, according to the restriction map, the probes were derived. Moreover, each probe has the same deletion-mapping profile as the proximal-end probe RN247 (Figs. 9B and 12), indicating that each also maps proximal to the A1 deletion. Based on this information and the restriction map, probe RN226.3 should contain the A1 breakpoint, since it is at the distal end of the 4.6-kb *Eco*RI fusion fragment. As a result, this probe should also detect *Eco*RI fragments located distal to the deletion. Although no distinct hybridization bands indicative of DNA sequences located distal to the A1 breakpoint could be identified with this probe (Fig. 12), it is probable that a repetitive-sequence portion of probe RN226.3 is located distal to the A1 breakpoint or is interfering with the detection of a faint, unique-sequence signal generated by distally located sequences.

To help confirm this breakpoint location, additional probes (i.e., probes 226.4 and 283; Table 2) were isolated from portions of the RN226 insert located distal to probe RN226.3, using the restriction map as a guide (Fig. 8), and the probes were hybridized to the mapping blots. Unfortunately, probe 226.4 contains highly repetitive sequences, and, as a result, was not helpful in confirming the breakpoint location. However, probe 283 was single copy and did, in fact, map distal to the breakpoint (data not shown). Therefore, the A1-deletion breakpoint can be mapped within the segment of the RN226 insert highlighted by the vertically

striped box above the restriction map (Fig. 8), and, most probably, within the 0.55-kb RN226.3 fragment itself (see above discussion).

The Portion of RN226 Located Proximal to the A1 Breakpoint Contains Evolutionarily Conserved DNA Sequences

Since the A1 breakpoint most likely occurs within the 0.55-kb 226.3 fragment, this fragment, as well as fragments derived from the 6.5 kb of RN226 insert located proximal to it (see Fig. 8), were screened for the presence of evolutionarily conserved DNA sequences by low-stringency hybridization to human, cat, dog, and rodent genomic DNA digested with *EcoRI*. Probe RN226.2 cross-hybridizes distinctly to 7.8-kb and 8-kb *EcoRI* fragments in human and mole-rat DNAs, respectively, and faintly to a 3.8-kb fragment in cat DNA (Fig. 13). Both probes RN226.1 (Fig. 13) and RN216 also hybridize faintly to *EcoRI* fragments in DNAs from the human, cat, and dog and intensely to a 9-kb fragment in mole rat. Probe 226.3 fails to hybridize significantly to non-rodent DNA, and 226.5 (Table 2) hybridizes considerably to these DNAs throughout the lengths of the lanes, possibly indicative of cross-species hybridization to repetitive sequences (data not shown). Probe 226.5 also appears to hybridize to 1.7-kb and 1.9-kb fragments in cat and mole rat, respectively, with greater intensity than the background signal in these DNAs.

Probes RN226.1 and RN226.2 Hybridize To Abundant Transcripts in Liver and Kidney of Perinatal, Juvenile, and Adult Mice

The presence of evolutionarily conserved DNA sequences within that portion of the RN226 insert located proximal to the A1 breakpoint

implied that this DNA segment could contain at least a portion of a gene, which is in a position to be a candidate for *hsdr-1*. Therefore, the 226.1 and 226.2 probes were hybridized consecutively to Northern blots containing both total RNA and poly-(A+) RNA (mRNA) from the livers of 18.5-day fetuses and from the livers and kidneys of 8.5-day-old and adult mice, the tissues mainly affected by deletion of *hsdr-1*. Each probe hybridizes to an abundant 1.6-kb transcript in both poly-(A+) RNA and total RNA derived from the livers of 18.5-day fetuses and from the livers and kidneys of 8.5-day-old and adult mice (Fig. 14A, B; Table 6). Probe 226.2 also hybridizes to an abundant transcript of 3.5-kb, a slightly less-abundant transcript of 5.3-kb, and low-abundance transcripts of 6.2-, 7.0-, and 8.9-kb in both liver and kidney (Fig. 14A). The identification of abundant transcripts in liver and kidney with these probes is consistent with the identification of a gene affected by deletion of the *hsdr-1* locus.

Interestingly, the 3.5-kb transcript is labelled only faintly by probe 226.1 in liver, and is labelled only slightly better in kidney (Fig. 14B). 226.1 also detects a 5.3-kb transcript of medium abundance, and a less-abundant transcript of 8.9-kb (Fig. 14B). Also, in some blots, it detects a low-abundance transcript of ~0.12 kb in liver and kidney (data not shown). It should be noted that when the blot hybridized with probe 226.2 and washed at 60°C (Fig. 14A) is washed at the more stringent temperature used in the hybridization with probe 226.1 (65°C), the signal intensity of the 3.5-kb transcript diminishes significantly relative to the 1.6-kb transcript (data not shown). Conversely, however, probe 226.1 does not label the 3.5-kb transcript more intensely at 60°C. The abundant 3.5-kb transcript is, therefore, uniquely identified by probe 226.2. Probes

226.3, 226.5, and 216 were also hybridized to Northern blots containing total RNA from the livers and kidneys of 8.5-day-old and adult mice. Probes 226.3 and 216, like 226.1, detected only the 1.6-kb transcript significantly, and probe 226.5 failed to identify any transcript (data not shown). One interpretation of the differential detection of the 3.5-kb transcript is that alternative splicing of a single gene may be occurring to produce the two major transcripts, which may have distinct functions in the liver and kidney. However, there is also a possibility that two overlapping genes may be present at this locus.

Probes 226.2 and 226.1 were also hybridized to Northern blots containing total RNA (data not shown) and poly-(A+) RNA from other tissues. In the adult brain, probe 226.2 also hybridizes to a major transcript of 3.5 kb and the same higher-MW transcripts as are found in both the liver and kidney, but the level of these transcripts in the brain is only about 25% of the levels found in liver and kidney (Fig. 14F, lanes 18 and 19). The 1.6-kb transcript is also detectable with this probe in the brain, but only at a small fraction of the level seen in liver and kidney. No significant amounts of these transcripts were identifiable with probe 226.2 in the brains of 18.5-day fetuses, or with probe 226.1 in adult or fetal brains (data not shown). The detection of significant amounts of the 3.5-kb transcript, but not the 1.6-kb transcript, in the adult brain (Fig. 14F), but not in 18.5-day fetal brains (data not shown), and the absence of reports of morphological or biochemical abnormalities in the brains of newborn mice homozygous for deletions of *hsdr-1*, implies that the 1.6-kb transcript may not function significantly in the brain, and that the 3.5-kb transcript may function at a later point in development.

The other tissues analyzed are from whole 16.5-day fetuses and from 8.5-day-old and adult mice. No transcripts are visible in spleen or thymus of 8.5-day-old mice with probe 226.2 (Fig. 14F, lanes 14 and 17), but this probe does detect very low levels (clearly visible on overexposure of the blots to X-ray film) of the 1.6-kb transcript in whole 16.5-day fetuses (Fig. 14A, lane 1) and of the 1.6-kb and 3.5-kb transcripts in adult testis, salivary gland, and small intestine (Fig. 14F, lanes 20-22). Partial degradation of the 16.5-day fetal mRNA, as evidenced by the intensely staining smear below the lane-control transcript in Fig. 14E, precludes a conclusion about the presence or absence of the 3.5-kb transcript in this RNA. Probe 226.1 labels the 1.6-kb transcript in whole 16.5-day fetus more intensely than does probe 226.2 (Fig. 14B), barely detects this transcript in testis, salivary gland, and small intestine, and, like probe RN226.2, fails to identify it at all in spleen or thymus (data not shown). These preliminary hybridization results suggest that the gene is expressed at low levels in some tissues (possibly constitutively), and, as expected from the *hsdr-1* phenotype, is considerably induced in liver and kidney.

Despite the absence of the 1.6-kb and 3.5-kb transcripts in thymus and spleen, a novel, low-abundance transcript of 1.4 kb is detectable in thymus with probe 226.1 (but not probe 226.2), and, to a lesser extent, possibly in spleen (data not shown). It is possible that this transcript is also present in liver and kidney, where it may be partially masked by the much more abundant 1.6-kb transcript. The identification of this 1.4-kb transcript predominantly in thymus is consistent with the observation,

in newborn mice homozygous for deletions of *hsdr-1*, of morphological abnormalities of the thymus (Erickson *et al.*, 1968).

Absence of Detectable Transcripts in Liver mRNA from A1/A1 and A5/A5 18.5-Day Fetuses Indicates that the Identified Coding Sequence Is a Candidate Gene for *hsdr-1*

Because probes RN226.1 and RN226.2 (and therefore at least part of the gene identified by these probes) are deleted by all the deletions with proximal breakpoints located between *msd* and *hsdr-1* except A1 and A5 (e.g., Figs. 9B and 12), these two deletions are the only ones of this class that may not disrupt expression of the identified gene. The identification of only normal transcripts in A1/A1 and A5/A5 fetuses would rule out the possibility that the identified coding sequences are part of the gene(s) contributing to the *hsdr-1* phenotype. In fact, Northern-blot hybridization of probes 226.1 and 226.2 to total and poly-(A+) RNA from the livers of 18.5-day fetuses of the genotypes A1/A1, A5/A5, and E/E (null-expression control) failed to identify any transcripts (Fig. 14A and B, lanes 3-5), lending further support to the possibility that the gene identified is the *hsdr-1* gene. The absence of any detectable transcripts in A5/A5 RNA with either probe implies that the gene extends distally at least as far as the A5 breakpoint.

**Liver cDNA Clones Isolated With Probe RN226.2 Have the Same
Hybridization Profiles in Wild-Type and Mutant RNAs
as Probes RN226.1 and RN226.2**

Isolation of cDNA Clones

A λ gt11 cDNA library, made from adult BALB/c liver poly-(A+) RNA, was screened with probe RN226.2, resulting in the identification of numerous cDNA clones. The inserts (ranging in size from 1.1-1.9 kb) of several plaque-purified, positive cDNA clones were amplified by PCR using the lambda gt11 sequencing primers, and the amplified inserts were confirmed as true positives by Southern-blot analysis with the RN226.2 probe. Two general classes of cDNA clones were identified, with one class hybridizing much more intensely to the probe than the other. The intensely hybridizing cDNAs also comprise the class of cDNAs that detect only a single *Eco*RI fragment of 6.5 kb when used as probes (see below). The amplified inserts were used directly as probes and sequencing templates, and several were subcloned into the pGEM4 vector (Promega) to facilitate DNA sequencing analyses.

Hybridization of cDNA Probes to Northern Blots

Several of the cDNA inserts were hybridized to Northern blots containing poly-(A+) and total RNA from wild-type mice and mice homozygous for the deletions affecting *hsdr-1*. As with DNA probes RN226.1 and RN226.2, none of the cDNA probes (e.g., RN271 and RN273, Table 2) hybridized to detectable transcripts in mRNA isolated from the livers of A1/A1 or A5/A5 18.5-day fetuses, although they all recognized an

abundant transcript of 1.6 kb in wild-type mRNA (e.g., Fig. 14C and D). In addition, some of the cDNAs (e.g., pRN273) label the 1.6-kb transcript much more intensely than others (e.g., pRN271; Figs. 14C,D). On the other hand, cDNA RN271 hybridizes to the lower-abundance transcript of 8.9 kb with somewhat greater intensity than cDNA RN273 (Fig. 14C,D). Most of the cDNAs also recognize other transcripts of higher molecular weight (also detected with the genomic DNA probe RN226.2) that are less abundant than the 1.6-kb transcript. None of the cDNAs analyzed labels the 3.5-kb transcript as intensely as does the genomic DNA probe RN226.2 (Fig. 14B), however. This implies that these cDNAs may not contain one or more of the exons that are located within RN226.2 and contribute to this transcript.

**The cDNA Clones Are Of Two Classes and Identify a Structural Locus
(Or Loci) That Is Deleted Or Disrupted By All the Deletions
Defining *hsdr-1*, With the Exception of the A5 Deletion**

The PCR-amplified cDNA inserts were used as hybridization probes on Southern blots containing *Eco*RI-digested DNAs of mice homozygous for the deletions affecting *hsdr-1*. This analysis revealed that, with the exception of the A5 deletion, these deletions either completely or partially delete the exon-containing *Eco*RI fragments normally detected by the cDNA probes (e.g., probes RN271 and RN273; Fig. 9B). The extent of the structural locus identified by these cDNA probes was completely intact in A5/A5 DNA (Fig. 9B). As cDNA RN273 was found to contain a complete open reading frame (see below), the A5

deletion most likely either removes the gene promoter or an essential *cis*-acting regulatory element(s).

The results of this analysis, together with the expression analysis, also identified two classes of cDNA clones: (1) cDNAs that recognize a single *Eco*RI fragment of 6.5 kb in wild-type DNA and a single 4.6-kb breakpoint-fusion fragment in A1/A1 DNA (e.g., cDNA probe RN271; Fig. 9B); and (2) cDNAs that identify several *Eco*RI fragments in wild-type DNA, including a 6.5-kb fragment which is also rearranged to a 4.6-kb fragment in A1/A1 DNA (e.g., cDNA probe RN273; Fig. 9B). cDNAs of the first class label the 1.6-kb transcript much less intensely than cDNAs of the second class. It is interesting that, of the cDNAs analyzed, the one with the largest insert (RN271, 1.9 kb) is of the former class. The sequences of the RN271 insert are, therefore, derived entirely from within a single 6.5-kb genomic *Eco*RI fragment. The significant abundance of this class of cDNAs (3 of 10 analyzed) suggests that they may be functional, despite the complete absence of exons derived from three other *Eco*RI fragments, present in the other class of cDNAs.

The Identified cDNA Clones Encode the Enzyme Fumarylacetoacetate Hydrolase (FAH)

The inserts of two cDNA clones (pRN273 and pRN274, with 1.45-kb and 1.1-kb inserts, respectively), of the type that hybridize to several *Eco*RI fragments in wild-type DNA, were sequenced by the chain-termination method (as described in Materials and Methods). Initially, sequencing reactions were performed with lambda gt11 sequencing

primers, oligonucleotide primers for the T7 and SP6 promoters of the pGEM4 vector, and primers for pGEM4 vector sequences located internal to the promoters (Table 3). Subsequently, primers corresponding to internal cDNA sequence (Table 3) were used to confirm and extend the initial sequence. This analysis determined that the smaller cDNA, RN274, is an abbreviated form of the larger cDNA, RN273. The preliminary sequence of the 1.45 kb of RN273 insert, which contains an entire open-reading frame (ORF; see below), is shown in Fig. 15.

A comparison of this cDNA sequence with DNA sequences deposited in the GenBank and EMBL sequence databases (Devereux *et al.*, 1984) revealed ~80% nucleotide-sequence homology with a recently cloned cDNA encoding human fumarylacetoacetate hydrolase (FAH) (Agsterribe, *et al.*, 1990). However, the RN273 sequence contains an additional 118 nucleotides beyond the putative 5' end of the human cDNA reported by Agsterribe *et al.* (1990). Most of these additional 5'-end sequences also are highly homologous to the 5'-end sequences recently reported for the longer, full-length human *FAH* cDNA isolated by Phaneuf *et al.* (1991).

The 3' end of cDNA RN273 was identified by the discovery of a poly-dT stretch near the SP6 primer. A potential initiation codon was identified near the 5' end, which is in-frame with a potential termination codon found near the 3' end, and an ORF is present between these sites (Fig. 15). This ORF is ~84% homologous in sequence to the ORF reported for the human *FAH* cDNA analyzed by Phaneuf *et al.* (1991). The amino-acid sequences predicted from these ORFs have even greater homology (~88% identity and ~93% similarity of amino acids, after correction of the

mouse sequence for frameshifts with respect to the human sequence; Fig. 16), strongly suggesting that the mouse cDNA also encodes FAH. The 5' untranslated portion of the mouse cDNA sequence ends 12 nucleotides from the 5' end of the full-length human cDNA sequence. Whether or not the mouse cDNA RN273 is full-length, remains to be determined. The reported human cDNA and encoded amino-acid sequences were used as guides in determining the preliminary mouse sequences shown in Figs. 15 and 16.

**The *Fah* (*D7Rn2*) and *D7Rn4* Loci Will Contribute To Molecular Analyses
of the *jdf* Locus and Genetic Loci at the Distal Ends
of the Di4 and Fp1 Deletions**

Hybridization of Probe RN216 to Size-Altered *Sma*I Fragments in A5/A5 DNA Suggests the A5 Breakpoint May Occur Within 16.5 kb Distal To the *D7Rn2* Locus

The A5 deletion belongs to the subgroup of A deletions that extend distally beyond the *Emv-23* locus and help to define the proximal limit of the *jdf* locus (Fig. 11). Therefore, cloning a breakpoint-fusion fragment resulting from the A5 deletion might make the *jdf* locus accessible at the molecular level. To investigate whether this might be possible with a probe from the *D7Rn2* locus, A5/A5 DNA was digested with several additional restriction enzymes, electrophoresed through 0.8% or 0.4% agarose gels, blotted to nylon, and hybridized with probe RN216. This probe, a 1.45-kb *Bam*HI-*Eco*RI fragment, is also derived from the proximal end of the RN226 insert (see Fig. 8 and Table 2). Of the enzymes

used for A5/A5 digestion --*KpnI*, *XhoI*, *SmaI*, *ClaI*, *BstEII*, *EcoRV*, *SacI*, *PvuII*, *BglII*, *AccI*, *HindIII*, and *TaqI*-- only *SmaI* yielded an apparent size-altered fragment (21 kb) visible with probe RN216 (Fig. 17).

Although the partially digested *SmaI* fragments identified with probe RN216 are faintly labelled (Fig. 17), there seems little doubt that the wild-type 16.5-kb fragment is absent and a new size-altered fragment of 21 kb is present in A5/A5 DNA. If this is due to disruption of the 16.5-kb wild-type fragment by the A5 breakpoint, then this breakpoint must be within 16.5 kb of probe RN216, well within reach for cloning a breakpoint-fusion fragment. The identification of a size-altered *ClaI* fragment of 23 kb in A5/A5 DNA may be additional evidence for proximity to the A5 breakpoint (Fig. 17), but the presence of an additional fragment of the same size as the wild-type fragment (17 kb) in this DNA indicates that the 23-kb fragment may be the result of poorer digestion in A5/A5 DNA vs. wild-type. Irrespective of whether or not the A5 breakpoint is within the 16.5-kb *SmaI* fragment, large-fragment (PFGE) analysis with probe RN216 of A5/A5 DNA digested with either *SacII* or *BssHII* has revealed that the proximal A5 breakpoint is located no more than 220 kb distal to the A1 breakpoint (data not shown).

The Proximal End-Clone (RN254) of a Cosmid Clone (RN224) Isolated From The *D7Rn2* Locus Maps Proximal To the *Di4* and *Fp1* Breakpoints and the *Fah* coding sequences.

With the availability of probes RN247 and RN216 at the *D7Rn2* locus, the distal end of the *msd* locus is also potentially accessible at the molecular level. Chromosome-walking proximally from the *D7Rn2* locus

[using both cosmid and yeast-artificial-chromosome (YAC) clones] should eventually yield a DNA clone that maps proximal to all the deletions that break between the *msd* and *hsdr-1* loci (i.e., deletions of the A, Di, Dp, and Fp complementation groups; see Fig. 11). Such a clone would then, by definition, be derived from the distal end of the *msd* locus, as defined by deletion breakpoints. Moreover, the isolation of DNA clones that map proximal to these breakpoints will permit the cloning of the breakpoints (by cloning breakpoint-fusion fragments), and, thereby, the molecular characterization of the developmentally important genes defined by the distal ends of these deletions (see Fig. 1). Isolation of a cosmid clone with a probe from the *D7Rn2* locus will also facilitate the molecular analysis of the *hsdr-1* locus, as well as the cloning of the A5 breakpoint.

To initiate the chromosome walk, a cosmid library of wild-type DNA was screened with the RN216 probe, and a cosmid clone (pRN224) of 46 kb of DNA was isolated. Restriction fragments containing the ends of the RN224 insert, as well as portions of flanking vector sequence, were gel-isolated and subcloned. The proximal and distal ends of the cosmid insert were identified by using the subcloned end fragments as probes to screen a panel of deletions by Southern-blot analysis. This analysis revealed that the end clones pRN254 and pRN259 (Table 2) contain proximal- and distal-end fragments of cosmid RN224, respectively (data not shown). The distal-end probe, RN259, exhibits a deletion-mapping profile identical to that shown by probe RN247 (Fig. 9 and data not shown). Therefore, although this cosmid is not very useful for cloning

the A5 breakpoint, it constitutes at least a 40 kb "step" in the chromosome walk toward *msd*.

Fortunately, the proximal end-probe (RN254) identifies a DNA locus (*D7Rn4*) that maps proximal to, and within 46 kb of, the breakpoints of both the Di4 and Fp1 deletions (Figs. 18 and 19). RN254 recognizes an *EcoRI* RFLV between *Mus musculus* and *M. spretus* DNA, which permitted it to be mapped with respect to the deletions of the Di, Dp, and Fp groups by looking for the presence or absence of the *musculus*-specific fragment in mice carrying a deletion opposite a *spretus* chromosome. This type of mapping eliminates the ambiguity involved in attempting to distinguish between one- and two-copy signal intensity in mice heterozygous for prenatally lethal deletions. Since cosmid pRN224 contains the sites of both the Di4 and Fp1 breakpoints, DNA probes derived from the cosmid insert can be used to clone breakpoint-fusion fragments from DNAs containing these deletions. This will facilitate molecular analyses of the loci, at or near the distal ends of these deletions, involved in prenatal survival and postnatal fitness (*eed*, *fit-1*, *pid*, and *l(7)-1Rn*; see Figs. 1 and 19, and Table 1).

The isolation of cosmid pRN224 also permitted an estimation of the physical extent of the *Fah* locus. The *Fah* locus, as defined by hybridization with the cDNA probe RN273 (which contains a complete ORF encoding the FAH enzyme), is completely deleted by both the Di4 and Fp1 deletions (Fig. 9B), which have proximal breakpoints located at most 46 kb proximal to the *D7Rn2* locus. Moreover, the *Fah* coding sequences end proximal to the proximal A5-deletion breakpoint, probably located no further than 16.5 kb distal to the *D7Rn2* locus (Fig. 17). Therefore, the

physical extent of the *Fah* coding sequences is presently delimited by the proximal breakpoints of the Di4 and Fp1 deletions (on the proximal side of *Fah*) and the proximal breakpoint of the A5 deletion (on the distal side of *Fah*), a DNA segment no greater than ~62 kb (the approximate sum of the 46-kb and 16.5-kb maximum distances discussed above).

Attempts to Access the *msd* Locus at the Molecular Level By Cloning Ai-Deletion Breakpoints

Cloning an Ai-deletion breakpoint is one way of beginning a molecular analysis of the *msd* locus, since an Ai-breakpoint-fusion-fragment clone would contain DNA sequences located within or proximal to the *msd* locus (see Figs. 1 and 2). The proximal end of an Ai-fusion fragment, in conjunction with the *D7Rn4* locus (located proximal to the Di4 and Fp1 deletions, see Fig. 19), would provide molecular approaches to *msd* from both sides of the locus, and would permit the estimation of the physical size of the A-Ai-breakpoint interval. This, in turn, will help to define the extent of the *msd* locus.

An End-Probe (RN202) Derived From a Breakpoint-Fusion Fragment (λ 617-4A1) Cloned From *c*^{16DFiOD}/*c*^{16DFiOD} Maps Within the A1 and A2 Deletions, But Distal to the *c*^{16DFiOD} and Ai1 Deletions

The discovery that the *D7Rn1* locus (for probe RN204) maps distal to the breakpoints of the A2, Ai1, and *c*^{16DFiOD} deletions (Figs. 9A and 10), together with the breakpoint locations identified by the long-range mapping, indicates that the Ai1 breakpoint, and possibly the A2

breakpoint, is closer to the *c*^{16DFiOD} breakpoint than is the *D7Rn1* locus (Fig. 11). Isolation of a probe from a portion of a *c*^{16DFiOD}-breakpoint-fusion fragment located distal to the *c*^{16DFiOD} breakpoint will, therefore, facilitate the cloning of breakpoint-fusion fragments associated with the Ai1 and A2 deletions. Cloning of the Ai1 breakpoint would be of particular significance for studying the *msd* locus, as discussed above.

The feasibility of this approach was explored with the cloning of the 10-kb *Eco*RI breakpoint-fusion fragment (λ clone 617-4A1) identified in *c*^{16DFiOD}/*c*^{16DFiOD} DNA with the MTY811 probe (Fig. 3B; Machanoff and Rinchik, unpublished data). The subcloned ends of this fusion fragment were mapped against the panel of deletions to identify the proximal and distal ends, as had been done for the *c*^{24R75M} deletion (above). A distal-end clone (pRN202; Table 2) was identified that maps distal to the A4, Ai1, and Ai2 deletions but within the other A and Ai deletions (Fig. 20; see also Fig. 19). Therefore, the distal-end clone, pRN202 (locus *D7Rn3*), can be used as a probe in an attempt to identify and clone the Ai1 breakpoint at the molecular level. Despite the possibility that the *D7Rn3* locus may map near the A2 breakpoint, its deletion by A2 precludes its use for identifying the breakpoint at the molecular level, unless additional DNA clones (e.g., cosmids), with sequences located distal to the A2 breakpoint, can be isolated with the RN202 probe.

Searching For the Ai1-Deletion Breakpoint With Probe RN202

To try to identify the Ai1 breakpoint, Ai1/Bem1 DNA was digested with a number of standard restriction enzymes (not including enzymes

with more than one CpG dinucleotide in their recognition sequence), electrophoresed through 0.8% or 0.4% agarose gels, blotted to nylon, and hybridized with probe RN202. The Bem1 deletion, like the E deletion used above, was placed opposite the peri-implantation-lethal Ai1 deletion in order to generate viable albino mice that are hemizygous for the *D7Rn3* locus. Despite the identification of wild-type fragment sizes as large as ~30 kb, no size-altered fragments were detected in Ai1/Bem1 DNA with the restriction enzymes *KpnI*, *BstEII*, *EcoRV*, *BamHI*, *SacI*, *PstI*, *PvuII*, *XbaI*, *EcoRI*, *BglII*, *HindIII*, *MspI*, or *TaqI* (data not shown). An attempt to identify the Ai1 breakpoint by PFGE of Ai1/Bem1 DNA partially digested with *KpnI*, Southern-blotting, and hybridization with RN202, was also unsuccessful, despite the identification of a wild-type 150-kb partially digested *KpnI* fragment in this DNA (data not shown). In light of these results, it was deemed unfeasible to attempt to clone the Ai1 breakpoint with the RN202 probe at this time. Instead, the Ai2 breakpoint became the primary target for molecular cloning, by chromosome walking from a probe isolated from DNA flanking the distal end of the tyrosinase gene.

Attempts to Chromosome Walk to the Ai2-Deletion Breakpoint From the *c* Locus

Long-range mapping in and around the *c* locus has determined that the distal end of the *c* locus maps about 60 kb proximal to the distal end of the 370-kb *BssHII-SacII* fragment containing the locus (G. Kelsey, personal communication; also see Fig. 5). Moreover, large-fragment analysis of the Ai2 and A4 deletions indicated that their distal

breakpoints are located between *c* and the distal end of this 370-kb fragment (Fig. 7). Therefore, the Ai2 and A4 breakpoints must be located within the 60-kb segment of DNA between the distal ends of the *c* locus and this 370-kb fragment (Fig. 19). In light of this finding, a distal chromosome walk, starting from probes isolated from DNA flanking the distal end of the *c* locus (probes 210.3 and Tyr14.EX.9, Table 2), was initiated in an attempt to reach the Ai2 breakpoint molecularly. Four separate cosmid libraries (2.3×10^6 total cosmids, unamplified) were constructed (with *mcr*⁻ cells and packaging extracts) and screened with probes 210.3 and Tyr14.EX.9, in attempts to begin the chromosome walk in cosmids. One positive cosmid clone was isolated with probe 210.3 from a library of 129/R1 DNA. However, the observation that one of the end-fragments of this clone mapped outside of the albino deletions indicated that the clone was a product of DNA rearrangement. Therefore, it was discarded. Unfortunately, no other positive clones could be isolated from any of these libraries. Since it appeared that this DNA region is very difficult to clone in cosmids, attempts to walk in cosmids to the Ai2 breakpoint were abandoned.

Mapping the Mouse *Fes* and *D7Rn4* Loci By Interspecific Backcross

Analysis Indicates That *Fes* May Facilitate Molecular Access

To the *msd* Locus From the Proximal Side

The mouse homolog of the feline sarcoma oncogene (*Fes*) locus (Sodroski *et al.*, 1984) has recently been mapped about 4.4 cM proximal to the *c* locus by *M. spretus*-*M. musculus* backcross analysis (Saunders and

Seldin, 1990). Also, *Fes* has been shown to map proximal to all of the albino deletions affecting *msd* (Rinchik and Taylor, unpublished data). Therefore, of all the currently available molecular markers that map proximal to *msd*, *Fes* is the nearest to it. As no information is available on the genetic distance of the *msd* or *hsdr-1* loci from the *c* locus, it is possible that *Fes* could be near enough to *msd* to help facilitate molecular access to the *msd* locus from the proximal side. Mapping the *Fes* and *D7Rn4* loci with respect to each other and additional flanking markers by interspecific backcross analysis should, therefore, help to determine the feasibility of using probes for these loci to study *msd* by long-range mapping of the *Fes-D7Rn4* interval, chromosome-jumping, and YAC cloning. Moreover, this mapping exercise could define the extents of some of the albino deletions in the genetic map, particularly by estimating the proximity of the *Fes* locus to the A1-deletion breakpoint (i.e., *D7Rn4* and *Fah* loci).

The chromosome 7 haplotypes of segregants of the backcross (129/Rl-*p c^{ch}/p c^{ch}* x *M. spretus*)F1 X 129/Rl-*p c^{ch}/p c^{ch}* have been defined previously by typing them for RFLVs at several molecular marker loci, including insulin-like growth-factor-1 receptor (*Igf1r*), *c*, and beta-globin (*Hbb*) (Committee for mouse chromosome-7 report, 1991; Rinchik, unpublished data). Of the proximally mapping markers typed in this backcross, *Igf1r* is currently the nearest one to the *c* locus. Recently, however, the *Fes* locus has been mapped between *Igf1r* and *c* with another interspecific backcross panel (Saunders and Seldin, 1990). *Hbb* is an anchor locus that maps about 3.5-6.2 cM distal to *c* (Doolittle *et al.*, 1990). To place the *D7Rn4* and *Fes* loci on the genetic map, probes RN254

(*D7Rn4* locus) and 80.BE1.3 (*Fes* locus, Table 2) were used to screen *Eco*RI- and *Bam*HI-digested DNAs, respectively, from 184 segregants of the interspecific backcross by Southern-blot analysis. This analysis identified 2 recombinants between *Fes* and both *D7Rn4* and *c*, 20 recombinants between *Fes* and *Igf1r*, and no recombinants between *D7Rn4* and *c* (Table 7).

The new haplotype information obtained with the 80.BE1.3 probe confirms a map position for *Fes* proximal to *c*, but at a significantly lesser distance ($1.1 \text{ cM} \pm 0.8 \text{ cM}$; Table 7) than the $4.4 \text{ cM} \pm 1.9 \text{ cM}$ estimated by previous analyses (95% confidence limits of 1.1 and 9.0 cM; Saunders and Seldin, 1990). The mapping analysis indicates that *Fes* maps 1.1 cM proximal to the *D7Rn4* locus, as well, since *c* and *D7Rn4* were inseparable by recombination in this set of 183 backcross segregants. In light of this new, and lower, estimate of the *Fes-c* genetic distance, there appears to be an even greater possibility that the *Fes* locus may be useful for the molecular analysis of *msd*. There is evidence, however, that the 1.1 cM distance estimate may be too low (see below and discussion). Interestingly, the $10.9\text{-cM} \pm 2.3\text{-cM}$ distance found between *Igf1r* and *Fes* (Table 7) is much greater than the $6.1 \pm 2.2 \text{ cM}$ reported by Saunders and Seldin (1990).

Although deletion mapping has positioned *D7Rn4* proximal to *c*, the absence of recombinants between *D7Rn4* and *c* precludes confirmation of this map position by linkage analysis. The absence of any recombinants out of 183 progeny does indicate that the *D7Rn4* and *c* loci may be less than 0.5 cM from each other. This is considerably less than the distance expected to be found between these loci based on long-

range-mapping (PFGE) analyses with probe RN216 (from the *D7Rn2* locus) and probes alb18 and N9.AH.8 (from the *D7Cw18* and *Pmv-31* loci, that map between the *hsdr-1* and *c* loci; Table 2 and Fig. 19). These PFGE analyses revealed that the *D7Rn2* locus maps at least 2 megabases proximal to the *c* locus (data not shown). Additional, and more detailed, long-range mapping analyses in this region of chromosome 7 have revealed that the *c* and *D7Rn4* loci are at least 2.8 megabases (mb) apart (G. Kelsey, pers. comm.), which, in light of the currently accepted estimate of 1-2 megabases/cM, translates into roughly 1.4-2.8 cM.

IV. DISCUSSION

Gaining Molecular Access to the *hsdr-1* and *msd* Loci By Cloning Deletion Breakpoint-Fusion Fragments

This study demonstrates that a panel of deletion mutations that has been used to define a region of the mouse genome functionally can be successfully exploited to gain molecular access to specific breakpoint-defined loci (i.e., functional units) known to contain genes essential for various stages of development. Of the extant lethal *c* deletions, those of the A (neonatally lethal) and Ai (prenatally lethal) complementation groups have proximal breakpoints that help to delimit the physical extents of the hepatocyte-specific developmental regulation-1 (*hsdr-1*) and mesoderm-deficient (*msd*) loci (Figs. 1 and 2). Cloning deletion-breakpoint-fusion fragments caused by the A and Ai deletions with probes mapping distal to the deletions will, therefore, provide DNA probes that map near or within the essential coding sequences of *hsdr-1* and *msd*, respectively. These probes will, in turn, facilitate the isolation of the genes within these loci that are necessary for normal regulation of liver-specific enzymes and proteins and for mesoderm formation.

In this study, refined mapping of the anonymous locus *D7Was12* with respect to the panel of albino deletions (Fig. 3A) positioned that locus between the *c* and juvenile development and fertility (*jdf*) loci, the same map position identified earlier for the *Emv-23* locus (Rinchik *et al.*, 1989; Fig. 2). These results suggested that the 12A and 23.3 probes (from the *D7Was12* and *Emv-23* loci, respectively) could possibly be used to initiate

molecular analyses of the *hsdr-1* and/or *msd* loci by identifying and cloning deletion breakpoints of the A and Ai complementation groups. Because conventional Southern-blotting analyses with these probes failed to identify deletion-breakpoint-fusion fragments associated with any A and Ai deletions, long-range mapping analyses were performed (Table 4, and Figs. 4 and 6) to estimate distances from the *Emv-23* and *D7Was12* loci to the distal breakpoints of the A and Ai deletions, as well as to the breakpoints of two viable *c* deletions (complementation-group Va) that break within the tyrosinase gene. A long-range (1.7-megabase) restriction map (Fig. 5) of the wild-type DNA segment containing the *c*, *D7Was12*, and *Emv23* loci was constructed, and the breakpoints of the A, Ai, and Va deletions were localized on this long-range map (Fig. 7) by the identification of structurally altered restriction fragments in DNAs containing these deletions (Fig. 6). The detection of rearranged *Sac*II fragments with probe 12A, but not probe 23.3, in carriers of several of the deletions demonstrated conclusively that the proximal to distal order of the cloned DNA loci is: *c* - *D7Was12* - *Emv-23*. This is in contrast to the *c* - *Emv-23* - *D7Was12* order recently reported for these loci by Saunders and Seldin (1990), which was based on the identification of one recombination event (in a double recombinant) between the *D7Was12* and *Emv23* loci in 114 *M. spretus*-*M. musculus*-backcross progeny.

As a result of the large-fragment analysis of the A1, A2, A4, Ai1, and Ai2 deletions, it was possible to estimate the approximate distances of their deletion breakpoints from the *c* and *D7Was12* loci (Table 5) based on a comparison of wild-type fragment sizes with the rearranged fragment sizes associated with these deletions (Fig. 6). The A1 (*c*¹⁴CoS)

breakpoint has the shortest maximum possible distance from a cloned locus (250-kb maximum distance from the *D7Was12* locus), and the Ai2 and A4 breakpoints have the shortest minimum possible distance from a locus (~2.5 kb from the *c* locus; Table 5). The potentially great distances of the *D7Was12* and *c* loci from the A- and Ai-deletion breakpoints, however, argued against attempting to clone the breakpoints by chromosome walking from these loci with cosmid clones.

In addition, the large-fragment analysis made it possible to estimate the approximate distances of the A and Ai deletion breakpoints from either or both of the *c*^{24R75M} and *c*^{16DFiOD} distal breakpoints (Table 5). It was evident from these distances that both the A1 and A2 breakpoints are potentially closer to the *c*^{24R75M} breakpoint than to the *D7Was12* locus. Moreover, the Ai1 breakpoint is potentially closer to the *c*^{16DFiOD} breakpoint than to the *c* locus. In light of the fact that both the *c*^{24R75M} and *c*^{16DFiOD} deletions break within the *c* locus (Fig. 3 and Rinchik *et al.*, in preparation) and extend distally into the vicinity of the breakpoints of the A1, A2, and Ai1 deletions (Fig. 7), breakpoint-fusion fragments were cloned from *c*^{24R75M}/*c*^{24R75M} and *c*^{16DFiOD}/*c*^{16DFiOD} DNAs (clones RN207 and λ 617-4A1, respectively; Fig. 8 and Rinchik and Machanoff, unpublished data) using the tyrosinase cDNA probe (MTY811). The distal-end clone, RN202 (locus *D7Rn3*), derived from the λ 617-4A1 fusion fragment, maps within all the A and Ai deletions, except Ai1, Ai2, and A4 (Figs. 19 and 20). Among the latter deletions, Ai1 is the one with a distal breakpoint that maps closest to *D7Rn3*, based on the results of large-fragment analysis (Fig. 7). Probe RN204, a distal-end clone derived from the RN207 fusion fragment (Fig. 8), maps within the

A3, A5, A6, and Ai3 deletions, but distal to the A1, A2, A4, Ai1, and Ai2 deletions (Figs. 9A and 10).

Fortunately, RN204 detects size-altered *Bam*HI, *Kpn*I, and *Bst*EII fragments in A1/A1 DNA (e.g., Fig. 9A), a result strongly suggesting that the A1 breakpoint is responsible for the aberrant fragments. The 21.5-kb size-altered *Bam*HI fragment was, therefore, cloned from A1/A1 DNA (clone RN226; Fig. 8) to isolate DNA sequences located proximal to the A1 deletion. The proximal 160-bp *Hind*III-*Bam*HI end of this fragment (clone RN247; locus *D7Rn2*) was, in fact, found to map proximal to the A1 and A5 deletions, but distal to the breakpoints of all the other deletions with proximal breakpoints located between the *msd* and *hsdr-1* loci (Figs. 9A,B and 11). Moreover, failure to identify a size-altered *Bam*HI fragment in A5/A5 DNA with probe RN247 indicates that the proximal A5-deletion breakpoint maps distal to the proximal A1 breakpoint. Probe RN247 is the first DNA clone isolated from a locus (*D7Rn2*) that maps distal to the *msd* locus but proximal to or within the *hsdr-1* locus, and should be useful for initiating the physical mapping of these loci.

The Fumarylacetoacetate Hydrolase (*Fah*) Gene Maps at the Proximal Breakpoint of the *c*¹⁴CoS (A1) Deletion

Because the *D7Rn2* locus maps within or near the *hsdr-1* locus, the region of DNA surrounding this locus was also screened for the presence of a transcription unit. The position of the A1 breakpoint was localized on the restriction map of the fusion-fragment clone RN226 (Fig. 8), and

probes, isolated from that portion of RN226 located proximal to the breakpoint, were found to contain evolutionarily conserved DNA sequences (e.g., Fig. 13). These probes, as well as other probes from the *D7Rn2* locus, do indeed detect transcripts of several different sizes and abundances, the most abundant being transcripts of 1.6 and 3.5 kb, in the livers of 18.5-day fetuses and juvenile and adult mice (Fig. 14A,B). The considerable expression of the identified gene(s) in the liver and kidney is provocative evidence that this gene may be the one responsible for the *hsdr-1* function. The 1.6-kb transcript is already apparent in whole 16.5-day fetal mRNA, and all the liver transcripts are also detected in adult kidney and brain, although the 1.6-kb transcript is not abundant in the adult brain, and neither it nor the 3.5-kb transcript is abundant in the brains of 18.5-day fetuses (Fig. 14A,B,F). Neither transcript is significantly present in any of the other tissues examined (Fig. 14F). However, probe 226.1, but not probe 226.2, does appear to detect a lower-abundance, 1.4-kb transcript in the thymuses of 8.5-day-old mice. It is possible that the function of this transcript in the thymus may be different from the function of the abundant transcripts in the liver and kidney. A function of this transcript in the thymus would be consistent with the observation of morphological abnormalities of the thymus in newborn mice homozygous for deletions of *hsdr-1* (Erickson *et al.*, 1968).

The identification of an abundant 3.5-kb transcript with DNA probe RN226.2 (Fig. 14A), but not with other probes that flank 226.2 on the restriction map (Figs. 14B and 8, and data not shown), implies that probe 226.2 contains exons specific for the 3.5-kb transcript. This indicates that RNAs derived from this transcription unit may be alternatively spliced

or, perhaps, that two overlapping genes may be present at this locus [e.g., see Xu *et al.* (1990) and Huang *et al.* (1990)]. In fact, the only abundant transcript identifiable in the adult brain was the 3.5-kb transcript (Fig. 14F). Moreover, it was detected only with the 226.2 probe, and not with 226.1. This suggests that the 1.6-kb transcript does not play an important role in the adult brain, whereas the 3.5-kb transcript may. It is, therefore, possible that these two transcripts may also have distinct roles in both the liver and kidney. A comparison of the sequences of cDNA clones derived from each of these transcripts should indicate if these are likely possibilities.

Since the proximal breakpoint of the A5 deletion is more distal than any of the other deletion breakpoints located between the *msd* and *hsdr-1* loci, A5 is the deletion that is critical for determining whether the identified gene may be the *hsdr-1* gene or one of the genes contributing to *hsdr-1* function. Indeed, the failure to detect any transcripts in liver RNA from 18.5-day A5/A5 and A1/A1 fetuses with the 226.1 and 226.2 probes (Fig. 14A,B) considerably strengthens the possibility that this gene may contribute to *hsdr-1* function.

Numerous BALB/c liver cDNA clones were isolated with probe RN226.2. All of the cDNAs analyzed label the 1.6-kb transcript most intensely, although several of the cDNAs label it much more intensely than do others (e.g., Fig. 14C,D). None of the cDNAs analyzed labels the 3.5-kb transcript as intensely as does probe 226.2 (Fig. 14A-D), implying that the 226.2 fragment contains exons for this transcript that are not present in these cDNAs. Interestingly, none of the cDNA probes detects size-altered or deleted fragments in A5/A5 DNA (Fig. 9B). Since

sequence analysis showed that at least one of these cDNAs (RN273) contains a complete open reading frame (ORF; see Figs. 15 and 16, and discussion below), the failure to detect deletion of exon-containing fragments with this cDNA suggests that the A5 deletion removes either the gene promoter or other essential *cis*-acting regulatory sequences. Therefore, isolation of the wild-type DNA sequences disrupted by the proximal breakpoint of the A5 deletion may lead to the identification of some of these regulatory sequences.

Mapping the cDNAs with respect to the deletions by Southern blotting identified two classes of cDNAs. One class (e.g., cDNA RN271) hybridizes to a single *Eco*RI fragment in wild-type DNA, and the other class (e.g., RN273) detects several *Eco*RI fragments (Fig. 9B). cDNAs of the former class also label the 1.6-kb transcript less intensely than do cDNAs of the latter (Fig. 14D and 14C, respectively). cDNAs of the latter type are probably derived from a greater portion of the 1.6-kb transcript than are those of the former type. One of the cDNAs (cDNA RN273), of the class that labels the 1.6-kb transcript very intensely (Fig. 14C), has an insert size (1.45 kb) similar to the estimated size of the transcript, and has been shown to contain an entire ORF (Figs. 15 and 16; also see discussion below). Therefore, the 1.6-kb transcript detected by the Northern blot analyses is probably mature, functional, and the source of the RN273-like cDNAs. Those cDNAs that label this transcript less intensely and that are derived from exons contained within a single 6.5-kb *Eco*RI fragment may be derived from immature precursors of this transcript, or from one of the larger transcripts that are identified by these cDNA probes (Fig. 14D).

The nucleotide sequence of one of the cDNAs that detects several *Eco*RI fragments (cDNA RN273) is ~80% homologous (over the entire length of the cDNA) to a recently cloned cDNAs encoding human FAH (Agsterribe *et al.*, 1990; Phaneuf *et al.*, 1991). The mouse cDNA RN273 contains a complete ORF that is the same length as the ORF identified in the full-length human cDNA clone (Fig. 15 and Phaneuf *et al.*, 1991). The mouse and human ORFs have ~84% homology at the nucleotide-sequence level, and ~88% identity and ~93% similarity (considering conserved amino-acid changes) at the level of the encoded-amino-acid sequences (Figs. 15 and 16). Whether or not the mouse cDNA RN273 is full-length, remains to be determined. However, it is clear that the cDNA RN273 encodes mouse FAH. Therefore, the *D7Rn2* locus, from which the cDNA is derived, will hereafter be referred to as the structural locus, *Fah*.

The ORFs of both RN273 and the full-length human cDNA extend beyond and do not contain the putative upstream (5') termination codon reported for the human cDNA analyzed by Agsterribe *et al.* (1990). Although no in-frame stop codons are present upstream of the putative initiation codon in either the mouse cDNA RN273 or the human cDNA analyzed by Phaneuf *et al.* (1991), the loss of nucleotide sequence homology shortly upstream of the initiation codons and the absence of other possible upstream and in-frame initiation codons (Fig. 15) argue against a continuation of the ORF. The most compelling evidence that the observed ORF is complete has been provided by Phaneuf *et al.* (1991). They have shown that this ORF yields a protein (via *in vitro* translation) of molecular weight identical to that of purified human FAH, and with

comparable level of activity. Phaneuf *et al.* (1991) also demonstrated that, when the internal Met codon of the human gene is used as an initiation codon, as proposed by Agsterribe *et al.* (1990), for *in vitro* translation from abbreviated cDNA inserts, lower-molecular-weight proteins with no FAH activity are produced. In light of this information, it is probably unlikely that the cDNA sequence reported by Agsterribe *et al.* (1990) contains the 5' end of the open reading frame encoding FAH. However, it is possible that the internal Met codon of the cDNA sequence may be used as an alternative initiation site to generate, for example, a protein of different function. In this context, it should be noted that human cDNA clones of a type represented by the mouse cDNA RN271 (that hybridize weakly to the transcript encoding *Fah*) have not been reported.

The human *FAH* locus has been mapped to the q23-q25 region of human chromosome 15 by screening somatic-cell hybrids and performing *in situ* hybridization with the *FAH* cDNA (Phaneuf *et al.*, 1991). The isolation and mapping of the mouse *Fah* locus proximal to the *c* locus (by deletion mapping, Fig. 9B) but distal to the *Fes* locus (by interspecific-backcross analysis, Table 7), extends the region of homology between human chromosome 15 and mouse chromosome 7 that had previously been identified at and around the *Igf1r* and *Fes* loci (both of which map to human bands 15q25-qter; Lalley *et al.*, 1989). Therefore, human homology with this region of mouse chromosome 7 changes from human 15q23-q25 at the *Fah* locus to human 11q14-q21 at the *c* (*Tyr*) locus (Lalley *et al.*, 1989). These refined mouse-human linkage conservation relationships will be helpful for mapping the human homologs, if any, of other mouse genes that may be discovered in this chromosome-7 region.

Could *Fah* Deficiency Be Responsible For the *hsdr-1* Phenotype?

Maximal expression of the *Fah* gene in liver and kidney and absence of *Fah* gene expression in the livers of 18.5-day fetuses homozygous for the A5 deletion, the only *hsdr-1*-associated deletion that does not remove some of the coding sequences of the gene, indicate that, based on physical analyses only, *Fah* is a strong candidate for the *hsdr-1* gene. But, it is not readily apparent why deficiency of a metabolic enzyme such as FAH should lead to the failure of induced expression of the liver-specific genes affected by deletion of *hsdr-1*. The characteristics of hereditary tyrosinemia type 1 (*HT*), a human disorder caused by FAH deficiency, do suggest a possible explanation.

All *HT* patients display liver and renal dysfunctions, but acute and chronic forms of the disease have been observed (Tanguay *et al.*, 1990; Phaneuf *et al.*, 1991; and references therein). The acute form is most frequently reported, and is characterized by severe liver dysfunction, failure to thrive, vomiting, diarrhea, and, usually, death from liver failure within the first year, if untreated. Patients with the chronic form exhibit renal-tubular dysfunctions, vitamin D-resistant rickets, hepatic cirrhosis, and, often, adolescent hepatocarcinoma (Tanguay *et al.*, 1990; Phaneuf *et al.*, 1991; and references therein). In addition, about 40% of *HT* patients have been found to experience episodes of acute, severe peripheral neuropathy, similar to the neurological crises of porphyria patients (Mitchell *et al.*, 1990). The generalized liver and kidney failure associated with *HT* is consistent with the phenotype observed in newborn

mice homozygous for deletions of *hsdr-1*. *HT* patients excrete large amounts of the abnormal metabolite succinylacetone (SAC), which suggested that the primary defect in *HT* is FAH deficiency (Lindblad *et al.*, 1977). FAH deficiency was then documented in *HT* patients by assaying FAH activity in liver extracts (Fallstrom *et al.*, 1979; additional references in Tanguay *et al.*, 1990 and Phaneuf *et al.*, 1991) and immunoscreening with an antibody to the purified human enzyme (Tanguay *et al.*, 1990). Immunoblot analyses were also able to distinguish between the acute and chronic forms of *HT*, with no detectable FAH in patients with the acute form and about 20% of normal levels in patients with the chronic form (Tanguay *et al.*, 1990).

Failure to induce specific metabolic enzymes and proteins in the liver and kidney, as is observed in mice homozygous for deletions of *hsdr-1*, has not been reported in *HT* patients. Studies addressing whether or not these enzymes and proteins are induced in the livers of *HT* patients would be very helpful in determining whether or not inactivation of the *Fah* gene is responsible for the failure of induction in mouse mutants. The results of analyses of the effects of SAC accumulation in *HT* patients do indicate a possible role for SAC in the phenotype observed in *hsdr-1* mutants, which, most likely, also excrete SAC. SAC is known to be a potent inhibitor of δ -aminolevulinic acid dehydratase (δ -ALD), an enzyme in the heme biosynthetic pathway (Sassa and Kappas, 1983). The resulting inhibition of heme biosynthesis probably accounts for the acute neurological symptoms, resembling those of acute intermittent porphyria, that are often observed in *HT* patients (Sassa and Kappas, 1983; Mitchell *et al.*, 1990). Moreover, SAC reacts nonenzymatically with

amino acids, proteins, and reduced glutathione (GSH) to form stable adducts (Manabe *et al.*, 1985). This may account for some of the disorders found in *HT* patients, which include aminoaciduria, proteinuria, and reduced levels of GSH in the liver (Manabe *et al.*, 1985). Analyses of SAc levels, as well as other biochemical consequences of a block in the tyrosine degradation pathway, in *hsdr-1* mutants would help to correlate the phenotypes of *hsdr-1* mutants and *HT* patients and would contribute to an investigation of whether *Fah* could be the *hsdr-1* gene.

The preference exhibited by SAc, in the formation of adducts, for specific amino acids (the most reactive being, in descending order, lysine, glycine, methionine, and phenylalanine; Manabe *et al.*, 1985) could possibly be invoked to account for preference for particular enzymes and proteins as well. If SAc does act preferentially on certain proteins, then it may be contributing to some of the specific biochemical abnormalities observed in mice homozygous for deletions of *hsdr-1*. One possibility is that SAc may be binding to and inactivating some of the proteins involved in membrane architecture and in the induction or repression of the liver-specific genes secondarily affected by the albino deletions. Such a scenario may possibly account for the defects observed in mice homozygous for deletions of the *hsdr-1* locus, without invoking the presence of a gene encoding a trans-acting regulatory factor at the locus. Additional studies investigating the specificity of action, if any, of SAc for enzymes and proteins may help address this possibility.

Despite evidence, from Southern-blotting, expression analyses, and *HT* patients, that the mouse *Fah* gene is a candidate for a gene involved in *hsdr-1* function, a causal link between inactivation of *Fah* and

the phenotype resulting from homozygous deletion of *hsdr-1* is not yet possible. In this light, it is interesting that "saturation mutagenesis" with N-ethyl-N-nitrosourea (ENU) of DNA corresponding to that deleted in the Fp1 deletion has failed to identify a single mutation with the *hsdr-1* phenotype, whereas repeat mutations have been isolated for several other loci within the region (Rinchik *et al.*, 1990; Rinchik, 1991). There are two basic scenarios that could explain the failure to induce *hsdr-1*-like mutations with ENU. First, the *hsdr-1* locus may contain only the *Fah* gene, which, for some reason, is not readily mutable with ENU. For instance, the *Fah* gene may be able to tolerate a variety of point mutations without yielding a visible phenotype. Alternatively, the 1.25-kb ORF encoding the FAH enzyme may not be a large enough target for efficient ENU-induced mutagenesis. It may be useful to determine if there is a relationship between the size of a gene's ORF and its susceptibility to ENU-induced mutagenesis. Second, the *hsdr-1* phenotype may be due to the combined effects of deletion of *Fah* and an additional gene(s) located more distally than *Fah*.

Can a causal link between deletion of *Fah* and the *hsdr-1* phenotype be established? A positive result in a phenotype-rescue experiment with the *Fah* gene would demonstrate that inactivation of this gene alone leads to the *hsdr-1* phenotype. The availability of SV40-transformed mutant (i.e., A1/A1) liver-cell lines, that do not express the liver-specific enzymes regulated by *hsdr-1* (Paul *et al.*, 1989), may permit phenotype rescue to be demonstrated *in vitro* by transfection of these cells with the wild-type *Fah* gene. However, it would be difficult to conclude from a negative result (i.e., the failure of *Fah* to correct the deficiencies

in liver-specific enzymes *in vitro*) that *Fah* is not *hsdr-1*, as SV40-transformation and *in-vitro* culturing of the hepatocytes may have altered their original *in vivo* functions. Phenotype rescue *in vivo*, i.e., by the generation of rescued transgenic mice following microinjection of the wild-type gene into fertilized eggs from the cross *c^{ch}/A1* x *c^{ch}/A1*, would be a more definitive experiment, as it may permit the demonstration of complete or partial rescue of the *in vivo* phenotype resulting from deletion of *hsdr-1*. However, a negative result from such an experiment may also not be conclusive, since the absence of an essential regulatory element in the microinjected construct could possibly be invoked.

A definite answer to the question of whether inactivation of the *Fah* gene alone can lead to the *hsdr-1* phenotype could be given by the inactivation of the wild-type gene by homologous recombination in embryonic-stem (ES) cells, and the subsequent generation of homozygotes for the inactivated gene. If inactivation of *Fah* fails to cause the *hsdr-1* phenotype, then deletion of a different gene or a combination of genes would have to be invoked to account for the phenotype.

Irrespective of whether or not *Fah* functions as *hsdr-1*, the isolation of the gene has redefined the extent of *hsdr-1* on the physical map (Fig. 19). The *hsdr-1* locus is now bounded proximally by *Fah* and distally by *D7Cw18*, a locus, identified by a microdissection clone, that maps proximal to all deletions with breakpoints located between the *hsdr-1* and *c* loci, except the Bem1 deletion (Table 2 and Fig. 19). Additional large-fragment analyses of wild-type DNA between these two loci should result in the construction of a long-range map of the *hsdr-1* locus and, thereby, in the localization of other "CpG islands" (Bird, 1986;

see below). Identification of such islands will help to determine whether other genes, in addition to *Fah*, may exist in the interval between these two probes. If the *Fah* gene fails to correct the *hsdr-1* phenotype, then it will be necessary to conduct phenotype-rescue and/or gene-inactivation experiments on any additional transcription units (or possibly even on combinations of them) identified in the *Fah-D7Cw18* interval.

The *Fah* and *D7Rn4* Loci Will Contribute To Molecular Analyses of Additional Genetic Loci

In addition to its use in identifying the *Fah* gene, the *D7Rn2* locus was used to identify the A5 breakpoint at the molecular level and to initiate a chromosome walk proximally toward the *msd* locus. The failure to identify a size-altered *Bam*HI fragment in A5/A5 DNA with probe RN247 (from the *D7Rn2* locus; Fig. 9A) indicated that the proximal A5-deletion breakpoint maps distal to the proximal A1 breakpoint (Fig. 11). Further analysis of A5/A5 DNA with probe RN216 (also from the *D7Rn2* locus) revealed the disappearance of a wild-type 16.5-kb *Sma*I fragment and the appearance of a size-altered *Sma*I fragment of 21 kb in this DNA (Fig. 17). This strongly suggests that the A5 breakpoint is located within 16.5 kb of the *D7Rn2* locus. Isolation of a breakpoint-fusion fragment from A5/A5 DNA should, therefore, be possible with a probe derived from DNA near this locus. A distal portion of such an A5-fusion-fragment clone would be derived from DNA sequences located distal to the A5 breakpoint and, therefore, at or near the proximal end of the juvenile-development and fertility (*jdf*) locus (Fig. 11 and Table 1).

[Molecular access to the gene(s) involved in the *jdf* syndrome will require the cloning of DNA sequences located distal to all the A, Ai, and Vb deletions, which collectively define, genetically and physically, the proximal border of the *jdf* locus (Figs. 1 and 11).] A DNA probe derived from the distally mapping portion of the fusion fragment may permit a molecular analysis of the *jdf* locus to be initiated. For instance, chromosome walking with yeast-artificial-chromosome (YAC) clones (Burke *et al.*, 1987) from such a probe may lead to the isolation of a DNA probe from a locus that maps distal to the C deletion, the only deletion that breaks within or distal to the *jdf* locus (Fig. 11). This YAC contig would, by definition, contain at least a portion of the *jdf* locus.

In addition to the identification of the A5 breakpoint, probe RN216 was used to isolate a cosmid clone (pRN224) of 46 kb of wild-type DNA at the *D7Rn2* locus. This, in turn, led to the isolation of a proximal-end clone from a locus (*D7Rn4*) that maps at least 40 kb closer to the distal end of *msd* and proximal to the Di4- and Fp1-deletion breakpoints (Figs. 18 and 19). This mapping information, together with the finding that the A5 breakpoint is located within 16.5-kb of the *D7Rn2* locus (Fig. 17), delineates a ≤ 62 -kb DNA segment containing the entire structural *Fah* gene (Fig. 19). This definition of the *Fah* structural locus is possible, because *Fah* coding sequences (identified by cDNA probe RN273, which contains a complete ORF) are completely deleted by both the Di4 and Fp1 deletions but are not affected by the A5 deletion on the genomic level (Fig. 9B). In addition to closing the gap between the proximal extent of the physical map and the *msd* locus, the RN224 cosmid will provide DNA probes that can be used to clone breakpoint-fusion fragments from DNAs

containing the Di4 and Fp1 deletions. These clones will, in turn, provide DNA probes that can be used to facilitate molecular analyses of the genetic loci at or in the vicinity of the distal breakpoints of the Di4 and Fp1 deletions (Fig. 1).

The recent isolation of a breakpoint-fusion fragment containing the Bem1-deletion breakpoint (T. Magnuson, personal communication) led to the cloning of a locus (*D7Cw11D*), from the portion of the fusion-fragment located distal to the Bem1 breakpoint, that maps distal to the Di4 deletion. Both the Bem1 and Di4 deletions affect the *eed* locus (Table 1; Fig. 1), while only Bem1 deletes the fitness-1 (*fit-1*) locus, defined by a fitness/runting mutation recently induced by ENU mutagenesis and mapped between the *eed* and *exed* (Table 1) loci (Rinchik et al., 1990; Rinchik, 1991; Fig. 1). Therefore, isolation of a Di4-breakpoint clone (starting from the *D7Rn4* locus) will be valuable for beginning a molecular analysis of not only the *eed* locus, but also the *fit-1* locus. This probe will be particularly useful, in conjunction with probes from the *D7Cw11D* locus, for physically mapping the Di4-Bem1-breakpoint interval and identifying candidate transcription units for the *fit-1* gene.

The distal Fp1-deletion breakpoint is located between the shaker-1 (*sh-1*) and *Hbb* loci (Fig. 1). Recently, an ENU-induced prenatally lethal mutation has been induced in a gene, designated *l(7)-1Rn*, located between the distal border of the *pid* locus (Table 1) and the distal breakpoints of the Fp1 and Fq1 deletions (Fig. 1). The *D7Or1* and olfactory marker protein (*Omp*) loci have also been mapped into this DNA interval (Johnson et al., 1989; Rinchik et al., 1991). Therefore, an Fp1-breakpoint clone will provide a molecular marker for the distal border of this

interval (Fig. 1), and, together with *D7Or1* and *Omp*, should facilitate the physical mapping of this DNA segment and the identification of candidate transcription units for the *sh-1* and *l(7)-1Rn* genes.

Strategies Used to Facilitate Molecular Access to the *msd* Locus

The isolation of the RN224 cosmid clone resulted in about a 40 kb molecular "step" toward the distal end of the *msd* locus. Chromosome-walking proximally from the *D7Rn4* locus at the proximal end of the cosmid insert (using both cosmid and YAC clones), in conjunction with chromosome jumping (Collins and Weissman, 1984; Poustka and Lehrach, 1986) (if CpG-island clones become available as a result of the chromosome walk), should eventually yield a DNA clone that maps proximal to all the deletions that break between the *msd* and *hsdr-1* loci. This clone would, by definition, be derived from the distal end of the *msd* locus (see Fig. 19). Cloning the deletion breakpoints of the A2 and/or A4 deletions, which are now known to extend further proximally than the A1 deletion (Figs. 9A and 19), with probes that map distal to the deletions (using the same type of approaches as outlined above) would also facilitate molecular access to *msd* from the distal side. Pursuing these strategies for accessing *msd* will, at the same time, provide molecular access to the proximal breakpoints of many of the other deletions of the A, E, Di, and Dp complementation groups (Fig. 19). Cloning the breakpoint-fusion fragments associated with these deletions would generate DNA clones from within or near the functional loci at the distal ends of those deletions (i.e., essentially all of the *c*-region loci located distal to the *c*

locus; Fig. 1), thus permitting molecular analyses of these developmentally important DNA loci to be initiated. This would also expedite the construction of a long-range map of the entire albino deletion complex.

The strategy of cloning breakpoint-fusion fragments, used successfully to clone the breakpoints of the *c^{16DFiOD}* (Machanoff and Rinchik, unpublished data), *c^{24R75M}*, and *c^{14CoS}* (A1) deletions (Fig. 8), was continued, in this study, in an attempt to clone an Ai-deletion breakpoint. Cloning an Ai-breakpoint-fusion fragment would result in the isolation of a DNA probe from within or near the proximal end of the *msd* locus. This probe should render the *msd* locus amenable to molecular analysis, and may permit the *msd* locus to be linked to the *Fah* locus on a long-range physical map.

Large-fragment analyses of the deletions (Fig. 6; Table 4) and deletion-mapping of the *D7Rn1* locus (Fig. 9A) suggested that the distal breakpoint of the *c^{16DFiOD}* deletion may be near the distal Ai1 breakpoint (Fig. 10). Therefore, the distal end of the *c^{16DFiOD}*-breakpoint-fusion fragment (probe RN202) was mapped with respect to the deletion panel, and found to identify a locus (*D7Rn3*) that maps between the Ai1 and A2 distal breakpoints (Figs. 19 and 20). Since this map position suggested that the RN202 probe may be useful for cloning the Ai1 breakpoint, the probe was used to try to identify a breakpoint-fusion fragment by Southern-blot analysis with several standard restriction enzymes. Unfortunately, this analysis, as well as PFGE analysis of Ai1/Bem1 DNA partially digested with *Kpn*I, failed to identify the Ai1 breakpoint.

In light of the finding that the distal end of the tyrosinase (*c*-locus) gene maps within 60 kb of the Ai2 breakpoint (Fig. 19, see explanation in Results), attempts were made to clone the Ai2 breakpoint by chromosome-walking distally from the *c* locus. Despite attempts, with four different cosmid libraries (2.3×10^6 total cosmids screened), to isolate cosmid clones with probes derived from DNA flanking the distal end of *c*, no useful clones were identified. Thus, perhaps this region is not clonable in cosmids. In this context, it is noteworthy that YAC clones now exist for tyrosinase (Rinchik, Stubbs, and Walkowicz, unpublished). Future analysis of such clones may yield probes suitable for identifying the Ai2 breakpoint.

In addition to pursuing the Ai-breakpoint-cloning strategy for gaining molecular access to *msd*, the feasibility of using the *Fes* locus to access *msd* molecularly from the proximal side (see Fig. 1) was investigated. The *Fes* and *D7Rn4* loci were mapped by interspecific-backcross analysis with respect to each other and to flanking markers. The haplotypes observed in the backcross progeny indicated that *Fes* maps about 1.1 cM proximal to both the *D7Rn4* and *c* loci, which were not separated from each other by recombination in any of the backcross segregants (Table 7). The 1.1-cM distance between *Fes* and *c* is considerably less than the 4.4-cM distance recently estimated by Saunders and Seldin (1990), who also used interspecific-backcross analysis. These two backcross panels have also yielded significant differences in genetic-distance estimates between other loci proximal to *c*. This may be due to chromosomal differences between the different *M. musculus* strains used and/or between the *M. spretus* strains used in

these two backcross panels [C3H/HeJ-*gld/gld* (generalized lymphoproliferative disease) and *M. spretus* (Spain) used by Saunders and Seldin (1990)].

The observation from long-range mapping that the *c* and *D7Rn4* loci are at least 2.8 megabases (mb) apart (G. Kelsey, personal communication) should also be considered in a comparison of the different estimates for the genetic distance between *Fes* and *c*. This physical distance translates into about 1.4-2.8 cM based on the rough estimate of 1-2 mb per cM. In light of this estimated genetic distance and the observation in this study of no recombinants between *D7Rn4* and *c* out of 183 backcross segregants, it appears likely that recombination between *M. spretus* and 129/R1 inbred mice may be somewhat suppressed in this DNA region. Therefore, the 1.1-cM distance estimated between *Fes* and *D7Rn4* in this study may also be a low estimate. Obtaining another estimate of the distance between these loci by mapping *D7Rn4* with the backcross panel of Saunders and Seldin (1990), which gives larger genetic distances between loci proximal to *c*, would, therefore, be helpful in determining whether *Fes* may facilitate molecular access to *msd*.

A very rough estimate of 1.6-3 cM can be made for the genetic distance between *Fes* and *D7Rn4* based on the 4.4 cM *Fes-c* distance reported by Saunders and Seldin (1990) and the 1-2.5 cM *D7Rn4-c* distance estimated from physical-mapping analyses. Therefore, the *Fes-D7Rn4* interval, which contains the *msd* locus, is potentially a very large physical size (1.6-6 mb). Large-fragment analysis of deletions of the *Ai*, *Dj*, *Dq*, and *Fq* complementation groups with a probe from the *Fes* locus, may, however, result in the molecular identification of the breakpoints of

some of these deletions (see Fig. 1). Additionally, this approach may permit physical distances to the proximal breakpoints to be estimated, as was done successfully in this study. This, in turn, may indicate whether or not *Fes* may be near enough to *msd* to facilitate molecular access to *msd* from the proximal side.

Potential CpG Islands and Transcription Units Identified by Long-Range Mapping

Several candidate CpG-rich islands (also referred to as 'HTF' islands; Bird, 1986), which are often landmarks for the 5' ends of genes (Bird, 1987; Lindsay and Bird, 1987), have been identified by the long-range mapping around the *c*, *D7Was12*, and *Emv-23* loci (see Fig. 5). A good candidate for a CpG island (see Lindsay and Bird, 1987) is the *Bss*HII/*Sac*II/*Nae*I site located distal to the *Emv-23* locus. The functional importance of the DNA sequences near this potential island is currently unknown, since it is not yet known whether this site is removed by any of the viable albino deletions of the Vb group (Fig. 19). In addition, *Sac*II, *Nae*I, and *Sma*I cleavage sites occur very close together between the *D7Was12* and *Emv-23* loci, and a *Bss*HII/*Nae*I site occurs proximal, but 3', to the tyrosinase gene (Fig. 5). If there actually are unidentified genes located near these possible CpG islands, the null phenotypes associated with these genes must be very subtle, since the genomic region spanning both of the islands can be homozygously deleted with no apparent reduction in viability or fertility (Rinchik *et. al.*, in preparation). In fact, two viable albino deletions (*c*^{51DTD} and *c*^{31HATh}) exist that, at the very

least, delete 1.5 mb of DNA physically mapped in this study (including the *c*, *D7Was12*, and *Emv-23* loci; Fig. 19), as well as additional genomic DNA located proximal to this segment (Rinchik et al., in preparation; Fig. 19). These deletions may remove as much as another 2.4 mb of DNA located proximal to the most proximal *Bss*HII/*Nae*I site mapped in this study, based on a long-range map recently derived for this region (G. Kelsey, personal communication).

Cloning of these candidate CpG islands might aid in the identification of new transcription units; moreover, it should provide DNA probes useful for chromosome jumping to other restriction sites cleaved by the above enzymes (Collins and Weissman, 1984; Poustka and Lehrach, 1986). Chromosome jumping, in combination with the use of YACs, should facilitate the molecular identification and analysis of nearby loci and deletion breakpoints. For example, cloning the *Bss*HII/*Sac*II/*Nae*I site located distal to the *Emv-23* locus and jumping distally from this site, in conjunction with cloning the A5 breakpoint (using probes from the *Fah* locus, as discussed above), could be very useful for future identification of the gene(s) within the *jdf* locus (see Fig. 19) that is (are) necessary for normal spermiogenesis, completion of pregnancy, and overall fitness (Lewis et al., 1978; Russell et al., 1982). Molecular access to the proximal side of the *jdf* locus will require the cloning of DNA sequences located distal to all the A, Ai, and Vb deletions, which collectively define, genetically and physically, the proximal border of *jdf* (Figs. 2 and 19). Moreover, it should be noted that, since the *Bss*HII/*Sac*II/*Nae*I site located distal to the *Emv-23* locus is not

yet known to be deleted by any of these deletions (Fig. 19), it may, in fact, be the 5' CpG island for a gene at the *jdf* locus.

If additional coding sequences can be isolated from the region physically mapped in this study, the human homolog(s) to the gene(s), if they exist, can be isolated and mapped. This analysis should expand the synteny maps between the human and mouse genome in the mouse-chromosome-7 region between the *c* and *Mod-2* loci, which map within human chromosome regions 11q14-q21 and 6p21.3-pter (Lalley *et al.*, 1989). In addition, ongoing physical and functional mapping of the *Fes-Hbb* region of mouse chromosome 7, which is being facilitated by the use of the panel of albino deletions (Rinchik and Russell, 1990; Rinchik *et al.*, 1990; and Rinchik, 1991), will eventually result in the identification and cloning of many functionally defined mouse genes and their human counterparts.

LIST OF REFERENCES

- Agsteribbe, E., Van Faassen, H., Hartog, M. V., Reversma, T., Taanman, J.-W., Pannekoek, H., Evers, R. F., Welling, G. M., and Berger, R.: Nucleotide sequence of cDNA encoding human fumarylacetoacetase. *Nucl Acids Res* 18: 1887, 1990.
- Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A., and Struhl, K. (eds.): *Current Protocols in Molecular Biology: Volume 1*, John Wiley and Sons, New York, 1989.
- Aviv, H. and Leder, P.: Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc Natl Acad Sci USA* 69: 1408-1412, 1972.
- Baier, L. J., Hanash, S. M., and Erickson, R. P.: Mice homozygous for chromosomal deletions at the albino locus region lack specific polypeptides in two-dimensional gels. *Proc Natl Acad Sci USA* 81: 2132-2136, 1984.
- Bates, P. F. and Swift, R. A.: Double *cos* site vectors: simplified cosmid cloning. *Gene* 26: 137-146, 1983.
- Bird, A. P.: CpG-rich islands and the function of DNA methylation. *Nature* 321: 209-213, 1986.
- Bird, A. P.: CpG islands as gene markers in the vertebrate nucleus. *Trends Genet* 3: 342-347, 1987.
- Bowes, C., Danciger, M., Kozak, C. A., and Farber, D. B.: Isolation of a candidate cDNA for the gene causing retinal degeneration in the *rd* mouse. *Proc Natl Acad Sci USA* 86: 9722-9726, 1989.
- Brown, W. R. A. and Bird, A. P.: Long-range restriction site mapping of mammalian genomic DNA. *Nature* 322: 477-481, 1986.

- Burke, D. T., Carle, G. F., and Olson, M. V.: Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236: 806-812, 1987.
- Carle, G. F. and Olsen, M. V.: An electrophoretic karyotype for yeast. *Proc Natl Acad Sci USA* 82: 3756-3760, 1985.
- Collins, F. S., and Weissman, S. M.: Directional cloning of DNA fragments at a large distance from an initial probe: a circularization method. *Proc Natl Acad Sci USA* 81: 6812-6816, 1984.
- Committee for mouse chromosome-7 report: Maps of chromosome 7: First report. *Mamm Genome*: in press, 1991.
- Davisson, M. T., Roderick, T. H., and Doolittle, D. P.: Recombination percentages and chromosomal assignments. In M. F. Lyon and A. G. Searle (eds.); *Genetic Variants and Strains of the Laboratory Mouse*: 2nd. edition, pp. 432-505, Oxford University Press, Oxford, England, 1989.
- DeFranco, D., Morris, S. M., Jr., Leonard, C. M., and Gluecksohn-Waelsch, S.: Metallothionein mRNA expression in mice homozygous for chromosomal deletions around the albino locus. *Proc Natl Acad Sci USA* 85: 1161-1164, 1988.
- Devereux, J., Haeberli, P., and Smithies, O.: A comprehensive set of sequence analysis programs for the VAX. *Nucl Acids Res* 12: 387-395, 1984.
- Disteche, C. M., and Adler, D.: Localization of cloned mouse chromosome-7-specific DNA to lethal albino deletions. *Somatic Cell Mol Genet* 10: 211-215, 1984.

- Donner, M. E., Leonard, C. M., and Gluecksohn-Waelsch, S.: Developmental regulation of constitutive and inducible expression of hepatocyte-specific genes in the mouse. *Proc Natl Acad Sci USA* 85: 3049-3051, 1988.
- Doolittle, D. P., Davisson, M. T., Roderick, T. H., and Hillyard, A. L.: The locus map of the mouse. *Mouse News Lett* 87: 14-27, 1990.
- Erickson, R. P., Gluecksohn-Waelsch, S., and Cori, C. F.: Glucose-6-phosphatase deficiency caused by radiation-induced alleles at the albino locus in the mouse. *Proc Nat Acad Sci USA* 59: 437-444, 1968.
- Fallstrom, S. P., Lindblad, B., Lindstedt, S., and Steen, G.: Hereditary tyrosinemia-fumarylacetoacetase deficiency. *Pediatr Res* 13: 78, 1979.
- Frankel, W. N., Stoye, J. P., Taylor, B. A., and Coffin, J. M.: Genetic identification of endogenous polytropic proviruses by using recombinant inbred mice. *J Virol* 63: 3810-3821, 1989.
- Friedman, K. D., Rosen, N. L., Newman, P. J., and Montgomery, R. R.: Screening of λ gt11 libraries. In M. A. Innis, D. H. Gelfand, J. J. Sninsky, and T. J. White (eds.); *PCR Protocols: A Guide to Methods and Applications*, pp. 253-258, Academic Press, Inc., San Diego, CA, 1990.
- Frischauf, A. M., Lehrach, H., Poustka, A., and Murray, N.: Lambda replacement vectors carrying polylinker sequences. *J Mol Biol* 170: 827-842, 1983.
- Gatmaitan, Z., Lewis, S., Turchin, H., and Arias, I. M.: Premature development of ligandin (GSH transferase B) in mice with an inherited defect in endoplasmic reticulum-Golgi structure and function. *Biochem Biophys Res Commun* 75: 337-341, 1977.

- Gluecksohn-Waelsch, S.: Genetic control of morphogenetic and biochemical differentiation: lethal albino deletions in the mouse. *Cell* 16: 225-237, 1979.
- Gluecksohn-Waelsch, S.: Regulatory genes in development. *Trends Genet* 3: 123-127, 1987.
- Gluecksohn-Waelsch, S.G.: Private communication. *Mouse News Lett* 83: 149, 1989.
- Goldfeld, A. E., Rubin, C. S., Siegel, T. W., Shaw, P. A., Schiffer, S. G., and Gluecksohn-Waelsch, S.: Genetic control of insulin receptors. *Proc Natl Acad Sci USA* 78: 6359-6361, 1981.
- Goldfeld, A. E., Firestone, G. L., Shaw, P. A., and Gluecksohn-Waelsch, S.: Recessive lethal deletion on mouse chromosome 7 affects glucocorticoid receptor binding activities. *Proc Natl Acad Sci USA* 80: 1431-1434, 1983.
- Green, M. C.: Catalog of mutant genes and polymorphic loci. In M. F. Lyon and A. G. Searle (eds.); *Genetic Variants and Strains of the Laboratory Mouse*: 2nd. edition, pp. 12-403, Oxford University Press, Oxford, England, 1989.
- Gunning, P., Ponte, P., Okayama, H., Engel, J., Blau, H., and Kedes, L.: Isolation and Characterization of full-length cDNA clones for human α -, β -, and γ -actin mRNAs: skeletal but not cytoplasmic actins have an amino-terminal cysteine that is subsequently remove. *Mol Cell Biol* 3: 787-795, 1983.
- Harland, R.: Growth factors and mesoderm induction. *Trends Genet* 4: 62-63, 1988.

- Herrmann, B. G., Barlow, D. P., and Lehrach, H.: A large inverted duplication allows homologous recombination between chromosomes heterozygous for the proximal t complex inversion. *Cell* 48: 813-825, 1987.
- Herrman, B. H., Labeit, S., Poustka, A., King, T. R., and Lehrach, H.: Cloning of the *T* gene required in mesoderm formation in the mouse. *Nature* 343: 617-622, 1990.
- Huang, A., Campbell, C. E., Bonetta, L., McAndrews-Hill, M. S., Chilton-MacNeill, S., Coppes, M. J., Law, D. J., Feinberg, A. P., Yeger, H., and Williams, B. R. G.: Tissue, developmental, and tumor-specific expression of divergent transcripts in Wilms tumor. *Science* 250: 991-994, 1990.
- Johnson, D. K., Hand, R. E., Jr., and Rinchik, E. M.: Molecular mapping within the mouse albino-deletion complex. *Proc Natl Acad Sci USA* 86: 8862-8866, 1989.
- Killary, A. M. and Fournier, R. E. K.: A genetic analysis of extinction: trans-dominant loci regulate expression of liver-specific traits in hepatoma hybrid cells. *Cell* 38: 523-534, 1984.
- Kunkel, L. M., and coauthors: Analysis of deletions in DNA from patients with Becker and Duchenne muscular dystrophy. *Nature* 322: 73-77, 1986.
- Kusukawa, N., Uemori, T., Asada, K., and Kato, I. Rapid and Reliable Protocol for Direct Sequencing of Material Amplified by the Polymerase Chain Reaction. *Biotechniques* 9: 66-72, 1990.

- Kwon, B. S., Haq, A. K., Pomerantz, S. H., and Halaban, R.: Isolation and sequence of a cDNA clone for human tyrosinase that maps at the mouse *c*-albino locus. *Proc Natl Acad Sci USA* 84: 7473-7477, 1987.
- Kwon, B. S., Wakulchik, M., Haq, A. K., Halaban, R., and Kestler, D.: Sequence analysis of mouse tyrosinase cDNA and the effect of melanotropin on its gene expression. *Biochem Biophys Res Commun* 153: 1301-1309, 1988.
- Lalley, P. A., Davisson, M. T., Graves, J. A. M., O'Brien, S. J., Womack, J. E., Roderick, T. H., Creau-Goldberg, N., Hillyard, A. L., Doolittle, D. P., and Rogers, J. A.: Report of the committee on comparative mapping. *Cytogenet Cell Genet* 51: 503-532, 1989.
- Lem, J., Chin, A. C., Thayer, M. J., Leach, R. J., and Fournier, R. E. K.: Coordinate regulation of two genes encoding gluconeogenic enzymes by the trans-dominant locus *Tse-1*. *Proc Natl Acad Sci USA* 85: 7302-7306, 1988.
- Lewis, S. E., Turchin, H. A., and Wojtowicz, T. E.: Fertility studies of complementing genotypes at the albino locus of the mouse. *J Reprod Fertil* 53: 197-202, 1978.
- Lindblad, B., Lindstedt, S., and Steen, G.: On the enzymic defects in hereditary tyrosinemia. *Proc Natl Acad Sci USA* 74: 4641-4645, 1977.
- Lindsay, S., and Bird, A. P.: Use of restriction enzymes to detect potential gene sequences in mammalian DNA. *Nature* 327: 336-338, 1987.
- Manabe, S., Sassa, S., and Kappas, A.: Hereditary tyrosinemia: formation of succinylacetone-amino acid adducts. *J Exp Med* 162: 1060-1074, 1985.

- Maniatis, T., Fritsch, E. F., and Sambrook, J. (eds.): *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York, 1982.
- Margolis, F. L.: Olfactory marker protein: from PAGE band to cDNA clone. *Trends Neurosci* 8: 542-546, 1985.
- McKnight, S. L., Lane, M. D., and Gluecksohn-Waelsch, S.: Is CCAAT/enhancer-binding protein a central regulator of energy metabolism? *Genes Dev* 3: 2021-2024, 1989.
- Mitchell, G., Larochelle, J., Lambert, M., Michaud, J., Grenier, A., Ogier, H., Gauthier, M., Lacroix, J., Vanasse, M., Larbrisseau, A., Paradis, K., Weber, A., Lefevre, Y., Melancon, S., and Dallaire, L.: Neurologic crises in hereditary tyrosinemia. *N Engl J Med* 322: 432-437, 1990.
- Monaco, A. P., Neve, R. L., Collati-Feener, C., Bertelson, C. J., Kurnit, D. M., and Kunkel, L. M.: Isolation of candidate cDNAs for portions of the Duchenne muscular dystrophy gene. *Nature* 323: 646-650, 1986.
- Morris, S. M., Jr., Moncman, C. L., Kepka, D. M., Nebes, V. L., Diven, W. F., Dizikes, G. J., Cederbaum, S. D., and DeFranco, D.: Effects of deletions in mouse chromosome 7 on expression of genes encoding the urea-cycle enzymes and phosphoenolpyruvate carboxykinase (GTP) in liver, kidney, and intestine. *Biochem Genet* 26: 769-781, 1988.
- Niswander, L., Yee, D., Rinchik, E. M., Russell, L. B., and Magnuson, T.: The albino deletion complex and early postimplantation survival in the mouse. *Development* 102: 45-53, 1988.
- Niswander, L., Yee, D., Rinchik, E. M., Russell, L. B., and Magnuson, T.: The albino-deletion complex in the mouse defines genes necessary

- for development of embryonic and extraembryonic ectoderm. *Development* 105: 175-182, 1989.
- Niswander, L., Kelsey, G., Schedl, A., Ruppert, S., Sharan, S. K., Holdener-Kenny, B., Rinchik, E. M., Edstrom, J.-E., and Magnuson, T. Molecular mapping of albino deletions associated with early embryonic lethality in the mouse. *Genomics* 9: 162-169, 1991.
- Paterno, G. D., Gillespie, L. L., Dixon, M. S., Slack, J. M. W., and Heath, J. K.: Mesoderm-inducing properties of INT-2 and kFGF: two oncogene-encoded growth factors related to FGF. *Development* 106: 79-83, 1989.
- Paul, D., Kwon, B. S., Hohne, M., Tonjes, R., Haq, A. K., and Hoffmann, B.: Establishment and partial characterization of SV40 virus-immortalized hepatocyte lines of normal and lethal mutant mice carrying a deletion on chromosome 7. *J Cell Physiol* 139: 599-609, 1989.
- Petersen, D. D., Gonzalez, F. J., Rasic, V., Kozak, C. A., Lee, J.-Y., Jones, J. E., and Nebert, D. W.: Marked increases in hepatic NAD(P)H:oxidoreductase gene transcription and mRNA levels correlated with a mouse chromosome 7 deletion. *Proc Natl Acad Sci USA* 86: 6699-6703, 1989.
- Phaneuf, D., Labelle, Y., Berube, D., Arden, K., Cavanee, W., Gagne, R., and Tanguay, R. M.: Cloning and expression of the cDNA encoding human fumarylacetoacetate hydrolase, the enzyme deficient in hereditary tyrosinemia: assignment of the gene to chromosome 15. *Am J Hum Genet* 48: 525-535, 1991.

- Poustka, A. and Lehrach, H.: Jumping libraries and linking libraries: the next generation of molecular tools in mammalian genetics. *Trends Genet* 2: 174-179, 1986.
- Rinchik, E. M.: Chemical mutagenesis and fine-structure functional analysis of the mouse genome. *Trends Genet* 7: 15-21, 1991.
- Rinchik, E. M., Russell, L. B., Copeland, N. G., and Jenkins, N. A.: Molecular genetic analysis of the *dilute-short ear (d-se)* region of the mouse. *Genetics* 112: 321-342, 1986.
- Rinchik, E. M., Machanoff, R., Cummings, C. C., and Johnson, D. K.: Molecular cloning and mapping of the ecotropic leukemia provirus *Emv-23* provides molecular access to the albino-deletion complex in mouse chromosome 7. *Genomics* 4: 251-258, 1989.
- Rinchik, E. M., and Russell, L. B.: Germ-line deletion mutations in the mouse: tools for intensive functional and physical mapping of regions of the mammalian genome. In K.E. Davies and S.M. Tilghman (eds.); *Genome Analysis Volume 1: Genetic and Physical Mapping*, pp. 121-158, Cold Spring Harbor Laboratory Press, Plainview, New York, 1990.
- Rinchik, E. M., Carpenter, D. A., and Selby, P. B.: A strategy for fine-structure functional analysis of a 6- to 11-centimorgan region of mouse chromosome 7 by high-efficiency mutagenesis. *Proc Natl Acad Sci USA* 87: 896-900, 1990.
- Rinchik, E. M., Johnson, D. K., Margolis, F. L., Jackson, I. J., Russell, L. B., and Carpenter, D. A.: Reverse genetics in the mouse and its application to the study of deafness. *Proc NY Acad Sci*: in press, 1991.

- Rogers, K. E., Dasgupta, P., Gubler, U., Grillo, M., Khew-Goodall, Y. S., and Margolis, F. L.: Molecular cloning and sequencing of a cDNA for olfactory marker protein. *Proc Natl Acad Sci USA* 84: 1704-1708, 1987.
- Royer-Pokora, B., Kunkel, L. M., Monaco, A. P., Goff, S. C., Newburger, P. E., Baehner, R. L., Cole, F. S., Curnutte, J. T., and Orkin, S. H.: Cloning the gene for an inherited human disorder-chronic granulomatous disease-on the basis of its chromosomal location. *Nature* 322: 32-38, 1986.
- Ruppert, S., Muller, G., Kwon, B., and Schutz, G.: Multiple transcripts of the mouse tyrosinase gene are generated by alternative splicing. *EMBO J* 7: 2715-2722, 1988.
- Ruppert, S., Boshart, M., Bosch, F. X., Schmid, W., Fournier, R. E. K., and Schutz, G.: Two genetically defined *trans*-acting loci coordinately regulate overlapping sets of liver-specific genes. *Cell* 61: 895-904, 1990.
- Russell, L. B.: Definition of functional units in a small chromosomal segment of the mouse and its use in interpreting the nature of radiation-induced mutations. *Mutat Res* 11: 107-123, 1971.
- Russell, L. B. and Cacheiro, N. L. A.: The *c*-locus region of the mouse: genetic and cytological studies of small and intermediate deficiencies. *Genetics* 86: s53-s54, 1977.
- Russell, L. B. and DeHamer, D. L.: Complementation analysis of *c* locus lethals in the mouse. *Genetics* 74: s236, 1973.
- Russell, L. B., Montgomery, C. S., and Raymer, G. D.: Analysis of the albino-locus region of the mouse: IV. Characterization of 34 deficiencies. *Genetics* 100: 427-453, 1982.

- Russell, W. L.: X-ray-induced mutations in mice. *Cold Spring Harbor Symp Quant Biol XVI*: 327-336, 1951.
- Sanger, F., Niklen, S. and Coulson, A. R.: DNA-sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 74: 5463-5467, 1977.
- Sassa, S., and Kappas, A.: Hereditary tyrosinemia and the heme biosynthetic pathway: profound inhibition of δ -aminolevulinic acid dehydratase activity by succinylacetone. *J Clin Invest* 71: 625-634, 1983.
- Saunders, A. M., and Seldin, M. F.: A molecular genetic linkage map of mouse chromosome 7. *Genomics* 8: 525-535, 1990.
- Slack, J. M. W.: Growth factors as inducing agents in early *Xenopus* development. *J. Cell Sci Suppl* 13: 119-130, 1990.
- Sodroski, J. G., Goh, W. C., and Haseltine, W. A.: Transforming potential of a human protooncogene (*c-fps/fes*) locus. *Proc Natl Acad Sci USA* 81: 3039-3043, 1984.
- Tanguay, R. M., Valet, J. P., Lescault, A., Duband, J. L., Laberge, C., Lettre, F., and Plante, M.: Different molecular basis for fumarylacetoacetate hydrolase deficiency in the two clinical forms of hereditary tyrosinemia (type I). *Am J Hum Genet* 47: 308-316, 1990.
- Thaler, M. M.: Genetically determined abnormalities of microsomal enzymes in liver of mutant newborn mice. *Biochem and Biophys Res Commun* 72: 1244-1250, 1976.
- Thonsen, G., Woolf, T., Whitman, M., Sokol, S., Vaughan, J., Vale, W., and Melton, D. A.: Activins are expressed in *Xenopus* embryogenesis and can induce axial mesoderm and anterior structures. *Cell* 63: 485-493, 1990.

- Trigg, M. J., and Gluecksohn-Waelsch, S.: Ultrastructural basis of biochemical effects in a series of lethal alleles in the mouse: neonatal and developmental studies. *J Cell Biol* 58: 549-563, 1973.
- Viskochil, D., Buchberg, A. M., Xu, G., Cawthon, R. M., Stevens, J., Wolff, R. K., Culver, M., Carey, J. C., Copeland, N. G., Jenkins, N. A., White, R., and O'Connell, P.: Deletions and a translocation interrupt a cloned gene at the neurofibromatosis type 1 locus. *Cell* 62: 187-192, 1990.
- Wilkinson, D. G., Peters, G., Dickson, C., and McMahon, A. P.: Expression of the FGF-related proto-oncogene *int-2* during gastrulation and neurulation in the mouse. *EMBO J* 7: 691-695, 1988.
- Wilkinson, D. G., Bhatt, S., and Herrmann, B. G.: Expression pattern of the mouse *T* gene and its role in mesoderm formation. *Nature* 343: 657-659, 1990.
- Xu, G., O'Connell, P., Viskochil, D., Cawthon, R., Robertson, M., Culver, M., Dunn, D., Stevens, J., Gesteland, R., White, R., and Weiss, R.: The Neurofibromatosis type 1 gene encodes a protein related to GAP. *Cell* 62: 599-608, 1990.
- Yanagisawa, K. O., Fujimoto, H., and Urushihara, H.: Effects of the Brachyury (*T*) mutation on morphogenetic movement in the mouse embryo. *Develop Biol* 87: 242-248, 1981.

APPENDIXES

APPENDIX A

TABLES

TABLE 1. Genetic Loci Defined by Complementation Analysis of Albino Deletions

Locus ^a Abbreviation	Locus Name	Previous Locus ^b Designation	Complementation ^c Groups Defining Extent of Locus
<i>msd</i>	mesoderm-deficient	IS	Ai, Dj, Dq, Fq
<i>hsdr-1</i>	hepatocyte-specific developmental regulation-1	NS	A, E, Dp, Di, Fp
<i>jdf</i>	juvenile development and fertility	JS	C
<i>eed</i>	embryonic-ectoderm development	EE	Bem, Di, Dj ^d
<i>exed</i>	extraembryonic-ectoderm development	ExEE	Bex, Di, Dj ^d
<i>pid</i>	preimplantation development	PS	Bp, Dp, Dq
<i>Mod-2</i> ^e	Malic enzyme; mitochondrial form	<i>Mod-2</i>	E
<i>sh-1</i> ^e	shaker - 1	<i>sh-1</i>	Fp, Fq

^a Locus designations established in CMC7 report (1991).

^b Previous "functional unit" or "DNA subregion" designations.

^c Complementation groups containing deletions that may break within the locus or help to define its position on the functional map.

^d It is uncertain whether Dj-group deletions include the *exed* locus.

^e Genetic loci identified before the complementation analysis.

TABLE 2. Hybridization Probes

Probe	Plasmid Clone	Locus	Restriction Fragment Used as Probe	Reference ^a
MTY811	pMTY811 (<i>Tyr</i>)	<i>c</i>	1.4-kb <i>Eco</i> RI (partial cDNA)	Kwon <i>et al.</i> (1988)
23.3	pRN175	<i>Emv-23</i>	1.4-kb <i>Eco</i> RI - <i>Hae</i> III ^b	Rinchik <i>et al.</i> (1989)
12A	p12A	<i>D7Was12</i>	5.2-kb <i>Eco</i> RI	Disteche & Adler (1984)
94.1	p94.1	<i>d^c</i>	0.5-kb <i>Eco</i> RI- <i>Hind</i> III	Rinchik <i>et al.</i> (1986)
NA ^d	NA ^d	Chicken tubulin	1.2-kb <i>Pst</i> I (cDNA) ^b	NA ^d
HF1	pHF1	human γ -actin	1.5-kb <i>Pst</i> I- <i>Bam</i> HI (partial cDNA)	Gunning <i>et al.</i> (1983)
RN202	pRN202	<i>D7Rn3</i>	1-kb <i>Xho</i> I- <i>Eco</i> RI	Machanoff & Rinchik (unpublished)
RN204	pRN204	<i>D7Rn1</i>	1.25-kb <i>Bam</i> HI- <i>Eco</i> RI	-
RN247	pRN247	<i>D7Rn2^e</i> (<i>Fah</i>)	0.16-kb <i>Bam</i> HI- <i>Hind</i> III	-
RN216	pRN216	<i>D7Rn2</i> (<i>Fah</i>)	1.45-kb <i>Bam</i> HI- <i>Eco</i> RI	-
RN254	pRN254	<i>D7Rn4</i>	0.9-kb <i>Eco</i> RI- <i>Hind</i> III ^b	-
RN259	pRN259	<i>D7Rn2</i> (<i>Fah</i>)	1.45-kb <i>Xba</i> I- <i>Eco</i> RI ^b	-
RN207.1	-	<i>c</i> (<i>Tyr</i>)	0.7-kb <i>Bam</i> HI- <i>Hind</i> III	-
RN226.1	-	<i>D7Rn2</i> (<i>Fah</i>)	2.3-kb <i>Eco</i> RI- <i>Sma</i> I	-
RN226.2	-	<i>D7Rn2</i> (<i>Fah</i>)	1.75-kb <i>Sma</i> I	-
RN226.3	-	<i>D7Rn2</i> (<i>Fah</i>)	0.55-kb <i>Sma</i> I- <i>Eco</i> RI	-

Table 2 Continued

Probe	Plasmid Clone	Locus	Restriction Fragment Used as Probe	Reference ^a
RN226.4	-	ND ^f	2.0-kb <i>Eco</i> RI	-
RN226.5	-	<i>D7Rn2</i> (<i>Fah</i>)	0.9-kb <i>Eco</i> RI	-
RN283	pRN283	<i>D7Rn1</i>	0.95-kb <i>Hind</i> III	-
RN271	pRN271	<i>D7Rn2</i> ^e (<i>Fah</i> ?)	1.8-kb <i>Eco</i> RI (cDNA) ^g	-
RN273	pRN273	<i>Fah</i> (<i>D7Rn2</i>)	1.45-kb <i>Eco</i> RI (cDNA)	-
210.3	pTY-1	<i>c</i> (<i>Tyr</i>)	0.35-kb <i>Eco</i> RI- <i>Hind</i> III ^b	Kwon <i>et al.</i> (1988)
Tyr14.EX.9	pTyr14:E6 ^h	<i>c</i> (<i>Tyr</i>)	0.9-kb <i>Eco</i> RI- <i>Xba</i> I ^b	Ruppert <i>et al.</i> (1988)
80.BE1.3	p80	human <i>Fes</i>	1.3-kb <i>Bgl</i> II- <i>Eco</i> RI ^b	Sodroski <i>et al.</i> (1984)
alb18	palb18	<i>D7Cw18</i>	0.145-kb <i>Eco</i> RI	Niswander <i>et al.</i> (1991)
N9.AH.8	pN9	<i>Pmv-31</i>	0.8-kb <i>Acc</i> I - <i>Hind</i> III ^b	Frankel <i>et al.</i> (1990) Rinchik <i>et al.</i> (unpublished)

^a Probes with no reference were derived in this study.

^b Restriction fragment derived from a portion of the plasmid insert.

^c No locus designation given. The locus maps proximal to the dilute (*d*) locus and within a number of prenatally lethal *d* deletions.

^d NA = not available. The prepared fragment was a gift from R. Woychik.

^e This study indicates that the *D7Rn2* locus is part of the *Fah* locus.

^f ND = not determined

^g It is not yet known whether this cDNA encodes a portion of the FAH enzyme.

^h 3.6-kb *Eco*RI subclone derived from l clone gTYR14.

TABLE 3. Sequences of Primers Used For cDNA Sequencing^a

Primer Name	Sequence (5'-3')
RW18	AGTCAGTGAGCGAGGAA
RW19	CGAGAGAGATGATAGGG
ORN148	TCCACTTACCTGTGGGATAC
ORN151	GATGAAGTCATCATAACAGGTCAC
ORN174	GTGTATGGTGCCTGCAGACT
ORN180	CTACCATCTGCAGGTCTAAC

^a Primers with sequences complementary to the ORN148, 151, 174, and 180 primer sequences were used for sequencing both strands of the cDNA.

Table 4. Wild-type Restriction-fragment Sizes Used to Construct a Long-range Map of the Region of Chromosome 7 Containing the *c*, *D7Was12*, and *Emv-23* Loci.^a

Probe	Locus	<i>Bss</i> HI	<i>Sac</i> II ^b	<i>Sac</i> II+ <i>Bss</i> HI	<i>Sac</i> II+ <i>Nae</i> I	<i>Nae</i> I ^c	<i>Nae</i> I+ ^d <i>Bss</i> HI	<i>Sma</i> I	<i>Sma</i> I+ <i>Sac</i> II	<i>Sma</i> I+ ^d <i>Nae</i> I	<i>Xho</i> I	<i>Xho</i> I+ <i>Sac</i> II	<i>Xho</i> I+ ^d <i>Nae</i> I	<i>Xho</i> I+ ^d <i>Bss</i> HI
12A	<i>D7Was12</i>	1080	550 (≥1200) ^e (≥1500) ^e	550	550	920 ≥1500 ^e	920	(65) ^f (100) (140) 350 (400) (430) ^f 550 (665) ^{d,g} (850) ^{d,g}	(65) ^f (100) (140) 350 (215) ^f (230) ^f 350 (400) (475) ^f 550	(100) (140) (290) (315) (375) 405 (455) (475) ^{d,f} (490) ^d 515 ^d (620) ^d 665 ^g (≥700) ^{d,h}	240 265 (290) (315) (375) 405 (455) (475) ^{d,f} (490) ^d 515 ^d (620) ^d 665 ^g (≥700) ^{d,h}	140 ^d 165 ^d (280) (305) (240) (280) (290) (305) (315)	140 165 (240) (265) (280) (290) (305) (315)	240 265 (290) ⁱ (300) ⁱ (315) ⁱ (325) ⁱ (375) ⁱ (405) ⁱ (455) ^{g,i}
23.3	<i>Emv-23</i>	1080	160	160	160	160 (1080) ≥1500 ^e	160	320 (430) (665) ^{d,g} (850) ^{d,g}	160	160 320	240 265 (290) (315) (375) 405 (455) (475) ^{d,f} (490) ^d 515 ^d (620) ^d 665 ^g (≥700) ^{d,h}	100 ^d (150) ^{d,j} (160) ^{d,j} (240) (265) (325) (375)	100 (150) ^j (160) ^j (240) (265) (325) (375)	240 265 (290) ⁱ (300) ⁱ (315) ⁱ (325) ⁱ (375) ⁱ (405) ⁱ (455) ^{g,i}
MTY811	<i>c</i>	1080	820 ≥1200 ^e ≥1500 ^e	370	370 (820)	920 ≥1500 ^e	920	(70) (100) (140) (240) (285) (310) 370	(70) (100) (140) (240) (285) (310) 370	(70) (100) (140) (240) (285) (310) 370	190 ^d (240) ^d (265) ^d (305) ^d (345) ^d	190 ^d (240) (265) (305) (345)	190 (240) (265)	

^a The restriction fragment sizes shown were obtained from DNAs of the outbred strain 2A (*c^{ch}/c^{ch}*), inbred strain C3H/RI, and F1 progeny of the cross C3H/RI × 101/RI. More than one fragment listed under a particular enzyme or combination of enzymes for a single probe indicates that the probe identifies a series of partially digested fragments. Parentheses designate fragments labelled with lesser intensity. The size values reported here were obtained by combining the average fragment sizes observed in both single and double digestions, and are based on data obtained from two or more different pulsed-field-gel (PFG) blots, except where indicated. Fragments below 1000 kb ranged in size an average of about 35 kb (between different DNAs and PFGs) with the parameters used in these experiments. Fragments larger than this had greater variability in size.

^b Additional *Sac*II fragments with much less hybridization signal than the fragments listed in the table are occasionally detectable with probes 12A and MTY811.

^c Several *Nae*I fragments with much less hybridization signal than the fragments listed in the table are sometimes detectable with each probe.

^d Fragment sizes obtained from a single PFG blot.

^e Partial-digestion products too large to be sized with the PFGE parameters used.

^f Fragment with very low hybridization signal; not always detectable.

^g Fragment located within a PFG region of poor resolution; therefore, its size is relatively inaccurate.

^h Probes 12A and 23.3 also hybridized to two or more fragments ≥700 kb that were not able to be sized with the PFGE parameters used.

ⁱ The resolution of these fragments in this PFG blot was not sufficient enough to exclude the possible existence of other fainter fragments in this region of the blot.

^j The 150-kb *Xho*I-*Sac*II and 160-kb *Sac*II fragments usually comigrate as a single, diffuse band. The 300- and 325-kb *Xho*I-*Bss*HII fragments reinforce the existence of these separate 150- and 160-kb fragments.

Table 5. Approximate Physical Distances (in kb) of the Distal Breakpoints of the A and Ai Deletions From the *c* and *D7Was12* Loci and From the Distal Breakpoints of the *c*^{24R75M} and *c*^{16DFiOD} Deletions.

	<i>D7Was12</i>	<i>c</i> ^{24R75M}	<i>c</i> ^{16DFiOD}	<i>c</i>
A1	40-250	0-210	0-285	N.D. ^a
A2	40-345	0-210	0-285	N.D.
A4	N.D.	N.D.	125-575	2.5 ^b -300
Ai1	N.D.	N.D.	0-395	15-450
Ai2	N.D.	N.D.	125-710	2.5 ^b -300

^a ND, not determined

^b Rinchik and Machanoff, unpublished data.

Table 6. RNAs Analyzed in Fig. 14.

Lane	Tissue	Genotype	Age
1	Whole fetus	C3Hf/Rl (+/+)	E16.5d ^a
2	Liver ^b	<i>cch</i> / <i>cch</i> or <i>cch</i> / <i>c</i> *, ^c	E18.5d
3	Liver ^b	A5/A5	E18.5d
4	Liver ^b	A1/A1	E18.5d
5	Liver ^b	E/E	E18.5d
6	Liver	2A (<i>cch</i> / <i>cch</i>)	Adult
7	Liver	C3Hf/Rl (+/+)	Adult
8	Liver	C3Hf/Rl (+/+)	8.5d
9	Kidney	2A (<i>cch</i> / <i>cch</i>)	Adult
10	Kidney	C3Hf/Rl (+/+)	Adult
11	Kidney	C3Hf/Rl (+/+)	8.5d
12	Liver	C3Hf/Rl (+/+)	Adult
13	Kidney	C3Hf/Rl (+/+)	Adult
14	Thymus	C3Hf/Rl (+/+)	8.5d
15	Liver	C3Hf/Rl (+/+)	8.5d
16	Kidney	C3Hf/Rl (+/+)	8.5d
17	Spleen	C3Hf/Rl (+/+)	8.5d
18	Brain	FVB/N (<i>c</i> / <i>c</i>)	Adult
19	Kidney	FVB/N (<i>c</i> / <i>c</i>)	Adult
20	Testis	FVB/N (<i>c</i> / <i>c</i>)	Adult
21	Salivary Gland	FVB/N (<i>c</i> / <i>c</i>)	Adult
22	Small Intestine	FVB/N (<i>c</i> / <i>c</i>)	Adult

^a "E" designates days of gestation.

^b each liver sample from two or more pooled 18.5-day fetal livers.

^c * designates either the A1, A5, or E deletion.

TABLE 7. Interspecific Backcross Mapping of *D7Rn4* and *Fes* With Respect to Mouse Chromosome-7 Markers *Igf1r*, *Tyr* (*c*), and *Hbb*.

Cross: (129/Rl X <i>M. spretus</i>)F ₁ X 129/Rl										
I	II	III	IV							
<i>Igf1r^s</i>	<i>Fes^s</i>	<i>D7Rn4^s</i>	<i>Tyr^s</i>	<i>Hbb^s</i>	X	<i>Igf1r^m</i>	<i>Fes^m</i>	<i>D7Rn4^m</i>	<i>Tyr^m</i>	<i>Hbb^m</i>
<i>Igf1r^m</i>	<i>Fes^m</i>	<i>D7Rn4^m</i>	<i>Tyr^m</i>	<i>Hbb^m</i>		<i>Igf1r^m</i>	<i>Fes^m</i>	<i>D7Rn4^m</i>	<i>Tyr^m</i>	<i>Hbb^m</i>
Allele Configuration*					Genotype†			Number of Progeny		
P					<i>Igf1r^s</i>	<i>Fes^s</i>	<i>D7Rn4^s</i>	<i>Tyr^s</i>	<i>Hbb^s</i>	78
					<i>Igf1r^m</i>	<i>Fes^m</i>	<i>D7Rn4^m</i>	<i>Tyr^m</i>	<i>Hbb^m</i>	76
I					<i>Igf1r^s</i>	<i>Fes^m</i>	<i>D7Rn4^m</i>	<i>Tyr^m</i>	<i>Hbb^m</i>	6
					<i>Igf1r^m</i>	<i>Fes^s</i>	<i>D7Rn4^s</i>	<i>Tyr^s</i>	<i>Hbb^s</i>	14
II					<i>Igf1r^s</i>	<i>Fes^s</i>	<i>D7Rn4^m</i>	<i>Tyr^m</i>	<i>Hbb^m</i>	1
					<i>Igf1r^m</i>	<i>Fes^m</i>	<i>D7Rn4^s</i>	<i>Tyr^s</i>	<i>Hbb^s</i>	1
III					<i>Igf1r^s</i>	<i>Fes^s</i>	<i>D7Rn4^s</i>	<i>Tyr^m</i>	<i>Hbb^m</i>	0
					<i>Igf1r^m</i>	<i>Fes^m</i>	<i>D7Rn4^m</i>	<i>Tyr^s</i>	<i>Hbb^s</i>	0
IV					<i>Igf1r^s</i>	<i>Fes^s</i>	<i>D7Rn4^s</i>	<i>Tyr^s</i>	<i>Hbb^m</i>	2
					<i>Igf1r^m</i>	<i>Fes^m</i>	<i>D7Rn4^m</i>	<i>Tyr^m</i>	<i>Hbb^s</i>	5
Total										183

*P = parental configuration; I = crossover between *Igf1r* and *Fes*; II = crossover between *Fes* and *D7Rn4*; III = crossover between *D7Rn4* and *Tyr*; IV = crossover between *Tyr* and *Hbb*.

On the basis of these data, the order is as follows (with genetic distances):

Igf1r-- (10.9 ± 2.3 cM) -- *Fes*-- (1.1 ± 0.8 cM) --[*D7Rn4*, *Tyr*] -- (3.8 ± 1.4 cM) -- *Hbb*.

† All segregants also carry a *Igf1r^m Fes^m D7Rn4^m Tyr^m Hbb^m* chromosome.

APPENDIX B**FIGURES**

Figure 1. A genetic/functional map of the *c*-region in chromosome 7 and the associated albino deletion mutations.

The map is a consolidation of much of the complementation, genetic, molecular, and morphological analyses reported to date (Russell et al., 1982; Gluecksohn-Waelsch, 1979; Rinchik et al., 1989; Johnson et al., 1989; Niswander et al., 1989; Rinchik and Russell, 1990; Saunders and Seldin, 1990). No correlation with physical distance is implied. Genetic marker loci (references given in Green, 1989 except where indicated) shown above the chromosome are: *tp*, taupe; *Fes*, feline sarcoma virus protooncogene (Sodroski et al., 1984; Saunders and Seldin, 1990); *c*, albino; *Emv-23*, integration site of the endogenous ecotropic murine leukemia provirus-23 (Rinchik et al., 1989); *Mod-2*, malic enzyme (mitochondrial form); *fit-1*, fitness-1 (Rinchik et al., 1990; Rinchik, 1991); *sh-1*, shaker-1; *Omp*, olfactory marker protein (Margolis, 1985; Rogers et al., 1987); and *Hbb*, globin (β chain). The chromosomal locations defined to date by deletion-mapping for the anonymous DNA loci *D7Cw18* (formerly *D7TM18*; Niswander et al., 1991), *D7Was12* (Disteche and Adler, 1984), and *D7Or1* (Johnson et al., 1989) are also shown above the chromosome. The bracket below *Fes* indicates the region within which the locus may be situated, as it has been mapped 1.1 - 4.4 cM proximal to *c* by *M. spretus*-*M. musculus* backcross analysis (Saunders and Seldin, 1990; Rinchik et al., in preparation) and outside of the albino deletions (Rinchik et al., in preparation), but has not yet been mapped with respect to the *tp* locus. Recombination frequencies in centimorgans (cM) between marker loci on the standard genetic map are approximately as follows: *tp*--(3)--*c*--(2)--*Mod-2*--(4)--*sh-1*--(2)--*Hbb*. These map distances correspond to those on the most recent mouse genome linkage map (Doolittle et al. 1990; Davisson et al., 1989); however, the estimates for the size of map intervals between the standard loci do vary (see Davisson et al., 1989). The genetic distance between the *c* and *Emv-23* loci was estimated to be less than 0.5 cM by an intraspecific cross (Rinchik et al., 1989).

The positions of those genetic loci [previously referred to as "functional units" (Russell *et al.*, 1982) or DNA "subregions" (Rinchik and Russell, 1990)] that are defined solely by complementation analyses of lethal albino (*c*)- locus deletion mutations are indicated below the chromosome and are described in Table 1. More detailed descriptions of these loci are given in the text and elsewhere (Rinchik and Russell, 1990; Johnson *et al.*, 1989; and Niswander *et al.*, 1988, 1989). Horizontal solid lines below the map indicate the presumed extents of deletions with respect to the functional map, and the shaded box indicates the uncertain extents of the Dj1 and Dj2 deletions. All the deletions that have breakpoints between the same two loci are shown with their breakpoints in the same position, but they have not yet been localized on a physical map. The deletion classifications, derived from the original ones used by Russell *et al.* (1982), are shown directly above the lines with letters designating complementation groups and numerals specifying independent deletion mutations within each group. If no more than one deletion is available in a complementation group, then no numeral is used in its designation (e.g., deletion C). On the basis of information reported by Gluecksohn-Waelsch (1979), Russell *et al.* (1982) assigned the *c*^{3H}, *c*^{6H} and *c*^{25H} deletions to complementation groups E, Bi, and Dp (or Dq), respectively, and subsequent embryological studies showed *c*^{6H} to fit into the Bex group. However, these three mutations have not been tested in the large-scale complementation analysis.

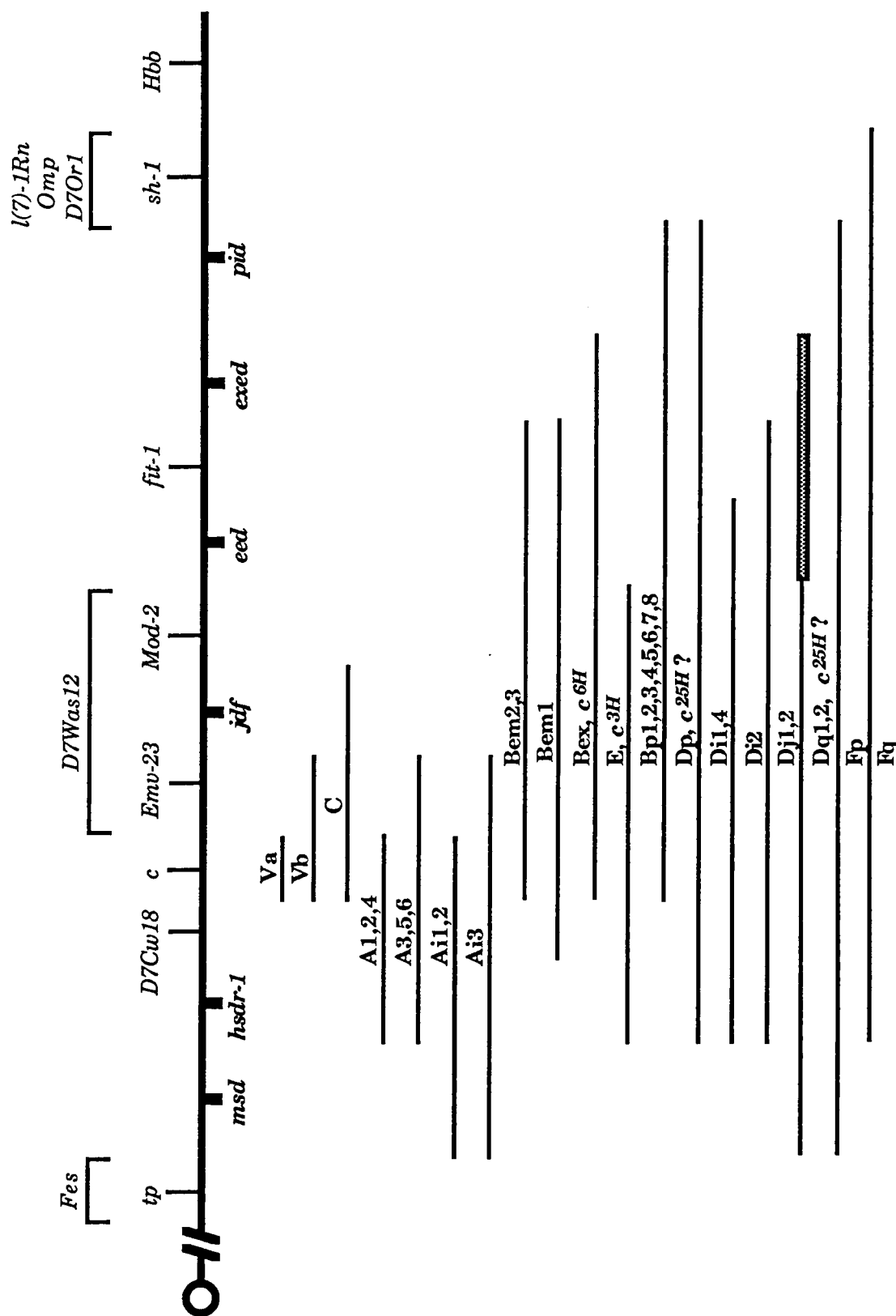


Figure 2. A genetic/functional map of a proximal portion of the *C*-region in chromosome 7 and some of the associated albino deletion mutations.

The map and scheme for representing the deletions have been described in Fig. 1. Only those deletions analyzed by large-fragment techniques and used for deletion-mapping newly derived breakpoint clones are shown.

Fes

D7Was12

tp

D7Cw18

c

Emv-23

Mod-2

Hbb

msd

hsdr-1

Va

Vb

A1, A2, A4

A3, A5, A6

Ai1, Ai2

Ai3

C

E, c^{3H}

Bem1

Figure 3. Deletion-mapping cloned DNA loci within the albino-deletion complex.

Panel A: Southern blot of wild-type (C3H γ /Rl, 101/Rl and 2A) and albino-deletion DNAs digested with *Bam*HI and hybridized consecutively with DNA probes 12A and MTY811, which recognize, respectively, the loci *D7Was12* and *c*. Genomic restriction-fragment sizes (in kb) are indicated to the right of the figures. The chromosome-9 probe 94.1 was used here as a lane control.

Panel B: Southern blot of *Eco*RI-digested DNAs hybridized with the MTY811 cDNA probe. Five *Eco*RI fragments are detectable with MTY811 in wild-type DNAs and are indicated to the right of the figure. Fragments of 0.7- and 4.7-kb and two 8.0-kb fragments (** indicating two comigrating fragments) are found in all the wild-type DNAs; whereas the largest fragment is polymorphic, i.e., 13.8 kb in C3H γ /Rl and 101/Rl DNAs and 10.5-kb in 2A (*c^{ch}/c^{ch}*) DNA. Both of the viable deletion mutants *c^{24R75M}* and *c^{16DFiOD}* are missing the 4.7-kb and 13.8-kb fragments and one of the 8-kb fragments, and a new rearranged 10-kb fragment is present in the *c^{16DFiOD}* mutant.

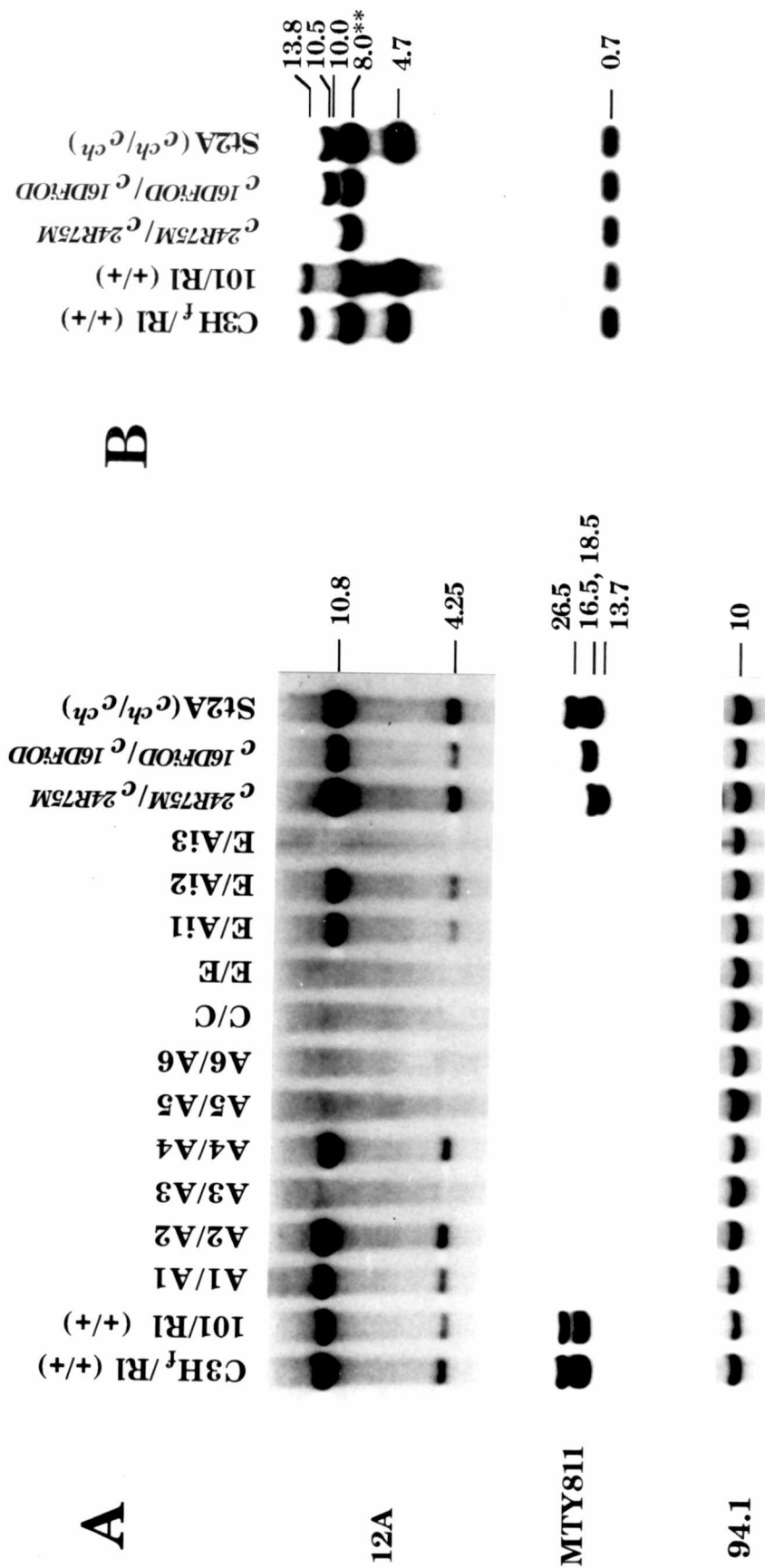
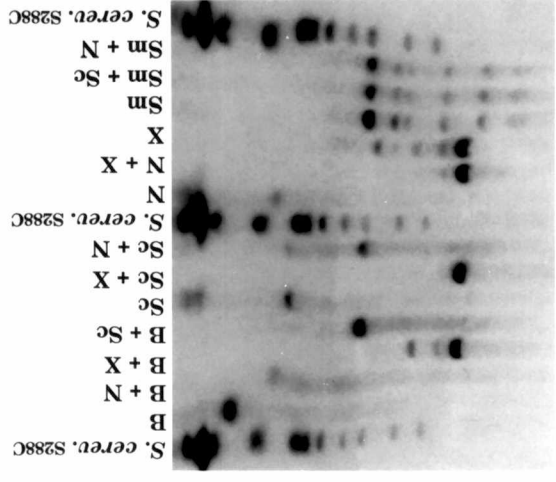


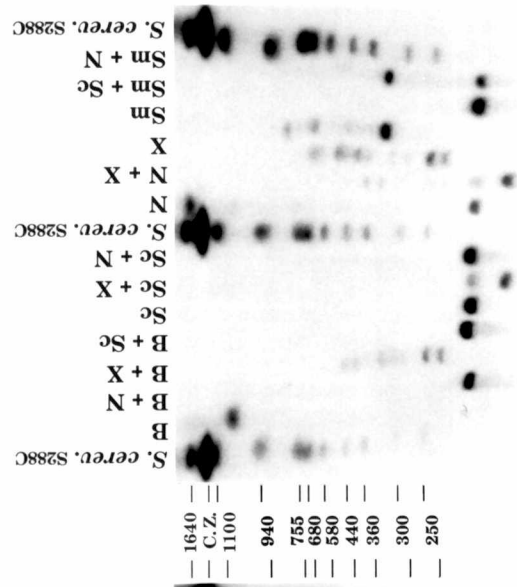
Figure 4. Pulsed-field gel analysis of the chromosome-7 region containing the *D7Was12*, *Emv-23*, and *c* loci.

A pulsed-field-gel-electrophoresis (PFGE) blot of single and double restriction-digests of outbred-stock 2A (*c^{ch}/c^{ch}*) DNA, used as a control DNA in this study, hybridized consecutively with DNA probes 12A (from *D7Was12*), 23.3 (from *Emv-23*), and MTY811 (from *c*). Restriction enzymes are *Bss*HII (B), *Sac*II (Sc), *Nae*I (N), *Sma*I (Sm), and *Xho*I (X). Molecular-weight markers are *Saccharomyces cerevisiae* strain S288C chromosomes with sizes indicated in kilobases (kb). The digested DNAs were electrophoresed in an LKB Pulsaphor apparatus with the hexagonal-array insert at 170 V (6 V/cm) for 18 hours with a pulse time of 25 seconds, followed by 26 hours with a pulse of 120 seconds. C.Z. indicates the beginning of the compression zone where DNA fragments of higher molecular weight are not well-resolved. The sizes of fragments below 240 kb were determined from other PFGE runs in which multimers of lambda DNA were used as size markers.

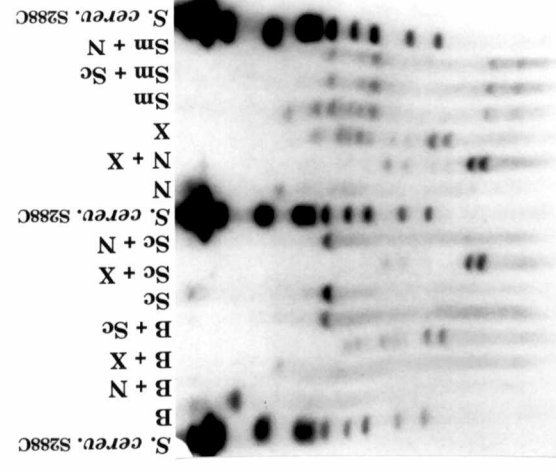
MTY811



23.3



12A



1640
 C.Z.
 1100
 940
 755
 680
 580
 440
 360
 300
 250

Figure 5. A long-range, wild-type restriction map of the region of chromosome 7 containing the *c*, *D7Was12*, and *Emv-23* loci.

A long-range restriction map of about 1700 kb of wild-type [strains 2A (*c^{ch}/c^{ch}*), C3H ϕ /Rl, and 101/Rl] mouse DNA constructed from the data obtained from pulsed-field gel blots (such as those in Fig. 4 and Table 4) hybridized with DNA probes recognizing the *c*, *D7Was12*, and *Emv-23* loci. Restriction enzyme sites are: B = *Bss*HII, S = *Sac*II, N = *Nae*I, A = *Sma*I, and X = *Xho*I. The circle at the left end of the map represents the centromere. The filled boxes above the map represent the restriction fragments containing the *c*, *D7Was12* and *Emv-23* loci. The horizontal bracketed lines indicate the sizes (in kb) of several restriction fragments referred to in the text. The *c* locus spans 70 kb of genomic DNA (Ruppert *et al.*, 1988), and is located somewhere within the 370-kb *Bss*HII-*Sac*II fragment, the *D7Was12* locus is located within the 100-kb *Sma*I-*Sac*II fragment, and the *Emv-23* locus within the 100-kb *Sac*II-*Xho*I fragment.

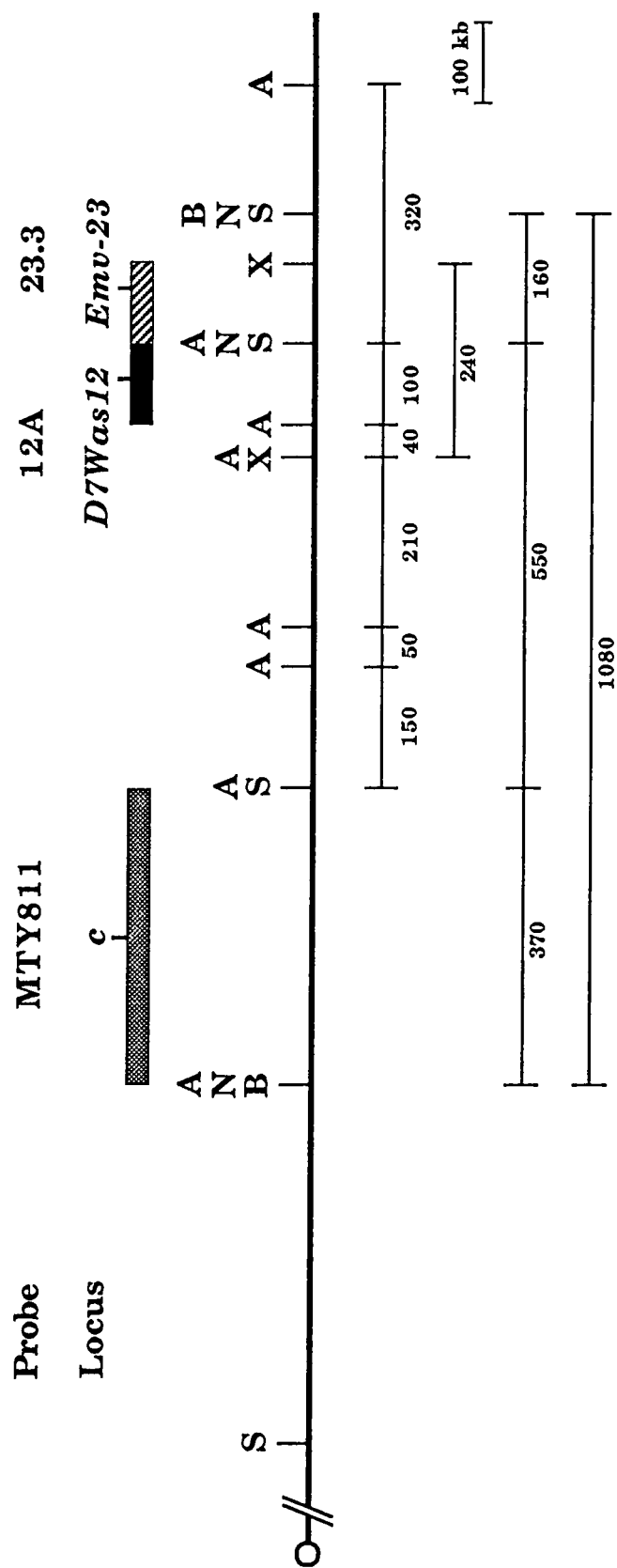


Figure 6. Molecular detection of the distal breakpoints of several viable, neonatally lethal, and prenatally lethal albino-deletion mutations.

Pulsed-field-gel-electrophoresis (PFGE) blots of DNAs of the indicated genotypes digested with the enzymes shown and hybridized with DNA probes 23.3 (panel A) and 12A (panels B,C and D), which recognize the *Emv-23* and *D7Was12* loci, respectively. The digested DNAs were separated with PFGE using either the BioRad CHEF-DRII apparatus (A and C) or the LKB Pulsaphor apparatus with the hexagonal-array insert (B and D). The electrophoretic parameters were as follows: (A) 170 V (5.1 V/cm), 225-second pulse-time for 4 hours, followed by 160-sec. pulse for 30 hrs.; (B) 170 V (6.2 V/cm), 110-sec. pulse for 48 hours; (C, left) 170V (5.1 V/cm), 225-sec. pulse for 6 hrs., then 160-sec. pulse for 30 hrs., then 80-sec. pulse for 10 hours; (C, right) 200 V (6 V/cm), 60-sec. pulse for 15 hrs., then 90-sec. pulse for 8 hrs.; (D) 170 V (6.2 V/cm), 60-sec. pulse for 46 hours (left) and for 44 hours (right). Molecular-weight (MW) markers are *Saccharomyces cerevisiae* chromosomes from the strains indicated or multimers of lambda-phage (c1857) DNA (in 48.5 kb steps). MW marker sizes are indicated in kilobases (kb) to the right of the blots shown in panels A and B and between the two blots shown in panel C. The sizes of the yeast chromosomes in panel D are the same as the six smallest chromosomes in panel A. The size of each genomic DNA fragment detected is also indicated in kb, followed in parentheses by the type of chromosome (wild-type or mutant) it is found in. Wild-type fragments are indicated by a +

in panels A, B and C and between the two blots of panel D. Duplicate DNA samples of each genotype were run adjacent to each other in the blot shown in panel B as an example of the variability in migration observed between different DNA preparations. Each lane contains genomic DNA from about 10^6 cells, except the lane marked with a * and the *c^{16DFiOD}* and *c^{24R75M}* lanes in panel A, which contain DNA from about 2×10^6 cells. ** in panel A indicates a degraded DNA sample. The * in panel D denotes a control DNA sample digested with *Sac*II alone. C. Z. = compression zone (see Fig. 3). The BioRad system, while giving straighter DNA lanes, was more susceptible to DNA overloading than the LKB system. As a result, DNA bands identified with the BioRad system were sometimes more diffuse (poorly resolved) and the estimated fragment sizes were sometimes greater than the bands and fragment sizes obtained with the LKB system. When this occurred, the fragment sizes reported were those obtained with the LKB system.

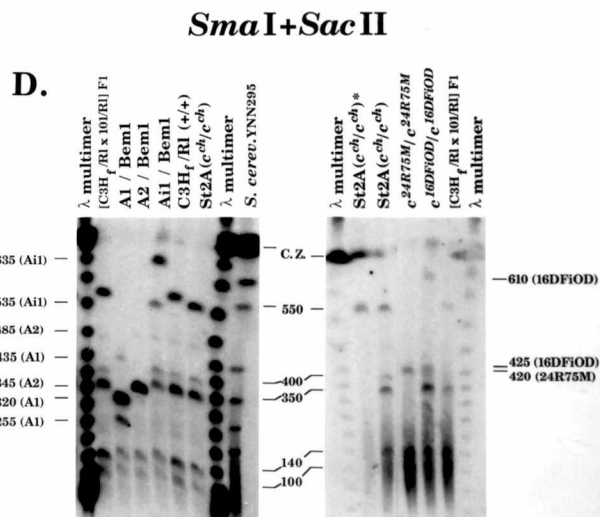
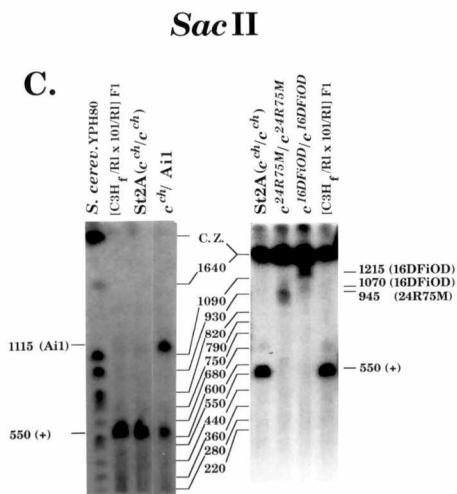
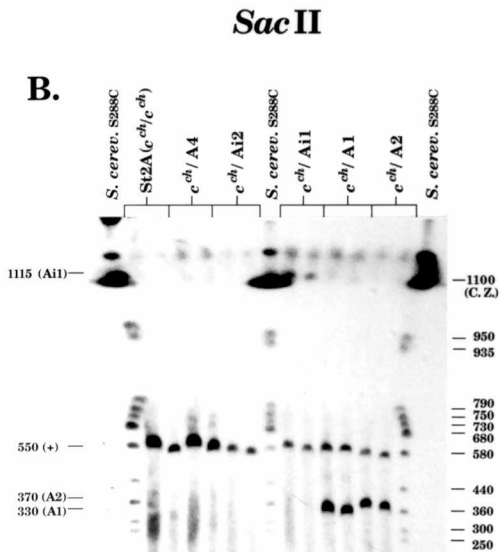
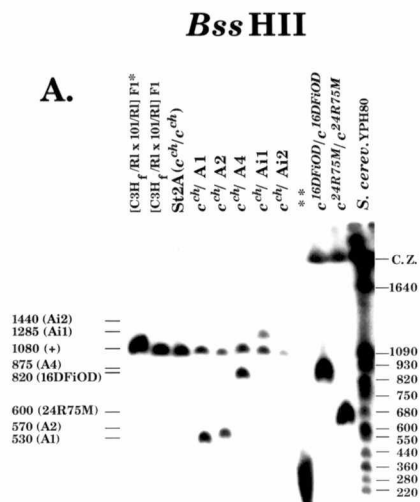


Figure 7. The approximate physical locations of the breakpoints of several viable (*c*^{24R75M} and *c*^{16DFiOD}), group-A, and group-Ai deletions on the long-range restriction map.

The long-range map has been described in the legend of Fig. 5. The horizontal bracketed lines indicate the sizes (in kb) of several restriction fragments referred to in the text. The dark, horizontal lines above the map represent the restriction fragments containing the *D7Was12* and *Emv-23* loci. The series of black boxes below the map represent the arrangement of genomic *Hind*III fragments that are recognized by the tyrosinase (*c*-locus) cDNA probe (from Ruppert *et al.*, 1988), and that are located in a 70-kb segment of DNA somewhere within the 370-kb *Bss*HII-*Sac*II fragment. The bars below the map represent the physical extents of the deletions that are labelled on the right. The unfilled portion of a bar indicates the extent of DNA known to be deleted by the particular mutation, and the filled portion indicates the smallest identifiable region of DNA known to contain the deletion breakpoint. The proximal breakpoint of the Bem1 deletion delimits the distal extent of the *hsdr-1* locus, and the Bem1 deletion extends distally beyond the physically mapped DNA segment shown here. The jagged lines extending vertically from the proximal ends of the two viable deletions, *c*^{24R75M} and *c*^{16DFiOD}, indicate the approximate positions of their breakpoints with respect to *Hind*III exon-containing fragments of the tyrosinase (*c*)locus (from Rinchik *et al.*, in preparation). The proximal ends of the group-A, -Ai, and -Bem deletions, although not physically mapped, are shown in order to orient the deletions with respect to the *msd* and *hsdr-1* loci, which have been included on the proximal end of the figure as separate (by hashmarks) from the DNA region that has been physically mapped.

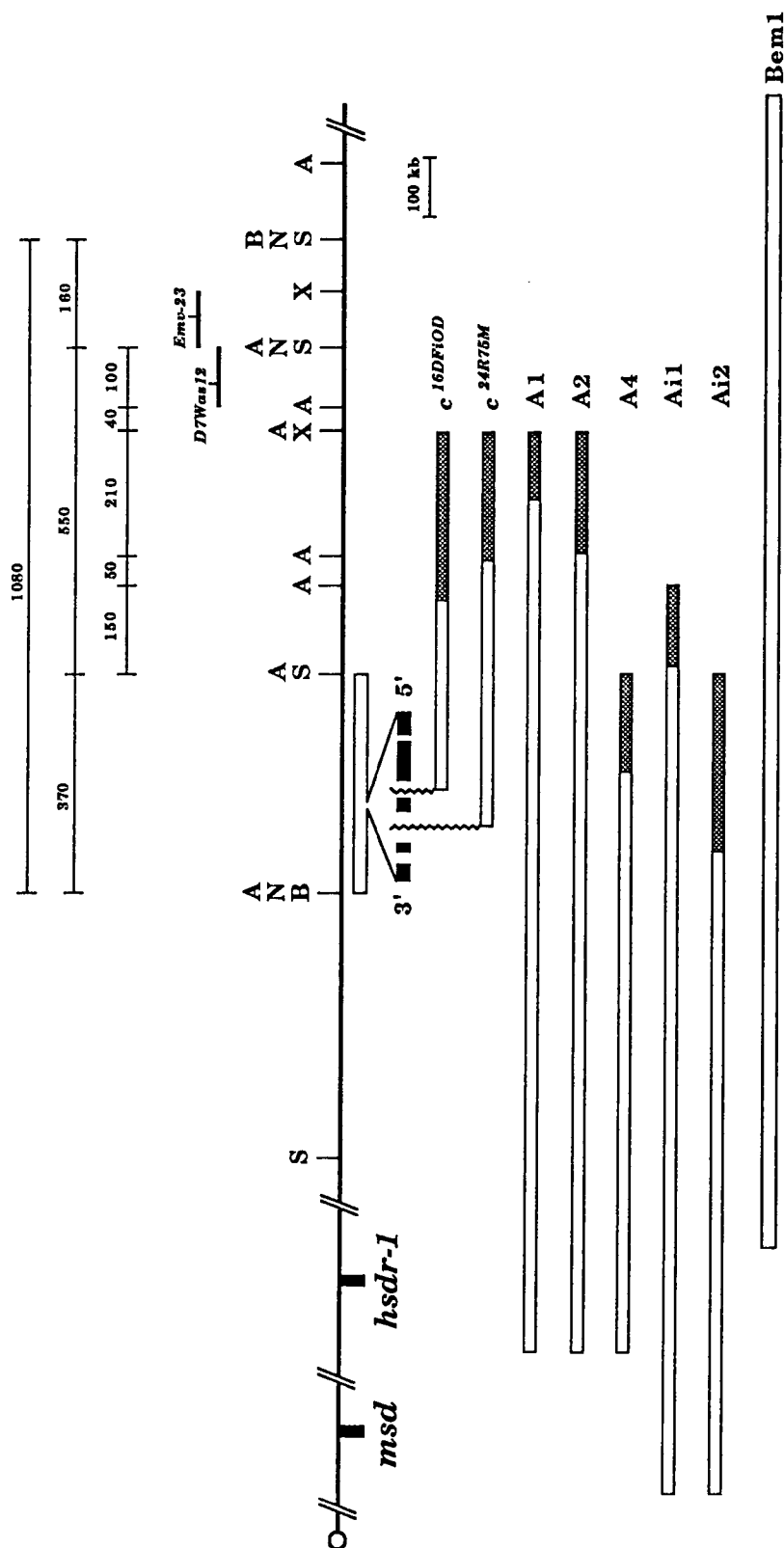
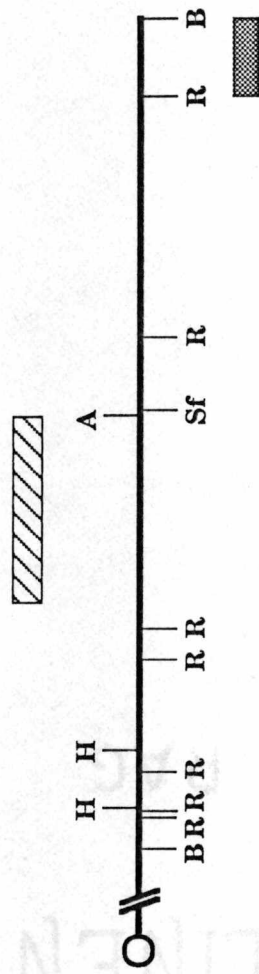


Figure 8. Restriction maps of deletion breakpoint-fusion fragments RN207 and RN226 cloned from *c*^{24R75M} and *c*^{14CoS} (A1) chromosomes.

RN207 and RN226 denote the 13.7-kb *Bam*HI breakpoint-fusion fragment cloned from *c*^{24R75M}/*c*^{24R75M} DNA and the 21.5-kb *Bam*HI breakpoint-fusion fragment cloned from *c*^{14CoS}/*c*^{14CoS} (A1/A1) DNA, respectively. The gray-shaded box below each chromosome highlights the 1.25-kb *Eco*RI-*Bam*HI fragment (probe RN204; locus *D7Rn1*) subcloned from the distal end of the 13.7-kb fragment and used to clone RN226. The black box delineates the 0.16-kb *Bam*HI-*Hind*III fragment (probe RN247; locus *D7Rn2*) subcloned from the proximal end of the 21.5-kb fragment. The unfilled boxes denote additional fragments which were used as probes. The abbreviated "RN" names of these probes are indicated below the boxes. Probes RN216 and RN283 have been cloned. Probes 226.1-226.5 have not been cloned, and were gel-isolated from the RN226 insert. Restriction sites are designated as follows: B = *Bam*HI, R = *Eco*RI, H = *Hind*III, A = *Sma*I, and Sf = *Sfi*I. The *c*^{14CoS} deletion breakpoint is located somewhere within the DNA region denoted by the vertically striped box, most likely within fragment 226.3. The *c*^{24R75M} deletion breakpoint occurs somewhere within the region designated by the diagonally striped box.

**$C^{24}R75M$
(RN207)**



**$C^{14}CoS$ (A1)
(RN226)**

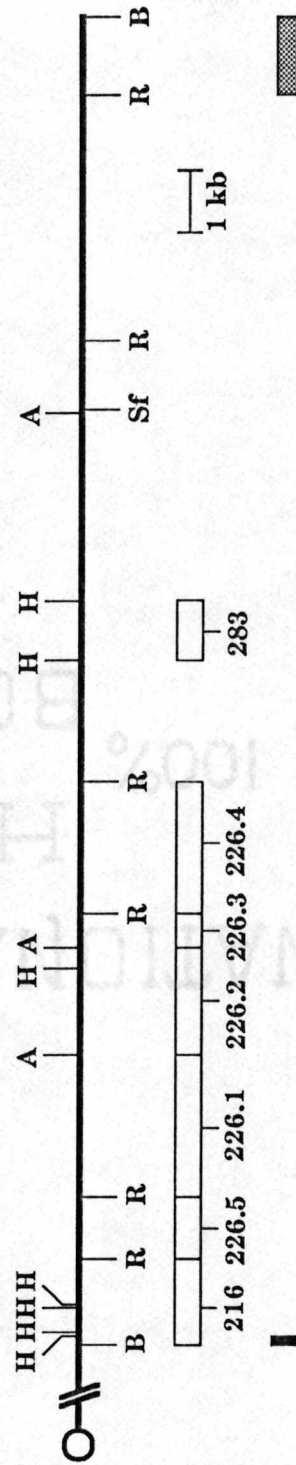
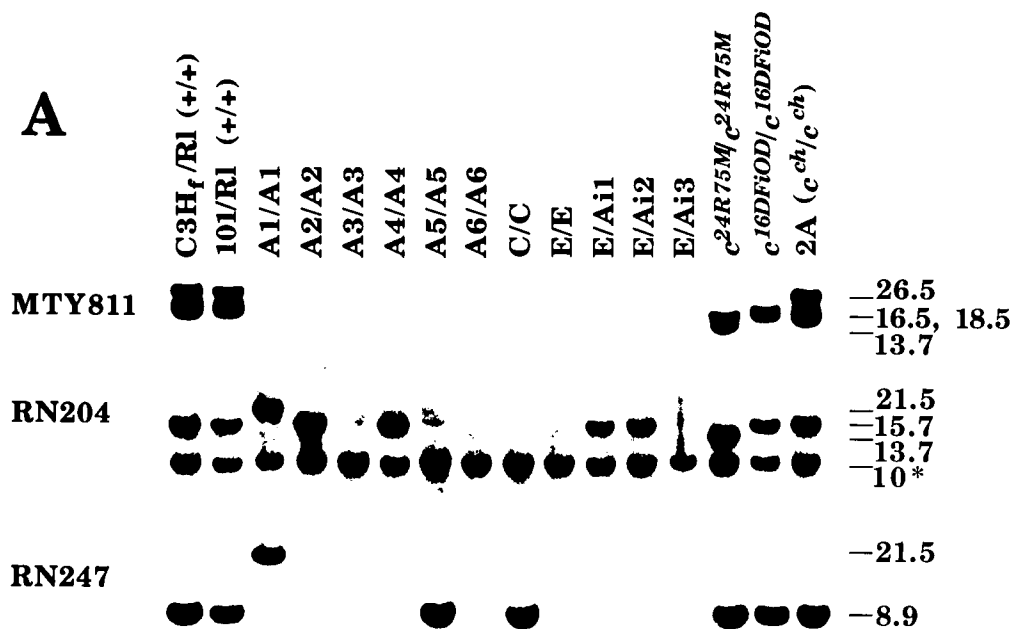


Figure 9. Deletion-mapping cloned DNA loci within the albino deletion complex.

Southern blots of wild-type (strains C3H ϕ /Rl, 101/Rl and 2A) and albino-mutant DNAs digested with *Bam*HI and hybridized consecutively with DNA probes MTY811, RN204, and RN247 (panel A), or digested with *Eco*RI and hybridized consecutively with DNA probes RN247, RN226.2, RN273, and RN271. Genomic restriction-fragment sizes (in kb) are indicated to the right of the figures. The * in panel A denotes the 10-kb fragment detected by the chromosome-9 probe, 94.1, used in both panels as a lane control. The * in panel B denotes the 4.6-kb deletion-breakpoint-fusion fragment observed in A1/A1 DNA. Probe 94.1 was hybridized to the blot shown in panel B concurrently with probe RN247.

A



B

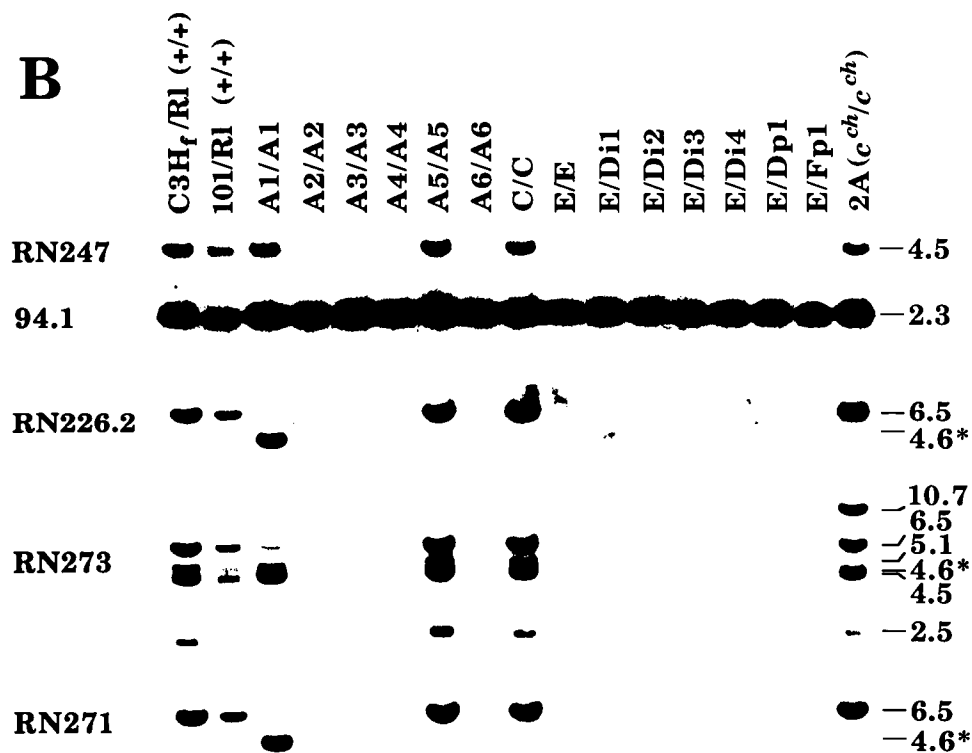


Figure 10. A proximal portion of the physical/functional map of the *c* region, as refined by deletion-mapping of the *D7Rn1* locus.

The map represents the portion of the functional map shown in Fig. 2, with an expanded view of the region between the *c* locus and the *jdf* locus that was physically mapped (Fig. 7). Loci, restriction-enzyme sites, and extents of deletions and restriction fragments are designated as in the legends to Figs. 1, 5, and 7. Descriptions of genetic loci are given in Table 1 and in the text, and descriptions of DNA loci are given in Table 2. The positions shown for the *D7Rn1* locus and the distal breakpoints of the *c*^{24R75M} and A1 deletions are merely to indicate the results of deletion-mapping of the locus and do not indicate exact physical locations. Because the locus maps distal to the A1 breakpoint, the best estimate of its physical location is shown by the possible location of the A1 breakpoint in Fig. 7. The possible distal limits of the *c*^{16DFiOD} and A2 deletions are shown slightly proximal to the *D7Rn1* locus, since these breakpoints were not detectable (by standard Southern-blot analysis) with a probe for this locus. The positions of the *c*^{24R75M} and *c*^{16DFiOD} deletion breakpoints within the *c* locus are also designated as in Fig. 7. The DNA and genetic loci located outside of the physically mapped region are positioned only relative to deletion breakpoints; definite physical distances are not implied.

Figure 11. The physical/functional map as refined by deletion-mapping of the *D7Rn2* and *Fah* loci.

The map has been described in the legend to Fig. 10. The proximal breakpoint of the Bem1 deletion defines the distal border of the *hsdr-1* locus, and the locus may extend proximally as far as the longest deletion of those that break between it and *msd*. The exon-containing fragments, identified by the cDNA probes derived from the *D7Rn2* locus, are completely contained within the DNA segment located between the proximal breakpoint of the A5 deletion and the nearest breakpoint located proximal to A1. The Di(4) designation indicates that there are four different deletions of the Di complementation group. The *c⁵¹DTD* deletion is a viable albino mutation, of the Vb complementation group, that has been mapped with respect to several of the loci (data not shown; Rinchik, *et al.*, in prep.).

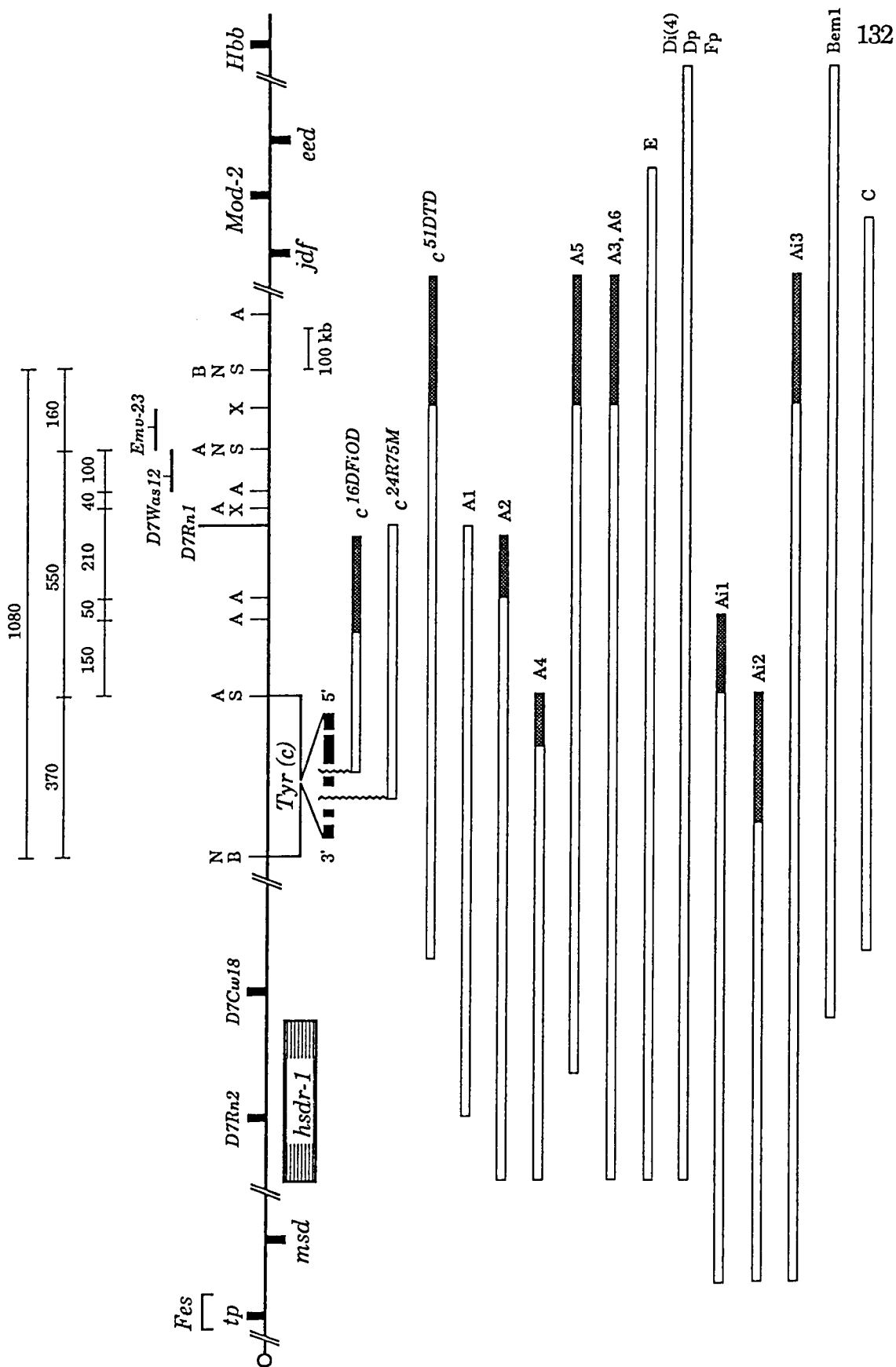
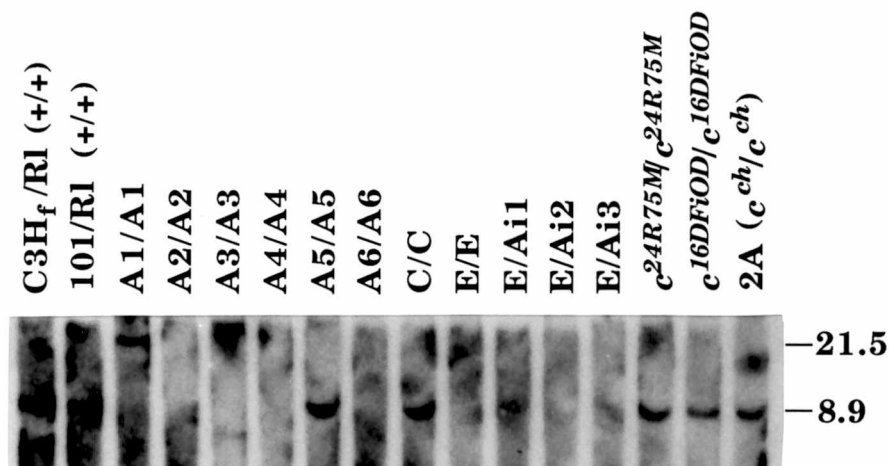


Figure 12. Probes RN226.1 and RN226.3 are derived from a portion of the A1-breakpoint-fusion fragment (RN226) located proximal to the breakpoint.

Southern blots of *Bam*HI-digested (panel A) and *Eco*RI-digested (panel B) wild-type (strains C3H₇/Rl, 101/Rl and 2A) and albino-mutant DNAs, of the genotypes indicated, were hybridized with the RN226.1 and RN226.3 probes, respectively. The blots are the same ones used in the hybridizations shown in Fig. 9B, which included lane controls. Therefore, lane controls were not needed in the hybridizations shown here. Genomic restriction-fragment sizes (in kb) are indicated to the right of the figures. The * denotes the 4.6-kb, *Eco*RI, deletion-breakpoint-fusion fragment observed in A1/A1 DNA. The blots were washed as usual, except that the wash temperature was raised to 72°C for the RN226.1 hybridization (panel A) to reduce the background signal obtained with this probe. Probe RN226.3 hybridizes stringently to a 1.5-kb *Eco*RI fragment, that does not map within the albino deletions, in addition to the 6.5- and 4.6-kb fragments, both of which contain the proximal A1 breakpoint.

A
(RN226.1)



B
(RN226.3)

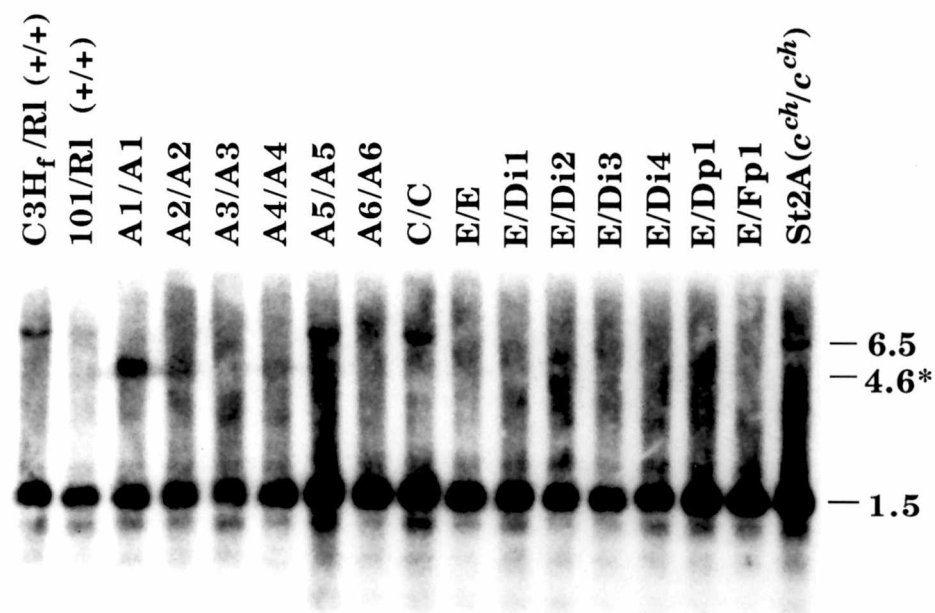
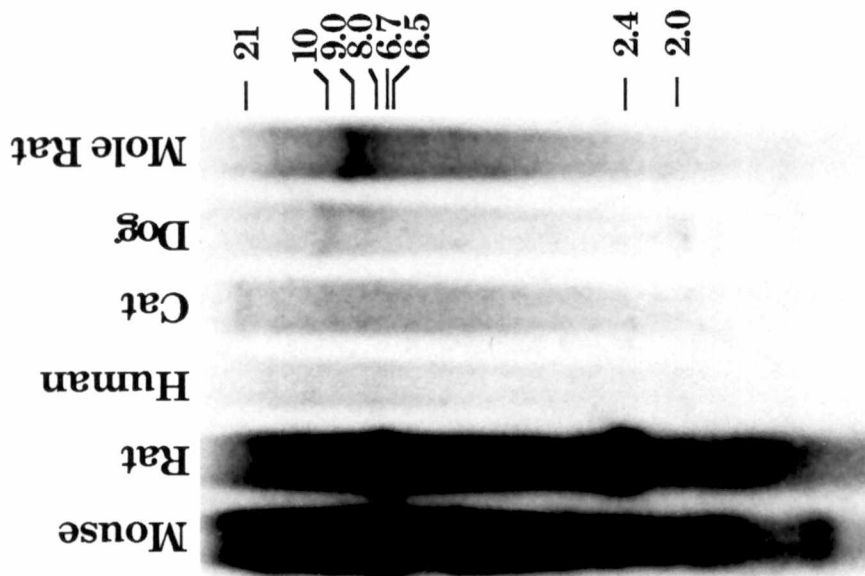
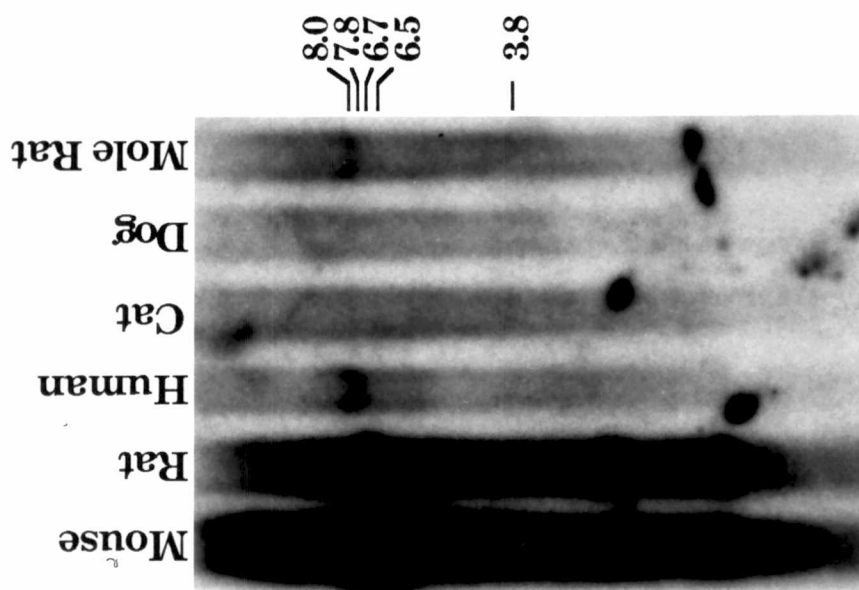


Figure 13. Probes RN226.1 and RN226.2 identify evolutionarily conserved DNA sequences.

Southern blots of *Eco*RI-digested DNAs from the species indicated were hybridized as usual with probes RN226.1 and RN226.2, except that the stringent-wash temperature was lowered to 50°C with probe 226.2 and 60°C with probe 26.1 to facilitate cross-species hybridization. The blots were exposed to X-ray film at -80°C for 6 days (RN226.2) and for 1.5 days (RN226.1). Genomic restriction-fragment sizes (in kb) are indicated to the right of the figures. The 8.0, 7.8, 6.7, 6.5, and 3.8-kb fragments in the RN226.2 panel are from mole rat, human, rat, mouse, and cat DNAs, respectively. The 21, 10, 9, 6.7, 6.5, 2.4, and 2.0-kb fragments in panel RN226.1 are from cat, dog, mole rat, rat, mouse, rat, and dog DNAs, respectively. Other fainter fragments are also seen in these, as well as human, DNAs.



RN226.1



RN226.2

Figure 14. Northern-blot analyses of the livers of wild-type and mutant mice, and of other wild-type tissues with genomic-DNA and cDNA probes.

The RNAs analyzed are summarized in Table 6. All samples are poly-(A+) RNAs (~2.5 ug per lane), except those in lanes 12 and 13, which contain total RNA (20 ug RNA per lane). Panels A-E show the results of consecutive hybridization of the same blot (lanes 1-11) with probes RN226.2 (A), RN226.1 (B), RN273 (C), RN271 (D), and chicken tubulin (lane control; E) (see Table 2). Panel F shows the results of hybridization of two different blots with probes RN226.2 and γ -actin (lane control, 2.1-kb transcript). The sizes (in kb) of the transcripts identified by each probe are indicated. The transcripts indicated are the most intensely labelled ones. Other more faintly labelled transcripts are also visible with some of the probes. The hybridized blots shown in panels A, B, and D were washed at stringent temperatures of 60-65°C, and were exposed to X-ray film, with an intensifying screen, at -80°C for 1 day. The blot shown in panel C (cDNA probe RN273) was exposed to film for 1 day at room temperature. Exposure of this blot to film, with an intensifying screen, at -80°C for 1 day also resulted in the identification of the larger transcripts that were detected with the cDNA RN271 (panel D). The blot in panel E was washed at 55°C and exposed to film at room temperature for 10 hours. The blots hybridized with RN226.2 in panel F were washed at 60°C and exposed to film, with an intensifying screen, at -80°C for 4 days (lanes 14-17) or 2 days (lanes 18-22). The same blots were stripped, rehybridized with the γ -actin (lane control) probe (2.1-kb transcript), washed at 60-65°C, and exposed to film at room temperature for 1-2 days.

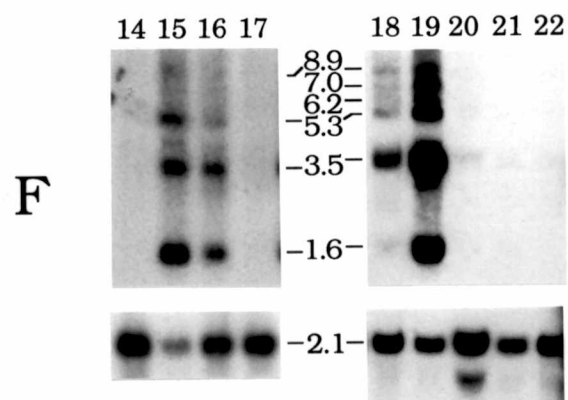
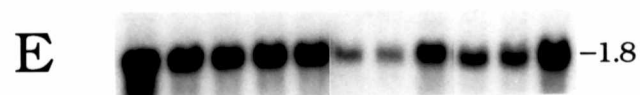
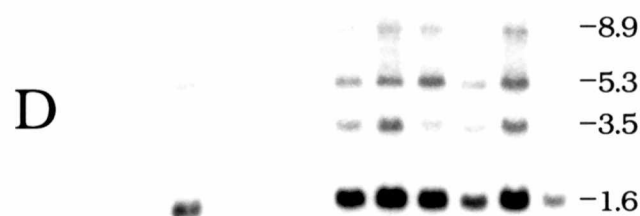
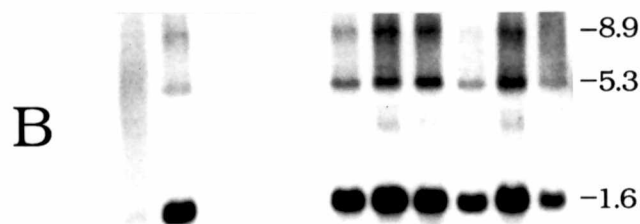
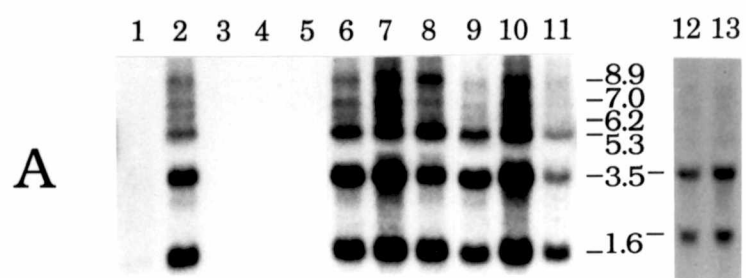


Figure 15. Comparison of the nucleotide sequence of mouse cDNA RN273 with the human *FAH* gene sequence reported by Phaneuf *et al.* (1991).

The partially completed mouse sequence is shown in the top line (designated by M before nucleotide #1), and the human sequence is shown in the bottom line (labelled H). Paired nucleotides that are identical in mouse and human are aligned by vertical lines. An "N" within the mouse sequence indicates that a nucleotide must be inserted in the mouse sequence at that position to prevent a frameshift with respect to the human nucleotide and amino-acid sequences (see also Fig. 16), with the following two exceptions. The "N" at position 967 of the mouse sequence represents either one or two "T"s. A second "T" here would cause a frameshift with respect to the human sequence, however. The "N" at position 1399 indicates that another "A" may or may not occur here. Other nucleotide ambiguity codes used are S (C or G), Y (C or T), and K (G or T). A colon between the mouse and human sequences at positions of ambiguity in the mouse sequence indicate possible nucleotide matches. The sequences are ~80% similar over their entire lengths and ~84% similar within the open-reading frame (ORF), which is delimited by the putative translation-initiation and -termination codons (enclosed within boxes). The underlined methionine codon at bases 267-269 of the human sequence was previously reported to be the translation-initiation codon of a shorter human cDNA (Agsterribe *et al.*, 1991), and the underlined GAC codon for aspartic acid (at bases 135-137 of the human sequence) was reported as a TAG (upstream stop codon) in the sequence of this shorter cDNA. The complete extent of the poly-(A+) tail in the mouse cDNA has not been determined.

```

M 1 .....CAACCCCTACCCGCCGTGGSTCTGC.....TGCCCGGTGCTC 38
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
H 1 AGTCCTGCTCTCCGCACGCCACCTTAGGCCCGCAGCCGTGCCGGGTGCTC 50
   . . . . .
39 GTCAGCATGTCCTTTATCCAGNTGGCNGAGGACTCCGACTTTCCCATCCA 88
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
51 TTCAGCATGTCCTTCATCCCGGTGGCCGAGGATTCCGACTTCCCATCCA 100
   . . . . .
89 AAACCTGCCCTATGGTGTTTTCTCCACTCAAAGCAACCCAAAGCCACGGA 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
101 CAACCTGCCCTACGGCGTCTTCTCGACCAGAGGCGACCCAAGACCGAGGA 150
   . . . . .
139 TTGGTGTAGCCATCGGTGACCAGATCTTGGACCTGAGTGTCAATAACAC 188
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
151 TAGGTGTGGCCATTGGCGACCAGATCCTGGACCTCAGCATCATCAAGCAC 200
   . . . . .
189 CTCTTTACCGGACCTGCCCTTTCCAAACATCAACATGTCTTCGATGAGAC 238
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
201 CTCTTTACTGGTCTGTCTCTCCAAACACCAGGATGTCTTCAATCAGCC 250
   . . . . .
239 AACTCTCAATAAACTTCATGGGTCTGGGTCAAGCTGCATGGAAGGAGGCAA 288
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
251 TAACTCAACAGCTTCATGGGCCTGGGTCAAGCTGCATGGAAGGAGGCGA 300
   . . . . .
289 GAGCATCCTTACAGAACTTACTGTCTGCCAGCCAAGCCCGGCTCAGAGAT 338
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
301 GAGTGTCTTCTGCAGAACTTGTCTGTCTGTGAGCCAAGCCAGGCTCAGAGAT 350
   . . . . .
339 GACAAGGAGCTTCGGCAGCGTGCATTACCTCCCAGGCTTCTGCGACAAT 388
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
351 GACACCGAACTTCGGAAGTGTGCATTATCTCCCAGGCTTCTGCCACGAT 400
   . . . . .
389 GCACCTTCCTGCTACCATAGGAGACTACACGGACTTCTACTCTTCTCGNC 438
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
401 GCACCTTCAGCCACCATAGGAGACTACACAGACTTCTATTCTCTCGGC 450
   . . . . .
439 AGCATCGCACCAATGTTGGCATTATGTTTCAAGGCAAGGAGANTGCGCTG 488
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
451 AGCATGCTACCAACGTCGGAATCATGTTTCAAGGACAAGGAGAATGCGTTG 500
   . . . . .
489 TTGCCAAATTGGCTCCACTTACCTGTGGGATACCATGGCCGAGCTTCTCTC 538
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
501 ATGCCAAATTGGCTGCCTTACCAGTGGGCTACCATGGCCGTGCCTCTCTC 550
   . . . . .
539 CATNGTGGTATCTGGAACCCGATTCTGAAGACCNATGGNGNAGATGAGAC 588
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
551 TGTCGTGGTGTCTGGCACCCCAATCCGAAGGCCCATGGGACAGATGAAAC 600
   . . . . .
589 CTGATAACTCAAAGCCTCCTGTGTATGGTGCCTGCAGACTYTTAGACATG 638
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
601 CTGATGACTCTAAGCCTCCCGTATATGGTGCCTGCAAGCTCTTGGACATG 650
   . . . . .
639 GAGTTGGAATGGCTTTTCTTCGTAGGCCCTGGGAACAGATTGCGAGAGCC 688
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
651 GAGCTGGAATGGCTTTTTTTGTAGGCCCTGGAAACAGATTGGGAGAGCC 700
   . . . . .
689 AATCCCCATTTCCAAAGCCCATGAACACATTTTCGGGATGGTCCTCATGA 738
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
701 GATCCCCATTTCCAAAGCCCATGAGCACATTTTGAATGGTCCTTATGA 750

```

```

739 ACGACTGGAGCGCACGAGACATCCAGCAATGGGAGTACGTCCCACTTGGG 788
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
751 ACGACTGGAGTGCACGAGACATTCAGAAGTGGGAGTATGTCCCTCTCGGG 800
    .      .      .      .      .      .      .      .
789 CCATTCTCTGGGAAAAGCTTTGGAACCACAATCTCCCGTGNGTGGTGCC 838
    ||||| ||||| ||||| ||||| ||||| ||||| |||||:|||||
801 CCATTCTCTGGGAAGAGTTTGGGACCACTGTCTCTCCGTGGGTGGTGCC 850
    .      .      .      .      .      .      .      .
839 TATGGATGCCCTCATGCCCTTTGTGGTGCCAAACCCAAAGCAGGACCCCA 888
    ||||| ||||| ||||| ||||| ||||| ||||| |||||
851 CATGGATGCTCTCATGCCCTTTGCTGTGCCCAACCCGAAGCAGGACCCCA 900
    .      .      .      .      .      .      .      .
889 AGCCCTTGCCATATCTCTGCCACAGCCAGCCCTACACATTTGATATCAAC 938
    |||| |||| |||| |||| | ||||| ||||| |||||
901 GGCCCTGCCGTATCTGTGCCATGACGAGCCCTACACATTTGACATCAAC 950
    .      .      .      .      .      .      .      .
939 CTGTCTGTCTCTTTGAAAGGAGAAGGAANGAGCCAGGCGNCTACCATCTG 988
    || ||||| ||||| ||||| ||||| ||||| |||||:||||| ||
951 CTCTCTGTTAACCTGAAAGGAGAAGGAATGAGCCAGGCGGCTACCATATG 1000
    .      .      .      .      .      .      .      .
989 CAGGTCTAACTTTAAGCACATGTACTGGACCATGCTGCAGCAACTCACAC 1038
    || ||| |||| ||||| ||||| ||||| ||||| ||||| |
1001 CAAGTCCAATTTTAAGTACATGTACTGGACGATGCTGCAGCAGCTACTC 1050
    .      .      .      .      .      .      .      .
1039 ACCACTCTGTAAATGGATGCAACCTGAGACCTGGGGACCTCTTGGCTTCT 1088
    ||||| |||| || ||||| |||| | ||||| ||||| |||||
1051 ACCACTCTGTCAACGGCTGCAACCTGCGGCCGGGGACCTCCTGGCTTCT 1100
    .      .      .      .      .      .      .      .
1089 GGAACCATCAGTGGATCAGACCCTGAAAGCTTTGGCTCCATGCTGGAAC 1138
    || ||||| || || || || |||| |||| ||||| |||||
1101 GGGACCATCAGCGGGCCGGAGCCAGAAAACCTCGGCTCCATGTTGGAAC 1150
    .      .      .      .      .      .      .      .
1139 GTCCTGGAAGGNAACAAAGGCCATCGATGTGGGGCAGGGGCAGACCAGGA 1188
    ||| |||||:||| ||| |||| || |||| | || |||||
1151 GTCGTGGAAGGGAACGAAGCCCATAGACCTGGGGAATGGTCAGACCAGGA 1200
    .      .      .      .      .      .      .      .
1189 CCTTCCTGCTGGACGGCGATGAAGTCATCATAACAGGTCACTGCCAGGGG 1238
    || ||||| ||||| ||||| ||||| ||||| |||||
1201 AGTTTCTGCTGGACGGGATGAAGTCATCATAACAGGTACTGCCAGGGG 1250
    .      .      .      .      .      .      .      .
1239 GACGGCTACCGTGTTGGCTTTGGCCAGTGTGCTGGGAAAGTGCTGCCTGC 1288
    || || |||| | ||||| ||||| ||||| ||||| |||||
1251 GATGGTTACCGCATCGGCTTTGGCCAGTGTGCTGGAAAAGTGCTGCCTGC 1300
    .      .      .      .      .      .      .      .
1289 CCTTTCACCAGCCTGA.....AGCTCCGGAAGTCACAAGACACA 1327
    || ||| |||| |||| |||| |||| |||| ||||
1301 TCTCCTGCCATCATGAGATTTTCTCTGCTCTTCTGGAAACAAAGGGCTCA 1350
    .      .      .      .      .      .      .      .
1328 CCCTTGCCCTTATGAGGATCATGCT..ACCACTGCATCAGTCAGGAATGAA 1375
    | |||| | | || || || || || || ||
1351 AGCACCCCTTTCAACCTGTGACTGGGGTCCTCCCTCGGGCTGTAGGCCT 1400
    .      .      .      .      .      .      .      .
1376 TAAA.GCTACTTTGATTGKGGGAANTGCCTTTTTTTTTTTTTTTTT.... 1420
    |||| || | ||: ||| || | || |
1401 GGTCCGCTA.TTCAGTGACAAATAAGCCATTGTGCTCTGAGGCCTGCAAA 1450

```

Fig. 15 (continued)

Figure 16. Comparison of the amino-acid sequence encoded by the open-reading frame (ORF) of mouse cDNA RN273 (see Fig. 15) with the human FAH sequence reported by Phaneuf *et al.* (1991).

The mouse sequence is shown in the top line (designated by M before amino-acid residue #1), and the human sequence is shown in the bottom line (labelled H). The one-letter IUB codes for amino acids are used. The two sequences are ~88% identical and ~93% similar (% similarity takes into account paired amino-acids that are different, but have similar function; see Devereux *et al.*, 1984). Paired amino-acid residues that are identical in mouse and human are aligned by vertical lines; residues that are similar, but not identical (e.g., hydrophilic residues, such as Q and H at position 15) are aligned by colons; residues with lesser degrees of similarity (e.g., hydrophobic and ambivalent residues, such as V and A at position 54) are aligned by single dots; and residues with no appreciable similarity (e.g., charged and ambivalent residues, such as R and C at position 105) are separated by an open space. A "?" indicates a mouse amino-acid residue that has not yet been identified conclusively due to nucleotide ambiguity. Many of the ?'s demarcate positions where N's were inserted in the mouse nucleotide sequence to prevent frameshifts in the mouse amino-acid sequence with respect to the human sequence.

```

M 1 MSFIQ?AEDSDFP IQNLPYGVFSTQSNPKPRIGVAIGDQILDLSVIKHLF 50
    |||. |||||:|||||.:|:|||||:|||||
H 1 MSFIPVAEDSDFP IHNLPYGVFSTRGDPRPRIGVAIGDQILDLSI KHLF 50
    .
51 TGPALSKHQHVFE TTLNNFMGLGQA AWKEARASLQNL LSASQARLRDDK 100
    |||. |||||. ||:|. |||. |||||:|||||. |||||. |||||.
51 TGPVLSKHQDVFNQPTLNSFMGLGQA AWKEARVFLQNL LSVSQARLRDDT 100
    .
101 ELRQRAFTSQASATMHLPATIGDYTD FYSSRQHRTNVGIMFRGKE?ALLP 150
    |||. ||. |||||:|||||:|||||:|||||:|||||:|||||:|||||
101 ELRKCAFISQASATMHLPATIGDYTD FYSSRQHATNVGIMFRDKENALMP 150
    .
151 NWLHLPVGYHGRASS?VVS GTP IRRPM??MRPDNSKPPVYGACRL LDME 200
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 NWLHLPVGYHGRASSVVVS GTP IRRPMGQMKPDDSKPPVYGACKLLDME 200
    .
201 EMAFFVGPGNRFGEPIPI SKAHEHIFGMVLMNDWSARDIQQWEYVPLGPF 250
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201 EMAFFVGPGNRLGEPIPI SKAHEHIFGMVLMNDWSARDIQKWEYVPLGPF 250
    .
251 LGKSFGTTISP?VVPMDALMPFVVPNPKQDPKPLPYLCHSQPYTFDINLS 300
    |||||:|| |||||:|||||:|||||:|||||:|||||:|||||
251 LGKSFGTTVSPWVVPMDALMPFAVPNPKQDPRPLPYLCHDEPYTFDINLS 300
    .
301 VSLKGEG?SQA?TICRSNFKHMYWTMLQQLTHHSVNGCNLRPGDLLASGT 350
    |. ||||| ||| |||:||||. |||||:|||||:|||||:|||||:|||||
301 VNLKGEGMSQAATICKSNFKYMYWTMLQQLTHHSVNGCNLRPGDLLASGT 350
    .
351 ISGSDPESFGSMLELSWK?TKAIDVGQGQTRTFLLDGDEVIITGHCQGDG 400
    |||. |:|. ||||| ||| |:|:|. ||||. |||||:|||||. |||||
351 ISGPEPENFGSMLELSWKGT KPIDLNGQTRKFLLDGDEVIITGYCQGDG 400
    .
401 YRVGFGQCAGKVLPA LSPA* 420
    ||:||||| |||. |. |
401 YRIGFGQCAGKVLPA LLPS* 420

```

Figure 17. Possible identification of the A5 breakpoint with a probe from the *D7Rn2* locus.

A Southern blot of (C3H₇/Rl x 101/Rl)F₁ (wild-type control indicated by H(+/+)) in the Fig.) and A5/A5 DNAs digested with the enzymes indicated. The digested DNAs were electrophoresed through a 0.4% agarose gel, blotted to nylon, and hybridized with the RN216 probe from the *D7Rn2* locus, the results of which are shown here. *Sma*I and *Cla*I restriction-fragment sizes (in kb) are indicated to the right of the figures. The *Kpn*I fragment is 19 kb in size. * denotes the size-altered *Sma*I fragment observed in A5/A5 DNA.

KpnI XhoI SmaI ClaI

H(+/+)

A5/A5

H(+/+)

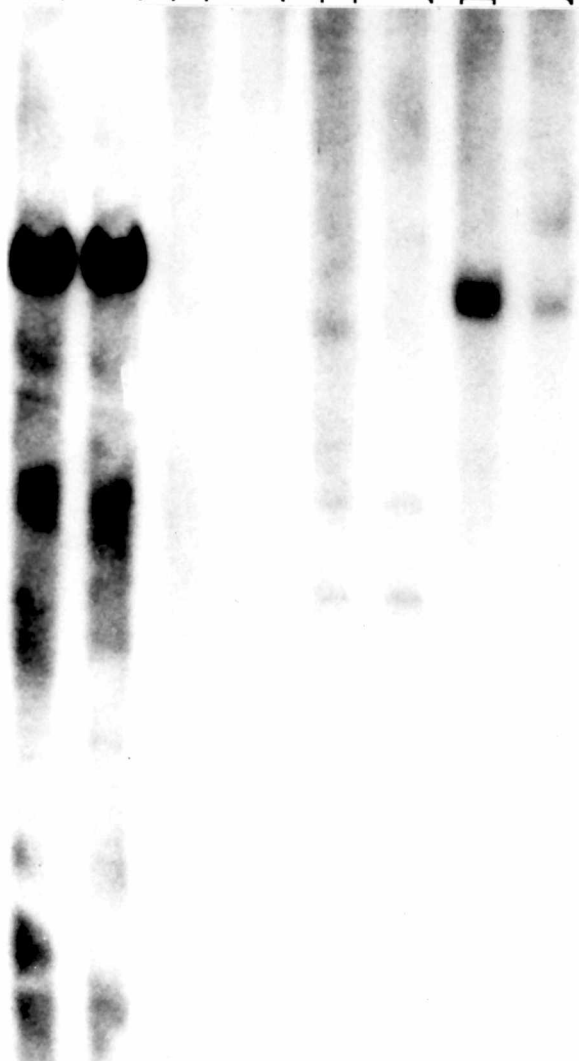
A5/A5

H(+/+)

A5/A5

H(+/+)

A5/A5



— 22.5 (Cla)
— 20.9 (Sma)*
— 17.6 (Cla)
— 16.5 (Sma)

— 11.2 (Sma)

— 9.5 (Sma)

Figure 18. The *D7Rn4* locus maps proximal to the proximal breakpoints of the Di4 and Fp1 deletions.

A Southern blot of *Eco*RI-digested DNAs of the genotypes indicated hybridized with the RN254 probe from the *D7Rn4* locus. Spt and *c^{ch}* designate *Mus. spretus* and *Mus. musculus* chromosomes, respectively. The *Eco*RI fragments specific to the Spt and *c^{ch}* chromosomes are indicated to the right of the figure (sizes in kb). The absence of the *c^{ch}*-specific fragment indicates that the *D7Rn4* locus is removed by the deletion.

c^{ch} / Di4

Spt / Di4

c^{ch} / Dp1

Spt / Dp1

c^{ch} / Fp1

Spt / Fp1

c^{ch} / Fq1

Spt / Fq1

Spt / Spt

-8.8 (Spt)

-3.7 (*c^{ch}*)

Figure 19. The physical/functional map as refined by deletion-mapping of the *D7Rn4* and *D7Rn3* loci.

Previous versions of the map have been described in the legends to Figs. 10 and 11. The *D7Rn4* locus maps proximal to (and within 46 kb of) the breakpoints of both the Di4 and Fp1 deletions. The coding sequences of the *Fah* gene are completely contained within the DNA segment (no greater than 62 kb in length) located between the proximal breakpoint of the A5 deletion and the most distal of the proximal breakpoints of the Di4 and Fp1 deletions. The *c* locus is positioned as defined by other long-range mapping analyses (G. Kelsey, pers. comm.). The positions shown for the *D7Rn3* locus and the distal breakpoints of the *c*^{16DFiOD}, A2, and Ai1 deletions are merely to indicate the results of deletion-mapping of the locus, and do not indicate exact physical locations.

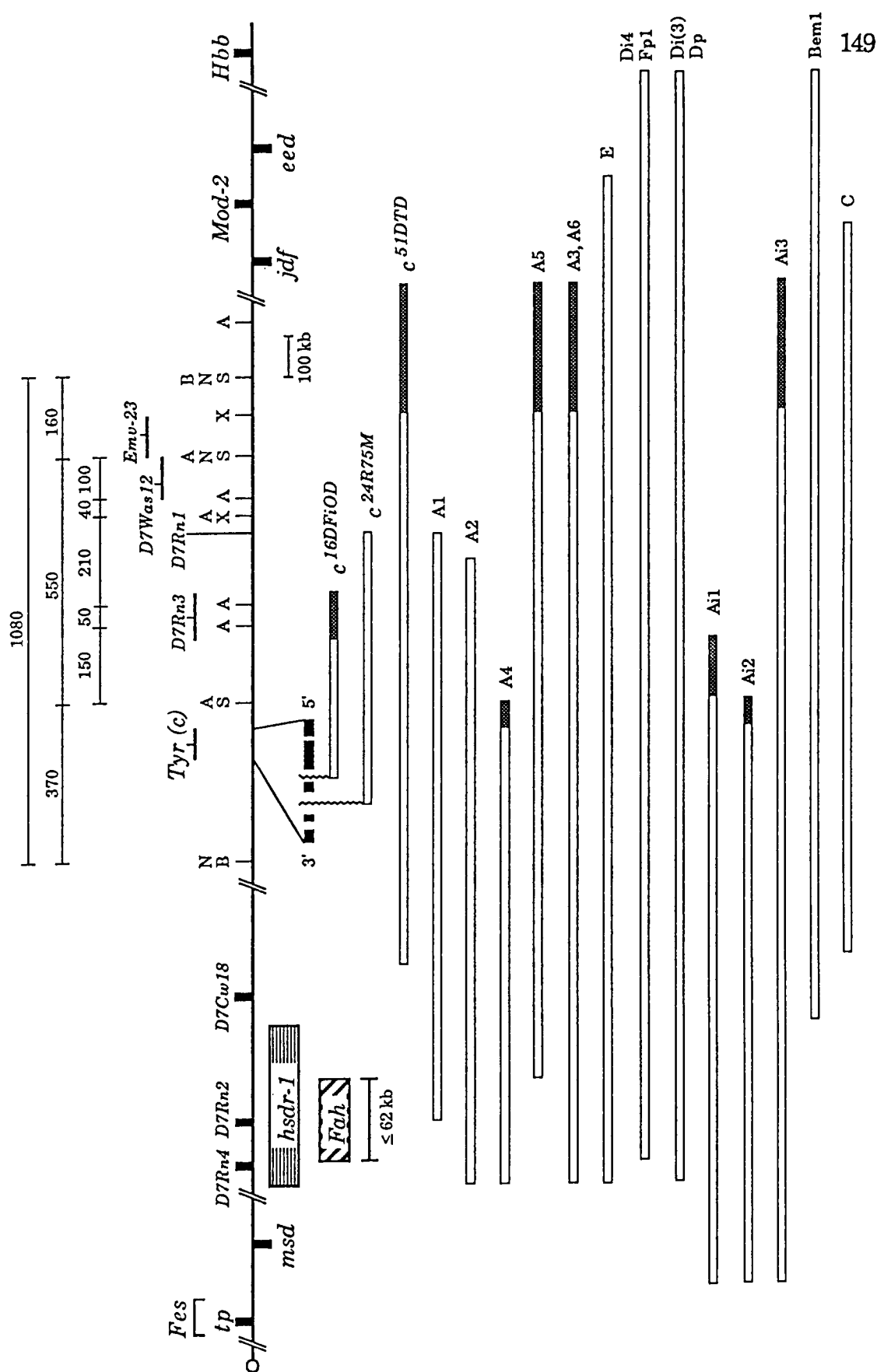


Figure 20. The *D7Rn3* locus maps distal to the *Ai1* and *c^{24R75M}* breakpoints, but proximal to the *A2* breakpoint.

A Southern blot of *Bam*HI-digested DNAs of the genotypes indicated hybridized with the RN202 probe from the *D7Rn3* locus. The probe hybridizes to a wild-type, 14.8-kb *Bam*HI fragment. The blot was washed after hybridization at a stringent temperature of 55°C, rather than the usual 65°C.

C3H_f/RI (+/+)

101/RI (+/+)

A1/A1

A2/A2

A3/A3

A4/A4

A5/A5

A6/A6

C/C

E/E

E/A!1

E/A!2

E/A!3

c^{24R75M}/*c*^{24R75M}

c^{16DF!OD}/*c*^{16DF!OD}

2A (*c*^{ch}/*c*^{ch})

VITA

Mitchell Louis Klebig was born in Winona, Minnesota on April 17, 1961, and was raised in Fountain City, Wisconsin. He graduated from Luther High School in Onalaska, Wisconsin, as salutatorian of his class, in 1979. He attended Luther College in Decorah, Iowa, was elected to Phi Beta Kappa, and received a B. A. degree (Magna Cum Laude) in Biology in May of 1983. He entered the University of Tennessee Graduate School of Biomedical Sciences at Oak Ridge in the fall of 1985, and was awarded the Doctor of Philosophy degree in August of 1991. He will begin postdoctoral study, under the direction of Dr. R. Woychik, at the Biology Division of Oak Ridge National Laboratory in June of 1991, and will continue to pursue a career in Mammalian Genetics research and instruction.