Genetic Manipulation of the Early Mammalian Embryo

Edited by

FRANK COSTANTINI

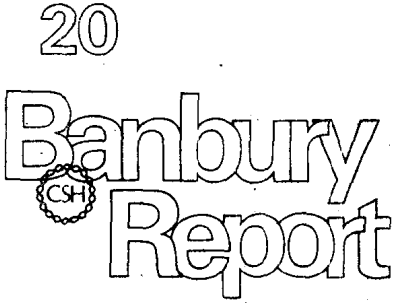
RUDOLF JAENISCH

Genetic Manipulation of the Early Mammalian Embryo

Edited by

FRANK COSTANTINI
Department of Human Genetics and Development
College of Physicians and Surgeons
Columbia University

RUDOLF JAENISCH
Whitehead Institute and Department of Biology
Massachusetts Institute of Technology and
Heinrich-Pette-Institut
Hamburg, Federal Republic of Germany



COLD SPRING HARBOR LABORATORY
1985

BANBURY REPORT SERIES

Banbury Report 1: Assessing Chemical Mutagens
Banbury Report 2: Mammalian Cell Mutagenesis
Banbury Report 3: A Safe Cigarette?
Banbury Report 4: Cancer Incidence in Defined Populations
Banbury Report 5: Ethylene Dichloride: A Potential Health Risk?
Banbury Report 6: Product Labeling and Health Risks
Banbury Report 7: Gastrointestinal Cancer: Endogenous Factors
Banbury Report 8: Hormones and Breast Cancer
Banbury Report 9: Quantification of Occupational Cancer
Banbury Report 10: Patenting of Life Forms
Banbury Report 11: Environmental Factors in Human Growth and Development
Banbury Report 12: Nitrosamines and Human Cancer
Banbury Report 13: Indicators of Genotoxic Exposure
Banbury Report 14: Recombinant DNA Applications to Human Disease
Banbury Report 15: Biological Aspects of Alzheimer's Disease
Banbury Report 16: Genetic Variability in Responses to Chemical Exposure
Banbury Report 17: Coffee and Health
Banbury Report 18: Biological Mechanisms of Dioxin Action
Banbury Report 19: Risk Quantitation and Regulatory Policy

Banbury Report 20: Genetic Manipulation of the Early Mammalian Embryo

Printed in the United States of America
Cover and book design by Emily Harste

Library of Congress Cataloging in Publication Data
Main entry under title:

Genetic manipulation of the early mammalian embryo.

(Banbury report, ISSN 0198-0068 ; 20)
Includes bibliographies and indexes.
1. Mammals–Genetics. 2. Genetic engineering.
3. Developmental genetics. 4. Embryology–Mammals.
I. Costantini, Frank, 1952- II. Jaenisch, Rudolf.
III. Series.
QL738.5.G44 1985 599'.015 85-13233
ISBN 0-87969-220-0

All Cold Spring Harbor Laboratory publications may be ordered directly from Cold Spring Harbor Laboratory, Box 100, Cold Spring Harbor, New York 11724. (Phone: 1-800-843-4388) In New York State (516) 367-8425

Participants

David Baltimore, Whitehead Institute for Biomedical Research, Massachusetts Institute of Technology

Ralph L. Brinster, School of Veterinary Medicine, University of Pennsylvania

C. Thomas Caskey, Departments of Medicine and Cell Biology, Baylor College of Medicine

Verne M. Chapman, Department of Molecular Biology, Roswell Park Memorial Institute

Neal G. Copeland, Department of Microbiology and Molecular Genetics, University of Cincinnati College of Medicine

Frank Costantini, Department of Human Genetics and Development, College of Physicians and Surgeons, Columbia University

Eric H. Davidson, Division of Biology, California Institute of Technology

Martin Evans, Department of Genetics, University of Cambridge, Cambridge, England

Ronald M. Evans, Salk Institute for Biological Studies

Jeanette S. Felix, Center for Research on Mothers and Children, National Institute of Child Health and Human Development

Kirsten Fischer-Lindahl, Basel Institute for Immunology, Basel, Switzerland

Rudolf Grosschedl, Whitehead Institute for Biomedical Research, and Department of Biology, Massachusetts Institute of Technology

Douglas Hanahan, Cold Spring Harbor Laboratory

Rudolf Jaenisch, Whitehead Institute for Biomedical Research, and Department of Biology, Massachusetts Institute of Technology

Nancy A. Jenkins, Department of Microbiology and Molecular Genetics, University of Cincinnati College of Medicine

Duane Kraemer, Department of Veterinary Physiology and Pharmacology, Texas A&M University

Elizabeth Lacy, Sloan-Kettering Institute

Arnold J. Levine, Department of Molecular Biology, Princeton University

Thomas Maniatis, Department of Biochemistry and Molecular Biology, Harvard University

Frederick Naftolin, Department of Obstetrics and Gynecology, Yale University School of Medicine

John Newport, Department of Biology, University of California, San Diego

Richard D. Palmiter, Howard Hughes Medical Institute, University of Washington

Christopher Polge, ARC Institute of Animal Physiology, Animal Research Station, Cambridge, England

Peter W. J. Rigby, Department of Biochemistry, Imperial College of Science and Technology, London, England

Frank H. Ruddle, Department of Biology, Yale University

Joseph D. Schulman, The Genetics and IVF Institute

Lee M. Silver, Department of Molecular Biology, Princeton University

Davor Solter, The Wistar Institute

Timothy Stewart, Department of Genetics, Harvard Medical School

Dan T. Stinchcomb, Department of Cellular and Developmental Biology, Harvard University

Ursula Storb, Department of Microbiology and Immunology, University of Washington

Azim Surani, AFRC Institute of Animal Physiology, Cambridge, England

Shirley M. Tilghman, Institute for Cancer Research, Philadelphia

Erwin F. Wagner, European Molecular Biology Laboratory, Heidelberg, Federal Republic of Germany

Keith Willison, Chester Beatty Laboratories, Institute for Cancer Research, London, England

Preface

Some new fields may begin to move so rapidly that it becomes important to draw the various strands together to gain an overall perspective on their current status and likely future directions. Perhaps no field exemplifies this more than that comprising recent developments in the experimental manipulation of the mammalian genome. Techniques for introducing cloned genes directly into the germline, introducing genetically altered cells into early embryos, or employing developmental mutations for identifying, isolating, and then cloning the specific genes affected are collectively contributing toward an understanding of mammalian gene regulation and genetic control of development at a level not previously approachable experimentally. The practical application of these techniques, especially in farm and ranch animal husbandry, is also being developed at several centers throughout the world and they have gained considerable popular publicity.

However, it is not sheep or cows, but the mouse which remains the undisputed mammal of choice for research in this area. It is thus not surprising that the October 1984 Banbury conference from which this publication emanates was dominated by experimental work being done in that animal. Several lines of research now involve the introduction and integration into mice of cloned genes, which are retained not only through embryogenesis, but also into later generations via normal germline transmission. In some cases these genes have also been appropriately expressed and regulated in recipient animals. Experimental approaches of this sort are beginning to probe key questions concerning the site specificity of gene integration and the identification and characterization of regulatory elements controlling gene expression during development.

These collected papers of the fall 1984 Banbury conference on the genetic manipulation of the mammalian ovum and early embryo should prove of value both in gauging present capabilities as well as for gaining a perspective on the likely near future progress of this rapidly moving new field.

I am pleased to take this opportunity to acknowledge the support received for this project from the March of Dimes Birth Defects Foundation, the Fogarty International Center, and the National Institute of Child Health and Human Development. It is also once again a great pleasure to express my appreciation to **Bea Toliver**, the Banbury Center administrative assistant, for her usual, highly competent and personable attention to the organization of this conference and to **Judith Blum**, the Banbury editor, for her continued standards of excellence and assiduous attention to all aspects of production of this volume.

Michael Shodell
Director
Banbury Center

Contents

Session 1: Developmental Genetics

Molecular Studies of the Structure and Evolution of Mouse t Haplotypes

LEE M. SILVER
Department of Molecular Biology
Princeton University
Princeton, New Jersey 08544

OVERVIEW

A particular region of mouse chromosome 17 occurs in a variant form known as a *t* haplotype which is present at a high frequency in wild populations. The *t* haplotype can be considered as a genomic parasite since it is able to maintain its own integrity and it functions to ensure its own survival without providing any obvious benefit to the animal in which it resides. A series of molecular probes have been obtained that distinguish *t* haplotype genomic sequences from wild-type counterparts. These probes have been used to clarify the relationships of *t* haplotypes to each other and to the normal form of the chromosome. Additionally, these probes allow a genetic dissection and characterization of the crucial *t*-specific property of transmission ratio distortion. Finally, the accumulated data allow speculation into the origin and evolution of these unusual chromosomal variants.

INTRODUCTION

Mouse *t* haplotypes have been studied intensely by many different investigators over the 50 years since they were first discovered as naturally occurring genetic variants (Gluecksohn-Waelsch and Erickson 1970; Bennett 1975; Sherman and Wudl 1977; Silver 1981). However, it is only within the last several years that we have begun to decipher the structure of *t* haplotypes and the rationale for their existence. In fact, much of the older literature on this topic is rather misleading. What workers now call *t* haplotypes were originally known as *t* alleles or *t* mutations that were thought to be alleles of each other and dominant *T* mutations at the *T* locus. The so-called *t* mutations appeared to express pleiotropic effects on tail length, fertility, embryogenesis, male transmission ratio, and meiotic recombination.

We now realize that the simple locus model of *t* haplotypes is incorrect. There is a well-defined *T* locus near the centromere on chromosome 17, and spontaneous and induced mutations at this locus cause a shortening of tail length in heterozygotes and are lethal in homozygous embryos. However, the *t* haplotypes are not single locus mutations, and they have never been observed to occur de novo in the laboratory. *t* Haplotypes are only derived from wild mice and are recognized by their interaction with *T* mutations to produce tailless *T/t* animals. A large body of evidence (to be discussed below) indicates that the *t* haplotypes are structurally variant forms of a major portion of mouse chromosome 17. Within this variant chromosomal region are many normally functioning genes interspersed with a num-

ber of independent "mutant loci" that mediate the *t* effects on sperm differentiation, embryogenesis, and tail length. The *t* haplotype maintains itself as a well-defined genomic entity by suppression of recombination along its length, and it propagates itself through mouse populations by means of a male-specific transmission ratio distortion in its favor. There is no convincing evidence that the individual *t*-associated "mutant genes" bear any structural or functional homology to each other. Therefore, a *t* haplotype cannot be considered to be a gene family like the H-2 complex, but rather, a *t* haplotype must be thought of as a structural entity.

A complete *t* haplotype is operationally defined as the common *t* form that exists in wild populations of *mus domesticus* and *mus musculus*. Complete *t* haplotypes extend from a point between the centromere and *T* to a point between the H-2 complex and *Pgk-2* (Fig. 1). Recombination is suppressed along the length of *t* DNA in +/*t* heterozygotes, but normal recombination will occur along regions of *t* DNA overlap in animals heterozygous for two *t* haplotypes in *trans* configuration (Silver and Artzt 1981). The t complex is defined as the genomic region over which recombination is suppressed in heterozygotes for complete *t* haplotypes. It is important to realize that the t complex would not exist apart from its *t* haplotype-dependent definition.

The *t* haplotypes provide a system that can be used to advantage for investigations into a variety of biological fields, including chromosome evolution, meiotic recombination, population dynamics, germ cell differentiation, early embryonic development, and immunogenetics. In this report, we describe studies that have combined the tools of classical genetics with molecular probes in order to understand the structure of *t* haplotypes. The accumulated data allow speculation into the origin and evolution of these unusual genomic parasites.

RESULTS AND DISCUSSION

Over the last 5 years, a variety of molecular probes for the t complex have been developed in order to investigate the relationships among independently-derived *t*

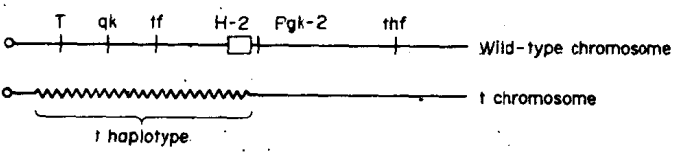


Figure 1
Schematic representation of a chromosome 17 genotype of an animal heterozygous for a complete *t* haplotype. The top line represents a wild-type form of chromosome 17 with well-characterized genetic marker loci. (*T*) Brachyury; (*qk*) quaking; (*H-2*) the *H-2* complex; (*Pgk-2*) Phophoglycerate kinase-2; and (*thf*) thin fur. The zigzag region of the bottom line represents the region of *t* DNA present in a complete *t* haplotype. A chromosome carrying a *t* haplotype is referred to as a *t* chromosome. The centromeres are on the left and the telomeres are on the *right*.

haplotypes and between *t* haplotypes and wild-type chromo omes. One set of probes is based upon the technique of high resolution two-dimensional gel electrophoresis. This technique has been used to identify a series of eight testicular cell polypeptides uniquely expressed in animals that carry a complete *t* haplotype (Silver et al. 1983). A comparison of testicular cell protein patterns expressed by mice carrying different *t* complex genotypes has allowed the identification of apparent allelic wild-type forms of five of these t complex proteins (TCP-1, TCP-3, TCP-4, TCP-7, and TCP-8). Cell-free translation experiments have demonstrated that the difference between the wild-type and mutant forms of each of these proteins is encoded within the mRNA sequences; this provides evidence for the location of the structural genes (e.g., *Tcp-1, Tcp-3*) for all five proteins within the t complex.

A second independent set of t complex probes was obtained through the direct cloning of random genomic sequences for microdissected fragments of the proximal portion of chromosome 17 (Roehme et al. 1984). These clones have been used to identify restriction fragment length polymorphisms (RFLPs) that are characteristic of *t* haplotypes. To date, we have identified 12 independent clones of this type that detect over 20 different regions across the length of *t* haplotype DNA (Fox et al. 1985; H. Fox et al., unpubl.). Other molecular probes for the t complex that are currently available include DNA clones of alpha globin psuedogene-4 (*Hba-ps4*) (Fox et al. 1984), the alpha crystallin I structural gene (H. Fox, unpubl.), and all of the genes located within the major histocompatibility (H-2) complex, which is an integral component of *t* haplotypes (Shin et al. 1982; Silver 1982).

A dramatic increase in our understanding of the biology of *t* haplotypes has emerged from analyses of *t* haplotypes and wild-type forms of chromosome 17 with each of the molecular probes available. One important conclusion derived from the accumulated data is that all *t* haplotypes originated from a single founder chromosome. In fact, the independent *t* haplotypes are nearly indistinguishable from each other at all coding and noncoding sequences outside of the H-2 complex (Silver et al. 1983; Fox et al. 1985). The significance of this result is emphasized by the observation that many of the DNA probes do detect polymorphisms among different wild-type chromosomes. These data suggest that the common ancestor of all currently existing *t* haplotypes occurred recently within an evolutionary timeframe.

A second conclusion derived from the accumulated data is that *t* haplotypes and wild-type chromosomes are not very different from each other at the level of primary DNA sequence (H. Fox and L.M. Silver, unpubl.). Rather, it would appear that the major structural differences between *t* and wild-type must occur at the level of gross chromosomal rearrangements. In fact, the available data provide evidence for the presence of two major non-overlapping inversions that distinguish *t* haplotypes from wild-type (Artzt et al. 1982; N. Sarvetnick et al., unpubl.). These gross rearrangements may be sufficient to explain the suppression of recombination observed between *t* and wild-type.

Although recombination is generally suppressed between *t* and wild-type, rare crossing over events do occur, and a number of the recombinant products resulting from such events have been recovered in the laboratory. These chromosomes are called partial *t* haplotypes since they retain only a portion of the original *t* haplotype DNA and express only a subset of the complete *t* haplotype properties. Partial *t* haplotypes have been used to great advantage in genetic experiments aimed at understanding the *t*-specific property of transmission ratio distortion.

Males heterozygous for a complete *t* haplotype and a wild-type chromosome can transmit their *t*-bearing chromosome to greater than 99% of their offspring. This transmission ratio distortion is crucial for the survival of *t* haplotypes in wild mouse populations. Only complete *t* haplotypes are transmitted at a high ratio.

Although mice that carry a single partial *t* haplotype cannot transmit it at a high ratio, distortion in favor of a *t* chromosome can be reconstituted in males that carry particular pairs of partial haplotypes in *cis* or *trans* configuration (Lyon 1984). With certain *trans* combinations, either the proximal, central, or distal partial *t* haplotype is transmitted at a high ratio; however, other haplotype combinations lead to near-equal ratios. Lyon and her colleagues have interpreted these data in terms of a model in which different partial *t* haplotypes carry different lengths of *t* DNA containing particular sets of "distortion loci." The basic tenet of this model is that transmission ratio distortion results from the action of a series of *t*-specific distortor loci (i.e., *Tcd-1, Tcd-2*) upon a single *t*-specific responder locus (*Tcr*). The effects of the distortor loci appear to be additive, and they can act in *cis* or *trans* to the *Tcr* locus which must be heterozygous (Fig. 2).

Molecular analysis of each of the partial *t* haplotypes used in these studies has provided strong evidence in support of the Lyon model for transmission ratio distortion (Fox et al. 1985). With additional breeding studies, it has become clear that there are at least four independent *Tcd* loci that are each correlated with a specific molecular marker (Fig. 2). The *Tcd-1* locus maps most proximally on *t* haplotypes with the T48 restriction fragment marker (Fox et al. 1985); the *Tcd-4* locus maps next with the *Tcp-1* structural gene for the p63/6.9 protein (L.M. Silver, unpubl.); the *Tcr* responder locus maps in the middle of the haplotype and is associated with the T66B restriction fragment (Fox et al. 1985); the *Tcd-3* locus maps next with several restriction fragment markers including T66C and T122; finally, the *Tcd-2* locus maps most distally with many molecular markers including T108 as well as the H-2 complex. A *t* haplotype must carry all of these loci in order to be transmitted at a high ratio. Since these loci are spread out along the entire complete *t* haplotype, it is clear that any intra-haplotype recombination event will produce recombinant chromosomes that can no longer express this property.

Although males heterozygous for a single complete *t* haplotype express a high transmission ratio, males that carry two complete *t* haplotypes are invariably and unconditionally sterile. It appears likely that this homozygous sterility is a recessive consequence of some or all of the same genes that are involved in heterozygous distortion (M.F. Lyon, pers. commun.).

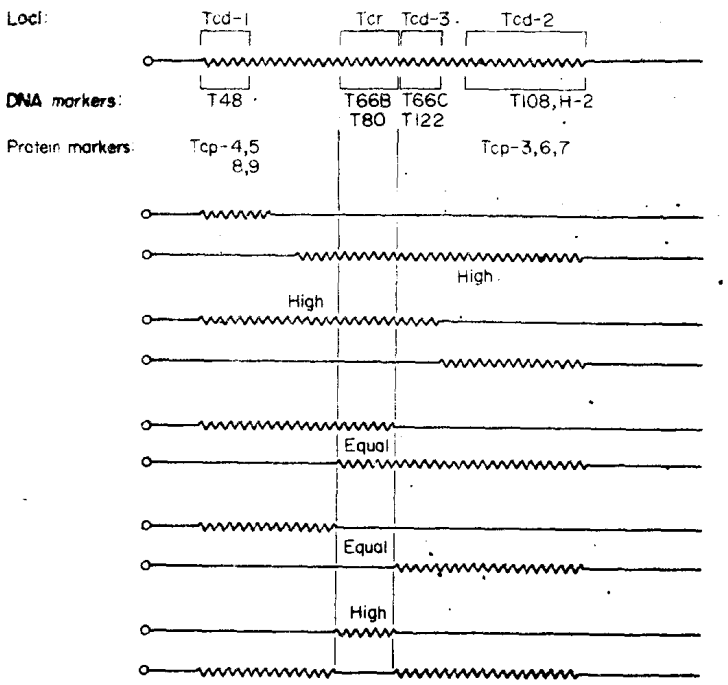


Figure 2
Genetic basis for transmission ratio distortion. The positions of three of the distorter loci (*Tcd-1*, *Tcd-2*, and *Tcd-3*) and the responder locus (*Tcr*) is shown at the top of the figure with associated molecular markers beneath the top chromosome representation. The *Tcd-4* locus was recently identified in association with *Tcp-1* and both are located between *Tcd-1* and *Tcr*. Four examples of genotypes with *trans* combinations of partial *t* haplotypes are shown. The chromosome transmitted at a high ratio is indicated in each case (Lyon 1984).

All complete *t* haplotypes express the same phenotypes of transmission ratio distortion, sterility, and recombination suppression. Some, but not all, express an additional phenotype of recessive embryonic lethality. To date, 15 complementing, recessive, lethal mutations have been found in association with *t* haplotypes from wild mice (Klein et al. 1984). The work of Artzt et al. (1982) clearly demonstrated that different, complementing *t* lethal mutations were nonallelic to each other. Furthermore, these studies and others (Condamine et al. 1984) show that with one exception (*tw32*), the individual lethal effects can be genetically mapped to single loci. These lethal loci are independent of the various genes involved in the other *t* haplotype properties.

With the recent advances described in this report, it has become possible to speculate on a model for the origin and evolution of *t* haplotypes that can account for nearly all of the seemingly unrelated properties characteristic of these variant genomic entities. This model is based upon the assumption that the proximal region

of all forms of mouse chromosome 17 (the so-called t complex) carries genes that play important roles in haploid germ cell differentiation or sperm maturation. This assumption is supported by the observation of a large number of *t*-encoded proteins found to be expressed in the testes but not in other tissues (L.M. Silver and M. Krangel unpubl.). In addition, several non-*t* haplotype-associated genes that affect male fertility have also been localized to the t complex. Most importantly, only chromosome 17, of all the autosomes, has mutated to a form that is capable of thwarting the machinery of germ cell differentiation to its advantage.

The first event in the evolution of *t* haplotypes would be the chance accumulation on a single chromosome of a set of alleles (haplotype) at the testes-expressing loci that acted synergistically to raise their own transmission frequency to a level greater than 50%. Once such a haplotype came together, selective forces would act at the testes-expressing (distorter) loci to continue to increase the transmission ratio to even higher levels (Lewontin and Dunn 1960). However, the transmission ratio could never effectively increase above 85% since recombination events would cause 15% of all offspring to lose at least one gene required for distortion. Therefore, at this stage of evolution, chromosomal rearrangements would be selected in order to lock the various distorter loci together by recombination suppression. With continued selection for extremely high transmission ratios (of 99% or greater), *t* haplotypes would increase their frequency to significant levels in mouse populations, and males homozygous for *t* chromosomes would begin to appear. Such males are sterile (probably as a consequence of homozygousity for *t* distorter alleles) and families dominated by sterile males will not reproduce. In other words, the successfulness of these haplotypes will result in their extinction at the family level in populations. At this stage of evolution, selection will act to eliminate self-extinction in favor of *t* haplotypes which carry spontaneous recessive embryonic lethal mutations. If a *t* haplotype is lethal, all *t*-carrying animals will be heterozygous and fertile. Interestingly, over 20 years ago, Lewontin (1962) demonstrated with computer modeling that lethal *t* haplotypes were at a selective advantage relative to sterile *t* haplotypes.

The rationale for the presence of lethal mutations on *t* haplotypes has not been previously explained. Only some of the naturally occurring *t* chromosomes carry lethal mutations, so the lethal phenotype is not an obligatory feature of this system. In fact, all *t* haplotypes are virtually identical except for the presence of distinguishing lethal mutations, and at least 15 different lethal t mutations are known to exist (Klein et al. 1984). Although t effects on sperm differentiation involve interactions among multiple loci, the t lethal effects appear to be the result of simple locus mutations (with one exception described above). All of these observations are consistent with the possibility that the lethal phenotypes were acquired spontaneously during the recent evolution of *t* haplotypes. Furthermore, it appears that the t system is still in a dynamic state and that the evolution of *t* haplotypes will continue in wild mice long after we are gone.