To the Graduate Council:

I am submitting herewith a dissertation written by Judith Harbron Moyer entitled "Molecular and Genetic Analysis of a New Insertional Mutation in the *TgN(Imorpk)737Rpw* Line of Mice that Causes Polycystic Kidney Disease". I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Biomedical Sciences.

Richard P. Woychik, Ph.D., Major Professor

We have read this dissertation
and recommend its acceptance:

Virginia L. Godfrey, D.V.M., Ph.D.

K. Bruce Jacobson, Ph.D.

Richard J. Mural, Ph.D.

Liane B. Russell, Ph.D.

J. Erby Wilkinson, D.V.M., Ph.D.

Accepted for the Council:

Associate Vice Chancellor
and Dean of the Graduate School

MOLECULAR AND GENETIC ANALYSIS OF A NEW INSERTIONAL MUTATION IN THE *TgN(Imorpk)737Rpw* LINE OF MICE THAT CAUSES POLYCYSTIC KIDNEY DISEASE

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

Judith Harbron Moyer
May 1994

DEDICATION

This dissertation is dedicated to my husband,

Dr. Bruce A. Moyer.

To weather a spouse's
pursuit of a doctoral degree and the inevitable accompanying
changes in the relationship is, to say the least, a challenge.
To do so with love, humor and patience is to demonstrate
the attainment of a separate kind of higher learning.

ACKNOWLEDGMENTS

First, I would like to thank my parents for the emphasis they placed on education, for as long as I can remember. They managed deftly to convey the message, without any attendant "shoulds" and "musts", that a college education would be provided to me, with the result that I accepted the notion without rebelliousness as a given. Putting me through my first four years of college was a financial struggle for them, and I have a deep appreciation of their efforts to provide an education for me, in addition to their constant encouragement and advice to never settle for "second best". The latter attitude has served me well in all aspects of my life.

I would like to thank my major professor, Dr. Rick Woychik, for his guidance and supervision over a four-year period. To have been a beneficiary of Rick's scientific abilities was a wonderful opportunity, and I am grateful for having had the chance to work in his lab on the Tg737 project. I am also grateful to the other committee members - Drs. Erby Wilkinson, Virginia Godfrey, Bruce Jacobson, Richard Mural and Lee Russell - for valuable assistance, constructive criticism and insightful discussions over the past three years. I also heartily thank Dr. Gene Rinchik for his interest in and active contribution to my academic progress, and for serving as a valued member of my dissertation committee until he relocated to New York several months ago.

I would like to gratefully acknowledge the following people for their indispensable contributions to this research project: B. Beatty, D. Andreadis, G. Barker, A. Ch'ang, W. Lee and W. McKinney for expert technical assistance; P. D'Eustachio for the mouse x Chinese hamster somatic cell hybrid DNAs; J.H. Nadeau for the *Tcrα* cDNA; B.L. Hogan and M.E. Dickinson for the *Rb-1* interspecific backcross DNAs and panels; L. Stubbs for the 14.5 dpc cDNA library; J. Bernstein and J. Rutledge for slides of human tissues; S.T. Reeders, L.F. Onuchic and J.J. Schrick for their contribution in the

computer analysis of the cDNA sequence; E. Michaud and B.L. Hogan for the 8.5 dpc cDNA library; J.J. Schrick for the cosmid and human genomic libraries and assistance with PFGE; M.L. Klebig for the genomic library from which the 5' genomic clones were isolated; E.D. Avner and W.E. Sweeney for the *bpk* mice; B. Yoder for probing a Northern blot with RW117/126 at my request; P. Galloway, M. Hauser, W. Lee and J. Cook for the ES cells, mouse fibroblasts, MDCK, LLC-PK_1, mouse keratinocyte and OK cell lines utilized in expression studies; P. Selby for providing C3H mice; V. Godfrey for maintaining and breeding the mice used in the *bpk-TgN737Rpw* complementation assays; and R. Isfort for thoughtful discussions about the possible role of the protein.

J. H. Moyer was supported by National Institutes of Health predoctoral National Research Service Award NR06789-01 from the National Center for Nursing Research.

ABSTRACT

A line of transgenic mice has been generated that contains an insertional mutation causing a phenotype similar to human autosomal recessive polycystic kidney disease. *TgN737Rpw* homozygotes display a complex phenotype that includes bilateral polycystic kidneys and an unusual liver lesion.

The mutant locus was cloned and characterized through use of the transgene as a molecular marker. A deletion of no more than 2.77 kb, and the insertion of 21.0 kb of genomic mouse DNA that stably cointegrated with the transgene, occurred at the integration site. An extended physical analysis of the mutant locus, which included pulsed field gel electrophoresis (PFGE), detected no additional rearrangements beyond the insertion site.

A candidate polycystic kidney disease (PKD) gene was identified whose structure and expression is directly associated with the mutant locus. A cDNA derived from this gene predicted a peptide containing multiple repeats of a degenerate 34-amino acid motif that defines members of the tetratricopeptide repeat (TPR) gene family, which are often involved in cell cycle control. The *TgN737Rpw* cDNA hybridizes to multiple transcripts during embryogenesis and in all wild-type adult tissues examined, and a cDNA isolated from embryonic day 14.5 may represent an alternatively-spliced form of the *TgN737Rpw* gene. Physical mapping data and the gene's expression pattern suggest that the mutant phenotype in *TgN737Rpw* mice is due to the disruption of a single, ubiquitously-expressed gene that functions during embryonic development and into adulthood.

The *TgN737Rpw* mutation is the first candidate gene accessed at the molecular level in any PKD model, and does not appear to be allelic with existing mouse models of PKD. The *TgN737Rpw* gene also represents the first *in vivo* TPR mutation in vertebrates. Moreover, the identification of the homologous human gene will

permit a thorough molecular study of whether this gene is directly associated with any of the various forms of PKD in humans.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADA	Adenosine deaminase [gene]
ADPKD	Autosomal dominant polycystic kidney disease
AP-1	Activator protein-1
ARPKD	Autosomal recessive polycystic kidney disease
BLAST	Basic local alignment search tool
bp	Base pair(s)
bpk	BALB/c polycystic kidneys
cAMP	3',5'-cyclic adenosine monophosphate
CAT	Chloramphenicol acetyltransferase
cDNA	Complementary DNA
CHEF-DR	Contour-clamped homogeneous electric field, dynamic regulation
cM	CentiMorgan (1% recombination, or one map unit)
cpk	Congenital polycystic kidneys
CREB	cAMP responsive element binding [protein]
dGTP	Deoxyguanosine-5'-triphosphate
dITP	Deoxyinosine-5'-triphosphate
DNA	Deoxyribonucleic acid
Ds	Disorganization
EDTA	Disodium ethylenediamine tetraacetate
EGF	Epidermal growth factor
EGF-R	Epidermal growth factor receptor
ES cells	Embryonic stem cells
ESP	2 mg/ml Proteinase K, 1% Sarkosyl and 0.5 mM EDTA, pH 8.0
GCG	Genetics computing group
jck	Juvenile cystic kidneys
kb	Kilobase(s)
LMP	Low melting point

MDCK	Madin Darby canine kidney
MOPS	3-(N-morpholino)-propanesulfonic acid
MW	Molecular weight
Na^+-K^+-ATPase	Sodium-potassium adenosine triphosphatase
OK	Opossum kidney
ORF	Open reading frame
^{32}P-dCTP	Deoxyadenosine-5'-[α-^{32}P] triphosphate
PCR	Polymerase chain reaction
pcy	Polycystic kidneys
pH	-log [H^+]
PFGE	Pulsed-field gel electrophoresis
PKD	Polycystic kidney disease
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
SCH	Somatic cell hybrid
SDS	Sodium dodecyl sulfate
SGP-2	Sulfated glycoprotein-2
SSC	Saline sodium citrate
SV40	Simian virus-40
TBE (1X)	89 mM Tris base, 89 mM boric acid, 2.0 mM EDTA, pH 8.0
TE	Tris-EDTA
TGF-α	Transforming growth factor-α
TPR	Tetratricopeptide repeat
UV	Ultraviolet
WBS	White belly spot

I. INTRODUCTION

Polycystic Kidney Disease

Description and Incidence

Polycystic kidney disease (PKD) is a serious heritable disorder characterized by the bilateral formation of multiple, nondysplastic cystic dilatations in the nephrons and/or collecting ducts of the kidney (Bernstein and Gardner, 1986; Gardner, 1990; Avner, 1994) that is associated with epithelial proliferation. Eventually the cysts lead to renal failure due to compression and ultimate destruction of the normal renal parenchyma. Although cystic kidneys are a well-recognized component of several complex genetically-inherited syndromes (Zerres *et al.*, 1984; Cideciyan *et al.*, 1993; Patau *et al.*, 1960; Hsai *et al.*, 1971; Saraiva and Baraitser, 1992; Tieder *et al.*, 1982; Reeders, 1990), two major clinical forms of PKD predominate in the human population. Autosomal dominant PKD (ADPKD), the third most common cause of renal failure in the United States (Iverson *et al.*, 1982), is most often associated with an adult onset and occurs at a frequency of up to 1 in 200 individuals (Gabow, 1990). The less common, but more severe, autosomal recessive PKD (ARPKD) typically affects infants and children at an estimated incidence of up to 1 in 6000 live births (Grantham, 1983), although prenatal deaths and stillbirths due to ARPKD increase that incidence by 27 times in some parts of the world (Dalgaard, 1971).

ADPKD and ARPKD exhibit many differences, which is consistent with linkage analysis demonstrating that the two are probably not allelic (Wirth *et al.*, 1987). One striking difference between the two disorders is the nearly invariable presence in ARPKD of noncystic lesions of the liver, ranging from biliary dysplasia to portal fibrosis (Kissane, 1974). In

addition, individuals affected with ADPKD generally do not experience symptoms until the third or fourth decade of life, and it has been estimated that approximately 50% of ADPKD patients develop end-stage renal disease by the age of 58 (Parfrey *et al.*, 1990). In contrast, ARPKD kills over half of its victims by 15 years of age (Kaplan *et al.*, 1989), and available treatments are only palliative (Avner, 1994).

Diagnosis and Search for the Molecular Basis of PKD

Renal cystic disease can be induced by chemicals (Gardner and Evan, 1983; Carone *et al.*, 1974) or traumatic injury (Evan and Gardner, 1985), and can also be acquired as a consequence of primary and secondary renal diseases (Grantham, 1991). However, the overwhelming majority of PKD has a genetic basis, and the search has intensified over the years for the molecular basis of heritable forms of PKD.

Rather extensive overlapping of clinical features and phenotypic variability as well as the relatively low individual frequencies of most cystic diseases have contributed to a lack of consensus among practitioners in the diagnosis of many of these disorders (Bernstein *et al.*, 1987b; Casamassima *et al.*, 1987; Breuton *et al.*, 1990; Cobben *et al.*, 1990). In addition, there is sometimes a lack of agreement as to whether kidneys are dysplastic [altered metanephric differentiation (Bernstein *et al.*, 1974)] or polycystic [cystic changes that occur after apparently normal development (Strayer and Kissane, 1979; Hsai *et al.*, 1971)], and past literature often failed to distinguish dysplastic from polycystic kidneys, an important distinction that is generally made today.

Clearly, the identification of the molecular bases for the various forms of PKD would dramatically improve diagnostic capabilities. Although the etiology of several syndromes associated with cystic kidneys has been traced to detectable chromosomal aberrations (Barakat and Butler, 1987),

attempts to localize the mutant genes have been largely unsuccessful. However, the search for cystic disease genes has yielded some promising results. The gene for recessive medullary cystic kidney disease has been mapped to human chromosome 2p (Antignac *et al.*, 1993), and recently the genes responsible for three disorders associated with renal cysts - von Hippel-Lindau disease (Latif *et al.*, 1993), Zellweger syndrome (Gärtner *et al.*, 1992), and tuberous sclerosis (European Chromosome 16 Tuberous Sclerosis Consortium, 1993) - were cloned. In addition, the locus responsible for at least 85% of ADPKD *(PKD1)* has been localized to human 16p13.3 (Reeders *et al.*, 1985; Reeders, 1992; Germino *et al.*, 1993), and non-*PKD1* forms of ADPKD have been linked to additional chromosomes (Nørby and Schwartz, 1990; Wright *et al.*, 1993; Daoust *et al.*, 1993), including human 4q13-q23 (Peters *et al.*, 1993; Kimberling *et al.*, 1993). The entire *PKD1* locus has been isolated on a series of overlapping clones (Germino *et al.*, 1992), and although at least 23 genes have been identified in this 750-kb region (Reeders, 1992), it still remains unclear which of these genes, if any, is associated with ADPKD. No chromosomal localization has yet been identified for ARPKD.

Because no molecular lesion has been directly associated with ADPKD or ARPKD, many investigators have attempted to inferentially identify PKD genes on the basis of the abnormal expression of certain genes in cystic kidneys from humans and/or animal models. In addition, the proliferative nature of the epithelial lesions associated with PKD (Grantham, 1983; Avner, 1994) has focused a great deal of interest on genes controlling cellular growth, such as growth factors and proto-oncogenes, as possible targets for investigation (Kelley *et al.*, 1991; Avner, 1994). Indeed, analysis of the pathology of PKD in both mice and humans has revealed a multitude of defects;

however, little has yet been learned about the specific initiating factor(s) in PKD.

Matrix remodeling, fluid accumulation and epithelial proliferation are generally recognized as constant features in polycystic kidneys, suggesting that these processes play critical roles during cystogenesis (Grantham, 1990). These and other insights have led to experiments demonstrating the correlation of cystic disease with the reversed polarity of surface proteins on cystic epithelium (Wilson *et al.*, 1991; Avner *et al.*, 1992; Avner and Sweeney, 1992), and the altered expression of growth factors, proto-oncogenes, basement membrane components and other genes (Ebihara *et al.*, 1988; Gattone *et al.*, 1990; Cowley *et al.*, 1991; Harding *et al.*, 1991; Torres *et al.*, 1992; Aziz *et al.*, 1993; Carone *et al.*, 1993).

Matrix remodeling in cystic disease. Basement membrane defects are common to diseases involving cellular proliferation (Wilson and Sherwood, 1991), a fact consistent with studies on cystic tissues that have revealed alterations in the structure and composition of the basement membrane. Changes in the synthesis and/or degradation of basement membrane components including fibronectin, heparan sulfate proteoglycan (Carone *et al.*, 1988; Wilson, 1991), type IV collagen and laminin B1 and B2 chains (Ebihara *et al.*, 1988) have been observed in various forms of PKD. Although it is uncertain whether such changes reflect a primary or secondary effect in cystic disease, it has been proposed that changes resulting in matrix-tubular epithelium interactions modulate cyst formation (Avner, 1993), and that these interactions, and not the cysts themselves, lead to the progression of pathologic changes in the kidney (Bennett *et al.*, 1992).

Fluid accumulation and composition in cystic disease. Studies showing that renal cysts vary in fluid composition and other characteristics have proved useful in developing an understanding of cystogenesis. Some cysts are able to maintain steep concentration gradients of K^+, Cl^-, H^+ and certain amino acids (Gardner, 1969), and in 1976, the ouabain-sensitive Na^+-K^+-ATPase was implicated in the etiology of ARPKD (McGeoch and Darmady, 1976). A few years later the first direct evidence for active transport in cysts from human polycystic kidneys was obtained (Perrone, 1985). These and other studies demonstrated that cysts have retained the ability to actively transport substances and are not simply passive, urine-filled sacs once they have "pinched off" from the tubule segment of origin.

Investigations into cyst fluid composition have revealed a heterogeneity of cyst contents among different cystic diseases and sometimes among cysts from the same kidney (Gardner, 1969; Gardner *et al.*, 1992). As evidence accumulated that active transport occurs in cystic epithelium, a concept of dynamic cystogenesis emerged, leading to detailed compositional analyses of cystic fluids. Cysts from symptomatic ADPKD patients have been found to contain cytokines including interleukin-1β, interleukin-2, tumor necrosis factor alpha, stromelysin and prostaglandin E_2 (Gardner, 1969; Gardner *et al.*, 1991; Gardner *et al.*, 1992). Cytokines are proteins, and as such, are not normally present in appreciable quantities in the normal kidney. Most plasma proteins are filtered in the glomerulus and the few that do enter the proximal tubule are reabsorbed by pinocytosis (Langley, *et al.*, 1974).

In addition, mitogenic concentrations of transforming growth factor alpha (TGF-α) and epidermal growth factor (EGF) have been found in cyst fluid (Wilson and Sherwood, 1991; Horikoshi *et al.*, 1991). Although the precise role of

these factors in cyst formation is unknown (Gardner *et al.*, 1991), evidence that they are involved in the proliferative process in PKD comes from studies showing that cyst fluid from ADPKD patients promotes cyst formation in renal epithelial cells *in vitro* (Ye *et al.*, 1992).

Apical mislocalization of the Na^+-K^+-ATPase: implications for fluid transport in cysts. The alterations in the normal fluid and ion transport that are associated with abnormal fluid accumulation in renal cysts have led some investigators to examine the polarity of cellular proteins involved in active transport across the tubular epithelium. For example, after embryogenesis, the Na^+-K^+-ATPase is normally localized to the basolateral surface of the cells in which it is contained. Mislocalization of the Na^+-K^+-ATPase to the apical surface of cystic epithelium has been reported in ADPKD (Wilson *et al.*, 1991) and in cystic collecting tubules from *cpk* mice (Avner *et al.*, 1992), but not in ARPKD (Avner and Sweeney, 1992). Na^+-K^+-ATPase localization was restricted exclusively to the apical surface of cystic epithelium in ADPKD cysts (Wilson *et al.*, 1991), but was detected apically and, to a lesser extent, basolaterally, in the *cpk* cysts (Avner *et al.*, 1992). Because Na^+ is the main osmotic determinant in the kidney (Wilson and Sherwood, 1991), apical mislocalization in these instances is hypothesized to result in fluid accumulation in cysts by mediating transport of Na^+, and hence, fluid, from the basal to the apical (lumenal) aspect of the cell (Avner, 1993). Moreover, because normal localization of the Na^+-K^+-ATPase is apical during embryogenesis, and basolateral in differentiated cells, (Holthofer, 1987; Avner *et al.*, 1992), the mislocalization may also reflect a less differentiated state of cystic epithelium (Avner, 1993).

Disparate findings may reflect the complex nature of PKD. It should be noted here that disparate results by independent investigators and seemingly contradictory findings among the various forms of PKD in both humans and mice demonstrate the complexity encountered among renal cystic disorders and argue against a single mechanism of cystogenesis in any species. For example, in contrast with reports of apical mislocalization in ADPKD kidneys (Wilson *et al.*, 1991), at least one investigator has found apical mislocalization of the Na^+-K^+-ATPase in ADPKD kidneys to be a rare event (Carone *et al.*, 1992). In addition, the Na^+-K^+-ATPase is not mislocalized in ARPKD kidneys (Avner and Sweeney, 1992), demonstrating that other factors can result in the abnormal fluid accumulation in cysts.

Another example of heterogeneity among cystic diseases is expression of sulfated glycoprotein-2 *(SGP-2)*. Whereas *SGP-2* expression fails to be down-regulated in *cpk* kidneys (Harding *et al.*, 1991), sulfated glycoprotein synthesis is decreased in ADPKD kidneys, and the Golgi processing of the proteins is delayed (Carone *et al.*, 1993). Clearly, PKD provides a useful, albeit complex, model system for the study of differentiation and pathogenesis throughout development and, more specifically, during nephrogenesis.

Epithelial proliferation in renal cystic disease. The recognition that the increased surface area of renal cysts could not be explained by a compensatory thinning and stretching of the resident cells sparked the notion that cystogenesis is a process involving cellular proliferation (Grantham, 1983; Welling and Welling, 1985). Further evidence for the proliferative nature of PKD came from the identification of the abnormal expression and/or localization in cysts of various cellular proteins known to be involved in

mitogenic stimulation. For example, although EGF expression is decreased in *cpk* kidneys, the concentrations of EGF and EGF-like peptides are at mitogenic levels in *cpk* cysts (Horikoshi *et al.*, 1991; Wilson and Sherwood, 1991). Furthermore, the epidermal growth factor receptor (EGF-R) is mislocalized to the apical surface in cysts from ADPKD and ARPKD patients as well as *cpk* mice (Avner *et al.*, 1991; Du and Wilson, 1990; Avner and Sweeney, 1992). The EGF-R also serves as a receptor for transforming growth factor-α (TGF-α), and because EGF-R is expressed at higher levels in ADPKD and ARPKD kidneys (Avner, 1993), recent research has focused on the potential proliferative role of TGF-αin PKD (Avner, 1994).

PKD may reflect a less differentiated state. Apical mislocalization of the Na^+-K^+-ATPase is but one example of growing genetic and morphological evidence that PKD involves abnormalities in cellular differentiation. Studies of PKD patients and animal models have identified several biochemical and physiological abnormalities, in addition to apical mislocalization of the Na^+-K^+-ATPase, that are consistent with a less-differentiated state.

Most early ADPKD cysts contain cells that can be recognized as having originated from specific tubular segments (Evan and McAteer, 1990). However, in one study of 387 cysts from 10 patients, the tubular origin of 84% of the cysts could not be identified (Grantham *et al.*, 1987).

Much of the work involving the role of cellular differentiation in PKD has been carried out with the *cpk* mouse. Sulfated glycoprotein-2 *(SGP-2)*, a protein normally down-regulated during renal maturation, continues to be expressed at high levels in renal cysts from *cpk* mice (Harding *et al.*, 1991). It has also been shown that *SGP-2* expression is restricted to specified cell types immediately

preceding apoptosis (Connor *et al.*, 1991), although it is unknown whether this is a cause or effect phenomenon. Only recently was it discovered that the kidney undergoes apoptosis as part of its normal development (Koseki *et al.*, 1992; Coles *et al.*, 1993); hence *SGP-2* may reflect a less-differentiated state, serving as a marker for metanephric apoptosis (Bettuzzi *et al.*, 1989; Connor *et al.*, 1991).

The proto-oncogene c-*myc*, normally expressed in immature tubules during cellular proliferation, continues to be expressed postnatally in *cpk* kidneys at abnormally high levels (Cowley *et al.*, 1987). Constitutive expression of c-*myc* has been shown to block differentiation in mouse erythroleukemia cells (Coppola and Cole, 1986), F9 teratocarcinoma cells (Griep and Westphal, 1988), pre-B-cells (Schmidt *et al.*, 1988) and rat myoblasts (Denis *et al.*, 1987), prompting speculation that c-*myc* acts to prevent differentiation in *cpk* kidneys (Harding *et al.*, 1992).

cpk kidneys also have reduced levels of N-cell adhesion molecule and E-cadherin transcripts, which are thought to guide the differentiation and polarization of renal epithelium (Rocco, *et al.*, 1992), and ultrastructural data from the *cpk* mouse are consistent with a problem in terminal differentiation (Calvet, 1993).

Finally, epidermal growth factor, which is thought to function as a differentiation agent during renal development, is down-regulated in *cpk* and *pcy* kidneys (Gattone *et al.*, 1990; Gattone and Lowden, 1992; Ebihara *et al.*, 1992). Furthermore, exogenous administration of EGF to *cpk* mice has been shown to retard cyst development (Gattone and Lowden, 1992).

The role of proto-oncogenes in renal cystic disorders. The epithelial proliferation and problems in terminal differentiation associated with PKD have drawn particular attention to the role proto-oncogenes may play in cystic disease (Avner *et al.*, 1990). To date, there have been reports of increased renal expression of c-*myc*, c-*fos*, c-Ha-*ras* and c-*Ki-ras* in *cpk* mice (Cowley *et al.*, 1987; Cowley *et al.*, 1991; Lindemann *et al.*, 1992); c-*myc* in the *Cy* rat model of ADPKD (Cowley *et al.*, 1993); and *C-erbB-2* in ADPKD (Herrera, 1991).

The existence of several transgenic mouse models for cystic disease has virtually removed any doubt that proto-oncogenes can play a role in cystogenesis. Trudel and colleagues (Trudel *et al.*, 1991) unexpectedly induced renal cysts in all 18 lines of mice carrying a transgenic construct expressing the mouse c-*myc* proto-oncogene under the control of the β-globin promoter and Simian virus (SV40) enhancer, evidence that c-*myc* overexpression is sufficient to induce PKD in mice. Renal cysts developed in two other transgenic lines of mice into which the SV40 early region was stably incorporated (MacKay *et al.*, 1987; Kelley *et al.*, 1991), suggesting an oncogenic role in PKD. In other experiments, mice carrying a construct that resulted in the constitutive expression of the c-*erbB-2* oncogene developed an unusual and pronounced epithelial proliferation with dilatation and microcystic changes in the cortical renal tubules (Stöcklin *et al.*, 1993). Renal cysts have also developed in transgenic mice that overexpressed growth hormone (Wanke *et al.*, 1991) or the *ras*T24 oncogene (Schaffner *et al.*, 1993). Finally, mice that fail to express the *bcl-2* oncogene, an inhibitor of apoptosis, develop polycystic kidneys (Veis *et al.*, 1993). The induction of cystic kidneys in all of these mouse lines not only demonstrates the considerable heterogeneity in the etiology of renal cystogenesis but is also consistent

with a proliferative role in PKD, since the overexpressed genes are normally involved in cellular growth regulation.

Animal Models of PKD

In addition to the models of PKD generated by the intentional disrupted expression of specific genes, many spontaneous mutations exist that cause PKD in animals. Animal models of PKD that resembles the various forms of human PKD have been extremely useful for studying the pathogenesis of cystic kidney disease. The recessive *cpk* (congenital polycystic kidneys) mutation (Russell and McFarland, 1977) has been examined in considerable detail (Avner *et al.*, 1987; Cowley *et al.*, 1991; Avner *et al.*, 1992; Avner and Sweeney, 1992). More recently, additional spontaneous recessive mutations in mice - *bpk* (BALB/c polycystic kidneys) (Natau *et al.*, 1993), *pcy* (polycystic kidneys) (Takahashi *et al.*, 1991), and *jck* (juvenile cystic kidneys) (Atala *et al.*, 1993) - have been reported. The cystic disease that develops in the *cpk*, *bpk* and *jck* models shows similarities to clinical features of human ARPKD, but the dual hepatorenal pathology that is a hallmark feature of ARPKD has only been reported as a consistent feature in the *bpk* mouse on the BALB/c genetic background (Natau *et al.*, 1993). The *pcy* locus has been mapped to a region of mouse chromosome 9 that corresponds with human chromosome 3 (Takahashi *et al.*, 1991) but, to date, a locus on chromosome 3 has not been associated with a human renal cystic disease. The *cpk* locus has been mapped to a region of mouse chromosome 12 with homology to human chromosome 2p24-2p25 (Guay-Woodford *et al.*, 1993), but it has been shown that the gene responsible for ARPKD is not linked to this region (Guay-Woodford *et al.*, 1990). In none of these models has the mutant gene been identified, primarily

because of a lack of DNA probes directly associated with the locus. Therefore, at present it is impossible to test directly whether the gene associated with any of these mutations corresponds to a PKD gene in humans.

Insertional Mutagenesis Facilitates Molecular Access to Disease-Causing Genes

Experimental mutagenesis is a straightforward and powerful approach to the correlation of a specific gene with its function. Because transgene sequences - introduced into the germ line by retroviral or pronuclear microinjection methods (Hogan *et al.*, 1986) - are not endogenous to the organism's genome, the genetic mutations they generate can often be accessed by exploiting the transgene as a molecular "tag" for the affected locus (Gordon *et al.*, 1980). Transgenic hemizygotes (animals carrying the transgene on one chromosome) are subsequently bred to homozygosity so that recessive mutations can be detected.

An abnormal phenotype segregating exclusively with the transgene for several generations is likely due to a genetic mutation at or near the integration site. DNA flanking the integration site can be isolated and cloned by utilizing the transgene as a molecular recognition "tag" for the integration site. Nonrepetitive genomic sequences at or near the transgene insertion can then be utilized as probes to detect restriction fragment length polymorphisms (RFLPs) between transgenic and wild-type mice. Such information can then be used to construct restriction maps of the wild-type and mutant locus. Furthermore, once RFLPs have been identified, the genotypic analysis of a transgenic line is independent of the phenotype. This is particularly advantageous when mutant effects appear late in development, or when the

mutation is recessive, as the need for progeny testing is negated and animals can be genotyped even before birth.

Once genomic flanking DNA from the locus has been isolated and cloned, the fragments can be used to probe Southern blots containing DNA from other species. This "zooblot" technique is based on the concept that the coding portions of genes are conserved throughout evolution and will be recognized among various species (McGinnis *et al.*, 1984).

Transgenic animals generated by insertional mutagenesis provide the opportunity to identify important disease-causing loci in humans by serving as physiologic and genetic models for many human diseases. Two examples of human disorders and their genetic equivalents in mice include human type II oculocutaneous albinism and the pink-eyed dilution locus (Rinchik *et al.*, 1993), and Waardenburg syndrome (type I) and *splotch* (Hill and Van Heyningen, 1992). In addition, a great number of spontaneous and experimentally-induced mutations in mice exist that cause phenotypes resembling human diseases (Winter, 1988; Hodgkinson *et al.*, 1993); whether such mutations represent alterations in homologous genes awaits the isolation and characterization of the mutant loci.

An animal with an insertional mutation provides both an *in vivo* model for a particular disorder as well as a "roadmap" for the affected locus. The isolation of several genes in transgenic mice including *Mov-13* (Schnieke *et al.*, 1983), *limb deformity* (Woychik *et al.*, 1990), *Mov-34* (Gridley *et al.*, 1991), *Mpv-17* (Weiher *et al.*, 1990), *legless* (Singh *et al.*, 1991), Hβ58 (Lee *et al.*, 1992), microphthalmia *(mi)* (Hodgkinson *et al.*, 1993) and others underscores the merits of insertional mutagenesis in expediting the journey from an abnormal phenotype to its molecular basis. Moreover, the identification of a gene whose expression has been altered by

transgene integration permits the subsequent study of the gene' s role during normal and pathologic development.

Mice from the Tg737 Line Contain an Insertional Mutation Causing Recessive Polycystic Kidney Disease

As part of a large scale mutagenesis program in the Woychik laboratory, a new insertional mutation was generated in the Tg737 line of mice that gives rise to a recessive form of PKD with striking similarities to human ARPKD. The cardinal features of both the kidney lesion (collecting tubule ectasia) and the liver lesion (biliary dysplasia and/or portal hepatic fibrosis) typical of human ARPKD are a constant finding of the disease in these mutant mice (Moyer *et al.*, 1994a). Unlike other mutations causing cystic kidneys in mice, this mutation arose by insertional mutagenesis and is called *TgN(Imorpk)737Rpw* (*Imorpk*: insertional mutation, Oak Ridge polycystic kidneys)[named in accordance with the new nomenclatural rules for transgenic mice (Gordon, 1992), and abbreviated here as *TgN737Rpw*]. A unique feature of this mouse line is that, for the first time in an animal model, the disruption of a specific, endogenous locus has been directly linked to PKD. This study involves the molecular and genetic characterization of the *TgN737Rpw* mutation and addresses the possibility that the *TgN737Rpw* phenotype represents the murine counterpart of a form of PKD in humans.

II. PROJECT BACKGROUND

Research on the *TgN737Rpw* mouse line documented in this thesis was a continuation of a significant amount of previous molecular work, conducted primarily by Drs. Richard P. Woychik and Monica Lee-Tischler, which has not been published in detail. Therefore, a concise description of the findings from earlier work, designated "Project Background", is necessary to establish a framework for an understanding of the doctoral project reported here.

The Mutant Phenotype Cosegregates with the Transgene in the Tg737 Line

The *TgN737Rpw* line was established from matings of the *TgN737Rpw* founder animal with a wild-type FVB/N mouse. Transgenic progeny were identifed by genomic Southern blot analysis of DNA from distal tail tissue biopsies (Hogan *et al.*, 1986), using a transgene-derived probe. Initially, quantitative Southern blots enabled genotypic discrimination between transgenic hemizygotes and homozygotes on the basis of signal intensity.

Crosses between animals heterozygous for the mutant *TgN737Rpw* allele (hemizygous for the transgene) produced some offspring in which polydactyly, scruffy fur and runting were observed externally. Necropsies performed by Dr. Virginia Godfrey revealed bilateral polycystic kidney disease and lesions associated with the intrahepatic biliary tract, two well-recognized features of human autosomal recessive polycystic kidney disease (ARPKD) (Bernstein, 1978; Bernstein and Gardner, 1986). Mutant progeny represented approximately 25% of the offspring in each litter, and never resulted from crosses between *TgN737Rpw* heterozygotes and wild-type animals, demonstrating that the mutant

phenotype always cosegregated with the transgene in a recessive fashion.

Identification of a Restriction Fragment Length Polymorphism Permits Genotypic Analysis

Genomic DNA from mutant animals was digested with several different restriction enzymes, and analyzed by the Southern blot technique (Sambrook *et al.*, 1989), utilizing a transgene-derived probe. Digestion with *Bam*HI cleaved the transgene once, indicating that approximately only one copy of the transgene (5.0-6.5 kb) had integrated at the mutant locus. Cleavage with *Bam*HI yielded two fragments (4.5 and 7.6 kb) that presumably contained genomic flanking DNA (Fig. 1A)*. The fragments were subsequently isolated from a genomic library using a transgene-derived probe, and designated pλA6-7.6 and pλB1-4.5. Both subclones were digested with several restriction enzymes and probed with transgene sequences to identify nontransgene-containing (flanking) genomic regions. Digestion of pλA6-7.6 with *Hind*III and *Eco*RI produced a 4.2-kb *Eco*RI chimaeric fragment (HRA) that contained transgene and flanking sequences, and *Eco*RI digestion of pλB1-4.5 generated 0.5, 0.8 and 3.1-kb fragments (B1-C, B1-B and B1-A, respectively) (Fig. 1A). B1-A was also a chimaeric fragment, containing both genomic and transgene sequences (Fig. 1A).

Hybridization of *Eco*RI-digested DNA from wild-type, heterozygous and homozygous *TgN737Rpw* mice with radiolabeled HRA identified a 4.2-kb mutant fragment and a corresponding 3.0-kb wild-type fragment (Fig. 1B). This restriction fragment length polymorphism (RFLP) permitted

* All figures are in the Appendix.

subsequent qualitative genotypic analysis of *TgN737Rpw* mice.

Isolation of Wild-Type Clones Corresponding with the *TgN737Rpw* Locus

DNA from the liver of a wild-type animal was partially digested with *Sau*3A1 and used in the construction of two EMBL3 libraries which were screened with HRA and B1-C, two probes containing DNA flanking the insertion site. This led to the isolation of four clones that hybridized either with HRA (20-3 and E'4) or B1-C (20-1 and 20-4) (Fig. 1A).

Stable Cointegration of Genomic Mouse DNA at the Transgene Insertion Site

Characterization of the four clones (20-3, E'4, 20-1, and 20-4) by restriction enzyme digestion with *Sal*I, *Eco*RI and *Bam*HI revealed that clones recognized by B1-C (20-1 and 20-4) shared no sequence overlap with clones recognized by HRA (20-3 and E'4) (Fig. 1A). Because HRA and B1-C represented sequences that *presumably* were almost adjacent to one another in the corresponding wild-type genomic locus, the lack of sequence overlap among the clones isolated with those probes suggested that a rearrangement such as a deletion, translocation or inversion had occurred as a result of the integration event.

A Southern blot containing mouse x Chinese hamster somatic cell hybrid (SCH) DNAs was probed with HRA and B1-B in separate experiments. Although HRA hybridized with one SCH that contained chromosomes 14 and 15, B1-B did not (data not shown), suggesting that a region of DNA not endogenous to chromosome 14 had cointegrated with the transgene sequences. This was consistent with the failure of

clones 20-1 and 20-4 to overlap with 20-3 and E'4. SCH analysis restricted the endogenous location of the cointegrated fragment to chromosome 1, 2, 3, 5, 6, 8, 12, 13, 18 or 19. Subsequent Southern blot results (data not shown) were also consistent with B1-B having originated from elsewhere in the genome, and stably cointegrating with the transgene.

Karyotypic analysis conducted by Dr. N.L. Cacheiro on metaphase chromosome preparations from a *TgN737Rpw* heterozygote revealed no cytogenetically-detectable rearrangements, suggesting that the cointegrated DNA fragment was not extremely large.

Derivation of Useful DNA Probes from λ Clones by Restriction Digestion

Large scale digestion of wild-type λ clone E'4 with *Sal*I and *Eco*RI released several insert fragments that were gel isolated (Hogan *et al.*, 1986) and labeled as fragments A, B, C, D and E (5.0, 4.0, 3.0, 2.1 and 2.0 kb, respectively) (Fig. 1A); these restriction fragments were later utilized as probes in the characterization of the *TgN737Rpw* locus. Southern hybridization studies of wild-type, *TgN737Rpw/+* and *TgN737Rpw/TgN737Rpw* DNA with radiolabeled fragment C revealed that the *Eco*RI site defining the 5' boundary of fragment C had been deleted from the mutant locus (Fig. 1A).

In summary, initial experiments with the *TgN737Rpw* line resulted in the identification of a complex mutant phenotype that segregated with the transgene in a recessive manner. A transgene-derived probe was utilized to isolate and clone genomic sequences flanking the transgene integration site. From these flanking sequences, several restriction fragments were derived which were useful in the initial characterization of the *TgN737Rpw* locus. One of these,

HRA, allowed mice from the Tg737 line to be genotyped on the basis of an RFLP. SCH analysis localized the transgene integration to chromosome 14 or 15, and Southern hybridization experiments revealed a deletion of unknown size, as well as the insertion of cointegrating DNA, at the insertion site. No rearrangements were detectable cytogenetically.

III. SPECIFIC AIMS

The specific aims of this research project involved the molecular and genetic analysis of the *TgN737Rpw* locus and its wild-type counterpart:

1. To characterize the molecular structure of the mutant and wild-type locus, including characterization of the cointegrated (translocated) DNA with regard to size, structure, and whether or not the translocated DNA originated from one or more chromosomal sites.

2. To identify a candidate PKD gene at or near the *TgN737Rpw* locus whose expression was disrupted, and then to isolate the gene's wild-type counterpart. Wild-type and mutant DNA clones from the *TgN737Rpw* locus were to be utilized in cross-species hybridization assays to facilitate the identification of exons on the basis of evolutionary conservation.

3. To identify transcripts by using exon fragments from the gene of interest as probes on Northern blots containing poly(A)$^+$ RNA from mutant and wild-type animals.

4. To utilize exon-containing genomic DNA to isolate a full-length murine cDNA representing the gene at the *TgN737Rpw* locus.

5. To gain insight into the gene's possible function by the identification of domains/motifs in the gene and/or homology with previously identified genes with known functions.

6. To apply the cDNA to temporal and spatial expression studies in cell lines and in tissues from wild-type and mutant animals.

7. To determine the location of transgene insertion relative to the gene (i.e., intragenic versus intergenic insertion).

8. To determine the intron/exon structure of the *TgN737Rpw* gene and orientation of transcription of the gene at the genomic level.

9. To develop a model explaining the possible role the identified gene plays in the development of the mutant phenotype.

10. To initiate the characterization of any homologous human gene using DNA probes from the mouse cDNA.

IV. MATERIALS AND METHODS

Generation of Mice

The Tg737 transgenic line was generated on the FVB/N inbred genetic background by pronuclear microinjection (Gordon *et al.*, 1980; Hogan *et al.*, 1986) of the *Hind*III-linearized pPyF9-1 construct, which contains the bacterial chloramphenicol acetyltransferase (CAT) gene under the control of a mutated version of the polyoma virus early region promoter (Böhnlein *et al.*, 1985). Because transgenic homozygotes do not breed successfully, the line has been maintained with transgenic hemizygotes, which have no detectable abnormalities.

Hybridization Probe Radiolabeling

All probes were prepared by the glass powder gel elution method (Hogan *et al.*, 1986), brought up in TE (10 mM Tris, 0.1 mM EDTA, pH 7.6) to a concentration of 50 ng/λ and radiolabeled according to the random hexamer-labeling method (Feinberg and Vogelstein, 1984) with [α-^{32}P] deoxycytidine-5'-triphosphate (dCTP).

Southern Analysis

Genomic DNA, obtained from tail tissue (Hogan *et al.*, 1986) or adult liver (Sambrook *et al.*, 1989), or cloned DNA was digested (where indicated) to completion with restriction endonucleases, and 10 μg (genomic DNA) or 1 μg (cloned DNA) per lane were electrophoresed through 0.8% agarose gels. Genomic or cloned DNA was transferred to GeneScreen (DuPont) or GeneScreen Plus (DuPont) membranes

respectively, using conventional procedures (Sambrook *et al.*, 1989; Ausubel *et al.*, 1991).

Somatic Cell Hybridization Experiments

DNA from mouse x Chinese hamster somatic cell hybrids was digested with *Bam*HI or *Eco*RI, electrophoresed through a 0.8% agarose gel and transferred to GeneScreen Plus (DuPont) using standard Southern blotting methods (Sambrook *et al.*, 1989). The hybrids and the mouse chromosomes they contain are as follows: BEM 1-4 (1, 2, 3, 5, 6, 8, 12, 13, 14, 15, 16, 17, 18, 19, X); MAE 32 (16 and X); ECm4e (14 and 15); R44-1 (17 and fragments of 3 and 18); E36 (none - hamster parent).

Mus musculus/Mus spretus Interspecific Backcross Mapping

M. musculus/M. spretus interspecific backcross panels and DNA from offspring of F1 hybrid females (offspring from $Sey^{MH}/+$ females x *M. spretus* males) that had been backcrossed with C57BL/10Nimr males as described in (Stubbs *et al.* 1990) were generously provided by B. Hogan. The DNA was digested with *Bgl*II, *Msp*I or *Bam*HI and hybridized in separate experiments with ^{32}P-dCTP-radiolabeled HRA (*Bam*HI panels), the 1372 base pair (bp) Tcrα-receptor cDNA (*Bgl*II panels) (Saito *et al.*, 1984), and/or a probe from the *Rb-1* locus (*Msp*I panels).

Pulsed-Field Gel Electrophoresis

High molecular weight (MW) DNA was prepared from splenic white blood cells from a wild-type and a *TgN737Rpw/+* animal as described (Herrmann *et al.*, 1987),

with slight modifications. Single cell suspensions containing approximately 10 μg DNA each were embedded in blocks of 0.75% low melting point agarose (UltraPure LMP, BRL) and digested *in situ* in ESP solution (2 mg/ml Proteinase K, 1% Sarkosyl and 0.5 mM EDTA, pH 8.0) to remove cellular proteins. After repeated dialysis with TE, DNA blocks were digested with restriction enzymes *Sal*I, *Sac*II, *Not*I or *Bss*HII, according to manufacturer's instructions overnight. One undigested block of wild-type and one of *TgN737Rpw/+* DNA were included as controls for degradation. The DNA samples and yeast chromosome size standard (*S. cerevisiae*, strain YNN295, Bio-Rad) were electrophoresed with the CHEF-DR™ II Pulsed-Field Electrophoresis System (Bio-Rad) through a 14 x 12.7 cm 1.0% agarose (UltraPure, BRL) gel in 0.5 X TBE (44.5 mM Tris base, 44.5 mM boric acid, 1.0 mM EDTA, pH 8.0) running buffer at a constant temperature of 14°C. Electrophoretic parameters: 200 V; pulse time was a ramp from 50-100 seconds over a total run time of 28 hours. The gel was UV-irradiated (Stratalinker) prior to membrane transfer.

Library Construction

Isolation of pλA6-7.6, pλB1-4.5 and Subsequent Derivation of B1-A, B1-B, B1-C and HRA

DNA from the liver of an (FVB/N)-*TgN737Rpw* homozygote was digested to completion with *Bam*HI and size-fractionated on a 10%-40% sucrose gradient (Sambrook *et al.*, 1989). An EMBL3 phage library was constructed with hybridizing fractions in the 3.0-8.0-kb range, packaged according to manufacturer's directions (Gigapack Gold Kit, Stratagene) and screened with a radiolabeled transgene-derived probe. The 7.6-kb (pλA6-7.6) and 4.5-kb (pλB1-4.5) fragments (Fig. 1A) were isolated according to standard

procedures (Sambrook *et al.*, 1989; Ausubel *et al.*, 1991). pλA6-7.6 was subcloned into the *Bam*HI site of the pSP64 vector (Promega) and cut with *Hind*III and *Eco*RI to generate the *Eco*RI fragment HRA (Fig. 1A). pλB1-4.5 religated upon itself and could not be subcloned, apparently due to the presence of internal vector sequences from the transgene portion of the fragment. *Eco*RI-digestion of pλB1-4.5 yielded fragments B1-A (3.1 kb), B1-B (0.8 kb) and B1-C (0.5 kb).

Isolation of Wild-Type Genomic Clones 20-1, 20-3, 20-4 and E'4

DNA from the liver of a wild-type FVB/N animal was partially digested with *Sau*3A and size-fractionated on a 10%-40% sucrose gradient (Sambrook *et al.*, 1989). An EMBL3 phage library was constructed with hybridizing fractions and packaged according to manufacturer's directions (Gigapack Gold Kit, Stratagene) and screened with radiolabeled HRA (recognized by 20-3 and E'4) and B1-C (recognized by 20-1 and 20-4) and isolated according to standard procedures (Sambrook *et al.*, 1989; Ausubel *et al.*, 1991).

Derivation of Genomic Clones pλA2-5.0-3 and pλA2-8.8-7

Genomic liver DNA from an (FVB/N)-*TgN737Rpw* homozygous mouse was digested to completion with *Eco*RI, size-selected by 10%-40% sucrose gradient fractionation, cloned into the *Eco*RI site of the EMBL4 replacement λ phage vector (Stratagene) using standard techniques (Sambrook *et al.*, 1989; Ausubel *et al.*, 1991) and screened with radiolabeled HR0.2, a probe derived from the extreme 3' end of fragment A (Fig. 3).

Isolation of cDNAs from 14.5 Days Post Coitum (dpc) Total Embryos

Poly(A)$^+$ RNA was isolated from whole 14.5 dpc wild-type mouse embryos and converted to double-stranded DNA using standard cDNA library construction techniques (Sambrook *et al.*, 1989; Ausubel *et al.*, 1991); *Sma*I adaptors and *Eco*RI ends were added and the double-stranded cDNAs ligated into the λgt10 vector (Stratagene) using standard techniques (Sambrook *et al.*, 1989; Ausubel *et al.*, 1991). All *TgN737Rpw* cDNAs were isolated with radiolabeled PS1.6, an 1.6-kb *Pvu*II/*Sac*I fragment from fragment B (Fig. 7).

Isolation of 8.5 dpc cDNA

RNA from whole 8.5 dpc wild-type C57BL embryos was prepared as described above, except that poly(A)$^+$-enrichment consisted of two oligo d(T)-cellulose column passages. Double-stranded cDNA was prepared using standard procedures (Ausubel *et al.*, 1991; Sambrook *et al.*, 1989), and ligated into the λgt10 vector (Stratagene) following the addition of *Eco*RI linkers. After the library was probed with radiolabeled cDNA5 (Moyer *et al.*, 1994a) or cDNA6 (identical to cDNA5 except for a deletion of 4 nucleotides at the extreme 5' end, a shorter poly(A) tail and an adenine instead of thymine at base 2027 in cDNA5), cDNA 3B-1 was purified and subcloned into the *Eco*RI site of pGEM4 (Promega) vector using conventional techniques (Sambrook *et al.*, 1989).

Isolation of Cosmid Clone c2RBA2

Cosmid clone c2RBA2 was isolated from a cosmid library prepared from wild-type mouse DNA which was constructed in an *Eco*RI-digested c2RB cosmid vector (Bates and Swift, 1983) using standard procedures (Sambrook *et al.*, 1989;

Ausubel *et al.*, 1991) and screened with radiolabeled BB3.5 (Fig. 4).

Genomic Clones from the 5' Region of the *TgN737Rpw* Gene

Genomic splenic DNA from a wild-type FVB/N mouse was partially digested with *Sau*3A, size-selected by 10%-40% sucrose gradient fractionation and cloned into the *Bam*HI site of the EMBL3 replacement λ phage vector (Stratagene) using standard techniques (Sambrook *et al.*, 1989; Ausubel *et al.*, 1991). To select genomic clones from the extreme 5' end of the *TgN737Rpw* gene, synthetically-prepared oligonucleotide primers RW117 (5'-GGATCAGGCGTCGCTTCTTC-3') and RW127 (5'-TCCTCAGTATCATAGGCTGG-3') (Fig. 12), representing cDNA5 (Moyer *et al.*, 1994b) bases 1-20 and 141-160, respectively, were used to PCR-amplify (Ausubel *et al.*, 1991) the 5'-most region (160 bp) of cDNA5. A 111-bp fragment, representing bases 1-111 of cDNA5, was prepared similarly with RW 117 and RW 126 (5'-CGTCCTCATCTGTTTCTGGT-3') (Fig. 12). Both fragments were utilized as library screening probes.

RNA Preparation and Northern Blot Analysis

Total RNA was isolated by a modified guanidinium isothiocyanate method (Ausubel *et al.*, 1991), poly(A)$^+$-selected by 5 passages through oligo d(T)-cellulose columns (Aviv and Leder, 1972), electrophoresed through formaldehyde gels in 1 X MOPS (3-(N-morpholino)-propanesulfonic acid) running buffer and blotted to nylon membranes (GeneScreen, DuPont). Probe RW 114/134 is a 104-bp PCR product prepared with oligonucleotides RW 114 (5'-GCCAGAACAATAGTGCCAGT-3') and RW 134 (5'-TATTTCTTTGTTACCGCTAC-3') (Fig. 12), representing bases

2320-2339 and 2404-2423, respectively, from cDNA5 (Moyer *et al.*, 1994b).

Polymerase Chain Reaction Experiments

PCR-amplification of cDNAs was performed with primers from the λgt10 vector (Stratagene), using standard procedures (Ausubel *et al.*, 1991). PCR conditions: 30 cycles of denaturation (94°C for 1.5 minutes), annealing (51°C for 2 minutes) and polymerization with *Thermus aquaticus* polymerase (72°C for 3 minutes), followed by one cycle of polymerization at 72°C for 5 minutes.

Oligonucleotide Synthesis

Oligonucleotides were synthesized in a Model 391 PCR-Mate DNA Synthesizer (Applied Biosystems), using the β-cyanoethyl phosphoramidite method.

DNA Sequencing and Analysis

A modified bacteriophage T7 DNA polymerase (Sequenase™, US Biochemical) (Tabor and Richardson, 1987) was used in the Sanger dideoxynucleotide method (Sanger *et al.*, 1977) to sequence all genomic and cDNA clones. Computer analysis of sequence was conducted with the University of Wisconsin Genetics Computing Group sequence analysis programs (Devereux *et al.*, 1984) and the BLAST (Basic Local Alignment Search Tool) search algorithm (Altschul *et al.*, 1990).

V. RESULTS

Analysis of the Phenotype in *TgN737Rpw* Homozygotes

Inheritance and Penetrance of the Abnormal Phenotype in *TgN737Rpw* Homozygotes

The abnormal phenotype arose in the Tg737 line only among transgenic homozygotes, derived from crosses between animals hemizygous for the transgene. In the FVB/N inbred genetic background, external examination revealed that mutant animals were severely growth retarded, and had preaxial polydactyly on all limbs. The mice also exhibited scruffy fur, which appeared sparse, thin and sometimes oily as it emerged in newborns. The fur eventually filled in, producing a normal-appearing coat in the few animals that survived beyond weaning. In an analysis of 159 live offspring from a cross of (C3H x FVB/N)-*TgN737Rpw* heterozygotes, mice from 33 complete litters were genotyped with probe HRA for the presence of the transgene. Thirty-five (22%) animals were wild-type, 35 were homozygous (22%) and 89 (56%) were hemizygous for the transgene, demonstrating expected Mendelian frequencies and the absence of an association of prenatal homozygous lethality with inheritance of the transgene. Essentially equal numbers of male (18/35) and female (17/35) mutants were observed. Animals hemizygous for the transgene in the Tg737 line are phenotypically indistinguishable from their wild-type littermates, consistent with a recessive mode of inheritance. The phenotype in homozygotes appears to be 100% penetrant since all of 113 animals identified by genotypic analysis as transgenic homozygotes were similarly affected. The fact that a detectable mutant phenotype is expressed only by offspring homozygous for the transgene, and not by

heterozygotes from either the Tg737 line or any of 67 other transgenic lines produced with the same transgene construct, suggests that the mutant phenotype in the Tg737 line arose from an insertional mutation caused by the integration of the transgene, and not from an intrinsic phenotypic effect of the pPyF9-1 transgene.

Characteristics of the Renal and Hepatic Lesions in *TgN737Rpw* Homozygotes

As in human ARPKD (Bernstein, 1978; Bernstein and Gardner, 1986), the pathologic features of the *TgN737Rpw* phenotype include a spectrum of disease with variable severity in which the primary lesions of the kidneys and liver are expressed to different degrees. All histological and pathological studies described in this manuscript were conducted by Drs. J.E. Wilkinson, V.L. Godfrey, and E.D. Avner (Moyer *et al.*, 1994a). Internal examination of homozygous *TgN737Rpw* mice on the FVB/N inbred genetic background revealed the following constant traits: bilateral polycystic kidneys and abnormalities associated with the intrahepatic biliary tract. Gross examination of mutant kidneys revealed that they were slightly enlarged, pale, and contained numerous cysts with concurrent destruction of the surrounding parenchyma (Fig. 2A and C). The livers were slightly pale with a prominent reticular pattern.

Histologically, the lesions in the kidney and liver of *TgN737Rpw* mutant mice were remarkably similar to those seen in human ARPKD. The cardinal features of both the kidney lesion (collecting tubule ectasia) and the liver lesion (biliary dysplasia and/or portal hepatic fibrosis) typical of human ARPKD (Bernstein, 1978; Bernstein and Gardner, 1986) were a constant finding of the disease in the mutant animals. In the kidneys, an initial mild, multifocal, microscopic dilation of the proximal tubules was followed

rapidly by marked dilation and cyst formation of the collecting tubules. Lectin profiling was performed as described (Nauta *et al.*, 1993) on a total of 12 (FVB/N)-TgN737Rpw mice ranging in age from day 1 to day 28. In newborn animals, all cystic lesions were localized to proximal tubules. By postnatal day seven, cystic lesions were detected in collecting tubules, and a progressive shift in cyst localization from proximal to collecting tubules was characteristic of all subsequent developmental stages. For example, in one animal that survived to day 28, 56% of cysts were localized to collecting tubules. Expansion of the dilated collecting tubules ultimately led to the destruction of the surrounding renal parenchyma (Fig. 2C). This pattern of early proximal tubule dilation followed by a shift to dilation of the distal nephron as the predominant renal lesion is similar to that seen in the *bpk* and *cpk* murine models of ARPKD (Avner *et al.*, 1987; Nauta *et al.*, 1993). Similarly, in human ARPKD, the advanced lesions are localized to the collecting tubules (Fig. 2D) (Bernstein, 1978; Bernstein and Gardner, 1986), while proximal tubular lesions may be demonstrated at early stages of the disease (Avner and Sweeney, 1992).

Abnormalities in the intrahepatic biliary tract involving primarily the bile ducts and ductules represent a constant and prominent feature of humans with ARPKD and *TgN737Rpw* mutant mice (Fig. 2H-J). In (FVB/N)-*TgN737Rpw* mice, the liver lesion is characterized by the proliferation of epithelial cells that arise from the portal triads and form primitive, dysplastic, tortuous ductules that expand the portal and periportal areas (Fig. 2H). The proliferating cells were poorly differentiated and possessed characteristics of oval cells, a cellular population that was initially recognized in experimentally-induced rat models of hepatocellular carcinoma (Farber, 1956). In the livers of

some *TgN737Rpw* homozygotes, these cells disrupted the limiting plate and isolated individual hepatocytes.

Lifespan and Fertility of *TgN737Rpw/TgN737Rpw* Mice on the FVB/N Inbred Genetic Background

The lifespan of *TgN737Rpw* homozygotes on the FVB/N genetic background is severely shortened. Of 380 mutant animals that were allowed to die naturally, 202 (53.2%) died before the eighth day of age, and an additional 10.8% died before day 15. The average age of death for females was the 26th day; the average age of death for males was 19 days. The longest-lived male mutant died naturally at 192 days of age, the oldest female at 201 days.

In a separate population of 394 mutant mice on the FVB/N background, 206 mice of unknown sex, 90 females and 98 males were identified. These mice were checked at least every other day to ensure accurate dates of birth and death. 34/394 (8.6%) mice (13 females, 20 males and 1 of unknown sex) lived to day 42 or beyond, which coincides with the onset of sexual maturity (Hogan *et al.*, 1986; Ponzetto and Wolgemuth, 1985). Assuming a gestational period of 21 days, a female mouse could be expected to give birth as early as 63 days of age. Of the ten females that survived to the potential litter-bearing age of 63 days or beyond, pregnancy was detected in four (40.0%). Three of the mice littered, and one died while pregnant. The earliest pregnancy occurred at approximately 86 days of age, and the latest at approximately 153 days. Only one of 20 mutant FVB/N males that survived to or beyond 42 days of age was observed to be fertile; that mouse was housed in a pathogen-free barrier facility and sired a total of 27 pups with two females (V. Godfrey, personal communication).

The *TgN737Rpw* Phenotype on the C3H Genetic Background

To evaluate how genetic background might influence the expression of the *TgN737Rpw* phenotype, the mutation has been bred onto the C3H inbred genetic background. *TgN737Rpw* heterozygotes were mated with C3H mice, and offspring carrying the transgene were then backcrossed to C3H mice. *TgN737Rpw* heterozygotes from each subsequent generation were backcrossed with C3H mice for a total of 10 generations. Because the line has only recently been made congenic on the C3H genetic background (10 backcross generations), all mice described in this section represent generations 2 through 4 (75% to 93.8% C3H genetic makeup) except for 2 animals, which were the 5th and 6th generations (96.9% and 98.4% C3H, respectively).

Most *TgN737Rpw* mutant mice on the FVB/N background died during the first week of life, although a few animals have lived for several months. In contrast, (FVB/N x C3H)-*TgN737Rpw* mice had a less severe phenotype that was maintained through at least 6 generations of backcrosses to the C3H strain. Initial experiments indicate that these animals live longer, have a less severe growth inhibition, and have polydactyly that is morphologically more variable and considerably less pronounced in most animals. In addition, the animals do not appear to exhibit any fur defect, except for the presence of a white belly spot, which is described below. The longest-lived male mutant backcross mouse died naturally at 301 days of age, the oldest female at 289 days of age; in contrast, no FVB/N animal has survived beyond 201 days of age. It was not possible to obtain an accurate statistical average age of death for mutant mice on the C3H background since most were sacrificed, but the overall increase in longevity was obvious. Although many fewer FVB/N mutant animals were sacrificed than those on the C3H

background, only 18/567 (3.2%) mutant FVB/N animals that were not sacrificed survived to day 100 or beyond, whereas 21/98 (21.4%) mutant C3H animals survived to or beyond day 100.

Fertility in the backcrossed mutants is reduced, as in (FVB/N)-*TgN737Rpw* mutants. A total of 82/83 genotyped (FVB/N x C3H)-*TgN737Rpw* homozygotes reached sexual maturity (42 days): 36 (43.9%) were female and 46 (56.1%) were male. Of 17 females that lived to potential litter-bearing age, 4 (23.5%) delivered one or more litters. Due to the small numbers of mice, it is uncertain whether this percentage truly reflects a statistically significant lower fertility rate in the backcrossed mutants compared with FVB/N mutants. No fertile (FVB/N x C3H)-*TgN737Rpw* male mutants were identified, although at least 7 were mated.

The polydactyly was more variable morphologically, and less pronounced on the C3H background, sometimes appearing as "stubby" digits rather than overt polydactyly. 154 mice were scored as mutants on the basis of genotype (83/154) and/or polydactyly. 22/154 (20.8%) had unilateral polydactyly: the polydactyly was on the left side in 3/22 (13.6%) mice, on the right side in 19/22 (86.4%); toes were described as "stubby" in 12/154 (0.8%). In contrast, on the FVB/N genetic background, only 4 of over 574 animals (<0.7%) identified as mutants on the basis of severe runting and polydactyly displayed unilateral polydactyly: one on the left side, 2 on the right, and one in which the side was undocumented. Of those mice, one (right polydactyly) was genotyped as a transgenic homozygote.

Another obvious difference on the C3H background was the presence of a white belly spot (WBS) in at least 11/83 (13.3%) genotyped mutant mice. Because FVB/N mice are albino, this trait had not been previously observed in mutant animals. A WBS has never been detected in 431 wild-type or

TgN737Rpw/+ littermates of the mutant backcross progeny. Chi-square analysis of 2 x 4 contingency table data (Zar, 1984) indicates that the WBS trait is not independent of the mutant phenotype ($P<<0.001$). Scruffy fur, a manifestation of the mutant phenotype on the FVB/N background, was never observed in the mutant C3H hybrids.

Histological analysis by Dr. Wilkinson of mutant animals through the 6th backcross generation has shown that the animals develop renal cysts at a slower rate, and have a less aggressive liver lesion. Although the kidney lesions were morphologically similar between backcrossed *TgN737Rpw* mutants and mutant FVB/N mice, with all cysts being localized to collecting tubules by postnatal day 70, the liver lesions were characterized by biliary hyperplasia, dysplasia, and portal fibrosis (Fig. 2I). The significant difference in the lesions when the *TgN737Rpw* gene is on a different genetic background suggests the presence of important modifier genes (Jacob *et al.*, 1991). This may also be the case in human ARPKD, in which the severity of the phenotype and the nature of the liver lesion can be quite variable (Bernstein, 1978; Bernstein and Gardner, 1986).

Molecular Characterization and Structural Analysis of the *TgN737Rpw* Locus

Identification of a Deletion at the Integration Site

The gel-isolated *Eco*RI fragments designated A, B, C and E, derived from λ clone E'4, contained interspersed regions of nonrepetitive DNA, and were useful in the construction of a physical map of the mutant and corresponding wild-type *TgN737Rpw* locus (Fig. 3). A combination of comparative restriction mapping and hybridization studies between the mutant and wild-type locus resulted in the elucidation of the correct orientation of the fragments relative to each other

(Fig. 3). During this analysis it was discovered that while wild-type and *TgN737Rpw* heterozygous DNA hybridized with fragment E, *TgN737Rpw* homozygous DNA did not (data not shown), suggesting that all or most of the 2.0-kb region of DNA corresponding with fragment E had in fact been deleted from the mutant locus as a consequence of the integration event (Fig. 3).

The deletion of a portion of fragment C was demonstrated by the absence, from the mutant locus, of the *Eco*RI site that normally defined its 5' boundary (Fig. 3), and hybridization analysis had previously shown that HRA, the mutant *Eco*RI fragment corresponding with fragment C, contained the 3' deletion breakpoint between transgenic and genomic DNA (Fig. 1A). The 5' deletion breakpoint was defined by the cointegrated fragment (Fig. 3A). A nonrepetitive *Hind*III/*Eco*RI fragment approximately 200 bp in length (HR0.2) from the extreme 3' end of the 5.0-kb fragment A (Fig. 3B) hybridized with a larger *Eco*RI fragment in mutant DNA (data not shown), indicating that the *Eco*RI site between fragments A and E had been deleted, and that the breakpoint must have occurred within the 200 bp-region defined by HR0.2. Additional Southern blot analysis comparing the wild-type and mutant loci, and using fragment B as a probe showed no evidence of rearrangement for at least 25.0 kb beyond the 5' junction breakpoint, and similar analysis using BB3.5 (Fig. 3B) as a probe showed no evidence of rearrangement for at least 23.0 kb beyond the 3' breakpoint. These data suggested that the integration event had caused no additional rearrangements at the locus.

Characterization of the Cointegrated Fragment at the Mutant Locus

Experiments were next designed to characterize the cointegrated DNA fragment with respect to size, composition

and structure, including the nature and location of the breakpoint. Previous karyotypic analysis of metaphase chromosome preparations from *TgN737Rpw* heterozygotes had shown no evidence of major rearrangement. Although this demonstrated that the cointegrated DNA was too small to be detected cytogenetically, it did not eliminate the possibility that the insertion could be quite large.

Determination of the size of the cointegrated fragment was approached by isolating and cloning a portion of its 5' end so that it could be examined for sequence overlap with the previously isolated wild-type clone 20-4 (Fig. 1A) that corresponded with the 3' region of the cointegrated DNA. The wild-type probe HR0.2 (Fig. 3B) was used to identify and isolate a 13.8-kb *Eco*RI fragment from a genomic library prepared with DNA from a *TgN737Rpw* homozygote. The 13.8-kb fragment contained the 5' deletion breakpoint as well as DNA from the 5' flanking sequence corresponding to a region of fragment A. Digestion of the isolated clone with *Bam*HI and *Eco*RI released the 5.0-kb *Eco*RI/BamHI fragment containing 5' flanking sequences and the 5' deletion breakpoint, and the 8.8 kb *Bam*HI/*Eco*RI fragment containing only cointegrated DNA (Fig. 4A). These were subcloned and designated pλA2-5.0-3 and pλA2-8.8-7, respectively (Fig. 4A). pλA2-8.8-7 was digested with several restriction enzymes to obtain a detailed restriction map of the clone and to generate fragments that, when hybridized with total mouse genomic DNA, identified nonrepetitive sequences that would be useful as probes in Southern analysis. A nonrepetitive *Kpn*I/ *Eco*RI fragment (KR1.8) from the 3' end of pλA2-8.8-7 (Fig. 4A) was gel-isolated and used to hybridize with DNA from wild-type, heterozygous and homozygous *TgN737Rpw* animals. When probing the same blots with B1-B failed to identify any hybridizing fragments in common between the two probes, a 1.0-kb *Sal*I/*Xba*I

fragment of DNA (SX1.0) was isolated from the very 5' end of λ clone 20-4 (Fig. 4A). SX1.0 hybridized with restriction fragments unique to transgenic animals, indicating that it was included in the cointegrated DNA. KR1.8 and SX1.0 both hybridized with common mutant fragments (approximate sizes: 20.0-kb *Pst*I fragment and 16.0-kb *Pvu*II). Because the size of the lambda clones from which KR1.8 and SX1.0 were derived was known, this information permitted the elucidation of the size of the cointegrated DNA. This and other Southern blot evidence allowed the construction of a detailed restriction map of the mutant and wild-type locus and ultimately revealed that the cointegrated fragment of mouse DNA was approximately 21.0 kb in length, and had translocated as a linear, intact fragment (Fig. 4).

Characterization of Junction Breakpoint Regions

The sequence of the 3' end of pλA2-5.0-3, known to contain the breakpoint between cointegrated and endogenous DNA, was compared with the sequence of the 3' end of the corresponding wild-type 5.0-kb *Eco*RI fragment, revealing that the breakpoint was 174 bp 5' to the wild-type *Eco*RI site that defined the normal 3' boundary of fragment A (Fig. 4). Sequence analysis of 172 bp from the extreme 5' end of the cointegrated fragment, using the BLAST search algorithm (Altschul *et al.*, 1990) at the National Center for Biotechnology Information (NCBI) at the National Library of Medicine, identified a repetitive sequence, with greater than 90% identity with regions contained in mouse L1 and B1 interspersed repeat elements (Hastie, 1989; Evans and Palmiter, 1991) as well as other highly repetitive sequences. Sequence analysis of 246 bp from the 3' end of fragment A, using the FASTA database (Pearson and Lipman, 1988) in the Genetics Computing Group's (GCG) GenBank (Devereux *et al.*,

1984) showed no significant homology with previously reported nucleotide sequences (data not shown).

Next, HRA was subcloned and partially sequenced since it was known to contain the 3' breakpoint. FASTA (Pearson and Lipman, 1988) sequence analysis of 234 bp from the 5' end of HRA revealed 98% identity with the CAT gene, consistent with the known composition of the pPyF9-1 construct. The 2% mismatch is most likely due to errors in the sequence analysis, as the region was not sequenced across both strands or with dITP, which can reduce ambiguities in sequence determination that result from GC compressions. Restriction enzyme analysis of HRA also indicated that at least 2.4 kb of the corresponding 3.0 kb-wild-type *Eco*RI fragment (fragment C) had been retained at the mutant locus. This result, together with sequence analysis of the 5' breakpoint and the knowledge that the entire 2.0-kb E fragment was deleted, established the size limit of the deletion event to no more than 2.77 kb. The precise 3' breakpoint contained within HRA has not been characterized but is known to have occurred within an approximately 600-bp region at the 5' end of the DNA fragment represented by fragment C.

The *TgN737Rpw* Locus Maps to Murine Chromosome 14

Previous work demonstrating that the *TgN737Rpw* locus was contained on either chromosome 14 or 15 led to experiments designed to unequivocally determine the subchromosomal assignment for the *TgN737Rpw* locus. The segregation patterns of a set of *M. musculus/M. spretus* interspecific backcross animals with respect to the retinoblastoma gene *(Rb-1)* were ascertained through DNA analysis by M. Dickinson and Dr. B. Hogan (Vanderbilt University). In separate experiments, these DNAs were

hybridized with HRA and with the T cell receptor α *(Tcrα)* cDNA (Saito *et al.*, 1984). Recombination frequencies were determined by analyzing pairs of markers in interspecies backcross mice: 6/70 mice had chromosomes recombinant for *Tcrα* and *Rb-1* (8.6 cM + 3.35) (centiMorgans + standard deviation), 1/84 mice was recombinant for HRA and *Tcrα* (1.19 cM + 1.18), and 5/66 mice were recombinant for *Rb-1* and HRA (7.58 cM + 3.26). These results established linkage between the *TgN737Rpw* locus and the chromosome 14 marker loci, and placed the *TgN737Rpw* locus between the *Tcrα* and *Rb-1* genes on chromosome 14 (Fig. 5).

In a collaborative effort with Dr. Joseph Nadeau at the Jackson Laboratory, several DNA sequences from the *TgN737Rpw* locus were sent to Dr. Nadeau for the purpose of mapping the *TgN737Rpw* mutation relative to *Disorganization (Ds)*, which is another mutation on chromosome 14 that is sometimes associated with cystic kidneys and polydactyly (Hummel, 1958). Results indicated that the two loci were linked, but that recombination can occur between the two loci. Based on this result, the *Ds* and *TgN737Rpw* sites have been positioned approximately 2.2 cM apart on the chromosome 14 map (J. Nadeau *et al.*, 1993).

Identification and Characterization of the *TgN737Rpw* Gene

Identification of a Gene from the *TgN737Rpw* Locus

To identify a gene whose expression was affected by the insertional mutation, the cloned wild-type genomic sequences corresponding to the mutant locus were systematically screened for the presence of exons by utilizing a cross-species hybridization approach (McGinnis, 1984). No discrete hybridization signal was observed when restriction fragment E, which represents 2.0 kb of the approximately

2.77-kb deletion at the mutant locus, was utilized as a probe in Southern blot analysis with DNA from other species, suggesting that it does not contain exonic sequences. However, a 1.6-kb *Pvu*II-*Sac*I fragment (PS1.6) from within the region corresponding with fragment B (Fig. 6A) hybridized under low stringency conditions with distinct bands in dog and human DNA (Fig. 6B). (The bands in dog and human, although distinct in the autoradiogram, do not appear distinct in Fig. 6B due to underexposure during the photographic process.) A λgt10 cDNA library prepared from whole 14.5 days post coitum (dpc) wild-type embryos was screened with PS1.6, and 17 cDNA clones were isolated. Using λgt10 vector primers, ten clones were amplified by the polymerase chain reaction (PCR) technique, revealing sizes ranging from approximately 1.2 kb to 3.0 kb. The two largest, cDNAs 5 and 6, were each 3.0 kb in size. cDNAs 5, 6, 7, 12, 13 and 18 were subcloned, and restriction map and sequence analyses indicated that all were overlapping versions of the same cDNA, and therefore represented the same gene (Fig. 7). Sequence analysis was performed on cDNAs 5, 6, 12, 13 and 18; however, only cDNA5 (Figs. 7 and 8) was sequenced with the dGTP analogue, dITP (Mills and Kramer, 1979) in its entirety and across both strands.

A 2472-nucleotide open reading frame beginning with the second in-frame ATG predicts a peptide that is 824 amino acids in length. The second ATG was chosen as the putative start site since it is in better agreement with the Kozak consensus (Kozak, 1991). The 824-amino acid peptide potentially represents the entire peptide encoded for by the predominant 3.2-kb mRNA species (Figs. 8, 9A). Sequence comparison of cDNA5 with a corresponding human cDNA (J. Schrick, personal communication) as well as mouse cDNAs 6, 12, and 13 revealed that cDNA5 contains a "silent" point

mutation - thymine in cDNA5 but adenine in the others - at base 2027 that does not change the amino acid sequence.

The Predicted *TgN737Rpw* Protein is a Novel Member of a Family of Genes Containing Tetratricopeptide Repeat (TPR) Domains

The 824-amino acid peptide predicted by cDNA5 is unique and has not been previously described in the Swiss-PROT protein sequence database (Bairoch and Boeckmann, 1993) (Fig. 9A). Analysis of the primary amino acid sequence and a search using the BLAST network service of the NCBI (Altschul *et al.*, 1990) revealed that the protein contains ten copies of an internally repeated degenerate 34-amino acid sequence referred to as the tetratricopeptide repeat (TPR) (Sikorski *et al.*, 1990) (Fig. 9B and C). The TPR was first described in lower eukaryotes as a motif associated with several genes involved in cell cycle control (Sikorski *et al.*, 1990). More recently, TPR-containing proteins in yeast and *Drosophila* have been discovered that function in various cellular processes including protein import, transcription and neurogenesis (Goebl and Yanagida, 1991; Van der Leij *et al.*, 1993). Although most of the TPR proteins characterized to date are located in the nucleus, at least two are associated with the mitochondrial membrane (Goebl and Yanagida, 1991; Van der Leij *et al.*, 1993). Until now, mutations involving genes with TPRs have been identified only in yeast and *Drosophila* (Goebl and Yanagida, 1991; Van der Leij *et al.*, 1993); therefore, the mutation in the *TgN737Rpw* line is the first vertebrate mutation in which the effects of a disrupted TPR gene can be examined.

Determination of Transcriptional Direction

A series of oligonucleotides that had been synthesized during the course of sequencing cDNA5 (Fig. 10) was utilized

to determine which would prime within DNA corresponding to fragments B and C in sequencing experiments. In this way, the correct transcriptional orientation of the genomic sequences was deduced (Fig. 11).

in vivo Expression Analysis of the *TgN737Rpw* Gene

Northern analysis revealed that the size of the predominant wild-type transcript detected by a radiolabeled cDNA5 probe is 3.2 kb (Fig. 12, A and B). cDNA5 is 3024 bp in length (excluding the 19-base poly(A) tail), which is in good agreement with the size of the polyadenylated transcript.

Northern blots with RNA from wild-type and mutant tissues were probed with cDNA5 to determine whether the expression pattern of the gene was altered in the mutant. The *TgN737Rpw* gene is expressed in all tissues and all developmental stages examined to date (Fig. 12, A and B). Most important, the expression of the 3.2-kb transcript is abolished in all mutant tissues examined, including kidney and liver, a finding consistent with the disrupted gene's potential involvement in the renal and hepatic pathology in *TgN737Rpw* homozygotes.

The highest levels of expression relative to other tissues were observed in kidney, brain and testis. In addition, at least two less abundant transcripts, approximately 5.4 and 7.0 kb in size, are consistently observed in wild-type tissues. The 3.2-kb and 7.0-kb wild-type transcripts have never been detected in mutant tissues, but a 5.4-kb transcript is present (Fig. 12A) which, in mutant brain, appears to be expressed at higher than wild-type levels. However, it remains to be determined whether the sequence of the 5.4-kb transcript is identical in mutant and wild-type tissues.

To determine whether transcription is able to access exons on either side of the integrated sequences, Northern

analysis was conducted with two PCR-amplified probes representing sequences from exons 5' (RW117/126) and 3' (RW114/134) (Fig. 10) relative to the transgene integration. Importantly, neither probe was derived from TPR-containing exons, to eliminate the possibility of nonspecific hybridization with other members of the TPR family. In separate experiments, both probes hybridized with a 2.7-kb mutant testis transcript, suggesting that transcription is able to proceed across the integration site (Fig. 12C) (data from the RW117/126 hybridization not shown).

The *TgN737Rpw* gene is expressed as early as embryonic day 8.5, as demonstrated by the isolation of a cDNA (3B-1) from an 8.5-dpc wild-type cDNA library. 3B-1 is approximately 1.0 kb in size, and preliminary sequence analysis has revealed that over 200 bp are homologous with the *TgN737Rpw* gene, but that over 200 bp from the opposite end of the clone are not (B. Yoder, personal communication). This suggests that the clone represents a chimaera produced during the cDNA library preparation in which cDNA sequences representing another gene fused with the *TgN737Rpw* cDNA sequences. It is also possible that 3B-1 may represent an alternatively-spliced form of the *TgN737Rpw* gene.

The *TgN737Rpw* Gene is Expressed in Proximal and Distal Tubule Cell Lines

The *TgN737Rpw* gene contains 10 copies of the TPR domain, a motif first identified in yeast genes that are often involved in mitotic control (Sikorski *et al.*, 1990). Therefore, expression studies were designed to determine whether *TgN737Rpw* expression might differ in proliferating versus quiescent cells. RNA was obtained from proliferating and quiescent (confluent) cell lines representing different segments of the nephron: Madin Darby canine (MDCK, distal

tubule) (Taub and Saier, 1979), porcine (LLC-PK_1, proximal tubule) (Hull *et al.*, 1976), American opossum (OK, proximal tubule) (Nakai *et al.*, 1988) and African green monkey (COS-7, "fibroblast-like"; confluent only tested) (Gluzman, 1981) kidney cells. RNA from a mouse keratinocyte line as well as embyronic stem cells plus mouse fibroblast ("feeder") cells were also included to investigate the cell-specificity of *TgN737Rpw* expression. A predominant 2.8-kb transcript was observed in each of these cell lines with the exception of COS-7 cells, and no significant qualitative or quantitative differences were detected in the levels of expression between quiescent (confluent) and proliferating cells (Fig. 12D). Interestingly, the smaller size of the predominant transcript in these species is the same as in humans (data not shown). In humans this is due to a deletion in the 3' untranslated region (J. Schrick, personal communication), which may also be the case in the other species.

Alternative Splicing as a Possible Regulatory Mechanism in *TgN737Rpw* Gene Expression

Evidence that regulation of the *TgN737Rpw* gene may involve alternative splicing emerged when one (cDNA12) of five 14.5 dpc murine cDNAs was found to lack a 111-bp region from the 3' end of the ORF (Figs. 7, 8). Subsequent sequence and Southern hybridization analysis of the corresponding genomic DNA, which had been isolated in a cosmid clone (c2RBA2) (Fig. 4) confirmed that the deleted region is located 3' to the transgene integration site, several kb downstream from its 5' neighboring exon (Fig. 11). An in-frame ORF spanning the region is maintained in cDNAs with or without the 111 bp, suggesting that the region is an alternatively-spliced exon. The 111-bp insert is also present in each of at least four human cDNAs corresponding with the *TgN737Rpw* gene that were obtained from an adult liver

cDNA library (J. Schrick, personal communication). Interestingly, sequence of the corresponding genomic region from several human clones revealed that the last two nucleotides of the putative 5' human intron follow the consensus AG, while the first two nucleotides of the 3' intron are GC instead of the consensus GT (J. Schrick, personal communication).

To investigate the temporal and spatial expression of the 111-bp putative exon, a probe (RW114/134) representing 104 of the 111 bp was utilized in Northern analysis. The probe hybridized with RNA from mutant and wild-type tissues, but to a lesser degree in the higher MW transcripts (Fig. 12C). This demonstrated that the region is expressed, and the hybridization intensities suggested that it may be differentially represented among the transcript species. Alternatively, the lack of detectable hybridization signals from the higher MW transcripts in every tissue may be related to the transcripts' lesser abundance and the small probe size.

Finally, alternative splicing of the *TgN737Rpw* gene may result in the formation of a unique testis-specific transcript. In wild-type testis, a 2.8-kb transcript is observed; however, a slightly smaller mutant species is observed, at approximately 2.7 kb (Fig. 12A). The 104-bp (RW114/134) probe from the putative alternatively-spliced exon hybridizes with the mutant transcript (Fig. 12C), demonstrating that the decrease in transcript size is not due to the absence of the 111-bp exon.

Isolation and Characterization of the 5' Genomic Region of the *TgN737Rpw* Gene

The putative translational start site within cDNA5 corresponds with the second ATG (Fig. 8), which is in closer agreement with the Kozak consensus sequence (Kozak, 1991).

The ORF in cDNA5 continues 71 bp from the putative translational start site to the 5' end of the clone. To isolate and characterize the *TgN737Rpw* promoter, PCR-amplified fragments (RW 117/126 and RW117/127) from the extreme 5' end of cDNA5 (Fig. 10) were used to isolate two genomic clones (Fig. 13) which span approximately 22 kb and include the first four exons and 5.0-6.0 kb of sequence upstream from the start of cDNA5. Exonic regions, and not the actual sizes of the exons, were determined (Fig. 13) by sequence analysis and hybridization of cDNA5 with the restriction-digested genomic clones. Analysis of these clones demonstrated that the putative translational start site is contained in the second exon of cDNA5 (Fig. 13); hence, if cDNA5 is a full-length cDNA, the first exon is untranslated. Preliminary sequence analysis of over 600 bp preceding the first exon in cDNA5 suggests that the ORF may extend beyond the 5' end of the cDNA, and has failed to identify Sp1-binding sites, TATAA or CAAT boxes - features that often characterize the promoter regions of genes transcribed by RNA polymerase II (Mitchell and Tjian, 1989). However, approximately 475 bp 5' to the region corresponding with the start of cDNA5 is an AP-1-like consensus sequence, TGCGTCA, which is a recognition site for binding transcription factor AP-1 ("activator protein-1", which has been shown to consist of various dimer combinations of the proto-oncogenes Fos and/or Jun) and the CREB (cAMP responsive element binding) protein - oncogenes which can stimulate gene expression (Gutman and Wasylyk, 1991).

In addition, a 200 bp region located 5' to the first exon has a G+C content of 67%, a characteristic of TATAA-box-deficient G+C-rich promoters (Ackerman *et al.*, 1993). A nucleotide stretch here exhibits a 40% identity with a 48-bp region that contains the minimum promoter element of the murine adenosine deaminase (ADA) gene (Figs. 13 and 14A).

Another G+C-rich area was identified in the first *TgN737Rpw* intron, in which 230 bp of sequence contain 73% G+C. Interestingly, a 48-bp stretch within this G+C-rich region shares 72% identity with a 47-bp region contained within an enhancer element located in the first intron of the human ADA gene (Aronow *et al.*, 1992) (Figs. 13 and 14B). This core enhancer region is necessary for cell-type-specific expression of the ADA gene in human lymphoid cell lines, possibly due to its ability to bind nuclear activating factors (Aronow *et al.*, 1992). Nineteen of the twenty-eight (68%) nucleotides that comprise the ADA gene's core enhancer element are present in the *TgN737Rpw* intron, suggesting that this region may be involved in the regulation of the *TgN737Rpw* gene.

Pulsed-Field Gel Electrophoresis Analysis of the *TgN737Rpw* Locus

All of the molecular evidence up to this point had indicated that no rearrangements had occurred at the mutant locus, other than the insertion of cointegrating DNA and the 2.77-kb deletion. Nevertheless, pulsed-field gel electrophoresis (PFGE) (Schwartz and Cantor, 1984) was performed to detect any relatively large but cytogenetically undetectable rearrangements that may have occurred at the *TgN737Rpw* locus. HR0.2 (Figs. 3B and 15) and later cDNA5 (data not shown) were used as probes on high molecular weight DNA cut with several different rare-cutting enzymes. Exons contained within cDNA5 span at least 50.0 kb of the genomic sequence. This was demonstrated by hybridization of the radiolabeled cDNA5 with restriction fragments generated from wild-type lambda (20-3 and E'4) and cosmid (c2RBA2) clones (Fig. 4), and with regions of DNA spanning over 21.0 kb in the 5' genomic clones II(1) and II(2) (Fig. 13). Therefore, the long genomic distance spanned by cDNA5 made it a particularly useful probe. No hybridization signal

was observed in the DNA digested with *SalI*, probably because the enzyme cut several times within the gene, creating fragments too small to be included in the gel. The hybridization pattern was the same with each probe, and revealed no restriction fragment length polymorphisms between the mutant and wild-type loci (Fig. 15), suggesting that no major rearrangements had occurred within the large restriction fragments detected in this analysis.

Investigation of Possible Allelism with Other PKD Mouse Models

Several mouse models exist that display a recessive form of PKD. The phenotype in these models does not match that of (FVB/N)-*TgN737Rpw* homozygotes, but that could be a result of the different genetic backgrounds upon which the mutations arose. Therefore, the possibility of allelism and/or interaction of the *TgN737Rpw* gene with the mutant genes in the same pathway that causes PKD in other mouse models was considered.

Complementation Testing to Determine Allelism Between the *TgN737Rpw* and *bpk* Mutations

Several genetic mouse models of recessive PKD exist that are not allelic with each other (*cpk*, Russell and McFarland, 1977; *pcy*, Takahashi *et al.*, 1991; *jck*, Atala *et al.*, 1993), but unlike ARPKD patients, none consistently exhibits a hepatic lesion. Liver abnormalities are a constant finding in only two PKD mouse models: the (BALB/c)-*bpk* mouse - recently proposed to be a new mouse model for ARPKD (Nauta *et al.*, 1993) - and in *TgN737Rpw* homozygotes. Additional similarities between the two models include the neonatal development of renal cysts, and the initial localization of renal lesions to the proximal tubules. The *TgN737Rpw* gene

has been mapped to the proximal region of chromosome 14, but because the *bpk* mutation arose spontaneously, no markers for the locus have yet been identified. The phenotypic parallels raised the possibility that the mutations were allelic, and experiments were designed to test this hypothesis.

To test whether the *bpk* and *TgN737Rpw* mutations were allelic, a complementation experiment was performed. This analysis was accomplished by a mating scheme designed to generate compound heterozygotes - mice heterozygous for the two mutant alleles being tested. Obligate (BALB/c)-*bpk*/+ males (identified by progeny testing) were mated with *TgN737Rpw* carrier females, and offspring were genotyped for the presence of the mutant *TgN737Rpw* allele by Southern analysis with HRA. To determine whether offspring hemizygous for the presence of the transgene also carried the mutation *bpk* allele, they were then crossed to known or suspected *bpk*/+ mice. Progeny were genotyped with respect to the *TgN737Rpw* locus and observed for polycystic kidneys (Fig. 16). Litters containing nontransgenic mice with polycystic kidneys (pathological analysis conducted by Dr. Virginia Godfrey) confirmed the *bpk*/+ status of both parents, and in this way, one male (mouse #37) proven compound heterozygote (*bpk*/+; *TgN737Rpw*/+) has been identified to date. It is interesting to note that two of the offspring with PKD were also heterozygous for the mutant TgN737Rpw locus, suggesting that the *bpk* and *TgN737Rpw* genes are not closely linked, and may reside on different chromosomes. However, larger numbers of animals must be evaluated for this to be shown more conclusively. Mouse #37 has produced at least three litters, has exhibited no signs of PKD, does not have polydactyly and is alive at 18+ months of age. Necropsy results will ultimately reveal whether this compound heterozygote has polycystic kidneys or other

phenotypic abnormalities that would suggest complementation between the two mutant alleles. It will also be important to generate and examine several compound heterozygotes, since the penetrance of the phenotype might vary due to the combination of the disparate genetic backgrounds upon which the mutations were originally generated. However, these preliminary results suggest noncomplementation of the *TgN737Rpw* and *bpk* mutations; hence, the genes are probably not allelic.

Northern Analysis of *bpk/bpk* and *cpk/cpk* Tissues with the *TgN737Rpw* cDNA

Although complementation analysis had indicated that the *bpk* and *TgN737Rpw* mutations were probably not allelic, the possibility was considered that the two genes may be involved in a common molecular pathway that leads to PKD. Northern analysis was conducted in which cDNA5 was used as a probe on kidney and liver poly(A)$^+$RNA from *bpk/bpk* mice. A disruption in the quantitative or qualitative expression patterns of the *TgN737Rpw* gene could indicate the involvement of both genes in such a common pathway, in which the *TgN737Rpw* gene is under the regulatory control of the *bpk* gene. No differences in the expression of the *TgN737Rpw* gene were detected in *bpk/bpk* kidneys or liver (Fig. 17A), suggesting that the *TgN737Rpw* gene is not subject to regulatory control by the *bpk* gene. These results were also consistent with the results from the complementation experiment.

The *cpk* mutation (Russell and McFarland, 1977) arose spontaneously in a line of C57BL/6J mice, causing a form of autosomal recessive PKD that is usually fatal by 3 weeks of age. Liver lesions are associated with the *cpk* mutation, although they are an inconstant, genetic background-dependent finding (Crocker *et al.*, 1987). Shortly before it

was reported that the *cpk* gene had been localized to a region on mouse chromosome 12 (Davisson *et al.*, 1991), heterozygotes from this line had been obtained for the purpose of complementation testing with the *TgN737Rpw* mutation. To address the possibility that the genes, though not allelic, might be involved in the same pathway that results in PKD, poly(A)$^+$ RNA from brain and kidneys was obtained from a known *cpk/cpk* animal and a phenotypically wild-type (*cpk/+* or *+/+*) mouse. Hybridization of the Northern blot with cDNA5 revealed no detectable quantitative or qualitative alterations in the expression of the *TgN737Rpw* gene in *cpk/cpk* tissues (Fig. 17B), suggesting that either the two genes are not involved in the same pathway of cystogenesis, or if they are, that the *TgN737Rpw* gene is not subject to upstream regulation by the defective *cpk* gene.

VI. DISCUSSION AND CONCLUSIONS

Gaining Access to the *TgN737Rpw* Gene via Insertional Mutagenesis

This study has described how molecular access to a candidate gene for PKD in the Tg737 line of mice was achieved through the insertional mutagenesis technique. The mutant locus was successfully identified and cloned, despite previous reports of rearrangements associated with this method that include deletions, duplications and translocations (Gridley *et al.*, 1987). Although insertional mutagenesis occasionally introduces complex rearrangements during the integration event (Singh *et al.*, 1991), and it has been reported that the microinjection method is especially prone to cause such changes (Gridley *et al.*, 1987), the *TgN737Rpw* mouse represents at least the 5th microinjection-generated mutant from which disrupted genes have been identified (limb deformity, Woychik *et al.*, 1990; legless, Singh *et al.*, 1991; Hβ58, Lee *et al.*, 1992; microphthalmia, Hodgkinson *et al.*, 1993).

Lack of Evidence for Structural Rearrangement of the *TgN737Rpw* Gene Beyond the Insertion Site

The initial molecular characterization of the *TgN737Rpw* locus revealed that at least two rearrangements had occurred as a result of transgene integration: a deletion of not more than 2.77 kb at the insertion site, and the insertion of cointegrated mouse DNA approximately 21.0 kb in length. To investigate the possibility that other rearrangements had occurred, the analysis of the structural gene was extended with PFGE and additional conventional Southern analysis. cDNA5 was utilized as a probe in PFGE and Southern analyses

because it includes exons spanning at least 50.0 kb of the structural gene, making it particularly useful in searching for rearrangements over a long distance. The PFGE data are consistent with previous cytogenetic analysis of the mutant chromosome 14 as well as Southern analysis of the mutant locus and indicate that, apart, from the 2.77-kb deletion and the cointegrated DNA, there is no evidence of major rearrangements within several hundred kilobases of the *TgN737Rpw* locus. Hence, it is unlikely that more than one gene was structurally affected by the integration event in these mice.

The Mutant *TgN737Rpw* Gene Remains Transcriptionally Active

Another potential complication in transgenesis is that the introduced sequences can induce widespread methylation and/or chromatin changes over a large distance (Jähner and Jaenisch, 1985; Breindl *et al.*, 1984), resulting in the inactivation of more than one gene. This, of course, complicates any attempt to correlate a single gene at the mutant locus with the phenotype in the mutant line. Although the predominant 3.2-kb transcript is undetectable by Northern analysis, the presence of the testis-specific and the minor 5.4-kb transcripts in mutant tissues suggested at first that the *TgN737Rpw* gene had not been inactivated by methylation. However, the possibility was considered that the less abundant transcripts detected with the cDNA5 probe actually represented members of a homologous gene family, such as the TPR family, and not the *TgN737Rpw* gene. The altered size of the mutant testis transcript strongly suggests that it represents the *TgN737Rpw* gene, and the absence of the 7.0-kb minor species in mutant tissues suggests that it does not represent a member of a related gene family.

Moreover, it has been shown that sequences from nonTPR-containing exons from both sides of the transgene integration site are included in a mutant 2.7-kb testis transcript. Because the sequence of interTPR-regions is not conserved between members of the TPR family (Honoré *et al.*, 1992; Goebl and Yanagida, 1991), it is unlikely that the hybridization observed with these probes in Northern blot analysis represents homologous TPR members. This result also suggests that the gene's extreme ends are accessible for post-transcriptional processing, although proof that both probes hybridize to the identical transcript awaits the cloning and sequencing of the mutant testis transcript(s). The nature of the 5.4-kb-transcript that continues to be expressed in mutant tissues remains to be determined. Nevertheless, the presence of the size-altered testis transcript, and hybridization of mutant mRNA with probes from opposite ends of the gene, provide good evidence that the *TgN737Rpw* gene is transcriptionally active in the mutant. Thus, the transgene integration specifically ablates the production of the 3.2-kb transcript and does not induce a widespread methylation effect that could inactivate neighboring genes.

Mechanism of Disruption of the *TgN737Rpw* Gene

The physical disruption of the *TgN737Rpw* gene prevents, by an unknown mechanism, the formation of the predominant wild-type transcript in all mutant tissues that have been examined. The failure of the 3.2-kb transcript to be formed is apparently not (at least in testis) due to the inability of transcription to access exons from both sides of the integrated sequences. Although no exons appear to have been deleted from the structural gene as a result of the transgene integration, the altered size of the testis mRNA and the absence of the 3.2-kb species demonstrate that the

transcript is spliced aberrantly in the mutant, and may be partially or completely nonfunctional. The possibility has been considered that, since a 14.5-dpc cDNA was used in experiments designed to detect exons in the wild-type region corresponding with the deletion at the mutant locus, an alternatively-spliced exon(s) not highly expressed at 14.5 dpc may reside within the deleted region of genomic DNA. No exon corresponding with the deleted region was found in at least 4 cDNA clones obtained from adult human liver (R.P. Woychik and J.J. Schrick, personal communication).

An alternative explanation for the failure to form the predominant transcript is that a conformational change in the secondary structure of the mutated precursor *TgN737Rpw* RNA might result in the utilization of cryptic splice sites, preventing the generation of the correct transcripts except, perhaps, in the case of the 5.4-kb species.

It is also possible that the deleted region contains sequences involved in proper splice-site selection for the production of the 3.2-kb transcript. It has been shown that the specificity and efficiency of splicing involves intronic interactions (Neel *et al.*, 1993); therefore the substantial increase in the size of the intronic region as a result of the inserted sequences might be responsible for the absence of the 3.2-kb transcript in mutant tissues.

The possibility has been considered that the 21.0 kb of cointegrated DNA contains a gene(s) that is expressed from the mutant locus and contributes to the *TgN737Rpw* phenotype. While formal resolution of this question awaits rescue of the mutant phenotype with the wild-type *TgN737Rpw* gene, it is improbable that the mutant phenotype is caused by the expression of a gene from within the translocated DNA. Because heterozygous *TgN737Rpw* mice do not display abnormalities, the argument for an effect due to the presence of a gene within the cointegrated

fragment would have to be based in the context of a genetic dosage effect. For example, if one assumes that expression of an exogenous gene from within the cointegrated DNA produces a mutant protein responsible for the phenotype, it follows that because transgenic hemizygotes are phenotypically indistinguishable from their wild-type littermates, the abnormal phenotype emerges only when two, but not one, aberrant genes are expressed in addition to the two endogenous copies. Alternatively, if the cointegration contains an intact gene that is expressed from the *TgN737Rpw* locus in a wild-type manner, the mutant phenotype must arise from the expression of four, but not three, normal copies. This author is unaware of any example in which partial quadrosomy, but not partial trisomy, evokes a mutant trait. Therefore, it is unlikely that the cointegrated sequences contain a gene(s) responsible for the mutant phenotype.

Finally, the possibility remains that the insertional mutation disrupts the expression of a gene that is transcribed from the strand opposite that of the *TgN737Rpw* gene. Clearly, definitive proof that the *TgN737Rpw* gene is responsible for the mutant phenotype in the Tg737 line of mice awaits phenotype rescue experiments, which are in progress (R.P. Woychik and B. Yoder, personal communication).

cDNA5 is Likely a Full-Length Clone

Although there is no conclusive evidence that cDNA5 represents a full-length cDNA, several observations suggest that this is the case. First, the 3024-bp size of cDNA5 is consistent with the 3.2-kb size of the mRNA which probably confirms a 200-300-base poly(A) tail. In addition, a second cDNA (cDNA 6) was isolated that is identical to cDNA5 in

length and composition except for a single-nucleotide "silent" point mutation corresponding with base 2027 of cDNA5, a shorter poly(A)-tail and a deletion of 4 nucleotides at the extreme 5' end. The fact that both cDNA5 and cDNA6 end within 4 nucleotides of each other suggests that the discrepancy at the 5' end may be due to interference of the cap site with reverse transcriptase during the first strand synthesis step in the cDNA library preparation. Alternatively, there may be a strong stop for reverse transcriptase at that point. Sequence analysis of human DNA corresponding with the region 5' to the mouse cDNA5 has shown that the ORF is maintained for at least 185 additional amino acids (J.J. Schrick, personal communication). The size of the human transcript (2.8 kb) observed in Northern analysis is not consistent with the addition of 555 or more nucleotides, and because higher MW transcripts have not been observed in Northern analysis of the human *TgN737Rpw* gene, it is unlikely that the extended ORF represents translated sequences. The maintenance of an ORF through the 5' end of the cDNA may simply be a function of the G+C-richness of the region and does not necessarily indicate that the region is translated. It is also possible that the gene utilizes alternate start sites.

It should be noted here that specific aim #8, which was the characterization of the *TgN737Rpw* gene's intron/exon structure, was not completed due to the large size of the gene. The last 824 nucleotides of cDNA5, which represent less than 28% of the cDNA are represented in more than 25.0 kb of genomic DNA, and at least 25 exons have been identified in the corresponding human gene (J.J. Schrick, personal communication). Therefore, the complete determination of the gene's intron/exon structure awaits future analysis.

Identification of Motifs that May be Involved in the Regulatory Control of *TgN737Rpw* Transcription

Similarities of genomic DNA located 5' to cDNA5 with the 5' end of the murine ADA gene (Ackerman *et al.*, 1993) suggest that G+C-rich sequences immediately 5' to the first cDNA5 exon may represent a promoter-like region. G+C-rich leader sequences are typical of many critical regulatory proteins, including oncoproteins, growth factors, and transcription factors (Kozak, 1991). In addition, as may be the case in the *TgN737Rpw* gene, critical regulatory genes are often "encumbered" with more than one potential AUG transcriptional start site, which is thought to effect a protective regulatory control at the translational level for proteins that, if overexpressed, might be harmful to the cell (Kozak, 1991). If the *TgN737Rpw* gene is involved in cell cycle control, deregulated expression might indeed have deleterious consequences. The homology of regions in the 5' portion of the *TgN737Rpw* gene with the minimal promoter element for the murine ADA gene and a core enhancer element from the human ADA gene's first intron suggests that these regions may participate in complex transcriptional control of the *TgN737Rpw* gene.

The presence of an AP-1-like binding site in the genomic region 5' to the start of cDNA5 suggests that transcriptional activation of the *TgN737Rpw* gene may be regulated in part by direct communication with the proto-oncogenes Jun and Fos as well as cAMP-response element binding proteins, by virtue of the proteins' potential ability to bind at the site. It has been shown that both Fos and Jun can activate transcription in response to mitogenic stimuli (Curran and Franza, 1988; Ransone and Verma, 1990). Similarly, CREB can interact with the c-AMP responsive element to stimulate transcription in response to elevated intracellular levels of cAMP and so is involved in signal transduction "cross-talk"

(Meek and Street, 1992). Therefore it is speculative but intriguing to consider the possibility that *TgN737Rpw* transcription is induced by mitogenic stimuli, a situation which would be consistent with the gene playing a role in cell cycle control. Finally, the putative transcriptional start site is located in the second exon of cDNA5, suggesting that the first exon is nontranslated, a feature that sometimes facilitates transcription (Kozak, 1991). Future analysis will be required to definitively determine the *TgN737Rpw* gene's promoter region and its regulatory control.

Potential Roles for the *TgN737Rpw* Gene

Expression of the *TgN737Rpw* Gene: Related to Cellular Proliferation?

TgN737Rpw expression appears to be ubiquitous *in vivo*, appearing in every tissue and during every developmental stage examined. In addition, it is expressed by both distal and proximal tubular cell lines, which represent a heterogeneous blend of functionally diverse, segment-specific cells (Kaissling and Kriz, 1992). These results suggest that *TgN737Rpw* expression is restricted to neither a single cell type nor a specific nephron segment in the kidney. If the *TgN737Rpw* gene is normally involved in the maintenance of differentiation or the inhibition of proliferation, its expression in quiescent versus proliferating cells might be expected to differ. However, no significant difference in expression relative to proliferative status was detected by Northern analysis. Only COS-7 cells fail to express the gene, and it is intriguing to speculate that this may be related to the fact that the cell line has been stably transformed with the SV40 virus (Gluzman, 1981), or because COS-7 cells are "fibroblast-like" (Gluzman, 1992), in contrast with the epithelial nature of the other cell types.

The proliferative lesions observed in *TgN737Rpw* livers and kidneys, and the identification of the *TgN737Rpw* gene as a member of a family of proteins often involved in cell cycle control, are consistent with the possibility that the *TgN737Rpw* gene is normally involved in the regulation of cellular proliferation. If no difference in the expression pattern of the *TgN737Rpw* gene exists between proliferating and quiescent cells, it may be that other factors which are expressed differentially interact with the *TgN737Rpw* gene in a cooperative manner that is dependent on the stage of cellular growth/differentiation.

Pleiotropic Effects of the *TgN737Rpw* Mutation and Implications for Gene Function

The pleiotropy observed in *TgN737Rpw* mutants suggests that if the disruption of a single gene is the cause, the gene participates in a complex pathway, and/or has multiple functions. A mechanism such as alternative splicing, which may be responsible for the generation of the lower MW *TgN737Rpw* testis transcript, and may occur during day 14.5 of embryogenesis, has the potential to create several unique proteins with diverse roles. The presence of a nonconsensus splice site in the region of human DNA corresponding with the putative 111-bp exon is consistent with reports of nonconsensus mRNA splice sites, in which GC in place of GT at positions 1 and 2 of the intron was the most common variant at that site (for review, see Jackson, 1991). Furthermore, it has been shown that GC is the only 5' splice site consensus other than GT that will be accurately cleaved to an active form, although much less efficiently than the prototypical GT (Aebi *et al.*, 1987). In fact, a variant GC at this position in the αA-crystallin gene is involved in alternative splicing in the mouse (King and Piatigorsky, 1983), and it has been shown that donor site mutations can

result in the skipping of the immediate upstream exon (Nasim *et al.*, 1990; Tacke and Goridis, 1991; Talerico and Berget, 1990). Therefore, it is likely that a nonconsensus splice site could be advantageous for the purpose of generating lower levels of an alternatively-spliced product.

At present it is unknown whether a similar variant 5' donor site exists in the mouse *TgN737Rpw* gene. The fact that no human cDNA was identified from which the 111-bp insert was missing suggests that the region may not participate in alternate splicing in adult liver and/or other human tissues, or that alternative splicing occurs at a different developmental stage. Alternatively, because only one of five mouse cDNAs that were examined lacked the insert, it is also possible that such transcripts are fairly rare in humans, and that additional screening will identify human cDNAs lacking the region.

Alternative splicing would theoretically expand the functional repertoire of the *TgN737Rpw* gene, which would be consistent with the pleiotropic effects observed in mutant animals. Such pleiotropy might also be explained in other ways. For instance, the *TgN737Rpw* gene product, while seemingly ubiquitously expressed in a consistent pattern from the embryonic stage into adulthood, may in fact interact with tissue- and developmentally-specific gene products and so be responsible, directly and/or indirectly, for varied functions throughout development. Additional levels of functional complexity could also be achieved through cell-specific post-translational modifications of the *TgN737Rpw* protein.

Although the *TgN737Rpw* gene is expressed in every tissue studied *in vivo*, an abnormal phenotype is not observed in every organ. Interestingly, the pathology in the kidneys and liver of FVB/N mutant animals - tissues most severely affected by the mutation - is proliferative in nature,

although, in the adult, these tissues normally have very low mitotic rates (Baserga, 1981; MacDonald, 1961). Conversely, several organs in which the cellular proliferation rate is normally rapid (small intestine, lung, spleen, salivary gland) do not display a consistently abnormal phenotype that has been characterized as a primary defect attributable to the *TgN737Rpw* mutation. Perhaps the *TgN737Rpw* gene product is required for growth control in a subset of cells normally associated with low proliferative rates, whereas in tissues with more rapid turnover such inhibition is unnecessary or is accomplished through another pathway. The lack of a detectable abnormality in the brains of mutant animals could thus be explained if the gene is normally expressed not in the nonreplicating neuronal cells but in the numerous, replicating glial cells that comprise the bulk of the brain tissue (Langley *et al.*, 1974).

Pathogenesis in PKD

All forms of PKD are associated with a localized proliferation of the epithelium lining the cysts (Avner *et al.*, 1990). Perhaps it is therefore not surprising that the abnormal activation of several oncogenes, including c-*myc*, is known to be associated with the development of renal cysts in at least 3 mouse models of PKD (Cowley *et al.*, 1991; Trudel *et al.*, 1991; Kelley *et al.*, 1991). Moreover, it has been shown that the von Hippel-Lindau gene, which when mutated sometimes gives rise to cystic kidneys, is a tumor suppressor gene (Latif *et al.*, 1993). Still another gene associated with renal cysts - the tuberous sclerosis (*TSC2*) gene - has been proposed to also act as a tumor suppressor (European Chromosome 16 Tuberous Sclerosis Consortium, 1993). It is particularly noteworthy that the inactivation of *bcl-2*, a proto-oncogene that promotes cell growth by inhibiting apoptotis, also results in PKD (Veis *et al.*, 1993). The

discovery that *bcl-2* can inhibit the apoptotic, but not the proliferative, effects of c-*myc* (Fanidi *et al.*, 1992) has led to speculation that the PKD observed in *bcl-2* knockout mice may be due to the deregulation of c-*myc* expression (Veis *et al.*, 1993), as is the case in mice transgenic for a construct that expresses the c-*myc* gene (Trudel *et al.*, 1991). While the mechanism of the deregulated epithelial proliferation in PKD is currently unknown, the fact that c-*myc* and *bcl-2* both function to modify normal apoptosis (Bissonnette *et al.*, 1992; Fanidi *et al.*, 1992; Wyllie, 1993) suggests that either an activation of the cellular proliferative response or a failure of apoptosis could be responsible for the epithelial hyperplasia in PKD. Alternatively, the rate of proliferation in renal tissues might exceed the apoptotic rate, leading to net proliferation. The fact that the interrupted gene at the mutant locus in the Tg737 line contains a motif common to several cell cycle control genes could suggest that this gene normally functions to regulate the cell cycle in the epithelial component of the kidney.

The *TgN737Rpw* Mouse as a Unique Model for PKD

Nonallelism of the *TgN737Rpw* Mutation with Other PKD Mutations in Mouse

The pleiotropic phenotype of the *TgN737Rpw* mouse is different from that observed in any of the mouse models for PKD described to date, suggesting that the *TgN737Rpw* mutation is not allelic with existing mutations. The chromosomal assignment of the *TgN737Rpw* gene to chromosome 14 rules out allelism with *cpk*, and pcy, which map to chromosomes 12 and 9, respectively (Davisson *et al.*, 1991; Takahashi *et al.*, 1991), and also rules out all allelism with *cph* (formerly *jpk*), a mouse mutation mapping to chromosome 15 that is characterized by congenital

hydronephrosis (Horton *et al.*, 1988). Northern analysis and complementation testing with *bpk* mice have demonstrated that although the hepatic and renal lesions in *bpk* and *TgN737Rpw* are strikingly similar, the two loci are probably not allelic. Northern analysis also indicated that the *TgN737Rpw* gene is probably not under regulatory control by the *cpk* or *bpk* genes. Therefore, it appears likely that the *TgN737Rpw* mutation is a novel PKD mutation in the mouse. However, until molecular markers for the other PKD mutations become available, it will not be possible to directly test whether the disrupted genes are subject to upstream regulatory control by the *TgN737Rpw* gene. The existence of numerous, apparently nonallelic mouse models of PKD suggests that PKD is a common pathological consequence of mutations in any one of several physiological pathways. Isolation of the genes responsible for these mutations will undoubtedly provide considerable insight into the complex processes involved in PKD.

The *TgN737Rpw* Mouse as a Model for Human PKD

The PKD in *TgN737Rpw* mutants is similar in many respects to ARPKD. The recessive mode of inheritance, the morphology, localization and progression of the renal cysts, and the presence of noncystic hepatic lesions in these mice parallel some of the unique features of ARPKD that distinguish it from ADPKD. The *TgN737Rpw* mouse may therefore be a model for ARPKD. However, it should be noted that considerable phenotypic variability exists, even among humans with ARPKD, leading Blyth and Ockenden (1971) to propose that ARPKD fell into four groups, each probably due to a mutation in a different gene. This concept has been largely rejected since phenotypic heterogeneity within families has become apparent (Kaplan *et al.*, 1988; Avner, 1994), but the fact remains that considerable phenotypic

overlap exists among patients with ARPKD and pediatric onset ADPKD (Avner, 1994), including the occasional presence of congenital hepatic fibrosis in several patients with chromosome 16-linked ADPKD (Cobben *et al.*, 1990). Moreover, several other heritable syndromes are associated with renal cysts, hepatic lesions and polydactyly, although this triad of features is not often seen consistently in any one disorder. One recessive disorder in which cystic kidneys and polydactyly are commonly observed is Meckel syndrome, and in recent years it has been proposed that a proliferative or fibrotic liver lesion is an additional hallmark feature (Salonen, 1984). Considering the influences of genetic background, and the possibility of phenotypic variances between the two species, it is appropriate to consider each human cystic disease individually as a potential homologue of the PKD in *TgN737Rpw* mutants.

The *TgN737Rpw* Mouse: First Animal Model for PKD in which the Mutant Locus was Accessed at the Molecular Level

Unlike the other genetic models for PKD, the unique feature of the *TgN737Rpw* mutation is the fact that the mutant locus was tagged with a molecular marker, which has allowed the isolation and characterization of an excellent candidate gene whose expression is directly affected by the inserted transgene at the mutant locus. The structural information gathered to date suggests that the mutant phenotype in transgenic homozygotes is due to the exclusive disruption of the *TgN737Rpw* gene, and the presence of alternative transcripts demonstrates that the gene is subject to complex regulatory control and that it may participate in multiple and diverse functions. Further characterization of the *TgN737Rpw* gene will undoubtedly provide more insight

into its regulation and potential role in the development of PKD in *TgN737Rpw* mice.

Finally, the *TgN737Rpw* gene may not only provide new insights into the cellular pathophysiology of the progressive cystic enlargement and the renal and/or hepatobiliary epithelial proliferation found in renal cystic disease, but may guide the molecular genetic search for homologous candidate genes in human PKD. Utilizing the mouse cDNA as a probe, three genomic human DNA clones were isolated for future analysis. In addition, using the same cDNA probe, most of the corresponding human gene has been cloned (J. Schrick, manuscript in preparation) and mapped to human chromosome 13 (J. Kornberg, personal communication). The isolation of this human gene will permit a thorough molecular study which should ultimately determine whether this gene is directly associated with any of the various forms of PKD in humans.

LIST OF REFERENCES

Ackerman, S.L., Minden, A.G. and Yeung, C.-Y. (1993) The minimal self-sufficient element in a murine G+C-rich promoter is a large element with imperfect dyad symmetry. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 11865-11869.

Aebi, M., Hornig, H. and Weissmann, C. (1987) 5' cleavage site in eukaryotic pre-mRNA splicing is determined by the overall 5' splice region, not by the conserved GU. *Cell* **50**, 237-246.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. and Lipman, D.J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403-410.

Antignac, C., Arduy, C.H., Beckmann, J.S., Benessy, F., Gros, F., Medhioub, M., Hildebrandt, F., Dufier, J.-L., Kleinknecht, C., Broyer, M., Weissenbach, J., Habib, R. and Cohen, D. (1993) A gene for familial juvenile nephronophthisis (recessive medullary cystic kidney disease) maps to chromosome 2p. *Nature Genet.* **3**, 342-345.

Aronow, B.J., Silbiger, R.N., Dusing, M.R., Stock, J.L., Yager, K.L., Potter, S.S., Hutton, J.J. and Wiginton, D.A. (1992) Functional analysis of the human adenosine deaminase gene thymic regulatory region and its ability to generate position-independent transgene expression. *Mol. Cell. Biol.* **12**, 4170-4185.

Atala, A., Freeman, M.R., Mandell, J. and Beier, D.R. (1993) Juvenile cystic kidneys (jck): A new mouse mutation which causes polycystic kidneys. *Kidney Int.* **43**, 1081-1085.

Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (1991) in "Current Protocols in Molecular Biology", John Wiley and Sons, NY.

Aviv, H. and Leder, P. (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. *Proc. Natl. Acad. Sci. U.S.A.* **69**, 1408-1412.

Avner, E.D. (1994) Polycystic kidney disease. In "Pediatric Nephrology" (A. Drukker and A. Gruskin, Eds.) in the series *Pediatr. Adolesc. Med.*, (D. Branski, Ed.), Vol. 5, pp. 467-490, Basel, Karger.

Avner, E.D., Sweeney, W.E. and Nelson, W.J. (1992) Abnormal sodium pump distribution during renal tubulogenesis in congenital murine polycystic kidney disease. *Proc. Natl. Acad. Sci. U.S.A.* **89**, 7447-7451.

Avner, E.D. and Sweeney, W.E. (1992) Epidermal growth factor receptor (EGFR), but not Na^+-K^+-ATPase, is mislocated to apical cell surfaces of collecting tubule cyts in human autosomal recessive polycystic kidney disease (ARPKD). *J. Am. Soc. Nephrol.* **3**, 292 [abstract].

Avner, E.D. (1993) Renal developmental diseases. *Semin. Nephrol.* **13**, 427-435.

Avner, E.D., Sweeney, W.E. and Norman, J.T. (1991) Abnormal EGF-R expression in congenital murine polycystic kidney disease. *J. Am. Soc. Nephrol.* **2**, 250 [abstract].

Avner, E.D., McAteer, J.A. and Evan, A.P. (1990) Models of cysts and cystic kidneys. In "The Cystic Kidney", (K.D. Gardner

and J. Bernstein, Eds.), pp. 55-98, Kluwer, Dordrecht, Netherlands.

Avner, E.D., Studnicki, F.E., Young, M.C., Sweeney, W.E., Piesco, N.P., Ellis, D. and Fettermann, G.H. (1987) Congenital murine polycystic kidney disease: I. The ontogeny of tubular cyst formation, in *The Original Article Series, Pediatr. Nephrol.* **1**, 587-596.

Aziz, N., Maxwell, M.M., St.-Jacques, B. and Brenner, B.M. (1993) Downregulation of Ke 6, a novel gene encoded within the major histocompatibility complex, in murine polycystic kidney disease. *Mol. Cell. Biol.* **13**, 1847-1853.

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

Barakat, A.Y. and Butler, M.G. (1987) Renal and urinary tract abnormalities associated with chromosome aberrations. *Int. J. Pediat. Nephrol.* **8**, 215-226.

Baserga, R. (1981) The cell cycle. *N. Engl. J. Med.* **304**, 453-459.

Bates, P.F. and Swift, R.A. (1983) Double *cos* site vectors: simplified cosmid cloning. *Gene* **26**, 137-146.

Bennett, W.M., Elzinga, L.W. and Barry, J.M. (1992) Polycystic kidney disease: II. Diagnosis and management. *Hospital Practice*, April 15, 61-72.

Bernstein, J. and Gardner, K.D. (1986) Renal cystic disease and renal dysplasia. In "Campbells' Urology" (P.C. Walsh, R.F.

Gittes, A.D. Perlmutter and T.A. Stamey, Eds.), Vol. 2, pp. 1760-1818, W.B. Saunders, Philadelphia, PA.

Bernstein, J., Brough, A.J. and McAdams, A.J. (1974) The renal lesion in syndromes of multiple congenital malformations-cerebrohepatorenal syndrome; Jeune asphyxiating thoracic dystrophy; tuberous sclerosis; Meckel syndrome. *Birth Defects: Original Article Series* **10**, 35-43.

Bernstein, J., Chandra, M., Creswell, J., Kahn, E., Malouf, N.N., McVicar, M., Weinberg, A.G. and Wybel, R.E. (1987b) Renal-hepatic-pancreatic dysplasia: a syndrome reconsidered. *Am. J. Med. Genet.* **26**, 391-403.

Bernstein, J. (1978) in "Pediatric Kidney Disease", (C.M. Edelmann, Ed.), pp. 557-570, Little, Brown and Co., Boston, MA.

Bettuzzi, S., Hiipakka, R.A., Gilna, P. and Liao, S. (1989) Identification of an androgen-repressed mRNA in rat ventral prostate as coding for sulphated glycoprotein 2 by cDNA cloning and sequence analysis. *Biochem. J.* **257**, 293-296.

Bissonnette, R.P., Echeverri, F., Mahboubi, A. and Green, D.R. (1992) Apoptotic cell death induced by c-*myc* is inhibited by *bcl-2*. *Nature* **359**, 552-554.

Blyth, H. and Ockenden, B.G. (1971) Polycystic disease of kidneys and liver presenting in childhood. *J. Med. Genet.* **8**, 257-284.

Böhnlein, E., Chowdhury, K. and Gruss, P. (1985) Functional analysis of the regulatory region of polyoma mutant F9-1 DNA. *Nucl. Acids Res.* **13**, 4789-4809.

Breindl, M., Harbers, K. and Jaenisch, R. (1984) Retrovirus-induced lethal mutation in collagen I gene of mice is associated with an altered chromatin structure. *Cell* **38**, 9-16.

Breuton, L.A., Dillon, M.J. and Winter, R.M. (1990) Ellis-van Crevald syndrome, Jeune syndrome, and renal-hepatic-pancreatic dysplasia: separate entities or disease spectrum? *J. Med. Genet.* **27**, 252-255.

Calvet, J.P. (1993) Polycystic kidney disease: primary extracellular matrix abnormalitiy or defective cellular differentiation. *Kidney Int.* **43**, 101-108.

Carone, F.A., Rowland, R.G., Perlman, S.G. and Ganote, C.E. (1974) The pathogenesis of drug-induced renal cystic disease. *Kidney Int.* **5**, 411-421.

Carone, F.A., Jin, H., Nakamura, S. and Kanwar, Y.S. (1993) Decreased synthesis and delayed processing of sulfated glycoproteins by cells from human polycystic kidneys. *Lab. Invest.* **68**, 413-418.

Carone, F.A., Makino, H. and Kanwar, Y.S. (1988) Basement membrane antigens in renal polycystic disease. *Am. J. Pathol.* **130**, 466-471.

Carone, F., Nakumara, S., Caputo, M., Jin, H., Nelson, J. and Kanwar, Y. (1992) Cell polarity in human renal cystic disease. *J. Amer. Soc. Nephrol.* **3**, 293 [abstract].

Casamassima, A.C., Mamunes, P., Gladstone, I.M., Solomon, S. and Moncure, C. (1987) A new syndrome with features of the Smith-Lemli-Opitz and Meckel-Gruber syndromes in a

sibship with cerebellar defects. *Am. J. Med. Genet.* **26**, 321-336.

Cideciyan, D., Rodriguez, M.M., Haun, R.L., Abdenour, G.E. and Bruce, J.H. (1993) New findings in short rib syndrome. *Am. J. Med. Genet.* **46**, 255-259.

Cobben, J.M., Breuning, M.H., Schoots, C., ten Kate, L.P. and Zerres, K. (1990) Congenital hepatic fibrosis in autosomal-dominant polycystic kidney disease. *Kidney Int.* **38**, 880-885.

Coles, H.S.R., Burne, J.F. and Raff, M.C. (1993) Large-scale normal cell death in the developing rat kidney and its reduction by epidermal growth factor. *Development* **118**, 777-784.

Connor, J. Buttyan, R., Olsson, C.A., D'Agati, V., O'Toole, K. and Sawczuk, I.S. (1991) *SGP-2* expression as a genetic marker of progressive cellular pathology in experimental hydronephrosis. *Kidney Int.* **39**, 1098-1103.

Coppola, J.A. and Cole, M.D. (1986) Constitutive c-*myc* oncogene expression blocks mouse erythroleukaemia cell differentiation but not commitment. *Nature* **320**, 760-763.

Cowley, B.D., Chadwick, L.J., Grantham, J.J. and Calvet, J.P. (1991) Elevated proto-oncogene expression in polycystic kidneys of the C57BL/6J *(cpk)* mouse. *J. Am. Soc. Nephrol.* **1**, 1048.

Cowley, B.D., Smardo, F.L., Grantham, J.J. and Calvet, J.P. (1987) Elevated c-*myc* proto-oncogene expression in autosomal recessive polycystic kidney disease. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 8394-8398.

Cowley, B.D., Gudapaty, S., Kraybill, A.L., Barash, B.D., Harding, M.A., Calvet, J.P. and Gattone, V.H. (1993) Autosomal-dominant polycystic kidney disease in the rat. *Kidney Int.* **43**, 522-534.

Crocker, J.F.S., Blecher, S.R., Givner, M.L. and McCarthy, S.C. (1987) Polycystic kidney and liver disease and corticosterone changes in the *cpk* mouse. *Kidney Int.* **31**, 1088-1091.

Curran, T. and Franza, B.R., Jr. (1988) Fos and Jun: the AP-1 connection. *Cell* **55**, 395-397.

Dalgaard, O.Z. (1971) in "Diseases of the Kidney", (M.B. Strauss and L.G. Welt, Eds.), Vol. 2, p. 1252, Little, Brown and Co., Boston, MA.

Daoust, M.C., Bichet, D.G. and Somlo, S. (1993) A French-Canadian family with autosomal dominant polycystic kidney disease (ADPKD) unlinked to ADPKD1 and ADPKD2. *J. Am. Soc. Nephrol.* **4**, 262 [abstract].

Davisson, M.T., Guay-Woodford, L.M., Harris, H.W. and D'Eustachio, P. (1991) The mouse polycystic kidney disease mutation (*cpk*) is located on proximal chromosome 12. *Genomics* **9**, 778-781.

Davisson, M.T. and Roderick, T.H. (1993) Chromosome 14. *Mouse Genome* **91**, 67.

Denis, N., Blanc, S., Leibovitch, M.P., Nicolaiew, N., Dautry, F., Raymondjean, M., Kruh, J. and Kitzis, A. (1987) c-*myc* oncogene expression inhibits the initiation of myogenic differentiation. *Exp. Cell. Res.* **172**, 212-217.

Devereux, J., Haeberli, P. and Smithies, O. (1984) A comprehensive set of sequence analysis programs for the VAX. *Nucl. Acids Res.* **12**, 387-395.

Du, J. and Wilson, P.T. (1990) Mislocation of high affinity epidermal growth factor (EGF) receptors in human autosomal dominant polycystic kidney disease (ADPKD). *J. Am. Soc. Nephrol.* **1**, 627 [abstract].

Ebihara, I., Killen, P.D., Laurie, G.W., Huang, T., Yamada, Y., Martin, G.R. and Brown, K.S. (1988) Altered mRNA expression of basement membrane components in a murine model of polycystic kidney disease. *Lab. Invest.* **58**, 262-269.

Ebihara, I., Nakamura, T., Tomino, Y., Nagao, S., Takahashi, H. and Koide, H. (1992) Abnormal gene regulation of growth factor and extracellular matrix (ECM) components in kidney of slowly progressive polycystic kidney disease (PKD) mouse. *Am. J. Soc. Nephrol.* **3**, 294 [abstract].

European Chromosome 16 Tuberous Sclerosis Consortium (1993) Identification and characterization of the tuberous sclerosis gene on chromosome 16. *Cell* **75**, 1305-1315.

Evan, A.P. and Gardner, K.D. (1985) Morphology of polycystic kidney disease in man and experimental models. In "Chronic Renal Disease: Causes, Complications, and Treatment" (N.B. Cummings and S. Klahr, Eds.) pp. 55-68, Plenum, NY.

Evan, A.P. and McAteer, J.A. (1990) Cyst cells and cyst walls. In "The Cystic Kidney" (K.D. Gardner and J. Bernstein, Eds.) pp. 21-42, Kluwer, Dordrecht, Netherlands.

Evans, J.P. and Palmiter, R.D. (1991) Retrotransposition of a mouse L1 element. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 8792-8795.

Fanidi, A., Harrington, E.A. and Evan, G.I. (1992) Cooperative interaction between c-*myc* and *bcl-2* proto-oncogenes. *Nature* **359**, 554-556.

Farber, E. (1956) Similarities in the sequence of early histological changes induced in the liver of the rat by ethionine, 2-acetylaminofluorene, and 3'-methyl-4-dimethylaminoazobenzene. *Cancer Res.* **16**, 142-148.

Feinberg, A.P. and Vogelstein, B. (1984) Addendum: a technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* **137**, 266-267.

Gabow, P.A. (1990) Autosomal dominant polycystic kidney disease. In "The Cystic Kidney", (K.D. Gardner and J. Bernstein, Eds.), pp. 295-326, Kluwer, Dordrecht, Netherlands.

Gardner, K.D. (1990) Cysts and cystic kidneys. In "The Cystic Kidney", (Gardner, K.D. and Bernstein, J., Eds.), pp. 3-20, Kluwer, Dordrecht, Netherlands.

Gardner, K.D. (1969) Composition of fluid in twelve cysts of a polycystic kidney. *N. Engl. J. Med.* **281**, 985-988.

Gardner, K.D. and Evan, A.P. (1983) Renal cystic disease induced by diphenylthiazole. *Kidney Int.* **24**, 43-52.

Gardner, K.D. (1969) Composition of fluid in twelve cysts of a polycystic kidney. *N. Engl. J. Med.* **281**, 985-988.

Gardner, K.D., Burnside, J.S., Elzinga, L.W. and Locksley, R.M. (1991) Cytokines in fluids from polycystic kidneys. *Kidney Int.* **39**, 718-724.

Gardner, K.D., Burnside, J.S., Skipper, B.J., Swan, S.K., Bennett, W.M., Connors, B.A. and Evan, A.P. (1992) On the probability that kidneys are different in autosomal dominant polycystic disease. *Kidney Int.* **42**, 1199-1206.

Gärtner, J., Moser, H. and Valle, D. (1992) Mutations in the 70K peroxisomal membrane protein gene in Zellweger syndrome. *Nature Genet.* **1**, 16-23.

Gattone, V.H., Andrews, G.K., Fu-Wen, N., Chadwick, L.J., Klein, R.M. and Calvet, J.P. (1990) Defective epidermal growth factor gene expression in mice with polycystic kidney disease. *Dev. Biol.* **138**, 225-230.

Gattone, V.H. and Lowden, D.A. (1992) Epidermal growth factor ameliorates infantile polycystic kidney disease in mice. *J. Am. Soc. Nephrol.* **3**, 295 [abstract].

Germino, G.G., Somlo, S., Weinstat-Saslow, D. and Reeders, S.T. (1993) Positional cloning approach to the dominant polycystic kidney disease gene, *PKD1*. *Kidney Int.* **43** (suppl. 39), S20-S25.

Germino, G.G., Weinstat-Saslow, D., Himmelbauer, H., Gillespie, G.A.J., Somlo, S., Wirth, B., Barton, N., Harris, K.L., Frischauf, A.-M. and Reeders, S.T. (1992) The gene for autosomal dominant polycystic kidney disease lies in a 750-kb CpG-rich region. *Genomics* **13**, 144-151.

Gluzman, Y. (1981) SV40-transformed simian cells support the replication of early SV40 mutants. *Cell* **23**, 175-182.

Gluzman, Y. (1992) in "American Type Culture Collection Catalogue of Cell Lines and Hybridomas" (R. Hay, J. Caputo, T.R. Chen, M. Macy, P. McClintock and Y. Reid, Eds.) 7th Ed., p. 158, Rockville, MD.

Goebl, M. and Yanagida, M. (1991) The TPR snap helix: a novel protein repeat motif from mitosis to transcription. *Trends Biochem. Sci.* **16**, 173-177.

Gordon, J. (1992) Standardized nomenclature for transgenic animals. *ILAR News* **34**, 45-52.

Gordon, J.W., Scangos, G.A., Plotkin, D.J., Barbosa, A. and Ruddle, F.H. (1980) Genetic transformation of mouse embryos by microinjection. *Proc. Natl. Acad. Sci. U.S.A.* **77**, 7380-7384.

Grantham, J.J. (1990) Polycystic kidney disease: neoplasia in disguise. *Am. J. Kidney Dis.* **25**, 110-116.

Grantham, J.J. (1983) Polycystic kidney disease: a predominance of giant nephrons. *Am. J. Physiol.* **244**, F3-F10.

Grantham, J.J. (1991) Acquired cystic kidney disease. *Kidney Int.* **40**, 143-152.

Grantham, J.J., Geiser, J.L. and Evan, A.P. (1987) Cyst formation and growth in autosomal dominant polycystic kidney disease. *Kidney Int.* **31**, 1145-1152.

Gridley, T., Jaenisch, R. and Gendron-Maguire, M. (1991). The murine *Mov-34* gene: full-length cDNA and genomic organization. *Genomics* **11**, 501-507.

Gridley, T., Soriano, P. and Jaenisch, R. (1987). Insertional mutagenesis in mice. *Trends Genet.* **3**, 162-166.

Griep, A.E. and Westphal, H. (1988) Antisense myc sequences induce differentiation of F9 cells. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6806-6810.

Guay-Woodford, L.M., D'Eustachio, P., Muecher, G., Uhlhaas, S. and Zerres, K. (1990) Testing the hypothesis that murine congenital polycystic kidney disease *(cpk)* and human autosomal recessive polycystic kidney disease (ARPKD) affect homologous genes. *J. Am. Soc. Nephrol.* **1**, 300A [abstract].

Guay-Woodford, L.M., D'Eustachio, P. and Bruns, G.A.P. (1993) Identification of the syntenic human linkage group for the mouse congenital polycystic kidney *(cpk)* locus. *J. Am. Soc. Nephrol.* **4**, 814 [abstract].

Gutman, A. and Wasylyk, B. (1991) Nuclear targets for transcription regulation by oncogenes. *Trends Genet.* **7**, 49-54.

Harding, M.A., Chadwick, L.J., Gattone, V.H. and Calvet, J.P. (1991) The *SGP-2* gene is developmentally regulated in the mouse kidney and abnormally expressed in collecting duct cysts in polycystic kidney disease, *Dev. Biol.* **146**, 483-490.

Harding, M.A., Gattone, V.H., Grantham, J.J. and Calvet, J.P. (1992) Localization of overexpressed c-*myc* mRNA in

polycystic kidneys of the *cpk* mouse. *Kidney Int.* **41**, 317-325.

Hastie, N.D. (1989) in "Genetic Variants and Strains of the Laboratory Mouse", (M.F. Lyon and A.G. Searle, Eds.), 2nd Ed., pp. 559-573, Oxford Univ. Press, Oxford.

Herrera, G.A. (1991) *C-erbB-2* amplification in cystic renal disease. *Kidney Int.* **40**, 509-513.

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

Hill, R. and Van Heyningen, V. (1992) Mouse mutations and human disorders are Paired [sic]. *Trends Genet.* **8**, 119-120.

Hirano, T., Kinoshita, N., Morikawa, K. and Yanagida, M. (1990) Snap helix with knob and hole: essential repeats in *S. pombe* nuclear protein nuc2^{+}. *Cell* **60**, 319-328.

Hodgkinson, C.A., Moore, K.J., Nakayama, A., Steingrimsson, E., Copeland, N.G., Jenkins, N.A. and Arnheiter, H. (1993) Mutations at the mouse microphthalmia locus are associated with defects in a gene encoding a novel basic-helix-loop-helix-zipper protein. *Cell* **74**, 395-404.

Hogan, B.L.M., Costantini, F. and Lacy, E. (1986) Manipulating the mouse embryo: a laboratory manual. Cold Spring Harbor Laboratory. Cold Spring Harbor, New York.

Holthofer, H. (1987) Ontogeny of cell type-specific enzyme reactivities in kidney collecting ducts. *Pediatr. Res.* **22**, 504-508.

Honoré, B., Leffers, H., Madsen, P., Rasmussen, H.H., Vandekerckhove, J. and Celis, J.E. (1992) Molecular cloning and expression of a transformation-sensitive human protein containing the TPR motif and sharing identity to the stress-inducible yeast protein ST11. *J. Biol. Chem.* **267**, 8485-8491.

Horikoshi, S., Kubota, S., Martin, G.R., Yamada, Y. and Klotman, P.E. (1991) Epidermal growth factor expression in the congenital polycystic mouse. *Kidney Int.* **39**, 57-62.

Horton, C.E., Davisson, M.T., Jacobs, J.B., Bernstein, G.T., Retik, A.B. and Mandell, J. (1988) Congenital progressive hydronephrosis in mice: a new recessive mutation. *J. Urol.* **140**, 1310-1315.

Hsai, Y.E., Bratu, M. and Herbordt, A. (1971) Genetics of the Meckel syndrome (Dysencephalia Splanchnocystica). *Pediatrics* **48**, 237-247.

Hull, R.N., Cherry, W.R. and Weaver, W. (1976) The origin and characteristics of a pig kidney cell strain, LLC-PK_1. *In Vitro* **12**, 670-678.

Hummel, K.P. (1958) The inheritance and expression of *Disorganization*, an unusual mutation in the mouse. *Exp. Zool.* **137**, 389-423.

Iverson, W.O., Fetterman, G.H., Jacobson, E.R., Olsen, J.H., Senior, D.F. and Schobert, E.E. (1982) Polycystic kidney and liver disease in Springbok: I. Morphology of the lesions. *Kidney Int.* **22**, 146-155.

Jackson, I.J. (1991) A reappraisal of non-consensus mRNA splice sites. *Nucl. Acids Res.* **19**, 3795-3798.

Jacob, H.J., Lindpaintner, K., Lincoln, S.E., Kusumi, K., Bunker, R.K., Mao, Y.-P., Ganten, D., Dzau, V.J. and Lander, E.S. (1991) Genetic mapping of a gene causing hypertension in the stroke-prone spontaneously hypertensive rat. *Cell* **67**, 213-224.

Jähner, D. and Jaenisch, R. (1985) Retrovirus-induced *de novo* methylation of flanking host sequences correlates with gene inactivity. *Nature* **315**, 594-597.

Kaissling, B. and Kriz, W. (1992) Morphology of the loop of Henle, distal tubule, and collecting duct. In "Handbook of Physiology" (E.E. Windhager, Ed.), Vol. 1, pp. 109-167, Oxford University Press, NY.

Kaplan, B.S., Fay, J., Shah, V., Dillon, M.J. and Barratt, T.M. (1989) Autosomal recessive polycystic kidney disease. *Pediatr. Nephrol.* **3**, 43-49.

Kaplan, B.S., Kaplan, P., de Chadarevian, J.-P., Jequier, S., O'Regan, S. and Russo, P. (1988) Variable expression of autosomal recessive polycystic kidney disease and congenital hepatic fibrosis within a family. *Am. J. Med. Genet.* **29**, 639-647.

Kelley, K.A., Agarwal, N., Reeders, S. and Herrup, K. (1991) Renal cyst formation and multifocal neoplasia in transgenic mice carrying the simian virus 40 early region. *J. Am. Soc. Nephrol.* **2**, 84-97.

Kimberling, W.J., Kumar, S., Gabow, P.A., Kenyon, J.B., Connolly, C.J. and Somlo, S. (1993) Autosomal dominant polycystic kidney disease: localization of the second gene to chromosome 4q13-q23. *Genomics* **18**, 467-472.

King, C.R. and Piatigorsky, J. (1983) Alternative RNA splicing of the murine αA-crystallin gene: protein-coding information within an intron. *Cell* **32**, 707-712.

Kissane, J.M. (1974) in "Pathology of the Kidney", (R.H. Heptinstall, Ed.), pp. 69-119, Little, Brown and Co., Boston, MA.

Koseki, C., Herzlinger, D. and Al-Awqati, Q. (1992) Apoptosis in metanephric development. *J. Cell Biol.* **119**, 1327-1333.

Kozak, M. (1991) An analysis of vertebrate mRNA sequences: intimations of translational control. *J. Cell Biol.* **115**, 887-903.

Langley, L.L., Telford, I.R. and Christensen, J.B. (1974) The neuron. In "Dynamic Anatomy and Physiology" (C. Dilworth and S.B. Mobley, Eds.), 4th Ed., pp. 209-224, McGraw-Hill, NY.

Latif, F., Tory, K., Gnarra, J., Yao, M., Duh, F.-M., Orcutt, M.L., Stackhouse, T., Kuzmin, I., Modi, W., Geil, L., Schmidt, L., Zhou, F., Li, H., Wei, M.H., Chen, F., Glenn, G., Choyke, P., Walther, M.M., Weng, Y., Duan, D.-S. R., Dean, M., Glavac, D., Richards, F.M., Crossey, P.A., Ferguson-Smith, M.A., Paslier, D.L., Chumakov, I., Cohen, D., Chinault, A.C., Maher, E.R., Linehan, W.M., Zbar, B. and Lerman, M.I. (1993) Identification of the von Hippel-Lindau disease tumor suppressor gene. *Science* **260**, 1317-1320.

Lee, J.J., Radice, G., Perkins, C.P. and Costantini, F. (1992) Identification and characterization of a novel, evolutionarily-conserved gene disrupted by the murine Hb58 embryonic lethal transgene insertion. *Development* **115**, 277-288.

Lindemann, G.W., Foo-Hurwitz, I., Magenheimer, B.S., Rankin, C.A. and Calvet, J.P. (1992) Isolation of differentially expressed cDNAs in polycystic kidneys of the *cpk* mouse by subtractive cloning. *J. Amer. Soc. Nephrol.* **3**, 298 [abstract].

MacDonald, R.A. (1961) "Lifespan" of liver cells. *Arch. Intern. Med.* **107**, 335-343.

MacKay, K., Striker, L.J., Pinkert, C.A., Brinster, R.L. and Striker, G.E. (1987) Glomerulosclerosis and renal cysts in mice transgenic for the early region of SV40. *Kidney Int.* **32**, 827-837.

McGeoch, J.E.M. and Darmady, E.M. (1976) Polycystic disease of kidney, liver and pancreas: a possible pathogenesis. *J. Pathol.* **119**, 221-228.

McGinnis, W., Hart, C.P., Gehring, W.J. and Ruddle, F.H. (1984) Molecular cloning and chromosome mapping of a mouse DNA sequence homologous to homeotic genes of *Drosophila*. *Cell* **38**, 675-680.

Meek, D.W. and Street, A.J. (1992) Nuclear protein phosphorylation and growth control. *Biochem. J.* **287**, 1-15.

Mills, D.R. and Kramer, F.R. (1979) Structure independent nucleotide sequence analysis. *Proc. Natl. Acad. Sci. U.S.A.* **76**, 2232-2235.

Mitchell, P.J. and Tjian, R. (1989) Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. *Science* **245**, 371-378.

Moyer, J.H., Lee-Tischler, M.J., Kwon, H.-Y., Schrick, J.J., Avner, E.D., Sweeney, W.E., Godfrey, V.L., Cacheiro, N.L.A., Wilkinson, J.E. and Woychik, R.P. (1994a) Cloning and characterization of a candidate gene associated with a new insertional mutation causing recessive polycystic kidney disease in mice. *Science*, in press.

Moyer, J.H. and Woychik, R.P. (1994b) Physical structure and expression pattern of a disrupted gene in the *TgN737Rpw* transgenic mouse model of polycystic kidney disease, manuscript in preparation.

Nadeau, J.H., McCarthy, L. and Cox, R. (1993) Mouse chromosome 14. *Mamm. Genome* **4**, S203-S210.

Nakai, M., Fukase, M., Kinoshita, Y. and Fuhita, T. (1988) Atrial natriuretic factor inhibits phosphate uptake in opossum kidney cells: as a model of renal proximal tubules. *Biochem. Biophys. Res. Comm.* **152**, 1416-1420.

Nasim, F.-U., Spears, P.A., Hoffmann, H.M., Kuo, H.-C. and Grabowski, P.J. (1990) A sequential splicing mechanism promotes selection of an optional exon by repositioning a downstream 5' splice site in preprotachykinin pre-mRNA. *Genes Dev.* **4**, 1157-1184.

Nauta, J., Ozawa, Y., Sweeney, W.E., Rutledge, J.C. and Avner, E.D. (1993) Renal and biliary abnormalities in a new murine model of autosomal recessive polycystic kidney disease, in *The Original Article Series: Pediatr. Nephrol.* **7**, 163-172.

Neel, H., Weil, D., Giansante, C. and Dautry, F. (1993) *In vivo* cooperation between introns during pre-mRNA processing. *Genes Dev.* **7**, 2194-2205.

Nørby, S. and Schwartz, M. (1990) Possible locus for polycystic kidney disease on chromosome 2. *Lancet* **336**, 323-324 [abstract].

Parfrey, P.S., Bear, J.C., Morgan, J., Cramer, B.C., McManamon, P.J., Gault, M.H., Churchill, D.N., Singh, M., Hewitt, R., Somlo, S. and Reeders, S.T. (1990) The diagnosis and prognosis of autosomal dominant polycystic kidney disease. *N. Engl. J. Med.* **323**, 1085-1090.

Patau, K., Smith, D.W., Therman, E., Inhorn, S.L. and Wagner, H.P. (1960) Multiple congenital anomaly caused by an extra autosome. *Lancet* **1**, 790-793.

Pearson, W.R. and Lipman, D.J. (1988) Improved tools for biological sequence comparison. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 2444-2448.

Perrone, R.D. (1985) *In vitro* function of cyst epithelium from human polycystic kidney. *J. Clin. Invest.* **76**, 1688-1691.

Peters, D.J.M., Spruit, L., Saris, J.J., Ravine, D., Sandkuijl, L.A., Fossdal, R., Boersma, J., van Eijk, R., Nørby, S. Constantinou-Deltas, C.D., Pierides, A., Brissenden, J.E., Frants, R.R., van Ommen, G.-J.B. and Breuning, M.H. (1993) Chromosome 4 localization of a second gene for autosomal dominant polycystic kidney disease. *Nature Genet.* **5**, 359-362.

Ponzetto, C. and Wolgemuth, D.J. (1985) Haploid expression of a unique c-*abl* transcript in the mouse male germ line. *Mol. Cell. Biol.* **5**, 1791-1794.

Ransone, L.J. and Verma, I.M. (1990) Nuclear proto-oncogenes fos and jun. *Annu. Rev. Cell Biol.* **6**, 539-557.

Reeders, S.T. (1990) The genetics of renal cystic disease. In "The Cystic Kidney", (K.D. Gardner and J. Bernstein, Eds.), pp. 117-143, Kluwer, Dordrecht, Netherlands.

Reeders, S.T., Breuning, M.H., Davies, K.E., Nicholls, R.D., Jarman, A.P., Higgs, D.R., Pearson, P.L. and Weatherall, D.J. (1985) A highly polymorphic DNA marker linked to adult polycystic kidney disease on chromosome 16. *Nature* **317**, 542-544.

Reeders, S.T. (1992) Multilocus polycystic disease. *Nature Genet.* **1**, 235.

Rinchik, E. M., Bultman, S.J., Horsthemke, B., Lee, S.-T., Strunk, K.M., Spritz, R.A., Avidano, K.M., Jong, M.T.C. and Nicholls, R.D. (1993) A gene for the mouse pink-eyed dilution locus and for human type II oculocutaneous albinism. *Nature* **361**, 72-76.

Rocco, M.V., Neilson, E.G., Hoyer, J.R. and Ziyadeh, F.N. (1992) Attenuated expression of epithelial cell adhesion molecules in murine polycystic kidney disease. *Am. J. Physiol.* **262**, F679-F686.

Russell, E.S. and McFarland, E.C. (1977) Cystic kidneys. *Mouse Newsletter* **56**, 40.

Saito, H., Kranz, D.M., Takagaki, Y., Hayday, A.C., Eisen, H.N. and Tonegawa, S. (1984) A third rearranged and expressed gene in a clone of cytotoxic T lymphocytes. *Nature* **312**, 36-40.

Salonen, R. (1984) The Meckel syndrome: clinicopathological findings in 67 patients. *Am. J. Med. Genet.* **18**, 671-689.

Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) in "Molecular Cloning: A Laboratory Manual", (N. Irwin, N. Ford, C. Nolan, M. Ferguson and M. Ockler, Eds.), 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sanger, F., Nicklen, S. and Coulson, A.R. (1977) DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 5463-5467.

Saraiva, J.M. and Baraitser, M. (1992) Joubert syndrome: a review. *Am. J. Med. Genet.* **43**, 726-731.

Schaffner, D.L., Barrios, R., Massey, C., Bañez, E.I., Ou, C.-N., Rajagopalan, S., Aguilar-Cordova, E., Lebovitz, R.M., Overbeek, P.A. and Lieberman, M.W. (1993) Targeting of the *ras*T24 oncogene to the proximal convoluted tubules in transgenic mice results in hyperplasia and polycystic kidneys. *Am. J. Pathol.* **142**, 1051-1060.

Schmidt, E.V., Pattengale, P.K., Weir, L. and Leder, P. (1988) Transgenic mice bearing the human c-*myc* gene activated by an immunoglobulin enhancer: A pre-B-cell lymphoma model. *Proc. Natl. Acad. Sci. U.S.A.* **85**, 6047-6051.

Schnieke, A., Harbers, K. and Jaenisch, R. (1983) Embryonic lethal mutation in mice induced by retrovirus insertion into the a1 (I) collagen gene. *Nature* **304**, 315-320.

Schwartz, D.C. and Cantor, C.R. (1984) Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* **37**, 67-75.

Sikorski, R.S., Boguski, M.S., Goebl, M. and Hieter, P. (1990) A repeating amino acid motif in *CDC23* defines a family of

proteins and a new relationship among genes required for mitosis and RNA synthesis. *Cell* **60**, 307-317.

Singh, G., Supp, D.M., Schreiner, C., McNeish, J., Merker, H.-J., Copeland, N.G., Jenkins, N.A., Potter, S.S. and Scott, W. (1991) *legless* insertional mutation: morphological, molecular, and genetic characterization. *Genes Dev.* **5**, 2245-2255.

Stöcklin, E., Botteri, F. and Groner, B. (1993) An activated allele of the *c-erbB-2* oncogene impairs kidney and lung function and causes early death of transgenic mice. *J. Cell Biol.* **122**, 199-208.

Strayer, D.S. and Kissane, J.M. (1979) Dysplasia of the kidneys, liver and pancreas: Report of a variant of Ivemark's syndrome. *Human Pathol.* **10**, 228-234.

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

Tabor, S. and Richardson, C.C. (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. *Proc. Natl. Acad. Sci. U.S.A.* **84**, 4767-4771.

Tacke, R. and Goridis, C. (1991) Alternative splicing in the neural cell adhesion molecule pre-mRNA regulation of exon 18 skipping depends on the 5' spice site. *Genes Dev.* **5**, 1416-1429.

Takahashi, H., Calvet, J.P., Dittemore-Hoover, D., Yoshida, K., Grantham, J.J. and Gattone, V.H. (1991) A hereditary model of

slowly progressive polycystic kidney disease in the mouse. *J. Am. Soc. Nephrol.* 1, 980-989.

Talerico, M. and Berget, S.M. (1990) Effect of 5' splice site mutations on splicing of the preceding intron. *Mol. Cell. Biol.* **10**, 6299-6305.

Taub, M. and Saier, M.H. (1979) An established but differentiated kidney epithelial cell line (MDCK). In "Methods in Enzymology", (W.B. Jakoby, Ed.), Vol. 58, pp. 552-560, Academic Press, NY.

Tieder, M., Levy, M., Gubler, M.C., Gagnadoux, M.F. and Broyer, M. (1982) Renal abnormalities in the Bardet-Biedl syndrome. *Int. J. Ped. Nephrol.* **3**, 199-203.

Torres, V.E., Donovan, K.A., Scicli, G., Holley, K.E., Thibodeau, S.N., Carretero, O.A., Inagami, T., McAteer, J.A. and Johnson, C.M. (1992) Synthesis of renin by tubulocystic epithelium in autosomal-dominant polycystic kidney disease. *Kidney Int.* **42**, 364-373.

Trudel, M., D'Agati, V. and Costantini, F. (1991) c-*myc* as an inducer of polycystic kidney disease in transgenic mice. *Kidney Int.* **39**, 665-671.

Van der Leij, I., Franse, M.M., Elgersma, Y., Distel, B. and Tabak, H.F. (1993) PAS10 is a tetratricopeptide-repeat protein that is essential for the import of most matrix proteins into peroxisomes of *Saccharomyces cerevisiae. Proc. Natl. Acad. Sci. U.S.A.* **90**, 11782-1176.

Veis, D.J., Sorenson, C.M., Shutter, J.R. and Korsmeyer, S.J. (1993) *Bcl-2*-deficient mice demonstrate fulminant lymphoid

apoptosis, polycystic kidneys, and hypopigmented hair. *Cell* **75**, 229-240.

Wanke, R., Hermanns, W., Folger, S., Wolf, E. and Brem, G. (1991) Accelerated growth and visceral lesions in transgenic mice expressing foreign genes of the growth hormone family: an overview. *Pediatr. Nephrol.* **5**, 513-521.

Weiher, H., Noda, T., Gray, D.A., Sharpe, A.H. and Jaenisch, R. (1990) Transgenic mouse model of kidney disease: Insertional inactivation of ubiquitously expressed gene leads to nephrotic syndrome. *Cell* **62**, 425-434.

Welling, L.W. and Welling, D.J. (1985) Kinetics of cyst development. In "Chronic Renal Disease", (N.B. Cummings and S. Klahr, Eds.), pp. 95-103, Plenum Medical Book Company, NY.

Wilson, P.D., Sherwood, A.C., Palla, K., Du., J., Watson, R. and Norman, J.T. (1991) Reversed polarity of Na^+-K^+-ATPase: mislocation to apical plasma membranes in polycystic kidney disease epithelia. *Am. J. Physiol.* **260**, F420-F430.

Wilson, P.D. and Sherwood, A.C. (1991) Tubulocystic epithelium. *Kidney Int.* **39**, 450-463.

Wilson, P.D. (1991) Aberrant epithelial cell growth in autosomal dominant polycystic kidney disease. *Am. J. Kidney Dis.* **17**, 634-637.

Winter, R.M. (1988) Malformation syndromes: a review of mouse/human homology. *J. Med. Genet.* **25**, 480-487.

Wirth, B., Zerres, K., Fischbach, M., Claus, D., Neumann, H.P.H., Lennert, T., Brodehl, J., Neugebauer, M., Muller-Wiefel, D.E., Geisert, J. and Gal, A. (1987) Autosomal recessive and dominant forms of polycystic kidney disease are not allelic. *Human Genet.* **77**, 221-222.

Woychik, R.P., Maas, R.L., Zeller, R., Vogt, T.F. and Leder, P. (1990) "Formins": Proteins deduced from the alternative transcripts of the limb deformity gene. *Nature* **346**, 850-853.

Wright, A.F., Teague, P.W., Pound, S.E., Pignatelli, P.M., Macnicol, A.M., Carothers, A.D., De May, R.J., Allan, P.L. and Watson, M.L. (1993) A study of genetic linkage heterogeneity in 35 adult-onset polycystic kidney disease families. *Human Genet.* **90**, 569-571.

Wyllie, A.H. (1993) Apoptosis (The 1992 Frank Rose Memorial Lecture). *Br. J. Cancer* **67**, 205-208.

Ye, M., Grant, M., Sharma, M., Elzinga, L., Swan, S., Torres, V.E. and Grantham, J.J. (1992) Cyst fluid from human autosomal dominant polycystic kidneys promotes cyst formation and expansion by renal epithelial cells *in vitro*. *J. Am. Soc. Nephrol.* **3**, 984-994.

Zar, J.H. (1984) Contingency tables. In "Biostatistical Analysis" (B. Kurtz, Ed.), pp.61-78, Prentice-Hall, NJ.

Zerres, K., Völpel, M.-C. and Weiβ, H. (1984) Cystic kidneys: genetics, pathologic anatomy, clinical picture, and prenatal diagnosis. *Hum. Genet.* **68**, 104-135.

APPENDIX:

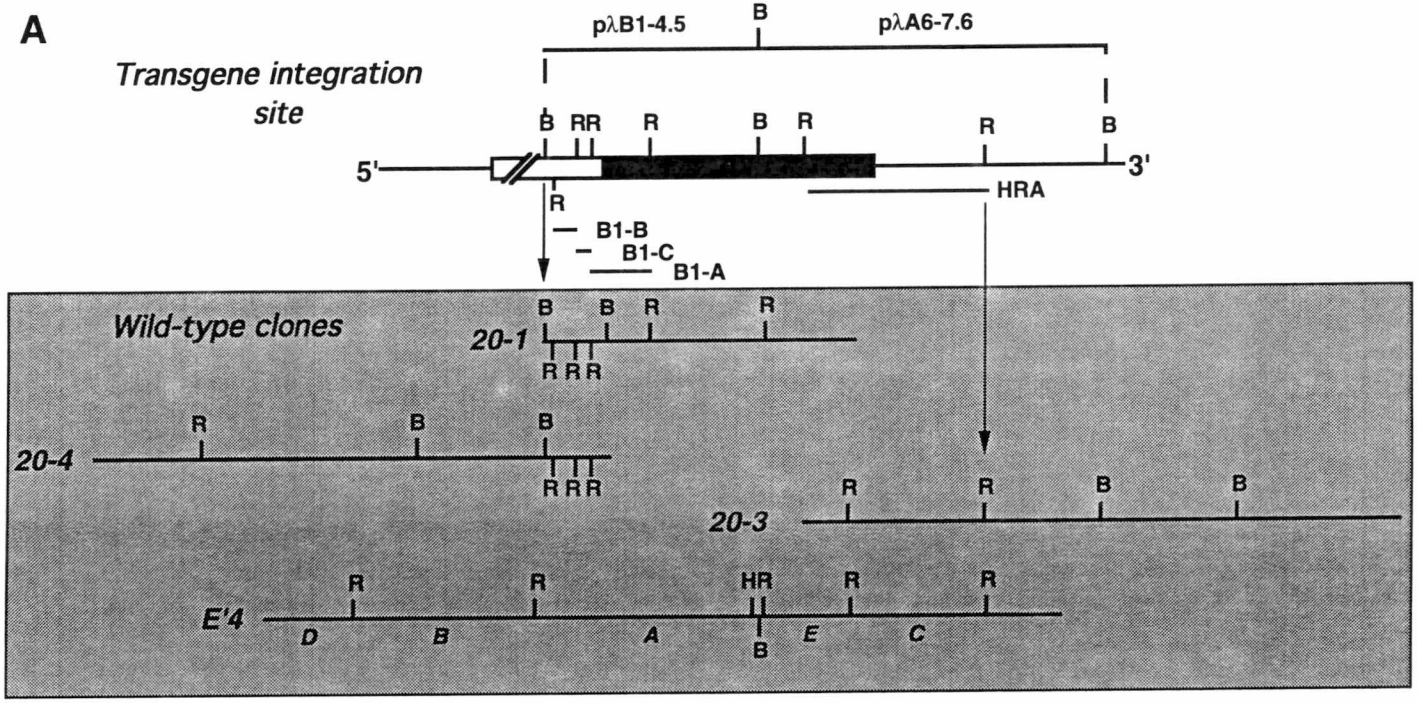
FIGURES

Figure 1. Isolation of mutant and wild-type clones from the *TgN737Rpw* locus and detection of a restriction fragment length polymorphism between transgenic and wild-type mice.

(A) Transgene integration locus and the corresponding wild-type clones (black box, transgene sequences; open box, cointegrated DNA). All DNA is shown in a 5'-3' orientation which was arbitrarily assigned initially, and later shown to be the correct orientation. Utilizing the transgene as a probe, two *Bam*HI (B) restriction fragments of 4.5 kb (pλB1-4.5) and 7.6 kb (pλA6-7.6) that contained sequences flanking the transgene insertion at the 5' and 3' junctions, respectively, were identified in *TgN737Rpw/TgN737Rpw* DNA. Subsequent *Eco*RI digestion of pλB1-4.5 produced B1-B, B1-C and B1-A, and *Hind*III/*Eco*RI digestion of pλA6-7.6 generated the *Eco*RI fragment HRA. Use of B1-C and HRA as hybridization probes led to the isolation of wild-type clones 20-1 and 20-4, and 20-3 and E'4, respectively (shaded area). Vertical arrows indicate alignment of two corresponding mutant and wild-type restriction sites. Several fragments (*A* through *E*, shown in italics) were derived from E'4 by digestion with *Sal*I and *Eco*RI. Unexpectedly, clones 20-1 and 20-4 did not share sequence overlap with clones 20-3 and E'4, suggesting a rearrangement at the integration site. Restriction enzyme sites: *Eco*RI (R), *Bam*HI (B), *Hind*III (H).

(B) Southern analysis showing restriction fragment length polymorphism between wild-type and transgenic mice. Southern blots containing *Eco*RI-digested genomic DNA from wild-type (+/+), *TgN737Rpw/TgN737Rpw* (Tg/Tg) and *TgN737Rpw/+* (Tg/+) mice were hybridized with radiolabeled HRA. A 3.0-kb wild-type *Eco*RI fragment and a 4.2-kb mutant

fragment are observed. Molecular sizes are shown on the left, in kb.


A
Transgene integration site
pλB1-4.5
pλA6-7.6
B
B RR R B R R B
5'
3'
HRA
R
B1-B
B1-C
B1-A
Wild-type clones
20-1
B B R R
R R R
20-4
R B B
R R R
20-3
R R B B
E'4
R R HR R R
D B A E C
B
1 kb

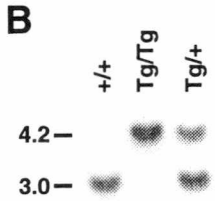

B
+/+
Tg/Tg
Tg/+
4.2
3.0

Figure 2. Gross appearance and histopathology of the kidney and liver of *TgN737Rpw* mice, and histopathology of the kidney and liver in human ARPKD.

(A) Typical gross lesions of bilateral polycystic kidneys in an FVB/N *TgN737Rpw* mouse.

(B) Cross section of a normal mouse kidney demonstrating cortex and medulla of a control FVB/N mouse.

(C) Cystic dilation of the collecting ducts in the medulla of the kidney of an FVB/N *TgN737Rpw* mouse.

(D) Renal medulla of a human ARPKD patient demonstrating a similar cystic dilation of the collecting ducts.

(E) Early lesion in the renal cortex of a *TgN737Rpw* mouse demonstrating dilation of collecting tubules.

(F) Similar early lesion in the cortex of a human ARPKD patient.

(G) Normal mouse liver.

(H) Hyperplastic and dysplastic bile ductules in the liver of an FVB/N *TgN737Rpw* mouse.

(I) Biliary hyperplasia, dysplasia, and portal fibrosis in the liver of an (FVB/N X C3H) F1 *TgN737Rpw* mouse.

(J) Abnormal intrahepatic biliary tract within the liver of a human ARPKD patient.

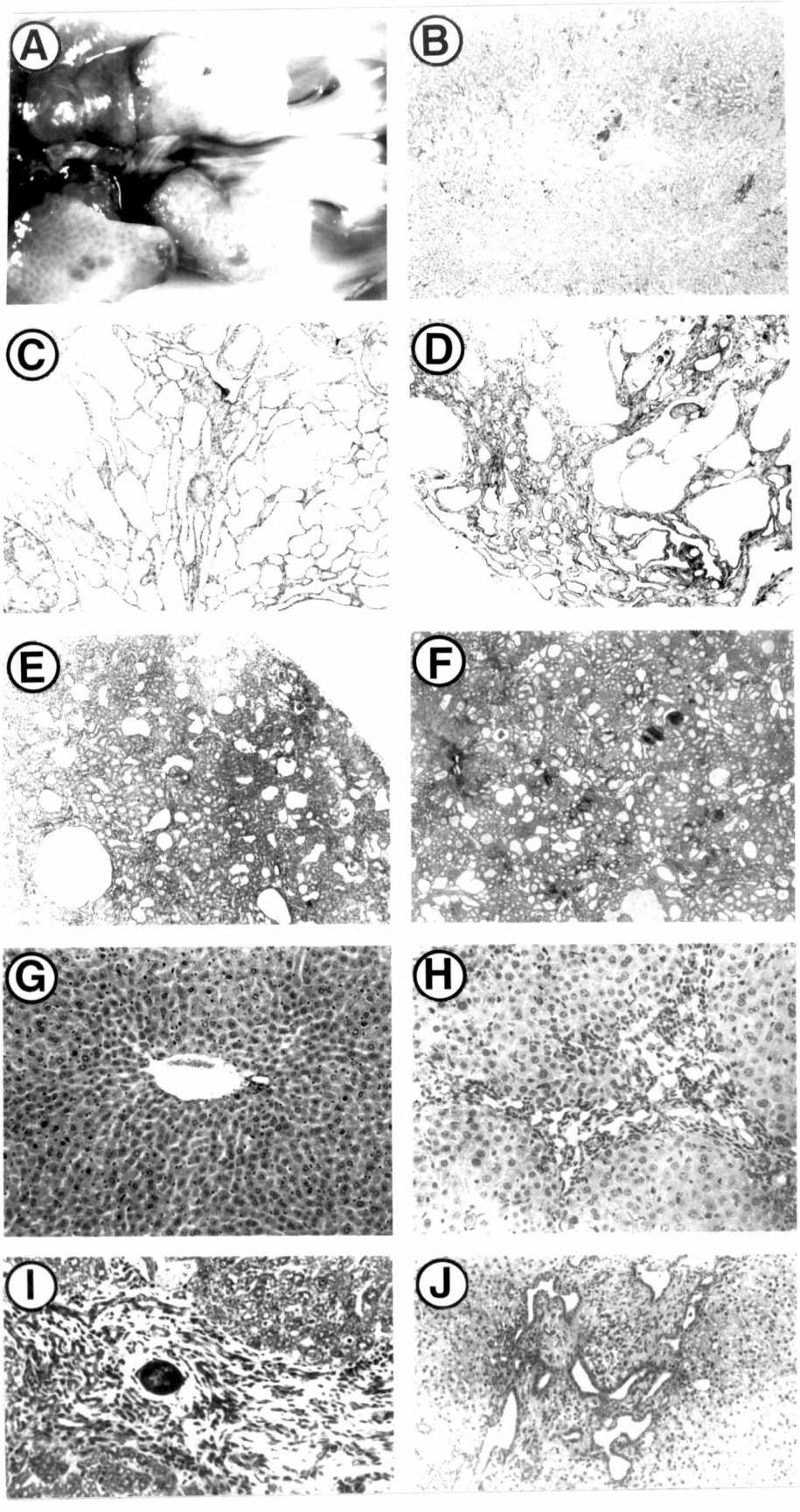

A
B
C
D
E
F
G
H
I
J

Figure 3. Composite maps of the *TgN737Rpw* mutant and corresponding wild-type locus.

(A) The mutant locus with the integrated sequences (black box, transgene sequences; open box, cointegrated DNA). Molecular characterization and comparison of the wild-type and mutant loci revealed that a deletion (designated by parentheses) occurred at the transgene integration site.

(B) Corresponding wild-type locus, with regions represented by *Eco*RI fragments A, B, C and E (large, bold lettering). The breakpoint between cointegrated and genomic DNA occurred within the region represented by the wild-type *Hind*III/*Eco*RI fragment HR0.2. BB3.5, utilized in Southern hybridization experiments, revealed no evidence of rearrangements at the mutant locus for approximately 23.0 kb 3' to the insertion.

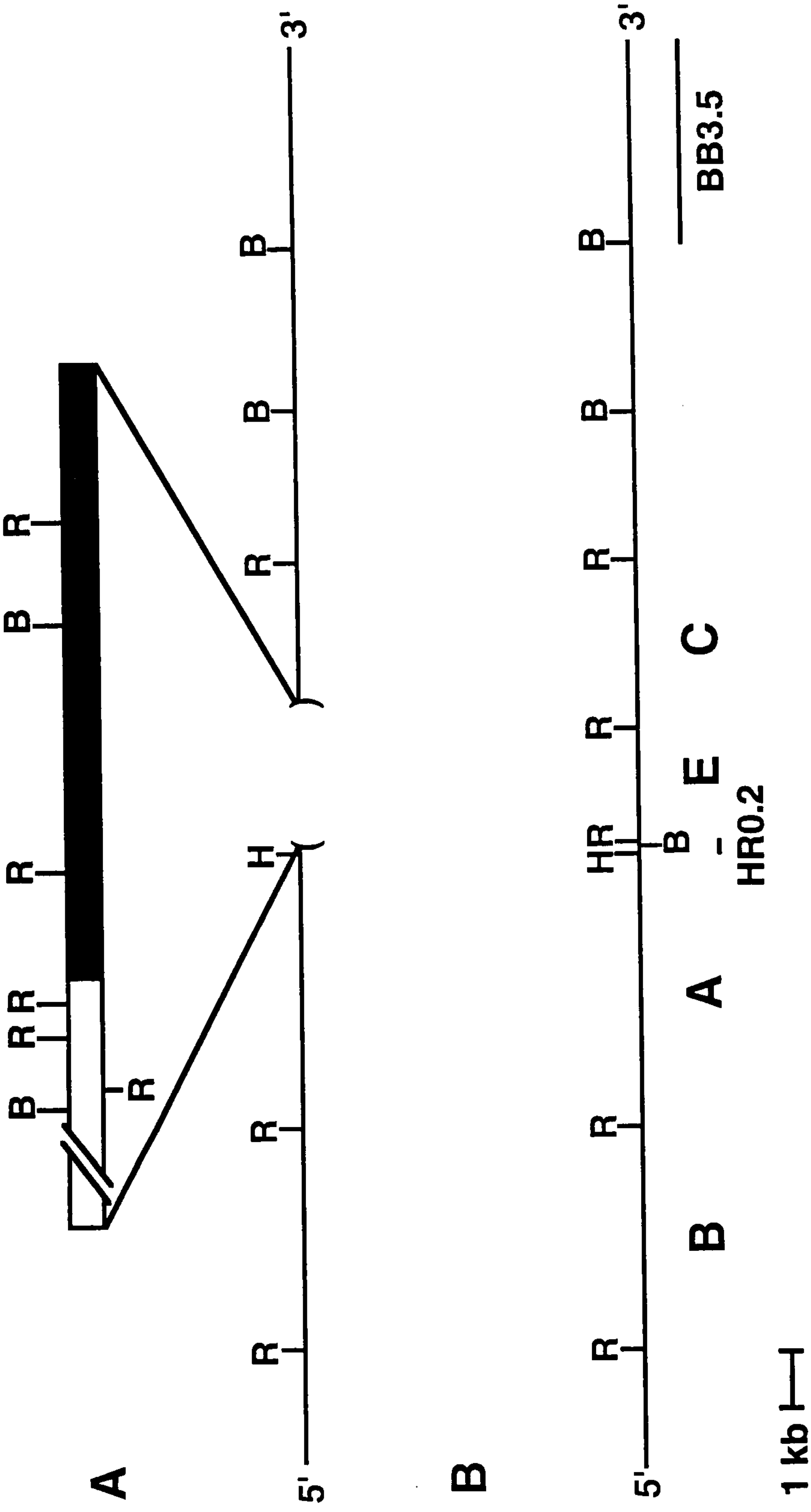
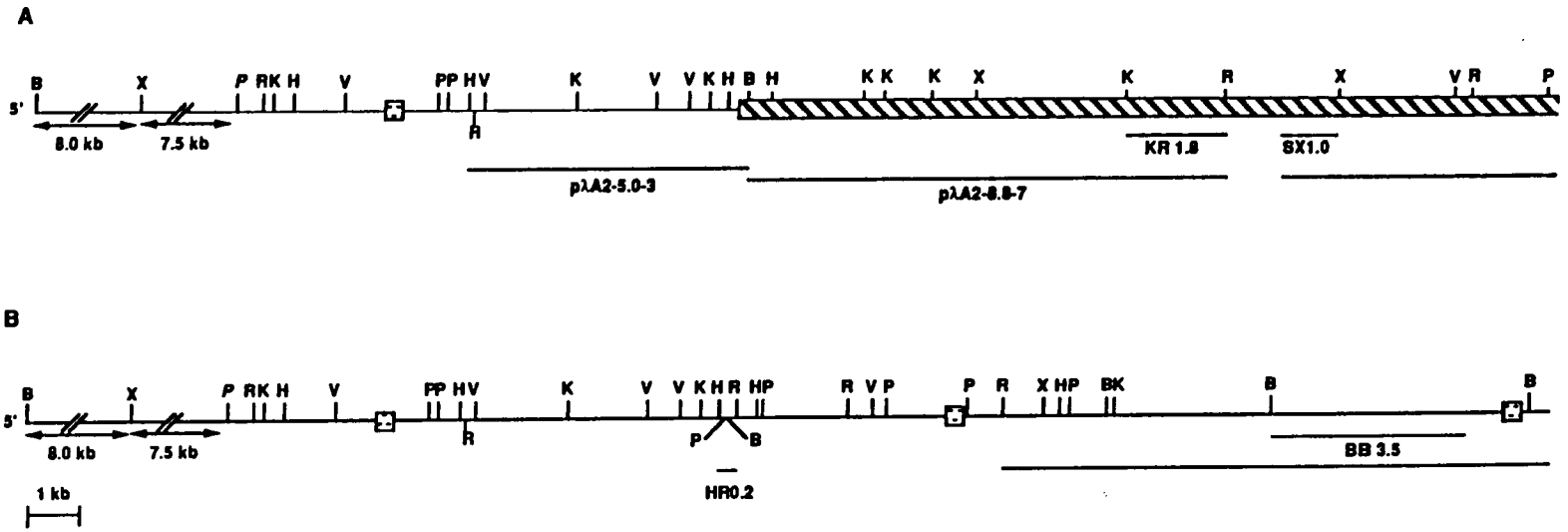

A
B
5'
3'
R
B
H
HR
HR0.2
BB3.5
A
B
E
C
1 kb

Figure 4. Detailed restriction map of the *TgN737Rpw* mutant locus and corresponding wild-type locus.

(A) *TgN737Rpw* locus with transgene sequences (black box), cointegrated DNA (hatched box) and regions showing approximate locations of three exons from the *TgN737Rpw* gene (stippled boxes). Approximately 21.0 kb of cointegrating DNA were identified flanking the 5' end of the transgene sequences. pλA2-5.0-3 and pλA2-8.8-7 are shown relative to clone 20-4. The junction breakpoint between the cointegrated and genomic DNA is 174 bp 5' from the corresponding wild-type *Eco*RI site. KR1.8 and SX1.0, critical probes in the elucidation of the size of the cointegrated DNA, are shown, in addition to fragments corresponding with BB3.5, HRA, B1-A, B1-B and B1-C.

(B) Wild-type locus, including the region represented by cosmid clone c2RBA2, which was isolated with probe BB3.5. Italicized restriction sites indicate approximate locations. Regions showing approximate locations of three exons from the TgN737Rpw gene are shown (stippled boxes). Restriction enzymes: *Bam*HI (B), *Hind*III (H), *Kpn*I (K), *Pst*I (P), *Pvu*II (V), *Eco*RI (R) and *Xba*I (X).


A
B X P RK H V PP HV K V V K H B H K K K X K R X V R P
5'
8.0 kb
7.5 kb
R
KR 1.8
SX1.0
pλA2-5.0-3
pλA2-8.8-7
B
B X P RK H V PP HV K V V K H R HP R V P P R X HP BK B B
5'
8.0 kb
7.5 kb
R
P
B
BB 3.5
HR0.2
1 kb

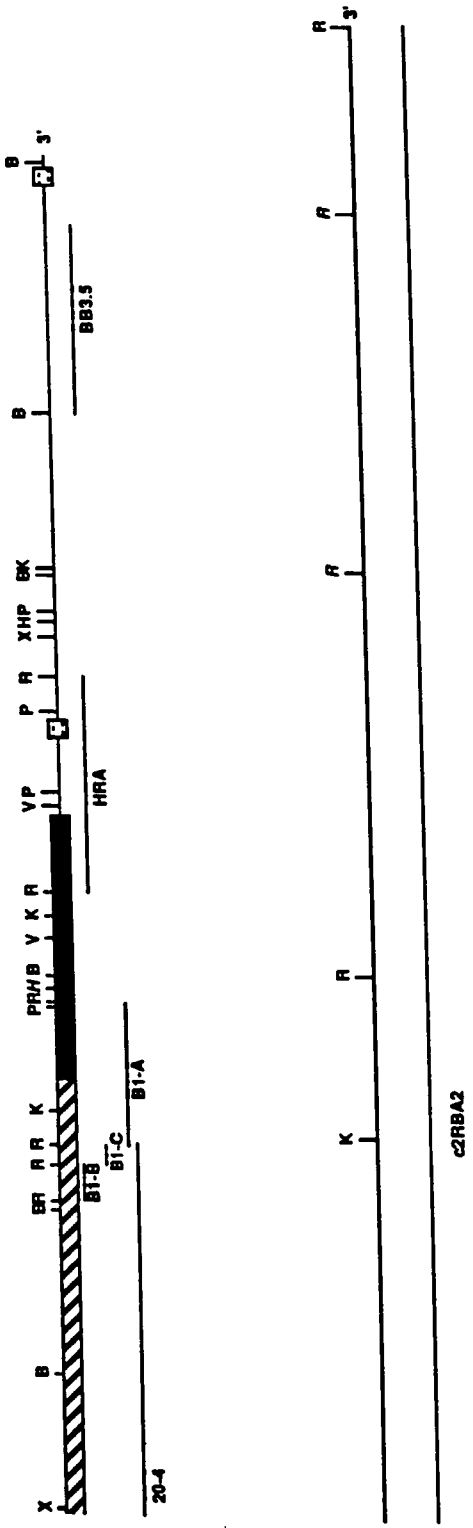
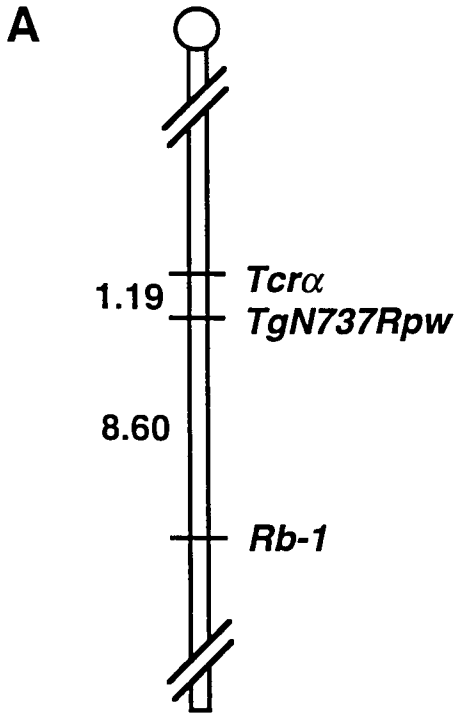


Figure 4 (continued).

Figure 5. Genetic map of the *TgN737Rpw* locus based on interspecific backcross analysis.

(**A**) The position of the *TgN737Rpw* locus on mouse chromosome 14 relative to the T cell receptor α *(Tcrα)* and retinoblastoma-1 *(Rb-1)* genes. HRA was used as a probe to identify the *TgN737Rpw* locus. Distances between loci are indicated on the left in centiMorgans.

(**B**) Summary of *Mus musculus/Mus spretus* interspecific backcross analysis results. The segregation pattern places the *TgN737Rpw* locus on chromosome 14 between the *Tcrα* and *Rb-1* genes. Boxes in columns represent the chromosomal linkage pattern inherited by backcross progeny that could be scored for all 3 markers. Backcross experiments were conducted as described (Stubbs *et al.*, 1990). The number of mice that inherited a specific haplotype is given at the top of each column. Open boxes represent *M. musculus* alleles; filled boxes represent *M. spretus* alleles.



B

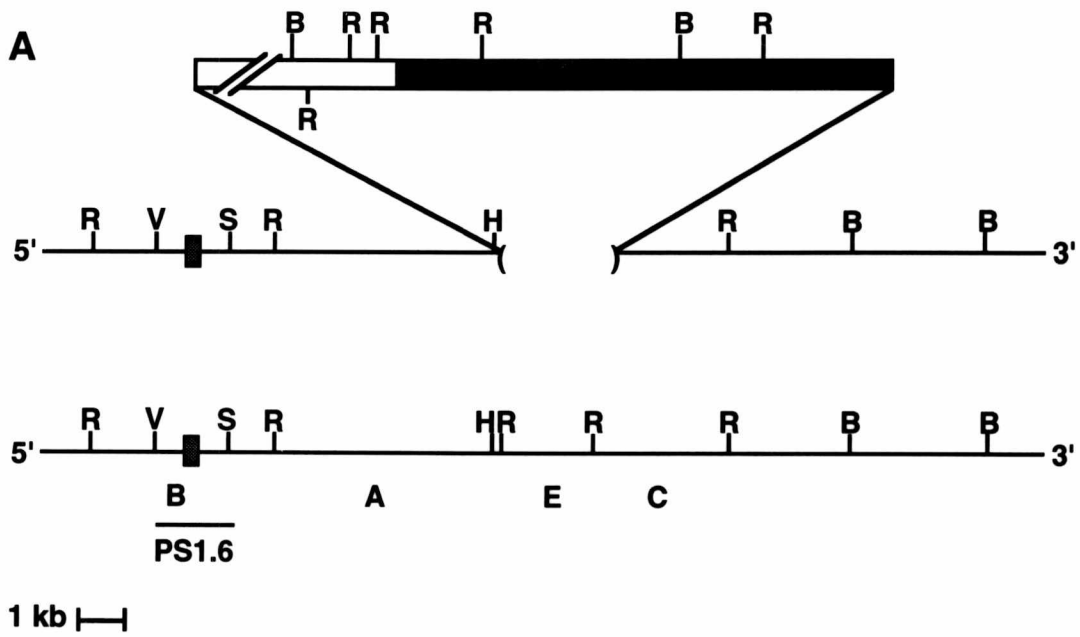
	Number of Backcross Animals							
	37	23	3	2	1	0	0	0
Tcrα	■	□	■	□	■	□	■	□
TgN737Rpw	■	□	■	□	□	■	□	■
Rb-1	■	□	□	■	□	■	■	□

Figure 6. Identification of evolutionarily-conserved DNA sequences within PS1.6.

(A) The 1.6-kb *Pvu*II/*Sac*I fragment PS1.6, derived from the *Eco*RI fragment corresponding with fragment B (fragments A, B, C and E shown under the locus), was used in cross-species hybridization experiments. The approximate location of the exon within PS1.6 is indicated.

(B) Hybridization of PS1.6 with DNA from dogs and humans under low stringency conditions. Final wash: 0.2 X SSC, 0.1% SDS at 50°C.

Restriction enzymes: *Bam*HI (B), *Eco*RI (R), *Hind*III (H), *Pst*I (P), *Pvu*II (V) and *Sac*I (S).


A
B R R R B R
R
R V S R H R B B
5'
3'
R V S R H R R R B B
5'
3'
B
A
E
C
PS1.6
1 kb

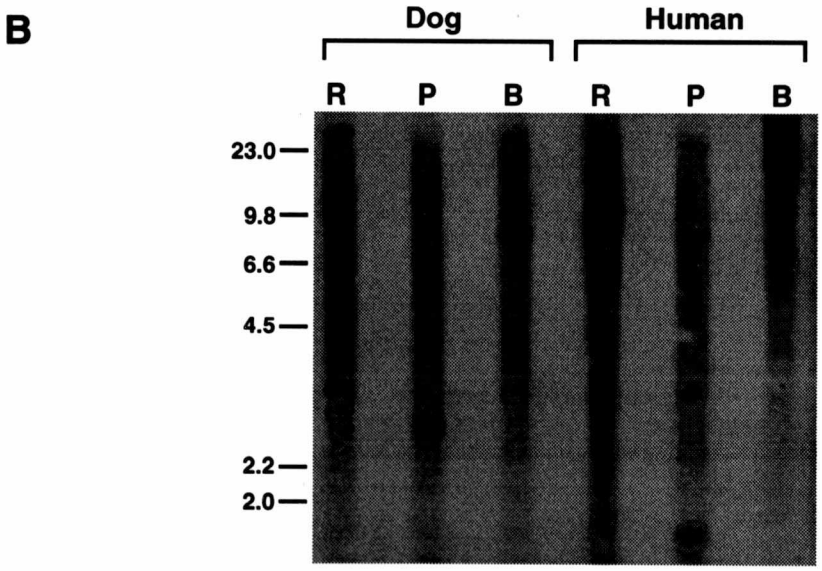

B
Dog
Human
R P B R P B
23.0
9.8
6.6
4.5
2.2
2.0

Figure 7. cDNA clones from the *TgN737Rpw* gene.

Restriction map analysis of six clones indicated that all represented overlapping versions of the same gene. The vertical dashed line aligns the clones at the poly(A) tail junction; differences at the 3' ends among clones reflect different poly(A) tail lengths. Restriction enzymes sites: *Pvu*II (V); *Bam*HI (B); *Pst*I (P). The asterisk (*) refers to the absence of a 111-bp exon from cDNA12. It has not been determined whether cDNA7 contains the 111-bp exon.

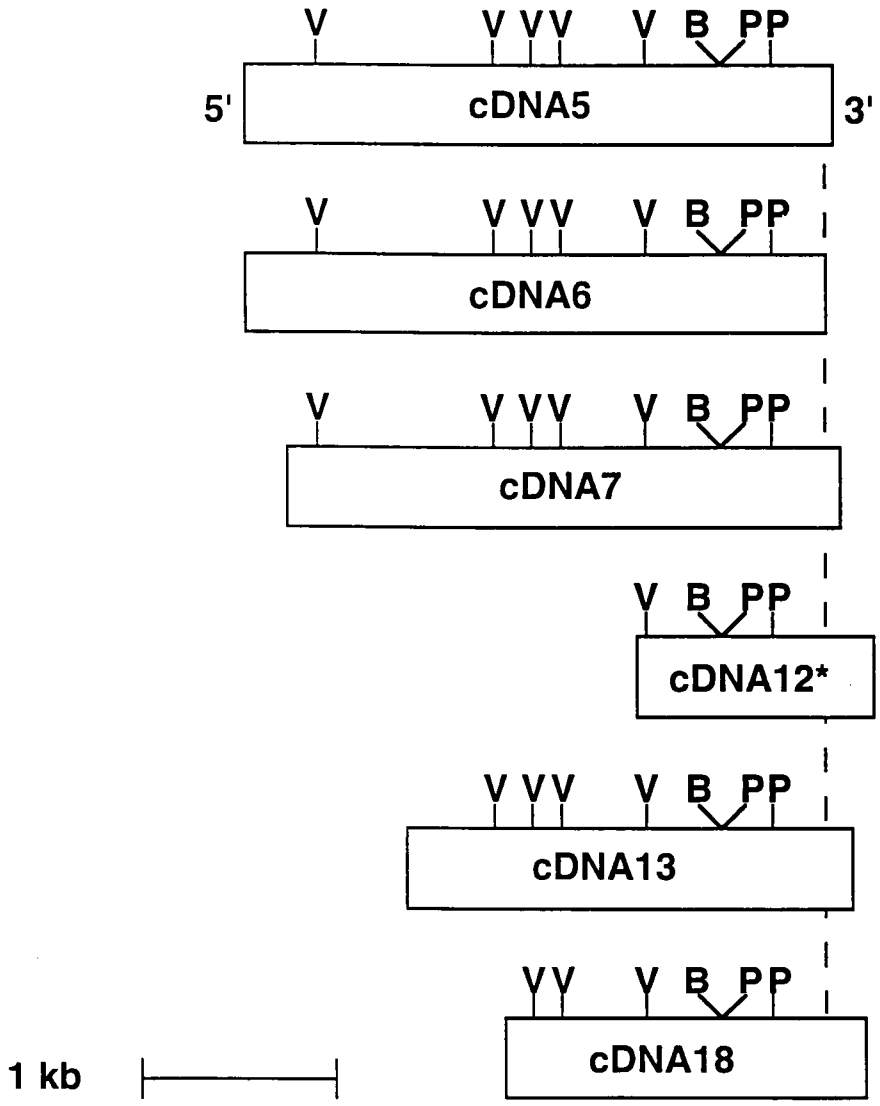
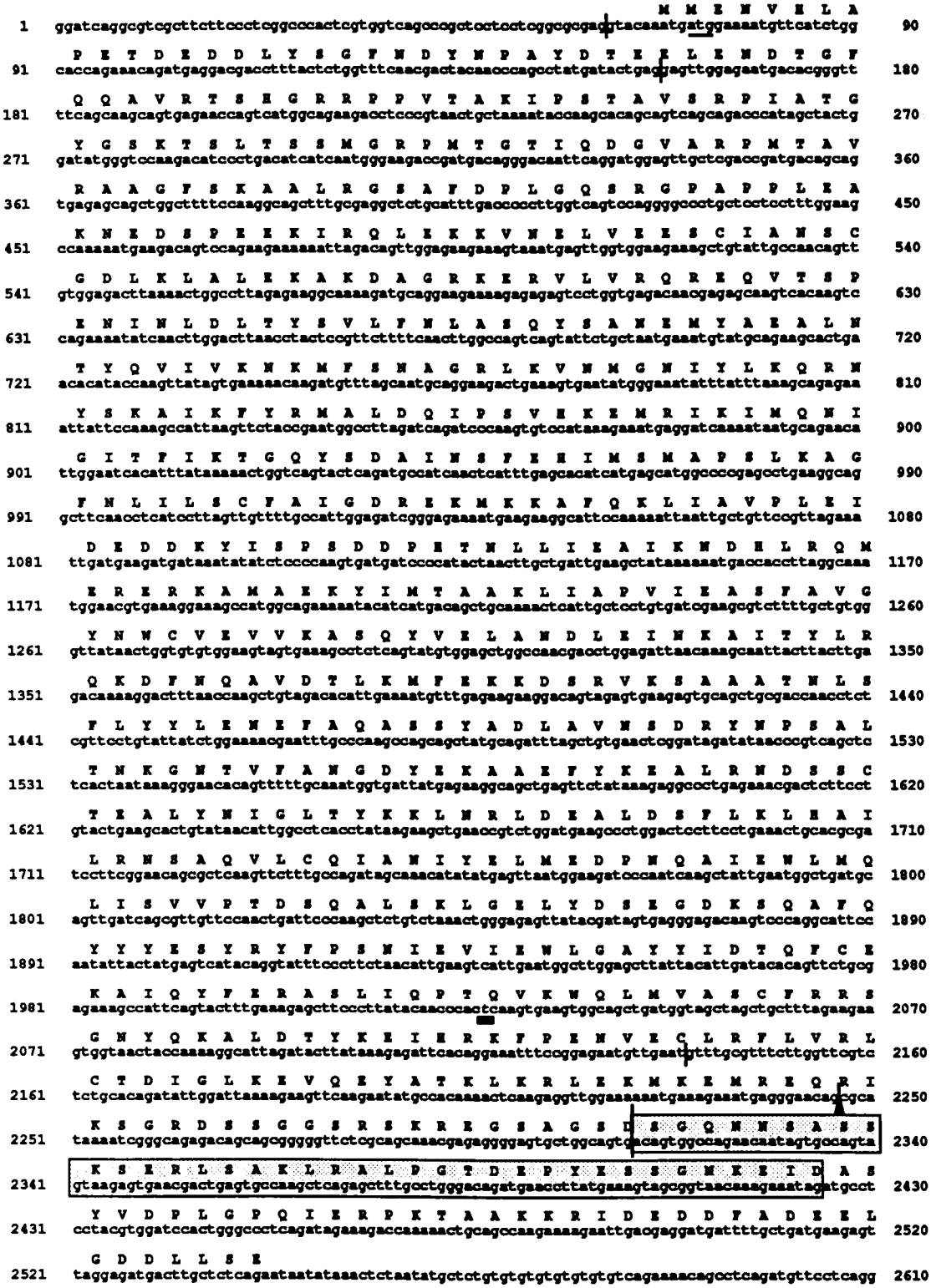

V
V V V
V
B
PP
5'
cDNA5
3'
cDNA6
cDNA7
cDNA12*
cDNA13
cDNA18
1 kb

Figure 8. Nucleotide and deduced amino acid sequences of cDNA5 from the *TgN737Rpw* gene.

The second ATG (underlined) within the 3024-bp cDNA is the putative start site, based on the Kozak context (Kozak, 1991). Vertical bars indicate known exon junctions. The transgene integration site occurred between two exons (arrow). The 111-bp region which is deleted from cDNA12 is shown (shaded box). The "silent" point mutation at base 2027 is italicized and underlined with a thick bar, and the polyadenylation signal is double-underlined.



2611 caccatccatcacctttttacatttgtggcagggttcctctggcctggaacttgccatataggcaagactggctggctggccaccccatg 2700

2701 ggtctgcttgtctctgctccctggctgcagagatttataagcacttgccaccacatccagctttctttaaatggactctagggcttgaac 2790

2791 tcaggtcctccacttggcaaggcaagcacttttccaattgtgctctctctagccaaataagtcatttaccaaagaaaagaaattgcctta 2880

2881 taaggtaatattcatgttaaacctggatcaaatatccaaactttaatttttgtacactggtgaaactttcaataagtgtgttttgttctct 2970

2971 gaggatgggtttgtgctgtcatttttttaaaatggaataaaaactatgaagctgc

Figure 8 (continued).

Figure 9. Predicted amino acid sequence of the *TgN737Rpw* cDNA and identification of tetratricopeptide repeat (TPR) domains.

(A) Amino acid sequence of the *TgN737Rpw* gene. The predicted sequence of the longest open reading frame is 2473 nucleotides (824 amino acids) long. The second in-frame ATG was used as the putative start site since it is in a better agreement with the Kozak consensus (Kozak, 1991). The cDNA nucleotide sequence has been deposited in GenBank (accession number L31959).

(B) The TPRs begin at amino acid 196, and amino acid numbers corresponding with the TPRs are shown on the left. A comparison of the ten TPRs with a unified consensus for several TPR proteins (Hirano *et al.*, 1990) is shown. Circled residues identify matches with conserved amino acids (asterisks represent large hydrophobic amino acids). The *TgN737Rpw* consensus shows the residues found most often at the designated positions.

(C) Schematic representation of the location of TPR motifs in the predicted protein. Each box represents a 34-amino acid repeat, and the shaded boxes represent the TPRs that most closely match the unified consensus. The 34-amino acid repeats of TPRs 1, 2, 3 and 4 are separated by 2, 5 and 145 amino acids, respectively, whereas the repeats in the remaining TPRs are uninterrupted.

A

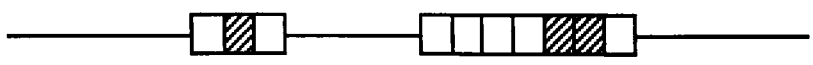
```
  1  MENVHLAPET DEDDLYSGFN DYNPAYDTEE LENDTGFQQA VRTSHGRRPP
 51  VTAKIPSTAV SRPIATGYGS KTSLTSSMGR PMTGTIQDGV ARPMTAVRAA
101  GFSKAALRGS AFDPLGQSRG PAPPLEAKNE DSPEEKIRQL EKKVNELVEE
151  SCIANSCGDL KLALEKAKDA GRKERVLVRQ REQVTSPENI NLDLTYSVLF
201  NLASQYSANE MYAEALNTYQ VIVKNKMFSN AGRLKVNMGN IYLKQRNYSK
251  AIKFYRMALD QIPSVHKEMR IKIMQNIGIT FIKTGQYSDA INSFEHIMSM
301  APSLKAGFNL ILSCFAIGDR EKMKKAFQKL IAVPLEIDED DKYISPSDDP
351  HTNLLIEAIK NDHLRQMERE RKAMAEKYIM TAAKLIAPVI EASFAVGYNW
401  CVEVVKASQY VELANDLEIN KAITYLRQKD FNQAVDTLKM FEKKDSRVKS
451  AAATNLSFLY YLENEFAQAS SYADLAVNSD RYNPSALTNK GNTVFANGDY
501  EKAAEFYKEA LRNDSSCTEA LYNIGLTYKK LNRLDEALDS FLKLHAILRN
551  SAQVLCQIAN IYELMEDPNQ AIEWLMQLIS VVPTDSQALS KLGELYDSEG
601  DKSQAFQYYY ESYRYFPSNI EVIEWLGAYY IDTQFCEKAI QYFERASLIQ
651  PTQVKWQLMV ASCFRRSGNY QKALDTYKEI HRKFPENVEC LRFLVRLCTD
701  IGLKEVQEYA TKLKRLEKMK EMREQRIKSG RDSSGGSRSK REGSAGSDSG
751  QNNSASSKSE RLSAKLRALP GTDEPYESSG NKEIDASYVD PLGPQIERPK
801  TAAKKRIDED DFADEELGDD LLSE
```

B

```
(196-229)  1-Y S V (L) F N (L) A S Q (Y) S A N E M (Y) A E (A)(L) N T (Y) Q V I (V) K N K M F S
(232-265)  2-G R L K V N (M)(G) N (I)(Y) L K Q R N (Y) S K (A)(I) K (F)(Y) R M (A)(L) D Q I (P) S V
(271-304)  3-I K I (M) Q N (I)(G) I T (F) I K T G Q (Y) S D (A)(I) N S (F) E H I (M) S M (A)(P) S L
(450-483)  4-S A A (A) T N (L) S F (L)(Y) Y L E N E (F) A Q (A) S S (Y) A D L (A)(V) N S D R Y N
(484-517)  5-P S A (L) T N K (G) N T V F A N G D (Y) E K (A)(A) E (F)(Y) K E (A)(L) R N D S S C
(518-551)  6-T E A (L) Y N (I)(G) L T (Y) K K L N R (L) D E (A)(L) D S (F) L K L H A (I) L R N S
(552-585)  7-A Q V (L) C Q (I) A N (I)(Y) E L M E D (P) N Q (A)(I) E (W) L M Q L (I) S (V) V (P) T D
(586-619)  8-S Q A (L) S K (L)(G) E (L)(Y) D S E G D K S Q (A)(F) Q (Y)(Y) Y E S (Y) R (Y) F (P) S N
(620-653)  9-I E V (I) E W (L)(G) A (Y)(Y) I D T Q F C E K (A)(I) Q (Y)(F) E R (A) S L (I) Q (P) T Q
(654-687) 10-V K W Q L M (V) A S C (F) R R S G N (Y) Q K (A)(L) D T (Y) K E I H R K F (P) E N

UNIFIED CONSENSUS   - - - * - - * G - * Y - - - - - * - - A * - * F - - A * - * - P - -
                                                  F                   Y
TgN737Rpw CONSENSUS - - A L - N L G N - Y - K - G D Y S K A I - Y Y - E A - R I - P S N
```

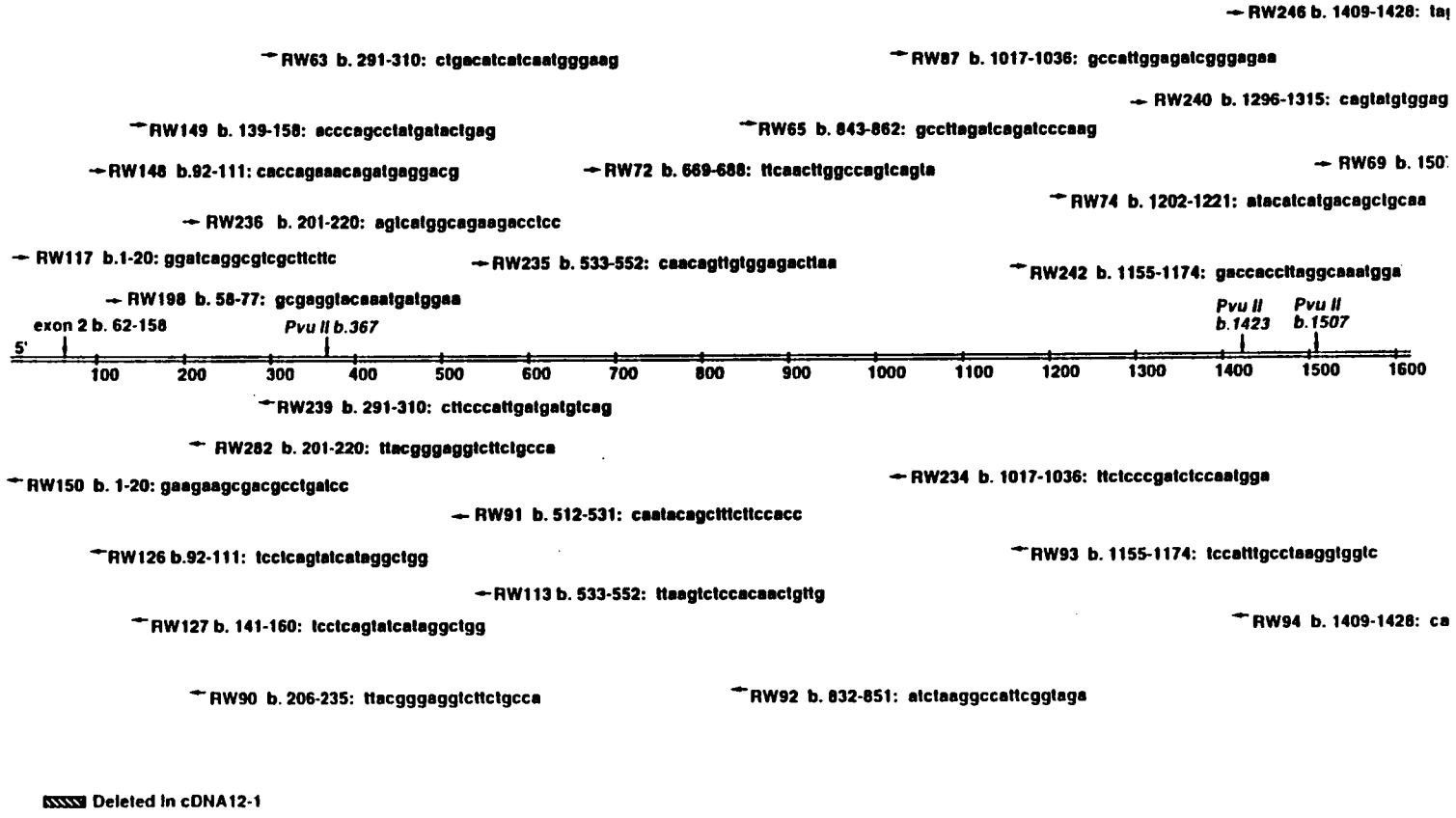
C



100 aa

Figure 10. Oligonucleotides synthesized during the sequencing of cDNA5.

Oligonucleotide primers derived from cDNA5 are shown relative to their positions in the cDNA. Oligonucleotides used to read in the more 3' direction of the cDNA are shown above the cDNA; oligonucleotides from the reverse (antisense) strand, used to read in the 5' direction, are shown below the cDNA. The transgene insertion site (open box) occurred between *Eco*RI fragments fragments B and C, which each contain an exon (exons B and C). Exon 2 is indicated. The 111-bp insert that is absent from cDNA12 is shown (hatched region). The regions corresponding with the beginning of cDNAs 12 and 18 are also indicated, as well as the polyadenylation signal site.


RW246 b. 1409-1428: ta
RW63 b. 291-310: ctgacatcatcaatgggaag
RW87 b. 1017-1036: gccattggagatcgggagaa
RW240 b. 1296-1315: cagtatgtggag
RW149 b. 139-158: acccagcctatgatactgag
RW65 b. 843-862: gccttagatcagatcccaag
RW148 b.92-111: caccagaaacagatgaggacg
RW72 b. 669-688: ttcaacttggccagtcagta
RW69 b. 150
RW74 b. 1202-1221: atacatcatgacagctgcaa
RW236 b. 201-220: agtcatggcagaagacctcc
RW117 b.1-20: ggatcaggcgtcgcttcttc
RW235 b. 533-552: caacagttgtggagacttaa
RW242 b. 1155-1174: gaccaccttaggcaaatgga
RW198 b. 58-77: gcgaggtacaaatgatggaa
exon 2 b. 62-158
Pvu II b.367
Pvu II b.1423
Pvu II b.1507
5'
100
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
RW239 b. 291-310: cttcccattgatgatgtcag
RW282 b. 201-220: ttacgggaggtcttctgcca
RW150 b. 1-20: gaagaagcgacgcctgatcc
RW234 b. 1017-1036: ttctcccgatctccaatgga
RW91 b. 512-531: caatacagctttcttccacc
RW126 b.92-111: tcctcagtatcataggctgg
RW93 b. 1155-1174: tccatttgcctaaggtggtc
RW113 b. 533-552: ttaagtctccacaactgttg
RW127 b. 141-160: tcctcagtatcataggctgg
RW94 b. 1409-1428: ca
RW90 b. 206-235: ttacgggaggtcttctgcca
RW92 b. 832-851: atctaaggccattcggtaga
Deleted in cDNA12-1

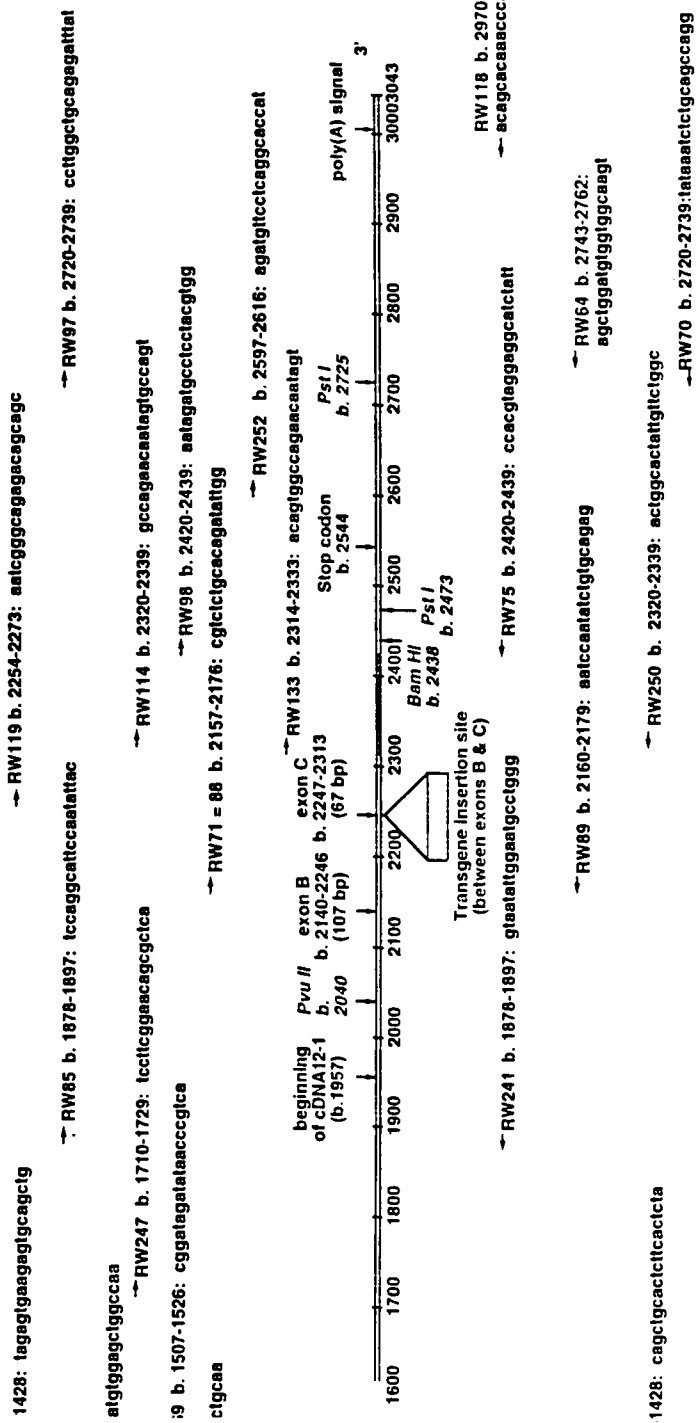


Figure 10 (continued).

Figure 11. The mutant and wild-type *TgN737Rpw* locus, with exons.

(A) The mutant TgN737Rpw locus. The integration site (transgene sequences, black box; cointegrated DNA, open box) is shown relative to the locations of three exons from the *TgN737Rpw* gene. The exons (hatched boxes) are adjacent to each other in cDNA5, and two are contained within DNA corresponding with the *Eco*RI fragments B and C (restriction fragments indicated in italicized, bold lettering under the wild-type locus). The 3'-most exon is a 111-bp region that is deleted from cDNA12.

(B) Corresponding wild-type TgN737Rpw locus.

Restriction enzyme sites: *Bam*HI (B), *Hind*III (H), *Eco*RI (R), *Pst*I (P), *Pvu*II (V), *Sac*I (S).

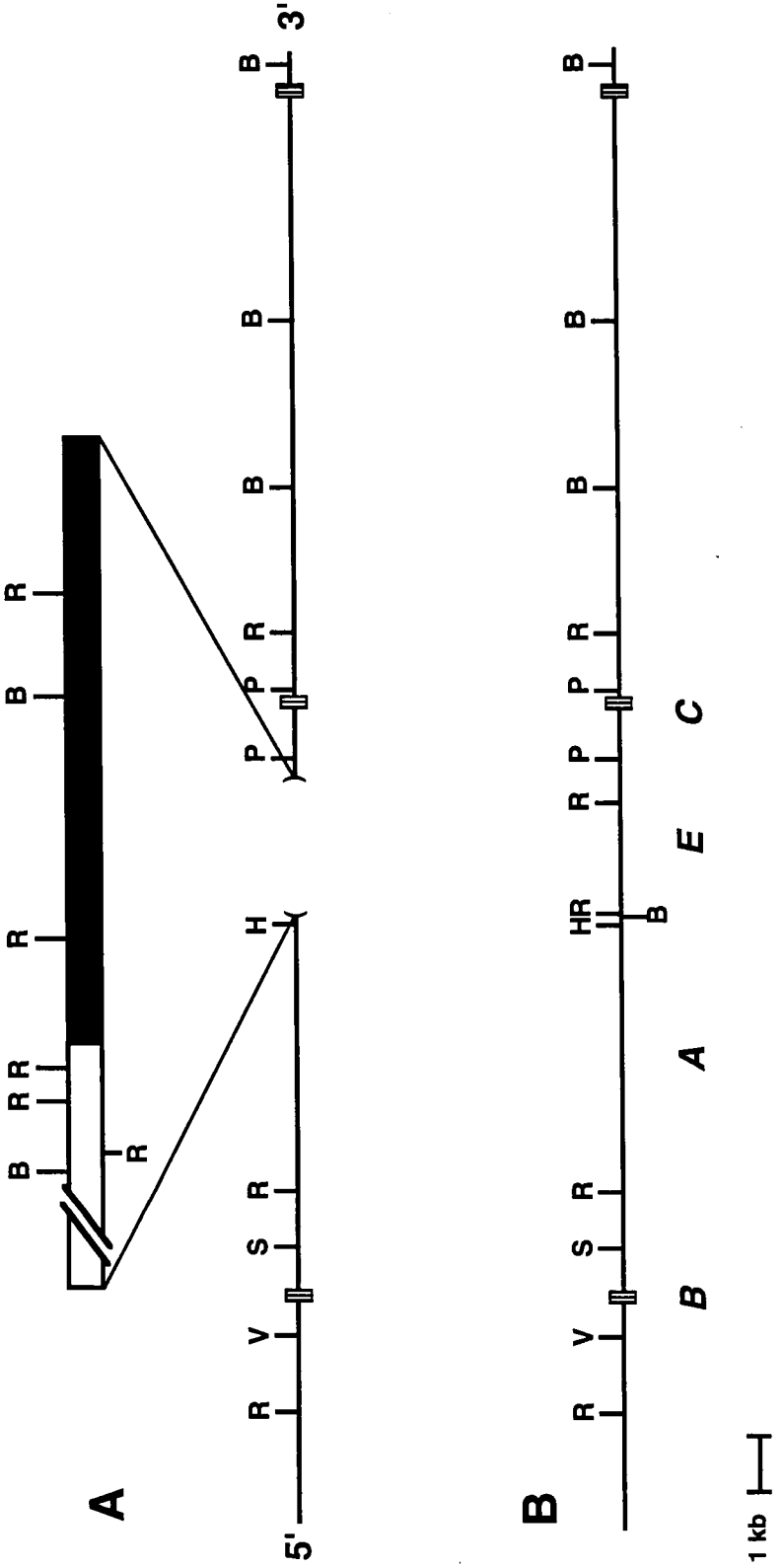

A
5'
3'
B
R
B
R
R
B
R
R
B
R
P
P
H
S
R
V
R
B
B
B
B
R
P
R
P
HR
B
S
R
V
R
A
B
C
E
1 kb

Figure 12. Northern analysis of the *TgN737Rpw* expression pattern.

RNA was poly(A)$^+$-selected in all cases. RNA molecular size standards (left) and the positions of the 28S and 18S ribosomal subunit (right) are indicated. After initial hybridizations, all blots were subsequently hybridized with a chick tubulin probe, and results are shown in the lower panels.

(A) Wild-type and mutant tissue expression pattern. Panels on the left: RNA from wild-type (+/+) skin, day 2; skin, day 3; brain; kidney; lung; liver; salivary gland; small intestine; and spleen. RNA from *TgN737Rpw* homozygotes brain; kidney; liver; salivary gland; and small intestine. Panel on the right includes testis RNA from wild-type (+/+), *TgN737Rpw/+* (Tg/+) and *TgN737Rpw/TgN737Rpw* (Tg/Tg) mice. Probe was radiolabeled cDNA5.

(B) *TgN737Rpw* expression pattern in RNA from whole embryos: days 10, 11, 12, 13, 14, 15, 16 and 18. Probe: radiolabeled cDNA5.

(C) Exons 3' to the integration site are transcriptionally accessible in the mutant. 1. (Tg/Tg) brain; 2. (Tg/Tg) kidney; 3. (Tg/Tg) liver; 4. (Tg/Tg) salivary gland; 5. (Tg/Tg) small intestine; 6. (Tg/Tg) testis; 7. blank; 8. (+/+) brain; 9. (+/+) kidney; 10. (+/+) lung; 11. (+/+) liver; 12. (+/+) salivary gland; 13. (+/+) small intestine; 14. (+/+) spleen; 15. (+/+) testis. Probe: radiolabeled RW114/134, which represents an exon that is probably alternatively-spliced in the mouse (see text).

(D) *TgN737Rpw* expression in cell lines. 1. Embryonic stem cells + mouse fibroblasts ("feeders"); 2. Feeders only; 3. MDCK cells (proliferating); 4. MDCK cells (confluent); 5. LLC-PK_1 cells

(confluent); 6. LLC-PK_1 cells (proliferating); 7. OK cells (confluent); 8. OK cells (proliferating); 9. mouse keratinocyte cells; 10. COS-7 cells.

A

skin d.2
skin d.3
brain
kidney
lung
liver
sal. gland
sm. int.
spleen
+/+
brain
kidney
liver
sal. gland
sm. int.
Tg/Tg
7.46
4.40
2.37
Tubulin
28S
18S
testis
+/+
Tg/+
Tg/Tg
+/+
7.46
4.40
2.37
28S
Tubulin

B

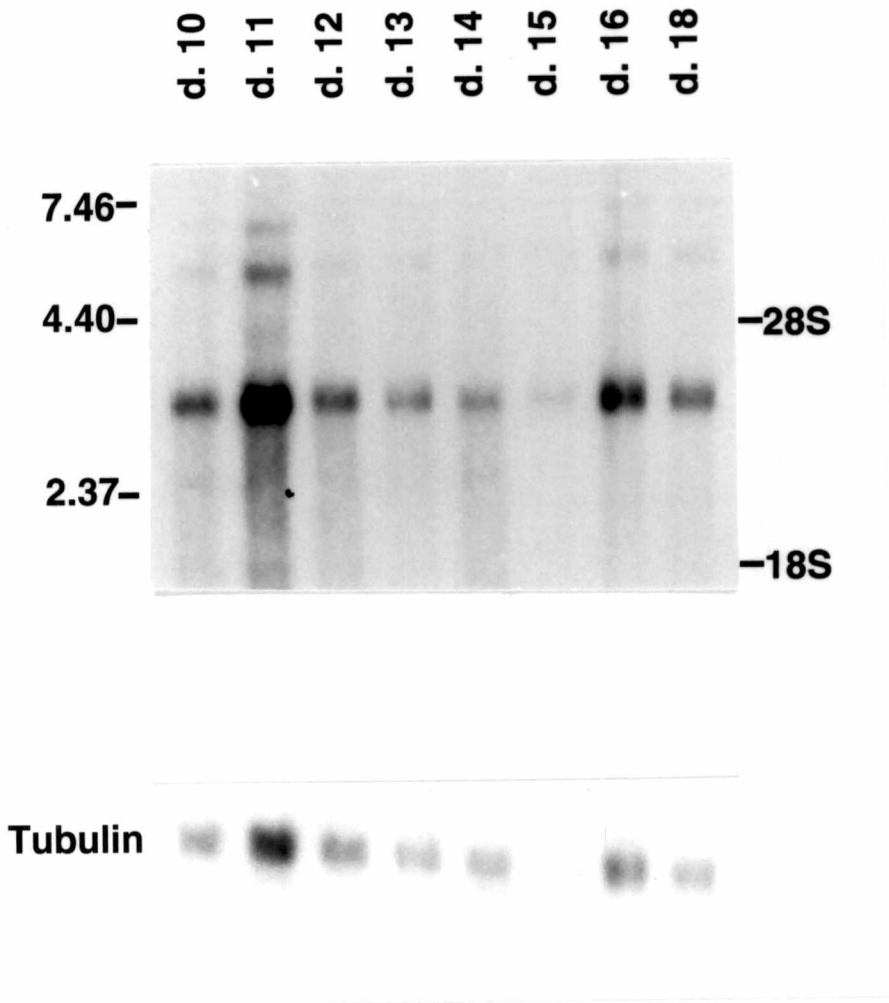


Figure 12 (continued).

C

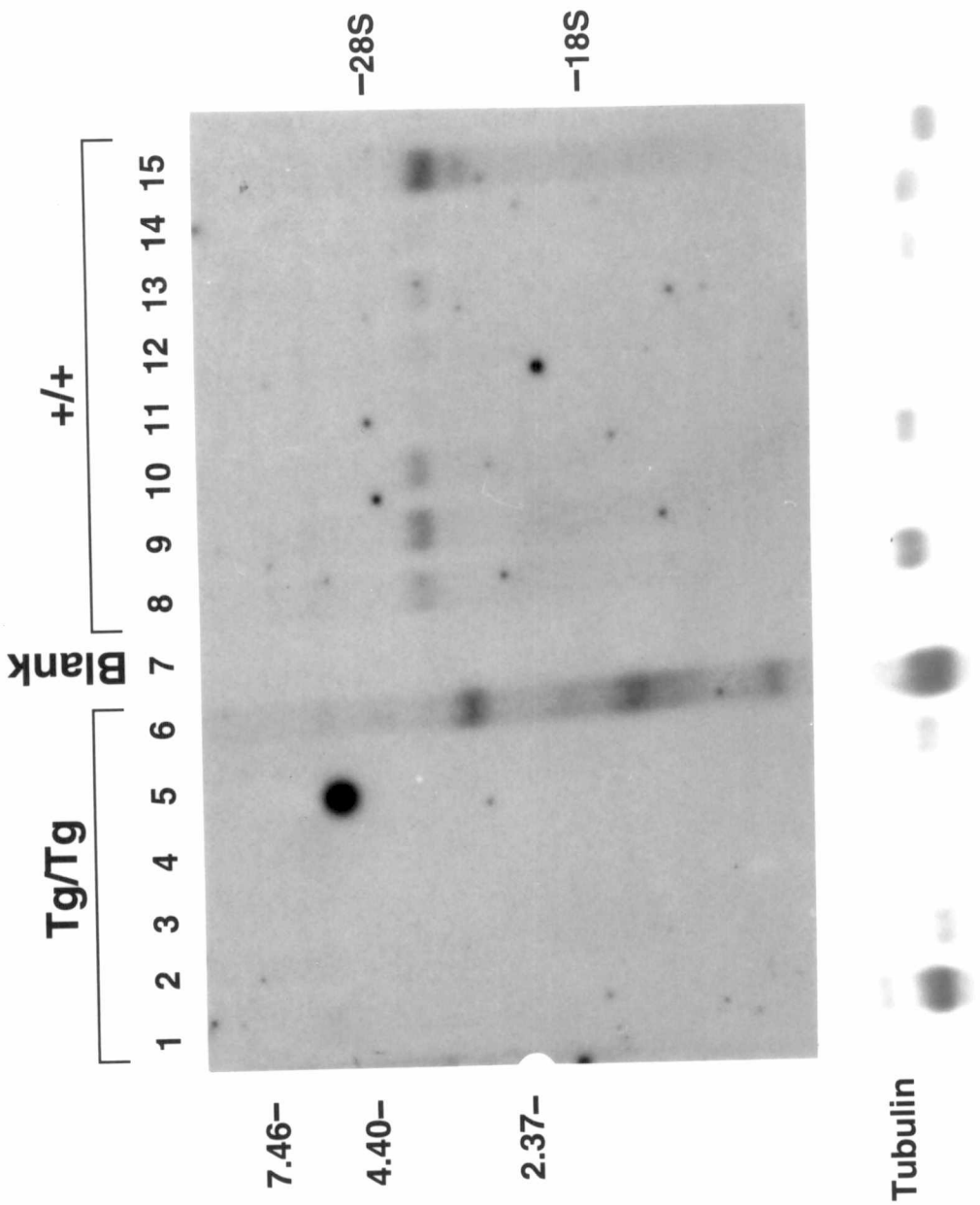


Figure 12 (continued).

D

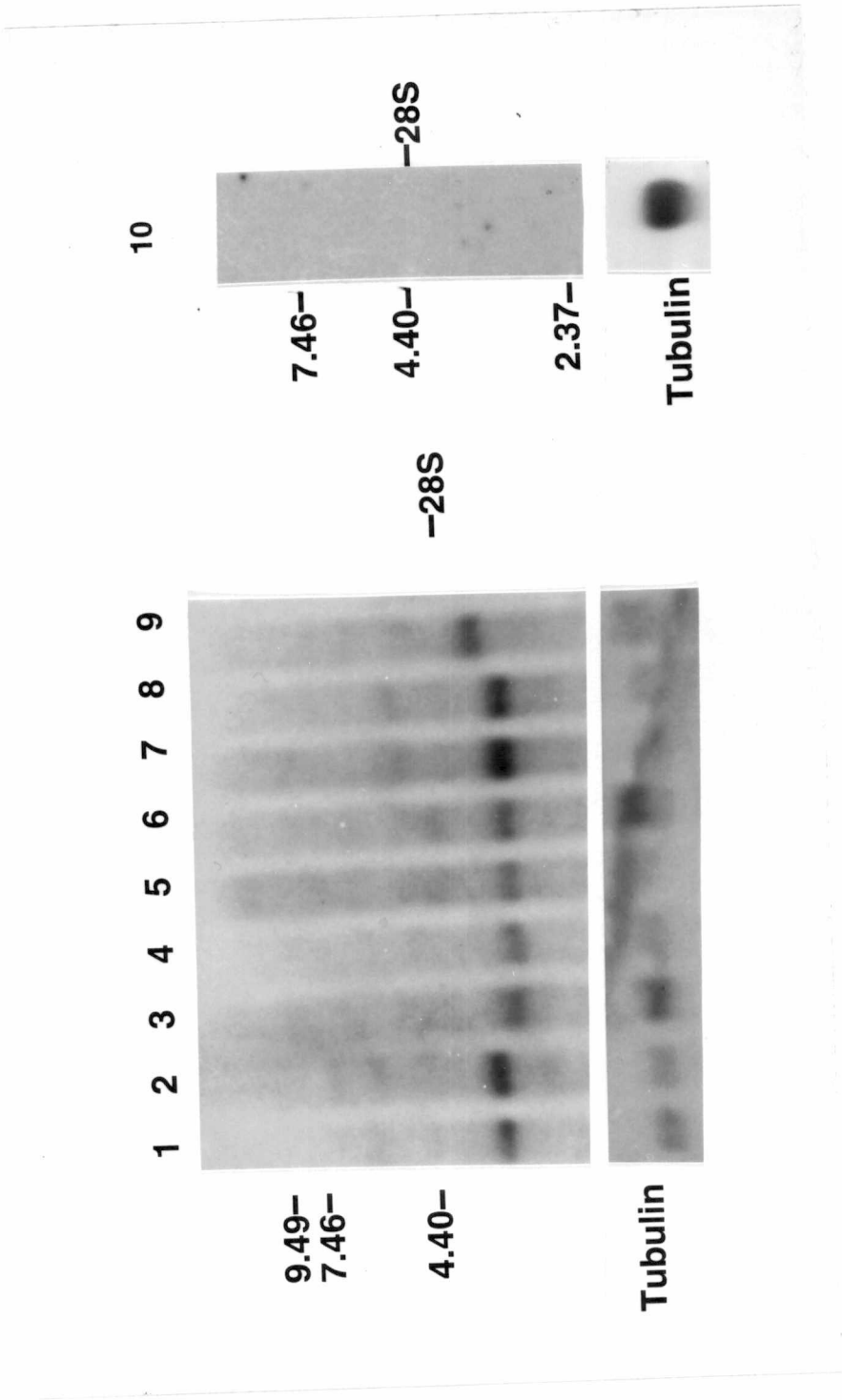


Figure 12 (continued).

Figure 13. Genomic clones representing the 5' region of the *TgN737Rpw* gene.

Clones II(1) and II(2) were isolated from a λ phage library using radiolabeled RW117/126, which includes sequences from the first two exons of cDNA5 (Fig. 10). The insert size is approximately 17.0 kb in clone II(1), and approximately 16.2 kb in II(2). Hatched regions denote approximate locations of exon regions from the *TgN737Rpw* gene, based on hybridization of restriction enzyme-digested clones with exon-specific PCR-amplified probes as well as the full-length cDNA5, and do not necessarily indicate the precise size or location of exons. Sequence analysis demonstrated that the first and second methionines predicted by cDNA5 reside within the second exon (Met). The boxed sequence and an arrow mark the site of the AP1-like binding site (nucleotides -477 through -471); the GC-rich areas are indicated with filled boxes. *Sal*I sites are vector-derived. Restriction enzymes: *Bam*HI (B), *Eco*RI (R), *Hind*III (H), *Pst*I (P), *Pvu*II (V), *Sal*I (L), *Sst*I (S) and *Xba*I (X).

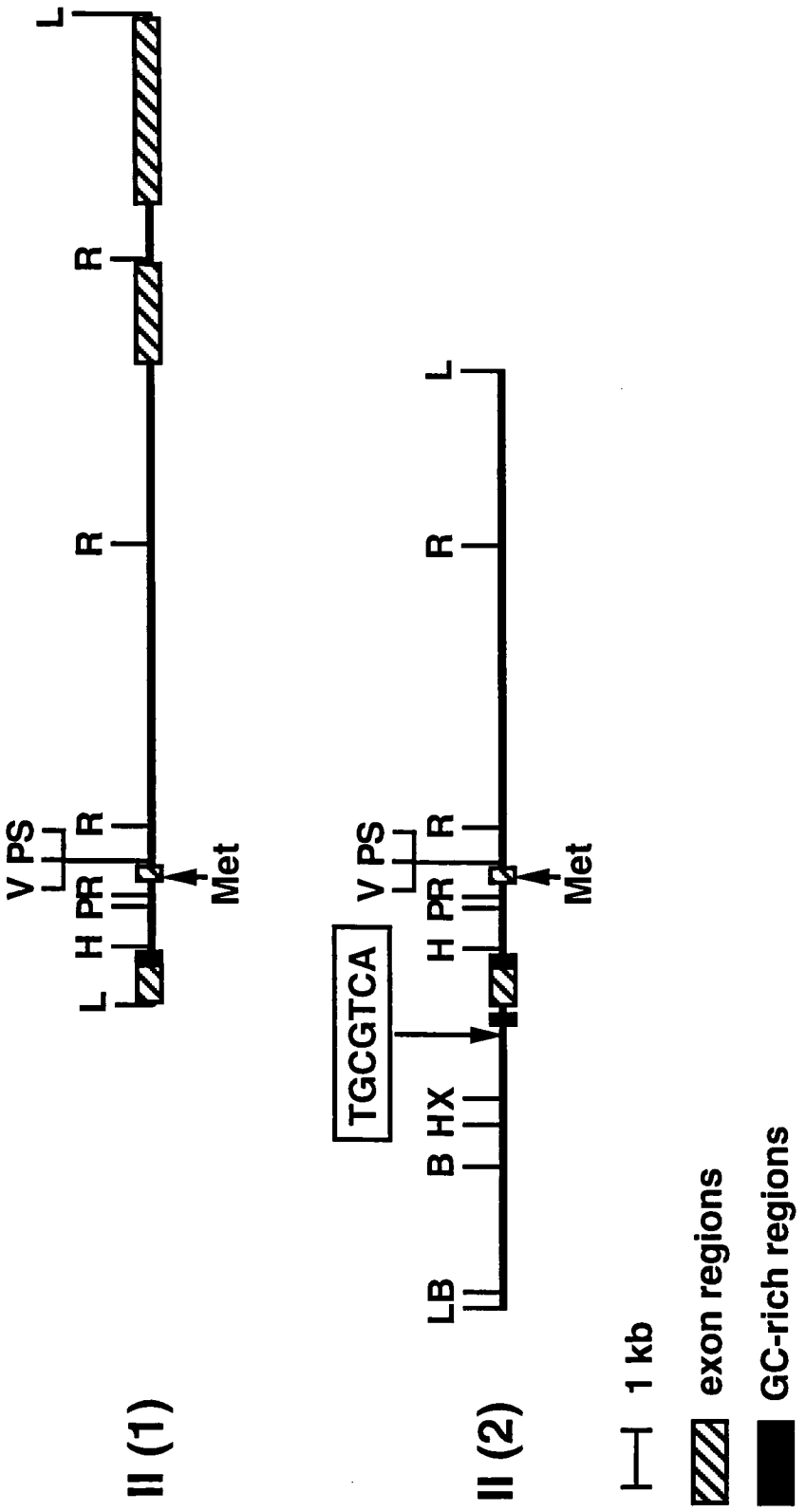

L
R
R
V PS
R
L H PR
Met
II (1)
L
R
R
V PS
R
TGCGTCA
H PR
Met
B HX
LB
II (2)
1 kb
exon regions
GC-rich regions

Figure 14. Homology of the 5' region of the *TgN737Rpw* gene with the promoter and enhancer regions of the murine adenosine deaminase gene.

(A) Regions of *TgN737Rpw* homology with the ADA promoter are in boxes. Nucleotides -175 through -140 relative to the beginning of cDNA5 are shown.

(B) Homology of sequence in the first *TgN737Rpw* intron with the ADA core enhancer region is shown in boxes and corresponds with nucleotides 273-320 relative to the beginning of the first exon of the gene.

A

Minimum promoter element: murine ADA gene

AGCGTGGCGGGCCGCGGTCTCTGAGAGCCATCGGGAAGCGACCCTGCC

***TgN737Rpw* gene**

TAAACATTGGGCC AGCCATCGAGAATCACTCCGCTG

B

Human ADA gene enhancer region

TC**ACATGGCAGTTGGTGGTGGAGGGGAACA**AAGGAGACT GAGTTTCA

***TgN737Rpw* gene**

TCACACGGC TTGGCGATGCATGGCAGCAGTGAA ACTTTTAGAGTTTCA

Figure 15. PFGE analysis of the *TgN737Rpw* locus.

Splenic DNA from a wild-type (+/+) and *TgN737Rpw* hemizygote (Tg/+) was digested to completion with the enzymes as indicated. One (Tg/+) sample was undigested so that the integrity of the DNA could be determined from the ethidium bromide-stained gel (not shown). No hybridization signal is observed in lanes containing DNA digested with *SalI*; most likely, *SalI* cleaves the locus into fragments smaller than the minimum size on contained in the gel. Yeast chromosome molecular weight standards (*S. cerevisiae*) are on the left, in kb.

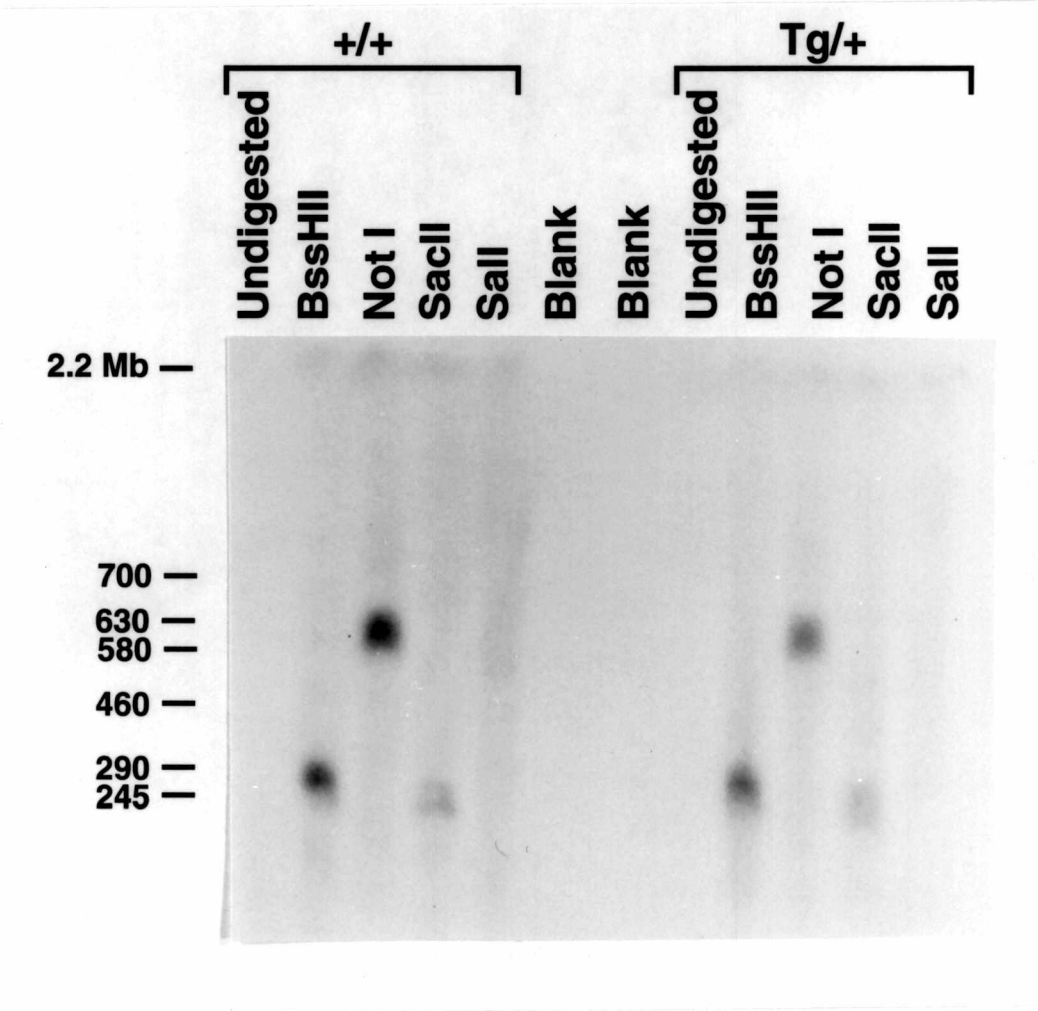

+/+
Tg/+
Undigested
BssHII
Not I
SacII
SalI
Blank
Blank
Undigested
BssHII
Not I
SacII
SalI
2.2 Mb
700
630
580
460
290
245

Figure 16. Complementation testing to determine allelism between the *bpk* and *TgN737Rpw* genes.

An actual pedigree is shown to demonstrate how mouse #37 was identified unequivocally as a compound heterozygote (*bpk*/+; *TgN737Rpw*/+) on the basis of its progeny's phenotypes and *TgN737Rpw* genotype. Stippled regions represent the mutant *bpk* allele (completely stippled circles and squares denote animals with the *bpk/bpk* polycystic kidney phenotype); hatched regions represent the mutant *TgN737Rpw* allele. Six additional progeny in the generation III litter (2 males and 4 females) are not shown; because they were not genotyped for the presence of the transgene, the mice were uninformative for the purposes of this experiment.

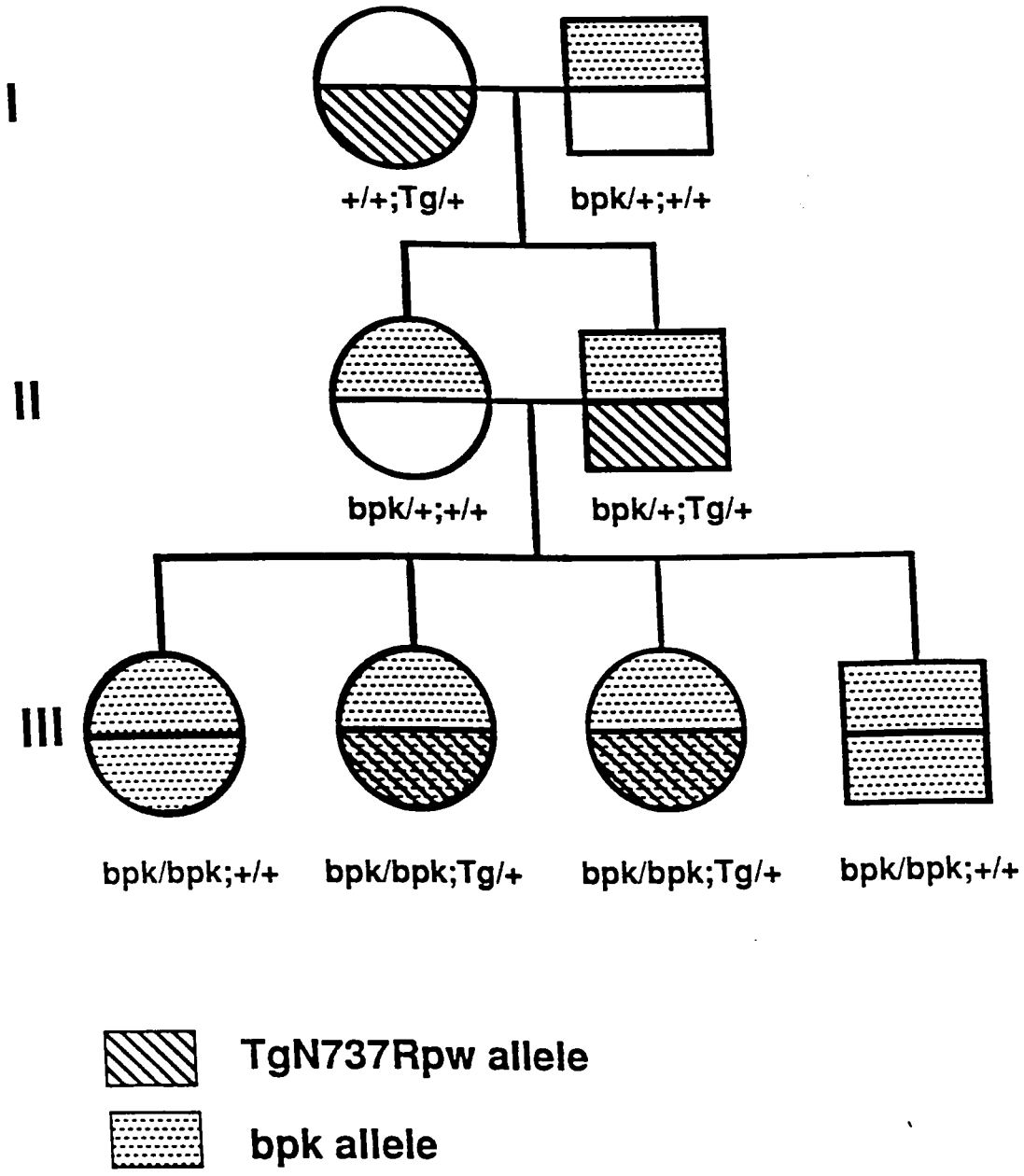
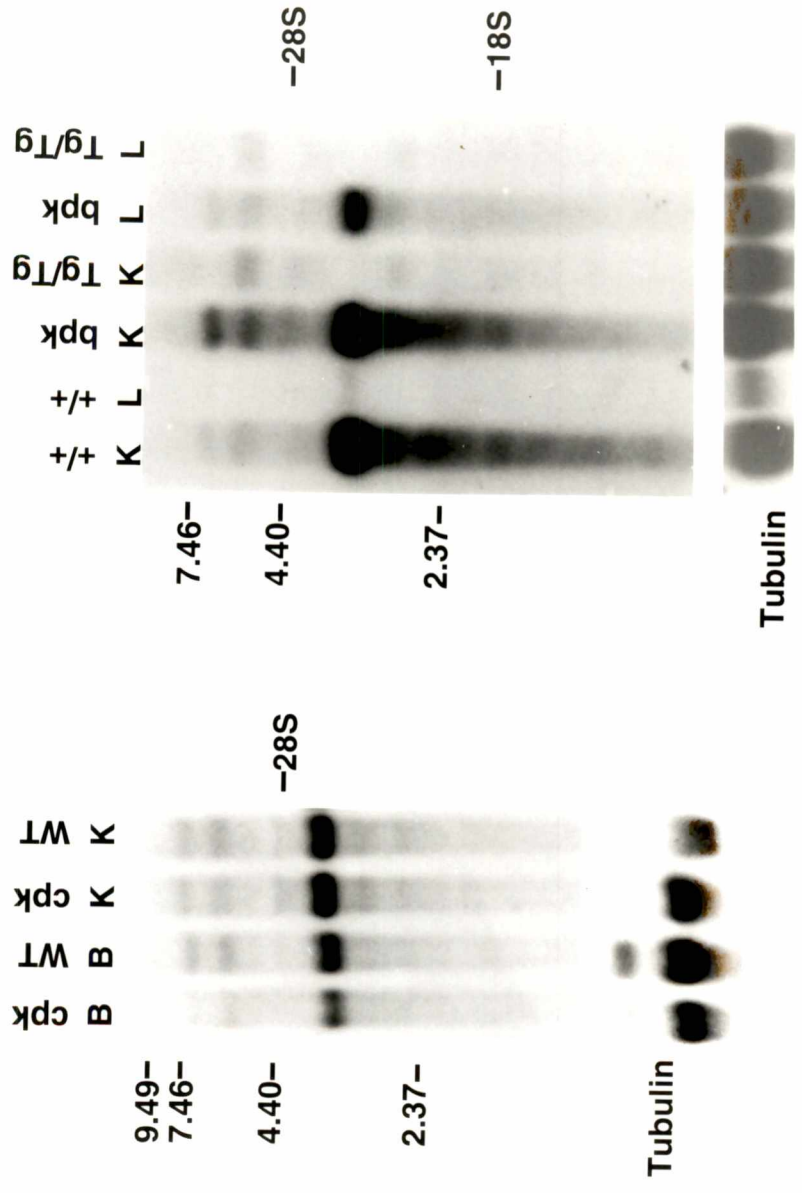

I
II
III
+/+;Tg/+
bpk/+;+/+
bpk/+;+/+
bpk/+;Tg/+
bpk/bpk;+/+
bpk/bpk;Tg/+
bpk/bpk;Tg/+
bpk/bpk;+/+
TgN737Rpw allele
bpk allele

Figure 17. Investigation of allelism of *TgN737Rpw* with *cpk* and *bpk* by Northern analysis.

Left panel: Northern blot with poly(A)$^+$-selected RNA probed with radiolabeled cDNA5. Lanes: 1. (*cpk/cpk*) brain, 2. (phenotypically) wild-type - (*cpk*/+) or (+/+) - brain, 3. (*cpk/cpk*) kidney, 4. wild-type - (*cpk*/+) or (+/+) - kidney.

Right panel: Northern blot with poly(A)$^+$-selected RNA probed with radiolabeled cDNA5. Lanes: 1. wild-type (+/+) kidney, 2. (+/+) liver, 3. (*bpk/bpk*) kidney, 4. (Tg/Tg) kidney, 5. (*bpk/bpk*) liver, 6. (Tg/Tg) liver. The apparent increase in the higher MW transcript in *bpk/bpk* kidney RNA was not observed in another Northern blot that had been prepared with the same RNA samples (data not shown), suggesting that the appearance of a quantitative difference here is artefactual.

The same blots were hybridized with chick tubulin as a control for RNA loading and integrity (lower panels).


cpk B
WT B
cpk K
WT K
9.49–
7.46–
4.40–
2.37–
–28S
Tubulin
+/+ K
+/+ L
bpk K
Tg/Tg K
bpk L
Tg/Tg L
7.46–
4.40–
2.37–
–28S
–18S
Tubulin

VITA

Judith Harbron Moyer was born June 19, 1956 in Hamilton, Ohio to Doris and Elmer Harbron. She was National Honor Society president at Stephen T. Badin High School and received numerous awards in piano composition and performance and English. In 1978 she graduated cum laude from the College of Mount St. Joseph in Cincinnati, Ohio, with the Bachelor of Science degree in nursing. During her college years she was inducted into the Alpha Chi Honor Society and received academic scholarships. In 1979 she moved to Oak Ridge, Tennessee and in 1982 married Bruce A. Moyer, a chemist and fellow musician.

An eleven-year career as a registered nurse in the hospital and public health fields preceded a return to school for Dr. Moyer. In 1989, she entered the University of Tennessee's Graduate School of Biomedical Sciences, and in 1991 she was awarded a predoctoral National Research Service Award from the National Institute of Health's National Center for Nursing Research, for her work under the direction of Dr. Richard P. Woychik. She received the Doctorate of Philosophy degree in Biomedical Sciences in May, 1994.

In the summer of 1994, Dr. Moyer will begin postdoctoral studies in protein biochemistry with Dr. Cynthia B. Peterson at the University of Tennessee. Her professional activities include membership in the American Association for the Advancement of Science and the Association for Women in Science.