Genomics **57,** 333–341 (1999)
Article ID geno.1999.5804, available online at http://www.idealibrary.com on IDEAL®

Contrasting Effects of ENU Induced Embryonic Lethal Mutations of the *quaking* Gene

Roger D. Cox,*[,1] Alison Hugill,* Alexandra Shedlovsky,† Janice K. Noveroske,‡ Steve Best,* Monica J. Justice,‡[,]§ Hans Lehrach,¶ and William F. Dove†

**Wellcome Trust Centre For Human Genetics, Oxford University, Windmill Road, Headington, Oxford OX3 7BN, United Kingdom; †McArdle Laboratory for Cancer Research, University of Wisconsin, 1400 University Avenue, Madison, Wisconsin 53706; ‡The University of Tennessee, Knoxville, Tennessee 37931; §Department of Molecular and Human Genetics, Baylor College of Medicine 410A, One Baylor Plaza, Houston, Texas 77030; and ¶Max Planck Institute for Molecular Genetics, Ihnestrasse 73, D-14195 Berlin (Dahlem), Germany*

Received November 2, 1998; accepted February 25, 1999

Multiple alleles of the quaking (*qk*) gene have a variety of phenotypes ranging in severity from early embryonic death to viable dysmyelination. A previous study identified a candidate gene, QKI, that contains an RNA-binding domain and encodes at least three protein isoforms (QKI-5, -6 and -7). We have determined the genomic structure of QKI, identifying an additional alternative end in cDNAs. Further we have examined the exons and splice sites for mutations in the lethal alleles $qk^{l\text{-}1}$, qk^{kt1}, qk^{k2}, and qk^{kt3}. The mutation in $qk^{l\text{-}1}$ creates a splice site in the terminal exon of the QKI-6 isoform. Missense mutations in the KH domain and the QUA1 domains in qk^{k2} and qk^{kt3}, respectively, indicate that these domains are of critical functional importance. Although homozygotes for each ENU induced allele die as embryos, their phenotypes as viable compound heterozygotes with qk^{v} differ. Compound heterozygous qk^{v} animals carrying qk^{kt1}, qk^{k2}, and qk^{kt3} all exhibit a permanent quaking phenotype similar to that of qk^{v}/qk^{v} animals, whereas $qk^{v}/qk^{l\text{-}1}$ animals exhibit only a transient quaking phenotype. The $qk^{l\text{-}1}$ mutation eliminates the QKI-5 isoform, showing that this isoform plays a crucial role in embryonic survival. The transient quaking phenotype observed in $qk^{v}/qk^{l\text{-}1}$ mice indicates that the QKI-6 and QKI-7 isoforms function primarily during myelination, but that QKI-5 may have a concentration-dependent role in early myelination. This mutational analysis demonstrates the power of series of alleles to examine the function of complex loci and suggests that additional mutant alleles of quaking could reveal additional functions of this complex gene.

[1] To whom correspondence should be addressed. Telephone: 44-1235-834393. Fax: 44-1235834776. E-mail: r.cox@har.mrc.ac.uk.

INTRODUCTION

The mouse *quaking* (*qk*) gene has several recessive alleles causing either embryonic lethality at approximately day 9 *in utero*, $qk^{lethal\text{-}1}$ ($qk^{l\text{-}1}$), qk^{kt1}, qk^{k2}, and qk^{kt3} (Shedlovsky *et al.*, 1988; Justice and Bode, 1986, 1988), or dysmyelination in the central nervous system (CNS) and peripheral nervous system (PNS), resulting in a quaking phenotype accompanied by severe tonic/clonic seizures in viable adult mice (qk^{v}; Sidman *et al.*, 1964). The lethal alleles were induced by *N*-ethyl-*N*-nitrosourea (ENU) while the qk^{v} allele arose spontaneously. Compound heterozygotes of the viable allele with each lethal allele exhibit a dysmyelinating phenotype similar to qk^{v} homozygotes (Justice and Bode, 1988), with the exception of the $qk^{l\text{-}1}$ allele, which exhibits a transient dysmyelinating phenotype (Shedlovsky *et al.*, 1988). Table 1 summarizes the different alleles of the *quaking* gene and their phenotypes. A candidate gene, QKI, encodes a KH domain-containing protein that binds RNA (probably as a dimer) and may be involved in some aspect of RNA metabolism (Ebersole *et al.*, 1996; Zorn and Krieg, 1997; Zorn *et al.*, 1997; Caslini *et al.*, 1997; Zaffran *et al.*, 1997; see Vernet and Artzt, 1997). Transcription of the gene is complex, producing at least three isoforms. The 5-kb isoform encodes a protein called QKI-5 that is localized in the nucleus of myelinating glia. QKI-5 is the only form seen in significant amounts in the early embryo, when homozygotes of the lethal alleles die (Ebersole *et al.*, 1996; Justice and Bode, 1988; Cox *et al.*, 1993, 1994). The 6- and 7-kb isoforms encode QKI-6 and QKI-7, and these are expressed together late in development and during myelination (Ebersole *et al.*, 1996; Hardy *et al.*, 1996). In the CNS these isoforms are located primarily in the cytoplasm of the glia (Hardy *et al.*, 1996).

Probes from this locus revealed that the molecular lesion in the qk^{v} mutation involved a >1-Mb deletion.

0888-7543/99 $30.00

TABLE 1

Phenotypes of Quaking Alleles

	qk^N/qk^v			qk^N/qk^N		Df/qk^N		
Allele	Quaking	Male fertile	Size	Time of embryonic death	Phenotype[a]	Time of embryonic death	Phenotype[a]	Reference
qk^v	Yes	No	Normal	NA	NA	NA	NA	
qk^{k2} or qk^{kt1}	Severe	Yes	Small	E8.5–11.5	Disorganized A-P axis; abnormal somites; cranial defects; heart defects	E9.5	Abnormal somites; cranial defects; heart defects	Justice and Bode (1988, and unpublished observations)
qk^{l-1}	Transient	Yes	Normal	E8.0–9.0	Development stops	ND		Shedlovsky *et al.* (1988); Cox *et al.* (1993; 1994)
qk^{kt4}	Yes	Yes	Normal	ND[b]		E9.5	Abnormal somites; cranial defects; heart defects	Justice and Bode (1988)

Note. qk^N denotes allele of quaking in column 1. Df denotes Tt^{Orl} deletion, which includes Brachyury (*T*) and *qk*. ND, not determined. NA, not applicable.

[a] Defects were examined morphologically. Homozygous embryos did not exhibit a single precise time of death. Disorganized A-P axes occurred only in embryos abnormal at 8.5 dpc. Abnormal morphology and position of somites frequently resulted in neural kinking. Cranial defects included small or open head folds, hemorrhaging, and abnormal ventricle formation. Heart defects included enlarged pericardial sacs.

[b] The qk^{kt1} and qk^{kt4} alleles were induced in t^{w5} chromatin. Mutations carried in *t*-chromatin will exhibit male transmission ratio distortion (about 95% of the gametes carry *t*-chromatin when transmitted from heterozygous males), fail to recombine with wildtype chromosomes, and will carry a resident *t*-lethal gene in *cis* with the new mutation. The qk^{kt1} allele was crossed away from the resident *t*-lethal gene to determine the homozygous embryo lethal phenotype. The resident *t*-lethal mutation prevented an analysis of the qk^{kt4} homozygous phenotype. However, its phenotype was determined in heterozygotes compound with qk^{k2}. The phenotype of compound qk^{k2}/qk^{kt4} heterozygotes was similar to Df/qk^{k2} embryos and was less severe than that of qk^{k2} homozygotes. In contrast, qk^{k2}/qk^{kt1} compound heterozygotes were similar to qk^{k2}/qk^{k2} homozygous embryos.

Thus confirmation of the candidacy of QKI was difficult. However, a nonconservative glutamic acid to glycine change was detected in the qk^{kt4} allele, in a residue conserved in the QUA1 domain of the *Xenopus* homologue and present in all three transcripts (Ebersole *et al.*, 1996; Zorn *et al.*, 1997).

We describe here the genomic structure of the QKI gene as well as the mutations in the qk^{l-1}, qk^{kt3}, and qk^{k2} lethal alleles. In addition, we have screened for mutations in the qk^{kt1} allele and can rule out mutation in the known alternative carboxy-terminal portions of the gene as well as the common coding and splice sequences. This work provides further evidence that the QKI DNA sequence is indeed the *quaking* gene and provides an explanation for the transient dysmyelinating phenotype of compound qk^v/qk^{l-1} heterozygotes.

MATERIALS AND METHODS

Mouse Strain Maintenance

The qk^{k2} allele was originally induced in C57BL/6J DNA by ENU mutagenesis but has since been maintained on a BTBR background linked to tufted (*tf*) for over 20 generations. Stocks are maintained by brother–sister matings, and offspring are genotyped using CA repeat primers D17MIT114 and D17MIT194 (Research Genetics, Huntsville, AL), which closely flank the *quaking* locus and are polymorphic between C57BL/6J and BTBR. For other strains and their maintenance, see Shedlovsky *et al.* (1988).

Isolation of BAC and P1 Clones

Clones were isolated from libraries by PCR screening of DNA pools obtained from Research Genetics (mouse BACs) or from Wellcome Trust Centre for Human Genetics DNA pools (mouse P1s, from a library constructed by Dr. Fiona Francis). Pools were screened according to the manufacturer's instructions. QKI gene PCR primers D and C (exon 3, Fig. 1 and Table 2) were used to isolate several BAC clones including 189e4, 219C22, 263d12, 288I13, 459g13, and 553n22 from the Research Genetics mouse BAC library. A P1 clone was isolated containing Exon I (ICRFP703J01286 (286J1)).

Isolation of D17WTC1

The microsatellite D17WTC1 was isolated from a mouse bacteriophage P1 clone (ICRFP703J243) from a library constructed by Dr. Fiona Francis and screened by hybridization using D17Leh502. The microsatellite was identified from a plasmid sublibrary of the P1 clone screened with simple sequence repeat oligonucleotides. The structure of the repeat unit, confirmed after generation by sequencing of subcloned (pGEM) BTBR PCR product DNA converted to single-stranded template DNA and using Sequenase, is (GT)12(GA)-18AA(GA)7AAG(GA)3G(GA)3AA(GA)3.

PCR amplification was carried out on 100 ng of DNA in standard Cetus buffer with a final concentration of 1.5 mM magnesium chloride, 1.25 mM dNTPs and using 2 U of AmpliTaq in a 50-μl reaction, with 10 pmol of each primer: D17WTC1F TGAGCCATGATAATACT-GGGTCA; D17WTC1R,GATATGTCTTGCAAGTGTAG. Cycling was carried out using a PTC100: 95°C for 1 min, 60°C for 45 s, 72°C for

45 s, for 30 cycles. The product sizes for BTBR and AJ are 215 bp and 220 bp, respectively. To resolve adequately this polymorphism on 3% NuSieve plus 1% LE agarose TAE gels, the PCR product is cut with the restriction enzyme *Ase*I, according to the manufacturer's instructions (NEB). The resulting fragments are 150 + 65 bp from BTBR and 155 + 65 bp from A/J.

Gene Structure Determination

Bubble PCR libraries were prepared by digestion of BAC DNA with 4- and 6-base-cutting restriction enzymes (NEB) followed by ligation of linker sequences incorporating a central noncomplementary region ("bubble") containing the bubble primer sequence, essentially as described by Munroe *et al.* (1994). Libraries were PCR amplified using cDNA-specific primers and bubble primers. PCR products were resolved on agarose gels by electrophoresis, and fragments were cut from the gel and purified using GeneClean III (Bio 101). Purified DNA was sequenced using a Dye Rhodamine *Taq* FS terminator sequencing kit (Perkin–Elmer-ABI) and analyzed on an ABI 377. Exon–intron boundaries were identified by comparison of the sequence generated with cDNA sequence. Splice sites were compared with consensus sequences. New PCR primers were designed to amplify exons and flanking sequence including the splice site from genomic DNA (Table 2); most primers were additionally tagged with M13 sequences according to Perkin–Elmer-ABI instructions, for Dye Primer *Taq* FS sequencing (Perkin Elmer–ABI) on an ABI 377. Identification of heterozygous bases was performed by manual inspection of the ABI fluorescence traces.

PCR Primers Used in Gene Structure Experiments

Primer names are in boldface: **C,** TAT CCC TCA TTG AGC CTT TGC CTC GG AC; **D,** GAG AAT CCT TGG ACC TAG AGG ACT TAC AGC; **E,** TCC TTG GAG CAG CCT GGG CAG TTG CTG CAA G; **G,** ATC AGG CAT GAC TGG CAT TTC AAT CCA CTC; **H,** TCCACATTGACTGGTCGTGTTATAACAGCTG; **K,** AGT TTC TCT TGT AAC TGA ACA ATG GGT C; **L,** CTT ACA TGT ACT AAT CAC TGT GGA AGA TGC; **M,** GGC ATG ACA GCG GTC TGT ATT TGT CTG ATC; **P,** GTG ATT TAA TGT TGG CGT CTC TGT AGG TGC C; **Q,** GAA GTG CAG AAT TGC CTG ACG CGG TGG GAC; **R,** GCA TCT TCC ACA GTG ATT AGT ACA TGT AAG; **QK1A,** GGC CTG AAG CTG GGT TAA TCT AC; **QK1B,** TAC ACA TTG GCA CCA GCT ACA TC; **GMAFTP-F1,** GTG GTA GTG AGG AGA TTG GTA TTA GC; **GAVA71-F1,** GCT GTC AAT GTT TCC CTT ATG TAT G; **GAVA-F2,** AAT GCA TCA TGC ATG AAC CTT CGG TC; **EWIMP-R1,** TTT TTC CAT CAG GTC TGG TGT TGC A; **490F,** AGC TGC GGA GCC TGG AAT AT; **908F,** AGG AGG AGC AAA ATA GAG GC. The Web site used for sequence searching is that for The Institute for Genome Research at http://www.tigr.org/.

RNA Preparation

Poly(A)$^+$ mRNA from T/qk^{kt1}, +/qk^{kt3}, T/qk$^{l\text{-}1}$, qk^v/qk^v and BTBR mice was prepared from whole mouse brains using an Invitrogen FastTrack mRNA 2.0 kit according to the manufacturer's instructions. For qk^{k2}, embryos from a heterozygous cross, qk^{k2} +/+ *tf* × qk^{k2} +/+ *tf*, were obtained at E10 and genotyped by PCR of yolk sac DNA with the same DMIT markers used for qk^{k2} stock maintenance. Poly(A)$^+$ mRNA was isolated from qk^{k2}/qk^{k2} homozygous and wild-type littermate control embryos using a Micro-FastTrack Kit (Invitrogen).

RT-PCR

Primers combinations. Primer names are in boldface: **GKYD F1,** GGT AAG TAT GAT TCC TGT ACT ATG TGA; **GKYD R1,** GAT GAA TTA ATC TCC CAG CAT CCG ATC, PCR conditions f (Table 2), Product size is 357 bp. **F,** GGA GCT TGC AAT TCT GAA TGG CAC CTA CAG; **GKYD R3,** as Table 2; PCR conditions e (Table 2). Product size is 446 bp. **N,** GTA CAC CTA CGC CAG CTG GCC CTA CCA TAA TG; **GMAFTP-R1,** as Table 2, PCR conditions e (Table 2). Product size is 349 bp. **N,** as Table 2; **GAVA R2,** as Table 2, PCR conditions e (Table 2). Product size is 388 bp. **N,** as Table 2; **GAVA R1,** GTA GTG ACT AAA CAA ACA GCA TGT C, PCR conditions e (Table 2). Product size is 832 bp. **490F,** AGC TGC GGA GCC TGG AAT AT; **1033R,** TCA CTT CTT CAA CCG CTC TC, PCR conditions g (Table 2). Product size is 543 bp. **938F,** GGG AGC ATC TAA ATG AAG AC; **1435R,** CAC CAC TGG GTT CAA TAG GG, PCR conditions g (Table 2). Product size is 497 bp. **938F,** as above; **1307R,** GTT CCG TTT GGC ATG ACA GC, PCR conditions g (Table 2). Product size is 369 bp. **1180F,** GCT CCA AGG ATC ATC ACT GG; **5kb1600R,** TAA CTA ATC AGC AGG CAG GC, PCR conditions g (Table 2). Product size is 420 bp. **1180F,** as above; **6kb1464R,** GCC TTT CGT TGG GAA AGC, PCR conditions g (Table 2). Product size is 284 bp. **1180F,** as above; **7kb1531R,** GGT CGT GTT ATA ACA GCT GC, PCR conditions g (Table 2). Product size is 351 bp.

Template and PCR. Poly(A)$^+$ RNA from T/qk^{kt1}, +/qk^{kt3}, T/qk $^{l\text{-}1}$, qk^v/qk^v and BTBR mice was reverse transcribed using random primers (Gibco Choice cDNA cloning kit) and Superscript Reverse Transcriptase II (Gibco) according to the manufacturer's instructions. PCR was carried out using AmpliTaq Gold according to the manufacturer's instructions (Perkin–Elmer Applied Biosytems), using the conditions indicated above for each primer pair. RT-PCR of qk^{k2} RNA was performed with a SuperScript Preamplification System (Gibco BRL) using oligo(dT) primers.

Qk^{k2} cDNA. PCR amplification with *quaking* cDNA-specific primers (above) was as follows: each 25-μl reaction in 1× PCR buffer contained 100 ng of each primer, 1 mM $MgCl_2$, 0.2 mM dNTPs, approximately 10 ng of RT cDNA, and 1.25 U *Taq* polymerase (Gibco BRL). After an initial denaturation step of 3 min at 94°C, reactions were subjected to 30 cycles of 30 s at 94°C, 2 min at 51°C, and 2 min at 72°C. Following a final extension time of 7 min at 72°C, samples were held at 4°C. PCR products were separated in a 0.7% gel (FMC BioProducts) in TAE buffer and purified using a GlassMAX DNA isolation spin cartridge system (Gibco BRL). If PCR amplification was low, a second round of PCR (using the same conditions as above) was performed using the GlassMAX purified product as a template prior to sequencing. Resulting PCR products were sequenced using a Dye Terminator Sequencing Kit (Perkin–Elmer/Applied Biosystems) according to the manufacturer's instructions with the following reaction conditions: 10 s at 96°C, 5 s at 50°C, and 4 min at 60°C with a final hold at 4°C. Products were analyzed on an ABI 373 automated sequencer.

cDNA Library Construction

Poly(A)$^+$ RNA (oligo(dT) cellulose purified from total RNA prepared by the guanidinium thiocyanate and cesium chloride method) was prepared from postnatal 10 to 15 day DBA mouse total brain tissue (peak myelination periods). cDNA was prepared using oligo(dT) priming essentially according to the Choice cDNA cloning kit protocol (Gibco). *Eco*RI linked and size fractionated cDNA was subcloned into Lambda ZAP II vector (Stratagene).

RESULTS

Genetic Mapping of the $qk^{l\text{-}1}$ Gene

Using a nonchimeric (FISH analysis, data not shown) yeast artificial chromosome (YAC 171C12) from the Whitehead Mouse YAC I library, we have extended the contig (Cox *et al.*, 1994) of the qk^v region to cover the entire deletion (over 1 Mb in size) and proximal flanking region. To determine the location of the *quaking* gene relative to D17Leh502, which is near the proximal *quaking* deletion breakpoint, we first isolated a mouse P1 clone and then used it to generate a

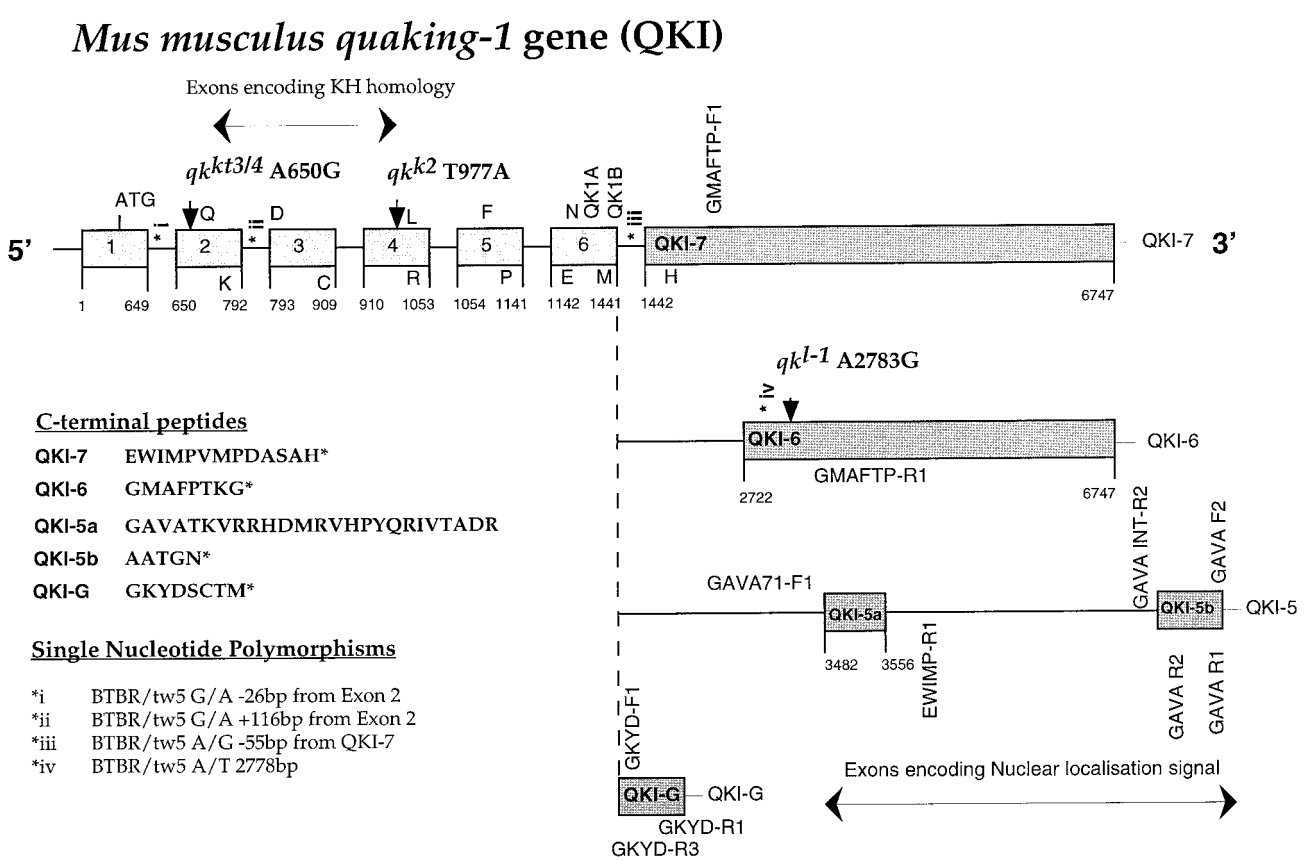

FIG. 1. Structure of the QKI gene and location of primer sequences used in determining the gene structure and RT-PCR. Exonic sequences are open boxes, and the peptide sequences of known alternative ends are shown. Numbers on the bottom line indicate the position of nucleotides in the published cDNA U44940 sequence. Single nucleotide polymorphisms are shown with the base found in BTBR shown first and tw5 second. Letters above and below the exons indicate PCR primer names and their orientations 5′ to 3′ (above, top strand) or 3′ to 5′ (below, bottom strand) relative to the gene, respectively.

polymorphic microsatellite repeat marker D17WTC1. In a genetic cross carrying $qk^{l\text{-}1}$, with a theoretical meiotic resolution of 0.05 cM (King *et al.*, 1989), we found no recombination between this marker and the $qk^{l\text{-}1}$ allele. Thus recombination is suppressed in this cross over more than a megabase, and we were unable to determine the qk^{v} deletion breakpoints relative to the gene (Cox *et al.*, 1994). Therefore we chose to look for mutations in the QKI gene, (Ebersole *et al.*, 1996), in the lethal alleles $qk^{l\text{-}1}$, qk^{kt1}, qk^{k2}, and qk^{kt3}.

Gene Structure

To determine the intron–exon structure of the *quaking* gene, we generated PCR products from the gene and sequenced them. This was carried out using primers designed to the cDNA (Ebersole *et al.*, 1996; Fig. 1 and Materials and Methods), in combination with bubble PCR primers (Munroe *et al.*, 1994) and template DNA from bubble linker libraries of BAC and P1 clones spanning the gene (see Munroe *et al.*, 1994; and Materials and Methods). The structure of the gene is shown in Fig. 1, and primer sequences for amplification of each exon and its splice sites are shown in Table 2.

To define the first exon of the gene, the cDNA sequence (U44940) was compared with published genomic sequence that spans 3.5 kb at the 5′ end of the gene (U44942) and includes the ATG start codon and 730 bases of the first intron.

We have identified six exons common to all transcripts, and all five of the internal exons are flanked by recognizable splice sites (see AJ012812–AJ012816). PCR amplification from genomic DNA verifies that all introns are accounted for (Table 2). These exons contain all the common coding sequence, and their organization is verified by RT-PCR and sequencing (data not shown). The genomic structure of the QKI-6 and QKI-7 isoforms was deduced from the cDNA sequence and amino acid sequences of the three isoforms of this gene (QKI-5, QKI-6, and QKI-7) given in Ebersole *et al.* (1996). The intron–exon boundary of QKI-7 was determined by bubble PCR (AJ012817). The QKI-6 terminal exon is wholly contained within the sequence of the longer C- terminal QKI-7 exonic sequence (Fig. 1, AJ012818). The first 25 amino acids of the QKI-5 C-terminus are encoded by an exon (QKI-5a, Fig. 1) wholly contained within the C-terminal QKI-7 exonic sequence (AJ012819). The remaining 5 amino acids and stop codon were identified as a result of searching

TABLE 2

PCR Primer Pairs, Conditions, and Product Sizes and Amplification of Exons and Flanking Sequence

Exon	Intron primer at 5′ end of exon	Intron primer at 3′ end of exon	Notes	Product size	Exon size
2	**INTK,** CGAATAGGTTGACAGCTCCTCACTAAAGTG	**INTQ,** CTGTGGTCAGAGTTGCACTTCATCTTGAAG	**a**	535	143
3	**INTC,** CGCATATGTATGCTTTCACATAAC	**INTD,** GTAGAAACCCTAAGACAATTACTGACAGGC	**b**	270	117
4	**INTR,** GTAGTTGGCACTAGGTGCCTGTTAC	**INTL2,** CTCTACCACATCACTTCCTTCCAC	**c**	≈500	144
5	**INTP,** GCCCTTGTAGTTAGATTGAGG	**INTF2,** GAGCAAACCACTGCTAAACAG	**d**	413	88
6	**INTM,** CTGTGTGGTCAGTGCATGTACTAATG	**INTN,** CGATAACTATCCTTCCAAAGTCTGGAGTAC	**a**	524	300
QKI-7-terminal	**INTG,** CTACTGTGGACTAGATGACCACATAC	**H,** TCCACATTGACTGGTCGTGTTATAACAGCTG	**b**	245	5306[a]
QKI-6-terminal	**GMAFTP-F1,** GTGGTAGTGAGGAGATTGGTATTAGC	**GMAFTP-R1,** AGTGCAAGTAAGTCACTGACTGCTAGGC	**e**	216	4026[a]
QKI-5a	**GAVA71-F1,** GCTGTCAATGTTTCCCTTATGTATG	**EWIMP-R1,** TTTTTCCATCAGGTCTGGTGTTGCA	**e**	219	75
QK1-5b-terminal	**GAVA INT-R2,** CATGTTGTCATTAAGCCGTCACTAGCTG	**GAVA R2,** GCAAAGGCGATTACCAGTTAACTGATCAGC	**e**	198	ND
QKI-G-terminal	**N,** GTACACCTACGCCAGCTGGCCCTACCATAATG	**GKYD R3,** TATACTGTCGCTCTGAAATCGATAAC	**e**	306	ND

Notes. The position of primers for exons 2–6 are not shown in Fig. 1 but flank each exon so that the splice sites can be sequenced. Annealing temperatures and magnesium concentrations for PCR amplification: **(a)** 55°C, 1.5 mM $MgCl_2$; **(b)** 50°C, 1.5 mM $MgCl_2$; **(c)** 50°C, 3.0 mM $MgCl_2$; **(d)** 50°C, 2.0 mM $MgCl_2$; **(e)** 60°C, 2.0 mM $MgCl_2$; (f) 50°C, 4.0 mM $MgCl_2$; (g) 51°C, 1.0 mM $MgCl_2$. **INT** within a primer name indicates primer is in an intronic context. Sequence Accession Nos. AJ012812–AJ012820.

[a] Based on U44940 sequence data.

The Institute for Genome Research (TIGR) mouse gene index (MGI) database with the cDNA sequence (U44940), which by sequence comparison identified a novel terminal exon (TC34592, QKI-5b in Fig. 1) containing the last 5 amino acids. Bubble PCR from this sequence allowed the genomic intronic sequence to be determined (Fig. 1, AJ012820). All of the identified intron–exon boundaries conform to splice site consensus sequences (Jackson, 1991; AJ012817-AJ012820). From these data, the order of the exons in genomic DNA is unambiguous (Fig. 1).

In the course of sequencing cDNA clones from a mouse 11- to 15-day postnatal brain cDNA library, and also in database searching, an additional putative 3′ sequence was found. This sequence (mo83c09, MGI AA210530) is contiguous in genomic DNA with the end of exon 6 (the last common exon). It crosses the splice site, includes sequence that is intronic in the context of the other alternative isoforms, and encodes a peptide, GKYDSCTM. To rule out the possibility of a cloning artifact, RT-PCR was carried out on brain RNA between a primer from exon 5 (F in Fig. 1) and a primer within the new sequence (GKYDR3 in Fig. 1), and the resulting PCR product was sequenced. This sequence indicates that the new polypeptide sequence is included in a correctly spliced (exon 5 to exon 6) transcript (data not shown). We have called this new C-terminal exon QKI-G.

Mutations in QKI

In the *lethal-1* strain, we have found an AT to GC transition at nucleotide 2783 of the published sequence (U44940) that was not found in the BTBR background strain. This mutation lies immediately downstream of the coding sequence of the alternative isoform QKI-6 (Fig. 1).

In homozygous qk^{k2}/qk^{k2} embryo RT cDNA, we found a TA to AT transversion at nucleotide 977 that changes a valine within the KH domain to a glutamic acid. This basepair change was not present in cDNA from wild-type littermates or in C57BL/6J cDNA.

In qk^{kt3} we identified an AT to GC transition that changes a glutamic acid to a glycine in the QUA1 domain, which is the same mutation described by Ebersole *et al.* (1996) for the qk^{kt4} allele. The qk^{kt3} and qk^{kt4} mutation isolates were littermates derived from the same ENU-mutagenized father, so they are suspected to be identical (Justice and Bode, 1988). This sequence analysis suggests that they were derived from the same mutagenized spermatogonial stem cell, and the sibling isolates of a single mutant allele will be referred to as $qk^{kt3/4}$ hereafter.

The defect in qk^{kt1} remains elusive, despite extensive sequencing of the coding region and splice sites. Although we have been unable to amplify the first exon–intron boundary due to the extreme intronic GC content, there is no evidence from RT-PCR of a splice variant.

Through sequencing $T/qk\text{-}t^{w5kt1}$ and $+/qk\text{-}t^{w5kt3}$, we have found a number of base changes that are single nucleotide polymorphisms in the *tw5* strain background (Fig. 1); these are as follows: at base −26 from exon 2 G/A (BTBR/*tw5*); at base +116 from exon 2 G/A (BTBR/tw5); at base −55 from QKI-7 A/G (BTBR/tw5), and in cDNA sequence A2778T (BTBR and *tw5*).

Consequences of the quaking$^{lethal\text{-}1}$ Mutation

The $qk^{l\text{-}1}$ mutation changes the sequence CAG ATC TGA to CAG GTC TGA, resulting in a probable exon–intron splice site, the consensus for which is (A/C)AG **GT**AAGT (the GT residues in boldface are invariant and are the first nucleotides in the intron (reviewed in Jackson, 1991). To determine whether this new splice site was utilized, cDNA prepared from adult brain poly(A)$^+$ RNA of each mutant line including $+/qk^V$ and a BTBR control was PCR amplified using primers from

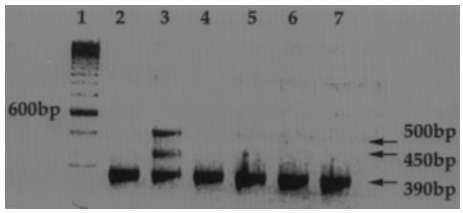

FIG. 2. Abnormal RNA splicing in the $qk^{l\text{-}1}$ strain. Primers N and GAVA-R2 (see Materials and Methods and Fig. 1) were used to PCR amplify random primed adult brain cDNA. Products were resolved on a 2% agarose TAE electrophoresis gel stained with ethidium bromide. Lanes **1,** 100-bp ladder; **2,** BTBR; **3,** $+/qk^{l\text{-}1}$; **4,** $+/qk^{v}$; **5,** $+/qk^{kt1}$; **6,** $+/qk^{k2}$; **7,** $+/qk^{kt3}$.

exon 6 and exon QKI-5b downstream of the termination codon (Fig. 1). The wildtype 390-bp product was found in all samples, whereas two RT-PCR products of 450 and 500 bp were detected only in the $T/qk^{l\text{-}1}$ sample (Fig. 2). The sequence of the two transcripts revealed that they utilized the new splice site (Fig. 3). The larger variant contained exon 6, the exon utilizing the new splice site, exon QKI-5a, and exon QKI-5b (Figs. 1 and 3). The smaller transcript was identical except for the omission of exon QKI-5a (Figs. 1 and 3). The coding potential of the QKI-5-specific exons is lost owing to the stop codon following the QKI-6 C-terminal polypeptide sequence.

To determine what effect this new splice site had on transcripts encoding the two alternative coding ends of this gene transcribed from the $qk^{l\text{-}1}$ allele, we exploited the fact that the mutation destroys a *Sau*3AI restriction site present in the wildtype sequence. RT-PCR of $T/qk^{l\text{-}1}$ cDNA between exon 6 and C-terminal exon QKI-6 using a primer downstream of the $qk^{l\text{-}1}$ mutation (Fig. 1), which distinguishes normal QKI-6 transcripts from those utilizing the mutant splice site, was digested with *Sau*3AI and resolved on agarose gels (Fig. 4). Clearly the normal QKI-6 transcript is being produced.

DISCUSSION

Structure of the quaking Gene

The *QKI* gene is composed of six common exons and four alternative ends, one of which, QKI-5, is composed of two exons. An additional transcript, designated QKI-G, that encodes a novel carboxy terminus of unknown function has been identified. Thus, this is a complex gene with alternative isoforms and a complex pattern of transcripts by Northern blot analysis (see Ebersole *et al.*, 1996). The sequences of the boundaries between intronic and exonic DNA all conform to more than the minimum consensus requirements.

Contrasting Phenotypes and Lesions of the qk Lethal Alleles

Previous studies of the QKI gene have revealed that the spatial and temporal patterns of expression of the QKI transcripts could account for the phenotypes of the various alleles. Further, wildtype QKI proteins are expressed in myelinating oligodendrocytes, but the QKI-6 and -7 protein isoforms are not expressed in homozygous qk^{v} brains (Ebersole *et al.*, 1996; Hardy *et al.*, 1996). An amino acid substitution in qk^{kt4} in QUA1 has suggested that this domain is functionally important (Ebersole *et al.*, 1996). Our analysis reveals an additional terminal exon in the embryonic isoform QK1-5 and demonstrates the molecular defects in two additional alleles of quaking that have contrasting phenotypic effects, $qk^{l\text{-}1}$ and qk^{k2}. We have screened the QKI gene for mutations in all the lethal quaking alleles with the exception of the first exon–intron boundary.

The $qk^{l\text{-}1}$ mutation generates a new splice site that potentially eliminates the QKI-5 isoform. The splice site sequence is a perfect match with the first 5 nucleotides of the consensus, including the obligatory GT; however, only one of the last 4 nucleotides (G) is

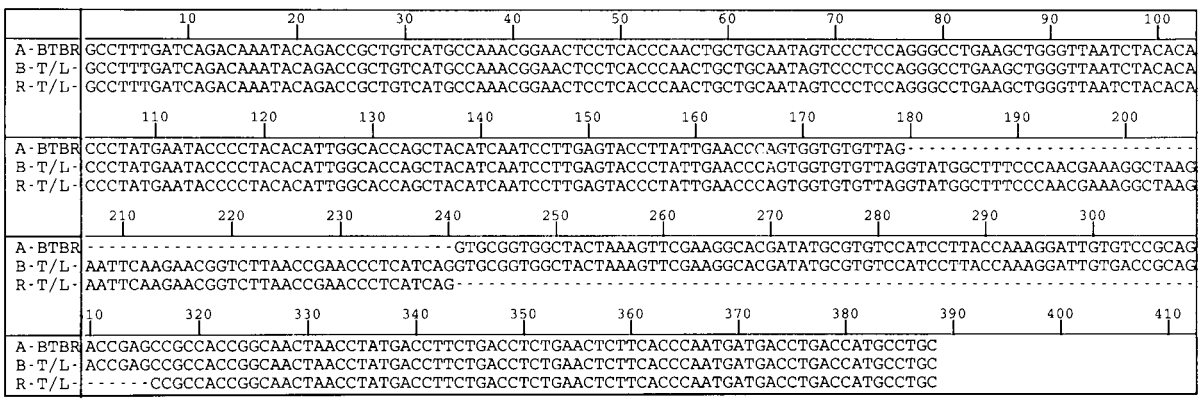

FIG. 3. Utilization of the splice site generated by the mutation in the $qk^{l\text{-}1}$ strain-aligned sequence of the RT-PCR products shown in Fig. 2. The top line of sequence is wildtype BTBR cDNA, and the two lines labeled B and R are $qk^{l\text{-}1}$ allele derived cDNA products. Dashes indicate that spaces have been inserted into the sequence to maintain the alignment. Bases up to 179 are from exon 6, the last common exon. Sequence from 180 to 240 (not found in BTBR) represents QKI-6 sequence (TAA at 223 bases is the stop codon) found in both of the splice forms shown here, and the G at 240 precedes the mutated base in genomic DNA. The sequence from 241 to 315 in BTBR and the B isoform sequence is exon QKI-5a and the sequence from 316 is the exon QKI-5b. The R isoform omits the first QKI-5-specific exon, splicing straight to exon QKI-5b. None of these new splice isoforms have any coding potential because of the in-frame stop codon at 223.

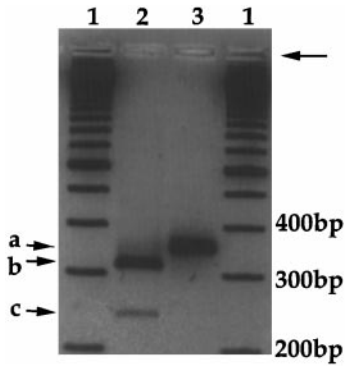

FIG. 4. Expression of normally spliced transcripts from the $qk^{l\text{-}1}$ allele. RT-PCR of brain cDNA from $+/qk^{l\text{-}1}$ with primers N and GMAFTP-R1. Lanes **1,** 100-bp ladder; **2,** *Sau*3AI-digested PCR product; **3,** undigested PCR product. Arrows: **a,** uncut PCR product; **b,** $qk^{l\text{-}1}$ allele derived product; **c,** BTBR allele derived product; unlabeled arrow denotes origin of agarose gel.

matched. This divergence is not unusual for GT splice sites, and the number of sites showing identity at each of 3 of the nucleotides 3′ to the GT has been reported (reviewed in Jackson, 1991) to be 59% (A), 71% (A), and 100% (G, identical in this splice site). The QKI-5 isoform is a nuclear localized protein and is the primary isoform found during embryonic development (Ebersole *et al.*, 1996; Hardy *et al.*, 1996). The affected region of the protein contains the nuclear localization signal as shown by deletion construct experiments in *Xenopus* (Zorn and Krieg, 1997). Experiments in *Xenopus* also provide some evidence that the carboxy end of the protein, outside of the KH domain, is also important in RNA binding. However, it was not determined whether all of the residues were required including those specific to QKI-5 (Zorn and Krieg, 1997; Zorn *et al.*, 1997). The $qk^{l\text{-}1}$ mutation specifically affects the QKI-5 isoform, but the QKI-6 isoform is produced from the $qk^{l\text{-}1}$ allele at levels similar to those of the wildtype allele. The $qk^{l\text{-}1}/qk^{v}$ compound heterozygotes differ from compound heterozygotes of qk^{k2}, qk^{kt1}, and $qk^{kt3/4}$ in that animals show a transient dysmyelinating phenotype (Shedlovsky *et al.*, 1988; Justice and Bode, 1988).

The precise molecular nature of the qk^{v} mutation is not known, although the close proximity of a megabase deletion is presumably significant (Ebersole *et al.*, 1996; Hardy *et al.*, 1996; Cox *et al.*, 1994). It has been suggested from immunohistochemistry experiments on brain sections that oligodendrocytes from qk^{v} homozygotes, unlike wildtype homozygotes, do not express the QKI-6 and QKI-7 protein isoforms. However, the mutant does express QKI-5, and strikingly all three isoforms are expressed in other glial cells (Hardy *et al.*, 1996). Assuming that the viable mutation acts in oligodendrocytes as reported by Hardy *et al.* (1996), this suggests that complementation of the qk^{v} deficit by $qk^{l\text{-}1}$ is achieved by expression from the $qk^{l\text{-}1}$ allele of the QKI-6 and QKI-7 proteins in oligodendrocytes. Two alternative hypotheses are additionally required to explain the transient myelin deficiency. The first is that the QKI-5 isoform is also required early in myelin formation but is less important later in the process. The second alternative is that this mutation reduces to limiting levels (below that of qk^{v} heterozygotes) the amount of QKI-6 and QKI-7 protein derived through expression from the $qk^{l\text{-}1}$ allele. Given the distance (1343 bases) from the terminal exon splice site for this isoform and the $qk^{l\text{-}1}$ mutation, it is unlikely that QKI-7 transcripts are affected by the mutation. In addition, we can show that correctly spliced QKI-6 transcripts are indeed made in the $qk^{l\text{-}1}$ allele in adult brain RNA. Further, the QKI-6 isoform-specific amino acids are effectively contained within a new exon in the $qk^{l\text{-}1}$ allele of the gene, resulting in some transcripts now carrying a noncoding sequence of one or two exons normally found with QKI-5. However, the QKI-5 isoform is also found in oligodendrocytes and therefore should not result in loss of the mutant splice isoform through mechanisms acting at the RNA level, although QKI-5 does undergo some down-regulation during the period of myelination (Hardy *et al.*, 1996). Therefore, we find the first hypothesis most attractive. The mechanism for the down-regulation of QKI-5 is unknown, but could include RNA stability determined by 3′ sequences. It is possible that lack of expression of QKI-5 from the $qk^{l\text{-}1}$ allele of the gene in $qk^{v}/qk^{l\text{-}1}$ compound heterozygotes results in QKI-5 protein being limiting during early myelination. Because levels are normally down-regulated (Hardy *et al.*, 1996), these cease to be limiting, and the animals recover their myelin completely. It has been proposed that $qk^{l\text{-}1}/qk^{v}$ may reflect either an ability of the $qk^{l\text{-}1}$ gene product to regulate homeostatically in ways lost to other alleles or else leakiness of this allele for normal function (Shedlovsky *et al.*, 1988). Perhaps the intact isoforms expressed by the $qk^{l\text{-}1}$ allele are able to carry out the function of the QKI-5 isoform inefficiently, allowing compensation over time.

The mutation in qk^{k2} changes a valine that lies within the KH domain to a glutamic acid. This valine is evolutionarily conserved in the QKI protein in human, *Xenopus*, and *Drosophila* and is also conserved in related KH-domain containing proteins, such as Sam68 (Zorn and Krieg, 1997; Vernet and Artzt, 1997). The KH domain is essential for RNA binding in *Xenopus* (Zorn and Kreig, 1997). In contrast to $qk^{l\text{-}1}$, the lesion lies within the common coding region of all three isoforms, explaining its effect on embyrogenesis and its inability to complement the myelination defect of qk^{v}. It is possible that this lesion affects RNA-binding capacity or specificity, since it appears to be essential for function for any of the protein isoforms.

The $qk^{kt3/4}$ lesion is also found within the coding region, but in a conserved amino acid in the QUA1 domain, changing a glutamic acid to a glycine. Interestingly, embryos compound for qk^{k4} and qk^{k2} die later in development (9.5–10.5 dpc) than qk^{k2}/qk^{k2} homozygotes (8.5 dpc, see Table 1), indicating a difference in

the two mutations (Justice and Bode, 1988). However, because $qk^{kt3/4}$ arose in tw5 chromatin, its homozygous embryo lethal phenotype is complicated by linkage with other lethal mutations and has not been determined. Even further, qk^{k2}/qk^{k2} homozygotes die earlier than qk^{k2} hemizygotes, indicating other unique features of this allele (Justice and Bode, 1988). Together, these alleles demonstrate that the KH and QUA1 domains are important for function in both embryonic survival and adult myelination. It would be interesting to assess the RNA-binding capacity of these two mutant alleles.

The qk^{kt3} mutant isolate was presumed to be a sibling to qk^{kt4}, and we found the same molecular lesion in these two strains. This is the first molecular confirmation of a cluster derived from an ENU-mutagenized male. Since ENU mutagenizes spermatogonial stem cells, repeat isolates, or clusters, of the same mutation are expected to occur. At present, the qk^{kt1} mutation is unknown but is not located in the coding region of the shared transcripts or in the embryonic isoform carboxy terminus, contrasting it with the other lethal mutations.

A Comparison of Mutation Spectra with ENU

The mutation in the $qk^{l\text{-}1}$ and $qk^{kt3/4}$ alleles is an AT to GC transition, which is a different spectrum from those usually seen with high-dose ENU in *Escherichia coli, Drosophila,* and the Big Blue assay system (reviewed in Marker *et al.*, 1997). However, evidence is accumulating that this lesion is as common as an AT to TA transversion in the mouse and that the most common lesions observed in *E.coli*, flies, and the Big Blue system, a GC to AT transition, is rarely seen in mouse germline and other mouse somatic assay systems (Marker *et al.*, 1997; Huang *et al.*, 1998a,b). In fact of 44 sequenced mouse ENU mutations, 18 are AT to TA and 19 AT to GC transversions (reviewed in Justice, in press).

Contrasting Functions of the QKI Isoforms Revealed by Mutagenesis

A comparison of the phenotypes of the five quaking alleles had led investigators to postulate that the quaking gene product played a role in both myelination and embryogenesis (Justice and Bode, 1988). Missense mutations identified in the qk^{k2} and $qk^{kt3/4}$ alleles reveal that the KH/QUA RNA-binding domains are essential for function of all three quaking protein isoforms, since these alleles fail to survive embryogenesis as homozygotes and are dysmyelinating in compounds with qk^{v}. These lesions are found in transcripts that produce all three protein isoforms. Notably, these alleles were isolated in screens designed to select alleles that would fail to complement qk^{v}. In contrast, the transient quaking phenotype of $qk^{l\text{-}1}/qk^{v}$ compounds reveals a unique function of the QK1-5 isoform in myelination. Homozygotes for this mutation also fail to survive embryogenesis, but this mutation was selected by virtue of its embryo lethality and was subequently found to be an allele of quaking (Shedlovsky *et al.*, 1988). It is likely that complementation occurs through expression of QKI-6 and QKI-7 transcripts from the $qk^{l\text{-}1}$ allele of QKI, in oligodendrocytes. However, QKI-5 must play a role in early myelination. Although it is down-regulated later as myelination is completed, this may be concentration dependent. The QKI-5 isoform must also play a crucial role in embryonic survival, since compounds with the qk^{k2} and qk^{kt1} alleles fail to survive embryogenesis. This mutational analysis demonstrates the power of series of alleles to examine the function of complex loci and suggests that additional mutant alleles of quaking could reveal additional functions of this complex gene.

ACKNOWLEDGMENTS

This work and A.H. were supported by an UK MRC project grant (G9418672) to R.D.C. This work was also supported by grants to W.F.D., A.S.D., and the McArdle Laboratory from the NIH (CA23076, CA63677and CA07075). M.J.J. was supported by the Oak Ridge National Laboratory, managed by Lockheed Martin Energy Research Corp. for the U.S. Department of Energy under Contract DE-AC05-96OR22464. We would like to thank Drs. Elaine Levy and Renata Hamvas for FISH and YAC Mapping.

REFERENCES

Caslini, C., Spinelli, O., Cazzaniga, G., Golay, J., Gioia, L., Pedretti, A., Breviaro, F., Amaru, R., Barbui, T., Biondi, A., Introna, M., and Rambadi, A. (1997). Identification of two novel isoforms of the ZNF162 gene: A growing family of signal transduction and activator of RNA proteins. *Genomics* **42:** 268–277.

Cox, R. D., Whittington, J., Shedlovsky, A., Connelly, C. S., Dove, W. F., Goldsworthy, M., and Lehrach, H. (1993). Detailed physical and genetic mapping in the region of plasminogen, *D17Rp17e,* and quaking. *Mamm. Genome* **4:** 687–694.

Cox, R. D., Shedlovsky, A., Hamvas, R., Goldsworthy, M., Whittington, J., Connelly, C. S., Dove, W. F., and Lehrach, H. (1994). A 1.2-Mb YAC contig spans the quaking region. *Genomics* **21:** 77–84.

Ebersole, T., Chen, Q., Justice, M., and Artzt, K. (1996). The quaking gene product necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins. *Nat. Genet.* **12:** 260–265.

Hardy, R., Loushin, C., Friedrich, V., Chen, Q., Ebersole, T., Lazzarini, R., and Artzt, K. (1996). Neural cell type-specific expression of QKI proteins is altered in quaking viable mutant mice. *J. Neurosci.* **16**(24): 7941–7949.

Huang, J. D., Cope, M. J. T. V., Mermall, V., Strobel, M. C., Kendrick-Jones, J., Russell, L. B., Mooseker, M. S., Copeland, N. G., and Jenkins, N. A. (1998a). Molecular genetic dissection of mouse unconventional myosin-VA: Head region mutations. *Genetics* **148:** 1951–1961.

Huang, J. D., Mermall, V., Strobel, M. C., Russell, L. B., Mooseker, M. S., Copeland, N. G., and Jenkins, N. A. (1998b). Molecular genetic dissection of mouse unconventional myosin-VA: Tail region mutations. *Genetics* **148:** 1963–1972.

Jackson, I. I. (1991). A reappraisal of non-consensus mRNA splice sites. *Nucleic Acids Res.* **19**(14): 3795–3798.

Justice, M. (in press). Mouse germline mutagenesis. *In* "Mouse Genetics and Transgenics: A Practical Approach" (I. Jackson and C. Abbott, Eds.), Oxford Univ. Press, London.

Justice, M., and Bode, V. (1986). Induction of new mutations in a mouse t-haplotype using ethylnitrosourea mutagenesis. *Genet. Res.* **47:** 187–192.

Justice, M., and Bode, V. (1988). Three ENU-induced alleles of the murine quaking locus are recessive embryonic lethal mutations. *Genet. Res.* **51:** 95–102.

King, T., Dove, W., Herrmann, B., Moser, A., and Shedlovsky, A. (1989). Mapping to molecular resolution in the T to H-2 region of the mouse genome with a nested set of meiotic recombinants. *Proc. Natl. Acad. Sci. USA* **86:** 222–226.

Marker, P. C., Seung, K., Bland, A. E., Russell, L. B., and Kingsley, D. M. (1997). Spectrum of Bmp5 mutations from germline mutagenesis experiments in mice. *Genetics* **145:** 435–443.

McDonald, J., and Charlton, C. (1997). Characterization of mutations at the mouse phenylalanine hydroxylase locus. *Genomics* **39:** 402–405.

Munroe, D., Haas, M., Bric, E., Whitton, T., Aburatani, H., Hunter, K., Ward, D., and Housman, D. (1994). IRE-bubble PCR: A rapid method for efficient and representative amplification of human genomic DNA sequences from complex sources. *Genomics* **19:** 506–514.

Shedlovsky, A., King, T., and Dove, W. (1988). Saturation germ line mutagenesis of the murine t-region including a lethal allele at the quaking locus. *Proc. Natl. Acad. Sci. USA* **85:** 180–184.

Sidman, R., Dickie, M., and Appel, S. (1964). Mutant mice (quaking and jimpy) with deficient myelination including in their central nervous system. *Science* **144:** 309–311.

Vernet, C., and Artzt, K. (1997). STAR, a gene family involved in signal transduction and activation of RNA. *Trends Genet.* **12**(12): 479–484.

Vogel, E., and Natarajan, A. T. (1979). The relation between reaction kinetics and mutagenic action of mono-functional alkylating agents in higher eukaryotic systems I. Recessive lethal mutations and translocations in *Drosophila*. *Mutat. Res.* **62:** 51–100.

Zaffran, S., Astier, M., Gratecos, D., and Semeriva, M. (1997). The held out wings: (how) *Drosophila* gene encodes a putative RNA-binding protein involved in the control of muscular and cardiac activity. *Development* **124:** 2087–2098.

Zorn, A., Grow, M., Patterson, K., Ebersole, T., Chin, Q., Artzt, K., and Krieg, P. (1997). Remarkable sequence conservation of transcripts encoding amphibian and mammalian homologues of quaking, a KH domain RNA-binding protein. *Gene* **188:** 199–206.

Zorn, A. M., and Krieg, P. A. (1997). The KH domain protein encoded by quaking functions as a dimer and is essential for notocord development in *Xenopus* embryos. *Genes Dev.* **11:** 2176–2190.