学位論文
Doctor's Thesis

Expression of Hqk Encoding a KH RNA Binding Protein Is Altered in Human Glioma
（ヒトグリオーマにおけるHqk遺伝子の発現の変化）

李　　正　哲
Li　　Zheng Zhe

指導教官名：臓器形成分野教授　　山村研一

審査委員名：腫瘍医学講座教授　　佐谷秀行
神経発生分野教授　　大久保博晶
分子遺伝学講座教授　　森　正敬
神経内科学講座教授　　内野　誠

2002年10月

CONTENTS

Summary I

Preface II

Publications list III

Acknowledgements IV

Abbreviations V

Chapter 1. Introduction 1

1.1 General introduction 2

1.2 *Quaking* mutant and *qkI* gene 4

1.3 KH motif and GSG Family Protein 9

1.4 QKI function in glial cell differentiation 14

1.5 Possible functions of other GSG/STAR family members 16

1.6 Biochemical functions of QKI protein 17

1.6.1 QKI may be involved in translational control 17

1.6.2 QKI may be involved in mRNA translation 18

1.6.3 QKI may be involved in mRNA splicing 18

1.6.4 QKI act as an apoptosis inducer? 18

1.7 Human *quaking* gene, *Hqk* 19

Chapter 2. Materials and methods 21

2.1 Chromosome mapping of *Hqk* 22

2.2 Cloning and sequencing of cDNA of *Hqk* 22

2.3 Determination of *Hqk* genomic structure 23

2.4 Northern blot analysis 24

2.5 Expeossion analysis and mutation screening of *Hqk* in human tumor samples 25

Chapter 3. Results 29

3.1 Chromosome mapping of *Hqk* 30
3.2 Cloning of human homolog of the *qkI* gene, *Hqk* 30
3.3 Exon-intron organization of *Hqk* 35
3.4 Expression of *Hqk* messages in human tissues 36
3.5 Alterations of *Hqk* expression in human tumors 40

Chapter 4. Discussion 47

4.1 Discussion 48

References 56

Summary

The *quaking* gene family encodes single KH domain RNA-binding proteins that play vital roles in cell differentiation, proliferation, and apoptotic processes. The human *quaking* gene, *Hqk*, maps to 6q25-q26, where cytogenetic alterations associated with a variety of human malignancies, including gliomas have been reported. To assess possible relationships of *Hqk* with human diseases such as glial tumors, we first isolated the *Hqk* gene, characterized its structure and expression pattern, and carried out mutational analysis of *Hqk* in primary tumor samples. The *Hqk* gene contains 8 exons spanning a ~200 kb genomic region, generating at least four alternatively spliced transcripts, *Hqk*-5, *Hqk*-6, *Hqk*-7 and *Hqk*-7B, of which *Hqk*-7 was abundantly expressed in brain. Analysis of primary tumors indeed demonstrated a high incidence of expression alterations of *Hqk* in gliomas (30%; 6/20), but not in other tumors such as schwannomas (0/3), or meningiomas (0/8). Among the tumor samples showing expression alterations, two were devoid of all three major transcripts, one was missing only the *Hqk*-5 message, and only the *Hqk*-7 message was absent in two cases. Our results thus imply the involvement of *Hqk* in the human glial tumor progression.

Preface

This thesis is the result of series of investigations that took place during my graduate and postgraduate studies at Kumamoto University, Japan.

This thesis will hopefully give some of the new information about human *quaking* gene and its relevance to tumorigenesis.

This thesis also reflects my own training, research interests, and biases.

Publications list

Expression of Hqk Encoding a HK RNA Binding Protein Is Altered in Human Glioma.

Japanese Journal of Cancer Research **93**, 167-177, 2002.

Acknowledgments

I would like to thank Professor Ken-ichi Yamamura, my mentor, and my Heads of Division allowing me the chance and freedom to complete this thesis. He also convinced me that I could write this thesis, and encouraged me when I was ready to give up.

I am grateful to acknowledge Dr. Kuniya Abe as the advisor. I learned a lot about biological science from him and this was a very worthwhile experience for me. He also kindly reviewed all parts of this thesis and made many suggestions making this thesis much better than it would otherwise be.

I also wish to thank all colleagues, who worked in the past and present in Institute of Molecular Embryology and Genetics, Kumamoto University.

Finally, I would like to acknowledge the support and help of my wife, Yushan Li.

Abbreviations

BAC	bacterial artificial chromosome
bp	base pairs
ENU	*N*-ethyl-*N*-nitrosourea
EST	expressed sequence tag
gld-1	germline development 1
kb	kilobase pairs
KH motif	heterogeneous nuclear ribonucleoprotein K homology motif
LOH	loss of heterozygosity
MAG	myelin-associated glycoprotein
MBP	myelin basic protein
MOG	myelin/oligodendrocyte glycoprotein
PAC	P1-derived artificial chromosome
PCR	polymerase chain reaction
PLP	proteoliptid protein
qk	*quaking*
qk^{v}	*quaking viable*
RT-PCR	reverse transcriptase polymerase chain reaction
Sam68	src-associated protein68 in mitotic
STAR	signal transduction and activation of RNA
STS	sequence tagged site
SVZ	subventricular zone
UTR	untranslated region
WHO/HOW	wings held-out/ held out wings
YAC	yeast artificial chromosoe

Chapter 1. Introduction

1.1 General introduction 2

1.2 *Quaking* mutant and *qkI* gene 4

1.3 KH motif and GSG Family Protein 9

1.4 QKI function in glial cell differentiation 14

1.5 Possible functions of other GSG/STAR family members 16

1.6 Biochemical functions of QKI protein 17

1.6.1 QKI may be involved in translational control 17

1.6.2 QKI may be involved in mRNA translation 18

1.6.3 QKI may be involved in mRNA splicing 18

1.6.4 QKI act as an apoptosis inducer? 18

1.7 Human *quaking* gene, *Hqk* 19

1.1 General introduction

The human genetic disease was first recognized in early 1900s by Garrod showing alcaptonuria as an inherited disorder. Since new and powerful technologies became available in early 1970s, more evidences emerged that some of the human diseases were caused by the defects in the specific genes. Such discoveries made genetics an integral part of modern medicine and geneticists in collaborations with clinicians made important contributions to understanding of the pathogenesis of human diseases. For obvious ethical reasons, it is impossible to perform experimental studies directly on human patients; thus model animals such as mouse have served as a valuable experimental model system for human diseases. Mouse models of human genetic disorders have a great potential for elucidating the underlying mechanisms of disease development and for devising effective therapeutic strategies in human. There are thousands of spontaneous and induced mouse mutations, some of which exhibiting phenotypes similar to human diseases. With advent of gene disruption strategies using homologous recombination in embryonic stem cells, it is now possible to design and induce specific mutations in any genes in theory, providing successful examples of mouse models for human diseases. However, it is also true that manipulating the mouse genome by gene targeting technique does not always guarantee the creation of mouse lines with expected phenotypes. There are many examples that mutations in mouse counter parts of human disease genes result in no visible phenotypes or unexpected phenotypes. On the other hand, spontaneous mutations with phenotypes related to known disease characteristics had provided clues to identification of disease-causing genes, and to understanding the etiology of specific diseases. For example, *jimpy* is one of the spontaneous mouse mutations

showing neurological phenotypes such as tremor or seizures, and was considered as a model of Pelizaeus-Merzbacher disease in human (Hodson et al., 1987; Morello et al., 1986). Macklin et al. discovered that the mutant phenotype was caused by a point mutation in a splice acceptor site of intron 4 in proteolipid protein (PLP) gene, which results in truncation of PLP protein at carboxyl terminus (Macklin et al., 1987). It was subsequently found that missense mutations in PLP or duplication of the PLP locus indeed resulted in the Pelizaeus-Merzbacher disease in human (Hodes et al., 1993). This is one of the landmark studies, in which mouse genetics made great contributions to understanding human genetics diseases. This finding prompted researchers to use other neurological mouse mutants for asking relationships with human demyelinating diseases. It is now known that mutations in PMP-22 gene leads to dysmyelination in peripheral nerve systems in mouse and human, and are responsible for human Charcot-Marie-Tooth type 1a and mouse neurological mutation, Trembler (Suter et al., 1993). Also, point mutations in protein 0 (P0) results in Charcot-Marie-Tooth type 1b (Hayasaka et al., 1993), clinically very similar to the phenotype of PMP22 mutant, Trembler or Charcot-Marie-Tooth type 1a. P0 deficient mice exhibit abnormal motor coordination and tremors. These examples clearly demonstrate the usefulness of mouse models in identification of genes causing demyelination diseases. *Quaking* is another, classical neurological mouse mutants showing pleiotropic phenotypes including characteristic tremor, epileptic seizures and infertility in male. It has been thus considered quaking mutation as a potential model for the genetic disorder in human.

1.2 *Quaking* Mutant and *qkI* Gene

Quaking viable (qk^v) mutation isolated by Sidman et al in1964 (Sidman et al., 1964) exhibits pleiotropic effects on myelination, spermatogenesis (male sterility). The Original *Quaking viable* (qk^v) is a spontaneous, autosomal recessive mutation characterized by hypomyelination in the central nervous system (CNS). In the CNS of the qk^v homozygotes (qk^v/ qk^v), reduced number of myelin lamellae and a lack of compacted myelin sheaths were observed (Fig. 1). The qk^v /qk^v survive to adulthood and exhibit a characteristic tremor or '*quaking*' of the hindlimbs by postnatal 10 day (Fig. 2). The identification of oligodendrocytes as the affected cells, with incompletely wrapped myelin, suggested that an arrested of myelin is causative of the quaking mutation. The gene has been mapped on proximal part of the mouse chromosome 17, in which no structural genes for myelin structural proteins are located. Thus the gene product of the qk^v may not be an integral component of the myelin, rather representing regulatory molecule of myelination or glial cell differentiation. In 1980s, four ENU (*N*-ethyl-*N*-nitrosoura) induced mutant mouse lines were shown to have similar phenotypes as in qk^v mice (Justice et al., 1988). Interestingly, homozygoue embryos showed embryonic lethality at around is embryonic lethal around 9-10 embryonic days of gestation. Furthermore, the compound heterozygotes (qk^e/qk^v) exhibit a dysmyelination phenotype similar to the qk^v/qk^v, but being able to fertile. These evidences suggested that *qkI* gene was required for mouse embryogenesis, and that the male sterility was due to mutation in other genes. In 1992, during the process of genetic mapping studies in the proximal region of the Chr.17, Ebersole et al. discovered a large deletion in the genome of the qk^v mice (Ebersole et al., 1992).

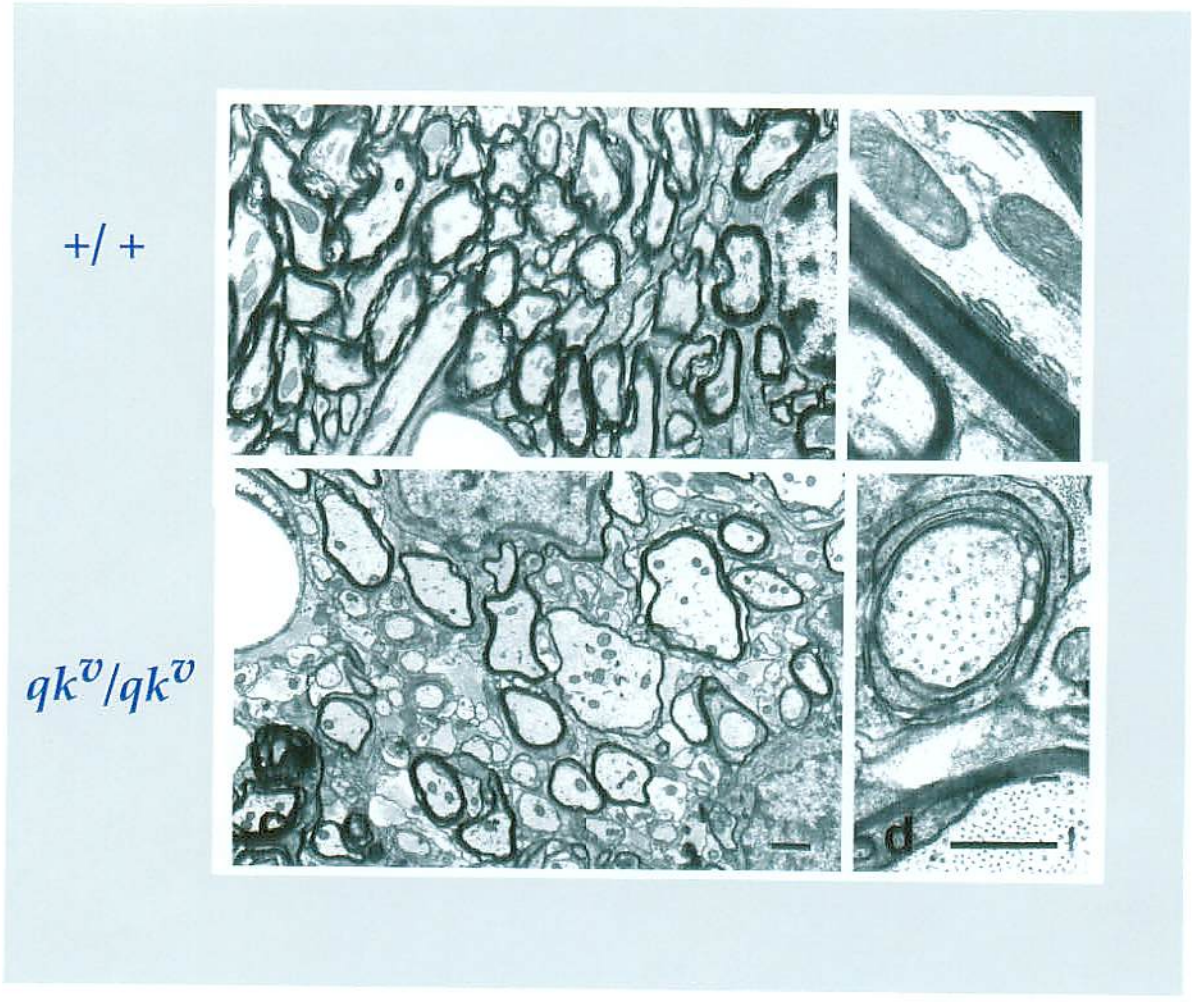

Fig. 1. Electron-micrograph of the hypomyelinated CNS from an adult *quaking* mouse (qk^v/qk^v) Compare with wild type (top), most axons thinly myelinated(low).

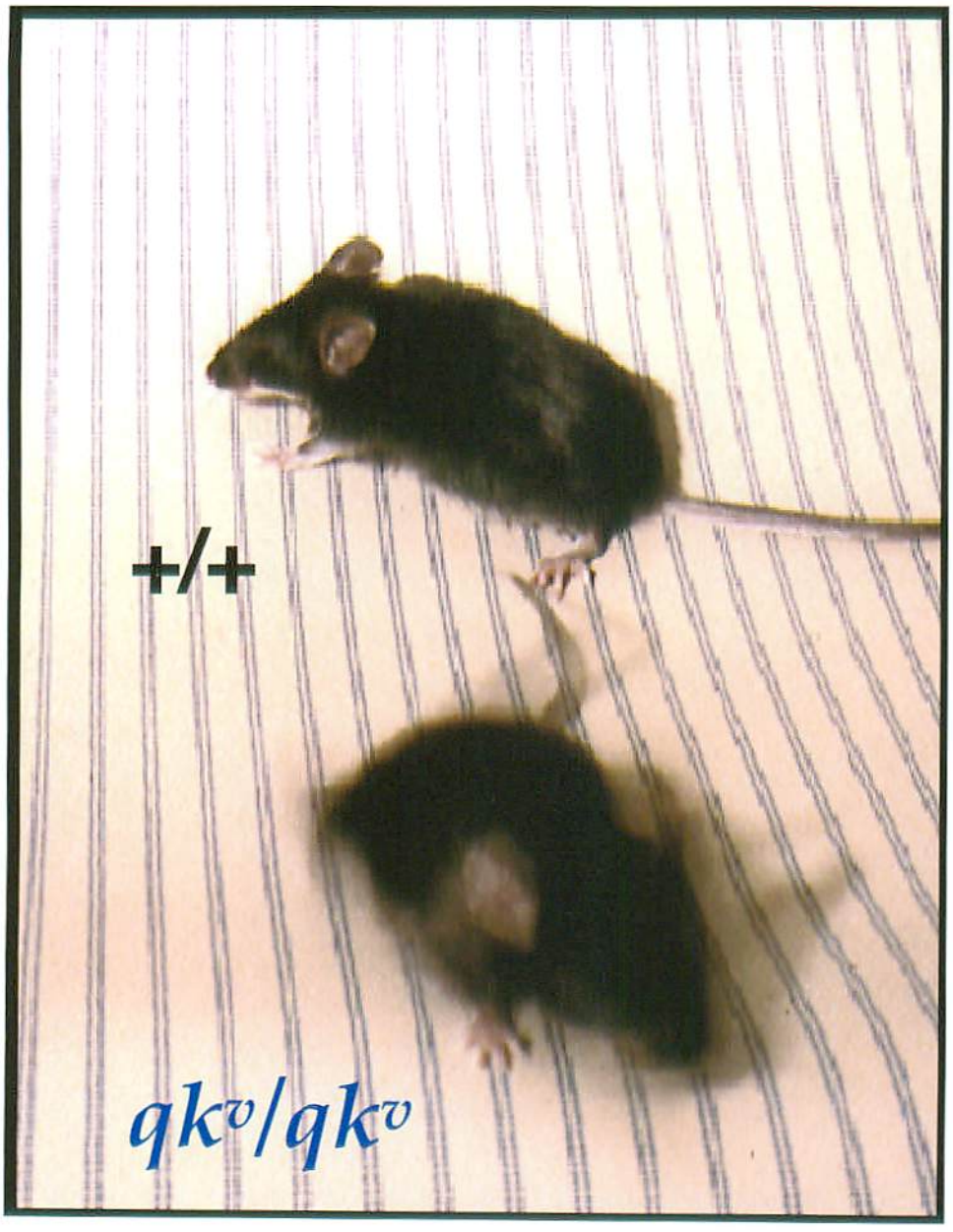

Fig. 2 *quaking viable* homozygous(qk^v) shows penotypes such as body tremor and seizures.

They searched transcription units in the deletion as well as the region adjacent to the deletion break point, and found one putative gene named as *qkI*, which was then considered as a candidate gene for *quaking* mutation (Fig. 3) (Ebersole et al., 1996). Mutational analyses and functional rescue experiments with BAC transgenesis now unambiguously showed that the qkI gene is the true responsible gene for the quaking mutation. Analysis of the expression of the *qkI* and the phenotype suggests that the ability of *qkI* gene to influence myelination is dose dependent (Kaname et al, in preparation). The *qkI* gene contains at least 9 exons spanning a genomic region of 65 kb DNA. Six distinct messages are produced from the *qkI* locus by alternative splicing of the primary transcript and at least three QKI protein isoforms have been identified known as QKI-5, QKI-6, QKI-7, which differ in their coding regions only at their carboxyl ends (Ebersole et al., 1996; Kondo et al., 1999). From immunohistochemical and cell biological analyses, these three QKI isoforms have distinct intracellular distributions; QKI-6 and QKI-7 are found predominantly in cytoplasm, whereas QKI-5 tends to be restricted to the nucleus. Such distinct localizations of QKI protein isoforms suggest that the each isoform may have unique role in cellular functions. In adult tissues, *qkI* gene is expressed more or less in all the tissues tested, but is highly abundant in brain or heart. In brain, the *qkI* gene is expressed in glial cells, both oligodendrocytes and astrocytes, but not in neurons. These glial cells share common precursors; the postnatal subventricular zone (SVZ) surrounding the lateral ventricles has long been recognized as a source of forebrain astrocytes and oligodendrocytes (Levison and Goldman, 1993). Immunohistochemical analysis revealed that subpopulation of the SVZ express QKI, suggesting that QKI may play a vital role in the decision of glial fate, or differentiation of glial cells (Hardy et al., 1996).

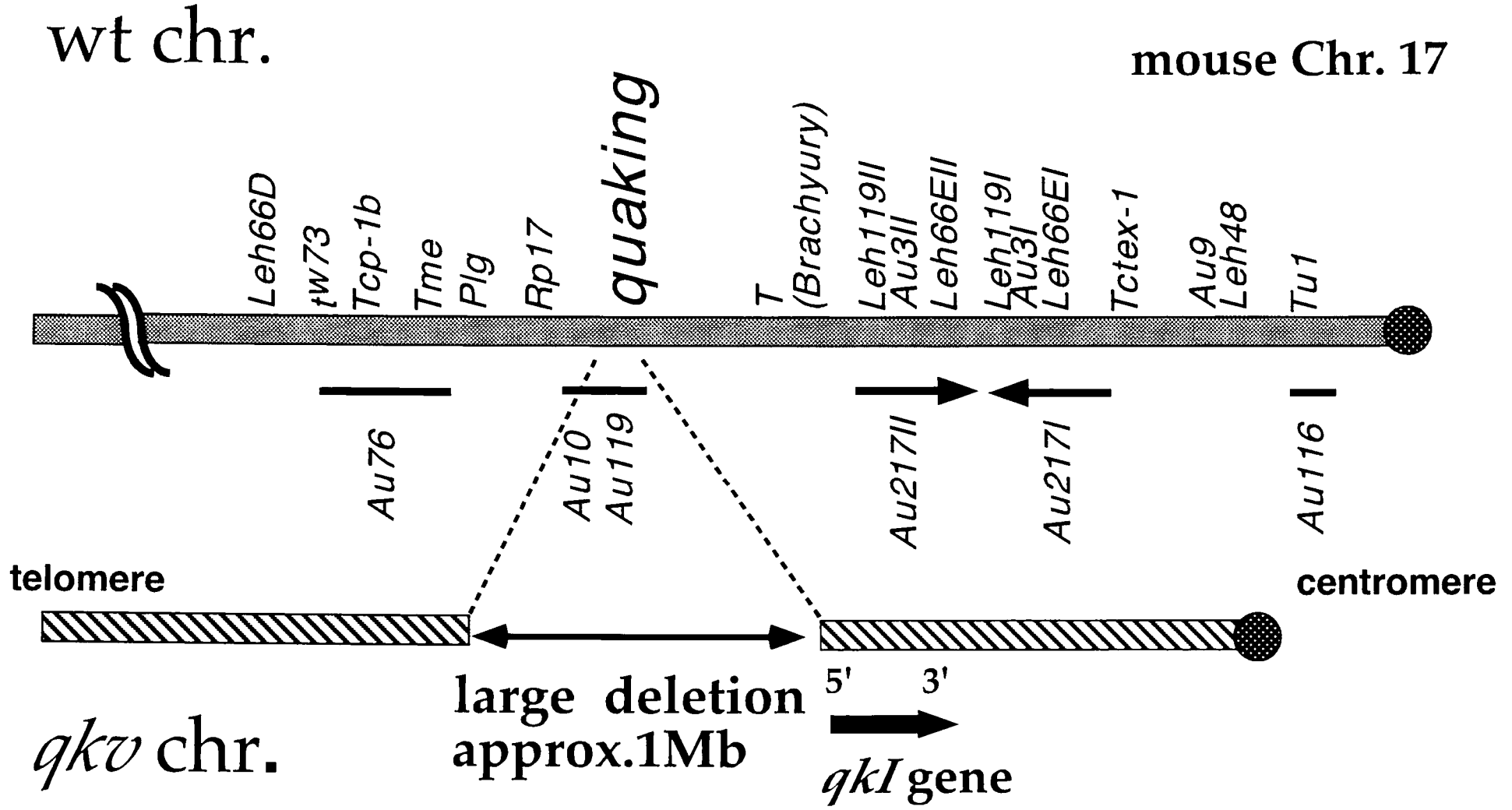

Fig. 3. A *quaking* candidate gene which is located at the break point of a large genomic deletion was idetified by reversed genetic method.

According to the sequence analysis and the subsequent homology search, the predicted amino acid sequence of the QKI contains a single KH domain, a well-known RNA binding protein motif (Siomi et al., 1993a). Among the RNA binding proteins with K homology (KH) domain, the QKI protein belongs to subfamily of the KH containing protein, i.e. GSG family proteins.

1.3 KH motif and GSG Family Proteins

The KH motif was originally identified as a conserved sequence represented in the human heterogeneous nuclear ribonucleoprotein (hnRNP) K protein, which a major pre-mRNA-binding protein (Siomi et al et al 1993a). From the analysis of more than 40 potential KH -containing proteins, a three-dimensional structure prediction for the KH motif that consists of three beta-strands and two a-helices in the order β - α - α - β - β - α -fold was provided (Fig. 4) (Gibson et al., 1993). All KH motif proteins of known functions can be associated with RNA and in fact many binds RNA in vitro (Siomi et al., 1993b; Siomi et al., 1994). The definitive role of the KH motif in RNA binding is not yet known, but recent observations demonstrate that it is essential for RNA binding and probably binds RNA directly. Mutations in highly conserved residues in the KH domain impairs the binding activity to single-stranded nucleic acids in vitro (Siomi et al., 1994). Like many other RNA-binding motifs, KH motifs are found in one or multiple sites in single protein molecule, e.g. one in *qkI* and 14 in chicken vigilin (Gibson et al., 1993). In hnRNP K containing three KH motifs and FMR-1 having two KH motifs, each KH motif in these proteins is necessary for in vitro RNA binding activity, suggesting that these KH motifs may function cooperatively. In the case of single KH motif proteins such as GRP33, SAM68 and GLD-1 they may form homo or heterodimer with other RNA binding protein with similar structure. Experiments with hnRNP

A

B

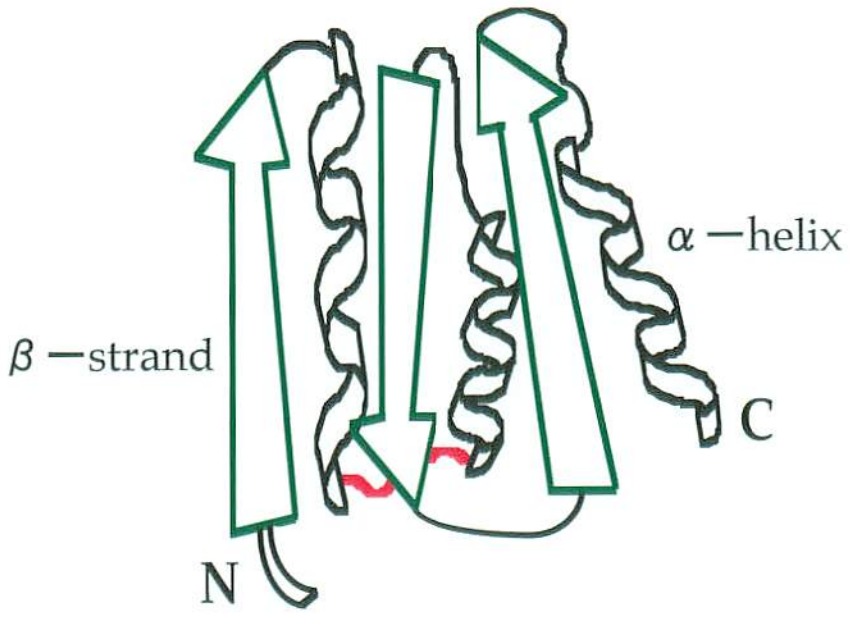

Fig. 4. A The secondary structure of KH motif.

B The structure of one of the human vigilin KH domain. Each arrow represent a β-strand and the curled ribbons a α helix. The helical side of the KH domain interacts with RNA and the loop between the first two α helix (indicated in red) which highly conserved also plays important roles in RNA binding.

K have demonstrated that this protein also binds DNA and that its KH motifs are required for DNA binding as well. The widespread presence of KH motifs in diverse organisms suggests that it is an ancient protein structure with important cellular functions. The KH motif in the QKI is embedded in a larger conserved domain of ~200 amino acids, and the whole conserved sequence is called GSG or STAR (signal transduction and activation of RNA) domain (Fig. 5) (Vernet et al., 1997). The GSG domain was initially identified by aligning the first three of its family members, Artemia salina GRP33, a hnRNP isolated from brine shrimp, SAM 68 (Taylor et al., 1994), a phosphoprotein involved in Src signaling (Courtneidge et al., 1994), and GLD-1, a protein required for germ cell differentiation in Caenorhabditis elegans (Francis et al., 1995a; Francis et al., 1995b). The GSG domain family members also include Drosophila (*Who/How*), Xenopus (*Xqua*) and mouse *qkI*. The GSG domain has a single KH domain that is longer than most of other KH domains (Musco et al., 1996). In addition to the KH domain, the GSG domain is composed of conserved sequences of unknown functions; approximately 70 amino acids sequence at N-terminal and ~25 amino acids at C-terminal relative to the KH domain. These regions in the GSG domain in QKI are called QUA1 and QUA2, respectively (Fig. 5) (Ebersole et al., 1996). The QUA1 and QUA2 regions are also referred to as NK(N terminus of KH region) and CK(C terminus of KH region) (Vernet et al., 1997). NK seems to be required for dimerization to help the KH domain binding target RNA. The KH domain and the CK are needed for splicing function in another GSG member, SF1 (Liu et al., 2001). Mutations in the GSG/STAR domain result in defects in cell proliferation and/or specific developmental transition in various organisms, suggesting the importance of these genes in cell differentiation and development (Fig. 6).

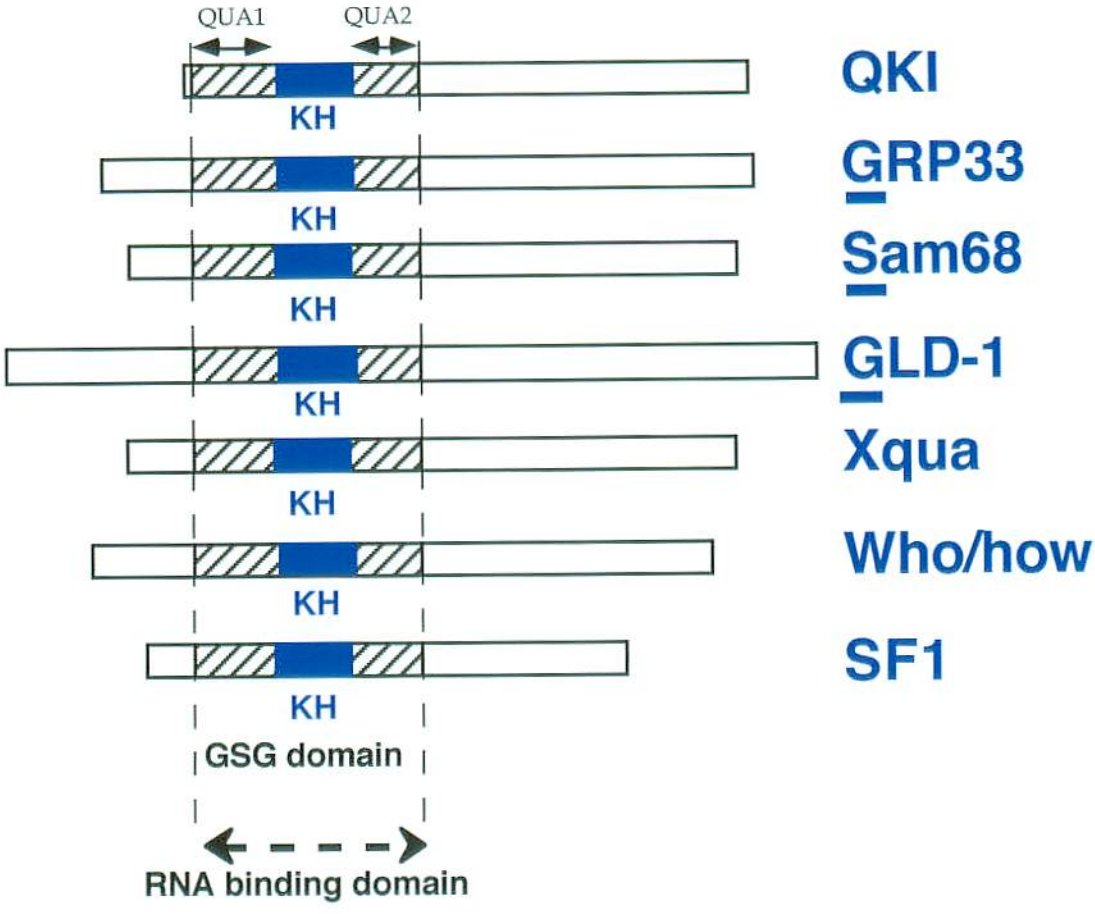

Fig. 5. The GSG domain was initially identified by aligning the first three of its family members, GRP33, Sam68, and gld-1(underlined). GRP33 is a hnRNP isolated from brine shrimp, SAM68, a phosphoprotein involved in Src signaling, and GLD-1, a protein required for germ cell differentiation in *Caenorhabditis elegans* .

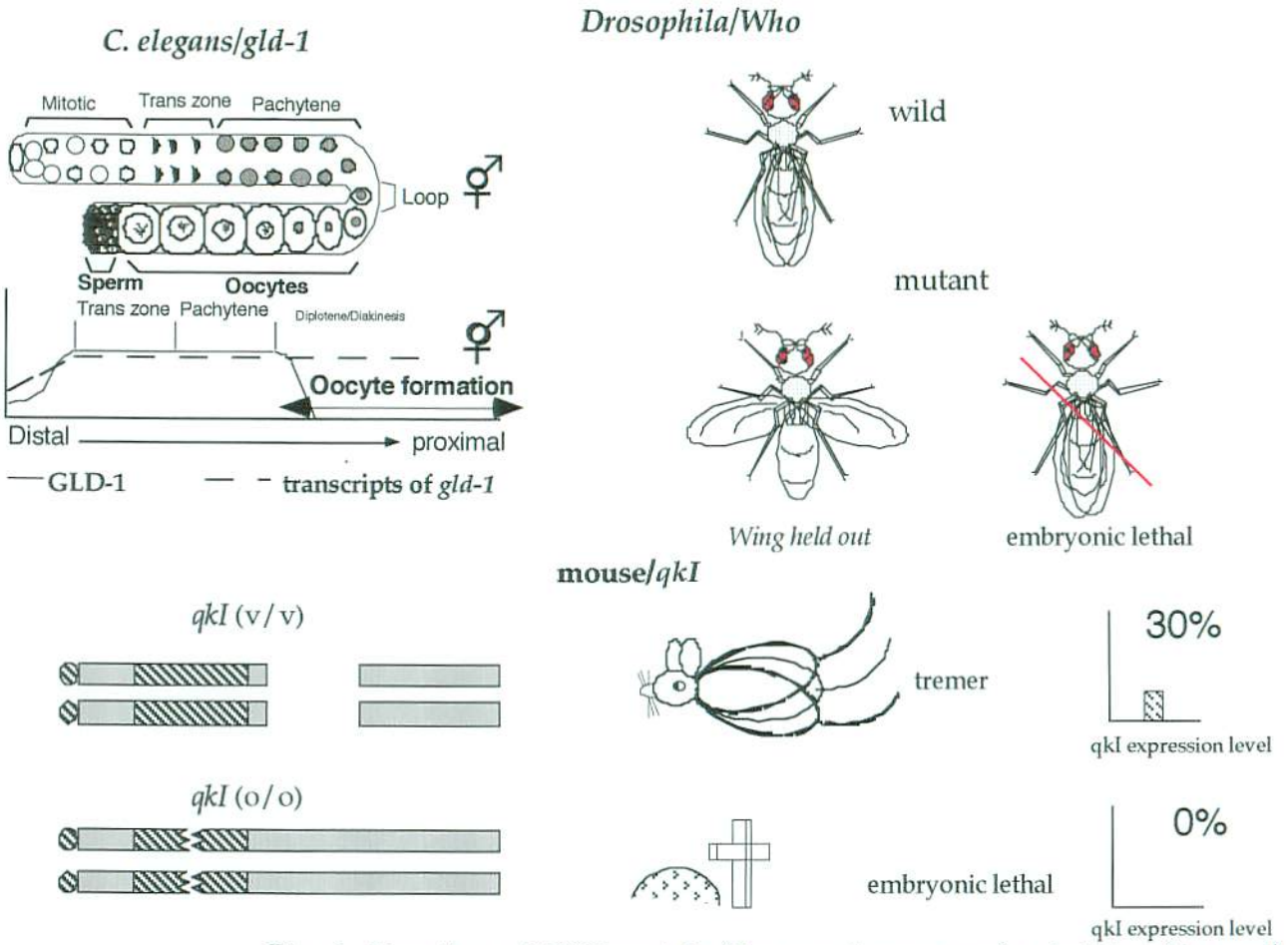

Fig. 6. Function of GSG protein (See previous page for detials of legend)

Functions of the GSG proteins in mouse and other organisms will be described in detail below.

1.4 QKI functions in glial cell differentiation

Analysis of myelination in qk^v mice suggests that QKI plays a vital role in glial cell differentiation. Myelin, named so by Virchow, is a spiral structure constituted of extensions of the plasma membrane of the myelinating glial cells, the oligodendrocytes in the CNS (Fig. 7) (1854). During myelination in the CNS, one type of glia, i.e. oligodendrocytes express a number of myelin proteins such as MAG (myelin associated glycoprotein), PLP (proteolipid protein), and MBP (myelin basic protein), which are structural components of myelin in the CNS. Morphologically, oligodendroglia extend their processes that enwrap axons and generate a compact multilamellar membrane spiral. With an electron microscopy, alternating intra- and extracellular membrane leaflet appositions can be visualized as major dense lines. In contrast to normal myelin structure, myelin of the qk^v/qk^v mice appears much thinner and is not compacted, suggesting that enwrapping process of myelination is arrested at its early stage. Expression of PLP and MBP is greatly reduced. In mouse, at least six MBP isoforms, ranging in size from 14kDa to 21.5kDa, are produced via alternate splicing of the primary MBP transcript (Takahashi et al., 1985; Newman et al., 1987). Relative ratio of 14kDa message (M14) to the total mRNA pool continuously increased as myelination proceeds, while the 21.5kDa message (M 21.5) decreased in mature myelin. Thus the ratio M14/M21.5 should be increased during myelination. However, the M14/M21.5 ratio was greatly reduced in qk^v homozygote when compared to the age-matched normal control (Li et al., 2000). These facts suggest that

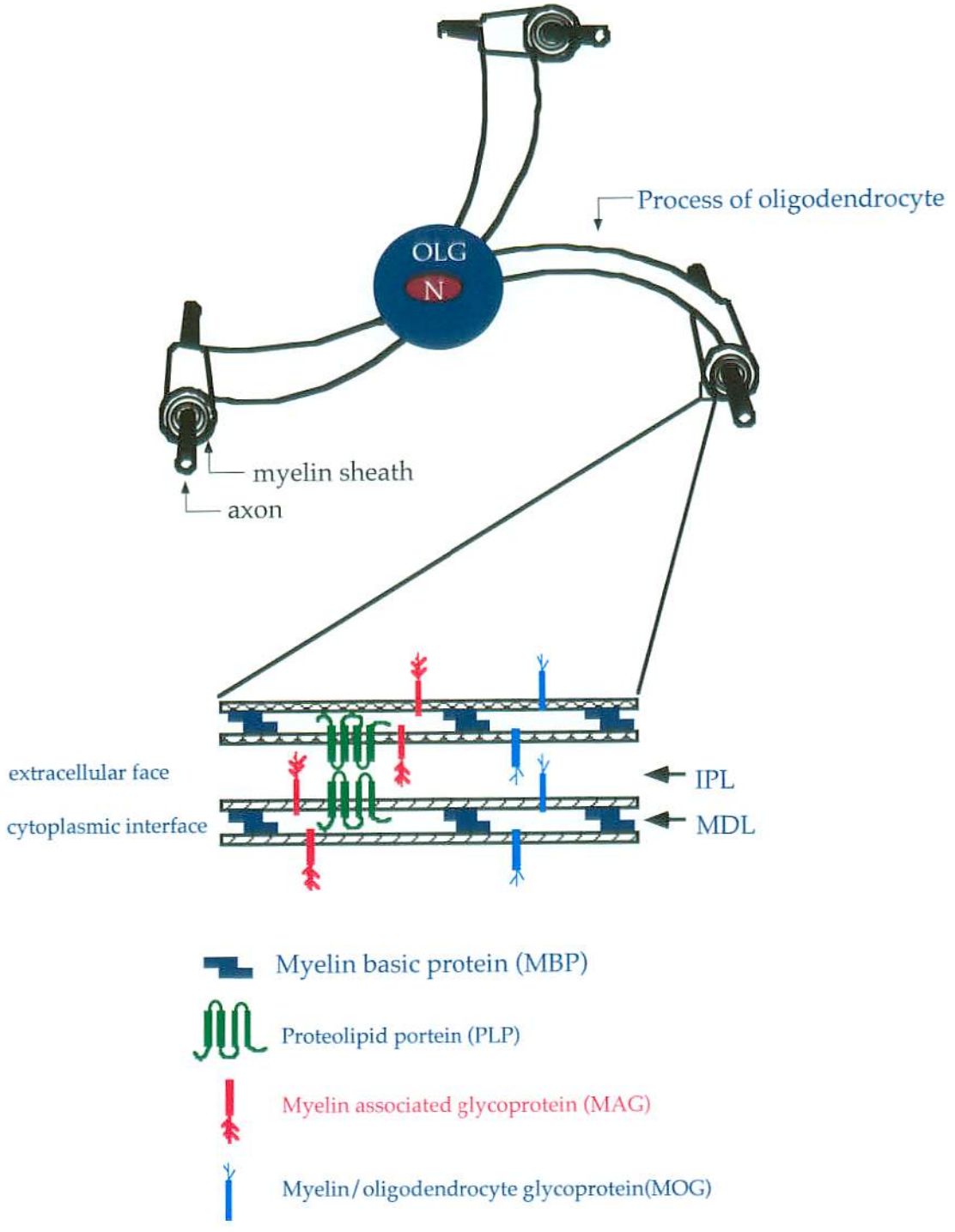

Fig. 7. The oligodendrocyte and the mature myelin sheath. An oligodendrocyte (OLG) processes spiral around the axon during myelination in the CNS to form compact myelin.

The apposing cytoplasmic and extracellular membrane surfaces compact to form the major dense line (MDL) and intraperiod line (IPL) of mature myelin.

hypomyelination in qk^{v} homozygote might result from incomplete differentiation of oligodendrocytes. We do not yet know how *qkI* is involved in these phenomenons. As *qkI* encodes an RNA binding protein, it is fair to say that *qkI* would regulate myelin structural gene expression at the level of mRNA, and expression alteration in *qkI* thus leads to misregulation of gene expression required for proper glial cell differentiation.

1.5 Possible functions of other GSG/STAR family members

Fig. 5 shows that the GSG domain is highly conserved in evolution and should play important roles in multiple developmental processes. The *gld-1*, a homolog of *qkI* in *C. elegans*, is required for normal oogenesis, meiotic prophase progression and sex determination in hermaphrodites (Jones, A.R. et al., 1995; Francis et al., 1995b; Jones, A.R. et al., 1996). Most of the *gld-1* alleles represent partial loss of function mutation. In hermaphrodites, the oogenesis were blocked, and female germ cells were arrested at pachytene stage in meiosis I. In *gld-1* null mutant, germ cells once develop into oocytes that enter the meiotic phase at least until the pachytene stage of meiotic prophase. However, these cells then exit meiosis, re-enter a mitotic cell cycle, and proliferate ectopically to produce an excessive oocytes leading to 'germ line tumor'. The result suggests that GLD-1 act as a negative regulator of oocyte proliferation in this case, and one might call the gene as a tumor suppressor (Francis et al., 1995a).

In *Drosophila, who /how* (*wing held-out/held out wings*) gene which encodes a protein homologous to the mouse QKI protein, was first identified by screening P-transposable enhancer trap mutations. Severity of aberration in these *who /how* hypomorphic mutants varies. Small number of *who/how* homozygotes are viable

and shows phenotype of abnormal wing morphology (Fig.5), and *who/how* is thought to be critical for muscle development (Baehrecke et al., 1997; Zaffran et al., 1997). Recently, it was reported that How functions as a regulator of Stripe, the key modulator of tendon cell differentiation, controlling the transition from premature tendon precursors into mature muscle-bound tendon cells (Nobel-Rosen et al., 2002).

The data provided by the analyses of QKI-like proteins in a variety of organisms indicated that the quaking protein family plays essential roles in cell differentiation/ proliferation during development through the regulation of target genes presumably at mRNA level.

1.6 Biochemical functions of QKI proteins

As described above, quaking protein family is important in many biological processes, and now question is raised. How this protein family actually functions in these processes? Possible functions of the QKI proteins in RNA metabolism or other cellular processes will be discussed below.

1.6.1 QKI may be involved in translational control:

GLD-1 is known to be a repressor of translation for *tra-2*, a sex-determining gene in *C. elegans*, and it binds to regulatory elements called TGEs (for *tra-2* and GLI elements) located in the 3'-UTR of the *tra-2* gene (Goodwin et al., 1993; Jan et al., 1999). Saccomanno et al. recently demonstrated that QKI-6 can act as a translational repressor of *tra-2* in *C. elegans* in the same manner as GLD-1, suggesting that the QKI-6 may have a conserved function in translational control in other organisms as well (1999).

1.6.2 QKI may be involved in mRNA translocation:

MBP mRNA is translocated from cell body of oligodendrocytes and is translated in the processes. There should be a machinery that transports the MBP mRNA to the site of translation, and QKI may be a component of that machinery because 1) QKI is an RNA binding protein, 2) QKI binds to specific sequence in the 3'-UTR of the MBP mRNA (Li et al., 2000), 3) greatly reduced amount of MBP mRNA is localized to the processes of oligodendrocytes in qk^v/ qk^v and majority of the mRNA remained in cell body in contrast to wild type control (Barbarese, 1991).

1.6.3 QKI may be involved in mRNA splicing:

Mouse MAG gene includes 13 exons (Milner et al., 1990), and by alternative splicing of exon 12, two MAG isoforms termed S-MAG and L-MAG are generated, which differ in the length of their carboxyl termini. L-MAG is expressed earlier than S-MAG, which is the predominant form in adult CNS myelin (Frail and Braun, 1984; Lai et al., 1987). However, in qk^v/qk^v mutant, L-MAG is almost absent even at early stage of myelination, while S-MAG is continuously expressed from the early step of glial cell differentiation, suggesting that alternative splicing of the MAG primary transcript was somehow impaired in qk^v (Fujita et al., 1988). Although there is no direct evidence for involvement of the QKI in splicing of MAG, it has been reported that SF1, another mammlian gene homologous to the *qkI* functions as splicing factor, and that the KH domain as well as QUA2 domain were required for this process (Liu et al., 2001). It is not surprising if the QKI act as a splicing regulator of developmentally-regulated transcript also in mouse.

1.6.4 QKI act as an apoptosis inducer?

Recently, Richard and Cheng found that the expression of QKI-7 in NIH3T3 cells induce apoptotic cell death (Chen and Richard, 1998). Subsequently, Pilotte

et al. dissected QKI-7 protein to determine sequences responsible for its apoptotic inducing activity and found that the activity resided in the unique C-terminal 14 amino acids sequence (2001). It is also interesting that heterodimerization of QKI-7 with other isoforms results in nuclear translocation of the QKI-7 and the suppression of apoptosis. Thus the balance between the QKI-isoforms may determine the 'life-or-death' fate of cells expressing *qkI*. Relative ratio of QKI-isoforms dynamically is changed during glial cell differentiation, and again the balance between the isoform levels is likely to regulate cellular physiology or differentiation status of myelinating oligodendrocytes.

Experimental results described above strongly suggest that QKI or its protein isoforms are vital in a number of cellular processes including RNA metabolism and apoptosis, and such multifunctionalities are likely to be generated by the presence of alternating isoforms with possible distinct functions. Relative expression levels of each isoform and subcellular localizations of these isoforms may determine the fate of *qkI*-expressing cells.

1.7 Human *quaking* gene, *Hqk*.

Since *qkI* gene family exists in diverse organisms, it is natural to assume that a *qkI* homolog is present in human, and that the homolog also plays important biological role in human and its deficit may cause disease conditions. Sequences homologous to the *qkI* were indeed found in human EST database, and designated as *Hqk* for human *qk* gene (Zorn et al., 1997). Although mutations in the *Hqk* gene have not been reported, there are increasing evidences that KH type RNA binding proteins are involved in human diseases. For instance, *FMR1*, a gene that encoding a protein containing two KH domains, is responsible for the fragile X syndrome (De Boulle et al., 1993; Ashley et al., 1993; Gibson et al., 1993; Siomi et al.,

1993b., Siomi et al., 1994). Expression of *KOC*, another KH-RNA binding protein, is elevated in human pancreatic cancer (Muller-Pillasch et al., 1997). It is also known that *HCC* (hepatocellular carcinoma), a homologous protein to the *KOC*, is overexpressed in hepatocellular carcinoma (Zhang et al., 1999). Furthermore, it is recently reported that one particular protein isoform of the QKI, QKI-7 is a potent inducer of apoptosis (Pilotte et al., 2001). It is thus possible that the *Hqk* gene is also involved in cell proliferation as well as disease development processes. As an initial step towards understanding the potential roles of *Hqk* gene in disease development, we have isolated both cDNA and genomic clones for the *Hqk*, characterized its genomic structure and determined its expression pattern as well as chromosomal location of *Hqk*. Possible involvement of *Hqk* in human glial tumors was also investigated.

Chapter 2. Materials and methods

2.1 Chromosome mapping of *Hqk* 22

2.2 Cloning and sequencing of cDNA of *Hqk* 22

2.3 Determination of *Hqk* genomic structure 23

2.4 Northern blot analysis 24

2.5 Expeossion analysis and mutation screening of *Hqk* in human tumor samples 25

2.1 Chromosome mapping of *Hqk*

The chromosomal location of *Hqk* gene was determined by PCR amplification of a human/rodent cell radiation hybrid DNA panel (Genebridge 4 hybrid panel; Research Genetics, Huntsville, AL), which consists of 90 hybrid cell clones covering the entire human genome. The DNAs were amplified with primers qk7C (5'-ATGAGGCCAAGAAATTCCATG-3') and qk5C (5'- CAAAGGCGATTACCAGTTAAC-3'). The products were digested with *AccI* to distinguish human and hamster sequences. Results of the panel typing was submitted and analyzed at the Radiation hybrid mapping page of the Whitehead Institute for Biomedical Research/MIT Center for Genome Research Web site (http://carbon.wi.mit.edu: 8000/cgibin/contig/rhmapper.pl).
YAC clones containing STS markers adjacent to *Hqk* were isolated. Presence of *Hqk* sequence in these clones were examined by PCR with qk7C and qk5C, or Southern blot analysis with a probe, HQ5K described below.

2.2 Cloning and sequencing of cDNA of *Hqk*

An oligo-dT primed cDNA was made from 5 microgram of total RNA isolated from human fetal heart, brain and adult peripheral blood leukocytes using the SuperScriptII reverse transcriptase (Life Technology, Bethesda, MD). Sequence of human ESTs that showed significant homology with mouse *qkI* were obtained from the EST database of Genebank with the BLASTN search program and were used to design primers for detection of alternative transcripts and filling the gaps of human ESTs sequence clusters.
Primers e2-4 (5'-TGAGCTGCGGAGCCTGCAATAT-3') and qk7E (5'-TGGTCGTGTTATAACAGCTGC-3') were used to amplify coding region for the 7 kb message, and 7-kd (5'-ATGCCAGTCATGCCTGATATT-3') and

qk7B (5'-CATGGAATTTCTTGGCCTCAT-3') for 3'-UTR of the 7 kb transcript. e2-4 (5'-TGAGCTGCGGAGCCTGCAATAT-3') and 6CT-1 (5'-TTTCGTTGGGAAAGCCATACC-3') were employed for amplification of the coding region of the 6 kb message and 6-kd (5'-GGTATGGCTTTCCCAACGAAA-3') and qk7B (5'-CATGGAATTTCTTGGCCTCAT-3') for the 6 kb 3'-UTR. Primers e2-4 (5'-TGAGCTGCGGAGCCTGCAATAT-3') and qk 5D (5'-CAAAGGCGATTACCAGTTAAC-3') were used for amplification of the 5 kb message coding region, while 5-kd (5'-AAAGTTCGAAGGCACGATATG-3') and *Hqk*5b (5'- CTGTGGGTTAATAGAAACAGC-3') were utilized for amplification of the 3'-UTR of the 5 kb transcript. Reactions were performed in a Perkin Elmer 9600 thermal cycler using conditions of 94°C for 5 min; 30 cycles of 94°C for 30 s, 60°C for 30 s, 72°C for 2 min. PCR products were then cloned into the pGEM-T vector (Promega, Madison, WI) and inserts were sequenced using the PRISM Dye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster city, CA). For comparative analysis of 5' and 3'-UTR sequences of *Hqk* and *qkI*, the PIP (percent identity plot) analysis was performed with PipMarker (http://bio.cse.psu.edu/) (Schwartz et al., 2000).

2.3 Determination of *Hqk* genomic structure

Human PAC and BAC libraries were screened with mouse *qkI* cDNA probe and positive clones were selected for subsequent analysis. BAC DNAs were isolated with a QIAGEN Plasmid Kit (QIAGEN , Tokyo) and completely digested with *Sau3AI* and cloned into a plasmid vector. This mini genomic library was further screened with the 3-XK cDNA probe containing exons 3~5 of mouse *qkI* gene (Ebersole et al., 1996). Positive clones were identified and inserts were sequenced

with T7 and T3 primers to determine exon-intron junctions. Boundaries of exon 5′ and 3′-UTR were determined by direct sequencing of PCR products using one of the PCR primers as a sequencing primer. BAC and PAC clones were also directly sequenced to verify the exon-intron junctions. Distances between exons were determined by measuring sizes of PCR products amplified with primers corresponding to each exon. Briefly, 100 ng of PAC DNA template was amplified with an LA-PCR kit (TaKaRa Shuzo, Kyoto) in a 50 μl reaction mixture containing 20 pmol of primers, using the following parameters; 94°C, 4 min for denaturation, then 25 cycles of 94°C, 30 s and 65°C, 12 min. Identities of the PCR products were verified by Southern blot hybridization or DNA sequencing. Sizes of introns longer than 20 kb were determined by restriction mapping of genomic fragments. The result of exon/intron boundaries and the size of introns except for intron 1 were also confirmed by comparing with genomic sequence data recently uploaded at the Sanger Center (http://www.sanger.ac.uk/HGP/Chr6/Chr6_blast_server.shtml).

2.4 Northern blot analysis

To look for expression of *Hqk* in human embryonic tissue samples, total RNAs were isolated from human fetal heart and brain (18-week) with Trizol reagent according to themanufacturer's instructions (Life Technology, Bethesda, MD). The 456-bpRT-PCRproduct was amplified by primers EH5-3 (5'- GAAGATGCAGCTGATGGAG-3') and qk7E (5'- GCAGCTGTTATAACACGACCA-3') for the 7 kb message, The 492-bpRT-PCRproduct was amplified by primers EH5-3 (5′- GAAGATGCAGCTGATGGAG-3′) and 6CT-1 (5′-TTTCGTTGGGAAAGCCATACC-3′) for the 6 kb message, and The 604-

bpRT-PCRproduct were amplified by primers
EH5-3 (5′- GAAGATGCAGCTGATGGAG-3′) and
qk5D (5′-CAAAGGCGATTACCAGTTAAC-3′) for the 5 kb message. PCR products were cloned into the pGEM-T vector (Promega, Madison, WI) and inserts were sequenced.

Human multiple tissue northern (MTN) blots (Clontech, Palo Alt, CA) were used for northern analysis with HQ2-6, a probe containing exon 2 to exon 6 that includes the conserved KH domain. HQ2-6 was made by amplification of human cDNA with primers e-3 (5′-CGGAAAGACATGTACAATGAC-3′) and e6-u (5′-GCATGACAGCGGTCTGTATTT-3′). The same blots were rehybridized with other probes corresponding to different alternative transcripts. Probes specific to the alternative transcripts were made by amplification of the *Hqk* cDNA. For the 5 kb transcript, a probe named HQ5K (Fig. 11B) was amplified using primer pair qk5a (5′-GTTCGTCTTACCATCTAA-3′) and qk5b (5′-AAGCATGGCTTTTACATCCT-3′). For the 7 kb specific probe named HQ7K (Fig. 11B) and primer pair 7-kd (5′-ATGCCAGTCATGCCTGATATT-3′) and qk7H (5'-TGTGCAATAGAATACAGCCTC-3') were used. A Fuji BAS2000 imaging analyzer was used to analyze the hybridization signals.

2.5 Expression analysis and mutation screening of *Hqk* in human tumor samples

We performed PCR amplifications of genomic DNA from a panel of primary tumors including twenty samples of glioma, eight meningioma samples and three schwannoma samples. For all the human samples used in this study, informed consents were obtained. Tumor tissues were taken from surgical operations conducted at the Kumamoto University Hospital and processed by the standard

methods as described by Miyakawa *et al* (2000). Each tumor specimen was first examined pathologically, and the specimens containing a high percentage of tumor cells, generally more than 95%, were used for the subsequent molecular studies. DNA extraction from the tumor tissues was performed as described by Liang *et al* (1994). Peripheral blood samples were obtained from patients with glial tumors, and DNAs were extracted with the DNA Extractor WB kit (Nippon gene, Toyama) according to the protocol provided by the supplier. To search for mutation in *Hqk* of these patients, each exon of the *Hqk* gene was amplified from genomic DNAs of both tumor and blood samples. Primers used for amplification of each exon are summarized in Table I. Sequences of the PCR products were determined by direct sequencing. For Southern analysis, probe HQAX-1, a 349 bp *ApaI-XhoI* genomic DNA fragment containing exon 1 was used. Probes HQ2-6 and HQ5K were also used for the Southern blot analysis.

To look for alteration in *Hqk* expression in human tumor samples, total RNAs were isolated from tumors with Trizol reagent according to the manufacturer's instructions (Life Technology). Five micrograms of total RNA was used for first-strand cDNA synthesis using a cDNA preamplification kit (Life Technology). RT-PCR was performed with the same cycling parameters as described for genomic DNA PCR. Primers used were e2-4 (5'-TGAGCTGCGGAGCCTGCAATAT-3') and qk7E (5'-TGGTCGTGTTATAACAGCTGC-3') for the 7 kb message, e2-4 (5'-TGAGCTGCGGAGCCTGCAATAT-3') and 6CT-1 (5'-TTTCGTTGGGAAAGCCATACC-3') for the 6 kb message, and e2-4 (5'-TGAGCTGCGGAGCCTGCAATAT-3') and qk5D (5'-CAAAGGCGATTACCAGTTAAC-3') for the 5 kb message. PCR

Table I. Details of the Primer Sequences Used to Amplify the Individual Exons of Hqk Coding Region

Exon No.	Forward primer	Reverse primer	Product size
2	ATGATAGAATAGGCCAGGAG	CACACTTCAGTTGATGACTGC	412bp
3	TTCTAGCTGTATACTGTTCCC	AAAGTGCTGGGATTACAGGCG	505bp
4	TTCAATCAAACTTTGAGGACCC	TTAGAAGTATCAGTCAGGTG	459bp
5	AGGCTCCATTATGTTAGCACC	AACAAGATAAGCTGCAAGCCC	576bp
6	GCATTTTTGTGTGATCAGCGC	ACTGAACCTCAGTTACCAAAGTG	673bp
7a	AAATACAGACCGCTGTCATGC	TGGTCGTGTTATAACAGCTGC	1203bp
7b	AAACAGATGCAGACATGTGTG	TGCAAGTAAGTCACTGACTGC	242bp
7c	AAACTACTGTGCCTTAACGTG	TTCCATCAGGTGTGGTGTTGC	294bp
8	ATGAGGCCAAGAAATTCCATG	CAAAGGCGATTACCAGTTAAC	748bp

products were cloned into the pGEM-T vector (Promega) and inserts were sequenced. Some of the tumor samples were analyzed by northern blot hybridization using probes described above.

The nucleotide sequence data reported in this paper have been submitted to DDBJ and have been assigned the accession numbers: AB067798 (human quaking gene, 5 kb mRNA), AB067799 (6 kb mRNA), AB067800 (7 kb mRNA), AB67801 (7 kb-B mRNA); AB067802 (exon 1), AB067803 (exon 2), AB067803 (exon 3), AB 067804 (exon 4), AB067805 (exon 5), AB067806 (exon 6), AB067807 (exon 7), AB067808 (exon 8).

Chapter 3. Results

3.1 Chromosome mapping of *Hqk* 30
3.2 Cloning of human homolog of the *qkI* gene, *Hqk* 30
3.3 Exon-intron organization of *Hqk* 35
3.4 Expression of *Hqk* messages in human tissues 36
3.5 Alterations of *Hqk* expression in human tumors 40

3.1 Chromosome mapping of *Hqk*

The chromosomal location of the *Hqk* gene was determined by use of Gene bridge 4 radiation hybrid Panel. The *Hqk* locus should be located on chromosome 6 in the region between 6q25 and 26 flanked by markers CHLC.GATA81B01 and WI-4442, as shown in Fig. 8. YAC clones positive for these STS markers were isolated, and examined for the presence of the *Hqk* sequence by PCR or Southern hybridization. These analyses confirmed that the *Hqk* was indeed contained in the YAC clones, 837-D-9 and 920-A-3. The date of Human/mouse homology maps that obtained from NCBI indicates the chromosal location for *Hqk* correspond to region of conserved syteny in the mouse genome. (NCBI website: http://www.ncbi.nlm.nih.gov/Omim/Homology/human6.html).

3.2 Cloning of human homolog of the *qkI* gene, *Hqk*

Zorn *et al.* previously reported that sequences homologous to the mouse *qkI* could be retrieved from a human EST database (1997). In order to obtain more complete information on *Hqk* cDNA sequences, we again searched the current human dbEST database. A total of 41 human EST sequences showed significant similarities to the mouse sequences, and were classified into 14 clusters (Fig. 9). For example, an EST sequence (Acc# EB893920) showed strong similarities to the exon 2 sequences of *qkI* , which encodes the amino-terminus of the *qkI*, while other ESTs were homologous to carboxyl-terminus of the *qkI* gene product. To fill gaps between the clusters and to obtain entire sequences for all alternative transcripts, we amplified cDNA templates derived from human RNA using primers specific for 5 kb, 6 kb, and 7 kb alternative messages, and sequenced the products to ensure that they were indeed derived from the *Hqk* gene. We obtained the entire coding

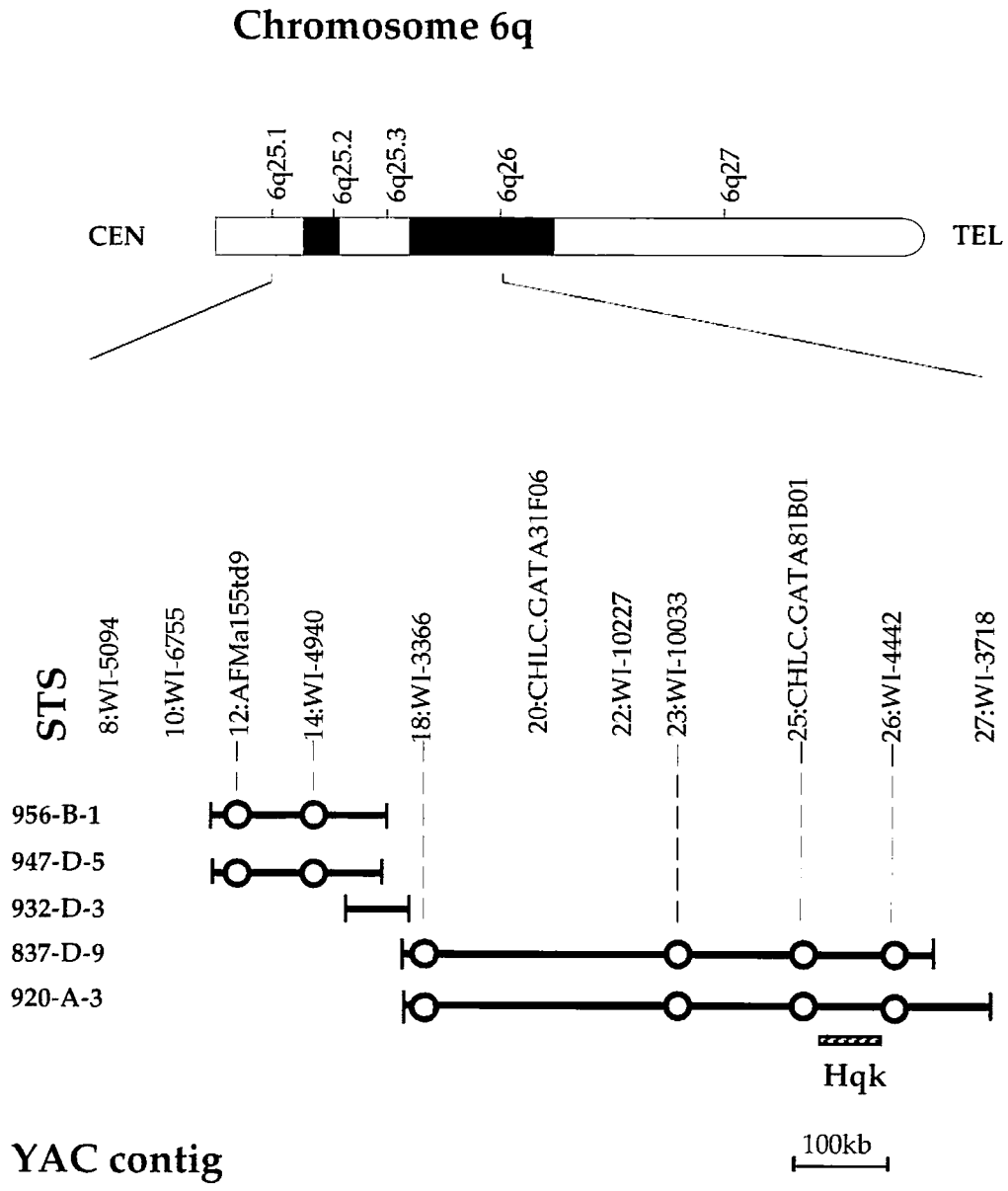

Fig. 8. Chromosomal location of Hqk gene. Location of the Hqk was determined by typing Gene Bridge 4 radiation hybrid panel. Top; long arm of human chromosome 6. Middle; STS markers in the vicinity of the Hqk. Hqk gene was mapped between the markers CHLC.GATA81B01 and WI-4442. Bottom; A YAC contig covering the Hqk genomic region. Bar represents each YAC clone and open circle indicates that YAC clone contains corresponding STS markers. It was confirmed by Southern analysis that the clones 837-D-9 and 920-A-3 contained the Hqk gene.

sequence and a portion of the 3'-UTR sequence for each alternative transcript. The predicted open reading frame is 1023 nucleotides in length for the 5 kb message, 957 nucleotides for the 6 kb message, 975 nucleotides for the 7 kb transcript and 1059 nucleotides for the 7kb-B message. The sequences showed strong similarities to the mouse *qkI* gene (Kondo et al., 1999). The two genes are highly homologous to each other; the amino acid sequence of each isoform of *Hqk* is identical to that of mouse *qkI* (Fig. 10), and 96% identity was found even at the nucleotide level (data not shown). Given the high homology between the human and mouse *qk* gene, it is likely that the ATG triplet found at nucleotide 553 in the human sequence (Acc#: AB067798) represents a start codon of the human *Hqk* gene. Genomic fragments upstream to the start codon were isolated from the *Hqk* -positive BAC clone and sequence of about 4 kb was obtained. The region approximately 800 bp region upstream to the start codon was highly similar to that of *qkI*, which corresponds to the putative intron 1 and a part of exon 2 of the *qkI* gene. Similarity in this 800 bp region is nearly 80%, while the homology sharply drops to ~50% in the further upstream region. Percent Identity Plot (PIP) analysis showed that this region is CpG or GpC-rich (see Fig. 11). We have not precisely mapped transcription initiation site yet. However, it is likely that the 5'-UTR is no larger than 500 bp, since the predicted message size is comparable with the length of the cognate transcript estimated by northern blot hybridization. PIP analysis also showed that the 3'-UTR regions of the *Hqk* were very similar to the corresponding mouse 3'-UTR sequences. The 3'-UTR nucleotide identities between human and mouse are 85%, 73% and 73% in the 5 kb, 6 kb and 7 kb transcripts, suggesting that the 3'-UTRs may share similar gene regulating functions among species.

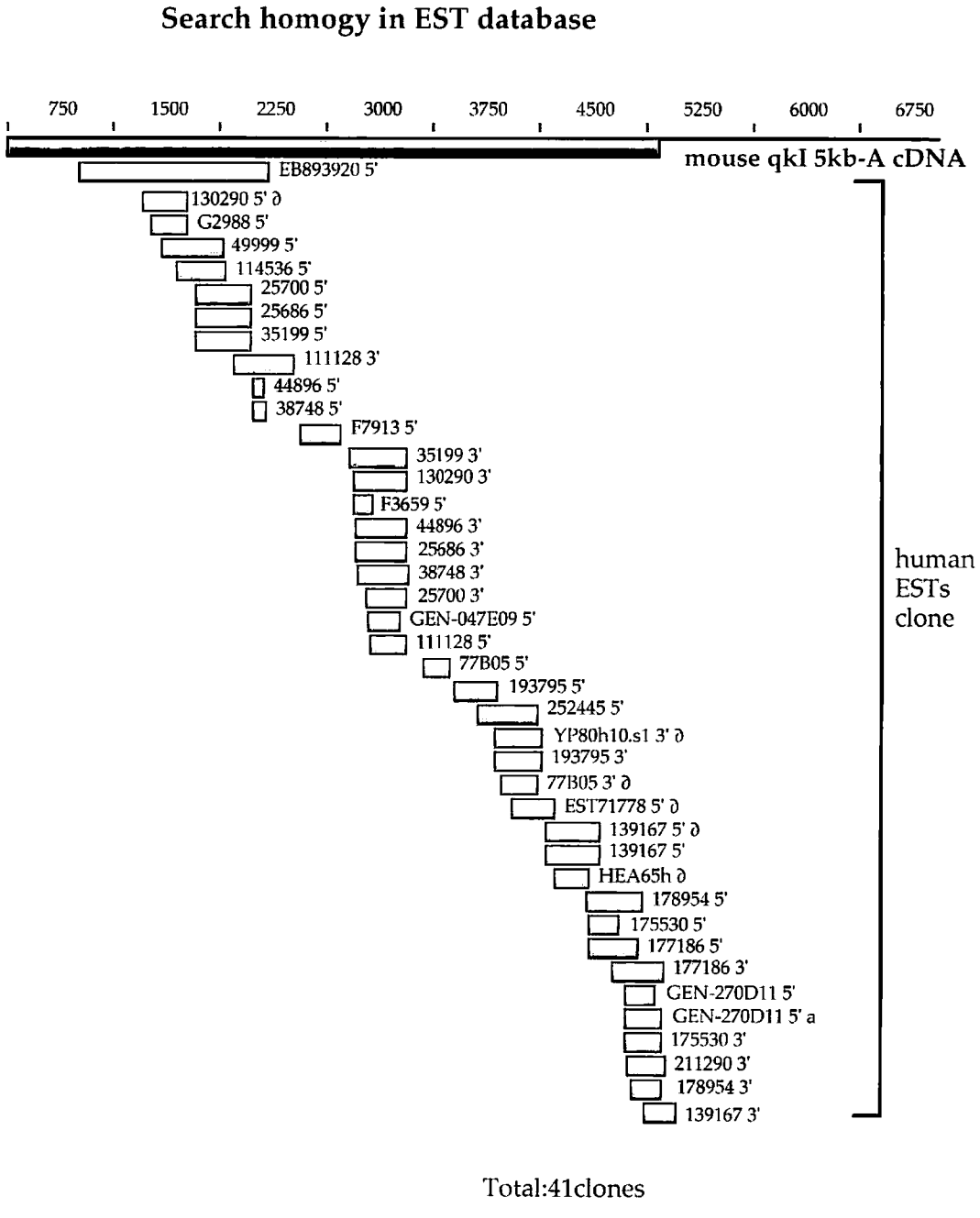

Fig. 9. BLAST search showed that databases have ESTs significant sequence with the 5kb transcript of *qkI* gene.

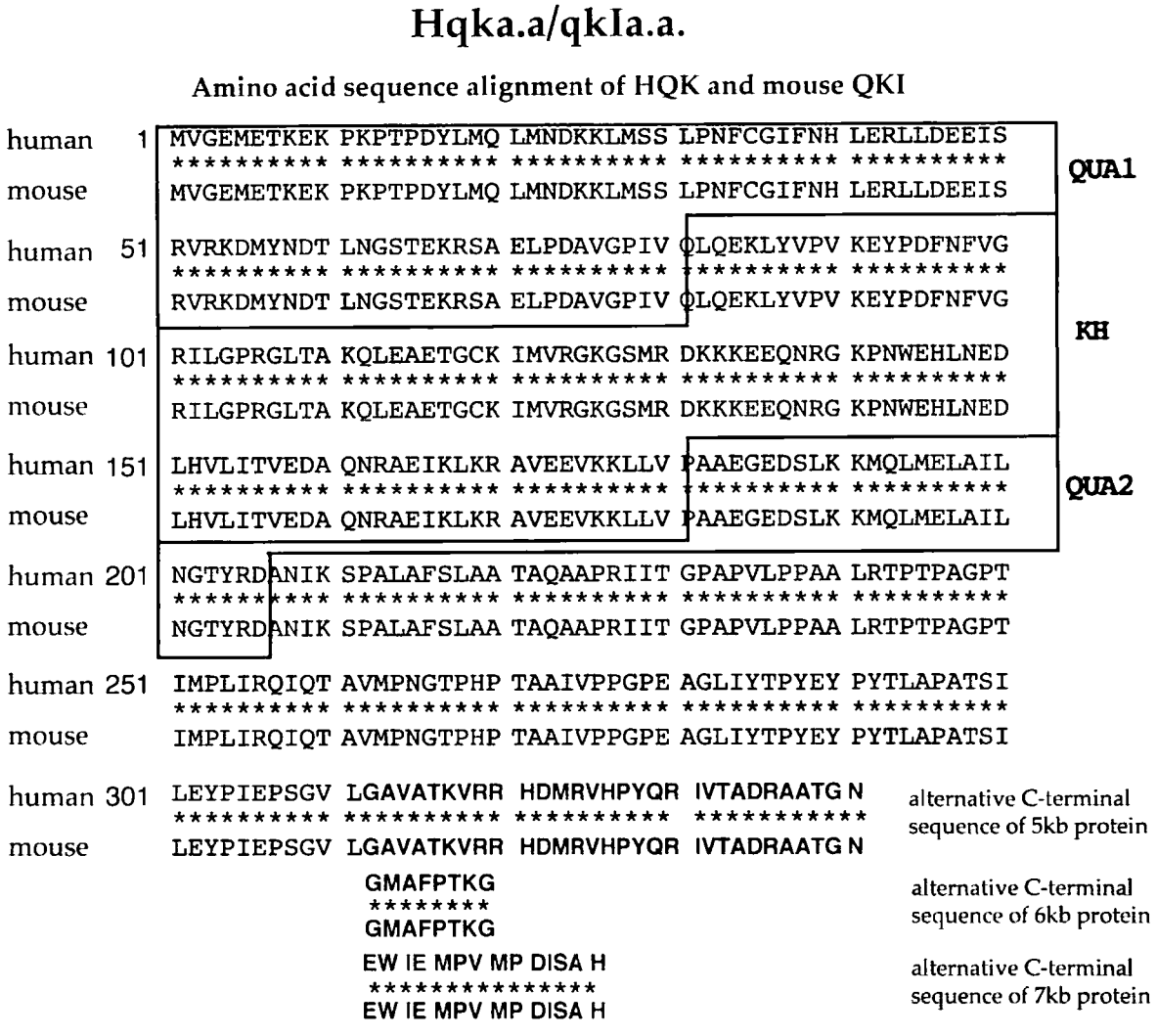

Fig. 10. Amino acid alignments of the predicted proteins of the human and mouse contigs, *Hqk* and *qkI*, respectively. Identities between the two polypeptides are asteristed. The box represents the GSG domain. Residues of the C-terminal amino acids conserved between two genes are indicated in **bold**.

3.3 Exon-intron organization of *Hqk*

To determine the genomic organization of the *Hqk* locus, human PAC or BAC libraries were screened with a mouse *qkI* cDNA probe containing the conserved GSG/STAR domain. We isolated three PAC and two BAC clones, which covered approximately 300 kb of a genomic region including the *Hqk* locus (Fig. 11). We constructed a library of short genomic DNA fragments derived from the PAC and BAC clones, and clones containing *Hqk* exons were selected and sequenced to determine exon-intron junctions. These clones were found to contain exons 2-4. By compiling such genomic analysis data, the exon-intron organization of the *Hqk* was determined to be as shown in Fig. 11. *Hqk* contains eight coding exons and two distinct UTRs spanning approximately 200 kb with large introns of 80 kb between the first and the second exon and of ~56 kb between the third and fourth exon (Fig. 11). The size of the *Hqk* locus is thus much larger than that of mouse *qkI* locus, which is about 70 kb, and the entire locus was covered by three BAC/PAC clones (Fig. 11). This size difference arises from differences in intron length between the two species. The exons range in size from 89 bp (exon 5) to 5687 bp (exon 7), and all of the exons possesses proper splicing donor/acceptor sequences (Table II). However, utilization of the seventh exon is unusual; the exon is used differentially for each alternative transcript (Fig. 12). For the 7 kb transcript, the entire, 5687 bp sequence is used as the exon, of which first 41 bases represents a coding exon followed by a termination codon, tga, and 3'-UTR of 5438 bp in size. For the 6 kb message, the first 1329 bp of exon 7 is spliced out, and a smaller exon of 4200 bp containing 23 bp coding exons and a 4177 bp 3'-UTR is present. For the 5 kb message, only a 75 bp sequence of the exon 7 (exon 7c) is utilized as a part of the coding exons. The first 14 bp sequence of exon 8 belongs

to the coding region of this message. In the case of the 7 kb-B, intron 6 is not spliced out, resulting in C-terminal amino acid sequence distinct from that of the other three protein products (data not shown; see Acc# AB067801). Apart from the intron length difference, the genomic structure of *Hqk* is similar to that of the mouse *qkI*. However, there are minor differences in exon-intron organization between the two. For example, the last 69 bp of the mouse exon 5 corresponds to the upstream part of a novel human exon designated as 5 and the first 20 bp of the mouse exon 6 is homologous to that of the human exon 5 (Fig. 11). From the *Hqk* locus, at least four alternative messages encoding four different protein products are generated (Fig. 12). All of the four alternative transcripts can be found in mice as well (Kondo et al., 1999).

3.4 Expression of *Hqk* messages in human tissues

The expression of *Hqk* was examined by northern blot analysis and RT-PCR. We detected the transcripts of *Hqk* in fetal heart and brain tissues by RT-PCR (data not shown), suggesting the possibility that *Hqk* may play an important role in development, as the *qkI* does in mouse. With a probe containing the highly conserved KH domain sequence, bands with several different sizes were detected on northern blots. Signals were found in most of the tissues examined, except for duodenum. On the other hand, three major transcripts of approximately 5, 6, and 7 kb in size were detected in a variety of human tissues (Fig. 13A), probably corresponding to the 5, 6 and 7 kb messages in mouse. In some of the tissues examined, there were also bands with sizes different from the 5, 6, and 7 kb transcripts, which might represent uncharacterized alternative transcripts. Based on the knowledge of *Hqk* genomic organization, we designed probes specific to

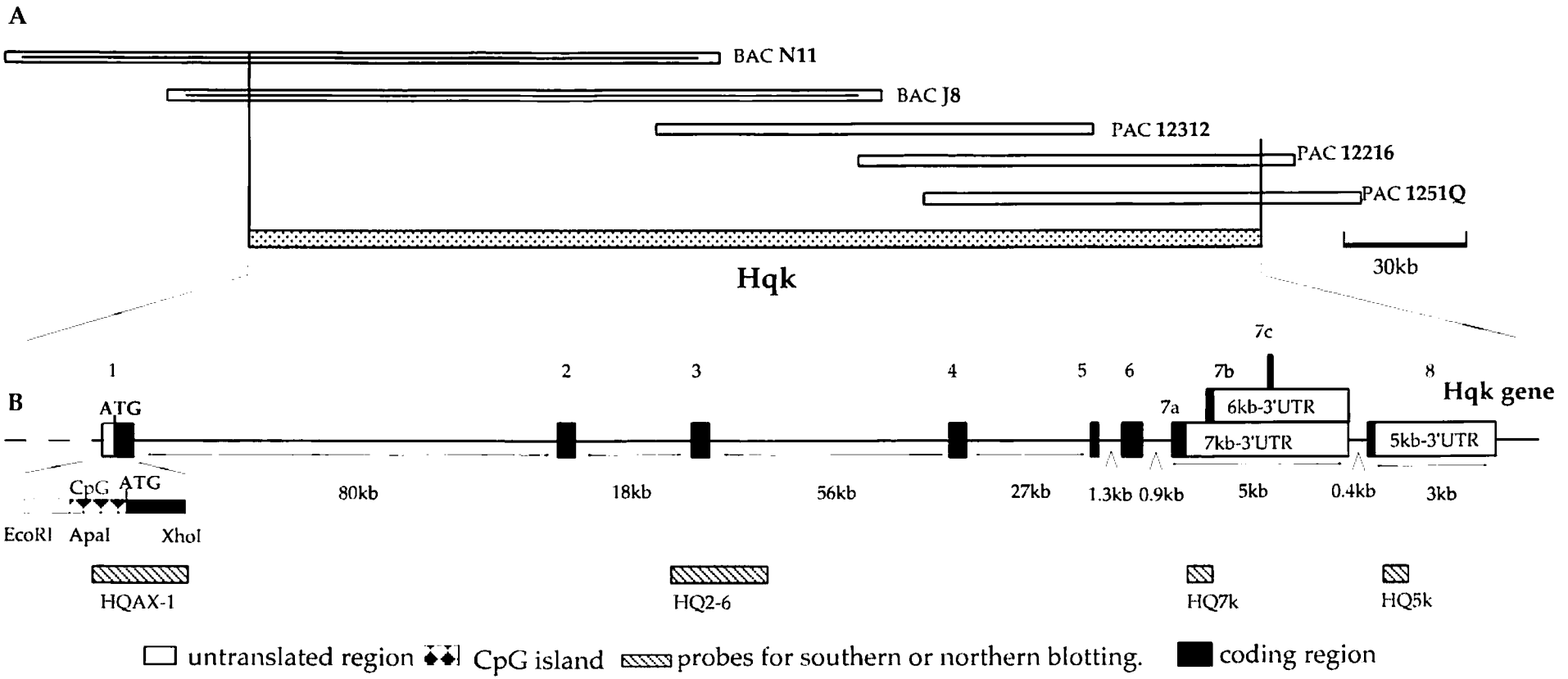

Fig. 11. A) A BAC/PAC contig which covers the entire Hqk locus.

B) Genomic organization of the Hqk gene. The Hqk gene consists of eight exons, of which exon 7 is differentially utilized for each alternative transcript. Boxes indicate exons, and closed boxes show coding exons. 5' UTR may be larger than the one shown in this fugure. CpG islands were found in the region upstream to the translation initiation site. Sizes of intron are shown in kb. Locations of the probes for southern or northern blot analysis are also shown.

Table II **Exon-intron boudary sequences of the Hqk gene**

Intron Sequence	Exon number and Sequence	Intron Sequence	Exon Size (bp)
	----------1----CTGCTGGACGAAG	**gt**gagcgtctccaggg	ND
ttttta**ctttaacag**	AAATTA---2----AGAATACCCAGAT	**gt**aagtatcatgttca	144
ttttttttttctc**ag**	TTTAAT---3----GGATAAAAAAAAG	**gt**aagtccttgaaaat	117
tctgtaaattttt**tag**	GAGGAG---4----GGTACCTCCTGCA	**gt**aagtaataatttcc	144
tgttttccatgta**tag**	GCAGAA---5----AACTTTAAATCAC	**gt**aagaatgagctctg	88
tctttgcttactg**tag**	CAGCCC---6----AGTGGTGTATTAG	**gt**aagttcttctcccc	300
tgccttttttta**tag**	AGTGGA---7a---TATTTCTGACT--	———	* 5687 (*41*+5646)
ttcctttatctt**tag**	GTATGG---7b---AAAGGCTAAGA--	———	* 4200 (*23*+4177)
attgcaatttaac**tag**	GTGCGG---7c---ACCGCAGACCGAG	**gt**tagtttagttctgc	75
gtttctaaccacc**cag**	CCGCCA---8----CGGAACTAACC--	———	* 3183 (*14*+3169)

Consensus sequence for splice acceptor and donor sites are shown in bold face. The stop codon are underlined.
*Exon 7a, 7b and 8 were last exons of the 7 kb, 6 kb, and 5 kb alternative transcripts, respectively. Sizes of the coding sequences in these exons were italicized and shown in the parentheses.

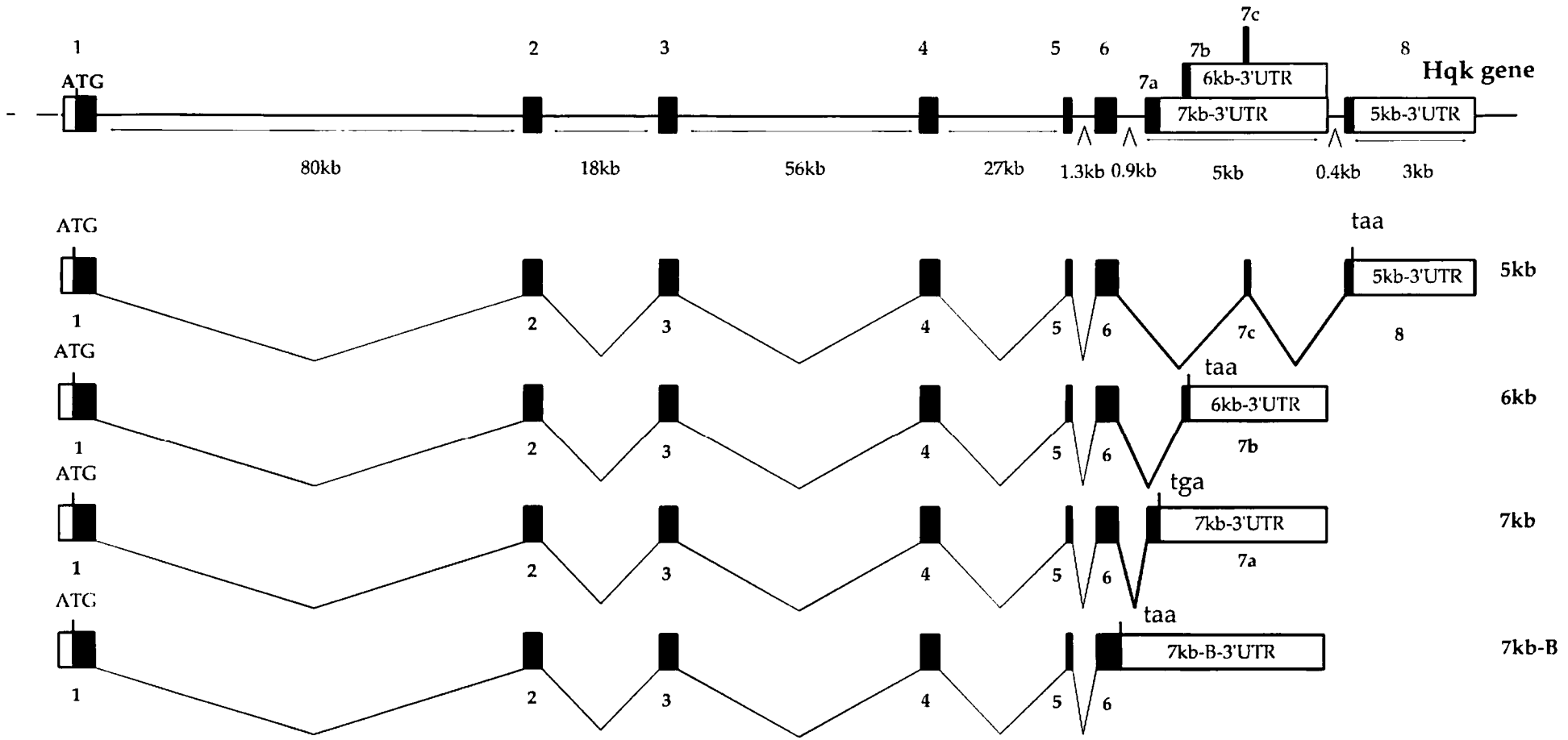

Fig. 12. Genomic organization of the Hqk(Top).Splicing patterns for each alternative transcript, i.e. 5 kb, 6 kb,7 kb and 7 kb-B are shown at th bottom panel.

each transcript and used them separately for northern analysis (Fig. 13, B and C). There were some variations in relative expression level of the alternative transcripts among different tissues. For example, the 5 kb message was abundant in heart, but scarcely expressed in brain. The situation was opposite for the 7 kb message. When hybridized with the probe specific to the 7 kb and the 7 kb-B message, it was found that a band of about 7kb was strongly expressed in brain, while only a weak band was detected in heart. Compared to the 5 kb message, the 7kb message showed a more restricted expression pattern. To look for transcripts, which were undetectable by northern analysis, RT-PCR was carried out. Based on the genomic structure of the mouse *qkI*, primers were designed to distinguish six alternative transcripts in human, if they exist. In addition to the major transcripts of the5, 6, and 7kb messages, a novel transcript designated 7kb-B were detected, while other minor transcripts found in mice were not detected in human fetal heart, brain, and adult leukocytes (data not shown). These data suggest that three major alternative transcripts are produced from the *Hqk* locus, and that their expression levels of each transcript vary among human tissues tested.

3.5 Alterations of *Hqk* expression in human tumors

Cytogenetic studies suggest a high incidence of genomic deletion in the long arm of human chromosome 6 in various tumors including glioblastomas and astrocytoma (Miyakawa et al., 2000; Liang et al., 1994; Mitelman et al., 1997). To investigate whether the *Hqk* is involved in tumor formation, we first looked for genomic deletion of *Hqk* in tumor samples including twenty gliomas, eight meningiomas and three schwannomas. PCR primer pairs that flank each exon were used to detect homozygous genomic deletions, but no such deletions were

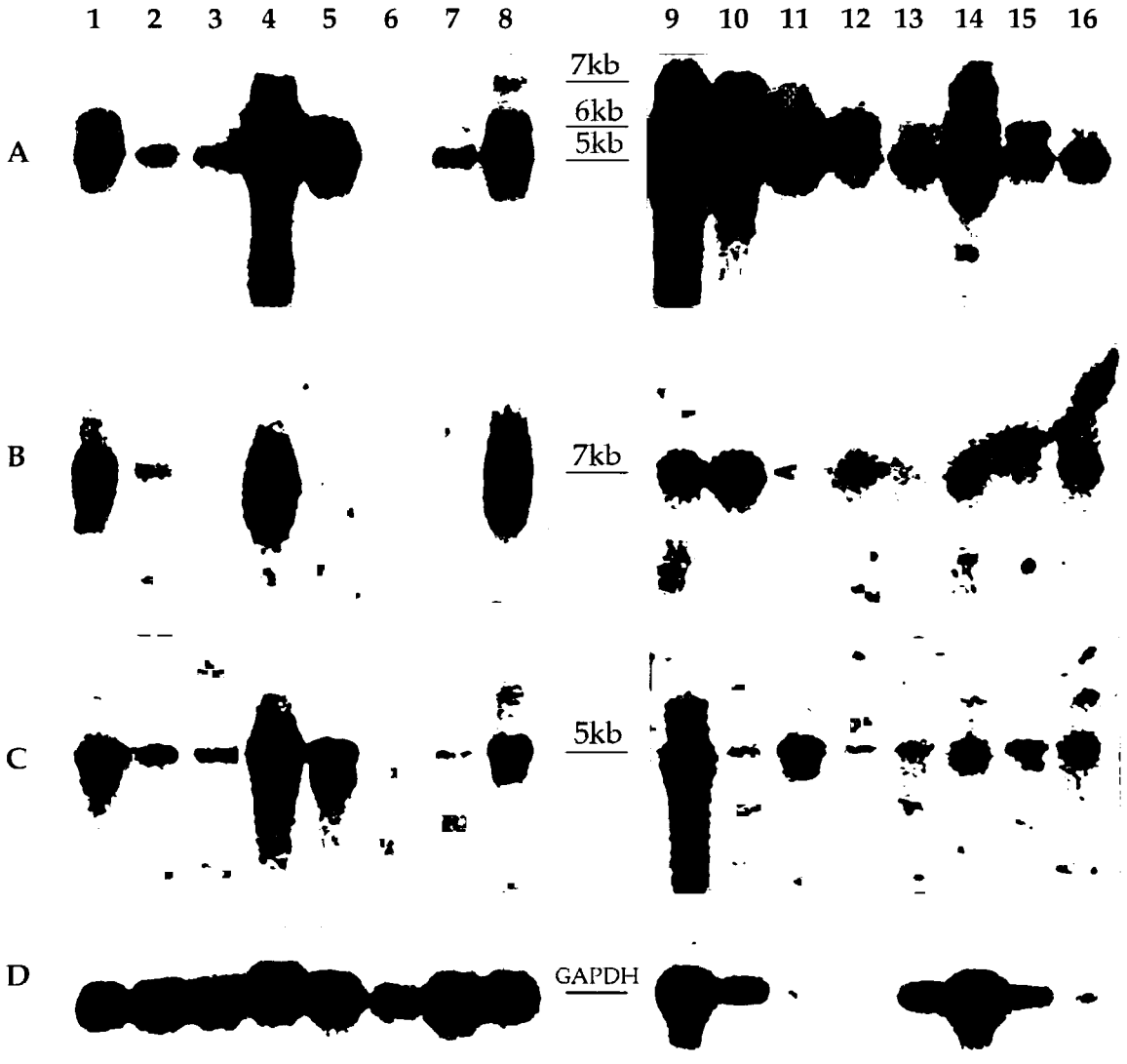

Fig. 13. Northern blot analysis of Hqk expression in various human tissues. Human multiple tissues Northern blots were hybridized separately with a Hqk cDNA probe containing conserved KH domain (A), a probe specific for the 7 kb message as well as 7 kb-B message (B), or a 5 kb message specific probe (C). Poly A+ RNAs (2 ug/lane) from spleen (lane 1), thymus (lane 2), prostate (lane 3), testis (lane 4), ovary (lane 5), small intestine (lane 6), colon (mucosal lining)(lane 7), peripheral blood leukocytes (lane 8), heart (lane 9), brain (lane 10), placenta (lane 11), lung (lane 12), liver (lane 13), skeletal muscle (lane 14), kidney (lane 15), and pancreas (16) were loaded. Integrity and amount of RNA loaded onto the Northern blots were examined by probing with a GAPDH cDNA. Position of the 7 kb message is indicated by triangle in lane 10 of panel A and B.

found in the tumor samples tested. Point mutations that would affect amino acid sequence or splice-site consensus sequences were not found in the PCR products. Genomic Southern blot analysis also failed to detect deletions in the *Hqk* locus. Then, we examined the *Hqk* expression in a panel of cDNA samples derived from the tumors described above. In human brain, the 7 kb transcript is highly expressed as shown in Fig. 13. RT-PCR assay detected all the major transcripts in normal brain samples. We, however, found that expression of some of the *Hqk* transcripts was altered in brain tumor samples (Fig. 14 and 16). In six specimens out of 20 glioma samples, *Hqk* expression was clearly downregulated. For example, in samples No. 5, 17, and 20, the 7 kb transcript was specifically missing, and all three major transcripts were not detected in the samples, No. 8 and 19. Only the 5 kb transcript was downregulated in sample No. 18.

In contrast, no alterations were found in three schwannomas and in eight meningioma samples (Fig. 15). Northern blot analyses using RNAs from samples No. 2, 5, 8, 12, and 17 unambiguously showed that the 7kb transcript was not expressed in the samples No. 5 , 8, and 17. (Fig.16). On the other hand, a band of almost the same size as those in samples 2 and 12 was clearly observed in lane M, on which mouse brain RNA was loaded. This band in lane M represents the 7 kb message of *qkI*, as the 7 kb-B transcript is scarcely expressed in mouse brain (Kondo et al., 1999).

The above data suggest that the *Hqk* messages, especially the 7 kb transcript, was downregulated in some of the glial tumors, but not in the tumors of other origins (Table III).

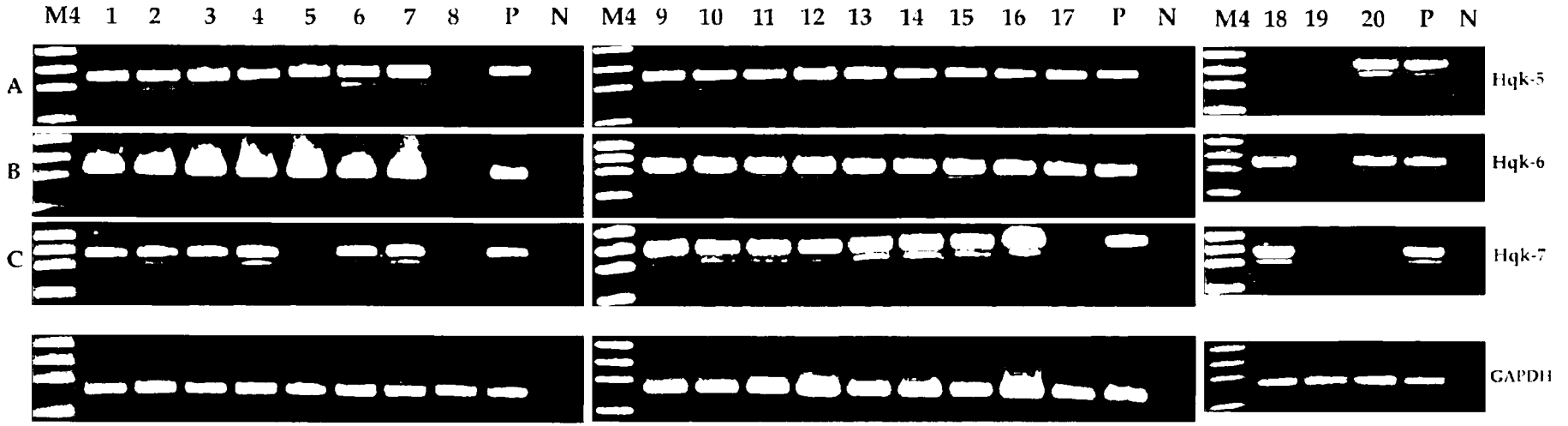

Fig. 14. Expression analysis of the Hqk in various brain tumors and meningioma.

A Detection of the Hqk messages in brain tumors by RT-PCR. cDNAs derived from twenty glioma samples as well as normal control brain RNA were amplified with primers e2-4/qk5D for the 5 kb message (panel a), e2-4/6CT-1 for the 6 kb message (panel b), and e2-4/qk7E for the 7 kb message (panel b). Sizes of the PCR products are 1199 bp, 1068 bp and 1043 bp for the 5 kb, 6 kb, and 7 kb transcript, respectively. Note that expression of the 5 kb message in sample No.18 (glioblastoma) was greatly reduced, while the 7 kb transcript was missing in the samples No. 5 and 20. All the three transcripts were absent in samples No. 8 and 19. P; As a positive control, cDNA template derived from normal brain tissues was used. N; Human genomic DNA was used as a negative control, since the corresponding genomic region is too large to amplify by PCR. M4; molecular weight marker, _X174 HaeIII digest. Primers for GAPDH were used for normalization of cDNA level in each sample. Tumor samples used in this experiment were classified as follows: Sample No. 1 and 10, pilocytic astrocytoma; No. 5, 8, and 12, anaplastic astrocytoma; No. 2 and 7, oligo-astrocytoma; No13, anaplastic oligodendroglioma; No. 3, 4, 6, 7, 9, 14, 16 and 17, glioblastoma.

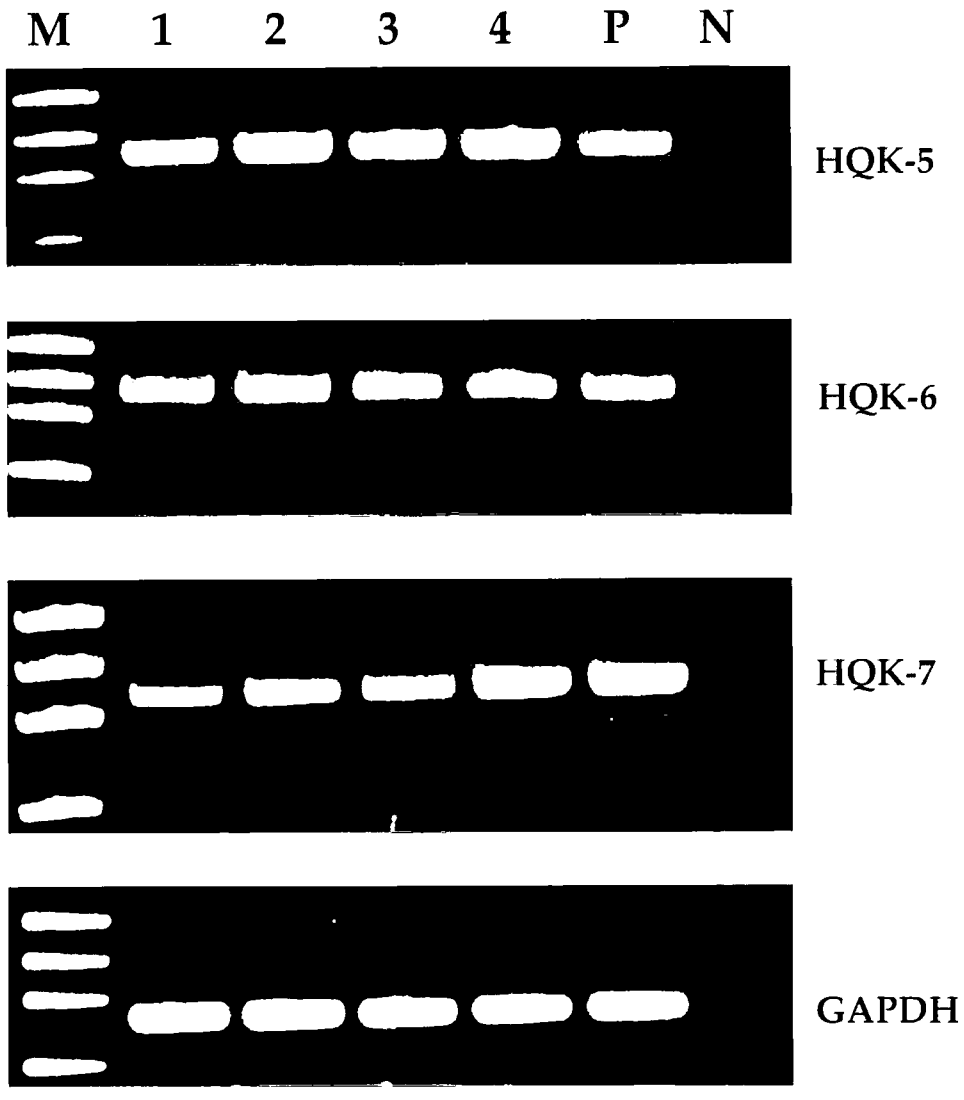

Fig. 15. RT-PCR analysis of the Hqk messages in meningioma.

cDNA synthesis and RT-PCR amplification were conducted for five meningioma samples as described above. No expression alterations were found in these non-glioma samples.

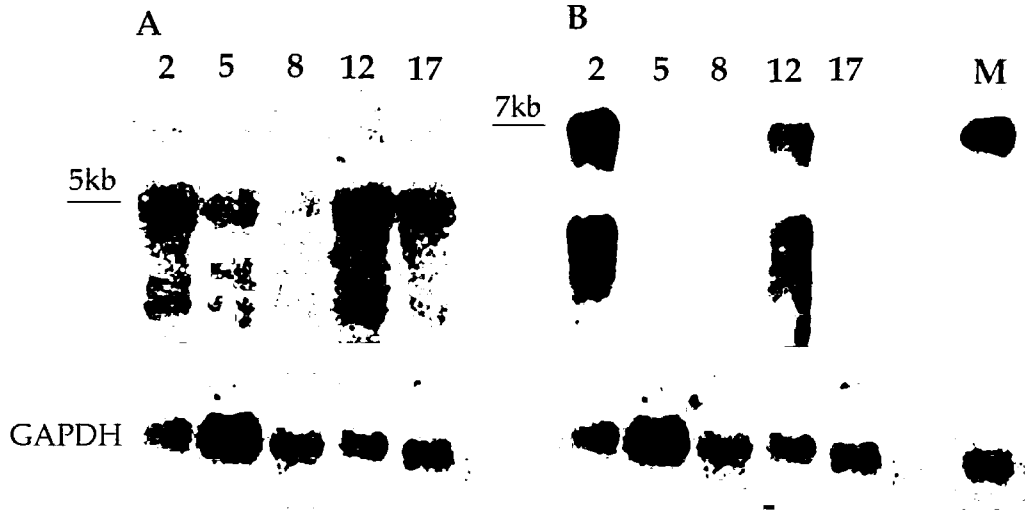

Fig. 16. Northern analysis of Hqk expression in brain tumor samples. For the seven glioma samples, in which Hqk expression alterations were detected by RT-PCR, Northern blot analysis was further conducted. Seven microgram of total RNA extracted from the tumor samples were loaded, and probed with the 5 kb specific probe (panel A) or the 7 kb probe (panel B). In lane M of panel B, mouse brain RNA was loaded as a positive control. The 7 kb probe detected the 7 kb message of mouse qkI and the 7 kb Hqk band in samples 2 and 12, while the 7 kb message was absent in the samples #5, 8 and 17. Lower band in lane 2, 12 or 17 probably represents cross hybridization with 28S ribosomal RNA. Hybridization results with the GAPDH probe (lower panel) revealed that comparable amount of RNA was loaded in each lane.

Table III. **Summary of Hqk Expression in Human Brain Tumors**

Tumor No.	Tumor type	Expression of Hqk-5		Expression of Hqk-6		Expression of Hqk-7	
		RT-PCR	Northern blot analysis	RT-PCR	Northern blot analysis	RT-PCR	Northern blot analysis
5	A A	+	+	+	N D	–	–
8	A A	–	–	–	N D	–	–
17	GMB	+	+	+	N D	–	–
18	GMB	–	N D	+	N D	+	N D
19	GMB	–	N D	–	N D	–	N D
20	GMB	+	N D	+	N D	–	N D

(+): normal level expression; (-): no Hqk expression; ND, not determined.

A A : Anapalstic astrocytoma. GBM : Glioblastoma multiforme.

Chapter 4.

4.1 Discussion 48

4.1 DISCUSSION

In this study, we described the isolation of *Hqk*, a human homolog of mouse *qkI* gene, and its genomic organization and expression pattern in human tissues. Comparison of the genomic structure of *Hqk* with that of *qkI* revealed remarkable similarities between the two genes. It is striking that there is no sequence differences at the amino acid level, and that there is 96% identity at the nucleotide level for the three major transcripts (Fig. 10). Moreover, both 5'-UTR and 3'-UTR sequences are highly conserved; for example, parts of the 3'-UTR of *Hqk* are 85-97% homologous to that of mouse *qkI* (data not shown). Such high sequence conservation suggests that this region may play an important role in the stability, localization, and translational control of the *Hqk* gene.

The *Hqk* gene consists of 8 exons distributed over a genomic region, of approximately 200 kb. The exon sizes and the exon boundary locations are conserved between human and mouse. However, minor differences in the exon/ intron organization were apparent. For example, comparison between the predicted *HQK7-B* and *QKI 7-B* sequences (Kondo et al., 1999) showed that the human protein is larger, with 4 additional amino acids at the carboxy-terminus (data not shown). Whether this particular isoform is translated in human tissues is not known at present. In the 5′ upstream region of the *Hqk* gene, we failed to find the intron corresponding to intron 1 of *qkI* (Kondo et al., 1999). Instead, larger exon 1 which includes sequences corresponding to mouse exon 1, intron 1 and exon 2 is present in the *Hqk* gene. We obtained a 4 kb sequence upstream to the putative translation start site. Comparative analysis demonstrated high sequence similarity in this region, especially in the ~800 bp region uppstream to the start codon (nearly 80%). Given this high similarity, it is possible that the 800

bp region might contain cis-element(s) required for proper *Hqk* expression or the initiation site for *Hqk* transcription. However, in the 4 kb genomic region containing the 800 bp sequence, no TATA or other sequences related to promoter functions could be detected, suggesting that the *Hqk* may be driven by TATA-less promoter (Kondo et al., 1999).

Northern blot analysis and RT-PCR (data not shown) showed that *Hqk* messages are expressed not only in adult, but also in fetal heart and brain, suggesting possible roles in human development. We noticed that the *Hqk* expression pattern was different from that of mouse, though this might be due to differences in age or sex of individuals from which the RNA samples were derived. In human brain, the 7 kb transcript is more abundant that the 5 kb transcript, while the 5 kb message is predominant in heart. In mouse, similar levels of the three major isoforms were detected in brain. In addition, we found that the 7 kb transcript in testis is shorter, when the northern blot analysis was performed using the 7 kb specific probe, indicating another tissue-specific alternative transcription in this gene. Similarly, a band larger than the 7 kb detected in heart (Fig. 13A and 13B) might represent an uncharacterized alternative transcript as this does not correspond to either the 7 kb or the 7 kb-B message.

Members of the *quaking* gene family are thought to play important roles in various biological processes such as development, cell differentiation and cell proliferation. Mutations in the *qkI* homolog, e.g. *gld-1* in nematode or *who/how* in fruit fly cause abnormal cell proliferation or cell migration. In the CNS of mice, reduction of *qkI* expression results in a hypomyelination phenotype (Hardy et al., 1998). Considering these facts, it is likely that alterations in the *Hqk* gene itself or its gene expression level may lead to disease conditions in human. We have

assigned the *Hqk* gene to the chromosomal region, 6q25-26, to which no humans neurological diseases have been mapped. However, a high incidence of chromosomal aberrations at 6q23.3 to 26 in astrocytic tumors has been reported (Miyakawa et al., 2000; Liang et al., 1994; Mitelman et al., 1997). Introduction of a fragment of normal chromosome 6 into a tumor cell line with 6q deletion resulted in suppression of tumorigenicity, suggesting the presence of tumor suppressor gene(s) on the corresponding chromosomal region (Trent et al., 1990). On the other hand, a null mutation in the *gld-1* gene, a *C. elegans* homolog of *qkI*, leads to excessive oocyte growth, and partial loss-of-function mutation in *gld-1* causes oocytes to arrest at the pachytene stage (Jones et al., 1995; Francis et al., 1995a). These results clearly suggest the involvement of the *gld-1* in the regulation of cell proliferation. More importantly, Pilotte *et al.* recently reported that the *QKI-7* isoform could act as a potent apoptosis inducer *in vitro.* (2001). *QKI* protein and possibly *Hqk* gene product function as RNA-binding proteins. There is increasing evidences for the involvement of RNA-binding proteins in tumorigenesis. *MCG10*, a p53 target gene, encodes a protein containing two KH domains, which can suppress cell proliferation by inducing apoptosis and cell cycle arrest at G2-M. Thus this KH domain RNA-binding protein is considered as a tumor suppressor (Zhu et al., 2000). On the other hand, a KH domain protein, *KOC*, was overexpressed in cancers (Zhang et al., 1999). *Musashi 1* containing two ribonucleoprotein (RNP)-type RNA-binding domains (RBDs) is also overexpressed in malignant gliomas (Tota et al., 2001). These results suggest that various type of RNA-binding proteins are involved in cell proliferation, and such expression alterations of those RNA binding proteins often cause tumor formation. This prompted us to examine the expression of *Hqk* in human tumors. Since *Hqk*-

7 is highly expressed in brain, and *Hqk-7* is likely to be an apoptosis inducer (Pilotte et al., 2001), we suspected that loss or reduction of *Hqk-7* expression would be detected in brain tumor samples. Our results indeed demonstrated a high incidence of expression alterations of *Hqk* in gliomas (30%; 6/20), but not in other tumors such as schwannomas (0/3), or meningiomas (0/8). Among the tumor samples showing expression alterations, two were devoid of all three major transcripts, one was missing only the 5 kb message, and only the 7kb message was absent in the last two samples (Fig. 14 and 15). Studies on mouse *qkI* genomic structure (Kondo et al., 1999) suggest that three major alternative messages are transcribed from a single transcription initiation site, and that the expression of each isoform is regulated at the level of splicing. Thus, genetic or epigenetic changes in the promoter region would result in loss of the three alternative transcripts. Specific loss of one of the alternative messages can be caused by alterations in the sequences required for proper splicing. Although abnormal expression of the Hqk is frequently found in glioma, it may not be enough to establish a causal link between Hqk and glioma progression. We failed to detect genetic abnormalities in the tested region of the Hqk gene derived from the patients whose Hqk expression were clearly downregulated. Knudeson's two hit theory predicts that tumor progression occur when both alleles of tumor suppressor gene is deleted or inactivated (Fig. 17) (Knudoson, A.G., 1971), and there are number of examples that lend support this hypothesis, e.g. loss of two copies of Rb or p53 leads to many types of cancer development (Weinberg ,R. A., 1991; Malkin, D., 1994; Srivastava et al., 1992). Genetic mutations/defects found in that these cases are point mutations within the genes or deletion of a portion of the corresponding genes. Although LOH analysis has often identified

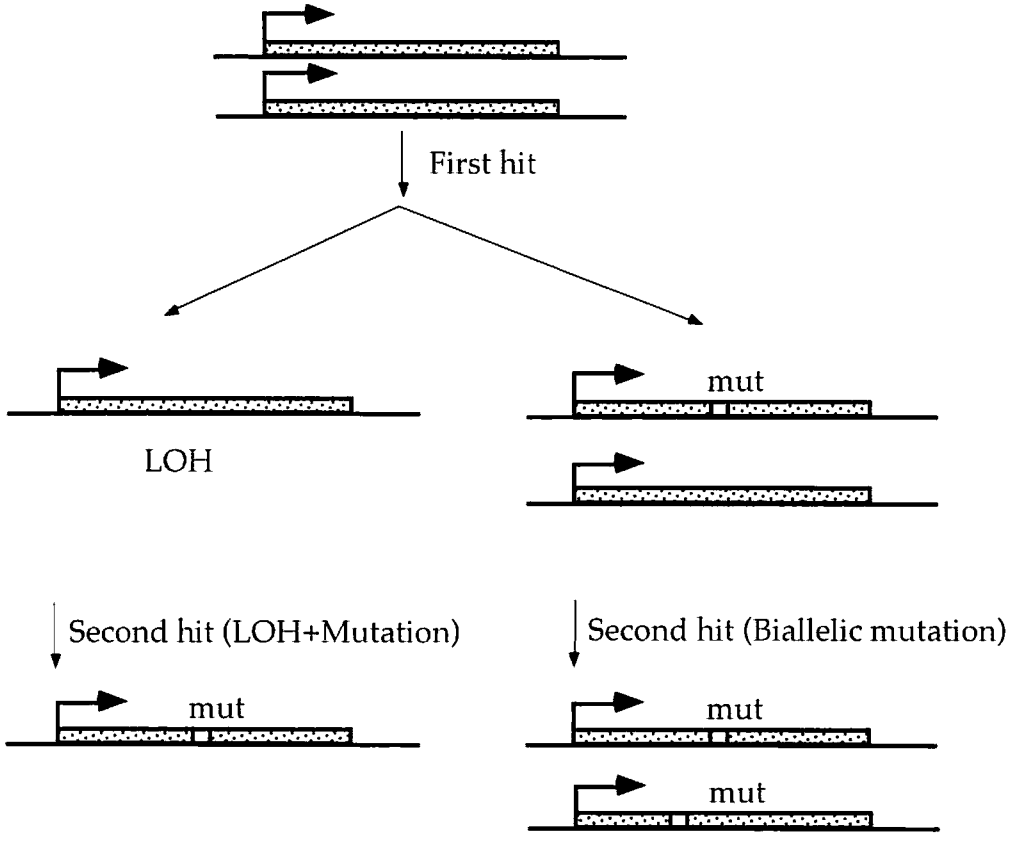

Fig. 17. Models for the effects of inherited tumor suppressor gene mutations in tumorigenesis. The upper model shows a simplified version of the Knudoson's' two-hit' hypothesis. In this model, loss of function in both tumor suppressor alleles is prerequisite for tumor formation.
the lower model is a modified version of the 'two hit' model which hypothesized that the first step of gene inactivation is shown as a localized mutation or LOH. The second hit is shown by LOH or another mutation.

chromosomal loss of tumor suppressor in many of the common cancers (Orborne et al., 2000), there are also many cases in which no obvious genetic mutations were detected in the genome. These may be explained by technical difficulties in finding a subtle, but important mutations in the genome, or some genes may have haploinsufficiency nature of gene activity (Faro et al., 1998). Also, there are growing evidences that 'epigenetic' silencing rather than genetic mutation is a common mechanism for loss of tumor suppressor gene activities (Ohtani-Fujita et al., 1997; Prowse et al., 1997). Epigenetic silencing of gene expression through DNA hypermethylation are apparently very important for generation of various types of cancer (Jones P. A. et al., 1996; Veigl et al., 1998). The fact that methylation of CpG islands located in the 5'-regulatory region of genes can cause transcriptional silencing, coupled with the observations that DNA methylation patterns are frequently perturbed in cancer cells, has lead to the notion that abnormal methylation of the promoters of tumor suppressor genes might be causally related to carcinogenesis. Recent studies combining DNA methylation study and LOH or mutational analyses indicated that constitutive hypermethylation on tumor suppressor gene is functionally equivalent to genetic loss of the gene activity (Fig.18). Frequency for finding promoter methylation contributing to gene silencing varies and ranges from about 9% for Rb gene (Ohtani-Fujita et al., 1997) to 84% for MLH1 gene (Herman et al., 1998). Therefore, it will be very important to study methylation status of the Hqk gene, particularly in the 5'-regulatory region, in which CpG islands have been identified. Because of the limitations in materials from the glioma patients, we could not perform detailed examinations on methylation patterns of the Hqk locus in this study. We, therefore, would like to conduct larger scale analyses on genetic and epigenetic

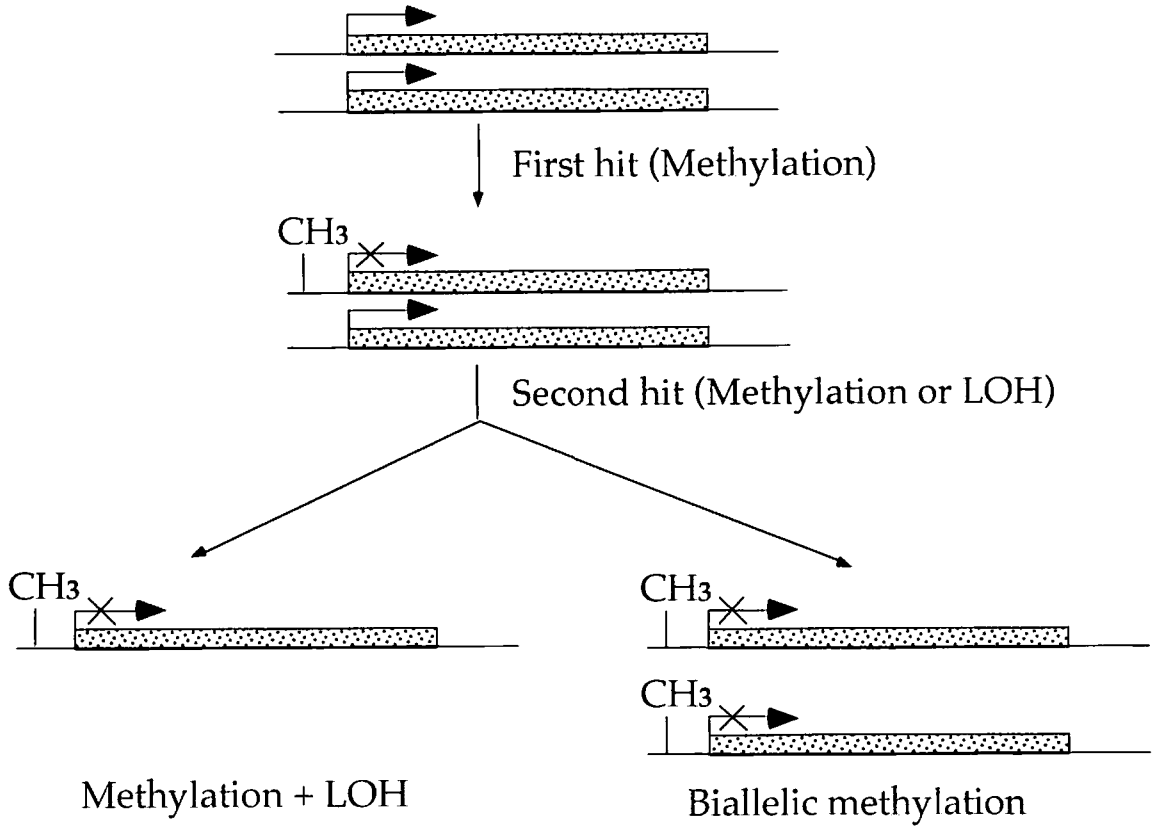

Fig. 18. The model is a modified version of the 'two hit' model which hypothesized that the first step of gene inactivation is shown as a transcriptional repression by DNA myethlation. The second hit is shown by LOH or transcriptional silencing.

mutations within the *Hqk* locus to gain more concrete evidences for the involvement of the *Hqk* in glioma-development. Our examinations on genomic DNA from the glioma patients failed to detect apparent genetic DNA rearrangements or point mutations in coding sequences or splicing donor/ acceptor sequences. However, epigenetic changes associated with tumorigenesis may cause silencing in *Hqk* gene expression or alterations in splicing regulation (Kane et al., 1997). In any case, frequent loss of *Hqk-7* in the high-grade glioma (Fig. 14; Fig 15) may suggest that this expression alteration is directly related to malignant glioma formation. Recent findings by Pilotte *et al.* (Pilotte et al., 2001) lends support this notion. The authors reported that among three major *qkI* isoforms, only *QKI-7* could induce apoptosis *in vitro* and that a life-or-death of *qkI* expressing cells would be determined by the balance between the *QKI* isoforms generated by alternative splicing. The authors claim that *QKI-7* is a sufficient apoptotic inducer if present in cytoplasm, while nuclear translocation of the *QKI-7* suppresses its apoptotic inducing ability. *QKI-5* predominantly localized in nucleus can form heterodimers with *QKI-7* (Pilotte et al., 2001) and could cause the nuclear translocation of *QKI-7*, resulting in cell survival. Relative level of each isoform would be crucial for cell death/proliferation switching. It is thus possible that not only the total loss of *Hqk* transcripts but alterations in balance between the isoform levels discovered in this study may result in progression of glial cell tumors. It will be important to screen *Hqk* gene abnormalities in glioma samples at a larger scale, and to elucidate its precise roles in tumorigenesis using mouse models. Isoform-specific knockout or conditional disruption of *qkI* gene only in glial cells should help understanding possible role of the *quaking* gene in tumor formation.

References

Ashley, C., Wilkinson, K. D., Reines, D. and Warren, S. T. *FMR* 1 protein: conserved RNP family domains and selective RNA binding. *Science,* **262,** 563-566 (1993).

Barbarese, E., Spatial distribution of myelin basic protein mRNA and polypeptide in Quaking oligodendrocytes in culture. *J. Neurosci.,* **29,** 271-281 (1991).

Barhrecke, E. H. *Who* encodes a KH RNA binding protein that functions in muscle development. *Development,* **124,** 1323-1332 (1997).

Chen, T. and Richard, S. structure-function analysis of Qk1: a lethal point mutation in mouse quaking prevents homodimerization. *Mol. Cell. Biol.,* **18,** 4863-4871 (1998).

Courtneidge, R. and S. Fumagalli. A mitotic function for *Src*?. *Trens Cell Biol.,* 4, 345-347 (1994).

Cruz-Alvarez, M. and Pellicer, A. Cloning of a full length complementary cDNA for a *Artemia salina* glycine-rich protein. *J. Biol. Chem.,* **271,** 13377-13380 (1987).

De Boulle, K., Verkerk, A. J., Reyniers, E., Vits, L., Hendrickx, J., Van Roy, B., Van, den Bos, F., de Grassff, E., Oostra, B. A. and Willems, P. J. A point mutation in *FMR-1* gene associated with fragile X mental retardation. *Nat. genet.,* 3, 31-35 (1993).

Ebersole ,T. A., Rho, O. and Artzt, K. The proximal end of mouse chromosome 17: new molecular markers identify a deletion associated with $quaking^{viable}$. *Genetics.* **131,** 183-190 (1992).

Ebersole ,T. A., Chen, Q., Justice, M. J. and Artzt, K. The *quaking* gene product

necessary in embryogenesis and myelination combines features of RNA binding and signal transduction proteins. *Nat. genet.*, **12**, 260-265 (1996).

Faro, M.L., Randel, E., Gurley, K.E., Roberts, J.M. and Kemp, C. The murine gene $p27^{Kip}$ is haplo-insufficient for tumour suppression. *Nature*, **396**, 177-180 (1998).

Frail, D.E. and Braun, P.E. Two developmentally regulated messenger RNAs differing in their coding region may exist for the myelin-associated glycoprotein. *J. Bio. Chem.*, **259**, 14857-14862 (1984).

Francis, R., Barton, M. K., Kimbel, J. and Schedl, T. *gld-1*, a Tumor suppressor gene required for oocyte development in *Caenorhabditis elegans*. *Genetics*, **139**, 579-606 (1995a).

Francis, R., Maine, E. and Schedl, T. Analysis of the multiple roles of *gld-1* in germline development: Interactions with the sex determination cascade and the *gpl-1* signaling pathway. *Genetics*, **139**, 607-630 (1995b).

Fujita, N., Sato, S., Kurihara, T., Inuzuka, T., Takahashi, Y. and Miyatake, T. Develomentally regulated alternative splicing of brain myelin-associated glycoprotein mRNA is lacking in the *quaking* mouse. *FEBS Lett.*, **232**, 323-327 (1988).

Gibson, T. J., Thompson, J. D. and Heringa, J. The KH domain occurs in a diverse set of RNA-binding proteins that include the antiterminator NursA and is probably involved in binding to nucleic acid. *FEBS Lett.*, **324**, 361-366 (1993).

Gibson, T. J., Rice, P. M., Thompson, J. D. and Heringe, J. KH domains within the *FMR1* sequence suggest that fragile X syndrome stems from a defect in RNA metabolism. *Trends Biochem.Sci.*, **18**, 331-333 (1993).

Goodwin, E. B., Okkema,P. G., Evans,T. C., and Kimble, J. Translational

regulation of *tra-2* by its 3' untranslated region controls sexual identity in *C. elegans*. *Cell*, **75**, 329-339 (1993).

Hardy, R. J., Loushin, C. L., Friedrich Jr, V. L., Chen, Q., Ebersole, T. A., Lazzarini, R. A. and Artzt, K . Neural cell type-specific expression of QKI proteins is altered in the *quakingviable* mutant mice. *J. Neurosci.*, **16**, 7941-7949 (1996).

Hardy, R. J. Molecular defects in the dysmyelinating mutant *quaking*. *J. Neurosci. Res.*, **51**, 417-422 (1998).

Hayasaka, K., Himoro, M., Sato, W., Takada, G., Uyemura, K., Shimizu, N., Bird, T.D., Conneally, P.M. and Chance, P.F. Charcot-Marie-Tooth neuropathy type 1B is associated with mutations of myelin P0 gene. *Nat. genet.*, **5**, 31-34 (1993).

Herman, J.G. et al. Incidence and functional consequences of hMHL1 promoter hypermethylation in colorecal carcinoma. *Proc. Natl Acad. Sci. USA*, **95**, 6870-6875 (1998).

Hodes, M. E., Pratt, V. W. and Dlouhy, S. R. Genetics of Pelizaeus-Merzbacher disease. *Dev. Neurosci.*, **15**, 383-394 (1993)

Hudson, L.D., Berndt, J.A., Puckett, C., Kozak, C.A. and Lazzarini, R.A. Aberrant splicing of proteolipid protein mRNA in the dysmyelinating *jimpy* mutant mouse. *Proc. Natl Acad, Sci. USA*, **84**, 1454-1458 (1987).

Jan, E., Motzny,C. K., Graves, L.E., and Goodwin, E.B. The STAR protein, GLD-1, is translational regulator of sexual identity in *Caenorhabditis elegans*. *EMBO J.*, **18**, 258-269 (1999).

Jones, A. R. and Schedl, T. Mutations in gld-1, a female germ cell-specific tumor suppressor gene in *Caenorhabditis elegans*, affect a conserved domain also found in Src-associated protein Sam68. *Genes Dev.*, **9**, 1491-1504 (1995).

Jones, A.R., Francis, R. and Schedl, T. GLD-1, a cytoplasmic protein essential for oocyte differentiation, shows stage- and sex-specific expression during *Caenorhabditis elegans* germline development. *Dev. Biol.,* **180,** 165-183 (1996).

Jones, P.A. DNA methylation errors and cancer. *Cancer Res.,* **56,** 2463-2467 (1996).

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

Kaname,T., Kondo, T., Imanaka, T., Suganuma, T., Suzuki, M., Ebersole, T. A., Artzt, K., Yamamura, K., Abe, K. *in preparation*

Kane, M., Loda, M., Gaida, G., Lipman, J., Mishra, R., Goldman, H., Jessup, J. M. and Kolodner, R. Methylation of the hMLH1 promoter correlates with lack of expression of hMLH1 in sporadic colon tumors and mismatch repair-defective human tumor cell lines. *Cancer Res.,* **57,** 808-811 (1997).

Knudoson, A. G. Mutation and cancer: statistical study of retinoblastoma. *Proc. Natl Acad. Sci. USA,* **68,** 820-823 (1971).

Kondo, T., Furuta, T., Mitsunaga, K., Ebersole, T. A., Shichiri, M., Wu, J., Artzt, K., Yamamura, K. and Abe, K. Genomic organization and expression analysis of the mouse *qkI* locus. *Mamm. Genome,* **10,** 662-669 (1999).

Lai, C., Brow, M.A., Nave, K-A., Noronha, A.B., Quarles, R., Bloom, F.E., Milner, R.J. and Sutcliffe, J.G. Two forms of 1B236/myelin-associated glycoprotein (MAG), a cell adhesion molecular for postnatal neural development, are produced by alternative splicing. *Proc. Natl Acad. Sci. USA,* **84,** 4337-4341(1987).

Levison, S.M. and Goldman, J. E. Both oligodendrocytes and astrocytes develop from progenitors in the subvetricular zone of postnatal rat forebrain. *Neuron,* **10,** 201-212 (1993).

Li , Z. H., Takakura, N., Oike, Y., Suda, T., Kaname, T., Artzt, K., Yamamura, K. and Abe, K. *in preparation*

Li, Z. Z., Zhang, Y., Li, D. and Feng, Y. Destabilization and mislocalization of myelin basic protein mRNAs in *quaking* dysmyelination lacking the QKI RNA-binding proteins. *J. Neurosci.,* **20,** 4944-4953 (2000).

Liang, B. C., Ross, D. A., Greenberg, H. S., Meltzer, P. S. and Trent, J. M. Evidence of allelic imbalance of chromosome 6 in human astrocytomas. *Neurology,* **44,** 533-536 (1994).

Liu, ZH., Luyten, I., Bottomley, M. J., Messias, A. C., Sophie, H. M., Spranger, R., Zanier, K., Krämer, A. and Sattler, M. Structural basis for recognition of the intron branch site RNA by splicing factor 1. *Science,* **294,** 1098-1101 (2001).

Macklin, MB., Gardinier, MV., King, KD. and Kamp, K. An AG-GG transition at a splice site in the myelin proteolipid protein gene in *jumpy* mice results in the removal of an exon. *FEBS Lett.,* **223,** 417-421(1987).

Malkin, D., Germline p53 mutations and heritable cancer. *Annu. Rev. Genet.,* **28,** 443-465 (1994).

Matthieu, J-M. An introduction to the molecular basis of inherited myelin disease. *J. Inherit. Metab. Dis.,* **16,** 724-732 (1993).

Milner, R. J., Lai, C., Nave, K-A., Montag, D., Farber, L. and Sutcliffe, J.G. Organization of myelin protein genes: myelin-associated glycoprotein. *Ann. NY Acad. Sci.,* **605,** 254-261 (1990).

Miyakawa, A., Ichimura, K., Schmidt, E. E., Vrmeh-Ziaie, S. and Collins, V. P. Multiple deleted regions on the long arm of chromosome 6 in astrocytic tumours. *Br. J. Cancer,* **82,** 543-549 (2000).

Mitelman, F., Mertens, F. and Johansson, B. A breakpoint map of recurrent

chromosomal rearrangements in human neoplasia. *Nat. Genet.*, **special issue, April,** 417-474 (1997).

Morello, D., Dautigny, A., Pham-Dinh, D. and Jolles, P. Myelin proteolipid protein (PLP and DM20) transcripts are deleted in jimpy mutant mice. *EMBO J.*, **5,** 3489-3493 (1986)

Muller-Pillasch, F., Lacher, U., Wallrapp, C., Micha, A., Zimmerhackl, F., Hameister, H., Varga, G., Friess, H., Buchler, M., Beger, H. G., Vila, M. R., Adler, G. and Gress, T. M. Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein. *Oncogene,* **14,** 2729-2733 (1997).

Musco, G., Stier, G., Joseph, C., Morelli, M. A. C., Nilges, M., Gibson, T. J. and Pastore, A. Three-dimensional structure and stability of the KH domain : molecular insights into the fragile X syndrome. *Cell,* **85,** 237-245 (1996).

Newman, S., Kitamura, K. and Campagnoni, A.T. Identification of a cDNA coding for a fifth form of myelin basic protein in mouse. *Proc. Natl Acad. Sci. USA.,* **84,** 886-890 (1987).

Nobel-Rosen, H., Volohonsky, G., Reuveny, A., Zaidel-Bar, R. and Volk, T. Two isoforms of the Drosophila RNA binding protein, How, act in opposing direction to regulate tendon cell differentiation. *Developmental Cell,* **2,** 183-193 (2002).

Ohtani-Fujita, N., Drya, T., Rapaport, J.M., Fujita, T., Matsmura, S., Ozasa, K., Watanabe, Y., Hayashi, K., Maeda, K., Kinoshita, S., Matsmura, T., Ohnishi, Y., Hotta, Y., Takahashi, R., Kato, M.V., Ishizaka, K., Sasaki, M.S., Horsthmeke, B., Minodo, K. and Sakai, T. Hypermethylation in retinoblastoma gene is associated with unilateral, sporadic retinoblastoma. *Cancer Genet. Cytogenet.,*

98, 43-49 (1997).

Orborne, R.J. and Hamshere, M.G. A genome-wide map showing common regions of loss of hetero zygosity/allelic imbalance in breast cancer. *Cancer Res.,* **60,** 3706-3712 (2000)

Pilotte, J., Larocque, D. and Richard, S. Nuclear translocation controlled by alternatively spliced isoforms inactivates the *QUAKING* apoptotic inducer. *Genes Dev.,* **15,** 845-858 (2001).

Prowse, A.H., Webster, A.R., Richards, F.M., Richard, S., Olschwang, S., Resche, F., Affara, N.A. and Maher, E. Somatic inactivation of the VHL gene in Von Hippel-Lindau disease tumors. *Am. J. Hum. Genet.,* **60,** 765-771 (1997).

Saccomanno, L., Loushin, C., Jan, E., Punkay, E., Artzt, K. and Goodwin, E.B. **The STAR protein QKI-6 is a translational repressor.** *Proc. Natl Acad. Sci. USA,* **96,** 12605-12610(1999).

Samorajski , T., Friede, R. L. and Reimer, P. R. Hypomyelination in the *quaking* mouse. A model for the analysis of disturbed myelin formation. *J. Neuropathol. Exp. Neurol.,* **29,** 507-523 (1970).

Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R. and Miller, W. PipMaker-a web server for aligning two genomic DNA sequences. *Genome Res.,* **10,** 577-586 (2000).

Sidman, R. L, Dickie, M. M. and Apple, S. H. Mutant mice (*quaking and jimpy*) with deficient myelination in the central nervous system. *Science,* **144,** 309-312 (1964).

Siomi, H., Matunis, M. J., Michael, W. M. and Dreyfuss, G. The pre-mRNA binding K protein contains a novel evolutionarily conserved motif. *Nucleic Acids Res.,* **21,** 1193-1198 (1993).

Siomi, H., Choi, M. C., Nussbaum, R. L. and Dreyfuss, G. The protein product of the fragile X gene, has characteristics of an RNA-binding protein. *Cell,* **74,** 291-298 (1993).

Siomi, H., Choi, M., Siomi, M. C., Nussbaum, R. L. and Dreyfuss, G. Essential role for KH domains in RNA-binding: impaired RNA binding by a mutation in the KH domain of FMR1 that causes fragile X syndrome. *Cell,* **77,** 33-39 (1994).

Srivastava, S., Tong, Y.A., Devadas, K., Zou, Z.Q., Sykes, V.W., Chen, Y., Blattner, W.A., Pirollo, K. and Chang, E.H. Deletion of both mutation and wild-type p53 protein in diverse tumors from a cancer-prone family with Li-Frau sydrome. *Oncogene,* **7,** 789-991 (1992).

Suter, U., Welcher, A.A. and Snipes G.J. Progress in the molecular understanding of hereditary peripheral neuropathies reveals new insights into the biology of the peripheral nervous system. *Trends. Neurosci.,* **16,** 50-57 (1993).

Takahashi, N., Roach, A., Teplow, D.B., Prusiner, S.B. and Hood, L. Cloning and characterization of the myelin basic protein gene from mouse: one gene can encode both 14kd and 18.5kd MBPs by the altenate use of exon. *Cell,* **42,** 139-148 (1985).

Taylor, S. J. and Shalloway, D. An RNA-binding protein associated with *Src* through its SH2 and SH3 domains in mitosis. *Nature,* **368,** 867-871 (1994).

Trent, J. M., Stanbridge, E. J., McBride, H. L., Meese, E. U., Casey, G., Araujo, D. E., Witkowski, C. M. and Nagle, R. B. Tumorigenicity in human melanoma cell lines controlled by introduction of human chromosome 6. *Science,* **247,** 568-571 (1990).

Tota, M., Iizuka, Y., Yu, W.J., Imai, T., Ikeda, E., Yoshida, K., Kawase, T., Kawakami, Y., Okano, H. and Uyemura, K. Expression of the neural RNA-

binding protein Musashi 1 in human gliomas. *Glia,* **34,** 1-7 (2001).

Veigl , M.L. *et al.* Biallelic inactivation of hMLH1 by epigenetic gene silencing, a novel mechanism causing hman MSI cancers. *Proc. Natl Acad. Sci. USA,* **95,** 8698-8702 (1998).

Vernet, C. and Artzt, K. STAR, a gene family involved in signal transduction and activation of RNA. *Trends Genet.,* **13,** 479-484 (1997).

Virchow, R. Ueber das ausgebreitete Vorkommen einer dem Nervenmark analogen substanz in den tierischen Geweben. *Virchows Arch Pathol Anat* **3,** 242-250 (1854).

Weinberg, R. A. Tumor suppressor genes. *Science,* **254,** 1138-1146 (1991).

Wong, G., Muller, O., Clark, R., Conroy, L., Moran, M. F., Polakis, P. and McCormick, F. Molecular cloning and nucleic acid binding properties of the GAP-associated tyrosine phosphoprotein p62. *Cell,* **69,** 551-558 (1992).

Zaffran, S., Astier, M., Gratecos, D. and Semeriva, M. 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 (1997).

Zhang, J. Y., Chan, E. K. L., Peng, X. X. and Tan, E. M. A novel cytoplasmic protein with RNA-binding motifs is an autoantigen in human hepatocellular carcinoma. *J. Exp. Med.,* **189,** 1101-1110 (1999).

Zhu, J., and Chen, X. MCG10, a novel p53 target gene that encodes a KH domain RNA-binding protein, is capable of inducing apoptosis and cell cycle arrest in G2-M. *Mol. Cell Biol.,* **20,** 5602-5618 (2000).

Zorn, A. M., Grow, M., Patterson, K. D., Ebersole, T. A., Chen. Q., Artzt, K. and Krieg, P. A. Remarkable sequence conservation of transcripts encoding amphibian and mammalian homologues of quaking, a KH domain RNA-binding

protein. *Gene,* **188,** 199-206 (1997).

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