



www.elsevier.com/locate/gene

# Genomic organization and chromosomal localization of the mouse IKBKAP gene<sup> $\ddagger$ </sup>

Gene 279 (2001) 81-89

Rocco Coli, Sylvia L. Anderson, Sabrina A. Volpi, Berish Y. Rubin\*

Department of Biological Sciences, Laboratory for Familial Dysautonomia Research, Fordham University, Bronx, NY 10458, USA

Received 11 June 2001; received in revised form 30 August 2001; accepted 24 September 2001 Received by A. Dugaiczyk

# Abstract

The autosomal recessive disorder familial dysautonomia (FD) has recently been demonstrated to be caused by mutations in the *IKBKAP* gene, so named because an initial report suggested that it encoded an I $\kappa$ B kinase complex associated protein (IKAP). Two mutations in *IKBKAP* have been reported to cause FD. The major mutation is a T  $\rightarrow$  C transition in the donor splice site of intron 20 and the minor mutation is a missense mutation in exon 19 that disrupts a consensus serine/threonine kinase phosphorylation site. We have characterized the cDNA sequences of the mouse, rat and rabbit *IKBKAP*-encoded mRNAs and determined the genomic organization and chromosomal location of mouse *IKBKAP*. There is significant homology in the amino acid sequence of IKAP across species and the serine/threonine kinase phosphorylation site altered in the minor FD mutation of IKAP is conserved. The mouse and human *IKBKAP* genes exhibit significant conservation of their genomic organization and the intron 20 donor splice site sequence, altered in the major FD mutation, is conserved in the human and mouse genes. Mouse *IKBKAP* is located on the central portion of chromosome 4 and maps to a region in which there is conserved linkage homology between the human and mouse genomes. The homologies observed in the human and mouse sequences should allow, through the process of homologous recombination, for the generation of mice that bear the *IKBKAP* mutations present in individuals with FD. The characterization of such mice should provide significant information regarding the pathophysiology of FD. © 2001 Elsevier Science B.V. All rights reserved.

Keywords: Familial dysautonomia; Gene structure; Exon/intron boundaries; Mapping panels

# 1. Introduction

The autosomal recessive disorder familial dysautonomia (FD) was recently demonstrated to be caused by mutations in the *IKBKAP* gene reported to encode the I $\kappa$ B kinase complex associated protein (IKAP) (Anderson et al., 2001; Slaugenhaupt et al., 2001). FD, also known as 'Riley–Day syndrome' or 'hereditary sensory neuropathy type III' (MIM 223900), affects the development and survi-

\* The nucleotide sequence data reported in this paper for the mouse, rat, and rabbit IKAP cDNAs have been submitted to GenBank and assigned Accession numbers AF387811, AF388201 and AF388202, respectively.

\* Corresponding author. Tel.: +1-718-817-3642; fax: +1-718-817-3828. *E-mail address:* rubin@fordham.edu (B.Y. Rubin). neurons (Riley et al., 1949; Axelrod et al., 1974). Individuals with FD are affected by a variety of symptoms, which include cardiovascular instability, decreased sensitivity to pain and temperature, recurrent pneumonias, an absence of overflow emotional tears, vomiting crises, and gastrointestinal dysfunction (Riley et al., 1949; Axelrod et al., 1974; Axelrod, 1996). This disorder is primarily confined to individuals of Ashkenazi Jewish descent (Brunt and McKusick, 1970) and the predicted carrier frequency of the defective gene is believed to be approximately one in 30 (Maayan et al., 1987). The major, or more common, FD-causing mutation of the IKAP-encoding gene is the result of a  $T \rightarrow C$ transition in the donor splice site of intron 20 that results in aberrant splicing, generating an RNA that lacks exon 20. The minor, or rarer, mutation of IKBKAP that causes FD is a missense mutation in exon 19 that disrupts a consensus serine/threonine kinase phosphorylation site.

val of sensory sympathetic and some parasympathetic

IKAP was originally identified as binding the  $I\kappa B$  kinases (IKKs) and the NF- $\kappa$ B-inducing kinase (NIK) and assembling these proteins into an active kinase complex (Cohen et

Abbreviations: aa, amino acid; bp, base pair(s); dNTP, deoxyribonucleoside triphosphate; DTT, dithiothreitol; EST, expressed sequence tag; FD, familial dysautonomia; HTGS, high throughput genome sequences; IKAP, I $\kappa$ B kinase associated protein; *IKBKAP*, gene encoding I $\kappa$ B kinase associated protein; nt, nucleotide(s); PCR, polymerase chain reaction; RACE, rapid amplification of cDNA ends; RT, reverse transcriptase; SSCP, singlestrand conformational polymorphism

al., 1998). More recent studies suggest that IKAP is not associated with IKKs and plays no specific role in NF- $\kappa$ B activation (Krappmann et al., 2000).

To facilitate an understanding of the IKAP-encoding gene, we have characterized the *IKBKAP*-encoded cDNAs of the mouse, rat and rabbit and determined the genomic organization and chromosomal location of the murine *IKBKAP*.

## 2. Materials and methods

# 2.1. EST and genomic database searches

The EST (expressed sequence tags), HTGS (high throughput genome sequences), and nr (non-redundant) databases at the National Center for Biotechnology Information were searched by BLAST software (Altschul et al., 1990).

# 2.2. RT-PCR amplification

cDNA was prepared from 1  $\mu$ g of DNase-treated RNA from the spleen of a 129/SvJ mouse, the cerebellum of a Sprague–Dawley rat and from the brains of New Zealand white rabbits at 42°C in a 20  $\mu$ l reaction containing 0.18 pM oligo dT primer, 500  $\mu$ M dNTPs, 10 mM DTT and 200 units of Superscript II RT (Life Technologies) according to the manufacturer's directions. PCR amplification of 50  $\mu$ l reactions containing 1.25 units of Taq polymerase (Life Technologies) and 10 pmol of primers was performed on these cDNAs using an initial denaturation step at 94°C for between 2 and 5 min followed by amplification for 45 cycles (30 s at 94°C, 30 s at 55°C, and 1–3 min at 72°C depending on the size of the product) and a final extension for 7 min at 72°C.

Amplification of the mouse cDNA was performed using the following primers whose design was based on the sequence of the human IKBKAP-encoded cDNA and mouse ESTs with homology to human IKBKAP-encoded cDNA sequence determined in this laboratory: MIKC-1forward 5'-TCCTTTCCAAACCCAGTGCG-3'; MIKC-1reverse 5'-CCAGTCACAGCATAGATACCG-3'; MIKC-2-forward 5'-ATCTGAAGCAAAGCCTGCC-3'; MIKC-2reverse 5'-AACCCCTTTCCACTTTCCG-3'; MIKC-3forward 5'-TGTGTCCTTGGTCTGACTG-3'; MIKC-3reverse 5'-ATGCTTCAAGTGCCTTCTCC-3'; MIKC-4forward 5'-TTATGGCGAGCACCTGATGC-3'; MIKC-4-5'-TGGACAAACGGTCTTTCC-3'; reverse MIKC-5forward 5-TAGCATCACAGCCTCTTACC-3'; MIKC-5reverse 5'-TATGTGGGTGCTGGGAAAC-3'.

Amplification of the rat cDNA was performed using the following primers that were designed based on the sequence of the human and mouse *IKBKAP*-encoded cDNA, rat ESTs with homology to human *IKBKAP* and newly derived rat cDNA sequence: RIKC-1-forward 5'-CACAGTTCCATG-GATCAGA-3'; RIKC-1-reverse 5'-CTGTGACTTCT-CAGCTAC-3'; RIKC-2-forward 5'-GTTTCTTTGGTGG-

CAGAAGG-3'; RIKC-2-reverse 5'-AAGGCAAACTCT-CGGTTCC-3'; RIKC-3-forward 5'-GCCAACAGAGTTC-ATCCACACC-3'; RIKC-3-reverse 5'-GACTATCCCA-CATCCCAGTTCTCC-3'; RIKC-4-forward 5'-GAAATA-CCTGCTGCTCCTG-3'; RIKC-4-reverse 5'-GGACAGG-TGTGGATGAACTC-3'; RIKC-5-forward 5'-TGTTTCT-CGTCTCCCGTGTG-3'; RIKC-5-reverse 5'-ATGACACA-GATACCACTGGC-3'.

Amplification of the rabbit cDNA was performed using the following primers that were designed based on the sequence of the human *IKBKAP*-encoded cDNA and newly derived rabbit cDNA sequence: RBIKC-1-forward 5'-GTTTCTTTGGTGGCAGAAGG-3'; RBIKC-1-reverse 5'-CTGTGACTTCTCAGCTAC-3'; RBIKC-2-forward 5'-TGTCACGAAGACCATGTACC-3'; RBIKC-2-forward 5'-TAGCACATTTGCTGAGGT-3'; RBIKC-3-forward 5'-CTATGACTTTGACTTGGTCCTC-3'; RBIKC-3-reverse 5'-CGAAGTCTTCTGTTGCTG-3'; RBIKC-4-forward 5'-GAATCACTTCATCATGCG-3'; RBIKC-4-reverse 5'-TGGTGTGTGTGCTGAGATTGC-3'.

## 2.3. Rapid amplification of cDNA ends

3'-rapid amplification of cDNA ends (RACE) was performed on the RNA samples described above using Superscript II RT (Life Technologies) with a 52 nt primer, termed  $Q_T$ , containing a 17 nt oligo-dT sequence at the 3' end, followed by a 35 nt sequence (5'-CCAGTGAGCA-GAGTGACGAGGACTCGAGCTCAAGCTTTTTTTTT-TTTTTTT-3') as described (Diffenbach and Kveksler, 1995). Following RNase H treatment, the cDNA synthesized by the RT was amplified using nested primers termed  $Q_0$  (5'-CCAGTGAGCAGAGTGACG-3') and  $Q_1$  (5'-GAGGACTCGAGCTCAAGC-3') encoded within the  $Q_T$ primer described above and the primers 5'-TGACT-CAGTGGTAAAGAGCG-3', 5'-ACTTGTGTCCCTGTT-CTCG-3' and 5'-CTCATAGCATCGCAGACA-3' located near the 3' end of the IKAP-encoding mRNA of the mouse, rat and rabbit, respectively.

5'-RACE was performed as described (Diffenbach and Kveksler, 1995). In brief, total RNA purified from the spleen of a 129/SvJ mouse was reverse-transcribed with Superscript II RT (Life Technologies) using the gene-specific primer 5'-TGGCAAGAAGCAGCAGTTC-3'. The product was poly (A) tailed by terminal deoxynucleotidyl transferase (Life Technologies). The above-described Q<sub>T</sub> primer was then annealed to the poly (A) tailed product and extended, followed by amplification with a gene-specific primer (5'-CCAGCAAGTCCTGAATACC-3') located 388 bases from the purported 5' end of the IKAP cDNA and the Q<sub>0</sub> primer. The DNA product generated was then subjected to nested PCR amplification using a gene-specific primer (5'-AACCTTCTGCCACCAAAG-3') located 337 bases from the purported 5' end of the IKAP cDNA and the Q<sub>I</sub> primer.

# 2.4. DNA sequencing

Nucleotide sequences were determined by the dideoxy chain termination method using the Amplicycle Sequencing Kit (Applied Biosystems).

# 2.5. Determination of genomic organization of the mouse *IKBKAP* gene

The serial primers designed, which correspond to regions spanning the entire IKAP-encoding cDNA, are presented in Table 1.

The PCR conditions were as follows: an initial denaturation step of 5 min at 94°C followed by amplification for 40 cycles (30 s at 94°C, 30 s at between 55 and 61°C, and 30 s to 5 min at 72°C) and a final extension for 7 min at 72°C. For the PCR template, genomic DNA was prepared from the liver of a 129/SvJ mouse using the Qiagen Blood & Cell Culture DNA Mini Kit. Comparison of the cDNA and genomic DNA sequences revealed the exon/intron organization of the gene.

# 2.6. Chromosomal mapping using backcross DNA panels

Linkage analysis was performed with the BSB and BSS backcross DNA panels (Jackson Laboratory). The BSB panel consists of 94 genotyped progeny of (C57BL/ $6J \times M.$  spretus) F1 × C57BL/6J and the BSS panel consists of 94 genotyped progeny of (C57BL/ $6Jei \times$  SPRET/Ei) F1 × SPRET/Ei (Rowe et al., 1994). Allele detection was performed using sequence polymorphisms that were amplified by PCR in the presence of [ $\alpha$ -<sup>33</sup>P]dATP and characterized for single-strand conformational polymorphisms (SSCP). PCR amplification was performed under the following conditions: an initial denaturation step of 2 min at 94°C followed by amplification for 40 cycles (30 s at 94°C, 30 s at 58°C, and 30 s at 72°C) and a final extension

Table 1

Oligonucleotides used for PCR to determine genomic organization of the mouse IKBKAP gene<sup>a</sup>

| Name | Sequence of oligonucleotide    | Intron | Sequence of oligonucleotide  | Name |
|------|--------------------------------|--------|------------------------------|------|
| M1F  | 5'-gtgtttctcgtctcccgt-3'       | 1      | 5'-atctacttccgtcagtccac-3'   | M2R  |
| M2F  | 5'-CAGTGTTTCTGTCTGCGAG-3'      | 2      | 5'-CCAGCAAGTCCTGAATACC-3'    | M3R  |
| M3F  | 5'-AGGAATCTGTGTGTGTGGC-3'      | 3      | 5'-ATGACAGAGATACCACTGGC-3'   | M4R  |
| M4F  | 5′-TGGAATGTGTTGGGAGTG-3′       | 4      | 5'-CAAAATCATCCTGGTGGA-3'     | M5R  |
| M5F  | 5'-CCACCAGGATGATTTTGG-3'       | 5      | 5'-TGGAACTGGGTCTGCTTAC-3'    | M6R  |
| M6F  | 5'-gcaagtttgtcactgttggatgg-3'  | 6      | 5'-aggtaatgtgtggtctgcggtc-3' | M7R  |
| M7F  | 5'-accgcagaccacacattacc-3'     | 7      | 5'-TGGCACAGACTCACTGGTTG-3'   | M8R  |
| M8F  | 5′-atggaaccgagagtttgcc-3′      | 8      | 5'-gctggttgggtttatcttga-3'   | M9R  |
| M9F  | 5'-caagataaacccaaccagc-3'      | 9      | 5'-gcttttcagagtggaactgtc-3'  | M10R |
| M10F | 5'-atgacctgctgtggaatgc-3'      | 10     | 5'-ccagtcacagcatagataccg-3'  | M11R |
| M11F | 5'-ATCTGAAGCAAAGCCTGCC-3'      | 11     | 5'-aaggtcattcccaaggtgag-3'   | M12R |
| M12F | 5'-gacgccagtaaccagatttc-3'     | 12     | 5′-ccaaatgaggtgttgtaagagg-3′ | M13R |
| M13F | 5'-cctcttacaacacctcatttgg-3'   | 13     | 5'-agtgtcgtcttcaacccagg-3'   | M14R |
| M14F | 5'-TTTCTCACCTGGGTTGAAG-3'      | 14     | 5'-tgacttggtcttagaacagcag-3' | M15R |
| M15F | 5'-gtcgtcattggtttatgctgc-3'    | 15     | 5'-tggctacttccatctgtgtgc-3'  | M16R |
| M16F | 5'-AGTCACCTTCTCTGGCTGTG-3'     | 16     | 5'-CACCTGTCAGTCAGACCAAG-3'   | M17R |
| M17F | 5'-TCCTTGGTCTGACTGACAGG-3'     | 17     | 5'-tgtgggaatgggttgttac-3'    | M18R |
| M18F | 5'-gctgtgtgtgtgctgactttctac-3' | 18     | 5'-AACCCCTTTCCACTTTCCG-3'    | M19R |
| M19F | 5′-aagtggaaaggggttcacg-3′      | 19     | 5'-TCCGAATCTGTGCCAAGAC-3'    | M20R |
| M20F | 5'-tggtcttggcacagattc-3'       | 20     | 5'-ggttaattctcagcttcctc-3'   | M21R |
| M21F | 5'-gaatgcatgaggaagctgag-3'     | 21     | 5'-CTGTTTTACGAAGGTTTCCAC-3'  | M22R |
| M22F | 5′-gtggaaaccttcgtaaaacag-3′    | 22     | 5'-tggacacctggacactcttg-3'   | M23R |
| M23F | 5'-TTGACCTCATCTGTGACGC-3'      | 23     | 5'-tgagatgtgagtattgacaggc-3' | M24R |
| M24F | 5'-gcctgtcaatactcacatc-3'      | 24     | 5'-tctacactcacactctcagg-3'   | M25R |
| M25F | 5′-AGAGTGTGAGTGTAGAGGAGG-3′    | 25     | 5'-CACTTGCTGAGGTGTCCAA-3'    | M26R |
| M26F | 5'-TTACCAGAGGTTCACCATAGAC-3'   | 26     | 5'-TACTGTGGTGAGTCTGGACG-3'   | M27R |
| M27F | 5′-ggacctgagtatttcacagaatgc-3′ | 27     | 5'-ATGCTTCAAGTGCCTTCTCC-3'   | M28R |
| M28F | 5'-TTATGGCGAGCACCTGATG-3'      | 28     | 5'-TACTGTTCCAGGACTGTGGC-3'   | M29R |
| M29F | 5'-agcagaggaagcacagtgag-3'     | 29     | 5'-gagcacagcctcttcataatc-3'  | M30R |
| M30F | 5'-TATGAAGAGGCTGTGCTCCTGC-3'   | 30     | 5'-TGGAAGGCTTTATGCTGGTTTC-3' | M31R |
| M31F | 5'-aaccagcataaagccttcc-3'      | 31     | 5′-ATGGCGAATGAATGTGGC-3′     | M32R |
| M32F | 5'-gacagccacattcattcgc-3'      | 32     | 5'-gtttccgagaagaggtctgac-3'  | M33R |
| M33F | 5'-TGAGTGGCAGTGAGATGAG-3'      | 33     | 5'-CTTTGAGGCTATGCTTCTTGCG-3' | M34R |
| M34F | 5'-aggtcatctaaaaaccgtcg-3'     | 34     | 5'-gtgtggttgaactctgttgg-3'   | M35R |
| M35F | 5'-ggtgcgtgctattttgaagg-3'     | 35     | 5'-TTCTGCTGCTGGTAAGAGGC-3'   | M36R |
| M36F | 5'-accagcagcagaagacttg-3'      | 36     | 5'-TGGACAAACGGTCTTTCC-3'     | M37R |

<sup>a</sup> Names ending in F refer to the forward direction and R to the reverse direction.

| mouse<br>human | 210 220 240 240 240 240 240 240 240 240 24   |
|----------------|--|
| mouse<br>human | 2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2510<br>2500<br>2500<br>2500<br>2500<br>2500<br>2500<br>2500<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600<br>2600 |
| mouse<br>human | 2600 2000 2000 2000 2000 2000 2000 2000  |
| mouse<br>human | 2770 2780 2780 2780 2780 2800 2800 2800  |
| mouse<br>human | 2600 2700 2700 2700 2700 2700 2700 2700  |
| mouse<br>human | 50/0 500 500 500 500 500 500 500 500 500   |
| mouse<br>human | 11/10 11/20  |
| mouse<br>human | 3339         360         370         380   |
| mouse<br>human | 1370 1380 1390 1400 1400 1400 1400 1400 1400 1400 14   |
| mouse<br>human | 400 1330 1330 1330 1330 1330 1330 1330 1   |
| mouse<br>human | MID MOD 100 100 100 100 100 100 100 100 100 10   |
| mouse<br>human | 1710         |
| mouse<br>human | M30         M40  |
| mouse<br>human | 3070       3000       4000   |
| mouse<br>human | 1000         1100         1120         1120         1140         1150         1160         1170         1180 <th< th=""></th<>  |
| mouse<br>human | 400 400 400 400 400 400 400 400 400 400  |
| mouse<br>human | 400 400 400 400 400 400 400 400 400 400  |
| mouse<br>human | 440 460 460 460 460 460 460 460 460 460  |
| mouse<br>human | 400 400 400 400 400 400 400 400 400 400  |

Fig. 1. Alignment of the nucleotide sequences of the mouse and human IKAP cDNAs. The mouse sequence is aligned with the comparable sequence of the human IKAP cDNA (nucleotides 142–4804 of Accession number NM\_003640).

| uIKAP<br>tuIKAP<br>tIKAP<br>bIKAP  | / MRNLKLFRTLEFRDIQUP GNPQCFSLKITEQGTVLIGSENGLIEVDFVSKEV, KNEVSLVALEGFLFEDGSGCIVGTQDLLDQESVCVATASGDVIVCNLSTQLECVGSVASGISVMSWSPDI<br>/ MRNLKLHRTLEFRDIQAPGKPQCFCLKAEQGTVLIGSENGLTEVDFVKREVKTEISLVAEGFLFEDGSGCIVGTQDLLDQESVCVATASGDVIVCNVSTQLECVGNVASGISVMSWSPDI<br>/ MRNLKLHQTLEFRDIQAPGKPQCFCLKAEQGTVLIGSENGLTEVDFVKREVKTEISLVAEGFLFEDGSGCIVGTQDLLDQESVCVATASGDVIVCNVSTQLECVGNVASGISVMSWSPDI<br>/ MRNLKLLQTLEFKDIQAPGKPQCFCLKAEQGTVLIGSENGLTEVDFVKREVKTEISLVAEGFLFEDGSGCIVGTQDLLDQESVCVATASGDVIVCNVSTQLECVGNVASGISVMSWSPDI<br>/ MRNLKLLQTLEFKDIQAPGKPQCFSLRTEFGTVLIGSENGLTEVDFVKREVKTEISLVAEGFLFEDGSGCIVGTQDLLDQESVCVATASGDVIVCNVSTQLECVGNVASGISVMSWSPDI<br>/ MRNLKLLQTLEFKDIQAPGKPQCFSLRTEFGTVLIGSENGLTEVDFVKREVKTEISLVAEGFLFEDGSGCIVGTQDLLDQESVCVATASGDVIVCNVSTQLECVGNVASGISVMSWSPDI   | 20<br>20<br>20           |
|------------------------------------|---|--------------------------|
| ulKAP<br>ulKAP<br>tlKAP<br>blKAP   | 121 Q E L V L L A T G Q Q T L I MM T K D F E P I L E Q Q I H Q D D F G E S K F I T Y G W G R K E T Q F H G S E G R Q A A F Q M Q M H E S A L P W D D H R P Q V T W R G D G Q F F A V S V V C P E T G A R K V R Y W N R E F A L Q S T S E P Y A G L 2<br>121 Q E L L L L A T A Q Q T L I MM T K D F E V I A E Q I H Q D D F G E G K F V T V G W G S K Q T Q F H G S E G R P T A F P V Q L P E N A L P W D D R R PH I T W R G D G Q F F A V S V V C R Q T E A R K I R Y W N R E F A L Q S T S E S V P G L 2<br>121 Q E L L L L A T A Q Q T L I MM T K D F E V I A E Q I H Q D D F G E G K F V T V G W G S K Q T Q F H G S E G R P T A F P V Q L P E N A L P W D D R R PH I T W R G D G Q F F A V S V V C R Q T E A R K I R Y W N R E F A L Q S T S E S V P G L 2<br>121 Q E L L L L A T A Q Q T L I MM T K D F E V I A E Q I H Q D D F G E G K F T T V G W G S K D T Q F H G S E G R P I T F P V Q M H E S A L F W D D H R P Q I T W R G D G Q F F A V S V V C S Q T G A R K I R Y W N R E F A L Q S T S E S V P G L 2<br>121 Q E L V L L A T A Q Q T L I MM T K D F E P I M E Q Q I H Q D D F G E G K F T T V G W G K K E T Q F H G S E G R Q A A F Q I Q T H E S A L F W D D H R P Q I T W R G D G Q F F A V S V V C S Q T G A R K I R Y W N R E F A L Q S T S E S V P G L 2<br>121 Q E L V L L A T G Q Q T L I MM T K D F E P I M E Q Q I H Q D D F G E S K F I T V G W G K K E T Q F H G S E G R Q A A F Q I Q T H E S A L F W D D H R P R V T W R G D G Q F F A V S V V C S Q T G A R K I R Y W N R E F A L Q S T S E P V F G L 2<br>121 Q E L V L L A T G Q Q T L I M M T K D F E P I M E Q Q I H Q D D F G E S K F I T V G W G K K E T Q F H G S E G R Q A A F Q I Q T H E S A L F W D D H R P R V T W R G D G Q F F A V S V V C P E T G A R K V R W N R E F A L Q S T S E P V F C L 2<br>121 Q E L V L L A T G Q Q T L I M M T K D F E P I M Q D D F G E S K F I T V G W G K K E T Q F H G S E G R Q A A F Q I Q T H E S A L F W D D H R P R V T W R G D G Q F F A V S V V C P E T G A R K V R W N R E F A L Q S T S E P V F C L 2<br>122 Q E L V L L A T G Q Q T L I M M T K D F E V I M Q D F G E S K F I T V G W G K K E F A L Q S T | 40<br>40<br>40           |
| uIKAP<br>tuIKAP<br>tIKAP<br>bIKAP  | M GPALAWKPSGSLIASTQDKPNQQDIVFFEKNGLLHGHFTLPFLKDEVKVNDLLWNADSSVLAVRLEDLQREK SSI 1 PKTCVQLWTVGNYHWYLKQSLSFSTCGKSKLUSLMWDPVTPYRLH<br>GPALAWKPSGSLIASTQDKPNQQDVVFFEKNGLLHGHFTLPFLKDEVKVNDLLWNADSSVLAVRLEDLPKEDSSTLKSYVQLWTVGNYHWYLKQSLPFSTTGKNQVVSLLWDPVTPYRLH<br>GPALAWKPSGSLIASTQDKPNQQDVVFFEKNGLLHGYFTLPFLKDEVKVNDLLWNADSSVLAVWLEDLPKEDSSTLKSYVQLWTVGNYHWYLKQSLPFSTTGKNQVVSLLWDPVTPYRLH<br>GPALAWKPSGSLIASTQNKPNQQDVVFFEKNGLLHGYFTLPFLKDEVKVNDLLWNADSSVLAVWLEDLPKEDSVLWTVGNYHWYLKQSLPFSTTGKNQVSLLWDPVTPYRLH<br>GPALAWKPSGSLIASTQNKPNQQDVVFFEKNGLLHGYFTLPFLKDEVKVNDLLWNADSSVLAVWLEDLQREEDSVLKVVQLWTVGNYHWYLNQLYFSTYGKSKIVSLMWDPVTPYRLH<br>GPALAWKPSGSLIASTQNKPNQQDVVFFEKNGLLHGYFTLPFLKDEVKVNDLLWNADSSVLAVWLEDLQREEDSVLKVVQLWTVGNYHWYLNGLLFSTYGKSKIVSLMWDPVTPYRLH  | 50<br>50<br>60<br>60     |
| udKAP<br>uulKAP<br>tlKAP<br>blKAP  | 5/1       V L C Q G WH Y LA Y D WH WT T D R S V G D N S S D L S N Y A Y I D G N R Y L Y T Y F R Q T Y Y P P P M C T Y Q L L F P H P Y N Q Y T F L A H P Q K S N D L A Y L D A S N Q I S Y Y K C G D C P S A D P T Y K L G A Y G G S G F K V C L R T P 4         5/1       Y L C Q G WH Y L A Y D WH WT T D R S S G N S A N D L A N Y A Y I D G N R Y L Y T Y F R Q T Y Y P P P M C T Y R L I I P H P Y N Q Y I P S A H L G - N D L A Y L D A S N Q I S Y Y K C G D K P N M D S T Y K L G A Y G G N G F K Y P L M T P P M C T Y R L I I P H P Y N Q Y I P S A H L G - N D L A Y L D A S N Q I S Y Y K C D D K P N M D S T Y K L G A Y G G N G F K Y P L M T P P M C T Y R L I I P H P Y N Q Y I B S N A H L G - N D L A Y L D A S N Q I S Y Y K C D D K P D M D S T Y K L G A Y G G N G F K Y P L M T P P M C T Y R L I I P H P Y N Q Y I B S A H L G - N D L A Y L D A S N Q I S Y Y K C D D K P D M D S T Y K L G A Y G G I G F K Y P L M T P P M C T Y R L I I P H P Y N Q Y I B S A H L G - N D L A Y L D A S N Q I S Y Y K C D K P D M D S T Y K L G A Y G G I G F K Y P L R T P 4         5/1       Y L C Q G WH Y L C Y D W R W T T D R S S G D N E S D L A N Y A Y I D G N R T L Y T Y P Q T T P P P M C T Y R L I I P H P Y N Q Y T E C A L P K K S N D L A Y L D A S N Q I S Y Y K C G D S P S M D P T Y K L G A Y G G N G F K Y S L R T P 4         5/1       Y L C Q G WH Y L C Y D W R W T T D R S G D N E S D L A N Y A Y I D G N R T L Y T Y P Q T Y P P P M C T Y R L L L P H P Y N Q Y T E C A L P K K S N D L A Y L D A S N Q I S Y Y K C G D S P S M D P T Y K L G A Y G G N G F K Y S L R T P 4  | 80<br>78<br>78<br>80     |
| ulKAP<br>uulKAP<br>tlKAP<br>blKAP  | 40 H LEKRYK I QFENN ED Q D VN PLK LG L LTWLEE D V FLAVSHSEFSPR - SVIHHLTA ASSEMD BEHG Q LN VSSSA AVDGVI I ISLCCN SKTKSVVL Q LAD G Q I FKYLWESPSLAIK PWKNS 3<br>40 H LEKRYSI QFG N - E EE EE EVALQUSFLATS (SHV) ELATS (SHV) ELATS (SHV) ELATS (SHV) ELATS (SHV) E E G Q Q LD VSSSV TVD G VVIG LCCC SKTKSLAV Q LAD G Q VLKYLWESPSLAV E PWKNS 3<br>40 H LEKRYSI QFG N E EE EE VIL Q LSFLTW (SHV) E LATS (SHV) E HHLTM ASSEMD E E G Q Q LD VSSSV TVD G VVIG LCCC SKTKSLAV Q LAD G Q VLKYLWESPSLAV E PWKNS 3<br>41 H LEKRYSKI QFE SN ED Q E TNPLKESLUSWIE ED I FLATS (SHV) I HHLTM ASSEMD E E G Q Q LSVSSSI SV VD G VVIG LCCC SKTKSVAL Q LAD G Q I LKYLWESPSLAV E PWKNS 3<br>41 H LEKRYSKI QFE SN ED Q E TNPLKESLUSWIE ED I FLATS (SHV) I HHLTM ASSEMD E E Q G Q LN VSSSV VD G VVIG LCCC SKTKSVAL Q LAD G Q I LKYLWESPSLAV E PWKNS 3<br>41 H LEKRYSKI QFE SN ED Q E TNPLKESLUSWIE ED I FLATS (SHV) I HHLTM ASSEMD E E Q G Q LN VSSSV VD G VVIG LCCC SKTKSVAL Q LAD G Q I LKYLWESPSLAV E PWKNS 3<br>41 HLEKRYSKI QFE SN ED Q E TNPLKESLUSWIE ED I FLATS (SHV) I HHLTM ASSEMD E E Q G Q LN VSSSV VD G VVIG LCCC SKTKSVAL Q LAD G Q I LKYLWESPSLAV E PWKNS 3<br>41 HLEKRYSKI I QFE SN ED Q E TNPLKESLUSWIE ED I FLATS (SHV) I HKLTVVPCEVD E E Q G Q LN VSSSV VD G VVIG LCCC SKTKSVAL Q LAD G Q I LKYLWESPSLAV E PWKNS 3<br>41 HLEKRYSKI I QFE SN ED Q E TNPLKESLO S I SV I I HKLTVVPCEVD E E VD E E Q G Q LN VSSSV I SVDGT I I SNCCN SKTKSVAL Q LAD G Q I LKYLWESPSLAV E PWKNS 3<br>41 HLEKRYSKI I QFE SN ED Q E TNPLKESLO S I SVDGT I I SNCCN SKTKSVAL Q LAD G Q I LKYLWESPSLAV E PWKNS 3<br>41 HLEKRYSKI I QFE SN ED Q E TNPLKESVAL I V I I SNCCN SKTKSVAL I SNCCN SKTKSVAL I V I I SNCCN SKTKSVAL I V I I SNCCN SKTKSVAL I SNCCN SKTKSVAL I SNCCN SKTKSVAL I V I I SNCCN SKTKSVAL I SNCCN SKTKSVAL I SNCCN SKTKSVAL I SNCCN S                                  | 96<br>97<br>93<br>97     |
| ndKAP<br>ndKAP<br>tlKAP<br>bIKAP   | 57 GGF PVRFP YPCTQTELAMIGEBECVLGLTDRCRFFINDIBVASNITSFAYYDEFLLLTTHSHTCQCFCLRDASFKTLQAGLSSNHVSHGEVLRKVERGSRIVTVVPQDTKLUQMPRGNL<br>58 EGIPVRFVHPCTQMEVATIGGEECVLGLTDRCRFFINDTEVASNITSFAYYDEFLLLTTHSHTCQCFCLRDASFKTLQAGLSGSHEASGETLRKVERGSRIVTVVPQDTKLILQMPRGNL<br>58 EGRPVRFARPCTQMEAAAIGGEECVLGLTDRCRFFINDTEVASNITSFAYYDEFLLLTTHSHTCQCFCLRDASIKMLQAGLCSSQMPSGELLRKVERGSRIVTVVPQDTKLILQMPRGNL<br>58 EGRPVRFARFCTQMEVAAAIGGEECVLGLTDRCRFFINDTEVASNITSFAYYDEFLLLTTHSHTCQCFCLRDASIKMLQAGLCSSQMPSGELLRKVERGSRIVTVVPQDTKLILQMPRGNL<br>58 EGRPVRFARFCTQMEVAAAIGGEECVLGLTDRCRFFINDTEVASNITSFAYYDEFLLLTTHSHTCQCYCLKDASIKTLQAGLSSSHVSNGEILRKVERGSRIVTVVPQDTKLILQMPRGNL<br>58 EGRPVRFARFVCTQMEVASNITSFAYYDEFLCAMIGGEECVLGLTDRCRFFINDTEVASNITSFAYYDEFLLLTTHSHTCQCYCLKDASIKTLQAGLSSSHVSNGEILRKVERGSRIVTVVPQDTKLILQMPRGNL   | 16<br>17<br>15<br>17     |
| wIKAP<br>wuIKAP<br>tIKAP<br>tIKAP  | 7/7       EVVHHRALVLAQIRKWLDKLMFKEAFECMRKLRINLNPTYDHNPKVFLG NVETFIKQIDSVNHINLFFTELKEEDVTKTMYPAPVTS SVYLSRDPDGNKIDLDLTCDAMRAVMESINPHKYC &         7/7       EVVHHRALVLAQIRKWLDKLMFKEAFECMRKLRINLNLIHDHNPKVFLG NVETFIKQIDSVNHINLFFTELKEEDVTKTMYPPPTTKS SVYLSRDPDGKKLDLICDAMRAVMESINPHKYC &         7/7       EVVHHRALVLAQIRKWLDKLMFKEAFECMRKLRINLNLIHDHNPKVFLG NVETFIKQIDSVNHINLFFTELKEEDVTKTMYPPTTKS SVYLSRDVG STHPDGKKLDLICDAMRAVMESINPHKYC &         7/7       EVVHHRALVLAQIRKWLDKLMFKEAFECMRKLRINLNLIHDHNPKVFLE NVETFTKQ (DSVNHINLFFTELKEEDVTKTMYPPTYKS VOUS TNPDGKKLDLICDAMRAVMESINPKKFC &         7/8       EVVHHRALVLAQIRKWLDKLMFKEAFECMRKLRINLNLIHDHNPKVFLG NVETFIKQIDSVNHINLFFTELKEEDVTKTMYPPTYFS SVOQSRDPGGTKLDLICDALWMETINLNHHKYC &  | 36<br>37<br>35<br>37     |
| nuIKAP<br>nuIKAP<br>tIKAP<br>tIKAP | 837 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS DPD A VSAEEALKYLLH LVDVN EL YDHSLGTYDFD LVLMVAEKSQKDPKEYLPFLNTLKKME TNYQRFTIDKYLKRYEKAIGHLSKCGPE 5 838 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS DPD A VSAEEALKYLLLUVDVN EL FNHSLGTYDFD LVLMVAEKSQKDPKEYLPFLNTLKKME TNYQRFTIDKYLKRYEKAIGHLSKCGPE 5 848 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS DPD A VSAEEALKYLLLUVDVN EL FNHSLGTYDFD LVLMVAEKSQKDPKEYLPFLNTLKKME TNYQRFTIDKYLKRYEKAIGHLSKCGPE 5 849 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS DPD A VSAEEALKYLLLUVDVN EL FNHSLGTYDFD LVLMVAEKSQKDPKEYLPFLNTLKKME TNYQRFTIDKYLKRYEKAIGHLSKCGPE 5 840 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS 940 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS 951 DKYLKRYEKAIGHLSKCGPE 5 953 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS 954 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS 954 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS 954 L\$ I LY LYNA A L\$ S QKDPKE Y LPFLNTLKKME TNYQRFTIDKYLKRYEKA I GHLSKCGPE 5 955 L\$ I LTS HVKKTTP E LE IV LQKVHE LQGNA PS   | 36<br>57<br>53<br>57     |
| nuIKAP<br>nuIKAP<br>tIKAP<br>tIKAP | 977 Y F P E C L N L I K D K N L Y N E A L K L Y S P S S Q Q Y Q D I S I A Y G E H L M Q E H M Y E P A G L M F A R C G A H E K A L S A F L T C G N W K Q A L C V A A Q L N F T K D Q L V G L G R T L A G K L Y B Q R K H 1 D A A M Y L E E C A Q D Y I<br>978 Y F T E C L N L I K D K N L Y K E A L K L Y R PD S P Q Y Q A V S M A Y G E H L M Q E H L Y E P A G L Y F A R C G A U E A F L A C G S W Q Q A L C V A A Q L N F T K D Q L V G L G R T L A G K L Y B Q R K H S E A A T Y L E Q Y A Q D Y I<br>978 Y F T E C L N L I K D K N L Y K E A L K L Y R PD S P Q Y Q A V S M A Y G E H L M Q E H L Y E P A G L Y F A R C G A U E K A L E A F L A C G S W Q Q A L C V A A Q L Q M S K M K D K V A G L A R T L A G K L Y B Q R K H S E A A T Y L E Q Y A Q D Y I<br>978 Y F T E C L N L I K D K N L Y K E A L K L Y R PD S P Q Y Q A V S W A Y G E H L M U P A G L Y F A R C G A H E K A L E A F L A C G S W Q Q A L C V A A Q L Q M S K M K D K V A G L A R T L A G K L Y B Q R K H S E A A T Y L E Q Y A Q D Y I<br>978 Y F S E C L N L I K D K N L Y N E A L K L Y P P T S Q E Y K D I S I A Y G E H L M E E H Q Y E P A G L Y F A R C G A H E K A L E S A F L T C G S W Q Q T L C M A A Q L N M T E E Q L A G L A R T L A G K L Y B Q R K H S E A A T Y L E Q Y A L D Y I<br>978 Y F S E C L N L I K D K N L Y N E A L K L Y P P T S Q E Y K D I S I A Y G E H L M E E H Q Y E P A G L Y F A R C G A H E K A L E S A F L T C G S W Q Q T L C M A A Q L N M T E E Q L A G L G R T L A G K L A E Q R K H S D A A I Y L E Q Y T Q D Y I<br>978 Y F S E C L N L I K D K N L Y N E A L K L Y P D T S Q E Y K D I S I A Y G E H L M E E H Q Y E P A G L Y F A R C G A H E K A L E S A L L A T L L Q Y L D Y M T E E Q L A G L G R T L A G K L A A L Y L E Q Y T Q D Y I D Y I  | 076<br>077<br>075<br>077 |
| ndKAP<br>nulKAP<br>tlKAP<br>blKAP  | 1077 E E A V L L L L L L L L L L L L L L V V K YN R L D I I E T N V K P S I L E A Q KN YM A F L D S Q T A T F S R H K K R L L V Y R B L K E Q A Q Q A G L D D E V P H G Q E S D L F S E T S S V V S G S E M S G K Y S H S N S R I S A R S S K N /<br>1078 E E A V L L L L L E G S A W E B A L R L V Y K YD R V D I I E T S I K P S I L E A Q KN YM D F L D S Q T A T F S R H K K R L L V Y R B L K E Q A Q Q A G L D D E V P H G Q E S D L F S E T S S T M S G S E M S G K Y S H S N S R I S A R S S K N /<br>1076 E E A V L L L L E G S A W E B A L R L V Y K YD R V D I I E T S I K P S I L E A Q KN YM D F L D S Q T A T F I R H K N R L K V Y R B K N E K S Q K P Y H D H E V A H G P E I S D L F S E T S S T M S G S E M S G R Y S H S N S R I S A R S S K N /<br>1076 E E A V L L L L E G S A W E B A L R L V Y K YD R V D I I E T S I K P S I L E A Q KN YM D F L D S Q T A T F I R H K N R L K V Y R B K N G R Y H S N S R I S A R S S K N /<br>1076 E E A V L L L L E G S A W E B A L R L V Y K YN R L D I I B T N I K P S I L E A Q KN YM A F L E S Q S A T F S R H K E R L L E Y R E L K E R A Q Q Y D L D D E M P H G Q E A D L F S E T S S I V S G S E M S G R Y S H S N S R I S A R S S K N /<br>1078 E E A V L L L L E G S A W E B A L R L V Y K YN R L D I I B T N I K P S I L E A Q KN YM A F L E S Q S A T F S R H K E R L L E Y R E L K E R A Q Q Y D L D D E M P H G Q E A D L F S E T S S I V S G S E M S G R Y S H S N S R I S A R S S K N /<br>1078 E E A V L L L L E G S A W E B A L R L V Y K YN R L D I I B T N I K P S I L E A Y K N YM A F L E S Q S A T F S R H K E R L L E Y R E L K E R A Q Q Y D L L D D E M P H G Q E A D L F S E T S S I V S G S E M S S K Y S H S N S R I S A R S S K N /  | 196<br>197<br>195<br>197 |
| ndKAP<br>nulKAP<br>tlKAP<br>tlKAP  | 1197 RRKAERKKHSLKEGSPLED LALLEALSEVVQN TENLKDEVYH ILKVLFLFEFD EQGRELQKAFED TLQLMERSLPE IWTLTYQQN SA TPVLGPN STANSIMASYQQQKTS VPVLDAELFI<br>1188 RRKAERKKHSLKEGSPLEG LALLEALSEVVQSVEKLKDEVRAILKVLFLFEFE Ø QAKELQRAFESTLQLMERAVPEIWTLTGQQSSSTPVLGPSSTANSIMASYQQQKTSVPVLDAELFI<br>1188 RRKAERKKHSLKEGSPLEG LALLEALSEVVQSI EKLKDEVRAILKVLFLFEFE Ø QAKELQRAFESTLQLMERAVPEIWTLTGQQSSSTPVLGPSSTANSIMASYQQQKTSVPVLDAELFI<br>1188 RRKAERKKHSLKEGSPLEG LALLEALSEVVQSI EKLKDEVRAILKVLFLFEFE Ø QAKELQRAFESTLQLMERSVQQN SA MPVLGPSSTANSIMASYQQQKTSVPVLDAELFI<br>1188 RRKAERKKHSLKEGSPLEG LALLEALSEVVQSI EKLKDEVRAILKVLFLFEFE Ø QAKELQRAFESTLQLMERSVQQN SA MPVLGPSSTANSIMASYQQQKTSVPVLDAELFI<br>1188 RRKAERKKHSLKEGSPLEG LALLEALSEVVQSI DKLKDEVRAILKVLFLFEFE Ø QAKELQRAFESTLQLVERSLFØR STEVITING<br>1199 RRKAERKKHSLKEGSPLEG LALLEALSEVVQSI DKLKDEVRAILKVLFLFEFED Ø GRELQK FFQ DTLQLVERSLFØR SLFEIWTLTYQQN SA MPVLGPSSTANSIMASYQQQKTSVPVLDAELFV   | 316<br>317<br>315<br>317 |
| ndKAP<br>nulKAP                    | 1377 P F K IN R R T O WK L S L L D 1332<br>1376 P F K M D P R S O WK L S L L E 1333<br>1376 P F K M D D D D S S S WK T S B L I I 1333   |                          |

 dKAP
 1316
 P P K I
 D Q R S
 Q W K L S L L E
 1331

 biKAP
 1318
 P P K I
 N R K T
 Q W K L S L L E
 1333

Fig. 2. Comparison of the amino acid sequences of human (huIKAP), mouse (muIKAP), rat (rtIKAP) and rabbit (rbIKAP) IKAP. A potential phosphorylation site conserved in all of these species is underlined.

for 7 min at 72°C. Each reaction was performed on 50 ng of genomic DNA using Taq polymerase (Life Technologies) and the following primers: 5'-TTGACCTCATCTGT-GACGC-3' and 5'-TGAGATGTGAGTATTGACAGGC-3', located in exons 23 and 24 of the mouse *IKBKAP* gene, respectively. For the SSCP analysis, the amplified products were denatured, fractionated on a nondenaturing 5% acrylamide gel at 4°C and detected by autoradiography.

### 3. Results and discussion

#### 3.1. cDNA amplification and characterization

cDNA generated from mRNA isolated from the spleen of a 129/SvJ mouse was subjected to PCR amplification using primers whose nucleotide sequences matched either that of the human IKAP-encoding mRNA or related mouse ESTs. Sequencing of these products and the sequencing of 3'- and 5'-RACE products generated from the mouse IKAP RNA revealed that the full-length IKAP-encoding mRNA is 5034 nucleotides in length and, at the nucleotide level, exhibits 77% identity with the human IKAP-encoding mRNA (Fig. 1). PCR amplification of the IKAP cDNA generated from mRNA of the cerebellum of Sprague-Dawley rats and the brain of New Zealand white rabbits allowed for the characterization of the rat and rabbit IKAP-encoding mRNAs. Comparison of the predicted amino acid sequence of the IKAP-encoding mRNAs reveals that the human, mouse, rat and rabbit genes encode proteins of 1332, 1333, 1332 and 1334 amino acids in length, respectively, with significant homology to each other (Fig. 2). This analysis further revealed that the amino acid arginine located at amino acid number 696 of human IKAP, which as a result of the missense mutation present in the minor form of the FD mutation is replaced by proline, is conserved in the mouse, rat and rabbit (Fig. 2).

# 3.2. Genomic structure of mouse IKBKAP

To examine the genomic organization of the *IKBKAP* gene, amplification was performed on the DNA of the 129/SvJ mouse using primers recognizing the mRNAencoding sequences. PCR products were sequenced to determine the intron/exon boundaries and the sizes of the introns were determined by either sequencing the smaller introns or estimating the sizes of the larger introns by comparison with DNA size standards. Mouse *IKBKAP*, like human *IKBKAP*, is organized into 37 exons and mouse *IKBKAP* is distributed over approximately 51 kb of genomic DNA (Table 2). Exon 2 contains the start codon and exon 37 contains the stop signal. The consensus donor splice site of intron 20, which is altered in the major FD mutation, is conserved in the mouse (Table 3).

| Table 2                                       |  |
|---|--|
| Exon/intron organization of mouse IKBKAP gene |  |

| Exon |     | Sequence at exon/intron junction and intron size $(nt)^a$ |                 |                        |  |
|------|-----|---|-----------------|------------------------|--|
| No.  | nt  | 5' splice donor   | nt <sup>b</sup> | 3' splice acceptor     |  |
| 1    | 140 | TAG <b>gt</b> gagcattc                                    | 2000*           | tttccctc <b>ag</b> AAA |  |
| 2    | 163 | GAA <b>gt</b> aggtcact                                    | 1150*           | tttgtgaa <b>ag</b> GTG |  |
| 3    | 153 | CAG <b>gt</b> aggtgtaa                                    | 2100*           | tctgatgc <b>ag</b> CTG |  |
| 4    | 82  | CAG <b>gt</b> aagctttg                                    | 900*            | aactccta <b>ag</b> CTC |  |
| 5    | 81  | AAG <b>gt</b> aagcgttt                                    | 1300*           | aaaactgt <b>ag</b> GCA |  |
| 6    | 86  | TTG <b>gt</b> aaggcggg                                    | 1650*           | ctctcttc <b>ag</b> CCT |  |
| 7    | 97  | CAG <b>gt</b> atggaaat                                    | 267             | tcctttgc <b>ag</b> AGG |  |
| 8    | 91  | GAA <b>gt</b> gagtgagc                                    | 1100*           | ctgctttc <b>ag</b> ACC |  |
| 9    | 124 | AAG <b>gt</b> aggggtca                                    | 600*            | tccctacc <b>ag</b> GTA |  |
| 10   | 94  | ATG <b>gt</b> atgacagc                                    | 2150*           | ccacacac <b>ag</b> TCC |  |
| 11   | 231 | GAA <b>gt</b> aagtcgct                                    | 1000*           | cattgtgt <b>ag</b> ACA |  |
| 12   | 165 | GTG <b>gt</b> aagtggaa                                    | 1900*           | tgttttct <b>ag</b> GTG |  |
| 13   | 100 | CTC <b>gt</b> aagttcct                                    | 2900*           | atttgaac <b>ag</b> GAT |  |
| 14   | 192 | CAG <b>gt</b> atcatggt                                    | 1200*           | tttgcttt <b>ag</b> TTC |  |
| 15   | 107 | GGG <b>gt</b> gaggatca                                    | 550*            | cttacaac <b>ag</b> AGT |  |
| 16   | 104 | GAG <b>gt</b> gaatagac                                    | 2000*           | ttctttgc <b>ag</b> GAA |  |
| 17   | 54  | GAG <b>gt</b> atgtaggc                                    | 89              | cctgttgc <b>ag</b> GTC |  |
| 18   | 106 | AAA <b>gt</b> aagctctc                                    | 1250*           | gtattttt <b>ag</b> TGC |  |
| 19   | 116 | CAG <b>gt</b> aagctgac                                    | 334             | ttattttg <b>ag</b> ATG |  |
| 20   | 74  | CAA <b>gt</b> aagtattt                                    | 252             | gtcctcac <b>ag</b> ACT |  |
| 21   | 79  | AAG <b>gt</b> acactttg                                    | 86              | tctttgat <b>ag</b> GTC |  |
| 22   | 80  | CAG <b>gt</b> aagtattt                                    | 1150*           | tggttctt <b>ag</b> GGA |  |
| 23   | 138 | AAA <b>gt</b> gggtgctg                                    | 97              | actacctc <b>ag</b> GTT |  |
| 24   | 86  | AAG <b>gt</b> agagacct                                    | 450*            | actccaac <b>ag</b> GAA |  |
| 25   | 149 | AAG <b>gt</b> atgtggag                                    | 850*            | tttttcct <b>ag</b> GAT |  |
| 26   | 124 | GTG <b>gt</b> aagggttt                                    | 350*            | tttttttc <b>ag</b> GAC |  |
| 27   | 98  | CAG <b>gt</b> atgtggtg                                    | 1800*           | cttgtcac <b>ag</b> GCG |  |
| 28   | 202 | CAG <b>gt</b> aagcaggg                                    | 1050*           | gtcttttc <b>ag</b> GAA |  |
| 29   | 62  | GAC <b>gt</b> gagctcct                                    | 5000*           | cccttgtc <b>ag</b> GAT |  |
| 30   | 63  | CTG <b>gt</b> aaggaagc                                    | 500*            | tccctctt <b>a</b> gGTC |  |
| 31   | 61  | AAG <b>gt</b> gaggatta                                    | 2200*           | gcatcctc <b>ag</b> CCC |  |
| 32   | 114 | TGG <b>gt</b> gagtgcct                                    | 650*            | tcttctct <b>ag</b> ATC |  |
| 33   | 112 | TGC <b>gt</b> acgtacga                                    | 600*            | tttctgac <b>ag</b> GAG |  |
| 34   | 128 | AAG <b>gt</b> atggcttc                                    | 900*            | tcttctct <b>ag</b> ATG |  |
| 35   | 155 | CCG <b>gt</b> aagcttcc                                    | 1800*           | ttctgctt <b>ag</b> GTC |  |
| 36   | 76  | TCG <b>gt</b> tagtgtct                                    | 3500*           | ttgcttcc <b>ag</b> ATC |  |
| 37   | 947 |   |                 |                        |  |

<sup>a</sup> Exon sequences are in uppercase letters; intron sequences are in lower case letters.

<sup>b</sup> Intron lengths are determined by nucleotide sequencing or by electrophoretic fractionating. An asterisk indicates an approximate size.

#### 3.3. Gene locus of mouse IKBKAP

To determine the chromosomal localization of mouse *IKBKAP*, sequence polymorphisms in intron 23 of this gene in C57BL/6J and *M. spretus* mice were identified and used to screen interspecific BSB and BSS backcross

| Table 3         |               |           |            |          |                            |
|-----------------|---------------|-----------|------------|----------|----------------------------|
| Splice junction | sequences for | intron 20 | of mouse a | nd human | <b>IKBKAP</b> <sup>a</sup> |

|           | 5' splice donor                    | 3' splice acceptor                   |
|-----------|------------------------------------|--------------------------------------|
| Consensus | <b>GT</b> AAGTA                    | YYYYYYYYYYYYYNC <b>AG</b>            |
| Human     | <b>GT</b> AAGTA<br><b>GT</b> AAGTA | TTTCCTGTCCTCACAG<br>CTCTgTCTTCTCACAG |

<sup>a</sup> Non-consensus nucleotides are represented by lower case letters.



Fig. 3. (A) Map figures for the Jackson BSB and BSS backcrosses showing part of chromosome 4 with loci linked to Ikbkap. The map is depicted with the centromere at the top. A 3 cM scale bar is shown to the right of the figures. Loci mapping to the same position are listed in alphabetical order. (B) Haplotype figure from the Jackson BSS backcross showing the central portion of chromosome 4 with loci linked to Ikbkap. Loci are listed in order with the most proximal at the top. The black boxes represent the C57BL6/ Jei allele and the white boxes the SPRET/Ei allele. The number of animals with each haplotype is given at the bottom of each column of boxes. The percent recombination (R) between adjacent loci is given to the right of the figure, with the standard error (SE) for each R. Missing typings were inferred from the surrounding data where assignment was ambiguous. Raw data from The Jackson Laboratory were obtained from the World Wide Web address http://www.jax.org/resources/documents/cmdata.

48 39 1 1 2 1 2

panels (Rowe et al., 1994). The mapping results reveal that mouse *Ikbkap* is located on the central part of chromosome 4 mapping to a region where conserved linkage homology has been identified between the human and mouse genome (DeBry and Seldin, 1996; Serikawa et al., 1998). The combined data from the two crosses give the position proximal - D4Mit112 - 2.66 cM + /-1.17 SE - Ikbkap -1.06 + /-0.75 - Gpcr26 - distal (Fig. 3).

#### 4. Conclusions

- 1. We characterized the mouse, rat and rabbit IKAPencoded cDNAs and determined their nucleotide sequences. The gene encodes a 3999, 3996 and 4002 bp open reading frame in the mouse, rat and rabbit, respectively.
- 2. The mouse, rat and rabbit IKAP mRNAs encode predicted proteins that share between 80 and 87% identity with human IKAP.
- 3. The mouse *IKBKAP* gene was localized to the proximalcentral portion of chromosome 4.
- 4. The amino acid residue that is altered in the minor FD mutated gene product is conserved in mouse, rat and rabbit IKBKAP.
- 5. The intron 20 donor splice site sequence that is mutated in the major FD mutation is conserved in mouse IKBKAP.
- 6. The presence in mouse IKBKAP of sequences that are homologous to the human sequences that are mutated in the FD-bearing genes should allow for the generation of mice that bear the mutations present in individuals with FD.

# Acknowledgements

We gratefully acknowledge Lucy Rowe, Mary Barter and Jennifer Johanson (The Jackson Laboratory, Bar Harbor, ME) for their help with analysis of the data generated using the BSS and BSB mapping panels and for their preparation of the Map-Manager figures. This work was funded in part by grants from Dor Yeshorim, the Committee for Prevention of Jewish Genetic Diseases, and Familial Dysautonomia Hope, Inc.

#### References

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403-410.
- Anderson, S.L., Coli, R., Daly, I.W., Kichula, E.A., Rork, M.J., Volpi, S.A., Ekstein, J., Rubin, B.Y., 2001. Familial dysautonomia is caused by mutations of the IKAP gene. Am. J. Hum. Genet. 68, 753-758.
- Axelrod, F.B., 1996. Familial dysautonomia. In: Robertson, D., Low, P.A., Polinsky, R.J. (Eds.). Primer on the Autonomic Nervous System, Academic Press, San Diego, CA, pp. 242-249.
- Axelrod, F.B., Nachtigal, R., Dancis, J., 1974. Familial dysautonomia: diagnosis, pathogenesis and management. Adv. Pediatr. 21, 75-96.
- Brunt, P.W., McKusick, V.A., 1970. Familial dysautonomia: a report of genetic and clinical studies, with a review of the literature. Medicine 49, 343-374.

- Cohen, L., Henzel, W.J., Baeuerle, P.A., 1998. IKAP is a scaffold protein of the IκB kinase complex. Nature 395, 292–297.
- DeBry, R.W., Seldin, M.F., 1996. Human/mouse homology relationships. Genomics 33, 337–351.
- Diffenbach, C.W., Kveksler, G.S., 1995. PCR Primer: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Krappmann, D., Hatada, E.N., Tegethoff, S., Li, J., Klippel, A., Giese, K., Baeuerle, P.A., Scheidereit, C., 2000. The IκB kinase (IKK) complex is tripartite and contains IKKγ but not IKAP as a regular component. J. Biol. Chem. 275, 29779–29787.
- Maayan, C., Kaplan, E., Shachar, S., Peleg, O., Godfrey, S., 1987. Incidence of familial dysautonomia in Israel 1977–1981. Clin. Genet. 32, 106–108.
- Riley, C.M., Day, R.L., Greely, D., Langford, W.S., 1949. Central auto-

nomic dysfunction with defective lacrimation. Pediatrics 3, 468-477.

- Rowe, L.B., Nadeau, J.H., Turner, R., Frankel, W.N., Letts, V.A., Eppig, J.T., Ko, M.S., Thurston, S.J., Birkenmeier, E.H., 1994. Maps from two interspecific backcross DNA panels available as a community genetic mapping resource. Mamm. Genome 5, 253–274.
- Serikawa, T., Cui, Z., Yokoi, N., Kuramoto, T., Kondo, Y., Kitada, K., Guenet, J.L., 1998. A comparative genetic map of rat, mouse and human genomes. Exp. Anim. 47, 1–9.
- Slaugenhaupt, S.A., Blumenfeld, A., Gill, S.P., Leyne, M., Mull, J., Cuajungco, M.P., Liebert, C.B., Chadwick, B., Idelson, M., Reznik, L., Robbins, C., Makalowska, I., Brownstein, M., Krappmann, D., Scheidereit, C., Maayan, C., Axelrod, F.B., Gusella, J.F., 2001. Tissue-specific expression of a splicing mutation in the *IKBKAP* gene causes familial dysautonomia. Am. J. Hum. Genet. 68, 598–605.