MOLECULAR AND CELLULAR BIOLOGY, Sept. 1987, p. 3107–3112
0270-7306/87/093107-06$02.00/0

Vol. 7, No. 9

The Testis-Specific Phosphoglycerate Kinase Gene *pgk-2* Is a Recruited Retroposon

POPPO H. BOER,[1] CHAKER N. ADRA,[1] YUN-FAI LAU,[2] AND MICHAEL W. McBURNEY[1]*

Departments of Medicine and Biology, University of Ottawa, Ottawa, Canada, K1H 8M5,[1] and Howard Hughes Medical Institute, University of California, San Francisco, San Francisco, California 94143[2]

Received 8 May 1987/Accepted 18 June 1987

In both humans and mice, two genes encode phosphoglycerate kinase, a key enzyme in the glycolytic pathway. The *pgk-1* gene is expressed in all somatic cells, is located on the X chromosome, and contains 10 introns. The *pgk-2* gene is expressed only in sperm cells, is located on an autosome, and has no introns. The nucleotide sequence of the *pgk-2* gene suggests that it arose from *pgk-1* more than 100 million years ago by RNA-mediated gene duplication. The *pgk-2* gene may, then, be a transcribed retroposon. Thus, gene duplication by retroposition may have been used as a mechanism for evolutionary diversification.

The mammalian genome contains many examples of multigene families (20). The members of each gene family generally contain the same structure of introns and exons and appear to have evolved from an ancestral gene by duplication followed by divergence of both the protein-coding regions and the sequences responsible for regulating gene expression. Frequently, those genes which duplicated relatively recently are present in the genome in tandem, suggesting that unequal crossing over is responsible for the gene duplication events (16).

The mammalian genome also contains families of related sequences in which family members are dispersed throughout the genome. In those families with relatively few members (usually 2 to 50), the structures of the sequences suggest that only one or a few members are functional genes containing introns, whereas the others are nontranscribed pseudogenes which arose by duplication through an RNA intermediate. These retroposons (33) generally possess no promoter, no introns, and a remnant of a 3′ polyadenylated sequence and are often flanked by short direct repeats. The number of intronless pseudogenes and their locations in the genome are different in different species (31), indicating that most arose relatively recently, a conclusion consistent with the high level of sequence homology between many of the retroposed pseudogenes and their presumed parental genes (34, 43, 44).

There are a few functional genes which have some of the characteristics of retroposons (6, 37), suggesting that retroposition may be a mechanism for gene duplication. We report here the sequence of the mouse *pgk-2* gene. The location and structure of this gene suggest that it arose from the *pgk-1* gene by retroposition in the ancestor common to marsupials and eutherian mammals.

MATERIALS AND METHODS

Molecular cloning. A cDNA library was constructed in the λgt11 expression vector (14, 15) by using RNA derived from the testes of BALB/c mice. The library was screened with a rabbit antibody specific for PGK-2 protein (19). A reactive clone was identified (pMPGK-2), and the cDNA insert was transferred into the sequencing plasmid pSP64 (23).

A mouse genomic library of *Bam*HI-digested DNA was prepared in the EMBL-3 derivative of bacteriophage lambda (17) and was screened initially by hybridization by using a human cDNA to *pgk-1* (25); positive clones were subsequently screened with the pMPGK-2 insert. From the 10.4-kilobase-pair (kbp) genomic clone homologous to the pMPGK-2 probe, the hybridizing *Pst*I-*Bgl*II region was subcloned into pSP64.

Hybridization analysis. Total RNA was isolated from adult mouse tissues (1), electrophoresed in formaldehyde-containing agarose gels (21), blotted onto nylon membranes, and cross-linked by UV irradiation (5). Southern blots of mouse DNA were prepared as described by Maniatis et al. (21). Hybridization to an oligonucleotide-primed [^{32}P]dCTP-labeled probe (11) was done overnight at 42°C in 50% formamide–5× SSC (1× SSC is 0.15 M NaCl plus 0.015 sodium citrate)–0.1% sodium dodecyl sulfate–2.5× Denhardt medium–50 μg of sheared denatured salmon sperm DNA per ml. All blots were washed at 65°C in 0.1× SSC–0.1% sodium dodecyl sulfate.

Sequence analysis. DNA subfragments were cloned into pSP64 or pSP65 plasmids (23), and minipreparations (2) were sequenced by using the chain termination technique (35) and the SP6 primer (18) or M13 reverse primer (New England BioLabs, Inc., Beverly, Mass.). Sequences were assembled and analyzed by computer programs (32).

Transfection assays and detection of PGK activity. DNA transfection of NIH 3T3 and HeLa cells was done by the calcium phosphate precipitation method (13). Cell extracts were prepared by freeze-thawing. Gel electrophoresis and phosphoglycerate kinase (PGK) staining were done essentially as described by Bucher et al. (4), as modified by Paterno and McBurney (30), except that electrophoresis was carried out for only 1 h at 200 V.

RESULTS

***pgk-2* cDNA.** A number of genes that are expressed in mammalian somatic cells have homologs whose expression is restricted to cells undergoing spermatogenesis (N. B. Hecht, *in* J. Rossant and R. A. Peterson (ed.), *Experimental Approaches to Mammalian Embryonic Development*, in press). The gene *pgk-2* is an example of a testis-specific gene (42) which is expressed postmeiotically in sperm cells (10, 12). To understand the relationship between sperm-specific genes and their somatic homologs, we set out to isolate the

* Corresponding author.

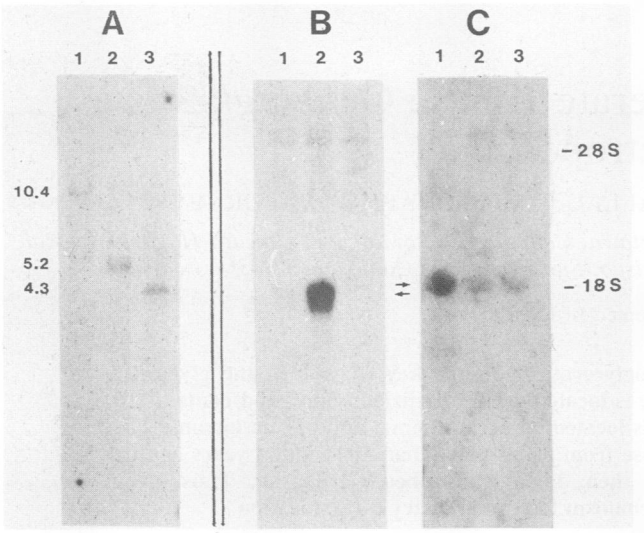

FIG. 1. The *pgk-2* gene is unique in the mouse genome and is expressed only in testes. (A) Southern blot analysis of mouse DNA hybridized with the subcloned *Rsa*I-*Bgl*II fragment from the 3′ noncoding region of the *pgk-2* cDNA (Fig. 3A). The DNA was digested with *Bam*HI (lane 1), *Eco*RI (lane 2), and *Hin*dIII (lane 3); the indicated sizes are in kilobase pairs. (B) RNA blot analysis of different adult mouse tissues: brain (lane 1), testis (lane 2), and spleen (lane 3); hybridization was with the same probe used in panel A. RNA isolated from heart and skeletal muscle also failed to hybridize to the *pgk-2*-specific probe. (C) The RNA blot in panel B was stripped and hybridized with a subclone of *pgk-1* cDNA corresponding to the 3′ noncoding region. The positions of mouse ribosomal RNAs (4.71 and 1.87 kb, respectively) are shown on the right. The arrows indicate the positions of the PGK-1 mRNA (top arrow) and the PGK-2 mRNA (bottom arrow).

mouse *pgk-2* gene to compare it to its somatically expressed counterpart, *pgk-1*. The first step was to isolate a cDNA homologous to *pgk-2* mRNA. RNA isolated from testes was used to construct a cDNA library in the λgt11 expression vector (14, 15). This library was screened with a monospecific antibody reactive with the mouse PGK-2 protein (19), and a cDNA clone (pMPGK-2) was identified. The cDNA insert in pMPGK-2 is 1.6 kbp long (see Fig. 3B for its restriction map). The 180-bp *Rsa*I-*Bgl*II fragment homologous to the noncoding 3′ end of the *pgk-2* mRNA was used to probe Southern blots of mouse genomic DNA (Fig. 1A). A single restriction fragment hybridized in each sample of digested DNA, suggesting the presence of a single copy of the *pgk-2* gene in the mouse genome.

Blot transfers of RNA isolated from a number of mouse organs were also probed with the *Rsa*I-*Bgl*II fragment (Fig. 1B), and only testes contained transcripts which hybridized. All organs contained RNA which hybridized to a probe specific for the *pgk-1* mRNA (Fig. 1C).

When inserted into the pSV2 expression vector (29) and transfected into HeLa cells, the pMPGK-2 cDNA insert directed the synthesis of a protein with PGK activity which had the electrophoretic characteristics of PGK-2A, the BALB/c isoenzyme of testis-specific PGK-2 (Fig. 2, lane 1). Thus, the pMPGK-2 insert appears to contain the entire coding region for PGK-2.

***pgk-2* gene.** The region of the mouse genome which hybridized to the pMPGK-2 probe was cloned as a 10.4-kbp *Bam*HI restriction fragment from a genomic library made from C3H/He mouse DNA (C. N. Adra, N. Ellis, and M. W. McBurney, manuscript submitted). To unequivocally establish the identity of the gene, this genomic fragment (Fig. 3A) was cotransfected with pSVneo (38) into NIH 3T3 cells, approximately 25 G418-resistant colonies were pooled, and extracts were prepared. The clones transformed with the genomic sequence contained the PGK-2B isoenzyme; that is, a PGK enzyme with electrophoretic mobility identical to that of enzyme isolated from testes of C3H/He mice (Fig. 2, lanes 3 and 6).

The region of the 10.4-kbp genomic fragment which hybridized to the pMPGK-2 probe was confined to a 2.1-kbp *Pst*I-*Bgl*II fragment. The restriction map of the genomic fragment is identical to that of the pMPGK-2 insert (Fig. 3).

Both the pMPGK-2 insert and the gene were sequenced (Fig. 4). The two sequences are identical except for the single nucleotide at position +452 (numbering is from the site of translation initiation). The gene, *pgk-2b*, predicts an arginine for codon 151, whereas the cDNA, derived from *pgk-2a*, predicts a glutamine. This single amino acid difference appears to be responsible for the electrophoretic difference between the allelic forms of PGK-2.

The predicted 417-amino-acid sequence of PGK-2 (Fig. 4) has all of the well-conserved stretches of amino acids of other PGK proteins (28). When compared over the coding region, the nucleotide sequence of the *pgk-2* gene is 80% identical to the mouse *pgk-1* cDNA (27). This level of homology drops to 51% in the 3′ noncoding region (Fig. 5). The upstream region of *pgk-2* bears no sequence homology to that of mouse *pgk-1* (C. N. Adra and M. W. McBurney, manuscript submitted) but contains a number of functional sequence motifs, including a CCAAT box at −160 relative to the translation start site and a presumptive transcription start site (7) at −57. There is no TATA box.

We compared the mouse *pgk-2* sequence with a human *pgk*-related sequence originally thought to be an intronless pseudogene (40). The sequence homology over the presumptive coding region was 85%. The levels of homology over the 5′ and 3′ flanking regions were 66 and 59%, respectively (Fig. 5). Homology extended well into the upstream region and was particularly striking in the region of the presumptive

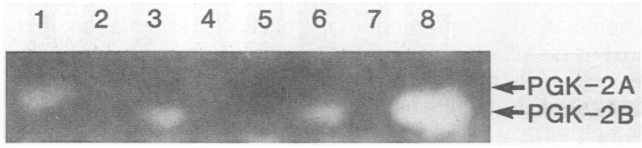

FIG. 2. PGK-2 activity in cells transfected with the *pgk-2* gene and cDNA. The full-length *pgk-2* cDNA insert of 1.6 kbp was inserted into the pSV2 vector, and the resulting plasmid, pCASV2*pgk-2*, was transfected into HeLa cells. After 48 h, a cell extract was prepared and run in lane 1. NIH 3T3 cells were transfected with pSV2neo alone (lanes 4 and 7), pCASV2*pgk-2* (lanes 2 and 5), and the 10.4-kbp *pgk-2* gene inserted into pSP64 (lanes 3 and 6). Stable transformants of the NIH 3T3 cells were selected in G418, the cells were pooled, and extracts were prepared. An extract prepared from the testis of a strain C3H/He mouse was run in lane 8. Gel electrophoresis conditions and staining for PGK activity have been described (4, 30). Only the portion of the gel containing PGK-2 is shown. The more slowly migrating PGK-1 human and mouse enzymes were approximately 10 times more active than the PGK-2 enzymes identified in lanes 1, 3, and 6. The arrows indicate the locations of the PGK-2A and PGK-2B allelic variants. The expression of the pCASV2*pgk-2* plasmid was observed in transiently (lane 1) but not stably (lanes 2 and 4) transfected cells, while expression of the stably integrated *pgk-2* gene (lanes 3 and 6) was higher than in transiently transfected cells (data not shown).

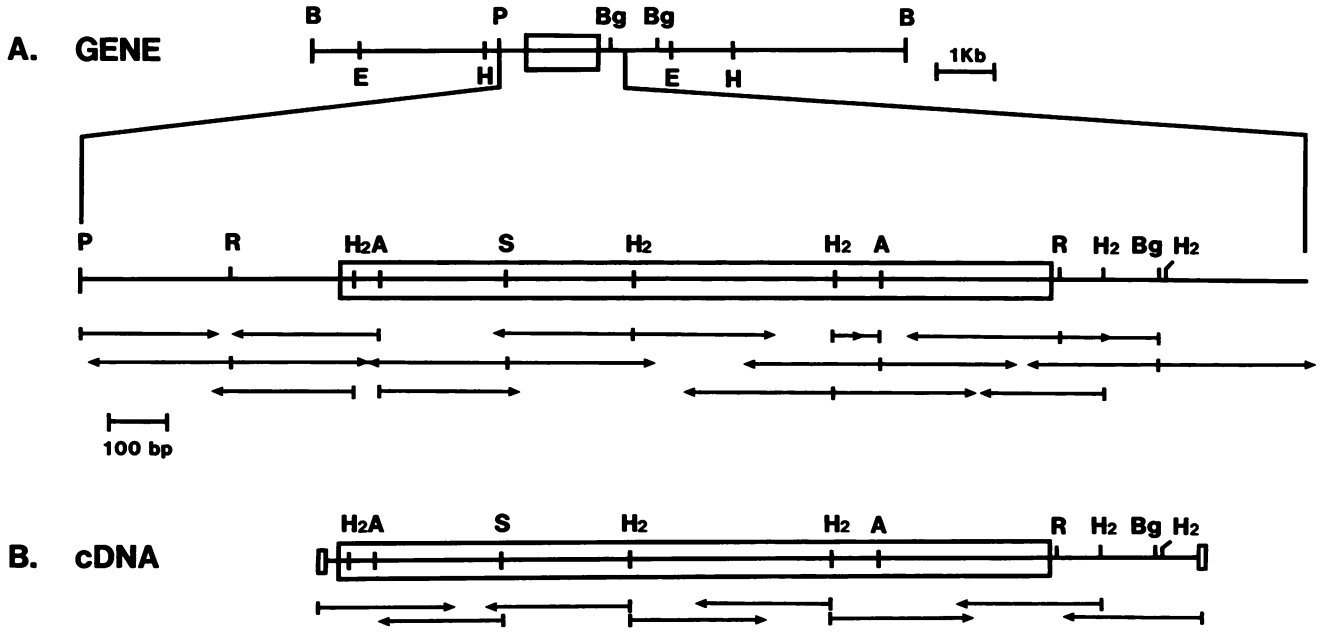

FIG. 3. Restriction maps of the *pgk-2* gene and cDNA are colinear. The coding regions of the *pgk-2* gene (A) and cDNA (B) are shown as open boxes; transcription is from left to right. Restriction sites: A, *Acc*I; B, *Bam*HI; Bg, *Bgl*II; E, *Eco*RI; H, *Hin*dIII; H_2, *Hin*cII; P, *Pst*I; R, *Rsa*I; S, *Sst*I. Shown below each restriction map is the sequencing strategy. The arrows indicate the direction and extent of sequence determined.

functional motifs: the CCAAT box, the transcription start site, and the relatively unusual ATTAAA (3) polyadenylation signal (Fig. 4). The high degree of homology between the mouse *pgk-2* sequence and the human sequence strongly suggests that the human sequence is, in fact, *pgk-2*. To confirm this suggestion, plasmid DNA containing the presumptive human *pgk-2* gene was transfected into mouse teratocarcinoma cells. The transformed cells expressed a PGK activity with the electrophoretic mobility of the mouse PGK-2A isoform (M. W. McBurney and C. N. Adra, unpublished results).

DISCUSSION

PGK-2 and PGK-1 were identified in eutherian mammals and marsupials as distinct isoforms on the basis of their different electrophoretic mobilities (41). In most species the PGK-2 isoform is expressed exclusively in male germ cells. A species survey has indicated that PGK-2 is present in most marsupials and eutherian mammals but absent from fish, reptiles, birds, and monotremes (41). This suggests that the *pgk-2* gene arose only once during evolution, between 100 and 200 million years ago, after the divergence of birds and monotremes from the lineage which gave rise to marsupials and eutherian mammals. Consistent with the idea that *pgk-2* arose only once is the fact that both the mouse (9) and human (39) *pgk-2* genes are linked to their major histocompatibility complexes. The results reported here further support this view. By comparing the 3′ noncoding region of the mouse and human *pgk-2* genes, a region thought to be under low sequence constraint, and using analyses described by Soares et al. (37), we deduced that the mouse and human *pgk-2* genes derived from a common gene and started diverging some 95 million years ago, approximately the time when the rodent and human lineages separated.

The most striking characteristic of the *pgk-2* gene is the complete absence of introns, particularly in view of the 10 introns present in human (24) and mouse (Adra and McBurney, submitted) *pgk-1* genes. The high degree of sequence homology between the mouse *pgk-2* gene and the mouse *pgk-1* cDNA over the coding region strongly suggests that *pgk-2* arose from *pgk-1* by gene duplication. The absence of introns in *pgk-2* and its autosomal location suggest that *pgk-2* originated as a retroposon derived from *pgk-1*. The sequence homology between the *pgk-2* gene and the *pgk-1* cDNA extends downstream of the *pgk-2* polyadenylation signal. This implies that a new polyadenylation signal evolved in the *pgk-2* retroposon, giving rise to a transcript with a 3′ untranslated region about 130 nucleotides shorter than that of *pgk-1*. The *pgk-2* mRNA is indeed shorter than the *pgk-1* mRNA (Fig. 1B and C). By comparing the sequence divergence between the mouse *pgk-2* gene and the *pgk-1* cDNA in the 3′ noncoding region and again using the analyses of Soares et al. (37), we deduced that these two genes started diverging some 130 million years ago, a figure in agreement with that deduced from the evolutionary considerations discussed above.

In addition to their lack of introns, retroposons generally are not functional genes, have a residual oligo(A) sequence at their 3′ ends, and are flanked by short (7 to 20 bp) direct repeats (34, 43). Neither the poly(A) region nor the flanking direct repeats are apparent in the *pgk-2* gene sequence. A short adenine-rich sequence is located downstream of the mouse and human *pgk-2* genes at the end of the sequence homologous to the *pgk-1* cDNA (Fig. 4). A possible direct repeat downstream of this adenine-rich tract shows only a low level of sequence homology to a number of regions upstream of the coding sequence. Because of the age of the *pgk-2* gene, it is not surprising that the adenine-rich tract and flanking direct repeats have diverged in sequence so much that they are no longer evident.

The development of *pgk-2* from a retroposon may have occurred by integration of the retroposed cDNA into a region already containing a promoter. However, it is also

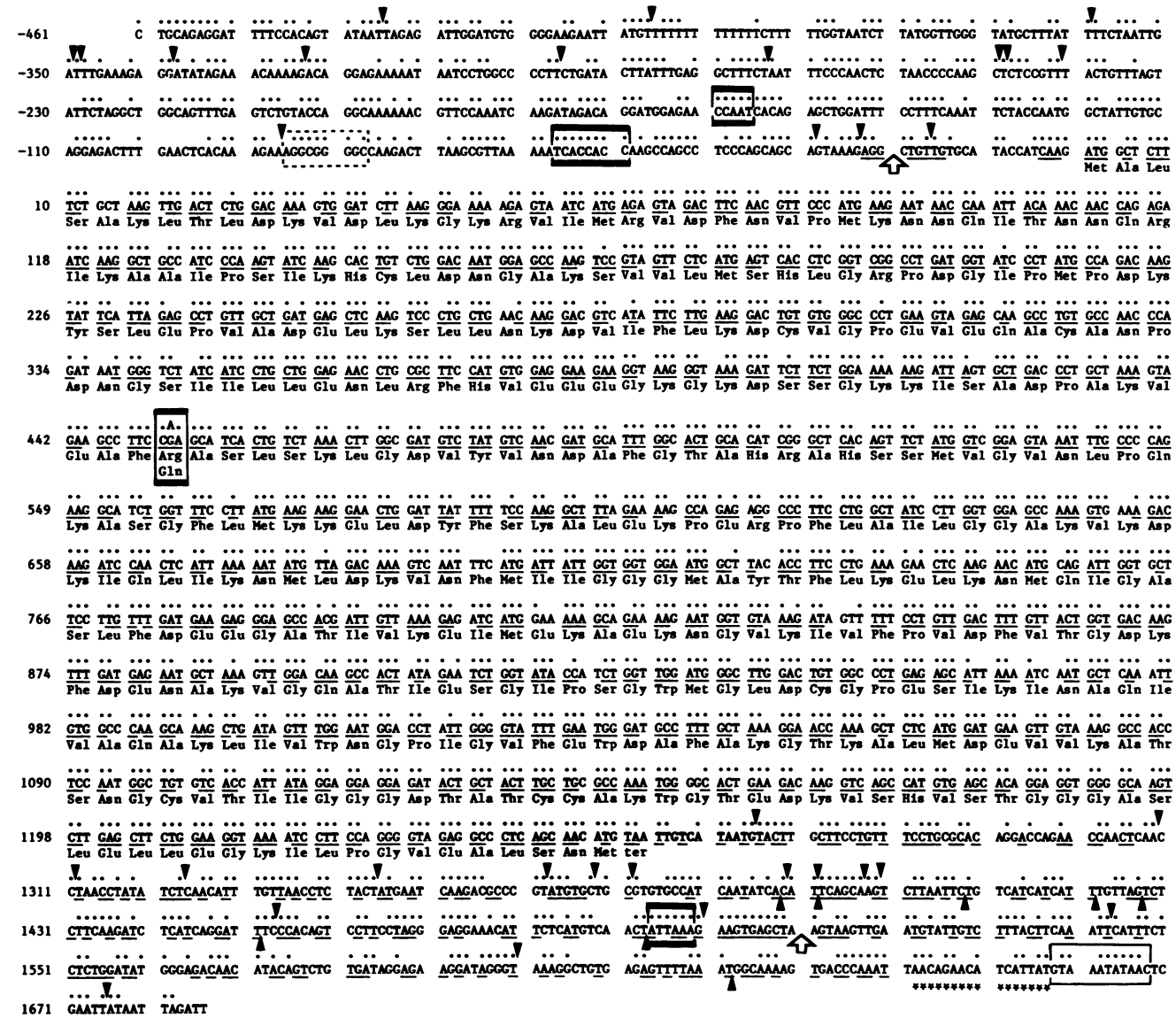

FIG. 4. Nucleotide sequence of the mouse *pgk-2* gene. The nucleotides are numbered from the start codon, and the deduced amino acids for PGK-2 are shown below. The entire sequence could be aligned with the human *pgk* sequence (40) to which we added G residues at positions 1021 and 1187. A dot over a nucleotide indicates identity with the human sequence, and small insertions of one to seven nucleotides in the human sequence are indicated by downward-pointing arrowheads. The boxed motifs are the common regulatory signals and include a CCAAT box at −160, a putative transcription start site at −57, and a polyadenylation signal at +1494. The poly(A) tail in the cDNA was added at position +1509 at the site of the large open arrow. The extent of the *pgk-2* cDNA is indicated by large open arrows at positions −20 and +1509. A mouse strain polymorphism occurs at position +452 (indicated by a box), at which the strain C3H/He codon is CGA and the strain BALB/c codon is CAA. Between nucleotides −23 and +1640, the *pgk-2* sequence could also be aligned with that of mouse *pgk-1* cDNA (28). Here, identical nucleotides are underlined, and insertions of one to seven nucleotides in the cDNA sequence are indicated by upward-pointing arrowheads. The asterisks at +1642 indicate the adenine-rich region, and the box at +1658 indicates the downstream direct repeat proposed by Tani et al. (40). The broken-line box at −77 is a GC box which represents a potential binding site for the Sp1 transcription factor (8).

possible that the retroposed cDNA carried its own promoter sequence with it. The promoter of the *pgk-1* gene (Adra and McBurney, submitted), like the promoters of many constitutively expressed genes, contains multiple GC boxes which function as important promoter elements (8). A processed transcript of *pgk-1* initiated at an upstream site and carrying one or more of these GC boxes could have served as the template for reverse transcription, resulting in a retroposon containing its own promoter. A GC box is, in fact, present at −86 upstream of the *pgk-2* coding region (Fig. 4) and may be a remnant of the retroposition event.

During spermatogenesis, the mammalian X chromosome becomes transcriptionally inactive, and all X-linked genes, including *pgk-1*, are turned off (26). Survival of sperm cells apparently requires glycolysis mediated by PGK-2 (41). Thus, the appearance of *pgk-2* must have preceded the evolution of X inactivation in spermatogenesis. However, the high degree of homology between the coding regions of *pgk-1* and *pgk-2* strongly suggests that both genes have been under continuous selective pressure. If X inactivation during spermatogenesis is the selective pressure acting on *pgk-2*, it seems likely that *pgk-2* became expressed in sperm very

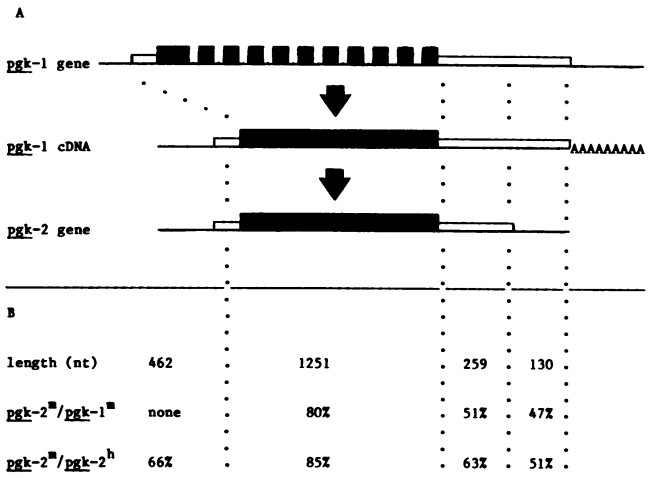

FIG. 5. Relationship between nucleotide sequences of *pgk* genes. (A) Model for the presumed origin of *pgk-2*. The ancestral *pgk-1* gene, like the human and mouse genes, is presumed to contain 10 introns. The processed *pgk-1* mRNA is thought to have been copied into a cDNA sequence which was integrated into an autosome. The new polyadenylation signal evolved some 130 nucleotides upstream of that used by *pgk-1*. Coding regions are shown as solid boxes, and nontranslated regions of mRNAs are indicated by open boxes. (B) Nucleotide (nt) sequence comparison over four regions of the *pgk-2* gene, the upstream, coding, 3′ noncoding, and downstream regions. The human *pgk-2* sequence is from Tani et al. (40), and the mouse *pgk-1* cDNA sequence is from Mori et al. (28). The mouse *pgk-1* 5′ region (Adra and McBurney, submitted), like that of human *pgk-1* (24, 36), is a G+C-rich sequence that shows no homology to the 5′ region of *pgk-2*.

shortly after its birth from the *pgk-1*-derived retroposon and that the process of X inactivation in spermatogenesis evolved within a relatively short period after the appearance of the *pgk-2* gene.

The results reported above strongly suggest that retroposition occurred very early in the evolution of mammals and that, at least occasionally, functional genes resulted from integrated reverse transcripts.

After this work was completed, the study of J. R. McCarrey and K. Thomas (22) was reported, in which they deduced that the human *pgk* sequence (40) is indeed the *pgk-2* gene and that its structure is reminiscent of a processed retrogene.

ACKNOWLEDGMENTS

We thank C.-Y. G. Lee for the PGK-2-specific antisera used to screen the λgt11 library, R. Korneluk for the SP6 sequence primer, and K. Jardine, K. Chan, and M. Elish for technical assistance.

This work was funded by grants from the Medical Research Council of Canada and the National Cancer Institute of Canada. C.N.A. was supported by a predoctoral scholarship from the National Sciences and Engineering Research Council of Canada, Y.-F.L. is an Associate Investigator of the Howard Hughes Medical Institute, and M.W.M. is a National Cancer Institute of Canada Terry Fox Cancer Research Scientist.

LITERATURE CITED

1. **Auffray, C., and F. Rougeon.** 1980. Purification of mouse immunoglobulin heavy chain messenger RNAs from total myeloma tumor RNA. Eur. J. Biochem. **107**:303–314.
2. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. **7**:1513–1523.
3. **Birnstiel, M. L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3′ processing: the end is in site. Cell **41**:349–359.
4. **Bucher, T., W. Bender, R. Fundele, H. Hofner, and I. Linke.** 1980. Quantitative evaluation of electrophoretic allo and isozyme patterns. FEBS Lett. **115**:319–325.
5. **Church, G., and W. Gilbert.** 1984. Genomic sequencing. Proc. Natl. Acad. Sci. USA **81**:1991–1995.
6. **Clark, B. D., K. L. Collins, M. S. Gandy, A. C. Webb, and P. E. Auron.** 1986. Genomic sequences for human prointerleukin 1 beta: possible evolution from a reverse transcribed prointerleukin 1 alpha gene. Nucleic Acids Res. **14**:7897–7914.
7. **Corden, J., B. Wasylyk, A. Buchwalder, P. Sassone-Corsi, C. Kedinger, and P. Chambon.** 1980. Promoter sequences of eucaryotic protein-coding genes. Science **209**:1406–1414.
8. **Dynan, W. S., and R. Tjian.** 1985. Control of eukaryotic messenger RNA synthesis by sequence-specific DNA-binding proteins. Nature (London) **316**:774–778.
9. **Eicher, E. M., M. Cherry, and L. Flaherty.** 1978. Autosomal phosphoglycerate kinase linked to mouse major histocompatibility complex. Mol. Gen. Genet. **158**:225–228.
10. **Erickson, R. P., A. M. Michelson, M. P. Rosenberg, E. Sanchez, and S. H. Orkin.** 1985. Post-meiotic transcription of phosphoglycerate kinase-2 in mouse testes. Biosci. Rep. **5**:1087–1091.
11. **Feinberg, A. P., and B. Vogelstein.** 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. **132**:6–13.
12. **Gold, B., H. Fryimato, J. M. Kramer, R. P. Erickson, and N. B. Hecht.** 1983. Haploid accumulation and translational control of phosphoglycerate kinase-2 messenger RNA during mouse spermatogenesis. Dev. Biol. **98**:392–399.
13. **Graham, F. L., and A. J. Van der Eb.** 1973. A new technique for the assay of infectivity of human adenovirus 5 DNA. Virology **52**:456–467.
14. **Gubler, U., and B. J. Hoffman.** 1983. A simple and very efficient method for generating cDNA libraries. Gene **25**:263–269.
15. **Huynh, T. V., R. A. Young, and R. W. Davis.** 1985. Construction and screening cDNA libraries in gamma gt 10 and gamma gt 11, p. 1–48. *In* D. M. Glover (ed.), DNA cloning techniques: a practical approach, vol. 1. IRL Press, Oxford.
16. **Jeffreys, A. J., and S. Harris.** 1982. Processes of gene duplication. Nature (London) **296**:9–10.
17. **Kaiser, K., and N. E. Murray.** 1985. The use of phage lambda replacement vectors in the construction of representative genomic DNA libraries, p. 49–78. *In* D. M. Glover (ed.), DNA cloning: a practical approach, vol. 1. IRL Press, Oxford.
18. **Korneluk, R. G., F. Quan, and R. A. Gravel.** 1985. Rapid and reliable dideoxy sequencing of double-stranded DNA. Gene **40**:317–323.
19. **Lee, C.-Y., D. Niesel, B. Pegoraro, and R. P. Erickson.** 1980. Immunological and structural relatedness of isozymes and genetic variants of 3-phosphoglycerate kinase from the mouse. J. Biol. Chem. **255**:2590–2595.
20. **Maeda, N., and O. Smithies.** 1986. The evolution of multigene families: human haptoglobin genes. Annu. Rev. Genet. **20**:81–108.
21. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
22. **McCarrey, J. R., and K. Thomas.** 1987. Human testis-specific PGK gene lacks introns and possesses characteristics of a processed gene. Nature (London) **326**:501–505.
23. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient in vitro synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7055.
24. **Michelson, A. M., C. C. F. Blake, S. T. Evans, and S. H. Orkin.** 1985. Structure of the human phosphoglycerate kinase gene and the intron-mediated evolution and dispersal of the nucleotide-binding domain. Proc. Natl. Acad. Sci. USA **82**:6965–6969.
25. **Michelson, A. M., A. F. Markham, and S. H. Orkin.** 1983. Isolation and DNA sequence of a full-length cDNA clone for human X chromosome-encoded phosphoglycerate kinase. Proc.

Natl. Acad. Sci. USA **80**:472–476.
26. **Monesi, V.** 1971. Chromosome activities during meiosis and spermiogenesis. J. Reprod. Fertil. **13**:1–14.
27. **Mori, N., J. Singer-Sam, C.-Y. Lee, and A. D. Riggs.** 1986. The nucleotide sequence of a cDNA clone containing the entire coding region for mouse X-chromosome linked phosphoglycerate kinase. Gene **45**:275–280.
28. **Mori, N., J. Singer-Sam, and A. D. Riggs.** 1986. Evolutionary conservation of the substrate-binding cleft of phosphoglycerate kinases. FEBS Lett. **204**:313–317.
29. **Mulligan, R. C., and P. Berg.** 1980. Expression of a bacterial gene in mammalian cells. Science **209**:1422–1427.
30. **Paterno, G. D., and M. W. McBurney.** 1985. X chromosome inactivation during induced differentiation of a female mouse embryonal carcinoma cell line. J. Cell Sci. **75**:149–163.
31. **Piechaczyk, M., J. M. Blanchard, S. Riaad-El Sabouty, C. Dani, L. Marty, and P. Jenateur.** 1984. Unusual abundance of vertebrate 3-phosphate dehydrogenase pseudogenes. Nature (London) **312**:469–471.
32. **Queen, C., and L. J. Korn.** 1983. A comprehensive sequence analysis program for the IBM personal computer. Nucleic Acids Res. **12**:581–599.
33. **Rogers, J. H.** 1983. Retroposons defined. Nature (London) **301**:460.
34. **Rogers, J. H.** 1985. The origin and evolution of retroposons. Int. Rev. Cytol. **93**:187–279.
35. **Sanger, F., S. Nicklen, and A. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
36. **Singer-Sam, J., D. H. Keith, K. Tani, R. L. Simmer, L. Shively, S. Lindsay, A. Yoshida, and A. D. Riggs.** 1984. Sequence of the promoter region of the gene for human X-linked 3-phosphoglycerate kinase. Gene **32**:409–417.
37. **Soares, M. B., E. Schon, A. Henderson, S. K. Karathanasis, R. Cate, S. Zeitlin, J. Chirgwin, and A. Efstratiadis.** 1985. RNA-mediated gene duplication: the rat preproinsulin I gene is a functional retroposon. Mol. Cell. Biol. **5**:2090–2103.
38. **Southern, P. J., and P. Berg.** 1982. Transformation of mammalian cells to antibiotic resistance with a bacterial gene under control of the SV40 early region promoter. J. Mol. Appl. Genet. **1**:327–341.
39. **Szabo, P., K.-H. Grzeschik, and M. Siniscalco.** 1984. A human autosomal phosphoglycerate kinase locus maps near the HLA cluster. Proc. Natl. Acad. Sci. USA **81**:3167–3169.
40. **Tani, K., J. Singer-Sam, M. Munns, and A. Yoshida.** 1985. Molecular cloning and structure of an autosomal processed gene for human phosphoglycerate kinase. Gene **35**:11–18.
41. **VandeBerg, J. L.** 1985. The phosphoglycerate kinase isozyme system in mammals: biochemical, genetic, developmental, and evolutionary aspects. Isozymes Curr. Top. Biol. Med. Res. **12**: 133–187.
42. **VandeBerg, J. L., D. W. Cooper, and P. J. Close.** 1973. Mammalian testis phosphoglycerate kinase. Nature (London) New Biol. **243**:48–50.
43. **Vanin, E. F.** 1985. Processed pseudogenes: characteristics and evolution. Annu. Rev. Genet. **19**:253–272.
44. **Weiner, A. M., P. L. Deininger, and A. Efstratiadis.** 1986. Nonviral retroposons: genes, pseudogenes, and transposable elements generated by the reverse flow of genetic information. Annu. Rev. Biochem. **55**:631–661.