Polymorphism in the *Dirofilaria immitis* immunodominant antigen gene

K. Sugane[1*], K. Nakayama[2] and H. Kato[3]

[1]Department of Parasitology, [2]Department of Anatomy, Shinshu University School of Medicine, Asahi 3-1-1, Matsumoto City, Nagano Prefecture 390, Japan: [3]Department of Bioscience, Fukui Prefecture University School of Bioscience, Fukui, Japan

Abstract

Dg2, a gene encoding a 34 kDa immunodominant antigen of *Dirofilaria immitis* was cloned and demonstrated to be specifically expressed in the larval stage. In this study, a newly constructed genomic DNA library was screened by hybridization with Dg2. One of the resulting positive clones was similar to Dg2 in the structure of its exonic regions but different in number, position, size and sequence of introns. This was designated DgK. Full-length cDNA was isolated using the rapid amplification of cDNA ends (RACE) method to study the transcript corresponding to DgK. Sequence analysis revealed that the mRNA corresponding to DgK is *trans*-spliced during post-transcriptional processing because the 5′ end of the amplified cDNA contains seven nucleotides of the nematode-spliced leader (SL) sequence.

Introduction

Dirofilaria immitis, a canine filarial nematode, is transmitted by mosquito vectors such as *Aedes togoi* and *A. albopictus* in Japan. It occasionally infects humans resulting in dirofilariasis with the formation of pulmonary or subcutaneous nodules. Diagnosis is made by histological identification of the worms in the nodules. However, immunodiagnosis using specific antigens is also useful. To this end, the genes encoding specific antigens of *D. immitis* have been cloned (Grandea III *et al*., 1989; Culpepper *et al*., 1992; Poole *et al*., 1992; Frank *et al*., 1996; Hong *et al*., 1996). We previously reported the characterization of a cDNA clone encoding a 34kDa immunodominant antigen (cD34) (Sun *et al*., 1991; Sugane & Sun, 1994), which is useful for immunodiagnosis of dirofilariasis with ELISA (Sun & Sugane, 1992). In addition, mRNA corresponding to cD34 was detected only in microfilariae in Northern blot assay (Sun *et al*., 1992). For elucidation of the mechanism of stage-specific expression of the gene, a genomic clone, Dg2, was cloned and its complete structure was analysed (Sun & Sugane, 1994). However, the 5′-upstream region of Dg2 contains only up to 151 bp upstream of the transcription initiation site. Therefore, a newly constructed genomic DNA library was screened with ^{32}P-cD34 probe to obtain the 5′ end of the gene. As a result, the genomic clone DgK was isolated. A comparison of the nucleotide sequences of DgK and Dg2 showed more than 91% identity in the exonic regions, while they differed from each other in the number, location, size and sequence of introns. We report here the molecular characterization of DgK and comparison with Dg2.

Materials and methods

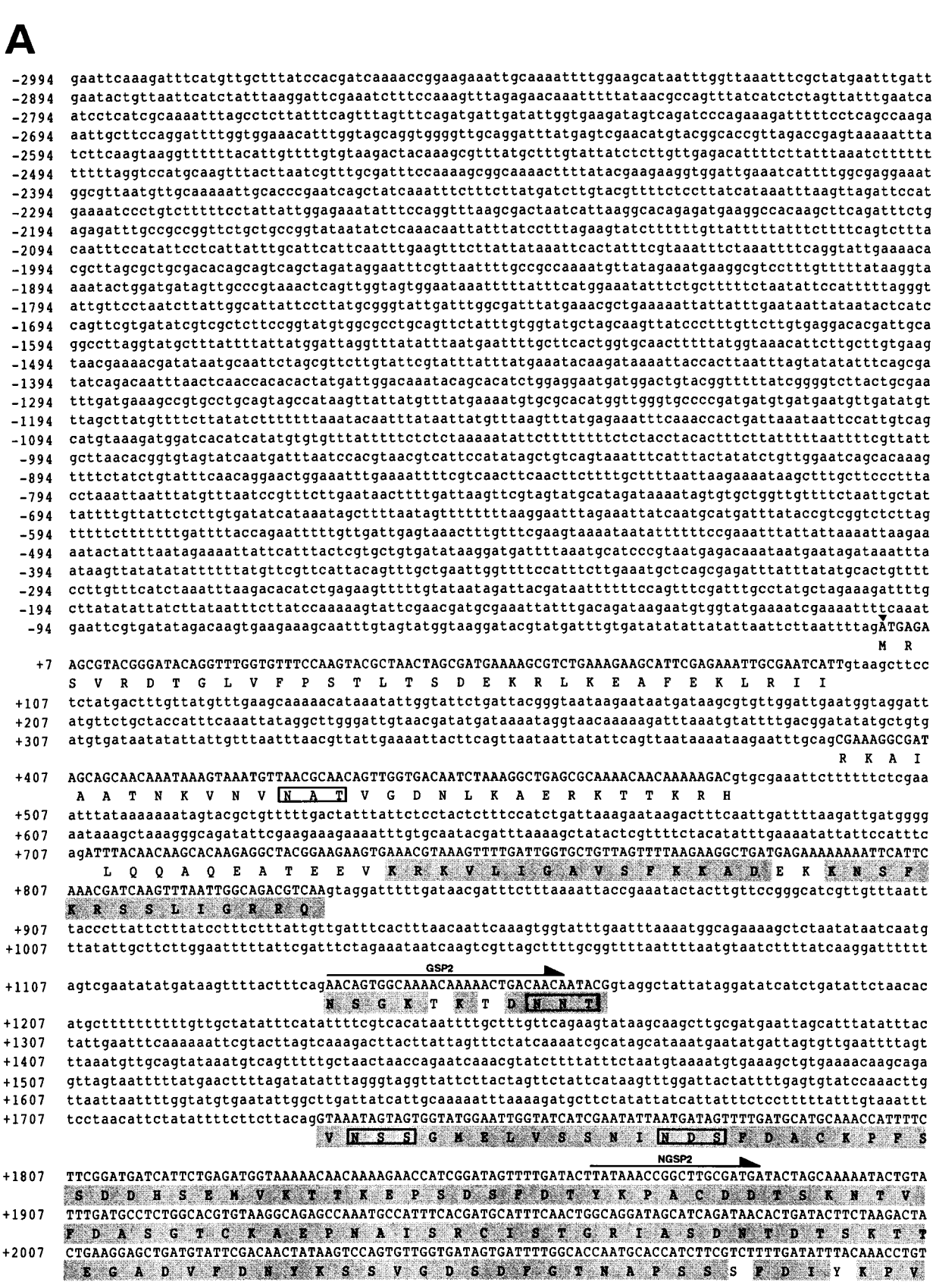
Parasites

Fresh male and female adult worms of *D. immitis* and whole blood from infected dogs were kindly provided by Dr J.X. Wang (1st Department of Physiology, Shinshu University School of Medicine). Microfilariae (mf) were harvested from whole blood of infected dogs by centrifugation after treatment with saponin (Ando *et al*., 1980). All worms were washed several times with sterile 0.85% saline solution and stored at –85°C until use.

*Fax: +81 263 37 3092

Note: The nucleotide sequence data reported in this paper are available in the DDBJ, EMBL and GenBank data bases under the accession numbers AB004252 and AB004253.

https://doi.org/10.1017/S0022149X0070054X Published online by Cambridge University Press



https://doi.org/10.1017/S0022149X0070054X Published online by Cambridge University Press

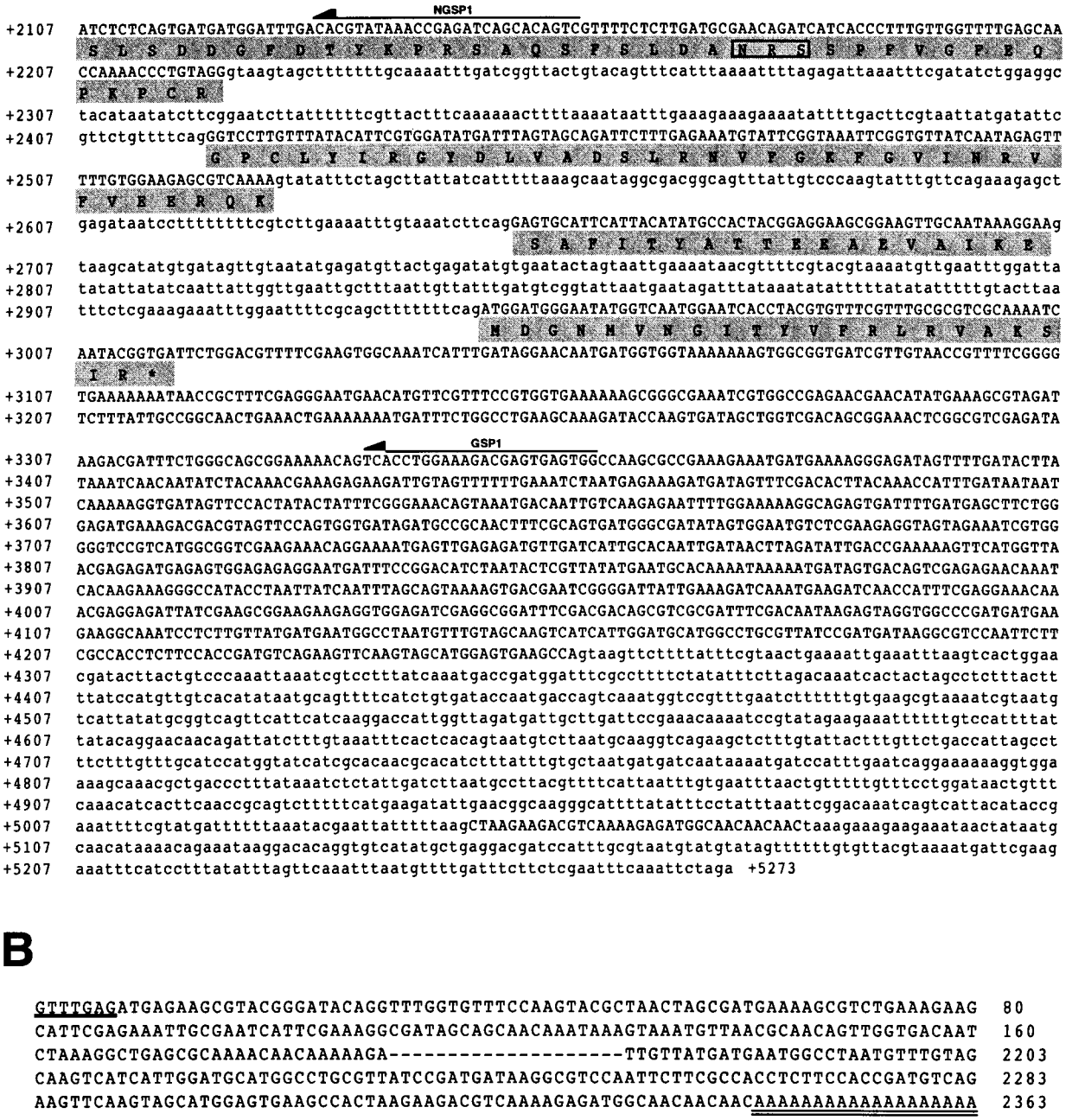


Fig. 1. Nucleotide sequence and predicted amino acid sequence of DgK and cDNA amplified with the RACE method. A, Nucleotide sequence of DgK. Nucleotides are numbered from the first position of exon 1 which is indicated by an arrowhead. Exon and intron composition were determined as described in Materials and methods. The amino acid sequence deduced from nucleotide sequence of exons is shown in one letter code below the coding DNA. Potential N-linked glycosylation sites are indicated by an open box and the amino acid sequences identical to cD34 are shaded. The arrows GSP2 and NGSP2 represent the degenerate sense primers used for amplification of the 3′ end. The arrows GSP1 and NGSP1 represent the gene-specific antisense primers used for amplification of the 5′ end. * stop codon. B, Nucleotide sequence of cDNA amplified with the RACE method. The combined nucleotide sequence of the 5′- and 3′-RACE products is 2363 bp in length. Putative SL sequence and poly(A) tail are indicated with single and double underlines, respectively. The nucleotide sequence of the middle region (191-2173, dashed line) is omitted.

Extraction of RNA and genomic DNA

Total RNA was extracted from 0.3 ml of packed mf using the phenol–chloroform extraction method (Chomczynski & Sacchi, 1987). Poly $(A)^+$ RNA was isolated with QuickPre® Micro mRNA Purification Kit (Pharmacia Biotech, Sweden). Genomic DNA was isolated from parasites as described by McReynolds *et al*. (1986).

https://doi.org/10.1017/S0022149X0070054X Published online by Cambridge University Press

Construction and screening of the genomic DNA library

Genomic DNA prepared from female adult worms was partially digested with Sau 3A1 and electrophoresed on 0.5% agarose gel. Then, DNA fragments ranging from 4 kb to 16 kb were isolated and inserted into the *Bam* HI site of Charomid 9–36 arm DNAs (Nippon Gene, Co., Ltd, Japan). Packaging was carried out using commercially available extracts (Gigapack Gold, Stratagene, California, USA). The recombinant phage particles were propagated on *Escherichia coli* (VOS 257). Approximately 3×10^4 recombinant phages were screened by colony hybridization using a ^{32}P-labelled *Eco* RI-*Sph* I fragment of cD34(320 bp in length, containing exon 1 and 2 of Dg2) as a probe (Sambrook *et al*., 1989). Colonies which gave a positive hybridization signal were taken through three rounds of purification.

Nucleotide sequence

Genomic DNA fragments were recovered from positive clones, digested with an appropriate restriction enzyme and subcloned into pBluescript II SK(+) and pGEM3Z. The nucleotide sequence of the plasmid inserts was determined by dideoxynucleotide chain termination method (Sanger *et al*., 1977) using Dye Terminator Cycle Sequencing FS Ready Reaction Kit (PE Applied Biosystems, USA).

cDNA amplification by the RACE method

cDNA synthesis from 1 μg poly $(A)^+$ RNA derived from mf and adapter addition to dscDNA ends were carried out using a Marathon^{TM} cDNA Amplification Kit (CLONTECH, USA) according to the manufacturer's protocol and 5′- and 3′-RACE reactions through the nested PCR were performed using KOD Dash DNA polymerase (TOYOBO, Japan). Gene specific primers GSP1 (5′-GGCCACTCACTCGTCTTTCCAGGTGA-3′: position from +3361 to +3338), GSP2 (5′-AACAGTGGCAAAACAAAA ACTGACAACA-3′: position from +1137 to +1164), NGSPI (5′-CGACTGTGCTGATCTCGGTTTATACGTG-3′: position from +2158 to +2131) and NGSP2 (5′-TATAAACCGGCTT GCGATGA-3′: position from +1868 to +1887) (see fig. 1A) were designed from the nucleotide sequence obtained from DgK. Adapter primers, AP1 and AP2 in the kit, were also used. The conditions for the first round of PCR were as follows: 94°C for 1 min and then 25 cycles as follows: 94°C for 30 sec and 68°C for 4 min. Nested PCR was carried out at 94°C for 1 min and then 25 cycles of the following scheme: 94°C for 30 sec, 60°C for 10 sec and 74°C for 2 min. PCR products were separated by 0.8% agarose gel electrophoresis, and specific fragments isolated from the gel were purified to use as templates for direct sequencing.

Southern blot analysis

Genomic DNAs of mf and adult worms were digested with restriction endonucleases, *Eco* RI, *Hae* III, *Hind* III or *Eco* RI + *Ssp* I, electrophoresed on 0.9% agarose gel, transferred onto Hybond N+ membrane (Amersham, UK) and prehybridized for 2 h at 65°C. The filters were then hybridized with ^{32}P-labeled Probe 1 (*Eco* RI-*Eco* RI fragment of DgK) or probe 2 (*Eco* RI-*Eco* RI fragment of Dg2) for 12 h at 65°C (see fig. 3A) and washed several times at 65°C. The filters were exposed for 12 h using an Imaging Plate Type BAS-IIIS (Fuji Photo Film, Japan) and visualized with a FUJIX Bio Imaging Analyzer BAS-1500 (Fuji Photo Film, Japan). After removing the probe, the filter was rehybridized with ^{32}P-labeled probe 2 or probe 1.

Results

Cloning and sequencing of DgK

A charomid 9-36 genomic DNA library derived from *D. immitis* female adult worms was screened using the ^{32}P-labeled *Eco* RI-*Sph* I fragment of cD34 as a probe. Five distinctively hybridized colonies were identified after the third round of screening. Restriction endonuclease mapping and Southern blot analysis showed that five positive clones contained different segments of the same genomic region (data not shown). The 8267 bp nucleotide sequence of two independent overlapping clones was determined. We designated the gene DgK (fig. 1A).

Structural analysis of DgK

For characterization of the mRNA corresponding to DgK, the cDNA amplification was performed using the RACE method. The first 5′-RACE (GSP1, AP1) and 3′-RACE (GSP2, AP1) reactions using an adapter ligated to the ds cDNA as template gave no specific RACE products. Therefore, an aliquot of the first RACE products were reamplified using specific nested primers. As a result, approximately 0.8 kb and 1.9 kb bands were prominently amplified by nested 5′-RACE (NGSP1, AP2) and 3′-RACE (NGSP2, AP2), respectively (data not shown). After isolation of the fragments, these overlapping RACE products were directly sequenced using appropriate primers which were designed from the nucleotide sequence of DgK. The combined nucleotide sequence of the 5′- and 3′-RACE products revealed that the cDNA was 2363 bp in length and contained seven nucleotides at the 5′ end which were not detected in DgK (fig. 1B). The cDNA corresponding to DgK was composed of an open reading frame encoding 353 amino acids followed by a long 3′ non-coding region of 1278 bp and contained five potential target for N-glycosylation(N-X-T/S). Comparison of the nucleotide sequence between cDNA and DgK revealed that DgK comprised nine exons separated by eight introns. The exonic regions were conserved between DgK and Dg2 with a 91% sequence identity. However, their introns differed in number, location, size and sequence (fig. 2A). Also, more than 80% amino acids was identical between DgK and Dg2 (fig. 2B).

Detection of DgK in mf and female or male adult worms

Nucleotide sequence analysis revealed that DgK is clearly different from Dg2 in its structure. To investigate whether Dg2 and DgK form the multigene family, Southern blot analysis of genomic DNA was carried out using probes 1 and 2 (fig. 3A). The results indicated that all prominent bands originated in DgK when the genomic DNA containing DgK was analysed using probe 1 (fig. 3B-1). Also, all the bands were derived from Dg2 in the

https://doi.org/10.1017/S0022149X0070054X Published online by Cambridge University Press

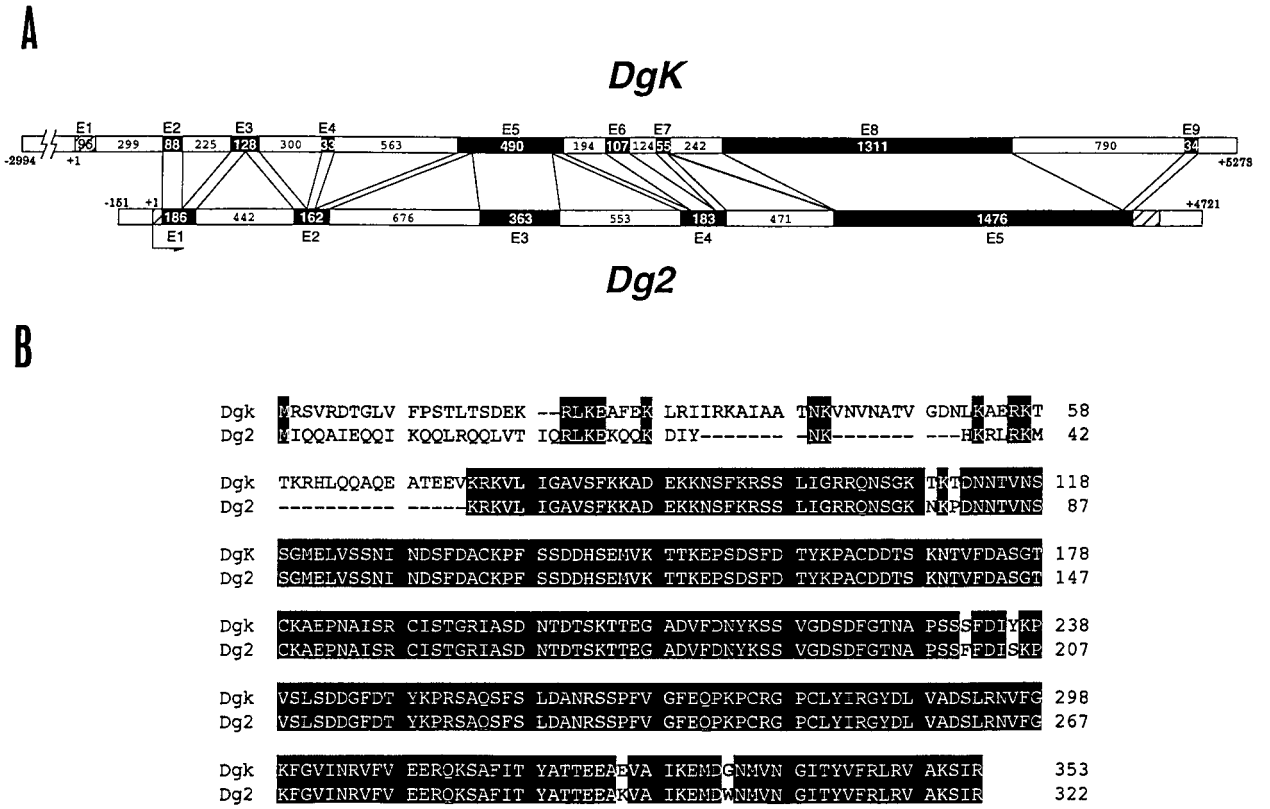


Fig. 2. A. Comparison of gene structure between DgK and Dg2. E1-E9 are exons, and inner numbers indicate the size (bp) of exons and introns. Black and striped boxes indicate highly conserved regions and not highly conserved regions between the exonic sequences of DgK and Dg2, respectively. Corresponding positions of exons are linked with lines. Position+1 of Dg2 indicates the transcription initiation site (Sun & Sugane, 1994). B. Comparison of amino acid sequence between DgK and Dg2. Identical amino acids are indicated in black boxes. Gaps(–) were inserted to maximize the homology.

analysis of the genomic DNA containing Dg2 (fig. 3B-2). Probes 1 and 2 cross-hybridized to Dg2 and DgK, respectively because of their structure (data not shown). The hybridization patterns were not incompatible with the restriction maps of DgK and Dg2 (fig. 3A). Furthermore, no detectable difference in hybridization pattern was observed between mf and male or female adult worms (fig. 3B-1, 3B-2).

Discussion

In order to elucidate the mechanism of stage-specific expression of Dg2, which encodes an immunodominant antigen of *D. immitis*, a genomic clone containing a longer upstream region compared to Dg2 was sought by screening a newly constructed genomic DNA library using ^{32}P-cD34 as a probe. As a result, DgK was cloned and its nucleotide sequence was determined. A comparison of the nucleotide sequences of DgK and Dg2 indicated that DgK is clearly different from Dg2 in structure. Therefore, the RACE method was used to confirm the existence of mRNA corresponding to DgK. The combined analysis of the nucleotide sequence of 5′- and 3′-RACE products revealed that the cDNA was composed of an open reading frame encoding 353 amino acids with a predicted molecular weight of 39.0 kDa with five potential targets for N-glycosylation (N-X-T/S). The region downstream of the 74th amino acid residue was almost identical in its amino acid sequence to cD34 (fig. 1A, shaded bar). Also, the amplified cDNA derived from DgK and cD34 had structural similarities; e.g. the length and long 3′ non-coding region suggested that the deduced proteins encoded by these two cDNAs may have a similar antigenicity and function. Determination of the transcription initiation site is very difficult because the transcripts from the DgK are *trans*-spliced during maturation of mRNA (Krause, 1995). PCR products that reflect the existence of pre-mature mRNA before the acquisition of the SL sequence were not detected by the RACE method. Since the nucleotide sequence, GTTTGAG at the 5′ end corresponds to the 3′ end of the nematode-specific spliced leader (SL) sequence of 22 nucleotides (Donelson & Zeng, 1990), the transcripts from DgK may be *trans*-spliced during pre-mRNA maturation. DNA repeated units containing the genes for 5S RNA and SL sequence were found in the genome of *D. immitis* (Zeng *et al*., 1990), and several mRNAs of *D. immitis* have an SL sequence at its 5′-end (Kuramochi *et al*., 1995; Ma *et al*., 1996). These results indicate that considerable mRNAs of *D. immitis* are *trans*-spliced like other nematodes. Comparison of the

https://doi.org/10.1017/S0022149X0070054X Published online by Cambridge University Press

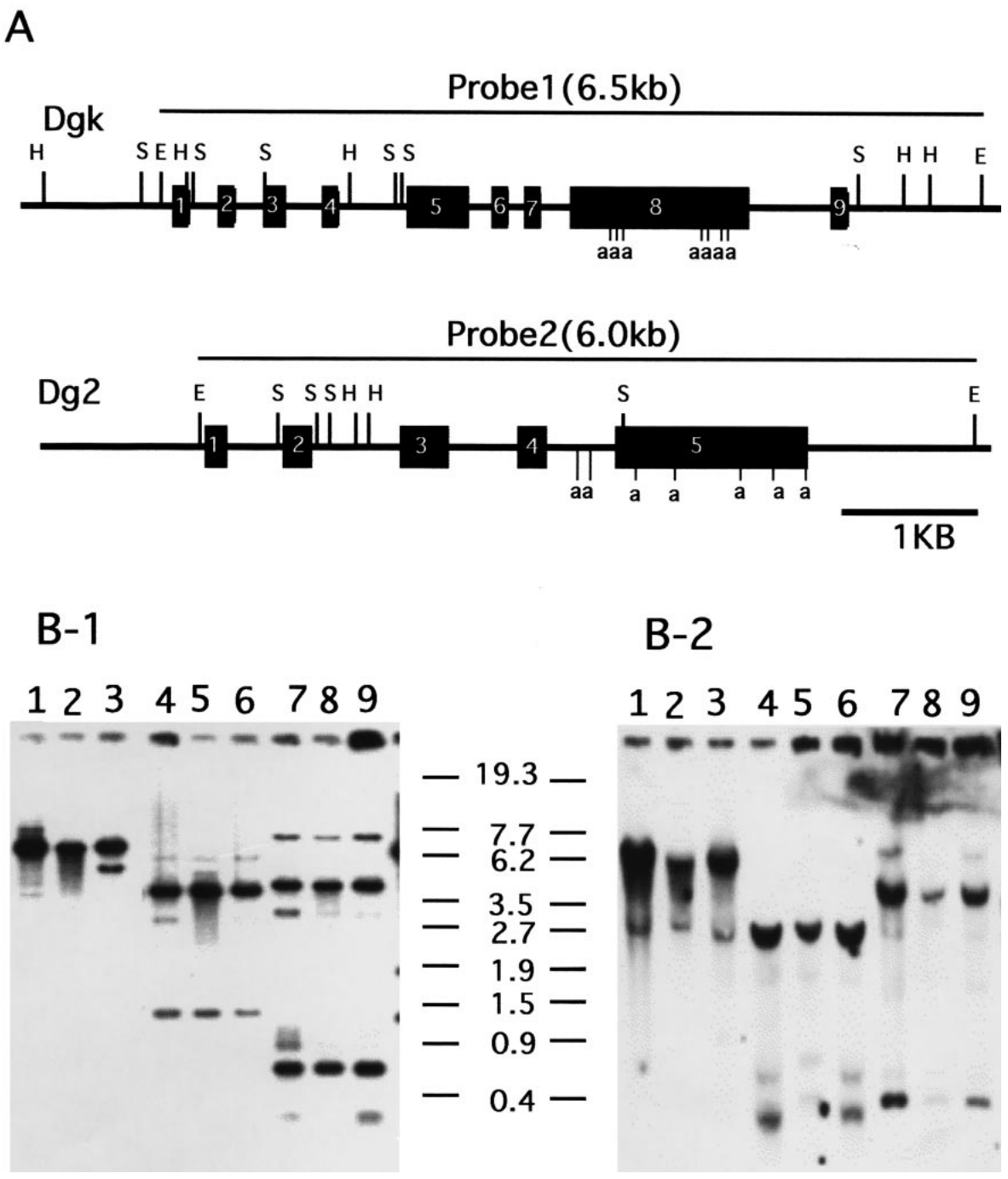


Fig. 3. Southern blot analysis of *Dirofilaria immitis*. A, Partial restriction maps of the DgK and Dg2 and the position of probes 1 and 2. E: *Eco* RI, H: *Hind* III, S: *Ssp* I, a: *Hae* III. Two micrograms of genomic DNA from adult females (lanes 1, 4, 7), adult males (lanes 2, 5, 8) and microfilariae (lanes 3, 6, 9) was digested with restriction endonuclease, *Eco* RI (lanes 1–3), *Hind* III (lanes 4–6) or *Hae* III (lanes 7–9) in fig. 3B-1 and *Eco* RI (lanes 1–3), *Eco* RI + *Ssp* I (lanes 4–6) and *Hae* III (lanes 7–9) in fig. 3B-2, separated by 0.9% agarose gel electrophoresis, blotted, and hybridized with ^{32}P-labelled probes as described in fig. 3A. Figures (kb) indicate the position of markers.

sequence between DgK and amplified cDNA indicated that the common sequence of A^A/Ggtaag^t/$_c^t$/$_a$···ttttcag^A/G at the splice junction of DgK conforms to the consensus sequence of AGgtaagtt···ttttcag^A/G found in *cis*-splicing of *Caenorhabditis elegans* (Krause, 1995). The similarity of the splice site sequence is found in the genes of other parasitic nematodes such as *Onchocerca volvulus* (Erondu & Donelson, 1990). Also, *trans*-splicing is common in the

https://doi.org/10.1017/S0022149X0070054X Published online by Cambridge University Press

gene of both free-living and parasitic nematodes such as *C. elegans* and *Toxocara canis* (Donelson & Zeng, 1990; Zeng *et al*., 1990; Nilsen, 1993; Gems *et al*., 1995). These results suggest that basic transcriptional systems of nematode genes may have been conserved during the process of evolution.

DgK is not a pseudogene, because its cDNA is synthesized. However, DgK is clearly different from Dg2 in structure. To clarify the possibility that Dg2 and DgK form a multigene family, a Southern blot analysis was carried out (fig. 3B-1, 3B-2). As a result, all prominent bands originated in either DgK or Dg2 and the same hybridization patterns were observed using probe 2 instead of probe 1, and vice versa. Five samples of *D. immitis* were examined from five individual dogs. Three samples showed the DgK type, two samples showed the Dg2 type and there was no sample in which both types were mixed. The gene is therefore considered to be highly polymorphic and does not form a multigene family in the worms. Further experiments using worm samples derived from many infected dogs are necessary to clarify whether Dg2 and DgK exist in worms as a single gene or a multigene family. In addition, although different hybridization patterns in the Southern blot analysis among mf, male and female adult worms were observed in our previous study (Sun *et al*., 1992), no differences were detectable in the current study (fig. 3B). Therefore, DgK may not be rearranged during the development of *D. immitis*.

A comparison of the structure of the genes between DgK and Dg2 indicated that the introns differed in number, location, size and sequence between them, while the exonic regions were conserved with more than a 91% sequence homology. The striking difference in the gene structure of Dg2 and DgK may be due to the unique evolutional process of the genes, since they had no homology in the nucleotide sequence with other genes. Therefore, it is necessary to examine a larger number of samples of the parasites from dogs in geographically different areas.

References

Ando, K., Mitsuhashi, J. & Kitamura, S. (1980) Uptake of amino acids and glucose by microfilariae of *Dirofilaria immitis* in vitro. *American Journal of Tropical Medicine and Hygiene* **29,** 213–216.

Chomczynski, P. & Sacchi, N. (1987) Single-step methods of RNA isolation by acid guanidinium thiocyanate–phenol–chloroform extraction. *Analytical Biochemistry* **162,** 156–159.

Culpepper, J., Grieve, R.B., Friedman, L., Mika-Grieve, M., Frank, G.R. & Dale, B. (1992) Molecular characterization of a *Dirofilaria immitis* cDNA encoding a highly immunoreactive antigen. *Molecular and Biochemical Parasitology* **54,** 51–62.

Donelson, J.E. & Zeng, W. (1990) A comparison of trans-RNA splicing in trypanosomes and nematodes. *Parasitology Today* **6,** 327–334.

Erondu, N.E. & Donelson, J.E. (1990) Characterization of a myosin-like antigen from *Onchocerca volvulus*. *Molecular and Biochemical Parasitology* **40,** 213–224.

Frank, G.R., Tripp, C.A. & Grieve, R.B. (1996) Molecular cloning of a developmentally regulated protein isolated from excretory–secretory products of larval *Dirofilaria immitis*. *Molecular and Biochemical Parasitology* **75,** 231–240.

Gems, D., Ferguson, C.J., Robertson, B.D., Nieves, R., Page, A.P., Blaxter, M.L. & Maizels, R.M. (1995) An abundant, *trans*-spliced mRNA from *Toxocara canis* infective larvae encodes a 26-kDa protein with homology to phosphatidylethanolamine-binding proteins. *Journal of Biological Chemistry* **270,** 18517–18522.

Grandea, A.G. III, Tuyen, L.K., Asikin, N., Davis, T.B., Philipp, M., Cohen, C. & McReynolds, L.A. (1989) A λgt 11 cDNA recombinant that encodes *Dirofilaria immitis* paramyosin. *Molecular and Biochemical Parasitology* **35,** 31–41.

Hong, X.Q., Santiago Mejia, J., Kumar, S., Perler, F.B. & Carlow, C.K. (1996) Cloning and expression of DiT33 from *Dirofilaria immitis*: a specific and early marker of heartworm infection. *Parasitology* **112,** 331–338.

Krause, M. (1995) Transcription and translation. *Methods in Cell Biology* **48,** 483–512.

Kuramochi, T., Kita, K., Takamiya, S., Kojima, S. & Hayasaki, M. (1995) Comparative study and cDNA cloning of the flavoprotein subunit of mitochondrial complex II (succinate-ubiquinone oxidoreductase: fumarate reductase) from the dog heartworm, *Dirofilaria immitis*. *Comparative Biochemistry and Physiology* **111,** 491–502.

Ma, D., Hong, X., Raghavan, N., Scott, A.L., McCarthy, J.S., Nutman, T.B., Williams, S.A. & Carlow, C.K.S. (1996) A cyclosporin A-sensitive small molecular weight cyclophilin of filarial parasites. *Molecular and Biochemical Parasitology* **79,** 235–242.

McReynolds, L.A., DeSimone, S.M. & Williams, S.A. (1986) Cloning and comparison of repeated DNA sequences from the human filarial parasite *Brugia malayi* and the animal parasite *Brugia pahangi*. *Proceedings of the National Academy of Sciences of the United States of America* **83,** 797–801.

Nilsen, T.W. (1993) Trans-splicing of nematode pre-messenger RNA. *Annual Review of Microbiology* **47,** 413–440.

Poole, C.B., Grandea, A.G. III, Maina, C.V., Jenkins, R.E., Selkirk, M.E . & McReynolds, L.A. (1992) Cloning of a cuticular antigen that contains multiple tandem repeats from the filarial parasite *Dirofilaria immitis*. *Proceedings of the National Academy of Sciences of the United States of America* **89,** 5986–5990.

Sambrook, J., Fritsch, E.F. & Maniatis, T. (1989) *Molecular cloning. A laboratory manual*. Cold Spring Harbor Laboratory, New York.

Sanger, F., Nicklens, S. & Coulson, A.R. (1977) DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74,** 5463–5467.

Sugane, K. & Sun, S. (1994) Detection of anti-helminth antibody by microenzyme-linked immunosorbent assay using recombinant antigen and anti-β-galactosidase monoclonal antibody. *Journal of Immunological Methods* **168,** 55–60.

Sun, S. & Sugane, K. (1992) Immunodiagnosis of human dirofilariasis by enzyme linked immunosorbent assay using recombinant DNA-derived fusion protein. *Journal of Helminthology* **66,** 220–226.

Sun, S. & Sugane, K. (1994) Complete sturucture of the gene encoding an immunodominant antigen of *Dirofilaria*

immitis and larva-specific synthesis of primary transcript. *Journal of Helminthology* **68,** 259–264.

Sun, S., Matsuura, T. & Sugane, K. (1991) Molecular cloning of the cDNA encoding an immunodominant antigen of *Dirofilaria immitis*. *Journal of Helminthology* **65,** 149–158.

Sun, S., Matsuura, T. & Sugane, K. (1992) Stage-specific expression of a developmentally regulated gene in *Dirofilaria immitis*. *Journal of Helminthology* **66,** 62–67.

Zeng, W., Alarcon, C.M. & Donelson, J.E. (1990) Many transcribed regions of the *Onchocerca volvulus* genome contain the spliced leader sequence of *Caenorhabditis elegans*. *Molecular and Cellular Biology* **10,** 2765–2773.

(Accepted 14 May 1999)