Original Article

Sequence and phylogenetic analysis of the *gp200* protein of *Ehrlichia canis* from dogs in Taiwan

Chia-Chia Huang, Yu-Chen Hsieh, Chau-Loong Tsang, Yang-Tsung Chung[*]

Department of Veterinary Medicine, College of Veterinary Medicine, National Chung Hsing University, Taichung 402, Taiwan

Ehrlichia (*E.*) *canis* is a Gram-negative obligate intracellular bacterium responsible for canine monocytic ehrlichiosis. Currently, the genetic diversity of *E. canis* strains worldwide is poorly defined. In the present study, sequence analysis of the nearly full-length 16S rDNA (1,620 bp) and the complete coding region (4,269 bp) of the *gp200* gene, which encodes the largest major immunoreactive protein in *E. canis*, from 17 Taiwanese samples was conducted. The resultant 16S rDNA sequences were found to be identical to each other and have very high homology (99.4～100%) with previously reported *E. canis* sequences. Additionally, phylogenetic analysis of *gp200* demonstrated that the *E. canis* Taiwanese genotype was genetically distinct from other reported isolates obtained from the United States, Brazil, and Israel, and that it formed a separate clade. Remarkable variations unique to the Taiwanese genotype were found throughout the deduced amino acid sequence of *gp200*, including 15 substitutions occurring in two of five known species-specific epitopes. The *gp200* amino acid sequences of the Taiwanese genotype bore 94.4～94.6 identities with those of the isolates from the United States and Brazil, and 93.7% homology with that of the Israeli isolate. Taken together, these results suggest that the Taiwanese genotype represents a novel strain of *E. canis* that has not yet been characterized.

Keywords: canine ehrlichiosis, *Ehrlichia canis*, *gp200* gene, phylogenetic analysis, sequence analysis

Introduction

Ehrlichia (*E.*) *canis*, which is the primary etiologic agent of canine monocytic ehrlichiosis (CME), is a Gram-negative obligate intracellular bacterium that replicates in monocytes and macrophages [5]. *E. canis* is primarily transmitted by the brown dog tick *Rhipicephalus sanguineus* [2]. CME has been reported throughout the world, with a higher frequency in tropical and subtropical regions [12]. Dogs infected with *E. canis* can present with a wide spectrum of symptoms, ranging from subclinical infection to death [1,6,11]. Clinical signs often include anorexia, dyspnea, hyperthermia, lethargy, weight loss, and bleeding disorders [3,4,22].

Although the genome of the *E. canis* prototype strain Jake was completely sequenced in 2006 [8], the genetic diversity of *E. canis* strains worldwide is still poorly defined. To date, most molecular epidemiology studies of *E. canis* have focused on the 16S ribosomal RNA gene (16S rDNA), while much less is known about the other genes. Unfortunately, molecular characterization of the 16S rDNA has provided little information regarding strain diversity and suggests a high level of conservation [16,20, 21]. It has been reported that the 16S rDNA partial sequences were 99.9～100% identical among isolates of *E. canis* from South America (VHE and VDE isolates), North America (Oklahoma isolate), Asia (Thai, Chinese, and Japanese Kagoshima isolates), and the Middle East (Turkish and Israeli isolates) [15,23]. These observations indicated that the 16S rDNA sequence might not be the best target for evaluation of the genetic diversity of *E. canis*.

Zhang *et al.* [24] recently analyzed and compared the *gp200* gene sequences of *E. canis* isolates from the United States, Brazil and Israel and found that *gp200* was highly conserved between the isolates from the USA and Brazil, but that substantial diversity was present in the Israeli isolate. The *gp200* gene encodes the largest major immunoreactive protein that has been identified in *E. canis* to date [10,14]. The *gp200* protein has been shown to be especially sensitive to immunodiagnositic antigen for *E. canis* infections and to provide species specificity [9]. Investigation of the extent of sequence variation in this antigen candidate may help elucidate the diversity of *E.*

*Corresponding author
Tel: +886-4-22840751; Fax: +886-4-22862073
E-mail: ytchung@dragon.nchu.edu.tw

canis isolates and pave the way for development of an effective vaccine for CME. Herein, we analyzed the sequence variations in the *gp200* gene of the *E. canis* parasite from naturally infected dogs in Taiwan. To define the genotype of the organism, the 16S rDNA was also analyzed.

Materials and Methods

Blood samples and microscopic examination

Eighty-seven blood samples were collected from dogs exhibiting clinical signs compatible with ehrlichiosis that presented to the Veterinary Teaching Hospital of National Chung Hsing University and eight veterinary clinics throughout Taiwan during 2009. Blood was collected into sterile tubes containing anticoagulant (EDTA) and then stored at 4°C until arrival at the laboratory for further processing. Giemsa-stained blood smears were examined by light microscopy (by observing 100 microscopic fields at 1,000× magnification) for the presence of morulae or inclusion bodies of *E. canis*. Genomic DNA was extracted from 200 µL of whole blood samples using a QIAamp DNA Blood Mini Kit (Qiagen, USA) according to the manufacturer's protocol. Nucleic acid was eluted into 100 µL of elution buffer and stored at −20°C until further use.

PCR amplification

The oligonucleotide primers used for the amplification of *E. canis* genes (16S rDNA and *gp200*) were designed using primer design software (PrimerSelect; DNAStar, USA) and information from the *E. canis* genome (GenBank accession number CP000107) [8]. All primers used in this study are shown in Table 1. For each PCR amplification, 5 µL of extracted DNA was used as the template in a 25 µL reaction mixture containing 1× PCR buffer (Promega, USA), 2.5 mM MgCl$_2$, 0.2 mM deoxynucleotide

triphosphates (dNTPs), 0.8 µM of each primer and 0.125U *Taq*Bead Hot Start Polymerase (Promega, USA). The reactions were conducted in an ABI2700 thermocycler (Applied Biosystems, USA) according to the following parameters: 94°C for 3 min followed by 35 cycles of 94°C for 1 min, 58°C for 1 min and 72°C for 1.5 min, followed by an extension step at 72°C for 5 min.

Cloning and sequencing

The resulting PCR products were electrophoresed on a 1.2% agarose gel that was stained with ethidium bromide to check the size of the amplified fragments by comparison to a DNA molecular weight marker (100-bp DNA Ladder; Promega, USA). In each case, the single amplified product of the expected size was column purified using the QIAquick PCR Purification Kit (Qiagen, USA) and then ligated into the pGEM-T vector (Promega, USA) for subsequent transformation in *Escherichia coli* DH5α competent cells. One plasmid vector containing the insert was purified from each clone using the QIAprep Spin Miniprep Kit (Qiagen, USA), after which it was sequenced using an ABI PRISM 3730 capillary sequencer (Applied Biosystems, USA) and the Dye Terminator Cycler Sequencing Kit (Applied Biosystems, USA) with the T7 or SP6 vector primer. Both the sense and antisense strands of each PCR-amplified product were sequenced, and the sequences were then manually edited to resolve ambiguities. A consensus sequence was obtained for each amplified PCR product by comparing both the sense and antisense sequences.

Sequence and phylogenetic analyses

The BLAST program (NCBI, USA) was used for comparison and analysis of sequence data obtained in this study versus those previously deposited in the GenBank database. Multiple sequence alignment was conducted

Table 1. Primers used for amplification and sequencing of the 16S rRNA and *gp200* genes from Taiwanese samples of *Ehrlichia canis*

Target	Primer	Sequence	Expected size
16S rDNA	EC16-F1	5´- AGG CGG CGA TGC ATA ATA GCT -3´	940 bp
	EC16-R2	5´- GGA GTG CTT AAC GCG TTA GCT -3´	
	EC16-F2	5´- GGT AGT CCA CGC TGT AAA CGA -3´	805 bp
	EC16-R1	5´- AGA CGG CGT AGC ACC TTA TCA -3´	
gp200	EC200-F1	5´- TGT CAT TCT AAA GCT TGG TTT ATC CT -3´	1,389 bp
	EC200-R1	5´- GTG GCA AAT GTT CAA CAG CAT G -3´	
	EC200-F2	5´- CAC AGA AGC CAG ATT GTT CTG T -3´	1,075 bp
	EC200-R2	5´- GGA GTA GGA GTT GCT GAA GTC T -3´	
	EC200-F3	5´- GGG ATT CGT TAC ACC TAC CAA GT -3´	1,286 bp
	EC200-R3	5´- GTC CGC AGC ACG ATA ATC AGA TA -3´	
	EC200-F4	5´- ACA TCC AAG ATA CGC AAA GGA GA -3´	1,292 bp
	EC200-R4	5´- ACC TGC ACC TCA TAT CCA ACT -3´	

using Clustal W version 1.8 [18]. Phylogenetic trees were inferred using the neighbor-joining method as implemented by the MEGA software version 4 [17]. The distance matrix of amino acid divergences was calculated according to Kimura's two-parameter model furnished by MEGA. A bootstrap resampling technique of 1,000 replications was conducted to statistically support the reliabilities of the nodes on the trees.

Results

The morulae or inclusion bodies of *E. canis* were detected by microscopic examination in 17 out of 87 blood smears (19.6%) of dogs exhibiting clinical signs compatible with ehrlichiosis (Fig. 1). Positive blood samples from these parasitemic dogs were subjected to DNA extraction, amplification and sequencing. All 16S rDNA sequences of 1,620 bp derived from 17 samples were found to be identical to each other. The representative sequence obtained from this study has been deposited in the GenBank database under accession number GU810149. The sequence had 100% identity with the corresponding sequences from *E. canis* isolates in Brazil (GenBank accession numbers EF195134), Greece (EF011110-1), Italy (EU439944), Turkey (AY621071), Venezuela (AF373612-3), and Thailand (EU263991 and EF139458), 99.9% homology with a Japanese isolate (AF536827) and a strain from the USA (M73221), 99.8% identity with a Chinese isolate (AF162860), 99.7% identity with a South American isolate (DQ915970), 99.7% homology with an Israeli isolate (U26740), and 99.4% identity with a Spanish isolate (AY394465). A phylogenetic tree of *E. canis* was inferred based on the 16S rDNA sequences obtained in this study and others (Fig. 2). Consistent with the results of previous studies [15,16,21], the *E. canis* 16S rDNA sequences generated in this study were very similar to those from other countries.

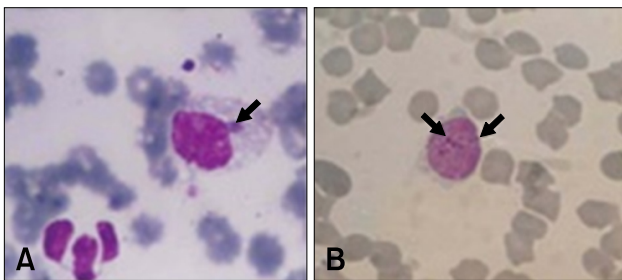Nevertheless, there was substantial variation in the *gp200*

gene sequences between the Taiwanese samples and other previously reported isolates. According to the *gp200* nucleotide sequences of the *E. canis* isolates from the United States (AF252298 and CP000107), the *gp200* open reading frame was 4,266 bp in length and encoded a protein of 1,421 amino acids. The complete *gp200* coding sequences derived from 17 Taiwanese samples were identical or nearly identical to each other (99.9~100% identities), even though the canine blood samples were obtained from different locations. The three representative sequences obtained from this study have been deposited in the GenBank database under accession numbers GU810148 ($n = 8$), HM067841 ($n = 4$) and HM067842 ($n = 5$). The only two nucleotide differences among these sequences occurred at positions 3591 and 3624. However, these variations did not cause any changes in amino acids. Notably, all of these sequences had three nucleotide insertions at positions between 4077 and 4078, 4081 and 4082, and 4082 and 4083 (based on the *E. canis* strain AF252298 numbering system) and thus consisted of 4,269 bp.

Comparative sequence analysis of the 4,269-bp products generated for the *E. canis* Taiwanese samples (TWN; provisionally designated as the TWN genotype) revealed that it had 97.9~98.0% homology with the corresponding sequences obtained from the USA Jake2 isolate (CP000107) and the Brazilian Sao Paulo isolate (EF636664), 97.8~97.9% homology with the sequence of the USA Jake1 isolate (AF252298), and 97.5~97.6% homology with the sequence of the Israeli 611 isolate (EF636665).
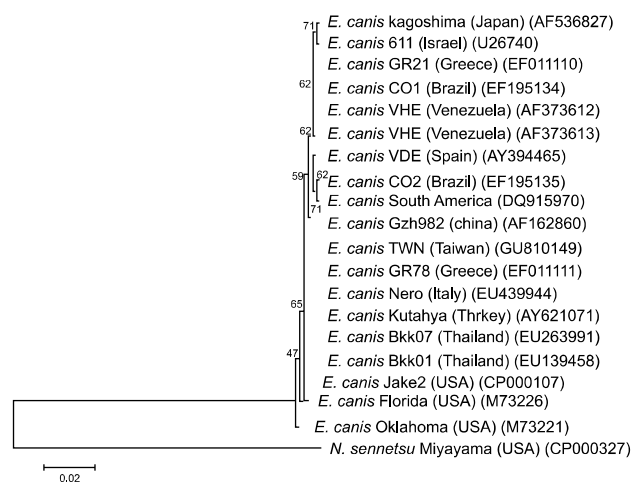
Fig. 2. Phylogenetic tree based on the partial 16S rDNA sequences of *E. canis* isolates. To root the tree, the corresponding sequence of *Neorickettsia* (*N.*) *sennetsu* was used as an outgroup. Accession numbers for *E. canis* isolates and the outgroup species *N. sennetsu* are given in parentheses. The scale bar indicates the number of substitutions per nucleotide position. The numbers at the nodes represent the percentage of 1,000 bootstrap resamplings.

Fig. 1. Photomicrographs of the Giemsa-stained thin blood smear showing (A) *Ehrlichia* (*E.*) *canis* morula and (B) *E. canis* inclusion bodies.

```
US1  MSDPKQGDPEQNQTNPSGDIQDQSQQDQQEQDQQQGAVGGAVGNSPIERERVAAPESEDLYTVIIPKGKRTAAPILERKS 080
US2  MSDPKQGDPEQNQTNPSGDIQDQSQQDQQEQDQQQGAVGGAVGNSPIERERVAAPESEDLYTVIIPKGKRTAAPILERKS 080
BRZ  MSDPKQGDPEQNQTNPSGDIQDQSQQDQQEQDQQQGAVGGAVGNSPIERERVAAPESEDLYTVIIPKGKRTAAPILERKS 080
ISR  MSDPKQGDPEQNQTNPSGDIQDQSQQDQQEQDQQQGAVGGAVGNSPVERERVAAPESEDLYTVIIPKGKRTAAPILERKS 080
TWN  MSDPKQGDPEQNQTNPSGDIQDQSQQDQQEQDQQQGAVGGAVGNSPIERERVAAPESEDLYTVIIPKGKRTAAPILERKS 080

US1  PTPEPKVEDDEDLPPTLPPRTFSGEGYDDVGVSMPTVSRGIYQPPIVQDSNLYSSIGGVPQEAQYDAAARAGGPRKFLYG 160
US2  PTPEPKVEDDEDLPPTLPPRTFSGEGYDDVGVSMPTVSRGIYQPPIVQDSNLYSSIGGVPQEAQYDAAARAGGPRKFLYG 160
BRZ  PTPEPKVEDDEDLPPTLPPRTFSGEGYDDVGVSMPIVSRGIYQPPIVQDSNLYSSIGGVPQEAQYDAAARAGGPRKFLYG 160
ISR  PTPEPKVEDDEDLPPTLPPRTFSGEGYDDVGVSMPTVSHGIYQPPIVQDSNLYSSIGGVPQEAQYDAAARAGGPRKFLYG 160
TWN  PTPEPKVEDDEDLPPTLPPRTFSGEGYDDVGVSMPTVSSGIYQPPIVQDSNLYSSIGGVPQEAQYDAAARAGEPRRFLYG 160

US1  PYTFSNGQEIMDFEFDTPWPDVRNAVLGNKEIKEEWLTTSGPVRDIADRIVASKGDLSEDQVEEILDIIFMNESEIAEGI 240
US2  PYTFSNGQEIMDFEFDTPWPDVRNAVLGNKEIKEEWLTTSGPVRDIADRIVASKGDLSEDQVEEILDIIFMNESEIAEGI 240
BRZ  PYTFSNGQEIMDFEFDTPWPDVRNAVLGNKEIKEEWLTTSGPVRDIADRIVASKGDLSEDQVEEILDIIFMNESEIAEGI 240
ISR  PYTFSNGQEIMDFEFDTPWPDVRNAVLGNKEIKEEWLTTSGPVRDIADRIVASKGDLSEDQVEEILDIIFMNESEIAEGI 240
TWN  PYTFSNGQEIMDFEFDTPWPDVRNAVLGNKEIKEEWLTTSGPVRDIADRIVASKGNLPEDQVEEILDIIYMNESEIAKGI 240

US1  SNPLHADVDNNPVKGAKNVMTLMHLVYACDVDPRIVKALGEVENDEGDLGANAYNVLDSEGNLPLHHAAKNCTGDKLKLC 320
US2  SNPLHADVDNNPVKGAKNVMTLMHLVYACDVDPRIVKALGEVENDEGDLGANAYNVLDSEGNLPLHHAAKNCTGDKLKLC 320
BRZ  SNPLHADVDNNPVKGAKNVMTLMHLVYACDVDPRIVKALGEVENDEGDLGANAYNVLDSEGNLPLHHAAKNCTGDKLKLC 320
ISR  SNPLHADVDNNPVKGAKNVMTLMHLVYACDVDPRIIKSLGEVENDEGDLGANAYNVLDSEGNLPLHHAAKNCTGDKLKLC 320
TWN  SNPLHADVDNNPVKGAKNVMTLMHLVYACDTDPRIIKALGEVENDEGDLGANAYNVLDSEGNLPLHHAAKNCTGDKLKLC 320

US1  MEKTKTDFIDTANFANQSPLHIITQKPDCSVLDIEEFTSRNLDFGLVDGDGKNPLHHAVEHLPPVILKGVMDHVKNSSEF 400
US2  MEKTKTDFIDTANFANQSPLHIITQKPDCSVLDIEEFTSRNLDFGLVDGDGKNPLHHAVEHLPPVILKGVMDHVKNSSEF 400
BRZ  MEKTKTDFIDTANFANQSPLHIITQKPDCSVLDIEEFTSRNLDFGLVDGDGKNPLHHAVEHLPPVILKGVMDHVKNSSEF 400
ISR  MEKTKTDFIDTANFANQSPLHIITQKPDCSALDIEEFTRRDLDFGLVDGDGKNPLHHAVEHLPPVILKGVMDCVKNSSEF 400
TWN  MEKTKSDFIDTANFDNQSPLHIITRKPDCSALDIEEFTSRNLDFGLVDGDGKNPLHHAVEHLPPVILKGVMDYVKSSTEF 400

US1  QDLVNDPDYFGNTIAHYAVKNKNADLTLFNMLKASGADLNVRNVVGRAPIHVASSNGKANAVSGLVSCGIDVNSQDVNGD 480
US2  QDLVNDPDYFGNTIAHYAVKNKNADLTLFNMLKASGADLNVRNVVGRAPIHVASSNGKANAVSGLVSCGIDVNSQDVNGD 480
BRZ  QNLVNDPDYFGNTIAHYAVKNKNADLTLFNMLKASGADLNVRNVVGRAPIHVASSNGKANAVSGLVSCGIDVNSQDVNGD 480
ISR  QDLVNDPDYLGNTIAHHAVKNQNADLTLFNMLKASGADLNVRNVVGRAPIHVASSNGNANAVAGLVSCGIDVNSQDVNGD 480
TWN  QDLVNDPDYLGNTIAHHAVKNQNADLTLFNMLKASGADLSVRNVVGRAPIHVASSDGKTNAVAGLVSCGIDVNSQDVNGD 480

US1  TPLHIAVEGGSMETVLAVLNQRGADVSVQNNDGVTPMLSAAKYGDIGVIKALGSAKPNIKGEDTVAKSLLMEDYKGFTPL 560
US2  TPLHIAVEGGSMETVLAVLNQRGADVSVQNNDGVTPMLSAAKYGDIGVIKALGSAKPNIKGEDTVAKSLLMEDYKGFTPL 560
BRZ  TPLHIAVEGGSMETVLAVLNQRGADVSVQNNDGVTPMLSAAKYGDIGVIKALGSAKPNIKGEDTVAKSLLMEDYKGFTPL 560
ISR  TPLHIAVEGGNMKTVLAVLNQRGADVSVQNNDGVTPMLSAAKYGDIGVIKALGSAKPNIKGEDTVAKSLLMEDYKGFTPL 560
TWN  TPLHIAVEGGNMETVLAVLNQRGADVSVQNNDGVTPMLSAAKYGDIGVIKALGSAKPNIKGEDTVAKSLLMEDYKGFTPL 560

US1  HFVAGGGSRDTFRVVRKNYEKCHDLATIRAALMQDRSGGELVNLGDFESENILGSPNAKFLQHIQSANFGFSPARRGIVS 640
US2  HFVAGGGSRDTFRVVRKNYEKCHDLATIRAALMQDRSGGELVNLGDFESENILGSPNAKFLQHIQSANFGFSPAHCAIVS 640
BRZ  HFVAGGGSRDTFRVVRKNYEKCHDLATIRAALMQDRSGGELVNLGDFESENILGSPNAKFLQHIQSANFGFSPAHCAIVS 640
ISR  HFVAGGGSRDTFRVVRKNYEKYHDLATIRAALTQDRSGGEIVNLGDFESENILGSPNAKFLQHIQSANFGFSPAHCAIVS 640
TWN  HFVAGGGSRDTFRVVRKNYEKYHDLATIGAALMQNRSGGELVNLGDFESENILGSPNAKFLQHIQSANFGFSPAHCAIVS 640

US1  SNHNVMKDILNFVGDSLHLPSERGYNAMQVAALFGDKEAVKMLAKSAKPSDLNFKTSATPTPLNLACLRGDNEVVRGLVG 720
US2  SNHNVMKDILNFVGDSLHLPSERGYNAMQVAALFGDKEAVKMLAKSAKPSDLNFKTSATPTPLNLACLRGDNEVVRGLVG 720
BRZ  SNHNVMKDILNFVGDSLHLPSERGYNAMQVAALFGDKEAVKMLAKSAKPSDLNFKTSATPTPLNLACLRGDNEVVRGLVG 720
ISR  SNHNVIKDILNFVGDSLHLPSERGYNAMQVAALFGDKEAVKMLAKSAKPSDLNFKTSATPTLLNLACLRGDNEVVRGLVG 720
TWN  SNHNVIKDILNFVGDSLHLPSERGYNAMQVAALFGDKEAVKMLAKSAKPSDLNFKTSATLTPLNLACLRGDNEVVRGLVG 720

US1  QHGIDINQRMGSDKNTVLHYAISKGDSFLVQKILAHTGVDVNCENNLGQTPLHLAVEGGDPKIVSSLLKAGAVVNRLDDN 800
US1  QHGIDINQRMGSDKNTVLHYAISKGDSFLVQKILAHTGVDVNCENNLGQTPLHLAVEGGDPKIVSSLLKAGAVVNRLDDN 800
BRZ  QHGIDINQRMGSDKNTVLHYAISKGDSFLVQKILAHTGVDVNCENNLGQTPLHLAVEGGDPKIVSSLLKAGAVVNRLDDN 800
ISR  QHGIDINQRMGSDKNTVLHYAISKGDSSLVRKILAHPEVDVNCENNLGQTPLHLAVEGGDPKIVSSLLKAGAVVNRLDDN 800
TWN  QHGIDINQRMGSDKNTVLHYAISKGDSSLVRKILAHTGVDVNCENDLGQTPLHLAVEGGDPKIVSSLLKVGAVVNRLDDN 800
```

Fig. 3. Alignment of the deduced amino acid sequences of *E. canis gp200*. Amino acids highlighted in grey represent residues divergent from the USA Jake1 isolate (US1) sequence, while dashes represent gaps. The underlined regions indicate known dominant species-specific antibody epitopes. The GenBank accession number for each sequence is given at the end of the sequence. Abbreviations of specific *E. canis* strains - US2: USA Jake2 isolate, BRZ: Brazilian Sao Paulo isolate, ISR: Israeli 611 isolate, TWN: Taiwanese sample.

```
US1 GRSVLSSAIVPGRKEKGVLGIVNKLLDRGADINLDGDHNILFDQCLRGGYNNVLDKLIQQGVEVNRNSEIRPMVYAAISG 880
US2 GRSVLSSAIVPGRKEKGVLGIVNKLLDRGADINLDGDHNILFDQCLRGGYNNVLDKLIQQGVEVNRNSEIRPMVYAAISG 880
BRZ GRSVLSSAIVPGRKEKGVLGIVNKLLDRGADINLDGDHNILFDQCLRGGYNNVLDKLIQQGVEVNRNSEIRPMVYAAISG 880
ISR GRSVLSSAIVPGRKEKEVLGIVDKLLNRGADINLDGDHNILLNKCLRGGYNNVLDKLIGQGVEVNRNSEIRPMVYAAISG 880
TWN GRSVLSSAIVPGRKEKGVLGIVNKLLDGGADIDLDGDHNILLDQCIRGGYNNVLDKLIAQGVEVNRNSEIRPMVYAAISG 880

US1 NEHAIKSLANAGGDVNEVVNNPSSRHSGNPLIMVAVADGNAGLLKTLVSEGCDVGKSGKDGNTALHYAVSHSDKEFGNKA 960
US2 NEHAIKSLANAGGDVNEVVNNPSSRHSGNPLIMVAVADGNAGLLKTLVSEGCDVGKSGKDGNTALHYAVSHSDKEFGNKA 960
BRZ NEHAIKSLANAGGDVNEVVNNPSSRHSGNPLIMVAVADGNAGLLKTLVSEGCDVGKSGKDGNTALHYAVSHSDKEFGNKA 960
ISR NKHAIKSLANAGGDVNEVVNNPSSRHSGNPLIMVAVADGNAGLLKTLVSEGCDVGKPGRDGNTALHCAVGHSDKDFGNKA 960
TWN NEHAIKSLANAGGDVNEVVNNPSSRHSGNPLIMVAVADGNAGLLKTLVSEGCDVGKPGRDGNTALHCAVGHSDKNFGNKA 960

US1 IKILISRNSVGTNRDILTQKNNAGDTPLHEALKSGNINSVQNILSAVHPRYAKEILTARDKEGYTPMHYTVGVNNVDVGR 1040
US2 IKILISRNSVGTNRDILTQKNNAGDTPLHEALKSGNINSVQNILSAVHPRYAKEILTARDKEGYTPMHYTVGVNNVDVGR 1040
BRZ IKILISRNSVRTNRDILTQKNNAGDTPLHEALKSGNINSVQNILSAVHPRYAKEILTARDKEGYTPMHYTVGVNNVDVGR 1040
ISR IKILTLRNSVRTNRDILTQKNNAGDTPLHEALKSGNINSVQNILSAVHPRYAKEILTARDKKGYTPVHCAVGVSNVDVGR 1040
TWN IKILTSRNSARTNRDVLTQKNNAGDTPLHEALKSGNINSVQNILSAVHTRYAKEILTARDKEGYTPMHHTVGVNNVDVSR 1040

US1 SILESMLSKGVNNLGEIVGAQDSNFRTPLHAAIKISDYRAADMIIGSLSKTELSKLSQLTDINGDTPLHLSCQSGNVEMT 1120
US2 SILESMLSKGVNNLGEIVGAQDSNFRTPLHAAIKISDYRAADMIIGSLSKTELSKLSQLTDINGDTPLHLSCQSGNVEMT 1120
BRZ SILESMLSKGVNNLGEIVGAQDSNFRTPLHAAIKISDYRAADMIIGSLSKTELSKLSQLTDINGDTPLHLSCQSGNVEMT 1120
ISR SILGSMLSKGVNNLGEIVEQDSNFRTPLHTAIKISDYRSVDMIICSLSKTELSKLSQLTDINGDTPLHLSCQSGNVEMT 1120
TWN SILESMLSKGVNNLGEIVGAQDSNFQTPLHSAIKISDYRAADMMSRLSKTELSKLSQLTDINGDTPLHLSCQSGNVYMT 1120

US1 QFFLGGLDKRELPKTLKIANKNGDTPLHDAIRNDDIKSAKMMIRNCNKEELANVLKCKDSFGNTVLHTIADQVIANPESK 1200
US2 QFFLGGLDKRELPKTLKIANKNGDTPLHDAIRNDDIKSAKMMIRNCNKEELANVLKCKDSFGNTVLHTIADQVIANPESK 1200
BRZ QFFLGGLDKRELPKTLKIANKNGDTPLHDAIRNDDIKSAKMMIRNCNKEELANVLKCKDSFGNTVLHTIADQVIANPESK 1200
ISR QFFLGGLDKRELPKTLKIANKNGDTPLHDAIRNDDTKSAKIMIRNCSKEELANVLKYKDNLGNTVLHAIANQVIANPESK 1200
TWN QFFLGGLDKRELPKTLKIANKNGDTPLHDAIRNDDVKSAKMMIRNCSKEELANVLKYKDSFGNTVLHVIADQVIANPESK 1200

US1 KDLDGLMNLAVKRLKNQDLKDLVNTRNNSDDTVAHCALLSDMKYAQKILKSCNHDTLVRGNSNNQSLSECIRDDSKYKKG 1280
US2 KDLDGLMNLAVKRLKNQDLKDLVNTRNNSDDTVAHCALLSDMKYAQKILKSCNHDTLVRGNSNNQSLSECIRDDSKYKKG 1280
BRZ KDLDGLMNLAVKRLKNQDLKDLVNTRNNSDDTVAHCALLSDMKYAQKILKSCNHDTLVRGNSNNQSLSECICDDSKYKKG 1280
ISR KDLDDLMNLAMKKLKSQDLKDLVNTRNNFEDTVAHCALLSDMKCAQKILKSCERDTLVRKNNSNQSLSECIRDDSQYKKG 1280
TWN KGLDGLMNLAMRRLKSQDLKDLVNTRNNFGDTVAHCALLSDMKCAQKILKSCERDTLVIRNSSNQSLSECIRDHSKYKKG 1280

US1 GIFSKSLFSKLKKLEARAASASYEELSSISSGSDVSSVSTNSTEVSAVPEVARSSGAVSFKHVQETGVDTSGPSDIESL- 1360
US2 GIFSKSLFSKLKKLEARAASASYEELSSISSGSDVSSVSTNSTEVSAVPEVARSSGAVSFKHVQETGVDTSGPSDIESL- 1360
BRZ GIFSKSLFSKLKKLEARAASASYEELSSISSGSDVSSVSTNSTEVSAVPEVARSSGAVSFKHVQETGVDTSGPSDIESLS 1360
ISR GIFKKSLFSKLKKLEARAASASYEELSSISSGSDVSSVSTNSTEVSAVPEMARSSGAVSFKHAQETGVDTSGPSDTES-- 1360
TWN GIFSKSLFSKLKKLEARAASASYEELSSISSGSDVSSVSTNSTEVSAVPEVARSSGAVLFKHVQETGVDTSGPSDTESLS 1360

US1 ERLSDTSLGSNDFDQRMADLDQEIANIVSGLPEVTQVAVSQQQAASPSSGQAAGVQQKEMQR 1421    AF252298
US2 ERLSDTSLGSNDFDQRMADLDQEIANIVSGLPEVTQVAVSQQQAASPSSGQAAGVQQKEMQR 1421    CP000107
BRZ ERLSDTSLGSNDFDQRMADLDQEIANIVSGLPEVTQVAVSQQQAASPSSGQAAGVQQKEMQR 1422    EF636664
ISR ERLSDTSLGSNDFDQRMAELDQEIADIVSGLPQVTQVAVSQQQAASPSSGQTAGVQQKEMQR 1420    EF636665
TWN DGLSDTSLGSNDFDQRMADLDQEIANIVSGLPEVTQVAVSQQQAASPSSGQAAGVQQKEMQR 1422    GU810148
```

Fig. 3. Continued.

The deduced amino acid sequence of *gp200* of the TWN genotype was aligned with the corresponding sequences of four other published isolates. When compared with the sequence of the USA Jake1 isolate, 79 amino acid changes scattered throughout the *gp200* sequence of the TWN genotype were observed (Fig. 3). Of these changes, 15 amino acid residues occurred in two of five known species-specific antibody epitopes [14]. Overall, the TWN
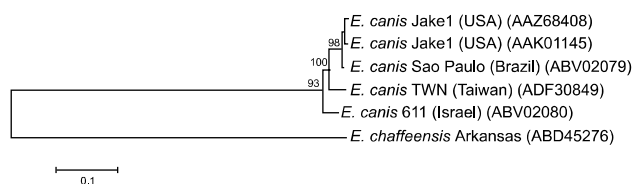
Fig. 4. Phylogenetic tree based on deduced amino acid sequences of the *E. canis gp200*. To root the tree, the sequence of an ortholog (ankyrin protein 200) in *E. chaffeensis* was used as an outgroup. Accession numbers for *E. canis* isolates and the outgroup species *E. chaffeensis* are given in parentheses. The lengths of the lines are proportional to the number of amino acid changes. The scale bar at the lower left indicates the number of substitutions per sequence position. The numbers at the nodes represent the percentage of 1,000 bootstrap resamplings.

genotype had 94.4% homology with the amino acids of the USA Jake1 isolate, 94.6% homology with the USA Jake2 isolate, 94.5% homology with the Brazilian Sao Paulo isolate, and 93.7% homology with the Israeli 611 isolate.

A phylogenetic tree of *E. canis* was inferred based on the *gp200* amino acid sequences obtained in this study and four other reported sequences [24]. In this tree, the *gp200* sequence from the TWN genotype differed distinctly from other corresponding sequences of *E. canis* isolates from North America (the United States), South America (Brazil), and West Asia (Israel), and formed a separate phylogenetic clade (Fig. 4).

Finally, the specificity and detection limit of the four sets of PCR primers used for *gp200* amplification (Fig. 5) were evaluated. For the purpose of specificity testing, the genomic DNA extracted from an *E. canis*-positive dog blood sample with 5% parasitemia was used to conduct PCR followed by agarose gel electrophoresis. The results showed that each of the four primer sets generated individual amplified products of the expected size, whereas no product was detected in reactions conducted without template DNA or using DNA that was separately extracted from blood samples of a healthy dog and dogs infected with different types of rickettsia, including *Ehrlichia chaffeensis*, *Ehrlichia ewingii*, *Anaplasma platys*, *Babesia canis vogeli*, and *Babesia gibsoni*. Among the four primer sets examined, EC200-F3/R3 exhibited the clearest pattern (Fig. 5A). Subsequently, the sensitivity of the PCR assay using this primer set was also determined. The blood sample with 5% parasitemia was serially diluted 10-fold with blood obtained from a healthy dog. The DNA extracted from these dilutions was used to determine the detection limit. As shown in Fig. 5B, the PCR results indicated that the lowest detection limit was 0.00005 parasitemia.
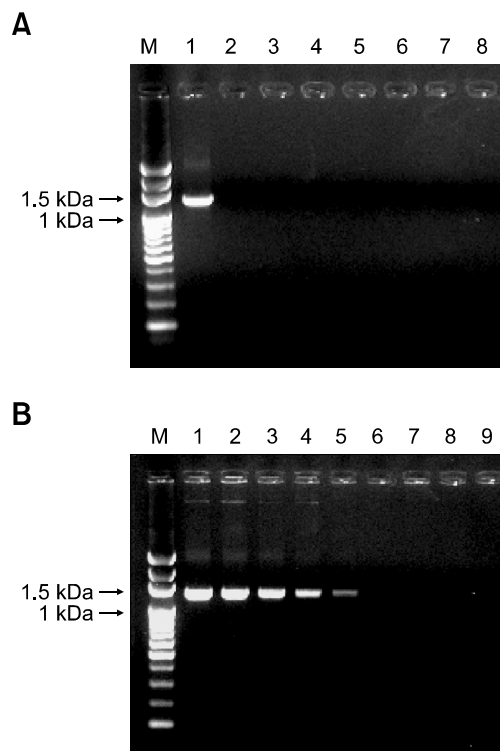

Fig. 5. The specificity and detection limit of the *gp200*-targeted PCR amplification using the primer set EC200-F3/R3. (A) The DNA extracted from blood samples of dogs infected by different rickettsia (Lane 1: *E. canis*, Lane 2: *E. chaffeensis*, Lane 3: *E. ewingii*, Lane 4: *A. platys*, Lane 5: *B. canis vogeli*, Lane 6: *B. gibsoni*) and a healthy dog (Lane 7) were subjected to PCR. The products were electrophoresed on a 1.2% agarose gel and stained with ethidium bromide. Lane M, 100-bp DNA ladder marker; Lane 8: no template DNA. (B) The blood sample with 5% parasitemia was serially diluted 10-fold from 10^0 to 10^{-7} (Lanes 1-8, respectively) with blood obtained from a healthy dog. Lane M, 100-bp DNA ladder marker; Lane 9: no template DNA.

Discussion

While canine *E. canis* infection is considered to be enzootic throughout Taiwan, this conclusion is based on clinical signs, hematological abnormalities and microscopic examination of Giemsa-stained blood smears or serological evaluations such as enzyme-linked immunosorbent assays. In the present study, the nearly full-length 16S rDNA of *E. canis* from 17 infected dogs was amplified and sequenced, providing molecular evidence that this etiologic agent is involved in canine disease in Taiwan. Although the generated 16S rDNA sequences showed a very low diversity when compared with those from other geographically distinct isolates, additional analysis of the antigen-encoding gene *gp200* helped identify genetic variations among *E. canis* isolates.

E. canis gp200 is a secreted nuclear translocated ankyrin repeat-containing protein that has five major species-

specific epitopes that are primarily located in terminal acidic domains [14]. The protein has been shown to elicit strong antibody responses in the acute phase of the infection [24]. The development of a widely applicable vaccine for CME is undoubtedly dependent on an understanding of genetic differences that may exist in geographically dispersed strains of *E. canis*, particularly with respect to the genes coding for immunodominant antigens, such as *gp200*. Previous studies have reported that *gp200* had a high level of conservation between the USA and Brazil isolates, but that substantial divergence was present in the Israeli isolate [14,24]. In the present study, we found that amino acid changes in *gp200* from the TWN genotype were distributed throughout the protein, and that some of the amino acid substitutions occurred in known *E. canis*-specific epitopes [14]. Notably, most of the amino acid substitutions in epitope-containing regions were dimorphic, containing only two different amino acids at positions wherever substitutions occur. With respect to the Israeli isolate, it has been suggested that amino acid substitutions in *gp200* epitopes are responsible for the lower immunoreactivity of the protein with heterologous antisera when compared with homologous Israeli antisera [24]. Accordingly, it would be interesting to determine if notable differences in the TWN genotype are involved in any antigenic variability.

Conversely, the sequence divergence in the *gp200* gene among geographically distributed *E. canis* strains examined in this study and others has provided useful information regarding a possible new target for genotyping of the organism. The results presented here also expand our knowledge of the genetic variability of *E. canis* and encourage further research for the analysis of genetic variation of *E. canis* strains worldwide using additional samples. Further studies of *E. canis* isolates with greater global distribution may enable inference of the phylogeographic patterns of these strains. Considering that the clinical outcome of *E. canis* infection vary widely from death to an asymptomatic chronic carrier state in untreated animals [4,6,13,19], it would be of interest to assess the relationship between genetic variations in *gp200* and strain infectivity or virulence of *E. canis*.

In summary, the present study documents infection of dogs with *E. canis* from Taiwan using molecular methods. Molecular evidence indicates that the *E. canis* genotypes circulating in dogs in Taiwan appear to be highly conserved. Based on phylogenetic analysis of the *gp200* sequences, it is clear that the Taiwanese genotype reported here represents a novel strain of *E. canis* that has not yet been characterized. Additionally, the divergence of *gp200* among geographically dispersed strains examined in this study and others supports the conclusion that this gene is useful for molecular genotyping of *E. canis*.

Acknowledgments

References

1. **de Castro MB, Machado RZ, de Aquino LP, Alessi AC, Costa MT.** Experimental acute canine monocytic ehrlichiosis: clinicopathological and immunopathological findings. Vet Parasitol 2004, **119**, 73-86.
2. **Groves MG, Dennis GL, Amyx HL, Huxsoll DL.** Transmission of *Ehrlichia canis* to dogs by ticks (*Rhipicephalus sanguineus*). Am J Vet Res 1975, **36**, 937-940.
3. **Harrus S, Aroch I, Lavy E, Bark H.** Clinical manifestations of infectious canine cyclic thrombocytopenia. Vet Rec 1997, **141**, 247-250.
4. **Harrus S, Waner T, Bark H, Jongejan F, Cornelissen AWCA.** Recent advances in determining the pathogenesis of canine monocytic ehrlichiosis. J Clin Microbiol 1999, **37**, 2745-2749.
5. **Huxsoll DL, Hildebrandt PK, Nims RM, Walker JS.** Tropical canine pancytopenia. J Am Vet Med Assoc 1970, **157**, 1627-1632.
6. **Komnenou AA, Mylonakis ME, Kouti V, Tendoma L, Leontides L, Skountzou E, Dessiris A, Koutinas AF, Ofri R.** Ocular manifestations of natural canine monocytic ehrlichiosis (*Ehrlichia canis*): a retrospective study of 90 cases. Vet Ophthalmol 2007, **10**, 137-142.
7. **Luo T, Zhang X, Nicholson WL, Zhu B, McBride JW.** Molecular characterization of antibody epitopes of *Ehrlichia chaffeensis* ankyrin protein 200 and tandem repeat protein 47 and evaluation of synthetic immunodeterminants for serodiagnosis of human monocytotropic ehrlichiosis. Clin Vaccine Immunol 2010, **17**, 87-97.
8. **Mavromatis K, Doyle CK, Lykidis A, Ivanova N, Francino MP, Chain P, Shin M, Malfatti S, Larimer F, Copeland A, Detter JC, Land M, Richardson PM, Yu XJ, Walker DH, McBride JW, Kyrpides NC.** The genome of the obligately intracellular bacterium *Ehrlichia canis* reveals themes of complex membrane structure and immune evasion strategies. J Bacteriol 2006, **188**, 4015-4023.
9. **McBride JW, Comer JE, Walker DH.** Novel Immunoreactive glycoprotein orthologs of *Ehrlichia* spp. Ann N Y Acad Sci 2003, **990**, 678-684.
10. **McBride JW, Corstvet RE, Gaunt SD, Boudreaux C, Guedry T, Walker DH.** Kinetics of antibody response to *Ehrlichia canis* immunoreactive proteins. Infect Immun 2003, **71**, 2516-2524.
11. **Mylonakis ME, Koutinas AF, Breitschwerdt EB, Hegarty BC, Billinis CD, Leontides LS, Kontos VS.** Chronic canine ehrlichiosis (*Ehrlichia canis*): a retrospective study of 19 natural cases. J Am Anim Hosp Assoc 2004, **40**, 174-184.
12. **Neer TM, Breitschwerdt EB, Greene RT, Lappin MR.** Consensus statement on ehrlichial disease of small animals from the infectious disease study group of the ACVIM. J Vet

Intern Med 2002, **16**, 309-315.

13. **Neer TM, Harrus S.** Canine monocytotropic ehrlichiosis and neorickettsiosis (*E. canis, E. chaffeensis, E. ruminatium, N. sennetsu* and *N. risticii infection*). In: Greene CE (ed.). Infectious Diseases of the Dog and Cat. pp. 203-217, Saunders, St. Louis, 2006.

14. **Nethery KA, Doyle CK, Zhang X, McBride JW.** *Ehrlichia canis gp200* contains dominant species-specific antibody epitopes in terminal acidic domains. Infect Immun 2007, **75**, 4900-4908.

15. **Pinyoowong D, Jittapalapong S, Suksawat F, Stich RW, Thamchaipenet A.** Molecular characterization of Thai *Ehrlichia canis* and *Anaplasma platys* strains detected in dogs. Infect Genet Evol 2008, **8**, 433-438.

16. **Siarkou VI, Mylonakis ME, Bourtzi-Hatzopoulou E, Koutinas AF.** Sequence and phylogenetic analysis of the 16S rRNA gene of *Ehrlichia canis* strains in dogs with clinical monocytic ehrlichiosis. Vet Microbiol 2007, **125**, 304-312.

17. **Tamura K, Dudley J, Nei M, Kumar S.** MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 2007, **24**, 1596-1599.

18. **Thompson JD, Higgins DG, Gibson TJ.** CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res 1994, **22**, 4673-4680.

19. **Unver A, Huang H, Rikihisa Y.** Cytokine gene expression by peripheral blood leukocytes in dogs experimentally infected with a new virulent strain of *Ehrlichia canis*. Ann N Y Acad Sci 2006, **1078**, 482-486.

20. **Unver A, Rikihisa Y, Kawahara M, Yamamoto S.** Analysis of 16S rRNA gene sequences of *Ehrlichia canis, Anaplasma platys*, and *Wolbachia* species from canine blood in Japan. Ann N Y Acad Sci 2003, **990**, 692-698.

21. **Vinasco J, Li O, Alvarado A, Diaz D, Hoyos L, Tabachi L, Sirigireddy K, Ferguson C, Moro MH.** Molecular evidence of a new strain of *Ehrlichia canis* from South America. J Clin Microbiol 2007, **45**, 2716-2719.

22. **Walker JS, Rundquist JD, Taylor R, Wilson BL, Andrews MR, Barck J, Hogge AL Jr, Huxsoll DL, Hildebrandt PK, Nims RM.** Clinical and clinicopathologic findings in tropical canine pancytopenia. J Am Vet Med Assoc 1970, **157**, 43-55.

23. **Yu XJ, McBride JW, Walker DH.** Restriction and expansion of *Ehrlichia* strain diversity. Vet Parasitol 2007, **143**, 337-346.

24. **Zhang X, Luo T, Keysary A, Baneth G, Miyashiro S, Strenger C, Waner T, McBride JW.** Genetic and antigenic diversities of major immunoreactive proteins in globally distributed *Ehrlichia canis* strains. Clin Vaccine Immunol 2008, **15**, 1080-1088.