Cervus elaphus 종의 sequencing과 BLAST search에 의한 감별

서정철[1,*], 김민정[1], 이찬[1], 임강현[2#]

1: 대구한의대학교 한방생명자원연구센터, 2: 세명대학교 한의과대학 본초학교실

Identification of *Cervus elaphus* Species by Sequencing Analysis and BLAST Search

Jung-Chul Seo[1,*], Min-jung Kim[1], Chan Lee[1], Kang-hyun Leem[2,#]

1: The Research Center for Biomedical Resources of Oriental Medicine,
College of Oriental Medicine, Daegu Hanny University,
2: Department of Herbal Pharmacology, College of Oriental Medicine, Semyung University

ABSTRACT

Objectives : *Cervus elaphus* species are some of the most medicinally important genera in the Oriental medicine. This study was performed to determine if *Cervus elaphus* species could be identified by sequencing analysis and to verify Basic Local Alignment Search Tool (BLAST) search, which was used to assess genetic identification.

Methods : The DNAs of *Cervus elaphus* species were extracted, amplified by PCR, and sequenced. The DNAs of Cervus species were identified by BLAST search in website.

Results : By BLAST search one of *Cervus elaphus* species was identified as Cervus elaphussibericus but the other was identified as *Cervus elaphus* nelsoni. This work showed that identification can efficiently be performed by BLAST search.

Conclusion : These results suggest that sequencing following BLAST search might be able to provide the identification of *Cervus elaphus* species.

Key words : sequencing, genetic identification, *Cervus elaphus* species

Introduction

Cervus species have been used as an important medicine in traditional Korean and Chinese medicines for a long time. They have been widely distributed in the world including East Asia. Their hornes have been used as tonic genera in general[1].

But a few dealers practiced a deception of disguising one antler from a certain country as other country. because a general consumer could not distinguish the provenance of antlers.

These Cervus species are very similar in appearance when they were sliced especially. The methods of identification among Cervus species, however, have not been recorded in traditional Korean and Chinese herbal literatures.

It is undesirable when different *Cervus elaphus* species might be used under the same name, especially it is very harmful in case of bovine spongiform encephalopathy (BSE).

And many commercial Cervus species products are extremely difficult to identify in the form of powder, or extracts. The authentication via analysing chemical profiles is also very difficult for many variables such as the analysing condition and nutritional factors.

This study was performed to determine if *Cervus elaphus* species could be identified by sequencing analysis and to verify Basic Local Alignment Search Tool (BLAST) search, which was used to assess genetic identification.

Materials and Methods

1. Samples and Purification of DNA

Two antlers of *Cervus elaphus* species were sent to the Research Center for Biomedical Resources of Oriental Medicine, Daegu Haany University (Korea) by two Oriental Medical doctors in September, 2005 and in March, 2006 respectively. We named two samples of antler as NY-C and Dongchangha. Three gram of NY-C and Dongchangha in the form of slice was firstly minced with a sterile scalpel and pulverized to a powder using a sterilized mortar and pestle. A DNA isolation kit (DNeasy, No. 69504) (Qiagen Inc., Valencia, CA) was used as described in the manufacturer's instructions with slight modifications. 300 mg of the powdered sample was used in the purification procedure. Before sample elution, the columns were dried at 37℃ for 5 minutes to evaporate residual ethanol. Samples were eluted in a total volume of 200 µl of TE buffer [10 mM Tris-HCl, 1 mM EDTA (pH 8.0)].

2. Preparation for sequencing

The extracted DNA was amplified by polymerase chain reaction (PCR). A 1200-bp region of mitochondrial D-loop of each NY-C and Dongchangha was amplified using 25 ng of DNA, 5 pmol of each primer; forward was 5'-TAATATACTGGTCTTGTAAACC -3' and reverse was 5'- GGGTCGGAAGGCTGGGACCAAACC -3'. The PCR amplification was performed by using 0.5 unit Taq polymerase (HT Biotechnology Ltd., Cambridge, United Kingdom). The 30 µl of PCR reaction mixtures were 10 mM Tris-HCl, pH 9.0, 1.5 mM magnesium chloride, 50 mM potassium chloride, 0.1% Triton-X 100, 0.01% [v/v] stabilizer, 0.25 mM of each deoxynucleotide triphosphate (dNTP), 0.1 M of each oligonucleotide primer. The PCR steps were denaturation of 5 minutes at 95℃, 30 cycles of 30 seconds at 95℃, 30 seconds at 6 0℃, and 30 seconds at 72℃ with a PCR System (Astec, Fukuoka, Japan). The quality of PCR products was controlled by 1.5% of agarose gel electrophoresis.

3. Sequencing

All amplicons were purified using the PCR-M Clean Up System (Viogene). The DNA fragments were sequenced using an ABI Prism 377 automated DNA sequencer (Applied Biosystems) with a BigDye Terminator cycle sequencing kit (version 3.1; Applied Biosystems). All amplicons were sequenced on both strands using primers of D-loop. We sequenced at least three times just to

make sure there were no errors in the sequence. The sequences were proofread and aligned and a consensus sequence was obtained with SeqMan II (DNASTAR).

4. BLAST Search

BLAST is a package of applications that can quickly and efficiently compare a specified nucleotide or amino acid sequence (or batch of such sequences) with an existing database of known sequences, translating, if necessary, between nucleotides and amino acids, and determine similarities among them. Various builds of blast are specialized to certain types of searches. This implementation used a build of the general version, blastn and bl2seq in National Center for Biotechnology Information (NCBI)[2,3].

Submenu in BLAST. we clicked on the link for DNA sequence data, copy and paste the DNA sequence data into the search box and clicked the BLAST button. Once the screen updated, we clicked the Format button. The screen eventually updated to the results screen. We scrolled down to the first "hit" in the BLAST results. The main features to look at were the E Value and the description. E Values of 0.0 are strong hits to sequences in the database.

The sequence was similar with the one that is in the database. To see this, we clicked on the score link next to the description, then could see how our sequence (called as a query) aligns to the sequence in the database, and look closely at the alignment and note any misalignments and gaps. One important data that we can get from the alignment is the number of identities. The ratio of identities to total number of bases give a percent identity. This is useful in determining the relatedness of our organisms to the one in the database.

Results

1. Sequencing Data

The sequences obtained in our study were shown in Figure 1 and 2.

```
TCGTAGCTCTCYAGAKCTCCCKAGGAGCM-KCCCTAtCCWCCCWAR
-G-ASTC-YAMTATTCCYTSAGACSTMYHAATAKRrcTCCMCAACcAYC
MAGAACTTTaTCAGAaTTAAaTTTCcAAAAaaTTtAATaTTTtAaTaCAGC
TTTcTACTCAACATCCAaTTTACATTTtATGTCcTACTAATtACACAGCA
AAACACATGATATAACTTTATGCACTCGTAGTACATAAAATCAATGT
GCTAGGACATACATGTATAACAGTACATGAGTTAGCGTATAGGACAT
ACTATGTATAATAGTACATAAATTAATGTATTAGGACATATTATGTA
TAATAGTACATTATATTATATGCCCCATGCTTATAAGCATGTACCCT
TCACTATCTAAAGTACATAGTACATAATGTTGTCCATCGTACATAGC
ACATTAAGTCAAATCAGTCCTTGTCAACATGCGTATCCCGTCCCCTAG
ATCACGAGCTTAATTACCATGCCGCGTGAAACCAGCAACCCGCTGGGC
AGGGATCCCTCTTCTCGCTCCGGGCCCATGGACCGTGGGGGTAGCTATT
TAATGAACTTT-ATCAGACATCTGG-TTCTTTTTT-CAGGGCCATCTC
ATCTAAAATCGCCCACTCCTTGCAATATAAGACATCTCGATGGACTAA
TGACTAATCAGCCCATGCTCACACATAACTGTGGTGTCATACATTTGG
TATTTTTAATTTTTGGGGGGATGCTTGgACTCAGCAATGgCCGTCTGA
GgCCCcGTCCCGGAGCATGAATTGTAgCTGGACTTAaCTGCATCTTGAG
CATCCCcATAATGGtTAGGCGCAGGgCATTACAGTCAaTGgTCACAGgACA
TAATTATTATTtCmTGAGtCAMbkgATAAGATCCAACCYCCCTACGGTT
CHACT-TAARCACCCWAMCAAASAMMCCAGRRSTCAAAATAGCACTCC
AGAGGRGCMKAASCAACCTCCCA-AASACTCCAGACG
```

Figure 1. Sequence of antler (Dongchangha). The other letters than A, G, C and T mean heterozygotic sites (IUPAC code). V is A/C/G. H is A/C/T. D is A/G/T. B is C/G/T. M is A/C. R is A/G. W is A/T. S is C/G. Y is C/T. K is G/T. N is either of the four bases (G/A/T/C). In the symbol the capital letters indicate that they are the main bases at each position, and the small letter means that it is the second main at the position.

```
TAWTCTGGTTTGGTcCCAGCCTTCCGACCCAGGAAGAAGCCATAG

CCCCAC-TATCAACACCCAAAGCTG-AAGTTCTATTTAAACTATTCCC

TGACGCTTATTAATATAGTTCCATAAAAATCAAGAACTTTATCAGT

ATTAAATTTCCAAAAAATTTAATATTTTAATACAGCTTTCTACTCA

ACAT-CC-AATTTACATTTTATGTCCTACTAATTACACAGCAAAACA

CGTGATATAACCTTATGCGCTCGTAGTACATAAAATCAATGTGCTAG

GACATGCATGTATAACAGTACATGAGTTAGCGTATAGGACATATTAT

GTATAATAGTACATAAATTAATGTATTAAGACATACTATGTATAAT

AGTACATTATATTATATGCCCCATGCTTATAAGCATGTACTTCTCAT

CATTTAAAGTACATAGTACATAATGTTGTTCATCGTACATAGCACAT

TAAGTCAAATCAGTCCTTGTCAACATGCGTATCCCGCCCCCTAGATCA

CGAGCTTAATTACCATGCCGCGTGAAACCAGCAACCC-GCTGGGCAGG

GATCCCTCTTCTCGCTCCGGGCCCATGAACCGTGGGGGTAGCTATTTA

ATGAACTTTATCAGACATCTGGTTCTTTTTTCAGGGCcATCTC-ATCT

AAAATCGCCC--ACTCC-TTGTAATATAAGACATCTCGATGGACTAAT

GACTAATCAGCCCATGCTCACACATAACTGTGGTGTCATACATTTGG

TATTTTTAATTTTTGGGGGGATGCTtGG-ACTCAGCAW-GGCCGTCTG

AGGGCCCGtCCcGGAGCAnGAATGTAgCTGGacTTACTGC-TCTGGGCAtCC

CATATGGTAGCGCAGGCATACGTCANGNCCNGNCAAATATATTCTGN

GCACCTAAACNTTTCCCCCTCTATTTTCCCCTAAAGTTC
```

Figure 2. Sequence of antler (NY-C). The other letters than A, G, C and T mean heterozygotic sites (IUPAC code). V is A/C/G. H is A/C/T. D is A/G/T. B is C/G/T. M is A/C. R is A/G. W is A/T. S is C/G. Y is C/T. K is G/T. N is either of the four bases (G/A/T/C). In the symbol the capital letters indicate that they are the main bases at each position, and the small letter means that it is the second main at the position.

2. BLAST Search

The sequences were compared to reference data available at the GenBank database by using BLAST to determine if species identification was possible by D-loop sequencing (Figure 3 and 4). By BLAST search one of *Cervus elaphus* species (Dongchangha) was close to Cervus elaphus sibericus but the other (NY-C) was close to *Cervus elaphus* nelsoni.

Sequences producing significant alignments: (Bits)   Value

		(Bits)	Value
gi|4678982|gb|AF058371.1|AF058371	Cervus elaphus sibericus is...	1546	0.0
gi|4678981|gb|AF058370.1|AF058370	Cervus elaphus sibericus is...	1521	0.0
gi|4678980|gb|AF058369.1|AF058369	Cervus elaphus sibericus is...	1505	0.0
gi|3327464|gb|AF016966.1|AF016966	Cervus elaphus nelsoni ROCK...	1464	0.0
gi|2245584|gb|AF005199.1|AF005199	Cervus elaphus manitobensis...	1464	0.0
gi|11138988|gb|AF291882.1|AF291882	Cervus canadensis nelsoni ...	1456	0.0
gi|3327477|gb|AF016979.1|AF016979	Cervus elaphus nelsoni TNP1...	1456	0.0
gi|3327478|gb|AF016980.1|AF016980	Cervus elaphus nelsoni TNP2...	1456	0.0
gi|3327469|gb|AF016971.1|AF016971	Cervus elaphus roosevelti R...	1456	0.0
gi|3327467|gb|AF016969.1|AF016969	Cervus elaphus roosevelti R...	1456	0.0
gi|3327465|gb|AF016967.1|AF016967	Cervus elaphus roosevelti R...	1456	0.0
gi|3327460|gb|AF016962.1|AF016962	Cervus elaphus nelsoni ROCK...	1456	0.0
gi|3327459|gb|AF016961.1|AF016961	Cervus elaphus nelsoni ROCK...	1456	0.0
gi|3327452|gb|AF016954.1|AF016954	Cervus elaphus manitobensis...	1456	0.0
gi|3327455|gb|AF016957.1|AF016957	Cervus elaphus manitobensis...	1450	0.0
gi|3327456|gb|AF016958.1|AF016958	Cervus elaphus manitobensis...	1450	0.0
gi|66062454|gb|AY970666.1|	Cervus elaphus canadensis D-loop, ...	1448	0.0
gi|3327462|gb|AF016964.1|AF016964	Cervus elaphus nelsoni ROCK...	1448	0.0
gi|3327475|gb|AF016977.1|AF016977	Cervus elaphus nannodes TUL...	1448	0.0
gi|3327463|gb|AF016965.1|AF016965	Cervus elaphus nelsoni ROCK...	1448	0.0
gi|2245581|gb|AF005196.1|AF005196	Cervus elaphus manitobensis is	1448	0.0
gi|2245582|gb|AF005197.1|AF005197	Cervus elaphus manitobensis is	1442	0.0
gi|3327468|gb|AF016970.1|AF016970	Cervus elaphus roosevelti R...	1440	0.0

Figure 3. BLAST search results of antler (Dongchangha). By BLAST search (the first "hit" in the BLAST results) Dongchangha was close to *Cervus elaphus sibericus*.

Sequences producing significant alignments: (Bits)   Value

		(Bits)	Value
gi|3327462|gb|AF016964.1|AF016964	Cervus elaphus nelsoni ROCK...	1479	0.0
gi|3327463|gb|AF016965.1|AF016965	Cervus elaphus nelsoni ROCK...	1479	0.0
gi|3327464|gb|AF016966.1|AF016966	Cervus elaphus nelsoni ROCK...	1471	0.0
gi|2245584|gb|AF005199.1|AF005199	Cervus elaphus manitobensis...	1463	0.0
gi|3327478|gb|AF016980.1|AF016980	Cervus elaphus nelsoni TNP2...	1455	0.0
gi|3327469|gb|AF016971.1|AF016971	Cervus elaphus roosevelti R...	1455	0.0
gi|3327467|gb|AF016969.1|AF016969	Cervus elaphus roosevelti R...	1455	0.0
gi|3327465|gb|AF016967.1|AF016967	Cervus elaphus roosevelti R...	1455	0.0
gi|3327460|gb|AF016962.1|AF016962	Cervus elaphus nelsoni ROCK...	1455	0.0
gi|3327459|gb|AF016961.1|AF016961	Cervus elaphus nelsoni ROCK...	1455	0.0
gi|3327452|gb|AF016954.1|AF016954	Cervus elaphus manitobensis...	1455	0.0
gi|3327455|gb|AF016957.1|AF016957	Cervus elaphus manitobensis...	1449	0.0
gi|3327456|gb|AF016958.1|AF016958	Cervus elaphus manitobensis...	1449	0.0
gi|3327475|gb|AF016977.1|AF016977	Cervus elaphus nannodes TUL...	1447	0.0
gi|2245583|gb|AF005198.1|AF005198	Cervus elaphus manitobensis...	1447	0.0
gi|2245581|gb|AF005196.1|AF005196	Cervus elaphus manitobensis is	1447	0.0
gi|3327458|gb|AF016960.1|AF016960	Cervus elaphus manitobensis...	1441	0.0
gi|2245582|gb|AF005197.1|AF005197	Cervus elaphus manitobensis is	1441	0.0
gi|66062454|gb|AY970666.1|	Cervus elaphus canadensis D-loop, ...	1439	0.0
gi|3327468|gb|AF016970.1|AF016970	Cervus elaphus roosevelti R...	1439	0.0
gi|3327477|gb|AF016979.1|AF016979	Cervus elaphus nelsoni TNP1...	1439	0.0
gi|3327453|gb|AF016955.1|AF016955	Cervus elaphus manitobensis...	1439	0.0

Figure 4. BLAST search results of antler (NY-C). By BLAST (the first "hit" in the BLAST results) NY-C was close to *Cervus elaphus nelsoni*.

The rates of correct identification by D-loop sequence analysis were 98% (781/789) and 98% (787/798), respectively (Table 1). In BLAST search

NY-C showed different rank pattern compared with and Dongchangha.

Results of this study leaved more to be investigated and answered, but they proposed the useful tool for identification of *Cervus elaphus* Species.

Table 1. List of identificated species of *Cervus elaphus* species

Sample name	Identificated species	Score(bits)	Matching(bp)	%identity
Dongchan gha	*Cervus elaphus sibericus*	1546	781/789	98
	Cervus elaphus nelsoni	1464	771/789	97
NY-C	*Cervus elaphus sibericus*	1479	787/798	98
	Cervus elaphus nelsoni	1384	775/798	97

Discussion

There are several types of DNA sequence variation, including single base pair difference-Single Nucleotide Polymorphisms (SNP), differences in the copy number of repeated sequences, and insertions and deletions. Characterization and scoring of genetic variations is increasingly important to correlate phenotype and genotype differences. We investigated the feasibility to determine *Cervus elaphus* species in DNA by using sequencing analysis and BLAST search, which was used to assess genetic identification.

Some deer horn sources as traditional Korean and Chinese medicines might have changed over times. It is undesirable when different *Cervus elaphus* species might be used under the same name, because some *Cervus elaphus* nelsoni in Canada having BSE since December, 2000. In these cases, genetic identification of traditional Korean and Chinese herbs should help to ensure the safe use of Korean and Chinese herbal materials. So, the method for identifying the origin is very important.

In BLAST search one of *Cervus elaphus* species was identified as *Cervus elaphus* sibericus but the other was identified as *Cervus elaphus* nelsoni. This work showed that identification can efficiently be performed by BLAST search. We designed specific primer to identify *Cervus elaphus* species. From above results we verified that our *Cervus elaphus* species-specific primers were well designed.

These results suggest that sequencing analysis and BLAST search methods are suitable for authentication of the concerned *Cervus elaphus* species. But in our study we couldn't identify reared places or areas of *Cervus elaphus* species. In conclusion, sequencing following BLAST search might be able to provide the identification of *Cervus elaphus* species.

Acknowledgement

References

1. Kang BS, Kim IR, Kim HC, Kuk YB, Park YK, Seo BI, Seo YB, Song HJ, Shin MK, Lee YJ, Lee YC, Leem KH, Cho SI, Chung JK, Joo YS, Choi HY. Boncho-Hak. Seoul : Young-Lim Press. 2004 : 591-3.
2. http://www.ncbi.nlm.nih.gov/BLAST/
3. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990 ; 215(3) : 403-10.