Full Length Research Paper

cDNA cloning of v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT) in silver fox (Vulpes vulpes) and its sequence polymorphism analysis among foxes and dogs

Chunyan Bai, Jiaxin Yu, Wei Zhao, Jiani Hou, Yumei Li and Shouqing Yan*

Department of Animal Science, College of Animal Science and Veterinary Medicine, Jilin University, Changchun, Jilin, China.

Accepted 25 April, 2012

As a critical receptor for basic cell functions, v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT) has been found to be closely related to many types of tumors and coat colour in animals. In this study, we cloned and sequenced the full-length coding sequence of KIT in silver fox (Vulpes vulpes). Alignment together with the public accessible mRNA sequences of dogs and transcriptome sequences of fox was run to predict the possible polymorphisms between dog and fox and within them. Results showed that the only difference between dog and fox in KIT amino acid sequence is a dog exclusive feature of an insertion of two Leu residues and the adjacent Gin/Leu substitution in signal peptide region. Twenty synonymous singe nucleotide polymorphisms (SNPs) were identified between dog and fox. An indel of 760-762 CAG and two transcript splice variants of 12bp difference in length were found among both dogs and foxes. The present study provides basic information for further studying on fox KIT and references for the comparative study on this gene in dog and other Canidae animals.

Key words: Silver fox, KIT, coding sequence, dog, polymorphism.

INTRODUCTION

The v-kit Hardy-Zuckerman 4 feline sarcoma viral oncogene homolog (KIT) receptor tyrosine kinase is a type 3 transmembrane receptor for mast cell growth factor (MGF) also known as stem cell factor). It was originally isolated as a retroviral oncogene (v-kit) from the acute transforming Hardy-Zuckerman 4 feline retrovirus. Its cellular homolog (c-kit) was identified in cat genomic DNA (Besmer et al., 1986). KIT is often used for the simplicity for both the protein and gene name of c-kit. KIT is also assigned as a member of the platelet-derived growth factor receptor (PDGFR) family (Yarden et al., 1987) and cluster of differentiation number117 (CD117). Activation of the receptor by its ligand also known as stem cell factor (SCF) or steel factor, is essential for stimulating the formation of many types of blood and other cells, including melanocytes, germ cells, gonadal and pigment stem cells (Brizzi et al.,1994; Galli et al., 1994; Serve et al., 1994; Linnekin et al., 1997; Price et al., 1997) and the interstitial cells of Cajal, which mediate intestinal motility (Besmer et al., 1986; Robertson et al., 2000). KIT also acts as an essential survival factor for migrating and proliferating melanoblasts (Blume-Jensen et al., 1991; Steel et al., 1992). In addition, KIT is involved closely with regulation of cell division and proliferation, and any alteration in its activity may play a very crucial role in neoplasms. Its mutations were found to be related to many types of tumors and different coat color variation in animals.

Domesticated dogs (Canis lupus familiaris) are very
ideal animal models for naturally occurring tumors, due to their easy accessibility and prominent status in diverse cultures (Rowell et al., 2011). As a member of the Canidae family, fox is considered as important farm fur-bearing animal and also useful model in comparative genomic studies of the canids (Switonski et al., 2009). Learning the information of fox KIT gene will provide important reference for the related studies in dogs. In addition, as an important candidate gene for animal coat color, fox KIT gene is necessarily to be studied for further exploring the molecular basis of coat color differences in this species, since little is known about the genetic basis for the plenty color variations in foxes, such as white, silver, platinum, grey, cross, blackish-brown, amber, samson and red in silver foxes (Vulpes vulpes). In this study, we obtained the cDNA sequence of KIT gene in silver fox and analyzed the coding sequences of this gene in dogs and foxes.

MATERIALS AND METHODS

Animal sample and cDNA preparation

The normal lung tissue in which the KIT has been shown to be expressed in dog (Ma et al., 1999), were taken from a seven-month-old silver fox (raised for fur products in Xingye Special Economic Animal Farm in Daan, Jilin, China), frozen in liquid nitrogen, and stored at -70°C until it was used. Total RNA was extracted using the TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the manufacturer instructions and then treated with RNase free-DNase I (Promega, Madison, WI, USA) to remove any contaminating genomic DNA. First-strand cDNA was synthesized using One Step reverse transcription-polymerase chain reaction (RT-PCR) Kit (PrimeScript®, Takara, Dalian, China). The reaction was performed at 42°C for 40 min and then 99°C for 5 min.

Cloning of fox KIT cDNA

According to the mRNA sequence of canine KIT (NM_001003181.1) and corresponding canine genomic sequence, three pairs of primers (Table 1) were designed to obtain the full length cDNA of fox KIT. P1 was designed to obtain the sequence containing the start codon while P3 was designed to obtain the sequence containing the stop codon. The forward primer of P1 was designed according to the genomic sequence (NC_006595.3), since the sequence located at the upstream of start codon in the mRNA sequence (GenBank® NM_001003181.1) is not identical with the genomic sequence and the primer designed according to the mRNA sequence failed to obtain product in our previous study. The reverse transcription polymerase chain reaction (RT-PCR) was performed with an Eppendorf Mastercycler (Eppendorf Limited, Hamburg, Germany). The amplification conditions were as follows: 94°C for 5 min, 30 cycles of 94°C for 30 s, 60°C and 72°C for 60 s, then completed by 72°C for 10 min. The amplified fragment of the predicted size was subcloned into the pMD18-T vector (TaKaRa, Dalian, China), and then sequenced on automated sequencers ABI 3730XL (Sangon Biotechnology Co., Ltd., Shanghai, China).

Sequence homology analysis

The sequences of three fragments obtained with P1, P2 and P3, overlapped with each other were assembled to obtain the full-length cDNA sequence of fox KIT using DNAMAN (version 1.0.7.176, Lynnon Corporation, Pointe-Claire, Quebec, Canada). The coding sequence (CDS) was isolated by being aligned to the canine CDS and translated to amino acid sequence with DNAMAN. The CDS and the deduced amino acid sequence of fox KIT were aligned against the reference RNA database of GenBank by BLAST (Sanger Biotechnology Co., Ltd., Shanghai, China).

Identification of fox KIT polymorphisms

The fox pre-frontal cortex transcriptome sequences from Kukekova et al. (2011) were downloaded from GenBank Sequence Read Archive (SRA) database (under accession number SRA029285.1). The transcriptome reads sequences were aligned to the fox KIT cDNA sequence obtained in this study with Bowtie2 (Langmead et al., 2009) and the matched reads sequences were collected for subsequent analysis. KIT cDNA sequences of three foxes (one obtained in this study and two from the study mentioned above) and five dogs (available in GenBank: NM_001003181.1, AY313776.1, AY296484.1, AF448148.1 and AF099030.1, respectively) were aligned together to identify the variations between dog and fox and among them with DNAMAN, including single nucleotide polymorphisms (SNPs) and insertions or deletions (indels). The nucleotide differences were considered as SNPs only when they appeared in more than three sequences for an allele, including the repeatedly sequenced transcriptome reads from the same animal. Different alleles consisting of SNPs or indels in signal peptide coding region were
translated protein sequence

(18: c > g)  (30: t > g)  (50: t > a)  (52-57 indel)

translated protein sequence

NM_001003181.1-dog
AY313776.1-dog
AY296484.1-dog
AF448148.1-dog
AF099030.1-dog
sequenced-fox
SRR094995.270685.2
SRR094995.461860.2

<table>
<thead>
<tr>
<th>NM_001003181.1-dog</th>
<th>AY313776.1-dog</th>
<th>AY296484.1-dog</th>
<th>AF448148.1-dog</th>
<th>AF099030.1-dog</th>
<th>sequenced-fox</th>
<th>SRR094995.270685.2</th>
<th>SRR094995.461860.2</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>A</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
</tr>
<tr>
<td>W</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>D</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>F</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>T</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>L</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>C</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>G</td>
<td>C</td>
<td>C</td>
<td>C</td>
</tr>
<tr>
<td>V</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
<tr>
<td>L/Q</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
<td>L</td>
</tr>
</tbody>
</table>

Figure 1. Fragmental KIT cDNA sequences containing the start codon. Three SNPs and the 52 to 57 indels between fox and dog and the deduced amino acids sequence are presented. The number for the polymorphic position is that of the CDS in NM_001003181.1 (as shown in Figures 2 and 3). The first five sequences are dog KIT cDNA sequences available online and marked by their GenBank accession numbers; the sixth is the fox KIT sequence obtained in this study, marked by "sequenced-fox"; the last two sequences are from the transcriptome sequencing reads, indicated by their reads codes (Kukekova et al., 2011). The amino acid is expressed by their one-letter abbreviation in capital letters. The Met coded by the start codon "ATG" is underlined.

RESULTS

The full-length cDNA sequence of fox KIT

The 3037 bp fox KIT cDNA sequence was obtained with 2922 bp CDS coding 973 amino acids. BLAST against the reference RNA database of GenBank showed that the fox KIT CDS shared high identity with its counterparts in dog (99%), giant panda (93%), cat (92%), horse (90%), chimpanzee (89%), Northern white-cheeked gibbon (89%), pig (88%), human (88%), cattle (88%) and Rhesus monkey (88%).

Sequence variations between dog and fox

The transcriptome reads from Kukekova et al. (2011) were aligned to the sequenced fox KIT cDNA. The result showed that each nucleotide position in the cDNA sequence were covered by at least two reads and therefore the matched reads sequences were quantified to be used for the further analysis. By alignment of KIT cDNA sequences from three foxes and five dogs, 21 fox specific variations were identified, including 20 SNPs and one indel. The SNPs represented by the position corresponding to the CDS of dog reference RNA (CDS in NM_001003181.1) are: 18 (c > g) (the former nucleotide is of dog and the latter is of fox, as in the following ones), 30 (t > g), 50 (t > a), 186 (c > t), 222 (t > c), 414 (c > t), 423 (c > t), 489 (g > t), 507 (a > g), 519 (g > c), 582 (c > g), 801 (t > c), 1260 (t > c), 1647 (g > a), 1866 (t > c), 1917 (c > a), 1932 (c > t), 1987 (g > a), 2154 (t > c), 2235 (g > a), 2259 (g > a) and 2271 (t > c), respectively. “CTGCTG” of 52-57 is an indel variation between dog and fox, existing in dogs but not in foxes. By searching in UniProt database (http://www.uniprot.org/), we also found that this 6 bp is a dog exclusive insertion among all 42 species with the sequence of this region in KIT available to date. The first two SNPs (as a demonstration for other SNPs) and the 52-57 “CTGCTG” indel are shown in Figure 1.
The deduced amino acid sequences analysis showed that all SNPs between dog and fox are synonymous, except for the one adjacent closely to the indel, 50 (t > a), which leads to a substitution of Leu17 in fox to Gln17 in dog. The indel of 52-57 CTGCTG leads to insertion of two Leu18 Leu19 in dog (shown in Figure 1). The ab initio prediction by SignalP4.0 showed that the 1 to 27 amino acid residues in dog KIT protein (1 to 25 in fox correspondingly) composed the signal peptide. Moreover, the difference of the indel and SNP between dog and fox does not affect the presence of the signal peptide with D value of 0.841 for dog and 0.561 for fox.

### Sequence variations among dogs and foxes

A 760 to 762 indel of 'CAG' corresponding to the indel of amino acid Gln254 was found in one dog sequence (AF099030.1) and two transcriptome reads sequences of fox (as shown in Figure 2). A 12 bp insertion of GGTAACAGCAAA between 1533 and 1534 was found in three dog sequences (AY296484.1, AF448148.1 and AF099030) and two of the fox transcriptome reads, compared to the other two dog sequences (NM_001003181.1 and AY313776.1), the fox sequence obtained in this study and three other matched fox transcriptome reads (Figure 3). Alignment with sequences of other species showed that this indel corresponds to the two transcript splice isoforms found in human and mouse. The longer sequence with the 12 bp insertion was assigned as transcript variant 1 in human and mouse (NM_000222.2 and NM_001122733.1, respectively), while the shorter sequence with the 12 bp deletion was assigned as transcript variant 2 (NM_002848.3 and NM_021099.3, respectively). Interestingly, these two types of transcript variants were both found present in the transcriptome reads from the same fox, indicating two transcript isoforms are expressed in same animal.
DISCUSSION

In this study, we obtained the full-length cDNA sequence of silver fox KIT by RT-PCRs and sequencing. As expected, the sequence of fox is of very high identity with that of dog, especially at amino acid sequence level. There is no difference between them except for the two Leu^{16}Leu^{19} insertion and the adjacent substitution of Gin^{17} in fox to Leu^{17} in dog, which is induced by the indel of 52-57 CTGCTG and 50 t-a, respectively. This sequence of Leu^{17}Leu^{18}Leu^{19} is exclusive in dogs compared to all the available sequences in the other species. SignalP4.0 prediction showed that the addition of these three hydrophobic amino acid residues in the dog KIT enhances the peptide’s character as signal peptide, judged according to the D value. Meanwhile, further studies are necessary to investigate the effect of this unique sequence feature on the function of the KIT protein. In dogs, the mast cell tumor (MCT) is the most common skin tumor, with an incidence of close to 20% in the canine population, much higher than in humans (London and Seguin, 2003). This dog exclusive sequence feature may provide a clue for studying the molecular genetic mechanism for the difference between species.

Since the sequences of dog and fox are highly identical, to predict the possible polymorphism between them and those within Canidae family, the transcriptome sequencing data available in public database was used for analysis in our study. Except for the indel and the non-synonymous SNP discussed above, 20 synonymous SNPs were found between fox and dog. In addition, two indels, 3 and 12 bp respectively, were showed to be present among both dogs and foxes. The 12 bp indel corresponds to the transcript variants found in human and mouse. Our results confirmed the existence of the different transcript variants in Canidae family. Interestingly, for both the indels, two types of variants were both identified in the transcript reads from same fox, indicating the simultaneous expression of two splice transcript variants in same animal.

One of the advantages of next generation sequencing is its use in identifying SNPs and small indels (Chan, 2009). The application of the public available next generation sequencing data in our study therefore showed the possibility of using the public available data of next generation sequencing in sequence analysis of a specific functional gene.

REFERENCES