Available online at www.sciencedirect.com
ScienceDirect
Gene 392 (2007) 151–156
GENE
SECTION
EVOLUTIONARY GENOMICS
www.elsevier.com/locate/gene

Aberrant termination of reproduction-related *TMEM30C* transcripts in the hominoids

Naoki Osada [a,*], Katsuyuki Hashimoto [a], Momoki Hirai [b], Jun Kusuda [a]

[a] *Division of Biomedical Research Resources, National Institute of Biomedical Innovation, 7-6-8 Saito-Asagi, Ibaraki, Osaka 567-0085, Japan*
[b] *International Research and Educational Institute for Integrated Medical Sciences, Tokyo Women's Medical University, 8-1 Kawada-cho, Shinjuku-ku, Tokyo 162-8666, Japan*

Received 14 July 2006; received in revised form 28 September 2006; accepted 28 November 2006
Available online 12 December 2006

Abstract

Finding genetic novelties that may contribute to human-specific physiology and diseases is a key issue of current biomedical studies. *TMEM30C* is a gene containing two transmembrane (TM) domains and homologous to the yeast CDC50 family, which is related to polarized cell division. It is conserved among mammals along with two other paralogs, *TMEM30A* and *TMEM30B*. We found that *TMEM30C* is expressed specifically in the testis of mammals, in contrast to the relatively wide expression distributions of the other paralogs. While macaques expressed two alternative splicing isoforms which include one or two TM domains, humans and chimpanzees predominantly expressed truncated transcripts because of the mutations in the splicing and/or poly(A) signal sites. The major transcript in humans harbored non-stop ORF (open reading frame) while the chimpanzee counterpart encoded a protein with one TM domain. The difference was due to the 1-bp indel upstream of the poly(A) signal site. In addition, both the hominoids expressed minor transcripts encoding short proteins with one TM domain. Phylogenetic analysis has showed the acceleration of amino acid substitution after the human and chimpanzee divergence, which may have been caused by a recent relaxation in functional constraints or positive selection on *TMEM30C*. Elucidating the precise reproductive function of *TMEM30C* in mammals will be important to the foundation of divergence in higher primates at a molecular level.

Keywords: Primates; Evolution; Transmembrane protein; Reproduction

1. Introduction

Genetic novelties that may contribute to the physiology and diseases of current human populations have fascinated many scientists for many years and have become an important subject of current biomedical studies (Gibbons, 1998; Hacia, 2001; Olson and Varki, 2003; Varki and Altheide, 2005). Chimpanzees (common chimpanzees and bonobos) are the closest relatives of humans, and a draft genome of the common chimpanzee (*Pan troglodytes*), which will be enormously valuable to evolutionary studies, is now publicly available (Chimpanzee Sequencing and Analysis Consortium, 2005). Consequently, extensive research has focused on genetic differentiation among higher primates, mainly based on single nucleotide substitutions and a set of genes has been uncovered in which changes might represent adaptations (Enard et al., 2002b; Clark et al., 2003; Hellmann et al., 2003; Watanabe et al., 2004; Evans et al., 2005; Glazko et al., 2005; Mekel Bobrov et al., 2005; Nielsen et al., 2005).

Undoubtedly, amino acid substitution is not the only factor responsible for the phenotypic diversity of species. In addition to amino acid substitution, divergence at a transcriptional level, such as *cis*-regulated gene expression divergence and the gain-and-loss of genes, is likely to be an important factor in the genomic evolution of organisms to create phenotypic complexity (King and Wilson, 1975; Enard et al., 2002a; Heissig et al., 2005; Marques et al., 2005; Rockman et al., 2005). Although gains-and-losses of genes might cause a larger phenotypic effect than single amino acid substitutions (Olson and Varki, 2003),

Abbreviations: TM, transmembrane; ORF, open reading frame; UTR, untranslated region.

* Corresponding author. Tel.: +81 72 641 9811; fax: +81 72 641 9812.
E-mail address: nosada@nibio.go.jp (N. Osada).

0378-1119/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.gene.2006.11.021

fewer studies have described lineage-specific gains-and-losses of genes among higher primates (Chou et al., 1998; Stedman et al., 2004; Hahn and Lee, 2005; Hayakawa et al., 2005; Wang et al., 2006), with the exception of the frequent births and deaths of genes in large multi-copied gene families (Meyer Olson et al., 2003; Fortna et al., 2004; Gilad et al., 2005; Go et al., 2005). Using the human and chimpanzee genome sequence, we could conduct a genome-wide survey of species-specific pseudogenes, which carry null mutations in the coding region either in the human and chimpanzee (Hahn and Lee, 2005; Wang et al., 2006). However, if the loss of function arose enough recently in the course of evolution, null mutation in the coding region is not a perfect criterion for nonfunctional genes. For example, mutation in a promoter region can lead null expression of the gene without coding disruption. Therefore, to explore lineage-specific transcript structures and expression patterns, valid and detailed experimental evidence is needed.

Recent studies have shown that the rate of transcriptome divergence was high in the testis, which might be involved in many biologically important phenomena such as reproductive isolation (Swanson and Vacquier, 2002). In the course of our transcriptome analysis of the macaque testis (Osada et al., 2005), we found that one of the genes, which showed homology to yeast gene CDC50, was expressed in macaques but not completely expressed in humans. In mammals, three types of CDC50 paralogs have been identified: transmembrane protein 30A (*CDC50A* or *TMEM30A*), 30B (*CDC50B* or *TMEM30B*), and 30C (*CDC50C* or *TMEM30C*). The CDC50 gene product, Cdc50p in yeast is a subunit of phospholipids-transacting P-type ATPase, which has been implicated in the asymmetrical localization of phospholipids within the plasma membrane (Saito et al., 2004) and hence is involved in polarized cell division system. The protein structure of this protein family is highly conserved from yeast to mammals; the structure comprises two transmembrane (TM) regions, with the head and tail sticking out on the cytoplasmic side. Although all three paralogous cDNA clones were present in mice, only *TMEM30A* and *30B* cDNAs have been cloned in humans, and *TMEM30C* has been predicted by only *in silico* analysis (Katoh and Katoh, 2004). Here, we describe that the *TMEM30C* gene is specifically expressed in mammalian testes and the transcript structure is highly diverse among higher primates.

2. Materials and methods

2.1. Sequence analysis

The putative coding sequence of chimpanzee *TMEM30C* was deduced from the chimpanzee draft genome sequence with the exception of the exon 6 sequence, which contained ambiguous base pairs and gaps in the public genome sequence. The sequence of chimpanzee exon 6 (DDBJ/EMBL/Genbank accession number: AB247157) was determined by sequencing the PCR product from genomic DNA, which was extracted from EB-transformed lymphocytes. The *TMEM30C* cDNA sequences of the mouse (AK161475), rat (XM_221533), dog (XM_545073), and bovine (BC111328) were obtained from the public databases. Note that the rat and dog sequences are hypothetical sequences that were predicted from the genome sequence. Human and chimpanzee *TMEM30C* cDNAs were amplified using the 3′-RACE method with an oligo-primer adaptor (GGCCACGCGTCGACTAG-TACTTTTTTTTTTTTTTTT) and forward primers in exon 2. The PCR products were cloned into pUC18 plasmids and the sequences of ≈10 clones from each library were then determined (AB249666; AB250297; AB265818; AB265819). *Macaca fascicularis* cDNA clones, QtsA-12626 (AB070082) and QtsA-16374 (AB070082), were isolated from the oligo-capped cDNA library described previously (Osada et al., 2005). The DNA samples were sequenced using an ABI 310 and 3730 sequencer (Applied Biosystems).

2.2. RT-PCR

The templates of the human and mouse total RNA were purchased from Clontech and Sawady technology (Japan), respectively. Chimpanzee testis sample was collected from an eight-year-old male chimpanzee (*P. troglodytes verus*) which died of natural causes. Total RNA samples of the cynomolgus monkey were obtained as described previously (Osada et al., 2005). One microgram of total mRNA was amplified using a One Step RNA PCR Kit (TakaraBio). The temperature and time schedules were 30 cycles of 94 °C for 20 s, 58 °C/60 °C for 30 s, and 72 °C for 1 min. The primer sequences were shown in Supplemental Table 1.

2.3. Tree construction of mammalian TMEM30C genes

For the phylogenetic analysis, we used 5′-sequences of ORFs (843 bp from the first ATG) that encode TMEM30C in the various mammals described above. The putative human and chimpanzee *TMEM30C* sequences corresponding to the other mammalian cDNAs were extracted from the human (NCBI build 35) and chimpanzee (NCBI build 1) genome sequence, respectively. The nucleotide sequences were aligned taking the translated amino acid sequences as guides and using ClustalW with default parameters (Thompson et al., 1994). The multiple alignments were used to construct a phylogenetic tree using the neighbor-joining method with Kimura's distance (Kimura, 1980; Saitou and Nei, 1987). The alignment and tree construction were performed using the MEGA 3.1 program package (Kumar et al., 2004). For the statistical test for positive selection, we used both PAML (Yang, 1997) and Hyphy (Pond et al., 2005) program packages but both programs failed to find both positively selected lineages and sites. The non-synonymous substitution rate (d_N) and the synonymous substitution rate (d_S) for each lineage were estimated using the maximum likelihood method implemented in PAML (Yang, 1997).

2.4. Tree construction of vertebrate CDC50 homologs

The amino acid sequences were used for the alignment and tree construction of vertebrate TMEM30 proteins and yeast CDC50. The phylogenetic tree was constructed using MEGA 3.1 program with *p*-distance and neighbor-joining method.