Genomics & Informatics Vol. 3(2) 61-65, June 2005

Comparative Genomics of T-complex protein 10 like in Humans and Chimpanzees

Il-Chul Kim[1], Dae-Soo Kim[1], Dae-Won Kim, Sang-Haeng Choi, Han-Ho Choi, Sung-Hwa Chae and Hong-Seog Park*

Genome Structure Research Laboratory, Korea Research Institute of Bioscience and Biotechnology, 52 Oun-dong, Yusong-gu, Daejeon 305-333, Korea

[1]These authors contributed equally to this work.

Abstract

Comparing 231 genes on chimpanzee chromosome 22 with their orthologous on human chromosome 21, we have found that 15 orthologs have indels within their coding sequences. It was rather surprising that significant number of genes have changed by indel, despite the shorter time since their divergence and led us hypothesize that indels and structural changes may represent one of the major mechanism of proteome evolution in the higher primates. Human T-complex protein 10 like (*TCP10L*) is a representative having indel within its coding sequence. Gene structure of human *TCP10L* compared with chimpanzee *TCP10L* gene showed 16 base pair difference in genomic DNA. As a result of the indel, frame shift mutation occurs in coding sequence (CDS) and human *TCP10L* express longer polypeptide of 21 amino acid residues than that of chimpanzee. Our prediction found that the indel may affect to dramatic change of secondary protein structure between human and chimpanzee *TCP10L*. Especially, the structural changes in the C-terminal region of *TCP10L* protein may affect on the interacting potential to other proteins rather than DNA binding function of the protein. Through these changes, *TCP10L* might influence gene expression profiles in liver and testis and subsequently influence the physiological changes required in primate evolution.

Keywords: chimpanzee genome, comparative genomics, *TCP10L*, evolution, insertion

*Corresponding author: E-mail hspark@kribb.re.kr, Tel +82-42-879-8132, Fax +82-42-879-8139
Accepted 2 June 2005

Introduction

Recently, as an international collaborative project, we reported the entire DNA sequence of the euchromatic region of chimpanzee chromosome 22 (PTR22) and compared the whole chromosome with human chromosome 21 (HSA21) (Watanabe *et al.*, 2004). We found that PTR22 and HSA21 differ at approximately 1.44% of their 33 million aligned nucleotides. In addition we reported nearly 68,000 insertion or deletions (indels). The number of indels is approximately one seventh the numbers of point mutations accumulated over the time since human and chimpanzees diverged. Since each indel involves the change of several nucleotides, this result confirms Britten's estimates (Britten, 2002) that indels are a major soruce of sequence divergence between human and chimpanzees. Indeed comparing 231 orthologous genes on the chromosome, we found that 15 orthologs have indels within their coding sequences and 32 have changes in the start codon or stop codon, which would potentially lead to gross structural differences between human and chimpanzee protein products. These results led us hyperthesize that indels and structural changes may represent one of the major mechanisms of proteom evolution in the higher primates.

Among 15 kinds of indels between human and chimpanzee orthologous genes, *TCP10L* is a representative having indel within its coding sequence (Watanabe *et al.*, 2004). The *TCP10L* showed 16 base pair differences in genomic DNA sequence between human and chimpzanzee. The primary structure of human *TCP10L* contains the representative character of the leucine zipper motif, with a leucine residue in every seventh position. The leucine zipper motif has been observed in a number of proteins thought to function as eukaryotic transcription factors (Amati *et al.*, 1993; Perez *et al.*, 2001; Sellers *et al.*, 1989). Especially, the gene in human is known to be specifically expressed in the liver and testis (Chen *et al.*, 2003) but not in 28 other tissues, and the protein is accumulated in nucli (Chen *et al.*, 2003). The subcellular localization and the leucine zipper structure of the protein suggested that it functions as a transcription factor. Indeed, human *TCP10L* has an activity of transcription repressor and the leucine zipper is indispensable to the biological functions of the transcription factors (Chen *et al.*, 2003). It has long been

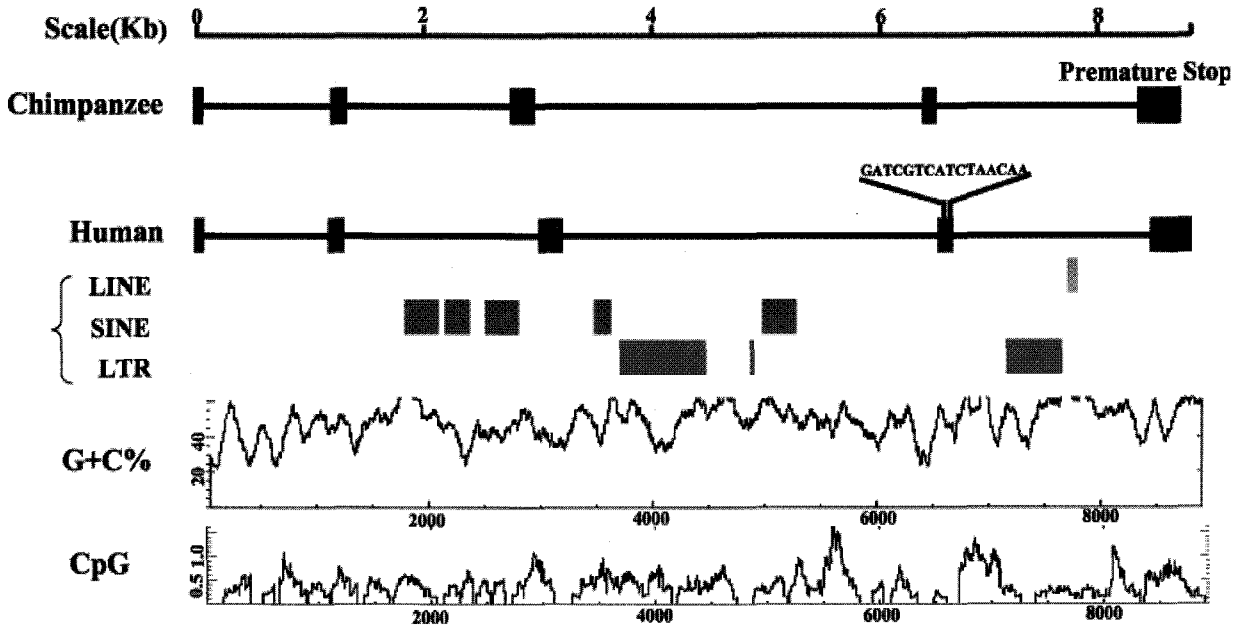

Fig.1. Comparative gene sturcture of *TCP10L* gene in Human and Chimpanzee

argued that changes in gene regulation may be more important to morphological and functional evolution than overall genomic divergence (Cherty *et al.*, 1978). To study this old issue, transcription factors can be a good candidate because the change of transcription factors could result in a global change in gene regulation. In this study, we analyze and compare genomic structure of *TCP10L* gene between human and chimpanzee and consequently discuss its implication in liver physiology and primate evolution.

Results and Discussion

Genomic structure of *TCP10L* gene in Chimpanzee and Human

Genomic organization of human *TCP10L* gene was reported by Chen et al. previously (Chen *et al.*, 2003). The *TCP10L* is localized at chromosome 21p22.11, and its genomic size is 8955 bp. As shown Fig. 1, human *TCP10L* was composed of five exons and four introns. To understand the genomic characters of chimpanzee *TCP10L*, we predicted genomic structure of chimpanzee *TCP10L* using SIM4 program (Florea *et al.*, 1998) that run by matching human Ref.Seq as a query sequence to chimpanzee genomic DNA sequence. The chimpanzee *TCP10L* showed genomic size of 8615 bp and similar exon and intron structure with human *TCP10L*. DNA sequence similarity of the *TCP10L* gene between human and chimpanzee was 97.8 % in total and similarities for each exons are shown in Table 1. Next, we tried to speculate the changes in CDS of human *TCP10L* caused by the genomic change. As shown in Fig. 2, human *TCP10L* has 9 base substitution and 16 bp indel in exon 4. The 16 bp indel in exon 4 resulted in frame shift of protein coding structure and also inactivation of stop codon. As a result, the human *TCP10L* showed longer CDS of 63 base pair than that of chimpanzees.

The consequence of genomic changes on the structure of *TCP10L*

Comparison of *TCP10L* CDS sequence between human and chimpanzee found that the human *TCP10L*

Table 1. Comparision of exon sequences between Human and Chimpanzee TCP10L

Exon	Exon Size		Identity	Mutation
	Chimp	Human		(2 base substitution)
1	86	86	98%	2 base substitution
2	145	145	96%	6 base substitution
3	216	216	99%	3 base substitution
4	121	138	97%	16 bp Human Insertion
5	372	372	100%	-

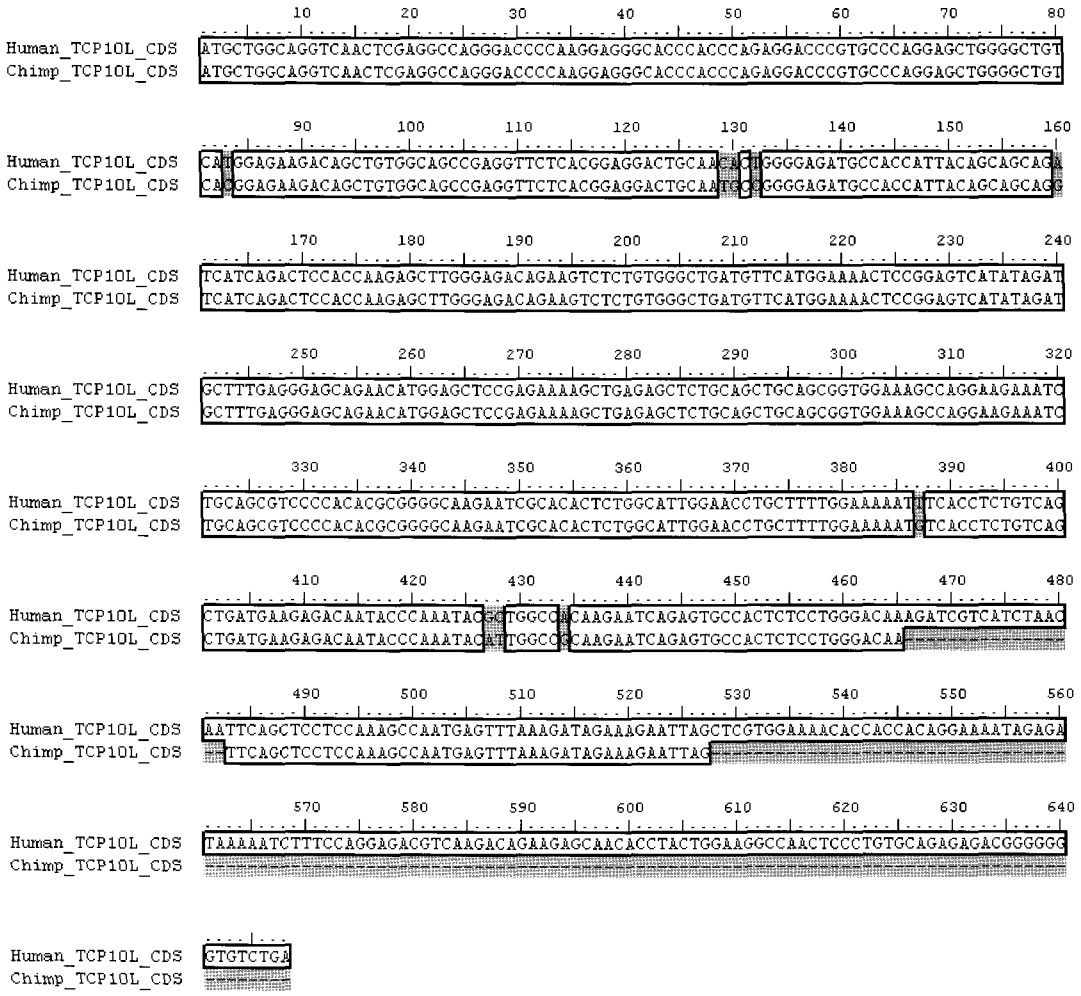

Fig. 2. Comparison of CDS between Human and Chimpanzee *TCP10L*

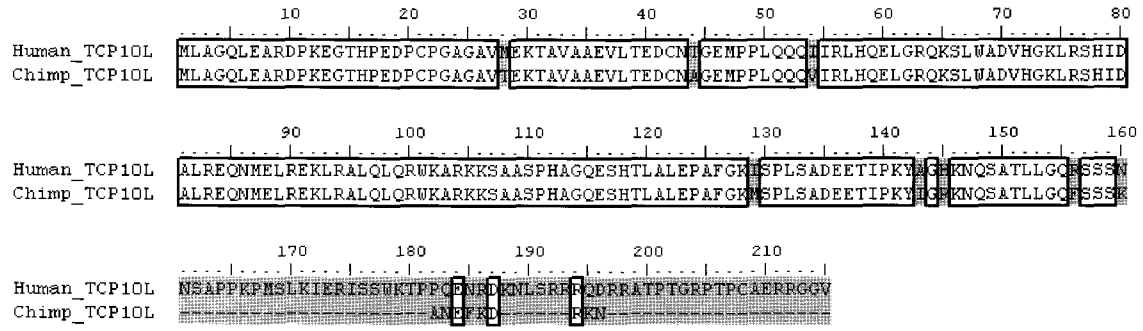

Fig. 3. Comparison of primary structure of not italic ; TCP10L between Human and Chimpanzee

deduced to be 215 amino acids polypeptide that is longer of 21 amino acid residues than that of chimpanzees (Fig. 3). We further predicted their secondary structure using PSIPRED server (McGuffin *et al.*, 2000) and found that the 16 base pair indel resulted in dramatic changes on the secondary structure between them. As shown in Fig. 4, the typical leucine zipper motif exits in the middle of peptides in both human

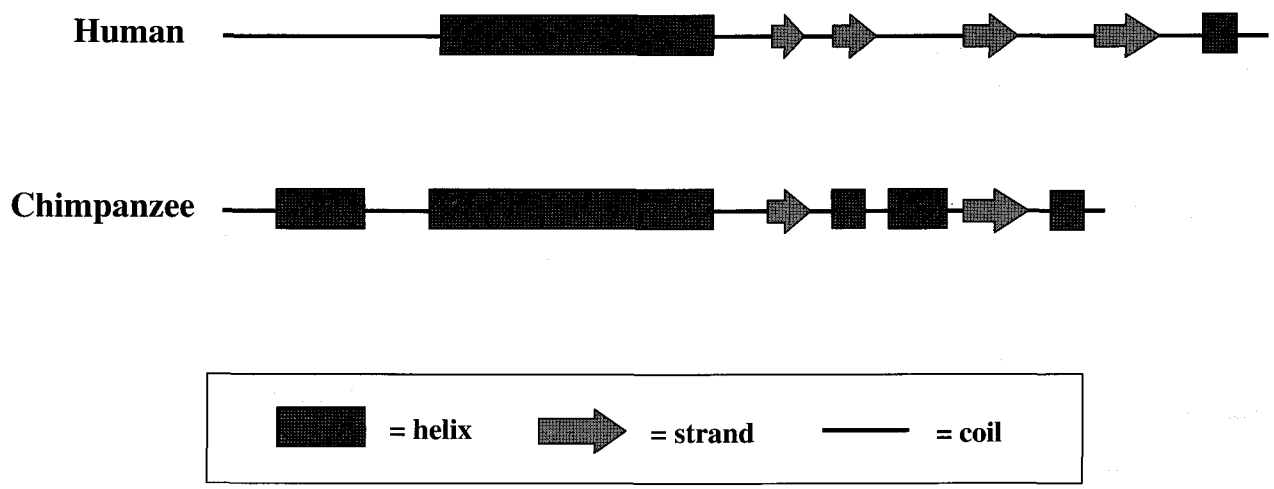

Fig. 4. Predicted secondary structure of not italic ; TCP10L in Human and Chimpanzee

and chimpanzees, but another helical structure in the chimpanzee *TCP10L* exist on the N-terminus of the typical leucine zipper. Dramatic changes of protein structure found out on the C-terminal region of the *TCP10L*. In spite of Chimpanzee *TCP10L* has two helixs and two beta-sheets in secondary protein structures, human *TCP10L* showed one helix and 4 beta-sheets structure on the C-terminal region. These dramatic changes on the secondary structure might influence on three dimensional structures and also in their biological activities of the *TCP10L* between two species.

Implication of *TCP10L* changes in physiological function of testis and liver

In this study, we show that human *TCP10L* has 16 base pair difference in genomic DNA compare to chimpanzee ortholog. As a consequence, frame shift mutation occurs in CDS of human *TCP10L* and affects to human express as longer protein of 21 amino acid residues than that of chimpanzee. Our further analysis found that the changes of primary structure of human *TCP10L* resulted in dramatic changes in secondary structure prediction. Functional study showed that human *TCP10L* is deduced to be 215 amino acids polypeptide and a typical leucine zipper motif exists in the middle of the peptide (Chen *et al.,* 2003). The leucine zipper is DNA binding motif and indispensable to the biological functions of these transcription factors through dimerization. Based on these facts, the structural changes in the C-terminal region of *TCP10L* protein may not affect the DNA binding function of this transcription repressor. More likely the C-terminal changes would affect some other function of human *TCP10L* such as interacting potential to other transcription factors or stability of the protein. It has been known that the *TCP10L* interact with MAD as a transcription factor (Jiang *et al.,* 2004). More recently, Yu et al. reported that *TCP10L* is expressed in nucleus of spermatogenic cell and bind to death associated protein kinase-3 (DAPK-3) (Yu *et al.,* 2005). Base on this, they suggested that *TCP10L* might play crucially important roles in spermatogenesis through its interaction with DAPK-3. However, it is not clear that C-terminal region of *TCP10L* is responsible for interaction with DAP-3. In this point, our finding through comparative study of chimpanzee and human *TCP10L* can supply new approach method to understand the roles of *TCP10L* in spermatogenesis.

The human *TCP10L* gene is specifically expressed in the liver and testis. In the female, *TCP10L* only expresses in the liver. In previous report, Enard *et al.* showed that gene expression patterns in human liver are more similar to those of the chimpanzee than to those of macaques (Enard *et al.* 2002). These results coincide with the evolutionary relationship of the species suggesting that gene expression patterns in liver reflect the evolutionary relationship in the primates. The liver is an essential organ playing a central role in metabolism and detoxification of poisonous substances. Therefore, these changes in liver physiology might be required to adapt new environment such as changes of food and xenobiotics. The various genes participating in these metabolic reactions are regulated on transcription revel by many transcription factors (Schiaffonati *et al.,* 1997). The implication of the changes in transcription factors to human and chimpanzee evolution are unclear. However, these changes in *TCP10L* gene might influence gene expression profiles in liver and testis and

subsequently influence physiological changes required in primate evolution. Further molecular biological studies on the *TCP10L* may help us gain insights into the molecular implication of the changes in the physiological functions of liver and primate evolution.

Acknowledgements

This work was supported by the Ministry of Science and Technology, and the Korea Research Institute of Bioscience and Biotechnology, Korea

References

Amati, B., Brooks, M.W., Levy, N., Littlewood, T.D., Evan, G.I., and Land, H. (1993). Oncogenic activity of the c-Myc protein requires dimerization with Max. *Cell* 72, 233-245.

Britten, R.J. (2002). Divergence between samples of chimpanzee and human DNA sequences is 5%, counting indels. *Proc. Natl. Acad. Sci. USA* 99, 13633-13635.

Chen, Z., Yu, L., Wu, H., Yu, J. Q., Zhang, L. S., Jiang, D. J., Ma, L., Li, D., and Zhao, S. (2003). Identification of a novel liver-specific expressed gene, TCP10L, encoding a human leucine zipper protein with transcription inhibition activity. J. *Hum. Genet.* 48, 556-563.

Cheryy, L.M., Case, S.M., and Wilson, A.C. (1978). Frog perspective on the morphological difference between humans and chimpanzees. *Science* 200, 209-211.

Enard, W., Khaitovich, P., Klose, J., Zollner, S., Heissig, F., Giavalisco, P., Nieselt-Struwe, K., Muchmore, E., Varki, A., Ravid, R., Doxiadis, G.M., Bontrop, R.E., and Paabo, S. (2002). Intra- and interspecific variation in primate gene expression patterns. *Science* 296, 340-343.

Florea, L., Hartzell, G., Zhang, Z., Rubin, G.M., and Miller, W. (1998). A computer program for aligning a cDNA sequence with a genomic DNA sequence. *Genome Res.* 8, 967-794.

Jiang, D.J., Yu, H.X., Sai-Yin, H., Guo, Z.K., Wang, X., Ma, L.J., Chen, Z., Zhao, S.Y., and Yu, L. (2004). Human liver specific transcriptional factor TCP10L binds to MDA4. *J. Biochem. Mol. Biol.* 37, 402-407.

McGuffin, L.J., Bryson, K., and Jones, D.T. (2000). The PSIPRED protein structure prediction server. *Bioinformatics* 16, 404-405.

Perez, S., Vial, E., van Dam, H., and Castellazzi, M. (2001). Transcription factor ATF3 partially transforms chick embryo fibroblast by promoting growth factor-independent proliferation. *Oncogene* 20, 1135-1141.

Schiaffonati, L. and Tiberio, L. (1997). Gene expression in liver after toxic injury: analysis of heat shock response and oxidative stress-inducible genes. *Liver* 17, 183-191.

Sellers, J.W. and Struhl, K. (1989). Change fos oncoprotein to a jun-independent DNA binding protein with GCN4 dimerization specificity by swapping "leucine zipper". *Nature* 341, 74-76.

Watanabe, H., Fujiyama, A., Hattori, M., Taylor, T.D., Toyoda, A., Kuroki, Y, Noguchi, H., BenKahla, A., Lehrach, H., Sudbrak, R., Taenzer, S., Galgoczy, P., Platzer, M., Scharfe, M., Nordsiek, G., Blocker, H., Hellmann, I., Khaitovich, P., Paabo, S., Reinhardt, R., Zheng, H.J., Zhang, X.L., Zhu, G.F., Wang, B.F., Fu, G., Ren, S.X., Zhao, G.P., Chen, Z., Lee, Y.S., Cheong, J.E., Choi, S.H., Wu, K.M., Liu, T.T., Hsiao, K.-J., Kim, C.G., Oota, S., Kitano, T., Kohara, Y.,Saitou, N., Tsai, S.-F., Park, H.S., Wang, S.-Y., Yaspo, M.-L., and Sakaki, Y. (2004). DNA sequence and comparative analysis of chimpanzee chromosome 22. *Nature* 429, 382-388.

Yu, H., Jiang, D., Guo, Z., Saiyin, H., Guo, J., Wang, X., and Yu, L. (2005). TCP1-L is expressed specifically in spermatogenic cells and binds to death associated protein kinase-3. *Int. J. Andrology* 28, 163-170.