

# ANALYSIS OF STK11/LKB1 GENE USING BIOINFORMATICS TOOLS

*Fuat Dilmeç<sup>1</sup>, Lokman Varışlı<sup>2</sup>, Abdullah Özgönül<sup>3</sup>, Osman Cen<sup>2</sup>*

Harran University University, Medicine Faculty, Departments of Medical Biology<sup>1</sup> and General Surgery<sup>3</sup>, Art and Science Faculty, Department of Biology<sup>2</sup>, Şanlıurfa, Turkey

---

**Aim:** The STK11 gene, also known as LKB1, encodes for a serine/threonine kinase with growth-suppressing activity, such as inhibition of cell cycle progression, cell growth retardation, apoptotic cell death, and cell polarity control. This study aimed to investigate some properties of the STK11 gene and its product, such as the homologous protein sequences in different species, the common transcription factor binding sites on their promoters, their phylogenetic relationship, their catalytic domains (S\_TKc), and their expression profiles.

**Methods:** We investigated the homology, conserved domain, promoter and expression profiles of the STK11/LKB1 genes in various species using bioinformatics approaches.

**Results:** Our results revealed that STK11/LKB1 molecules are conserved among all organisms investigated. The kinase domain (S\_TKc) of human STK11/LKB1 gene is closest to those of Pan troglodytes, Macaca mulatta and Macaca fascicularis. In contrast, the most diverse to the human S\_TKc domain is that of Bos taurus. With the multiple alignment strategy, protein and domain sequences of M. fascicularis and B. Taurus are predicted to have a truncation. The comparative screening of the promoters demonstrated that STK11/LKB1 genes do not seem to have any common conserved transcription factor binding sites.

**Conclusion:** This study demonstrated STK11 molecules in various species are well conserved throughout evolution. Comparative screening of the promoter sequences of the human STK11 and its homologues found in the NCBI database revealed that there was no any common transcription factor binding sites. Phylogenetic trees constructed using the neighbor-joining method (NJ) revealed a close evolutionary relationship of S\_TKc in various species.

**Key words:** LKB1, STK11, Bioinformatics, Comparative genomics, in silico biology

*Eur J Gen Med 2007;4(4):180-185*

---

## INTRODUCTION

The gene STK11, also known as LKB1, is considered to be a tumor suppressor gene that is mutated in the Peutz-Jeghers cancer syndrome (1). It encodes a serine/threonine kinase and is mapped to human 19p13.3 chromosome. Its protein product shows both nuclear and cytoplasmic localization. It is speculated that subcellular distribution of STK11/LKB1 is regulated by the balance between its importation into the nucleus and retention within the cytoplasm, and that the cytoplasmic retention ability is necessary for LKB1 to fulfill its normal function (2-5).

Human STK11/LKB1 shows strong homology to the cytoplasmic serine/threonine kinases of several organisms, including those of mouse and xenopus (6, 7). There is a 92.5% similarity between mouse and human STK11/LKB1 genes, and 97.5% in their core kinase

domain. Similarly, there is 84.5% and 96.2% similarity between the core kinase domains of LKB1 of mouse and xenopus, respectively (8,9).

It has been shown by RNase-protection assay that murine STK11/LKB1 mRNA is expressed in all tissues and cell lines (10). Strong expression of STK11/LKB1 in human fetal small intestine and stomach suggests that STK11/LKB1 has a regulatory function in the development and maintenance of the embryonic gastrointestinal tract (11).

In this study, we aimed to analyze the STK11/LKB1 genes in different species in silico. Specifically, their S\_TKc kinase domains, the transcription factor binding sites on their promoters, the tissue expression profile, homology level and phylogenetic tree among vertebrate STK11/LKB1 genes using bioinformatics tools.

---

**Correspondence:** Fuat DILMEC  
Harran University, Medicine Faculty, Department of  
Medical Biology, 63300 Sanliurfa-Turkey  
Tel: 904143128456/24 06, Fax: 904143139615  
E-mail: fdilmec@harran.edu.tr

**Table 1. BLASTp results of vertebrate STK11/LKB1 molecules and their homology**

Species	Protein Name	AC Number	Identity (%)
Homo sapiens	STK11	NP000446	100
Macaca mulatta	STK11	XP001093806	99
Pan troglodytes	STK11	XP524028	99
Macaca fascicularis	Hyp. Protein	BAE73059	99
Canis familiaris	STK11	XP542206	92
Gallus gallus	STK11	NP001039298	89
Rattus norvegicus	STK11	XP234900	89
Mus musculus	STK11	NP035622	89
Raja erinacea	STK11	AAL92113	82
Xenopus tropicalis	STK11	NP001016533	82
Danio rerio	LOC550537	NP001017839	81
Xenopus laevis	XEEK1	AAC59904	83
Bos taurus	STK11	XP593229	69

## MATERIALS AND METHODS

### Homology search

The search for homologous protein sequence to human STK11/LKB1 was carried out using the BLASTp program (12, 13) at NCBI (<http://www.ncbi.nlm.nih.gov>) using human STK11/LKB1 amino acid sequence (GI: 4507271) as query against the SwissProt protein databases. Full protein and kinase domain (S\_TKc) sequences of human and other species were downloaded and then aligned using the ClustalW (14) program at EBI (<http://www.ebi.ac.uk>).

### Promoter Analysis

We used Genomatix software (<http://www.genomatix.de>) for analysis of STK11/LKB1 gene promoters in various species. These nucleotide sequences were downloaded and then were aligned using the ClustalW program. Then common transcription factor binding sites were searched with the Dialign TF program in Genomatix software for all of STK11/LKB1 promoters present in the database.

### Evolutionary Analysis

We used amino acid sequences of S\_TKc domains to construct phylogenetic trees using the neighbor-joining method (NJ) with Jones-Taylor-Thomton (JTT) distances. NJ searches were conducted by using MEGA3 (15) and 500 bootstrap replicates were assessed for the reliability of internal branches; sites with gaps were ignored in this analysis.

### In silico Expression Analysis

The DigiNorthern database (16) was used

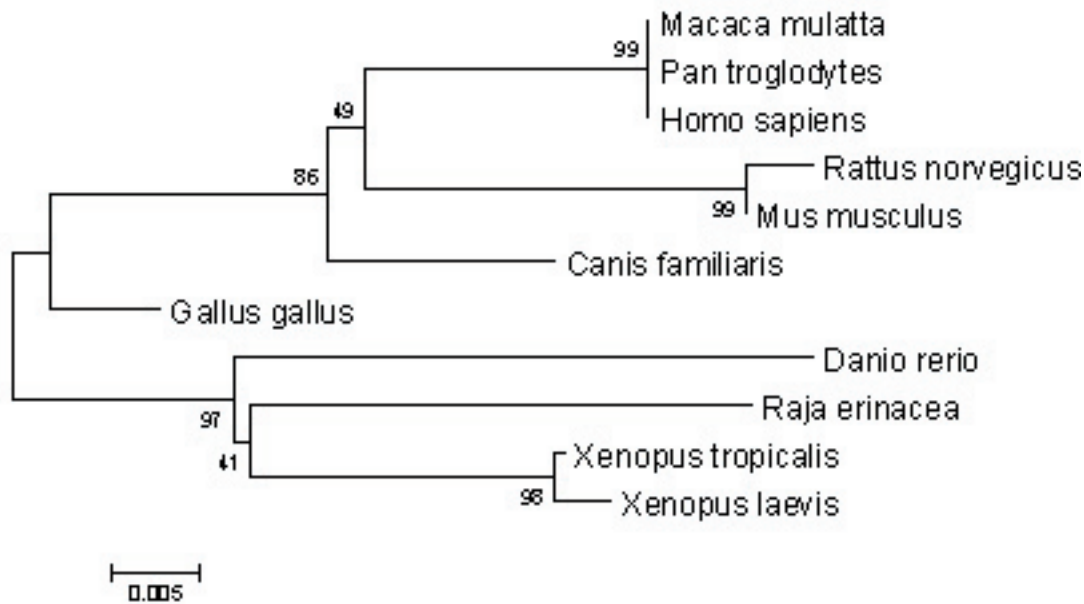
to analyze the expression of STK11/LKB1 mRNAs based on EST data. The DigiNorthern collects all ESTs for a query gene and categorizes these ESTs based on the types of tissues and their histological status. Pairwise comparisons of relative frequencies were performed with the Fisher's exact test using SPSS 11.0 for Windows.

## RESULTS

### Homology Search

BLASTp results revealed that STK11/LKB1 molecule is found in various species (Table 1). The homology search indicated that the STK11/LKB1 sequences of *P. troglodytes* (99%), *M. mulatta* (99%) and hypothetical protein of *M. fascicularis* (99%), respectively, have the highest homology to that of human. In contrast, the one with the lowest homology human STK11/LKB1 protein was that of *B. taurus* (69%) (Table 1).

Multiple alignment results of human STK11 and its homologous revealed that this molecule is yet uncharacterized in two species, which *B. taurus* and *M. fascicularis*, probably due their possible truncation. The sequence of *M. fascicularis* has a truncation at its N-terminus while that of *B. taurus* has a truncation possible of an exon in the C-terminus of its kinase domain. The rest of the molecule of the *B. taurus*'s molecule show a very diverse pattern compared to that of the other homologous. Ignoring the molecules of these two foregoing species, the S\_TKc domains of the other species are very well conserved through evolution.



**Figure 1. Phylogenetic tree of S\_TKc domains of vertebrate STK11/LKB1. Phylogenetic trees were constructed by the MEGA3 program. Species names are indicated on the figure. Branch lengths indicate evolutionary relationship.**

#### **Promoter Analysis**

We found by the database search that the orthologous STK11/LKB1 gene promoters do not include any common transcription factor binding sites (TFBs) among *H. sapiens*, *B. taurus*, *C. familiaris*, *R. norvegicus*, *M. musculus* and *G. gallus* in Database of Genomatix software. However, we observed that the similarity (value 1.000) and the number of identical nucleic acids (in % of short sequence segments) was 89% between the STK11/LKB1 promoters of *M. musculus* and *R. norvegicus* for each pairwise alignment; however, this does not necessarily mean that the two sequences are identical.

#### **Evolutionary Analysis**

From the phylogenetic trees constructed by MEGA3 we found that STK11/LKB1 S\_TKc domains are conserved among all organisms investigated. We have shown that the S\_TKc domains of *H. sapiens*, *P. troglodytes* and *M. mulatta* are more closely grouped (Figure 1). When we constructed phylogenetic tree, we ignored domain sequences of *M. fascicularis* and *B. taurus* due to their high diversity caused by possible truncation.

#### **In Silico Expression Analysis**

The distribution of STK11/LKB1 in the cDNA library database was analyzed using the DigiNorthern program. In columns 2 and

3 of Table 2, relative frequencies of STK11/LKB1 in the cDNA libraries from normal and tumor tissues respectively are shown both as absolute numbers as well as normalized values per 106 cDNAs (in parenthesis). Its normal as well as cancerous tissue expression profile was compared and the significance in its expression pattern was accessed with the Fisher's exact test (p value 0.05) (Table 2). The STK11/LKB1 gene seems to be expressed at a low level or not at all in some tissues while highly expressed in some others. The tissue distribution and differential expression pattern in normal and cancerous human tissues displayed somehow different frequencies. The expression of STK11/LKB1 gene in some of the normal human tissues, such as bone marrow, colon, muscle, nervous, ovary, and pancreas, was not detected. In contrast, its expression seems to be lost in some of the cancerous tissues such as cervix and placenta. Compared to the normal tissues, its expression is significantly increased in cancerous muscle tissues and decreased in cancerous skin, prostate and kidney tissues (Table 2).

Relative frequencies are normalized per 106 cDNAs; p-values are for comparison of relative frequencies of STK11/LKB1 in normal versus tumor tissues, using the Fisher's exact test. The data for tissues with significant or suggestive significant higher or

**Table 2. The distribution of relative frequencies of STK11/LKB1 cDNA in cDNA library database**

Tissue/organ type	Normal	Cancer	p-value
Bone	0/7929 (0)	4/45730 (87)	1.000
Bone marrow	0/21187 (0)	1/29586 (34)	1.000
Brain	13/257019 (51)	16/201219 (80)	0.263
Cartilage	0/13369 (0)	3/39893 (75)	1.000
Cerebrum	1/103896 (10)	0/0 (0)	0.578
Cervix	1/1157 (864)	0/44671 (0)	0.025
Colon	0/28085 (0)	13/220946 (59)	0.385
Eye	9/85966 (105)	4/49827 (80)	0.780
Genitourinary	0/1687 (0)	3/39698 (76)	1.000
Germ cell	0/0 (0)	21/56605 (371)	>1.000
Head and neck	1/55508 (18)	2/107902 (19)	1.000
Kidney	10/74917 (133)	2/96375 (21)	0.007
Lung	5/129822 (39)	20/207630 (96)	0.065
Lymph node	6/97096 (62)	6/54341 (110)	0.370
Lymphoreticular	0/15679 (0)	4/56791 (70)	0.584
Mammary gland	2/71315 (28)	1/124006 (8)	0.303
Muscle	0/90941 (0)	4/45799 (87)	0.013
Nervous	0/15506 (0)	2/63270 (32)	1.000
Ovary	0/11587 (0)	6/109344 (55)	1.000
Pancreas	0/8656 (0)	7/84351 (83)	1.000
Pancreatic islet	17/95891 (177)	0/0 (0)	>1.000
Peripheral nervous system	1/30154 (33)	0/1220 (0)	1.000
Placenta	1/248276 (4)	0/43818 (0)	1.000
Pooled tissue	11/373366 (29)	6/55060 (109)	0.016
Prostate	16/82545 (194)	3/81283 (37)	0.004
Salivary gland	0/414 (0)	5/20747 (241)	1.000
Skin	4/49729 (80)	2/137037 (15)	0.047
Stem cell	9/184378 (49)	0/0 (0)	>1.000
Stomach	5/26066 (192)	2/140405 (14)	0.001
Testis	7/122158 (57)	2/44649 (45)	1.000
Thymus	2/5359 (373)	0/201 (0)	1.000
Uncharacterized tissue	2/88784 (23)	6/105216 (57)	0.303
Uterus	3/36080 (83)	7/163186 (43)	0.402
Whole body	5/73648 (68)	0/0 (0)	>1.000
Total No. of ESTs Found	134/2508170	152/2470806	0.237

lower frequency of STK11/LKB1 in the tumor and normal tissues are shown in bold.

a: computed only for a 2x2 table

b: Two messages (50.0%) have expected count less than 5. The minimum expected count is 0.50. (SPSS)

## DISCUSSION

Human LKB1 gene (also called STK11) is located on chromosome 19p13.3 and recently identified as a tumor suppressor gene in which its mutation can lead to Peutz-Jeghers syndrome (PJS), characterized by gastrointestinal polyps and cancers of different organ systems. The STK11/LKB protein is present in both the cytoplasm and nucleus of living cells and translocates to

mitochondria during apoptosis. The defect in PJS may result directly in changes in the STK11/LKB1 gene expression (7).

Human, Mouse and Xenopus LKB1s (92.5% homology) and their core kinase domains (96.2% homology) are very similar. This high degree of conservation, in particular in regions outside the core kinase domain, strongly suggests that mouse, human and xenopus LKB1s are orthologues (6-8). Likewise, our BLASTp results indicate that LKB1 is found in various species of vertebrates and these molecules have 69-99% conservation degree in the total amino acid sequences (Table 1). The human LKB1 molecule has highest homology to those of *P. troglodytes* and *M. mulatta* (99%) and lowest

homology to that of *B. taurus* (69%). So, these results indicate that the STK11/LKB1 gene has been evolutionary well conserved (Table 1).

We also examined the phylogenetic trees of S\_TKc domain of LKB1 in different species using MEGA3 program. We observed that human S\_TKc domain protein shows closest homology to those of *M. mulatta* and *P. troglodytes*. Likewise, same figures show that S\_TKc domains of *R. erinacea*, *X. tropicalis*, *X. laevis* and *D. rerio* are most diverse from S\_TKc domain of *H. sapiens*.

The expression of STK11/LKB1 in different tissues was analyzed using the DigiNorthern program (Table 2). Its expression patterns in normal and cancer tissues displayed somehow different frequencies in human. In some tissues, such as muscle, nervous, ovary, and pancreas, their expressions seem to be expressed at a very low level or not at all. In cancerous cervix and placenta its expression is not observed while decreased in cancerous skin, prostate, and kidney tissues. In contrast, its expression is significantly increased in cancerous muscle tissues (Table 2). Its increased frequency in pooled cancerous tissue may be attributed to one of the tissues present in the pool.

We suspect that LKB1 the change in its expression pattern in different tissues may be related to the role for pathogenesis of some sporadic cancers. The availability of the comprehensive data generated by high-throughput functional genomic approaches, mainly expressed sequence tag (EST) and serial analysis of gene expression (SAGE), provides the feasibility to study gene expression through in silico analysis (17). Shen et al demonstrated that LKB1 protein expression levels were significantly associated with histological grade, tumor size, progesterone receptor status, and presence of the lymph node metastasis (18). It has been shown that the expression is higher in fetal than in adult tissues (19). Expression also appears to be higher in many malignant tumors than in normal tissues or benign lesions, although some cancers might have lost STK11/LKB1 expression, quite possibly as part of the process of tumorigenesis (18). It has been shown that STK11/LKB1 directly phosphorylates Thr-172 of AMPK (AMP-activated protein kinase) in vitro and activates its kinase activity. This suggests that LKB1/AMPK signaling plays a role in cell survival by protecting cells from apoptosis, specifically in response to agents that increase the cellular

AMP/ATP ratio (20).

We used Dialign TF program in Genomatix software for predicting transcription factor binding sites (transcriptional elements) of all orthologous STK11/LKB1 promoters that present in the database. Dialign TF results revealed that STK11/LKB1 orthologous promoters had no common conserved transcriptional elements. The conservation of transcriptional elements in promoter sequences may provide further evidence in support of functional conservation (5, 21-23). However, the element in the promoters or their vicinity may be more mobile than the genes themselves. Our results indicate that the binding sites of different transcription factors might have located on different parts of the promoter or promoter vicinity in various species.

Basic bioinformatics techniques are powerful tools in terms of leading to the discoveries and analysis of novel genes (24). Recently we identified and further characterized two novel genes using bioinformatics tools (25). Even though the results from bioinformatics studies are very helpful in directing and designing the experiments, they need to be supported and confirmed by further experimentation.

## REFERENCES

1. Hemminki A, Tomlinson I, Markie D et al. Localization of a susceptibility locus for Peutz-Jeghers syndrome to 19p using comparative genomic hybridization and targeted linkage analysis. *Nat Genet* 1997;15(1):87-90
2. Nezu J, Oku A, Shimane M. Loss of cytoplasmic retention ability of mutant LKB1 found in Peutz-Jeghers syndrome patients. *Biochem Biophys Res Commun* 1999;261:750-5
3. Rowan A, Bataille V, MacKie R et al. Somatic Mutations in the Peutz-Jeghers (LKB1/STKII) Gene in Sporadic Malignant Melanomas. *J Invest Dermatol* 1999;112:509-11
4. Tiainen M, Ylikorkala A, Makela TP. Growth suppression by Lkb1 is mediated by a G(1) cell cycle arrest. *Proc Natl Acad Sci USA* 1999;96(16):9248-51
5. Hearle NC, Tomlinson I, Lim W et al. Sequence changes in predicted promoter elements of STK11/LKB1 are unlikely to contribute to Peutz-Jeghers syndrome. *BMC Genomics* 2005; 6(38):1-5
6. Su JY, Erikson E, Maller JL. Cloning and characterization of a novel serine/



- threonine protein kinase expressed in early *Xenopus* embryos. *J Biol Chem* 1996;271(24):14430-7
7. Smith DP, Spicer J, Smith A, Swift S, Ashworth A. The mouse Peutz-Jeghers syndrome gene *Lkb1* encodes a nuclear protein kinase. *Hum Mol Genet* 1999;8(8):1479-85
  8. Waterson RH, Lindblad-Toh K, Birney E et al. Initial sequencing and comparative analysis of the mouse genome. *Nature* 2002;420(6915):560-2
  9. Boudeau J, Sapkota G, Alessi DR. LKB1, a protein kinase regulating cell proliferation and polarity. *FEBS Lett* 2003;546(1):159-65
  10. Collins SP, Reoma JL, Gamm DM, Uhler MD. LKB1, a novel serine/threonine protein kinase and potential tumour suppressor, is phosphorylated by cAMP-dependent protein kinase (PKA) and prenylated in vivo. *Biochem J* 2000;345(1):673-80
  11. Yee NS, Furth EE, Pack M. Clinicopathologic and Molecular Features of Pancreatic Adenocarcinoma Associated with Peutz-Jeghers Syndrome. *Cancer Biol Ther* 2003;2(1):38-47
  12. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. *J Mol Biol* 1990;215(3):403-10
  13. Gissen P, Johnson CA, Gentle D et al. Comparative evolutionary analysis of VPS33 homologues: genetic and functional insights. *Hum Mol Genet* 2005;14(10):1261-70
  14. Thompson JD, Higgins DG, Gibson TJ. CLUSTALW: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 1994;22(22):4673-80
  15. Kumar S, Tamura K, and Nei M. MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinf* 2004;5(2):150-63
  16. Wang J, Liang P. DigiNorthern, digital expression analysis of query genes based on ESTs. *Bioinformatics* 2003;19(5):653-4
  17. Lash AE, Tolstoshev CM, Wagner L et al. SAGEmap: a public gene expression resource. *Genome Res* 2000;10(7):1051-60
  18. Shen Z, Wen XF, Lan F, Shen ZZ, Shao ZM. The Tumor Suppressor Gene LKB1 is Associated with Prognosis in Human Breast Carcinoma. *Clin Cancer Res* 2002;8(7):2085-90
  19. Rowan A, Churchman M, Jefferey R, Hanby A, Poulson R, Tomlinson I. In situ analysis of LKB1/STK11 mRNA expression in human normal tissues and tumours. *J Pathol* 2000;192(2):203-6
  20. Shaw RJ, Kosmatka M, Bardeesy N et al. The tumor suppressor LKB1 kinase directly activates AMP-activated kinase and regulates apoptosis in response to energy stress. *Proc Natl Acad Sci USA* 2004;101(10):3329-35
  21. Vanpoucke G, Goossens S, De Craene B, Gilbert B, van Roy F, Berx G. GATA-4 and MEF2C transcription factors control the tissue specific expression of the alphaT-catenin gene CTNNA3. *Nucleic Acids Res* 2004;32(14):4155-65
  22. Doan LL, Porter SD, Duan Z et al. Targeted transcriptional repression of *Gfi1* by *GFI1* and *GFI1B* in lymphoid cells. *Nucleic Acids Res* 2004;32(8):2508-19
  23. Cartharius K, Frech K, Grote K et al. MatInspector and beyond: promoter analysis based on transcription factor binding sites. *Bioinformatics* 2005;21(13):2933-42
  24. Hu Z, Chen K, Wang L, Yao Q. Identification and characterization of *Bombyx mori* eIF5A gene through bioinformatics approaches. *In Silico Biol* 2005;5(5-6):573-80
  25. Varisli L, Cen O. Identification and Characterization of Rat GMDS Gene by Using Bioinformatics Tools. *Turk J Biochem* 2005;30(4):306-9