# *LY-6K* gene: A novel molecular marker for human breast cancer

JI WOO LEE[1], YEON SU LEE[2], KYUNG HYUN YOO[1], KEE HO LEE[3], KYUNGHEE PARK[2], TAEJIN AHN[2], CHRISTOPHER KO[2] and JONG HOON PARK[1]

[1]Department of Biological Science, Sookmyung Women's University, Chungpa-dong, Yongsan-gu, Seoul 140-742; [2]Bio lab, Samsung Advanced Institute of Technology, San14 Nonseo-ri, Giheung-eup Yongin-si, Kyoungki-do 449-712; [3]Laboratory of Molecular Oncology, Korea Institute of Radiological and Medical Sciences, Seoul 139-706, Korea

Received February 10, 2006; Accepted May 8, 2006

**Abstract.** A full-length cDNA was identified using one STS sequence containing an SNP (single nucleotide polymorphism) derived from genomic DNAs of breast cancer patients using a variety of bioinformatics tools. The cDNA encodes LY-6K, a novel member protein of the Ly-6/uPAR superfamily. It has been annotated as a target antigen for the HNSCC (head- and neck squamous cell carcinoma). We isolated the *LY-6K* gene from genomic DNAs obtained from breast cancer patients through large scale, case-control-screening. We performed Northern blot hybridization and semi-quantitative RT-PCR on a human multiple-tissue mRNA blot from several breast cancer patients. We investigated the expression level of the *LY-6K* gene in human breast cancer, and compared this to expression in human normal breast tissue. We found that *LY-6K* was more highly expressed in the mRNA of breast tumors compared to its expression in normal breast tissue. These results suggest that *LY-6K* is not only a target antigen for HNSCC but also a significant new molecular marker for diagnosis and gene therapy in patients with breast cancer.

## Introduction

The knowledge of human genetic variation derived from the human genome sequence makes feasible a polygenic approach to disease prevention. A major goal of the application of this knowledge is to identify individuals who are susceptible to disease based on their genotype and to prevent the development of disease by targeting interventions to those at risk (1).

Many studies have been performed to identify candidate genes that contribute to the risk for the development of sporadic breast cancer (2). The identified genes including: BRCA1 (3), BRCA2 (4), TP53 (5), ATM (6) and PTEN (7) are well known to contribute to familial cases of breast cancer. However, these high penetrance germ line mutations are responsible for less than 10% of all breast cancer cases. Genetic variation is estimated to contribute approximately 25% to the population risk for breast cancer; this is likely to be accounted for by a large number of yet undiscovered common, low penetrance alleles (1,8). It is possible that these more common low penetrance markers may serve as a useful tool for diagnosis and prognosis in the general population (9).

The Ly-6/uPAR superfamily of receptor and secreted proteins, is a group of lymphocyte antigens that have a carboxyl-terminal consensus sequence motif CCXXXXCN and one or several repeats of the Ly-6/uPAR domain, which is defined by a distinct disulfide bonding pattern between eight or 10 cysteine residues; (10) these molecules have a characteristic spacing pattern and attach to the cell surface by a glycosylphosphatidylinositol (GPI) anchor (11). The Ly-6/uPAR superfamily has two subfamilies: one represents the transmembrane protein, as mentioned above, and the other is a secretory protein without the GPI anchor (11). The identified human members of this group include: CD59, LY6H, LY6C, uPAR (urokinase-type plasminogen activator receptor), GML, PSCA (prostate stem cell antigen), RIG-E, SP-10, E48 (hLY6D), SLURP1 and SLURP2 (12). The *LY-6K* gene was identified two years ago and has been proposed as a potential target antigen for diagnosis and therapy of HNSCC (head- and neck squamous cell carcinoma) (13).

Using STS containing SNPs derived from genomic DNAs of breast cancer patients, we tried to identify the full-length cDNA. To investigate whether the human Ly-6 antigen (*LY-6K*) gene might be suitable as a novel molecular marker for diagnosis and therapy of human breast cancer, breast tumor cell-specific expression patterns were determined by Northern blotting and semi-quantitative RT-PCR.

## Materials and methods

*EST database searches for identification of cDNAs and genomic DNA clones*. To confirm EST, cDNAs or genomic

*Correspondence to*: Dr Jong Hoon Park, Department of Biological Science, Sookmyung Women's University, Chungpa Dong 2Ka, Yong San Gu, Seoul 140-742, Korea
E-mail: parkjh@sookmyung.ac.kr

*Key words:* human breast cancer, STS, Ly-6/uPAR superfamily, Northern blot, RT-PCR

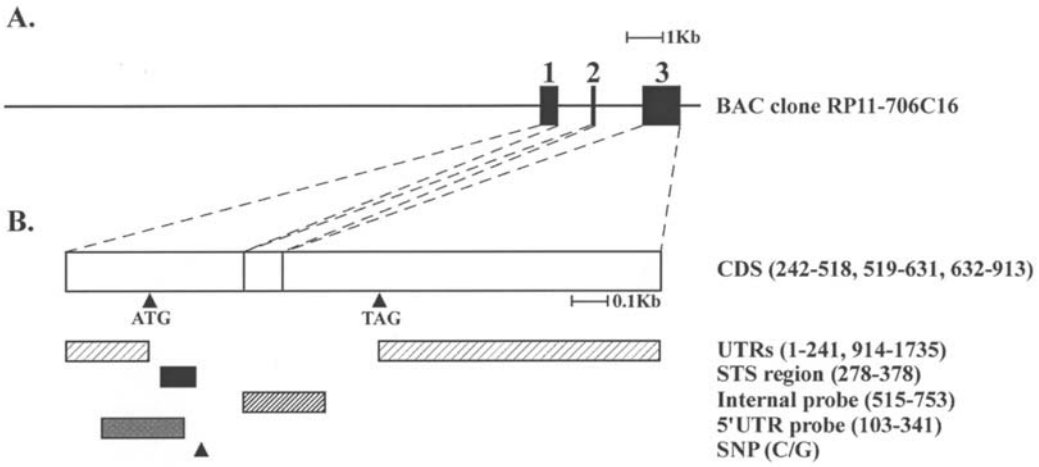


Figure 1. The human *LY-6K* gene organization. (A), Genomic locus of BAC clone (RP11-706C16). Exons are numbered and shown as black boxes. (B), The mRNA transcript of human *LY-6K* gene.

DNA matched with interesting STS (sequence tagged site, about 200 bp) were identified; the DNA sequence was compared with the redundant nucleotide and protein sequence database (Genbank, EMBL, DDBJ) using BLASTN and BLASTP. One mRNA transcript (GI:40068041) and several genomic DNA clones were isolated. *Homo sapiens* cDNA, HSJ001348, was identified as a candidate for encoding the new ly6 molecule. A genomic DNA BAC clone, RP11-706C16, was obtained from the BacPac resource and one cDNA I.M.A.G.E clone, BC001291, was obtained from Open Biosystems.

To identify the mRNA transcript start site (TSS), we used the DBTSS release 3.0 software (http://dbtss.hgc.jp/index.html). To identify exon-intron boundaries, expressed DNA and genomic sequences for the gene were aligned with each other using SIM4, a similarity-based tool (http://pbil.univ-lyon1.fr/sim4.php).

*Northern blot analysis*. Northern blot hybridization was performed with a probe using the internal PCR products of the BC001291 cDNA clone (239 bp-length, position 515-753). The internal PCR primer sequence of *LY-6K* is 5'-ACGGAC GAGGGTGACAATAG-3' and 5'-AGGTAAAAGAAGGG CATGGG-3'. The *GAPDH* gene probe was used as a control (5'-TGCACCACCAACTGCTTAGC-3' and 5'-GGCATGGA CTGTGGTCATGAG-3'). PCR products were purified with the gel extraction kit (D4001, Zymo research).

To detect the level of expression, in both normal and cancer breast tissue, a Northern Bound Human Single Tumor mRNA Blot (Sigma, N7159) was used. Northern blot hybridization was performed with a probe that is the internal PCR product labeled by [$\alpha$-$^{32}$P]-dCTP (NEN) using the Random Primed DNA Labeling Kit (1004760, Roche). Prehybridization and hybridization were performed using the protocol received from the Sigma Corp. *18S rRNA* was used as a control gene for normalization. The blot was exposed to X-ray film at -70˚C for 3-4 days according to the manufacturer's instructions as described in their manual (#PT-1200-1, Clontech Laboratories, Inc.). The bands were visualized and quantified as described above using the Las-3000 image analyzer (Life Science, Fujifilm, Japan) and Multigauge program (Fujifilm, Japan).

*Semi-quantitative RT-PCR*. Total-RNAs were extracted using Trizol (Invitrogen, USA) from non-lesion breast cancer tissue as well as lesion tissue from two human breast cancer patients. The *18S rRNA* gene (5'-GTAACCCGTTGAACCCCATT-3' and 5'-CCACCAATCGGTAGTAGCG) was used as a positive control. The *LY-6K* gene was amplified using primers (5'-GC GGCCGCGTTATCAGAGGTGAGCCCGT-3' and 5'-GAAC CTTCGCAGCCTCCGCG CG-3'). This primer set amplified the region (position 103-341) containing 5'-UTR. The PCR conditions for *LY-6K* were as follows. The initial denaturing phase for 5 min at 95˚C was followed by a 28-cycle amplification phase consisting of denaturation at 95˚C for 40 sec, annealing at 58˚C for 40 sec and elongation at 72˚C for 40 sec. Amplification was terminated at 72˚C for 10 min. The PCR conditions for the *18S rRNA* was the same as those for the *LY-6K* but the annealing temperature was 59˚C.

*Subcellular localization*. The PCR products containing the complete open reading frame of the *LY-6K* gene, with *Eco*RI and *BamH*I Linker, was purified and inserted into the *Eco*RI/*BamH*I restriction enzyme sites of pEGFP-N1 (Clontech). The full-length CDS region was obtained from the MDA-MB-231 cell line from RT-PCR. The primer sequences are 5'-CGGAATTCTGATGAGGCTCCAAAGACCCCGA-3' and 5'-CGGGATCCGCAGACAGGCTGAGGCCGGCTGC-3' containing the restriction enzyme linker. One microgram of pEGFP-N1-*LY-6K* and pEGFP-N1 expression plasmid DNA was transfected into 293T cells, respectively.

The transfection experiments were carried out in 6-well plates with a coverslip for each well and grown for 36-48 h, at 37˚C. Cells were washed with PBS twice for 5 min, and then fixed with cold fixing solution (4% paraformaldehyde in PBS). For nuclear staining, the fixed cells were incubated with DAPI (0.1 mg/ml, Sigma) for 30 min at room temperature. Fluorescent signals were revealed under the fluorescent microscope and images were recorded with a digital camera.

## Results

*Isolation of the full-length cDNA of LY-6K*. We have selected a few of the identified significant SNP markers from a case-

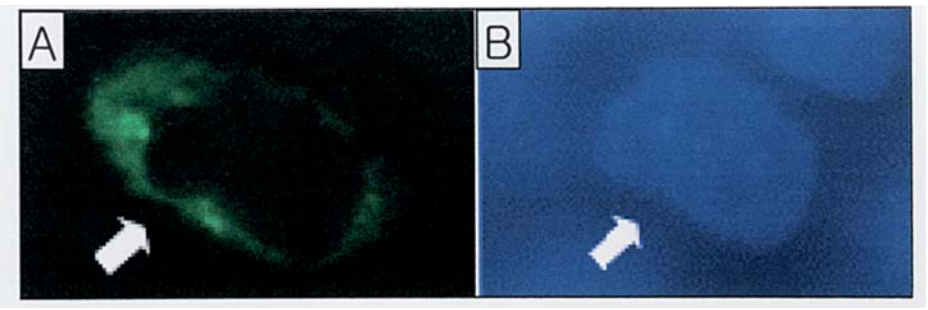


Figure 2. Subcellular localization of the *LY-6K* gene. (A), GFP fusion *LY-6K* expression in the plasma membrane of 293T cell lines. (B), DAPI staining of the nucleus.

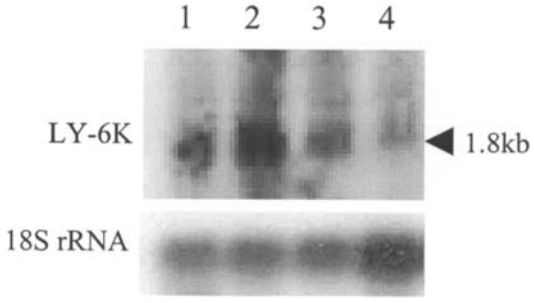


Figure 3. Northern blot analysis of the LY-6K gene using a human breast tumor blot (Sigma). Different breast tumor donors (lanes 1-3) and normal breast tissue (lane 4).

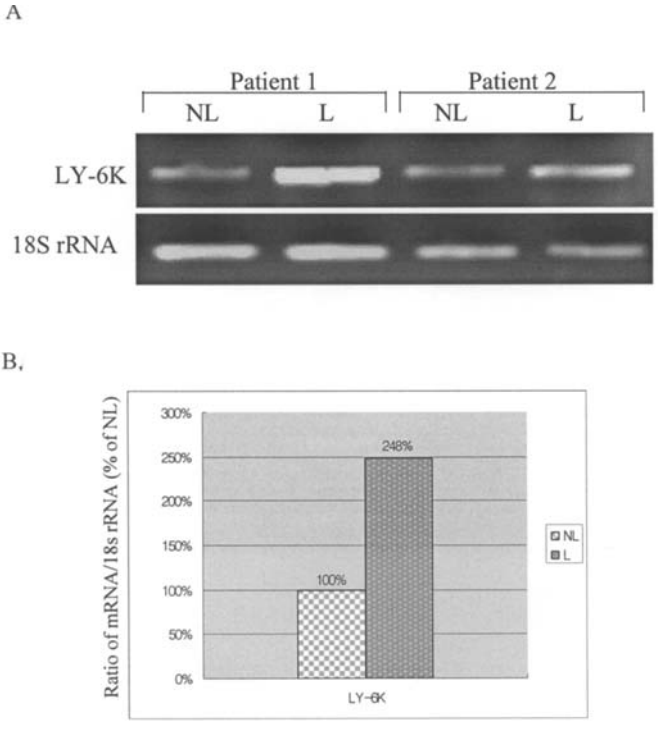


Figure 4. Expression analysis of the LY-6K gene in breast cancer tissues. (A), LY-6K mRNA was assayed in the lesion (L) and non-lesion (NL) tissues of two patients with breast cancer by RT-PCR. Expression level of LY-6K mRNA was normalized to signals detected with the 18S rRNA gene. (B), The amplified PCR products (position 103-341) were quantitated using the Multiguage program and expressed relative to the 18s rRNA, as a percentage of the NL (non-lesion tissue).

control study of 5000 SNPs obtained from breast cancer patients. Each STS (sequence tagged sites) containing one SNP marker was matched with the EST (expressed sequence tag) databases of NCBI using the BLASTn program; one full-length cDNA was detected, the *LY-6K* gene (*Homo sapiens* cDNA for differentially expressed *LY-6K* gene) encoding a novel member protein of the Ly-6/uPAR superfamily (13).

*Structural analysis of LY-6K*. In an effort to define the structure of this new gene, we identified the exon-intron boundaries using DBTSS, SIM4 and the Vector NTI Suite (Info max) program; then the mRNA sequence was translated into the amino acid sequence. As a result, we determined that the new gene having about 1735 bp-size mRNA contained three exons. The exon-intron boundary sequence was consistent with the GT-AG rule (Fig. 1).

*Human LY-6K gene expression in the nuclear membrane*. To analyze the subcellular localization of *LY-6K*, we constructed a clone containing *LY-6K* with a tag. The full-length CDS region of the new annotated mRNA was cloned into the pEGFP-N1. The constructed pEGFP-N1-*LY-6K* was transfected into the 293T cell line. Fig. 2B shows the signal of 293T cell counterstained with DAPI. The results showed that the new gene accumulated in the membrane region, as observed by nuclear counterstaining with DAPI (Fig. 2).

*Human LY-6K overexpression in human breast cancer*. The *LY-6K* gene was derived from STS of human genomic DNA obtained from breast cancer patients. Therefore, we investigated the possibility that the *LY-6K* gene might be a novel marker for human breast cancer. The expression level of *LY-6K*, in affected tissues from patients with breast cancer, was examined using Northern blotting and semi-quantitative RT-PCR. *LY-6K* was highly expressed in the three breast tumor lanes compared to the normal breast tissue lanes. Its transcript size was about 1735 bp (Fig. 3). The quantitative measurements of the *LY-6K* transcript level, in the breast tissues of breast cancer patients, were determined by quantitative RT-PCR. Total-RNAs were prepared from the skin of both lesion and non-lesion tissues from two breast cancer patients. *LY-6K* mRNA was found to be significantly up-regulated, about 2.5-fold, in the lesion compared to non-lesion skin (Fig. 4).

## Discussion

*Identification of the full-length cDNA and tissue specific expression of LY-6K gene*. In an effort to identify novel genes

involved in breast cancer susceptibility, we have conducted a large-scale, case-control study using 5000 SNPs from genomic DNAs of breast cancer patients. We detected and isolated the partial cDNA from genomic DNAs isolated from human breast cancer tumors using gene-screening consisting of BLAST searches. Furthermore, the mRNA (NM017527) sequence, containing the new full-length CDS (223 amino acids, 1735 bp), was newly annotated on the NCBI database following the mRNA clone (BC001291, 1373 bp). Therefore, we obtained the full CDS region from the human breast cancer MDA-MB-231 cell line, by semi-quantitative RT-PCR. A significant SNP that was isolated as a result of the case-control study had no amino acid change in the LY-6K CDS.

The nucleotides and amino acid sequences of the *LY-6K* gene were analyzed using several bioinformatic tools. The *LY-6K* gene is a member of the Ly-6 superfamily. It contains the Ly-6/uPAR common domain. We analyzed the structure of the *LY-6K* gene to identify exon-intron boundaries as well as to translate mRNA sequences to amino acid sequences. As a result, the size of new gene was found to span about 1735 bp, consisting of 3 exons; the third exon was identified as the largest one.

To analyze the subcellular localization of *LY-6K*, we constructed a clone containing *LY-6K* with a tag. The full-length CDS region of the new annotated mRNA was cloned into the pEGFP-N1. The constructed pEGFP-N1-*LY-6K* was transfected into the 293T cell line. The results showed that the *LY-6K* gene accumulated in the nuclear membrane region, as observed by nuclear counterstaining with DAPI.

*LY-6K as a novel molecular marker of human breast cancer*. The LY-6K gene was annotated as a novel member of Ly-6/uPAR family and a potential target antigen for HNSCC (13). We have isolated the gene from genomic DNAs of patients with breast cancer. Current reports have suggested that the uPA receptor (uPAR) can serve as diagnostic, therapeutic and prognostic markers in patients with metastatic breast cancer (14). Based on these facts, we investigated the expression level of the LY-6K gene in human breast cancer tissues. The *LY-6K* gene is a member of the Ly-6/uPAR family and was found to be expressed in both human breast normal and tumor tissues. In addition, the *LY-6K* gene is highly expressed in breast tumor tissues, and this was observed to be specific to the tumor tissue when compared to normal breast tissue. Our findings suggest that *LY-6K* might be a novel marker for human breast cancer.

The expression of the *LY-6K* gene was found to be about 2.5-fold up-regulated in breast cancer lesion skin compared to non-lesion skin; as identified with semi-quantitative RT-PCR. Therefore, LY-6K expression may be useful in establishing a diagnosis of human breast cancer, as it is for HNSCC. The consistency with which this receptor is overexpressed in breast cancer suggests that it may be a novel molecular marker; thus potential use in clinical practice may be possible for diagnosis and therapy of breast cancer.

## References


1. Pharoah PD, Antoniou A, Bobrow M, Zimmern RL, Easton DF and Ponder BA: Polygenic susceptibility to breast cancer and implications for prevention. Nat Genet 31: 33-36, 2002.
2. Hodgson SV, Morrison PJ and Irving M: Breast cancer genetics: unsolved questions and open perspectives in an expanding clinical practice. Am J Med Genet 129C: 56-64, 2004.
3. Lancaster JM, Wooster R, Mangion J, Phelan CM, Cochran C, Gumbs C, Seal S, Barfoot R, Collins N, Bignell G, Patel S, Hamoudi R, Larsson C, Wiseman RW, Berchuck A, Iglehart JD, Marks JR, Ashworth A, Stratton MR and Futreal PA: BRCA2 mutations in primary breast and ovarian cancers. Nat Genet 13: 238-240, 1996.
4. Futreal PA, Liu Q, Shattuck-Eidens D, Cochran C, Harshman K, Tavtigian S, Bennett LM, Haugen-Strano A, Swensen J and Miki Y: BRCA1 mutations in primary breast and ovarian carcinomas. Science 266: 120-122, 1994.
5. Malkin D, Li FP, Strong LC, Fraumeni JF Jr, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ and Tainsky MA: Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. Science 250: 1233-1238, 1990.
6. Thorstenson YR, Roxas A, Kroiss R, Jenkins MA, Yu KM, Bachrich T, Muhr D, Wayne TL, Chu G, Davis RW, Wagner TM and Oefner PJ: Contributions of ATM mutations to familial breast and ovarian cancer. Cancer Res 63: 3325-3333, 2003.
7. Liaw D, Marsh DJ, Li J, Dahia PL, Wang SI, Zheng Z, Bose S, Call KM, Tsou HC, Peacocke M, Eng C and Parsons R: Germline mutations of the PTEN gene in Cowden disease, an inherited breast and thyroid cancer syndrome. Nat Genet 16: 64-67, 1997.
8. Weber BL and Nathanson KL: Low penetrance genes associated with increased risk for breast cancer. Eur J Cancer 36: 1193-1199, 2000.
9. Hoyal CR, Kammerer S, Roth RB, Reneland R, Marnellos G, Kiechle M, Schwarz-Boeger U, Griffiths LR, Ebner F, Rehbock J, Nelson MR and Braun A: Genetic polymorphisms in DPF3 associated with risk of breast cancer and lymph node metastases. J Carcinog 4: 13, 2005.
10. Chimienti F, Hogg RC, Plantard L, Lehmann C, Brakch N, Fischer J, Huber M, Bertrand D and Hohl D: Identification of SLURP-1 as an epidermal neuromodulator explains the clinical phenotype of Mal de Meleda. Hum Mol Genet 12: 3017-3024, 2003.
11. Tsuji H, Okamoto K, Matsuzaka Y, Iizuka H, Tamiya G and Inoko H: SLURP-2, a novel member of the human Ly-6 superfamily that is up-regulated in psoriasis vulgaris. Genomics 81: 26-33, 2003.
12. Horie M, Okutomi K, Taniguchi Y, Ohbuchi Y, Suzuki M and Takahashi E: Isolation and characterization of a new member of the human Ly6 gene family (LY6H). Genomics 53: 365-368, 1998.
13. De Nooij-van Dalen AG, van Dongen GA, Smeets SJ, Nieuwenhuis EJ, Stigter-van Walsum M, Snow GB and Brakenhoff RH: Characterization of the human Ly-6 antigens, the newly annotated member Ly-6K included, as molecular markers for head-and-neck squamous cell carcinoma. Int J Cancer 103: 768-774, 2003.
14. Ravaioli A, Bagli L, Zucchini A and Monti F: Prognosis and prediction of response in breast cancer: the current role of the main biological markers. Cell Prof 31: 113-126, 1998.