**LRP16 gene function based on bioinformatic analysis**

Bo Yang,¹ Xue-Chun Lu,¹ Xiao-Hua Chi,² Wei-Dong Han,³ Li Yu⁴ and Fang-Ding Lou⁴

¹Department of Geriatric Hematology, PLA General Hospital, Beijing 100853, P. R. China; ²Department of Pharmacy, 262th Hospital of PLA, Beijing 100800, P. R. China; ³Department of Immunology, Institute of Basic Medical Sciences, PLA General Hospital, Beijing 100853, P. R. China; ⁴Department of Hematology, PLA General Hospital, Beijing 100853, P. R. China

**Abstract**

**Background and Objective:** The LRP16 gene is a new functional gene, which relates to the recurrence of leukemia. However, its biological function remains elusive. This study explored the biological function of human LRP16.

**Methods:** Bioinformatic prediction methods were used to analyze the structure and function of the promoter and coding proteins of LRP16. The recombinants of pGL3-Basic and LRP16 promoter subclones were constructed for luciferase activity analysis. The recombinant of LRP16 open reading frame coding sequence and pcDNA3.1 eukaryotic expression vector was established and transfected into HL-60 and K562 cell lines. DNA damage of HL-60 cells after ultraviolet irradiation was evaluated using single cell gel electrophoresis. Cell cycle of K562 cells was analyzed by flow cytometry.

**Results:** LRP16 promoter was a typical class II eukaryotic promoter and its core regulation sequence was located within upstream ~600bp of transcriptional start site. In addition, seven cis-acting elements, which may be implicated in cell cycle, hematopoiesis regulation, cell proliferation and repair of DNA damage, were identified. Long type LRP16 coding protein contained homologous sequences of histone H2A1C between 148 and 315 amino acid residue. The number of Comet cells and the length of comet tail HL-60 cells irradiated were significantly decreased, and the number of living cells was significantly increased in LRP16-overexpression group compared with empty plasmid control group. The proliferation rate of K562 cells and the proportions of cells at G2/M and S phases were significantly increased in LRP16-overexpression group compared with empty plasmid control group.

**Conclusions:** Promoter regulation prediction and protein domain analysis based on bioinformatics contribute to the study of gene function. LRP16 may play an important role in leukemia progression by promoting cell proliferation, regulating cell cycle, and antagonizing radiation-induced DNA damage.

**Key words:** LRP16, function, bioinformatics

The LRP16 gene (GenBank accession number: AF202922) is a novel human gene that was cloned using rapid amplification of cDNA end (RACE) technology from healthy human peripheral blood mononuclear cells in December 1999 by Yu et al.¹,² through comparing the status of differential DNA methylation of newly diagnosed and relapsed acute myeloid leukemia cells by methylation-sensitive restriction landmark genomic scanning (RLGS). Our previous studies have shown that LRP16 was closely correlated with pathogenesis and lymph node metastasis of breast cancer, and therefore was possibly an oncogene.³ To further explore the biologic functions of LRP16, we used bioinformatics methods to predict the structure and function of the LRP16 gene promoter and its encoded protein, as well as verified its roles in promoting tumor proliferation, regulating the cell cycle, and repairing DNA damage.

**Materials and Methods**

**Materials**

**Plasmids and bacterial strains** The pGL3 Luciferase Reporter Vectors, the pGEM-T Easy Vector, and E. coli strain JM109 were products of Promeg Inc. pcDNA3.1 (+) eukaryotic expression vector was provided by the Immunology Department at the Liberation Army General Hospital Basic Science Institute.

**Cells** Peripheral blood mononuclear cells were derived from a healthy male blood donor at the Liberation Army General Hospital and extracted by density gradient centrifugation. The human breast cancer cell line MCF-7 and the myeloid leukemia cell line...
K562 were generous gifts from Prof. Yi-Ming Mu from the Department of Endocrinology at the Liberation Army General Hospital. The acute promyelocytic cell line HL-60 was stock at the Hematology Laboratory of the Liberation Army General Hospital. MCF-7, K562, and HL-60 cells were cultured using Dulbecco's modified eagle medium (DMEM) and RPMI-1640 containing 10% fetal bovine serum, with 100 u/mL of penicillin and 100 μg/mL of streptomycin, respectively, at 37°C in 5% CO₂.

**Experimental methods**

**Bioinformatic prediction of the LRP16 gene promoter and its encoded protein**

Data were all from open-access free databases on the web. Promoter analysis software including Genomatix MatInspector Release professional 5.3, Regulatory Sequence Analysis Tools (RSA-tools), and Transcription Element Search System (TESS) were used to carry out the three-pattern analysis (that is, unknown cis-acting element screening, database analysis of known cis-acting elements, and frequency analysis of known cis-acting elements).

For the structure and function prediction of the encoded protein, we used the LRP16 gene transcription product as the target sequence to search ORFs in the human genome database and applied a computer-aided system to predict the primary, secondary, and tertiary structures of the LRP16 protein. Domain search was used to search for homologous proteins to the LRP16-encoded protein or those with similar structures. Software and related databases included Open Reading Frame Finder (ORF Finder), Clusters of Orthologous Groups of Proteins (COGs), Garnier-Osguthorpe-Robson (GOR), and Swiss-Prot.

**Construction of the LRP16 gene promoter or subcloned promoter-luciferase expression pGL3-Basic vector recombinants**

According to the method design in the literature, we synthesized a 3.0-kilobase (kb) promoter at the 5' flanking region of LRP16 and polymerase chain reaction (PCR) primers for promoter subcloning, of a total of 10 fragments. The sequences were 5'-GACATTTGGCAGGAGTTGCGA-3' for P1, 5'-TCCCAGTCGGGTGAGGCTGCTA-3' for P2, 5'-GAGCTC(SacI)CGCCGCTGCTTTGAACTGG-3' for Pu0 (2.6), 5'-GAGCTC(SacI)GGTGAAGCAGTGGAATATAACG-3' for Pu1 (2.2), 5'-GAGCTC(SacI)CTCTGTAAGCCACTTGGCT-3' for Pu2 (1.7), 5'-GAGCTC(SacI)TATATGTCAGCTGAGTCTTG-3' for Pu3 (1.4), 5'-GAGCTC(SacI)GCTACCCAACTCCTGCCGCTTC-3' for Pu4 (1.0), 5'-GAGCTC(SacI)AGATCTGCTCACAGTGGCCT-3' for Pu5 (0.6), 5'-GAGCTC(SacI)ACGAGTCGGTGGGCCCATCCTCCG-3' for Pu6 (0.2), and 5'-AAGCTT(HindIII)CCGCCACCTTGGACTCATTTI-3' for Pd.

Each subcloned promoter fragment differed by about 400 bp, with an identical 3' end and a different 5' end, that is, S0 (2.6 kb), S1 (2.2 kb), S2 (1.7 kb), S3 (1.4 kb), S4 (1.0 kb), S5 (0.6 kb) and S6 (0.2 kb) respectively. The construction of the LRP16 2.6-kb promoter-pGL3-Basic vector recombinant: Trizol was used to extract genomic DNA and a nested-PCR method was applied for DNA amplification. The PCR products were purified, recovered from the gel, and subjected to restriction enzyme digestion verification. The purified PCR products were ligated to the pGEM-T plasmid to construct the LRP16 promoter-pGEM-T vector recombinant, transformed into JM109 E. coli, and screened by the blue/white test plus a semi-nested PCR assay. The positive clones were sent for DNA sequencing. The 2.6-kb promoter of LRP16 with the fully matched sequencing result to the genome sequence was ligated to the pGL3-Basic plasmid to construct the LRP16 S0-pGL3-Basic vector recombinant. After blue/white screening and semi-nested PCR screening, the positive clones were sent for DNA sequencing again.

Construction of the subcloned LRP16 gene promoter-pGL3-Basic vector recombinants: the above subcloned promoter-pGL3-Basic vector recombinants were constructed based on the existing LRP16 gene-S0-pGL3-Basic vector.

**Luciferase activity analysis of the LRP16 gene promoter and the subcloned promoters**

MCF-7 cells of about 40% confluence were selected and transiently transfected with the LRP16 gene promoter and the subcloned promoter-pGL3-Basic plasmid recombinants using SuperFect transfection by Qiagen, following the manual instructions. At 24 h after transfection, Promega Luciferase Assay Kit Chemiluminescence luminol was used to detect Renilla luciferase intensity, following the instructions.

**Construction of the LRP16 gene-pcDNA3.1 (+) eukaryotic expression vector**

The one-step Trizol method was used to extract total RNA from the peripheral blood mononuclear cells, according to the instructions.

The full-length LRP16 gene ORF was amplified by reverse transcription-polymerase chain reaction (RT-PCR) and ligated to the pGEM-T plasmid to construct the full-length LRP16 gene ORF-pGEM-T vector recombinant. The primers for the full-length LRP16 gene ORF were designed and synthesized. The sequences were 5'-GG (Protector)GGTACC(GPI) for upstream primer and 5'-CG(Protector)GGATCC (BarHI) for downstream primer. The PCR reaction conditions were 94°C denaturation for 5 min, then 94°C 1 min, 62°C annealing 50 s and 72°C 1 min for 33 cycles, followed by 72°C final extension for 10 min, and the reaction was terminated at 4°C. The obtained PCR product was ligated to the pGEM-T Easy Vector using T4 DNA ligase and transformed into JM109 E. coli. Blue/white screening was used to search for positive clones, and restriction enzyme digestion and sequencing were used for verification. The GPI gene was selected as an internal control. The primer sequences for GPI were: GPI1: 5'-GACCCCCGAGTTTCAAGCTG-3' and GPI2: 5'-GACATTTGGCAGGAGTTGCGA-3', yielding a PCR product of 196 bp.

Construction of the full-length LRP16 gene ORF-pcDNA3.1 (+) plasmid: The full-length LRP16 ORF with a confirmed sequence was named LLRP16. The obtained positive LLRP16-pGEM-T plasmid and the pcDNA3.1 (+) plasmid were subjected to double digestion by KpnI and BarHI endonucleases. The digestion products were subjected to 1% agarose gel electrophoresis. The digested fragments were recovered from agarose gel. LLRP16 was inserted into the pcDNA3.1 (+) plasmid by T4 DNA ligase to construct the LLRP16-pcDNA3.1 (+) plasmid and transformed into E. coli JM109. The positive clones were verified by KpnI and BarHI double digestion and sequencing to ensure the correct sequence and orientation of the inserted LLRP16.
 Establishment of the HL-60 cell line stably overexpressing LRP16 and the detection of ultraviolet (UV) damage The LLRP16-pcDNA3.1 (+) clone with the correct sequence was selected to transfect HL-60 cells using the Superfect lipofection reagent by Qiagen according to the instructions. At 24 h after transfection, the culture medium containing 800 μg/mL G418 was supplied for 1 week to screen-transfected cells and 200 μg/mL G418 was supplied as a maintenance concentration. LRP16 mRNA expression was detected by RT-PCR. The HL-60 cell line stably overexpressing LRP16 was obtained after 1 month. At the same time, the pcDNA3.1 (+) plasmid transfected HL-60 cells were used as an empty plasmid control.

Detection of UV damage: Cell concentration was adjusted to 2 x 10^6 cells/mL and cultured in a 6-well plate in suspension. After removing the cover, the cells were irradiated immediately by UV light (360 nm, 40 W) for 10 min, with an irradiation distance of 40 cm. Samples were collected at 4 h and 24 h after the irradiation. Trypan blue staining was used to count living cells and single-cell gel electrophoresis (SCGE) was applied to detect nuclear DNA damage. Triplicate wells were tested for each sample and 20 cells in continuous fields were observed for each well. The experiment was repeated 3 times and the final result was calculated by averaging the 3 experiments.

Establishment of the K562 cell line stably overexpressing LRP16 and its effects on cell proliferation and the cell cycle LLRP16-pcDNA3.1 (+) was transfected into K562 cells to establish a K562 cell line that stably overexpressed LRP16, using the above method for the HL-60 cell line. At the same time, pcDNA3.1 (+) plasmid was transfected into the K562 cells as an empty plasmid control. Trypan blue staining was used to count the living cells and a growth curve was plotted. A serum-free culture was used for cell-cycle synchronization and then the serum was added in the culture medium. A solution of 75% ethanol was added to the cells at different culture stages. Cells were frozen at -20°C and stained by propidium iodide, and the cell cycle was tested using flow cytometry.

Statistical analysis Statistical analysis was performed using SPSS12.0 software. Differences were detected by t test and the level of significance was set as α= 0.05.

Results

Construction of the LRP16 gene promoter and the subcloned promoter-pGL3-Basic vectors A total of 7 subcloned LRP16 gene promoter-luciferase reporter recombinants were obtained and named pGL3-S0, pGL3-S1, pGL3-S2, pGL3-S3, pGL3-S4, pGL3-S5, and pGL3-S6, respectively. The correct ligation of the 7 subcloned LRP16 gene promoters with the pGL3-Basic vector was verified by Sac I and Hind III double-enzyme digestion and DNA sequencing.

Activity of the LRP16 gene promoter and the subcloned promoters The detection of luciferase activity showed that the subcloned LRP16 gene promoters could regulate the expression of luciferase, among which pGL3-S5 had the strongest regulatory function, indicating that the core regulatory sequence of LRP16 located at the -1 bp to -600 bp region upstream of ORF might execute a major regulatory function on gene expression. In addition, the activities of pGL3-S0, pGL3-S1, pGL3-S2, pGL3-S3, and pGL3-S4 were all relatively weaker than that of pGL3-S5, suggesting the existence of a negatively regulatory cis-structure (Fig. 1).

Bioinformatic characteristics of the LRP16 gene promoter

Basic structure analysis The DNA sequence at the 5' flanking region of LRP16 had characteristics of a typical type II eukaryotic promoter. The general constitutive promoter elements included a TATA box (-75 bp to -56 bp upstream to ORF), a GC box (-236 bp to -224 bp upstream), and a transcription initiation site (-52 bp upstream), to name a few.

Analysis of cis-acting elements In the promoter region of LRP16, some cis-acting elements had obvious region-specific distribution. For example, some cis-regulatory elements had multiple repeat sequences in the core regulatory region, a region equivalent to an approximately 400-bp fragment located between the subcloned promoter S5 and S6 (-623 bp to -200 bp), while no or little repeat sequences presented in the rest of the 3000 bp. This 400-bp region contained regulatory sequences with the

Figure 1 Effects of different regions of LRP16 promoter on luciferase activity

Statistical analysis

Differences were detected by t test and the level of significance was set as α= 0.05.
highest promoter activity. Since cis-acting elements such as T-Ag, IL-6, ZF5, Ets, estrogen receptor, retinoic acid receptor, and XPF-1 had multiple repeat sequences in the core regulatory region too, we speculated that LRP16 might be a functional gene closely related to some physiologic and pathologic processes such as the cell cycle, hematopoietic regulation, cell differentiation, tumorigenesis, stress responses, and others.

**Construction and transfection of the full-length LRP16 gene ORF-pcDNA3.1 (+) eukaryotic expression vector**

We successfully amplified a 975-bp-long target fragment, and an identical sequence of the full-length LRP16 ORF to the human genome was confirmed by sequencing. The full-length LRP16 gene ORF-pcDNA3.1 (+) eukaryotic expression vector was successfully constructed. That is, the LLRP16-pcDNA3.1 (+) and transfected into the HL-60 and K562 cell lines to obtain cells stably overexpressing LRP16 (Fig. 2).

**Structural and functional prediction of the LRP16-encoding protein**

The transcription product of LRP16 was 1030 bp long and its ORF-initiation site was located at 53 bp downstream to the transcription-initiation site. ORF Finder predicted 3 potential ORFs with a different N-terminus and an identical C-terminus, which were respectively 325 aa, 283 aa, and 195 aa in length. The full-length gene-encoded protein had a molecular weight of 35 505.11 and theoretical IPI value of 9.58, as well as was rich in glycine (8.3%) and leucine (11.1%). The secondary structure prediction showed that it was a protein with mixed secondary structures, α helix accounting for 36.0%, β folding accounting for 13.9%, and other structures accounting for 50.2% of the secondary structures. There were total of 6 known modification sites, including glycosylation sites, protein kinase C phosphorylation sites, casein kinase II phosphorylation sites, tyrosine kinase phosphorylation sites, N-14 (alkyl) acylation sites, and amidation sites. The US National Center for Biotechnology Information (NCBI) human protein database comparison and analysis found that at 148–315 amino acid residues, the LRP16-encoded protein had homologous sequences to the end of the human histone H2A1C including hismacro, COG2110, and A1pp sequences (Fig. 3A), which showed up to 80% homology. Based on that, it is possible that LRP16 might be closely related to chromosome structure stability and DNA damage repair. The LRP16-encoded protein is likely to belong to an ancient protein family in biologic evolution, since the homologous sequences of LRP16 were identified across ancient bacteria to rodents (Fig. 3B).

**Figure 2** Expression of LLRP16 gene in K562 cells transfected with LRP16-pcDNA3.1 (+) plasmid

Lane 1, DL2000 marker; lane 2, K562 cells transfected with LRP16-pcDNA3.1 (+); lane 3, K562 cells transfected with pcDNA3.1 (+).

**Figure 3** Conservative structure domain prediction of LRP16 protein and cladogram analysis

A, homologous sequences similar to human histone H2A1C terminal; B, biological cladogram.
**Effects of LRP16 overexpression on UV-induced DNA damage in HL-60 cells**

SCGE test results showed that the overexpression of LRP16 in HL-60 cells had an anti-DNA damage effect at 4 h after irradiation. Compared to the control group, the comet-like tail length in the overexpression group was shorter at 4 h and 24 h after irradiation (Fig. 4) and the number of comet-like cells were fewer (Table 1), while the number of living cells was greater (Table 1).

**Table 1 Changes in the percentage and number of comet cells in HL-60 cells after ultraviolet radiation**

<table>
<thead>
<tr>
<th>Group</th>
<th>Average percentage of comet cells (%)</th>
<th>Number of comet cell (x10⁶)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0 h</td>
<td>4 h</td>
</tr>
<tr>
<td>Control</td>
<td>0.33</td>
<td>52.33</td>
</tr>
<tr>
<td>LRP16 overexpression</td>
<td>0.67</td>
<td>29.00*</td>
</tr>
</tbody>
</table>

*p<0.01, *p<0.05, vs. control group.

**Effects of LRP16 overexpression on K562 cell proliferation and the cell cycle**

Compared with the empty vector control group, LRP16 overexpression had a proliferation/promoting effect on K562 cells (Fig. 5). Cell-cycle analysis revealed that at the same time point, the cell number and ratio in the M phase and the S phase were significantly more in K562 cells with LRP16 overexpression than in the control group. In addition, the 6th day growth plateau in the serum-containing culture was reached at an earlier time in the overexpression group. Accordingly, cells in the S phase were fewer than the control group (Fig. 6).

**Discussion**

LRP16 is a novel human gene cloned by researchers from China. Since some of the LRP16 gene expression sequence tags (ESTs) were derived from recurrent acute myeloid leukemia while the full-length gene was cloned from healthy human peripheral blood mononuclear cells, it suggests that LRP16 may be an important functional gene expressed in both healthy individuals and in patients with leukemia. Therefore, it is necessary to conduct in-depth studies on its function.

Studies by Han et al. found that LRP16 was subject to
regulation by estrogen and androgen and its high expression was closely correlated with the growth, invasion, and metastasis of breast and prostate cancers. However, currently the biologic functions of LRP16 in the pathogenesis and development of leukemia are not clear.8

In this study, we preliminarily predicted the function of human LRP16 using a combination of cis-acting element analysis of the promoter region and traditional gene function prediction methods, that is, protein structure and function prediction, functional domain homology analysis, and phylogenetic tree analysis of sequence similarity.9 In addition, we confirmed the predicted function of LRP16 in leukemia cell lines, and our results showed that LRP16 could promote K562 cell proliferation, regulate the cell cycle switch from the G0 phase to the S phase, and reduce DNA damage induced by UV.

The analysis of the cis-acting element in the promoter region improved the accuracy of traditional gene function analysis and could be used as a powerful complementary tool for gene-function prediction. Through establishing a 2.6-kb LRP16 gene promoter and subcloned promoters differing by 400 bp-luciferase reporter recombinants, we determined that the -1 bp to -600 bp fragment upstream of ORF was the core regulatory sequence of LRP16, which might play a major role in expression regulation. Further cis-acting element prediction indicated that multiple cis-acting elements, including T-Ag, IL-6, ZF5, Ets, estrogen receptor, retinoic acid receptor, and XPF-1 cis-acting elements, had more than 5 repeats in the -623 bp to -200 bp region.10 11 Thus, we speculated that LRP16 may be involved in many physiologic and pathologic processes such as the cell cycle, hematopoiesis regulation, cell differentiation, tumorigenesis, and stress reactions. At the same time, the LRP16 protein modification sites and structural predictions showed that LRP16 had multiple chemical modification sites involved in major biologic functions (eg, phosphorylation and glycosylation), as well as sequences homologous at the end of human histone H2A1C (an H2AX family member), such as hismacro, COG2110, and A1pp.

Since H2AX is related to DNA damage repair and chromosom e structural stability, mutation on this gene could lead to cancer.22 Therefore, it is likely that the LRP16 protein is related to cell-cycle regulation, gene-transcription activation and inhibition, DNA damage repair, and tumorigenesis. This is consistent with our analysis of the regulation-related cis-acting elements of the promoter region, which should direct further study of the LRP16 gene function.

Through establishing HL-60 and K562 cell lines that overexpressed LRP16, our study found that LRP16 overexpression could inhibit UV-induced DNA damage in HL-60 cells and promote proliferation of K562 cells. The proliferation/promoting effect was achieved through promoting cells entering the S phase from the G0 phase. This experimental result is consistent with the promoter regulatory analysis of LRP16 and the functional prediction of the encoded protein, which were also confirmed by experiments using the breast cancer cell line MCF-7 as an experimental model.3

In summary, this study explored bioinformatic prediction methods for gene function using a combination of gene-promoter analysis and protein structure and function prediction, which provided important guidance for clarifying the function of a novel gene and demonstrated advantages that were economical, accurate, and fast.

References


Chinese Journal of Cancer


