Sequencing of NFAT c2 interacting protein (Nip 45) coding gene.

Department of Science, Kumamoto University
073- s4106
Yuuya Uehara

Abstract

DNA sequencing is based on dideoxy method which was originals developed by Dr. Frederic Sanger in 1977, and still improving. In modern biology, DNA sequencing is a fundamental method to analyze genetic information of variety biological systems.

In this experiment, we analyzed a plasmid DNA by using fluoresceinated ddNTP and automated sequencer, and the cDNA involved in the plasmid was identified by homology search using internet.

Materials and Methods

The concentration of plasmid, named as pEGFP-007 (for nip), was determined by UV260nm absorbance using Eppendorf spectrophotometer (Biophotometer).

To analyze the DNA sequence, 16μg of DNA solution (mixture of 0.21μl of plasmid with 13.9μl of H_2O) was mixed with 2μl of sequence mix containing fluoresceinated ddNTP and 1μl of sequencing primer (primer sequence is " 5'CGAGAAGCGCGATCACAT3' ") and 1μg of buffer. The mixture was incubated as follows using thermal cycler : 20 seconds of denature at 96℃, 20 seconds of annealing at 50℃ and 3 minutes of extension at 60℃ (these 3 steps were performed 40 cycles) .

60μl of 99% ethanol and 5μg of stop solution (containing NaOAc, EDTA and glycogen) was added to the reaction and the DNA fragments were precipitated by centrifugation at 14000 rpm and 4℃ for 2 minutes using centrifuge. The precipitate was rinsed twice with 70% ethanol. After 15 minutes of desiccation, the precipitate dissolved in Sample Loading Solution (SLS). DNA sequencing was carried out using Beckman sequencer.

Results

The result of DNA sequencing and amino acid sequence which is translated based on the DNA sequence is shown in Figure 1.

From a result of homology search using Blast program, this plasmid DNA was identified as NFAT c2 interacting protein (Nip45). cRNA consists of 3388 nucleotides and cording region

consists of 1239 nucleotides (from 23 to 1261 nucleotides) . So it is deduced that Nip 45 protein consists of 413 amino acids.

Discussion

First, I want to discuss about content of stop solution which contains NaOAc, EDTA and glycogen. NaOAc contains sodium acetate, and plays in a role to form salt precipitate (DNA^{-}-Na^{+}). EDTA is a chelating agent, and it can keep enzyme inactivate state (Mg^{2+} is essential for activation of DNA polymerase). Glycogen is coprecipitated with DNA molecules.

Second, I want to discuss about Nip45. Definition of Nip45 is "*Mus musculus* nuclear factor of activated T-cells, cytoplasmic, calcineurin-dependent 2 interacting protein". The function of Nip45 is protein binding: interacting selectively with any protein or protein complex (a complex of two or more proteins that may include other non protein molecules). The process which Nip45 takes part in is cytokine production, protein modification and regulation of DNA dependent transcription. For example in T lymphocytes, covalent modification of Nip45 by arginine methylation is important mechanism of regulating the expression of nuclear factor of activated T cells (NFAT)-dependent cytokine genes. Nip45 substantially augments interleukin (IL)-4 gene transcription, whereas repression of IL-4 gene transcription is mediated through tumor necrosis factor receptor-associated factor (TRAF) family of proteins interaction with Nip45. TRAF is generally known to function as adapter proteins that transducer signals from the tumor necrosis factor receptor superfamily.

Acknowledgments

We are grateful to Teaching Assistants for doing this experiment. This experiment was supported by laboratory of Dr. Saito.

References

1. Novatchkova M, Bachmair A, Eisenhaber B and Eisenhaber F. Proteins with two SUMO-like domains in chromatin-associated complexes: the RENi (Rad60-Esc2-NIP45) family.BMC Bioinformatics. 2005; 6: 22
2. Mowen KA, Schurter BT, Fathman JW, David M and Glimcher LH (2004) Arginine methylation of NIP45 modulates cytokine gene expression in effector T lymphocytes. Mol. Cell. 2004; 15(4): 559-571
3. Lieberson R, Mowen KA, McBride KD, Leautaud V, Zhang X, Suh WK, Wu L and Glimcher LH. Tumor necrosis factor receptor-associated factor (TRAF)2 represses the T helper cell type 2 response through interaction with NFAT-interacting protein (NIP45). J Exp Med. 2001; 194(1):89-98

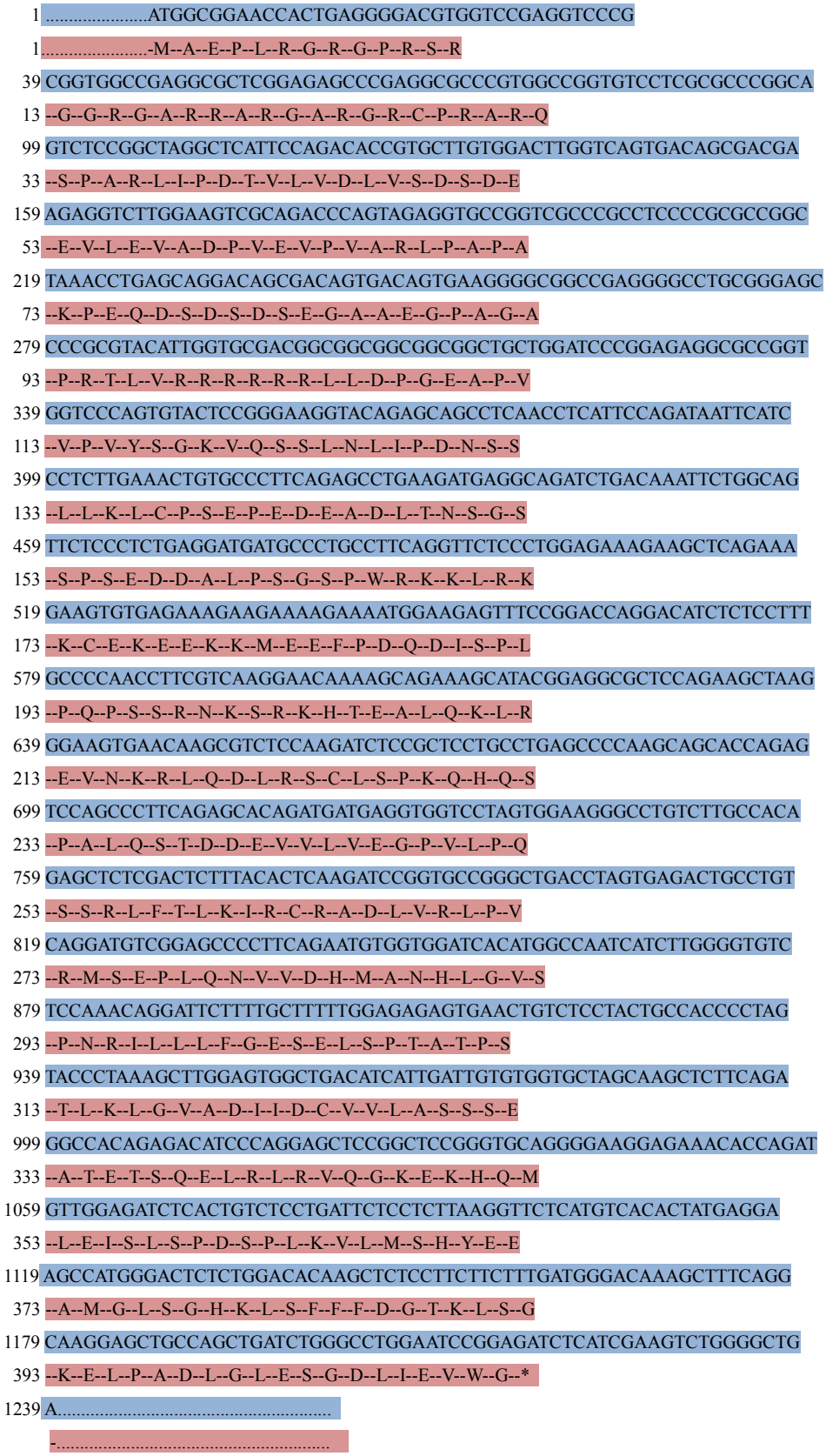


Figure1. Coding region sequence and amino acid sequence of Nip45. Amino acids are showed one-letter abbreviations. Bases are shaded blue, and amino acids are shaded red. mRNA sequence (3388 base pair) isn't showed here.

Identification of Small Ubiquitin-related Modifier 1 (SUMO1) in the plasmid named as pEGFP-hs

Kazuyo Hamada

Department of Biological Science, Faculty of Science, Kumamoto University

Material & Methods

The concentration of plasmid named as pEGFP-hs, was determined by UV260nm absorbance using the spectrophotometer (Eppendorf).

To analyze the DNA sequence, 0.35μg of plasmid DNA was mixed with 4μl of the sequence mix containing ddNTP, dNTP, DNA polymerase, and 2μl of sequencing primer(5'-CGAGAAGCGCGATCACAT-3'). The mixture was incubated as follows using thermal cycler (Applied Biosystems): denature (96˚C for 20 seconds), annealing (50˚C for 20 seconds), and extension (60˚C for 3 minutes).

99% ethanol was added to the reaction and the DNA fragments were precipitated by centrifugation at 14000rpm for 2minutes using centrifugal analysis. The precipitate was rinsed twice with 70% ethanol and dissolved in SLS. DNA sequence was carried out using Dye-terminator.

Results

The analysis gave the following DNA sequence, which consisted of about 680 base pairs (Fig.1). We carried out a homology search, and this sequence was identified as SUMO1.

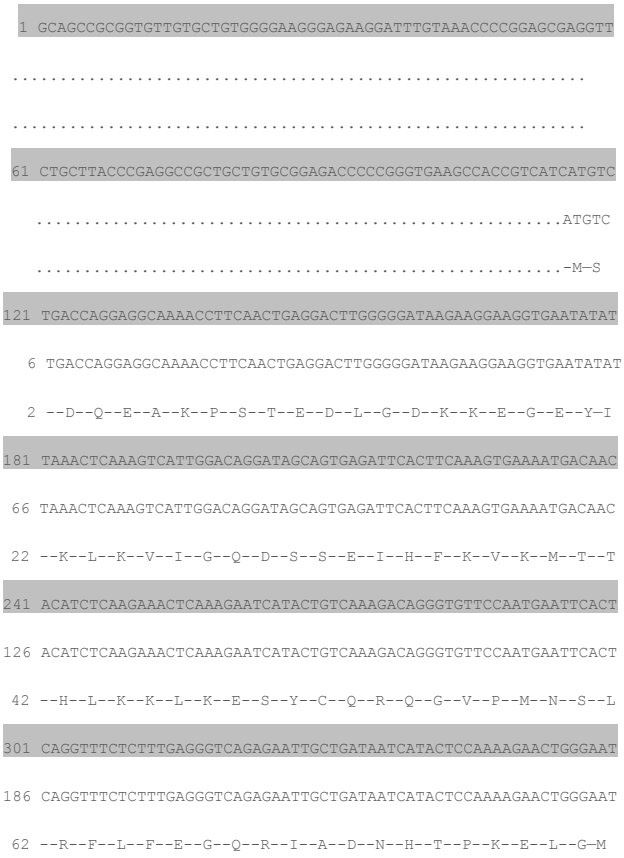
Discussion

Small Ubiquitin-related Modifer (SUMO) is a typical ubiquitin-like protein. SUMO protein looks like the Ubiquitin well. In the mammal, it is reported that four isoforms exist, SUMO-1, SUMO-2, SUMO-3, and SUMO-4, and form a protein structure similar to the ubiquitin very well. On the other hand, SUMO-1 shows the sequence similarity with the ubiquitin only by 18%, and the character on the surface of the protein is greatly different.

Ubiquitin is consisted of 76 amino acids and used to modify other proteins. The enzyme that takes part in making to the SUMO modification process is also analogous to that of the ubiquitin. While the ubiquitin becomes the tag of the degradation of proteins, such a function has not been repeated in SUMO. In the degradation process, SUMO may function as an antagonist of ubiquitin.

SUMO is consisted of about 100 amino acids. The length and weight are different according to the kind of the proteins and the originating in organisms. For example, human of SUMO-1 is composed of 101 amino acids. SUMO has many functions, and plays a role in a number of cellular processes such as nuclear transport, DNA replication and repair, mitosis and signal transduction.

SUMO is added to the target protein, lysine, though the covalent bound by a chain of enzyme reactions that look like making to ubiquitin well. In other worlds, SUMO is one of the modification factors after the translation of the protein.



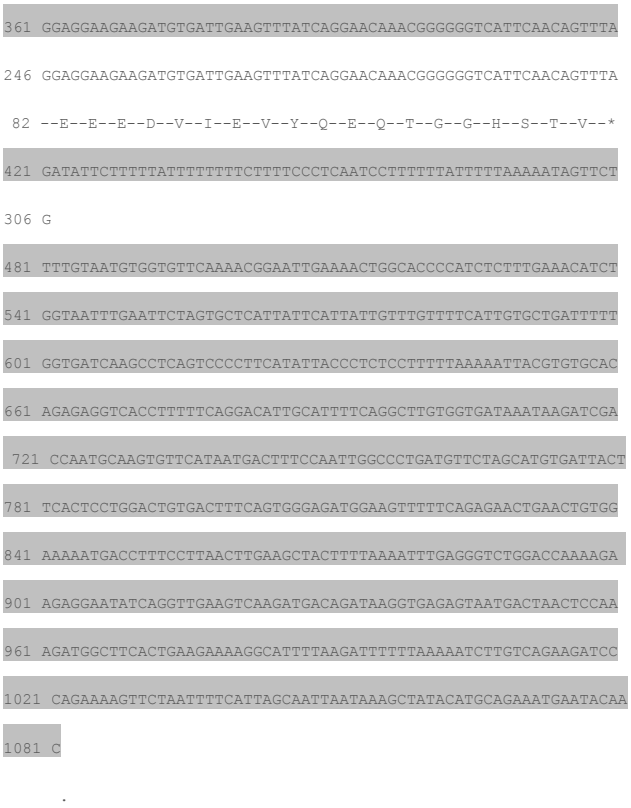


Fig.1. The nucleic acid sequence and amino acid sequence of human SUMO1 encoding gene. Top line is the nucleic acid sequence, which contained no coding and coding region of mRNA. Middle line is the nucleic acid sequence, which contained coding region of m RNA. Bottom line is amino acid sequence.

Acknowledgement

I would like to express my deepest appreciation to Mr. Saito and teaching assistants.

Reference

http://www.ncbi.nlm.nih.gov/

http://au.expasy.org/

http://www.ensembl.org/index.html

Sequencing of the human SUMO1 encoding gene

Momoyo Nomaguchi
Faculty of Science, Kumamoto Univercity.

Materials & Methods

[1] Determination of the plasmid DNA

We prepared the dilute solution of plasmid DNA, and then determined the concentration of the solution by absorptiometry. The name of this plasmid was 'pEGFP - hs SUMO1.'

[2] Sequencing reaction

After adjustment of 13μl of the DNA solution, we added to this solution the following reagents: 2μl of the primerDNA, 4μl of the sequence mix containing fluoresceinated ddNTP, dNTP, and DNA polymerase, and 1μl of Buffer. The sequence of the primer is 5'-CGAGAAGCGCGATCACAT-3', which correspond to the part of hs SUMO1 gene specifically. Amplification reaction was carried out with the thermal cycler (Applied Biosystems).

[3] Ethanol precipitation

To stop the sequencing reaction, we added 5μl of the STOP solution containing NaOAc, EDTA, and Glycogen, and 60μl of 99% ethanol to the reaction solution. After centrifugation at 14,000rpm for 15min, the supernatant was removed and the cDNA was precipitated. 200μl of 70% ethanol was added. After centrifugation at 14,000rpm for 2min, the supernatant was removed. When the centrifugation was repeated again, the cDNA pellet was collected. For the sequencing, we added 30μl of SLS solution, and scrambled.

[4] Sequencing

DNA sequencing was carried out with a DNA sequencer(Beckman). About 2 hours later, sequence was terminated.

[5] Input the sequence

For homology search, we input the DNA sequence to Sequence assistant (Vector) according to the manufacturer's protocol.

[6] Homology search

Homology search was carried out with NCBI's websites: http://www.ncbi.nlm.nih.gov/. We got the corresponding amino acids sequence. The function of analysis gene was estimated with literature.

Results

The analysis gene was found to be human SUMO1 encoding gene (Fig.1). The gene has 1081 bases in total. Base 116-421 of cDNA is coding region, but the others are non-coding region.

Discussion

'Small ubiquitin-like modifier' is formal name of SUMO. The SUMO protein's 3D structure is similar to ubiquitin. So, in a phase of post-translational modification, Sumoylation and ubiquitination are believe to function in similar but distinct biological processes.

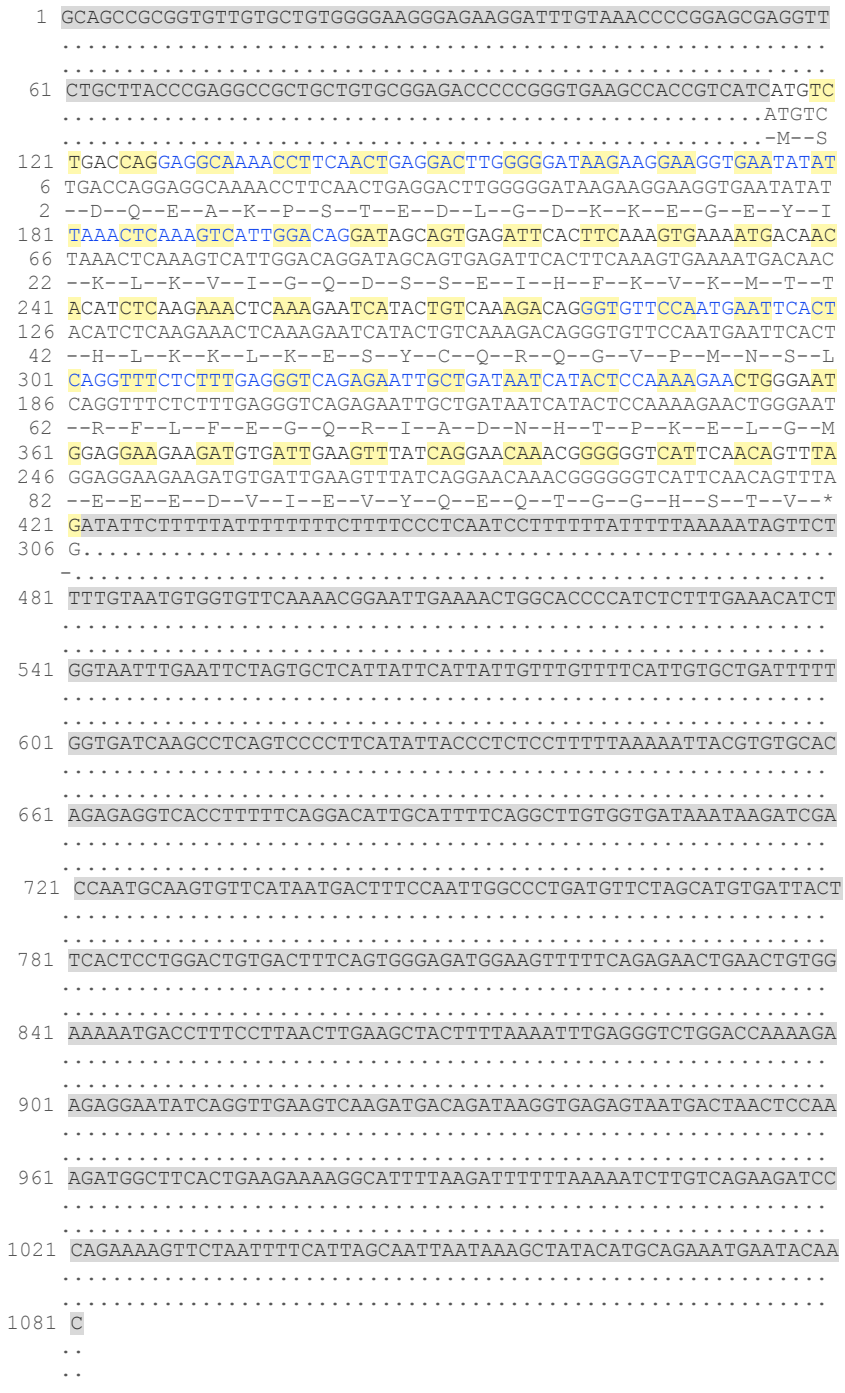
Acknowledgement

We thank T. Sasano and Y. Yasutake of graduate students of Kumamoto University for helpful technical guidance.

References

Hong-Lin Su and Steven S. -L. Li.
(2002) Molecular features of human ubiquitin-like SUMO genes and their encoded proteins. Gene 296:1-2, 21.

Fig.1
Nucleic acid sequence and corresponding amino acid sequence of SUMO1 encoding gene. Top line shows non-coding and coding region of SUMO1 encoding cDNA. Middle line shows coding region of SUMO1 encoding cDNA. Bottom line shows corresponding amino acids sequence to SUMO1 encoding cDNA.



Analysis of DNA sequence and gene function

Department of Science, Biology Selection Kumamoto University
Yamato Kojimoto

Materials and Methods

The concentration of plasmid, named as pEGFP-009, was determined by UV260nm absorbance using Bio photometer spectrophotometer (Eppendorf).

To analyze the DNA sequence, 0.4 μl of plasmid DNA was mixed with 2 μl of the sequence mix (containing dye-terminator ddNTP and dNTP and DNA polymerase), 1 μl of sequencing primer, 1 μl of primer DNA and 1 μl of reaction buffer. The mixture was incubated conditions using thermal cycler. 60 μl of 99% ethanol and 5 μl of STOP solution was added to the reaction and the DNA fragments were precipitated by centrifugation at 4℃, 14000rpm using MX-300 (Tomy). The precipitate was rinsed twice with 200 μl of 70% ethanol and dissolved in 30 μl of SLS. DNA sequencing was carried out using CEQ8000 (BECKMAN). A given sequence was searched by homology search with NCBI of web site.

Results

We got the DNA sequence of pEGFP-009 and found that it contains Homez cDNA sequence (Figure1). Homez cDNA consists of 5633 bases and the protein-coding region is from 96 to 1652 nucleotide. We found that Homez protein contains structure of homeodomain leucine zipper.

Discussion

Homez means homeodomain leucine zipper-encoding gene.

There are several regions within Homez in vertebrate homeobox gene including three atypical homeodomains, two leucine zipper-like motifs, and an acidic domain. The gene is ubiquitously expressed in human and murine tissues, although the expression pattern is more restricted during mouse development. Genomic analysis revealed that human and mouse genes are located at 14q11.2 and 14C, respectively, and are composed of two exons.

Acknowledgments

We are grateful to Ms.Muramatsu and Ms.Kawasaki for doing this experiment. This experiment was supported by laboratory of Dr.Saitoh.

References

・Proc Natl Acad Sci U S A. 2003 September 2; 100(18): 10358–10363.

GCTGTATGTAGCTTATTGTTAAAATGCTTGGCCACCGACTCCTGCCCAGTTTAGACTTCCCAGCTGTCTCTGAAGGGTACAAACCAGAGCACGACATGAGTCCTAA
TAAAGATGCCAGCAGTCTCAACAGCTCTGCAGCAGGGCTTGTCTGTCTCCCGCCAGTCTCAGAGGAGCTACAGCT
-M--S--P--N--K--D--A--S--S--L--N--S--S--A--A--G--L--V--C--L--P--P--V--S--E--E--L--Q--L
TGTGTGGACTCAAGCCATCCAAACCAGTGAGCTGGATGGAAATGAGCACCTGCTACAAGCCTTCAGCTACTTCCCCTATCCAAGCCTGGCAGATATCGCTCTCCTC
TGCCTGCGGCATGGGCTGCAGATGGAGAAGGTAAAGACCTGGTTCATGGCCCAACGTCTCCGCTG
--V--W--T--Q--A--I--Q--T--S--E--L--D--G--N--E--H--L--L--Q--A--F--S--Y--F--P--Y--P--S--L--A--D--I--A--L--L--
-C--L--R--H--G--L--Q--M--E--K--V--K--T--W--F--M--A--Q--R--L--R--C
CGGCATTAGCTGGTCATCTGAAGAGATAGAGGAGACTCGTGCCAGAGTGGTGTACCACCGAGATCAGCTCCTTTTCAAGTCTCTTCTCTCTTTCACACAGCAGTCA
GTGAGGCCCCCACAGGAGAGGCCTCCAGTGCTCCGTCCGGAGCAGGTTGCTCTTGGACTGAGTCC
--G--I--S--W--S--S--E--E--I--E--E--T--R--A--R--V--V--Y--H--R--D--Q--L--L--F--K--S--L--L--S--F--T--Q--Q--S--
-V--R--P--P--Q--E--R--P--P--V--L--R--P--E--Q--V--A--L--G--L--S--P
TCTGGCGCCTAGTGAGCAGCCCACTCATATGAAAGGATTGAAGGTCGAGCCTGAGGAGCCCTCTCAGGTATCACAGCTGCCACTGAATCATCAGAACGCTAAGGAG
CCTCTGATGATGGGCAGCAGAACATTCTCCCACCAATCAGACTGTCAGGATCTTCAGATCAGTGG
--L--A--P--S--E--Q--P--T--H--M--K--G--L--K--V--E--P--E--E--P--S--Q--V--S--Q--L--P--L--N--H--Q--N--A--K--E--
-P--L--M--M--G--S--R--T--F--S--H--Q--S--D--C--Q--D--L--Q--I--S--G
CCTCTCTAAGGAGCAGGCAGGCAGAGGTCCTGACCAGTCATGTGGAAAGACTGCTTCCTGGAATCACTTCACAGCTGTCCATCAGCCAGATAAGCCCGCATCGGTC
TCATTACTTGATAATAGTTGTAAGGAGGAATCCGAACCTAGTGGAATACCTCCATCTTCCTCTAC
--L--S--K--A--E--Q--A--G--R--G--P--D--Q--S--C--G--K--T--S--W--N--H--F--T--A--V--H--Q--P--D--K--P--A--S--V--
-S--L--L--D--N--S--C--K--E--E--S--E--P--S--G--I--P--P--S--S--S--T
CTCTTCTCCCTCTTTCCAGGCATTGGCTAATGGAACCACTGCCACCCCTAAGCCCCTCCAGCCTTTGGGTTGTATCTCACAGTCACTCTCACCTAGTAAGAAGGCA
CTGTCTCCACAAGTAGAACCACTCTGGCCCCAAAGGCTATGGAATAACTCAGAACCAAATTCAGC
--S--S--P--S--F--Q--A--L--A--N--G--T--T--A--T--P--K--P--L--Q--P--L--G--C--I--S--Q--S--L--S--P--S--K--K--A--
-L--S--P--Q--V--E--P--L--W--P--Q--R--L--W--N--N--S--E--P--N--S--A
TGGCCCCACAGAATACCTTTCCCCAGATATGCAACACCAGCGAAAGACTAAGCGGAAAACCAAAGAACAGTTGGCCATCCTCAAATCCTTTTTCCTACAGTGCCAA
TGGGCACGACGAGAAGATTACCATAAATTAGAACAGATCACTGGTTTACCTCGCCCGGAGATCAT
--G--P--T--E--Y--L--S--P--D--M--Q--H--Q--R--K--T--K--R--K--T--K--E--Q--L--A--I--L--K--S--F--F--L--Q--C--Q--
-W--A--R--R--E--D--Y--H--K--L--E--Q--I--T--G--L--P--R--P--E--I--I
TCAATGGTTTGGTGACACACGATATGCCTTGAAGCATGGACAGCTGAAATGGTTTCGGGACAATGCAGTACTTGGTACCCCTAGTTTTCAAGATCCAGCCATTCCT
ACACCTTCAACTCGTTCCTTGAAAGAATGGGCCAAGACACCACCTCTACCAGCCCCACCGCCCCC
--Q--W--F--G--D--T--R--Y--A--L--K--H--G--Q--L--K--W--F--R--D--N--A--V--L--G--T--P--S--F--Q--D--P--A--I--P--
-T--P--S--T--R--S--L--K--E--W--A--K--T--P--P--L--P--A--P--P--P--P
ACCAGATATACGACCTTTGGAGAAGTACTGGGCAGCCCACCAGCAGCTGCAGGAAGCTGATATCCTTAAACTGAGTCAGGCATCACGACTAAGCACTCAGCAAGTG
TTGGACTGGTTTGACTCTCGATTGCCTAAGCCAGCAGAAGTGGTAGTTTGTCTAGATGAAGAAGA
--P--D--I--R--P--L--E--K--Y--W--A--A--H--Q--Q--L--Q--E--A--D--I--L--K--L--S--Q--A--S--R--L--S--T--Q--Q--V--
-L--D--W--F--D--S--R--L--P--K--P--A--E--V--V--V--C--L--D--E--E--D
CGAGGAGGATGAGGAGGATGAACTGCCAGAAGATGGTGAGGAAGAGGAGGAGGAGGAGGAAGATGATGATGATGGTGATGATGATGTCATCATATGGGACTGA
GGGAGGCTGAGGCAAAGGGGAGTCTTTACTGAAATATTAACCAACTTAGTGACTGTTCAAACCTTTTGAAATTG
--E--E--D--E--E--D--E--L--P--E--D--G--E--E--E--E--E--E--E--E--D--D--D--D--G--D--D--D--V--I--I--W--D--*

CTCTTGTAAAAAGTTTTGTGTTCCTGTGTAGGCAAAGGGGGTAATCTGCAGTAAACTGGAGAGGGTGACCAGGACAAACTGAAGTATTGGATCAACACAAATCCAA
TTCTCCAGCCTCCGTGTAGAGAGAGTTTGGGATCTGGGCAGATACTGGCTGAAAAGAGCTCTTGCTTTGCTCTTGCTCCTTAGGACTCTTTAAGTCGTGTCTCAGC
CCCACATCTACAGGGAACCTGAATGGAGAGAAGTATTGTGATCCAATTGATCCCAGTGCACCCCAGCATTGTCAGAGACTCCTTGTGCCTGGCCTGTCTTTCTGAG
CACATGTTCAGGCACACTGCAGCTAGTGTTTGCAGGGTCTTGGCTTTATAACAGGCACATCGTCAGACCCCTGATAATGTCTGACTCGATTTGTCACTAATTCACA
GCAGCCTGTTCTTGCTTTTGACACTGGGGCCAATCTGCTCTCTACCATTTTATTTCTATGAAATCCCCTCTTTGCTTACTATTGTCTCAGGGTAAATCTTTACATA
GAGAGTTTGTGAAACTTTTTAATAATCAGAGTGTGAATATTAGCCTGTGTATTTATTGGAGGCGGGGATGAGCCGGCATAGTGACTGTGACTACTATACTTGGCAG
GTAGAGGCAGGAGGATCAAGGGGTTCAAGGTCAGCCTTGACTACATATGAAGTTAAAGGCCAACCTGGACTATGTGAGACTCTGTTTTGAAACATAACAGTGAGTC
AGAGGAGAAGGGCGGTGGGGTAGTTACAGGTTTGGATTTACAAACCTGACACTATTTTCTGCCCAGATTGATATAGATTGTTTGGTTGGTTTTTTAATTATTTTTT
CCTTTTTTAACTAGTGCTTCAATTTGTATCTAACAAGAGATGTTTCCCTCTTTATAGGATCTGTCTTGTGGACAGAAGGACCATGGTCTTCTAACTCAATGATGTT
CTAGAACTGGGAATTACTGGTAACAATTGGTTAGTCATCTGAAAATAGAGTGCAAGACCCAAATTGCTTAATAATAAAAAAGGCCAAAAATAAAATTGTTTCTTT
TCTCCACTTTGCCACTTCTGTGTTTCAAGATTTATCTCCGTCTATCTATCCATCCACTTTCCTTTAGGTATATGAGTACACTGTTGCTGTCTTCAGACACACCGGA
AGGGGGCATAAGATCCCACTACAGATGGTTGTGAGCCACCATGGGGTTGATGGGAATTGCACTCAGGACCTCTGGAAGAGCAGTGTTCTTAACTGCTGAGCCATCT
CTCTAGCCCCATCTTCACTTTCTTAAATGCTTGTCTTGAACTCACTGTGAAGCCATGGCTGCTGCTGTTCTTATCTTCCTGTTTCCATCTCCTAAACGCTGATTAG
AAGTATGTGCCATTATGCCCAGCTTATGCAGTACTGGGGATCAAACCAGGCCTTTATGTATGGTAGGCAGGTACTCTGCCAACTGGGCTACATCTGTACCTAGCCT
GAGTCTAGGAATTTAATGCTGGCATGTAGCCCAGTGGTAGACCTCTTACCTAGCATGTGTGACACCCTAGTTTCAAGCCCAGCATCACTTTAAAATCATTTCTGGT
TTAGTTTATTTTTTTAAAGAGTTAGTCAGGACACGCTGCAGTAGATAAAAATTGAAGACAGTAATTCAGCATCTATTGCTACATGCATGGTTTGCAGCCATCAGAT
GACAAGGTTTATTGGCCCTCCAGCAGTGACAACAGGTTGTAACTCCACTTGCTCATCCTACCATTTGAGAAACTATTTTAGCAAGAGTGGCAGAAGCCATAAAAGG
GTATGTTAATGATTTTTTTCCTCTTTATTAGTCTATATACCTTACTAAGGTGGAAATCTACCTGTTGTCAAAGCTGGAAATGTATTCTACCTATTGTTTGGATAAA
GGAAAATCAGAACATTTCATAGCAGCTGGGTGTAGAACAATAGGTAAAAAGTATAATACATGAGGGTGGGGAACAGTTGTTTATTAACATTGGCCACAGGATAGCA
ACGTCCAGAGCTAAGCAGTGAAATCCAGAAATGAACCAGGATAAATATATTATCTAGTCTTCCTATGCCAAAGCATGAGGCTAAGGATTTTTGCAACCTTAGTCTC
AAACTAGGAAATGAAAGGCCGCACATATGGGAGAGATGAAGATCTTCCTGAAGCCCTGGGGAAGACAGCCACTGAGATGGCACTTGCCACCAAGTCTAGAGACTTG
AGTTGGATCCCCAGGACTAACTTTGTTGGAAAGAACTGACTCCTGCAAGTTGTCCTCAGACCCCCCACAAGTACGTACACAAAGTAAACAAAACCCAGGGGGCTGG
AGACGTGGCTCAGCAGTTTAGTATGCTTTGTGCCGGGCTCCCCAGCACCCATGTCAGTCAGCTCACACGCTGCCTATAACCCTGGCTTTAGGGGATCATACACCCT
CTTCTGGTCTCTGAGGGCACCTGCATCACAGGCACCCATATGCAAGTAGTATACCCTCATACAACACACACAGTTTTTAAAAATTTTTTTAATGTTTGGAGAGAAA
GACATGTAAGATAACGTTCAAGAAAACAAGCTGCAGACTGAAACCAGGAACCAGCAAGTGCCCATAGCCCCTGAGAGAACTGGACACTGCAGGCCAACCTTGAGCA
TTATGCACAGAGCAGACCTGGAAGAGGGATGTCTATCTGGAAATGTTCCATCCGTCTTTCCTGGCTAGTCCCATCTGCCTCTGTTGGTAGAACGTCTGCAGCAGAG
AACAGATTTGCATGTTATGACAATACAGACTACTAAAGGTGTCCTACCAATAAAGTTTTATTGAACAAAATCAACTGCTACTGCTTCTCCATGGTTTTTCCTCATT
GTATGATGACTAACTTACAAACAAAAAGAAATTTTGTTTGAGACAGGGTCTCATTATTGCTATTGCTATGTAGACCATGGCTGGCCTTGAATTCACAGCAATCTGC
CTGCTTCTGTTTTCCAAAAGGGCAGGGATTAAAGGTGCCACCATGCCCAACCTAGATTAAGAGCTGATGCGTCCTGTGAGACAAGGCCTTTGTTGTTACTGAAGAA
ACTGTACGCTTCAGTGCTTGATGATGATGATGTAAAATATCTGAGGAACTTTGGGAAGATTTGGAACATCTGACTACTGTATCAAAAGACAATTACCAGGGAGGCA
AAGATCTTACACCTCTAGTGATGTGGTACTTTTCAGTAACTATTTTTAACTTTATAATGTTTTATTTTGAAGTTTCATATTTATGGATGTTTTGTTGAATGTATT
TCTGTGCTCAGAGACAACTCTGGAGATCCCCTGGGCTGGAGTTACAGAGGATTGTGATCCACCATGTGGGTGCTGGAATCGAACCCAGGTACCCTGGAAGAGCAGC
CACTATTCTTAACTGCTGAATCATCTCTCCAGTCCCAATAGGTTTTGTTTTAAAACAGTAGTTCTTAACCTGTTGGTCTCCACCCCTTTCGGGTCAAATGACCCTT
TCACAGGAGTCACATATCAGATATCCTGCATATCAGATATTTATATGATTCATAACAATAGCAGAGTTACAGTTATGAAGTAGTAATAATTCTATGGTTGGGGTCA
CCACACATGAGGGACTGTATTAAAGGGTTACAGCAGTAGGAAGGCTAGGAACTACTGTTTTAAAGGATAGACCAGAATAATGAACCTGTCATCCAATATCAACAAA
ACTCGTGGCTAATATTTATTGTTCTTTCTATTGCTTCCTTTTAGTAAATTATACTAAAGCAAAATCTAGATACCATTTAACTTCATAAAATTTTTATTTACATATA
TATGGATATGTACTTGTTTGTGTTTGTGTATTTGTGAGTAAATACAAAGTCATCAGACCTGAAAAAAAAAAAAAAAAAAAAAAAAAAA

Figure1. cDNA sequence and amino acid sequence of Homez.There are 5633 nucleotides in Homez cDNA, gray line indicates the coding region of Homez (96 – 1652 nucleotide). Blue line shows the deduced amino acid sequence.