ISSN 2250-3153

In Vitro and in Silico Study of Homologous Sequences in Leishmania Major

Thiruvarutchelvan S, Dr. Stephen Heath

Dr. Stephen Heath, School of Environment & Life Sciences, University of Salford, UK.

Abstract- Leishmaniasis is a widely spread disease caused by protozoan parasites belonging to the genus *Leishmania.* The unique life cycle of this parasite has two main stages of existence, amastigote and promastigote. The non-infective amastigotes, found in the primary host (example: human) when taken up by certain species of sand fly (Phlebotomine) whilst ingesting the blood of the infected person, were transformed into promastigotes which is the infective form of the parasite and this process takes place in the mid-gut of the sandfly. One of specific proteins in the parasite's genome in the transformation process, the adhesion protein gene was selected for this homology study. This *In Silico* approach to identify the homologous sequences in *Leishmania major* was to understand the evolutionary relationship among species, both in structural and functional aspects. The adhesion protein gene, which is responsible for the adhesion of the parasite in promastigote stage on to the midgut wall of the insect vector, was selected for the purpose of primer design. The DNA from the *leishmania major* culture was extracted and the designed primers were used against it in PCR. The product was purified from the gel and the PCR product was sequenced. The sequenced data was found to have homology with *Leishmania major* genome and the confirmations are produced.

Index Terms- Leishmaniasis, Homologous Sequences, Adhesion Protein Gene, Polymerase Chain Reaction, *In Silico* study.

Abbreviations Used

NCBL – National Centre for Biotechnology Information
nBLAST – Nucleotide Basic Local Alignment Search tool
MSA – Multiple Sequence Alignment
LmjF – *Leishmania major* Friedlin
PCR – Polymerase Chain Reaction
EtBr – Ethidium Bromide
DNA – deoxyribonucleic acid
mRNA – messenger Ribonucleic acid
DES – DNA Extraction Solution
dNTPs – deoxynucleotides
Taq – *Thermophilus aquaticus*
T.E. buffer – Tris EDTA buffer
rpm – rotations per minute

I. Introduction

Leishmaniasis is a disease caused by the protozoa of *leishmania* species. *Leishmania* is a parasite which belongs to the genus, trypanosome protozoa. It is an intracellular pathogen which has major affinity towards dendtritic cells and macrophages of the immune system. This disease is spread globally with variation in the strains across continents, for example, the causative agent of Visceral Leishmaniasis is L.donovani in India whereas it is L.infantum in Latin America (Berman 1997). This parasite is transmitted by the bite of female sand flies of genus, *phlebotomus*. Leishmaniasis can be clinically classified into three categories and they are cutaneous leishmaniasis, visceral leishmaniasis and mucocutaneous leishmaniasis in which visceral leishmaniasis is often fatal if not treated at earlier stages. The common symptoms are skin sores lasting for months and in some cases, it can also damage spleen, liver and may cause anaemia (Killick-Kendrick, R., 1987).

There are two different morphological forms in the life cycle of this parasite, leishmania. They are promastigotes, with an anterior flagellum and the amastigotes, without flagella. The intermediate stage of this parasite takes place within the vector organism, sandfly (Newton 1974). There is no known reproduction cycle and multiplication in both invertebrate and vertebrate hosts. The only multiplication cycle is by binary fission. The transmission of these parasite protozoa between mammalian reservoir hosts is predominantly by the bite of the insect vector (Berman 1997).

The spherical shaped amastigotes does not possess a flagellum and hence they are non-motile. The transformation of non-motile amastigotes into motile promastigotes takes place in the mid-gut of the sand fly. There are sequences of steps involved in the process of transformation of non-motile amastigote into motile promastigote. The promastigotes attaches to the mid-gut wall by inserting the long flagella between the microvilli which line the midgut during when they undergo metacyclogenesis and transform into metacyclic promastigotes which are responsible for an infection in a primary host (Vickerman 1974). According to Alexander, 1975 the inference from the *invitro* observations in mouse peritoneal macrophage cultures of different species of leishmania such as L. donovani, L. mexicana, L. major and L. enriettii, is that the transformation to spherical amastigotes within the macrophages takes within 12-24 hours of internalization of the macrophages (W.Peters & R.Killick-Kendrick, 1987). After the cell lysis of the macrophages, the amastigotes are released from the macrophage and accumulated on the skin sores and on its periphery. It is also present in the peripheral blood stream of the vertebrate host which can be transmitted to the insect vector through a bite of the insect vector (Molyneaux 1983).

Importance of Homologous Sequence Identification

The identification of homologous sequences in an organism paves the way to find out the accurate relationship between

sequences which are structurally similar (Balaji and Srinivasan 2001). Those sequences which have evidence of evolutionary relationship and functional similarity may also be non-homologous to each other (Aloy et al 2002). There are neighbour sequences which are well super imposed on each other and may also have some structural and functional similarity but they may not be homologous neighbours. The study of homologous sequences helps to obtain the data about the course of its evolutionary relationship with other sequences and also about the mutation of the genes with respect to time (Higgins 2003). The collection of sequences and the origin of the gene were completely stored in many nucleotide and protein databases such as NCBI, GenBank, SwissProt, etc, (Higgins 2003). The complete access of sequences through these databases and the tools in these databases enables a deep study about the sequences and also to identify the homologous sequences in particular organisms (Schleif 1986).

The aim of this project is to identify the homologous sequences in *Leishmania major* using *in silico* studies and PCR. The homologous sequence identified through this study will certainly lead to deeper understandings of the structural and functional relationship of *Leishmania major* with other species.

II. MATERIALS AND METHODS

NCBI

NCBI (National Center for Biotechnology Information) is established in the year 1988 as a national resource for molecular biology information. NCBI nucleotide database was used to identify the list of sequences and information on the sequences of our interest in a greater detail.

nBLAST and ClustalW

nBLAST (Nucleotide Basic Local Alignment Search Tool) was used to find the homologous sequences. About five homologues from different organisms with ~60% homology were selected and used for further study. MSA (Multiple Sequence Analysis) was performed in these five sequences to analyse the area of homology among the gene homologues of adhesion protein gene using clustalW (Altschul 1990 and Lesk 2008).

Primer Design

Primers were designed by Primer3 which is a tool available online that is a convenient primer designing software. The adhesion protein gene was chosen in leishmania major, for which NCBI database was used to access the information about the sequence. The adhesion protein gene was chosen for the reason that the point at which the amastigote form of the parasite were transferred to promastigote after attaching to the gut wall of the insect vector, phlebotomus thereby being one of the principal genes expressed during the transformation process in this parasite's life cycle. This sequence was used in nBLAST which is available online, to find its corresponding gene homologues. About five gene homologues were chosen for multiple sequence alignment in clustalW (Thompso 1994). Through this method, it was possible to identify the conserved regions among those five homologues. *Branchiostoma floridae* hypothetical protein (BRAFLDRAFT_102194) which is the one with more homology was used in *Leishmania major* database in order to find homologous sequence within LmjF database. One of the homologous sequences was used in its FASTA format in primer3 programme to design the corresponding primers. The reagents for DNA extraction process and Taq Polymerase, dNTPs were purchased from Cambridge Biosciences, UK.

DNA Extraction

The extraction buffer was added to the sample and centrifuged for about 10mins to form the pellet. The supernatant was discarded and the pellet is preserved which was then used to extract the DNA. The pellet preserved was defrosted and resuspended in 500µl of DNA Extraction Solution (DES). 5µl of proteinase K (20mg/ml) was added to the resuspended solution. Proteinase K is commonly used in nucleic acid purification because it inactivates the nucleases present in the mixture which can degrade the nucleic acid by breaking the phosphodiester bond between the nucleotides.

This mixture was incubated at 50ºC for 1hour in a hot water bath to enable the cell lysis. Then, the mixture was chilled on ice for 5minutes. To these mixtures, 275µl of propan-2-ol was added and mixed gently by inversion for several minutes. This mixture was then centrifuged at 6500rpm for 10minutes to form the pellet. The supernatant was then discarded which was then followed by the addition of 25µl of T.E.buffer containing 50µg/ml RNAse in order to dissolve the pellets.

An equal volume (100µl) of phenol/chloroform/isoamylalcohol was then added and repeatedly spun for 5 minutes using a vortexer, which was then followed by 2 minutes centrifugation at 13000rpm. The supernatant was then carefully transferred to a fresh tube and the above process was repeated. The transfer of supernatant was done carefully since there may be some presence of RNAse at the junction of the two layers which may degrade the DNA. The supernatant was then transferred to a fresh tube again and the DNA was precipitated by adding 1 volume (100µl) of 95% ethanol which was followed by centrifugation for 5minutes at 13000rpm. The supernatant was then discarded and the pellet was air dried for 5minutes. The air dried pellet was then dissolved in 25µl of T.E.buffer and stored at 4ºC over night.

DNA Concentration plate

The bottom of the plate was marked with two columns and six rows enabling twelve boxes. The left column was allotted for the known concentration of lambda DNA and the other column for the sample. In the left column, 1µl of known concentration of lambda DNA was added in all of the boxes in the order from 100 to 3.12 whereas the boxes in the other column was used to load 1µl of known concentration of sample.

The sample was loaded in the concentration order from neat sample to 1/125 dilution with water. The Petri plate was then viewed under UV transillumination and photographed in order to estimate the concentration of DNA in the sample.

PCR

PCR machine used in this reaction was from Stratgene Robocycler40®. This required set of components like Taq DNA Polymerase and dNTPs were purchased from Sigma-Aldrich® UK. The annealing temperature and the number of cycles were

optimized upon trials. The master mix preparation was done using the manufacturer's protocol and the conditions for PCR were optimized in trials. The PCR processed samples were loaded into the agarose for electrophoresis. Agarose gel and electrophoresis tank were purchased from Mini-Sub® Cell GT systems, UK. Electrophoresis was performed for the PCR processed samples that were loaded into the wells and for the intercalating agent, Ethidium Bromide (EtBr) was replaced by GelRed® purchased from Cambridge Bioscience, UK. After electrophoresis, the gel slices were placed in UV transilluminator that was docked with a computer in order to image the gel slices for document purposes. The UV transilluminator was purchased from Gentaur Ltd, France.

Gel Extraction

The PCR processed samples, once after gel electrophoresis was purified from the gel slice for further analysis, sequencing. The extraction process was done using QIAquick® Gel Extraction Kit following the recommended manufacturer's protocol. The extracted sample was vacuum centrifuged in order to concentrate the sample and enable sequencing.

III. Results

This project work on Identification of homologous sequences in leishmania major using *in silico* analysis and PCR consists of many phases. The primers were designed using the conserved region of sequences in the output page of clustal W. The details on the gene homologs of adhesion protein gene of *Leishmania major*, the homologous sequence in LmjF database and the results of ClustalW have been mentioned in the addendum and in Figure 3 respectively. The DNA from the parasite was extracted using the protocol explained in the methodology part (Fig.2(A)). The concentration of the DNA extracted from the sample was determined in order to know the amount of DNA present in the sample (Fig.2(B)). 1/25 dilution of the DNA sample was chosen for the experiments. The information on the designed primers was also included in the addendum. The PCR samples were added totally and electrophoresed to increase the concentration of the PCR product which was then followed by gel elution (Fig.4). The PCR product was purified using a purification kit by following the recommended protocol. Since the length of the PCR product is considerably small, the purified PCR product was vacuum centrifuged to increase the final concentration of the sample that was sent for DNA sequencing.

IV. Discussion

The conserved region of five homologs of adhesion protein gene in *Leishmania major* was identified using multiple sequence alignment (Fig.2). *Branchiostoma floridae* hypothetical protein (BRAFLDRAFT_102194) among the homologous gene was searched in LmjF database to obtain the homologous sequence for the purpose of primer design. This analysis was done to insist on the point, that the structural and functional properties among species are evolutionarily related.

As mentioned earlier, the primers were designed using the homologous sequence, LmjF09.0740 (Ubiquitin Ligase, putative) in LmjF database. It would be essential to insist on the point that LmjF09.0740 was one of the closest matches to one of the gene homologues of the gene of interest, the adhesion protein gene. The primers flanked their target sequence, thus representing the homology of *Leishmania major* with the particular homologue.

The sequencing of DNA is usually done by chain termination method and also by chemical degradation method. These methods of sequencing a DNA strand is very hard with small samples as there will be possibilities to miss part of sequence to which the primers are flanked and also some more after them during one way sequencing which is an influencing factor for small products. The amplified products are just small fragments and to separate them out in a better way, the concentration of the agarose gel was increased from 1% to 1.5%. This increased concentration of the gel resulted in better separation of the smaller fragments.

The sequenced data was aligned to check for the PCR product in order to confirm the sequenced data from the final result was the exact homologous sequence expected. MSA was performed for the sequenced data and the primary sequence, L. major ubiquitin ligase, putative mRNA from which the primers were designed. It was interesting to notice the homologous sequence which is the PCR product. However, there are some nucleotides missing which was because of the small size of the PCR product. Additionally, we encountered a compromise in setting up the annealing temperature because of the considerable variation among the forward and the reverse primers which was believed as a major influence in obtaining a better PCR product. The sequenced data obtained was viewed in Chromas 1.45 for the chromatogram of the sequence. This alignment program clearly proves the presence of the homologous sequence in *Leishmania major*. However, there is a strong need to delve deeper into understanding the structural and functional properties of the sequence in order to clarify their homology.

References

[1] Aloy P et. al (2002) Structural similarity to link sequence space: new potential superfamilies and implications for structural genomics. Protein Science 11(5):1101-16.

[2] Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. (1990) "Basic local alignment search tool." J. Mol. Biol. 215:403-410

[3] Berman, J.D. (1997) 'Human Leishmaniasis: Clinical Diagnostics and Chemotherapeutic Developments' in Clinical Infectious Diseases. Vol 24, Issue 4, pp 684-703.

[4] Bray, R.S. (1974) Trypanosomiasis and Lesihmaniasis. Associated Scientific Publishers: Amsterdam.

[5] Herwaldt, B.L. (1999) 'Leishmaniasis' in Lancet. Vol 354, Issue 9185, pp 1191-1199.

[6] Higgins, D. and Taylor, W. (2003) Bioinformatics: Sequence, Structure and Databanks. Oxford University Press: New York

[7] Killick-Kendrick, R. (1979) Biology of Leishmania in Phlebotomine Sandflies. Biology of Kinetoplastida, Vol. 2 (ed. Lumsden and Evans) Academic Press: New York.

[8] Lesk, A.M. (2008) Introduction to Bioinformatics, 3rd Edition. Oxford University Press: New York.

[9] Molyneaux, D and Ashford, R. (1983) Biology of Trypanosoma and Leishmania, Parasites of man and domestic animals. Taylor and Francis (London)

[10] Newton, B.A. (1974) Trypanosomiasis and Leishmaniasis. Associated Scientific Publishers: Amsterdam.

[11] Peters, W. and Killick-Kendrick, R. (1987) Leishmaniasis in Biology & Medicine. Vol 1. Academic Press: London.

[12] S.Balaji and N.Srinivasan (2001) Use of a database of structural alignments and phylogenetic trees in investigating the relationship between sequence and structural variability among homologous proteins. Protein Engineering vol.14 no.4 pp.219–226, 2001

[13] Schleif, R. (1986) Genetics & Molecular Biology. Addison-Wesley Publishing Company: Canada.

[14] Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. Nucleic Acids Res. 22(22):4673-80.

[15] Vickerman k. (1974) Trypanosomiasis and Leishmaniasis. Associated Scientific Publishers: Amsterdam.

AUTHORS

First Author – Thiruvarutchelvan S, BTech, MSc, M.S. (by Research), Dr. Stephen Heath, School of Environment & Life Sciences, University of Salford, UK, Email: Sbalaji21@gmail.com

Second Author – Dr. Stephen Heath, MSc, PhD, Dr. Stephen Heath, School of Environment & Life Sciences, University of Salford, UK., Email: s.heath@salford.ac.uk

Correspondence Author – Thiruvarutchelvan S, BTech, MSc, M.S. (by Research), Dr. Stephen Heath, School of Environment & Life Sciences, University of Salford, UK, Email: Sbalaji21@gmail.com

Figures

Figure 1

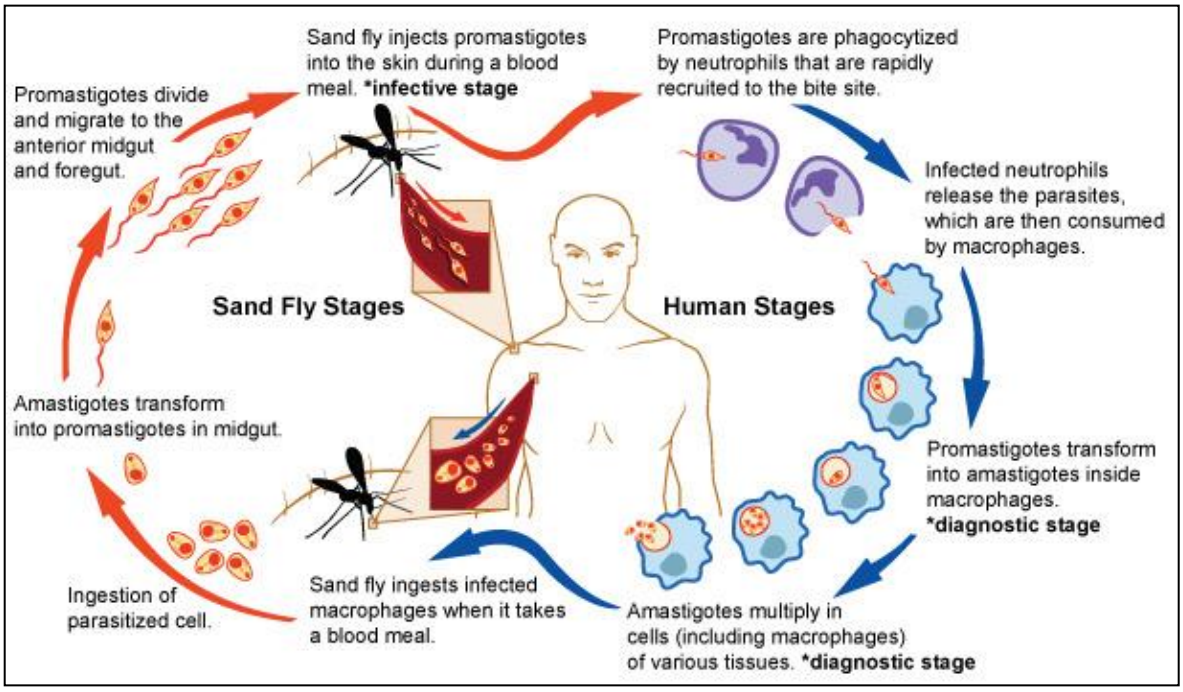
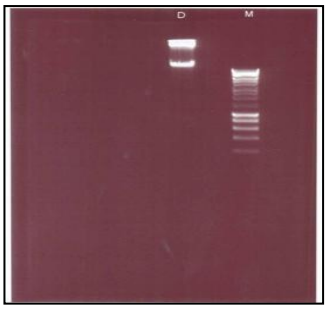


Figure 2



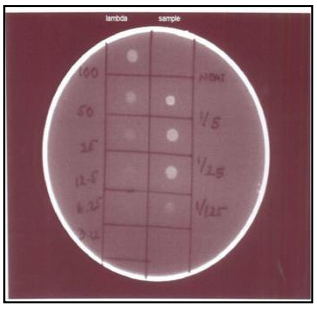
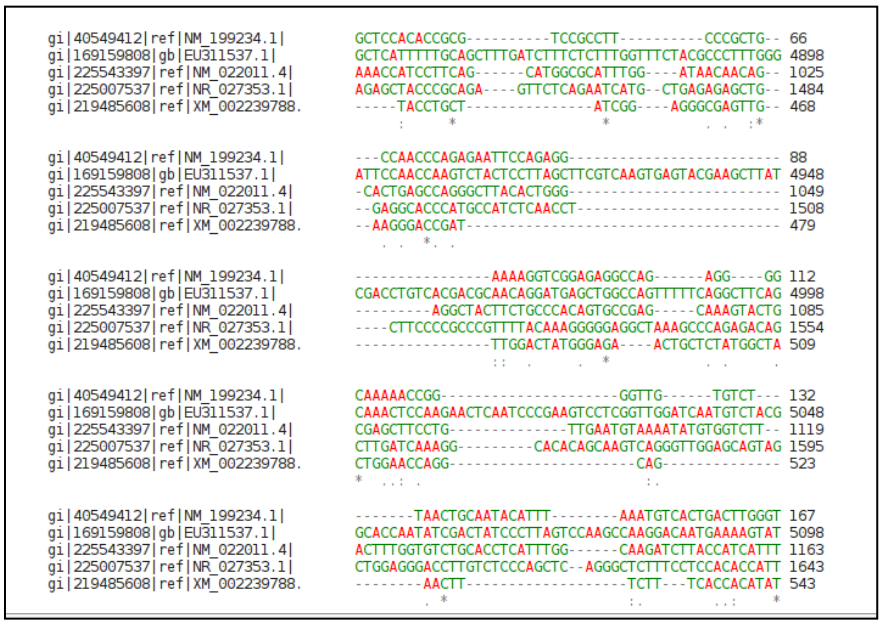


Figure 3



Legends for Figures

Figure 1

Sandfly and Human Stages of Leishmaniasis Logo (Image Courtesy: National Institute of Allergy and Infectious Diseases (NIAID))

Figure 2

2 (A) Image confirming the DNA Extraction from the parasite, *Leishmania major* Logo

2 (B) Determination of DNA concentration in the sample Logo

Figure 3

Identification of the conserved region in clustalW Logo

Appendices

Information on Primers:

Forward Primer	5’ – GTCTGGCTATCAGCGGCTAC – 3’
Melting Temperature (Tm)	61.4ºC
GC Content	60%
Molecular Weight	6108 g/mol
Length	20-mer

Reverse Primer	5’ – TTGCAGGCAGTCTCAGAAGA – 3’
Melting Temperature (Tm)	57.3ºC
GC Content	50%
Molecular Weight	6166 g/mol
Length	20-mer

The gene homologues of the adhesion protein genes and the information about them were obtained from NCBI database. FASTA format of the sequences was used in Multiple Sequence Alignment. The accession code in NCBI Database for the gene homologues were mentioned below.

>gi|225543397|ref|NM_022011.4| Mus musculus general transcription factor II H, polypeptide 2 (Gtf2h2), mRNA
GAATTTTGGCGGCGCACTCAGACTGCCGGTACTTCCGGTGAGGGTCGTCTCGTCGAGGCTGGTGGGTCTACGCGAGGAG
ACCTGCGGGGTCCTGAGGCTGTGGAGGGCTGGAGTGAGGCGGACAGGAGAAGAAGCCTGGAGCTGAGGACCGAGAGT
TGAATAAGCATATATATCATTATAAGCACCATGGATGAAGAACCTGAGAGAACCAAGCGGTGGGAAGGAGGCTATGA
GAGAACCTGGGAAATTCTTAAAGAAGATGAAACTGGATCACTTAAAGCTACAATAGAAGATATTCTCTTCAAGGCAAA
GAGGAAAAGAGTGTTTGAGCACCATGGACAAGTCCGACTTGGAATGATGCGCCACCTGTATGTGGTGGTGGATGGATC
GAGAACAATGGAAGATCAGGATTTAAAGCCCAATAGACTGACTTGCACTTTAAAGTTGCTGGAATACTTTGTAGAAGA
ATATTTTGATCAAAACCCTATCAGTCAGATTGGAATAATTGTAACTAAGAGTAAAAGAGCTGAAAAACTGACTGAACT
CTCAGGAAACCCAAGGAAACATATAACATCTTTGAAGAAAGCTGTAGATATGACCTGCCATGGAGAACCATCGCTCTA
TAATTCCTTAAGCATGGCTATGCAGACCCTAAAACACATGCCTGGACATACAAGTAGAGAAGTGCTCATCATCTTCAGC
AGCCTCACCACCTGTGATCCATCTAATATTTACGATCTCATCAAGACCCTGAAGACAGCTAAAATTAGAGTGTCTGTTA
TTGGATTATCTGCGGAGGTTCGAGTTTGTACTGTACTTGCTCGTGAAACTGGTGGCACATACCATGTTATCTTAGATGAA
ACCCATTACAAGGAGTTGTTGGCACATCATGTGAGCCCCCCTCCTGCCAGCTCAAGCTCCGAGTGCTCACTCATTCGCA
TGGGATTCCCTCAGCATACCATTGCTTCTTTGTCTGATCAGGATGCAAAACCATCCTTCAGCATGGCGCATTTGGATAAC
AACAGCACTGAGCCAGGGCTTACACTGGGAGGCTACTTCTGCCCACAGTGCCGAGCAAAGTACTGCGAGCTTCCTGTT
GAATGTAAAATATGTGGTCTTACTTTGGTGTCTGCACCTCATTTGGCAAGATCTTACCATCATTTATTTCCTTTGGATGCT
TTTCAAGAAATTTCCCTAGAAGAATATAAAGGAGAAAGGTTTTGTTATGGATGTCAGGGGGAATTGAAAGACCAACAT
GTCTATGTTTGCACAGTGTGCCAAAATGTTTTTTGTGTGGACTGTGATGTCTTTGTTCATGACTCTCTCCATTGTTGTCCT
GGCTGTATTCATAAGATCCCAACTCCTTCAGGTATTTAATACCAATATGAAGAACACACACATCGAGTGGGTTTAAAAA
CAAAACAAAAGTTGTAACTGTAAATAAAATGATTCTTTGAATAAAGTGATTCTCTAATTAAAACATGAAAGCACTTTGA
AAGAAGATCTGATTTAAGAAACTTTTAACTATGAAAATAGAGTATTTTTATTTGGATTCTATCACCAATATGCATTAAA
ATTTAATTAACTTTGTTTTCTGTTATTTTGAGCCATTTTTATTCTAACTCATTAGTGTGTGTGTGTATACACACACACACA
CACACACATATATAGATATTACTTGTAATGATTAAATTGATGTCATAATCT

>gi|40549412|ref|NM_199234.1| Homo sapiens glial cell derived neurotrophic factor (GDNF), transcript variant 3, mRNA
ATGAAGTTATGGGATGTCGTGGCTGTCTGCCTGGTGCTGCTCCACACCGCGTCCGCCTTCCCGCTGCCAACCCAGAGAA
TTCCAGAGGAAAAGGTCGGAGAGGCCAGAGGGGCAAAAACCGGGGTTGTGTCTTAACTGCAATACATTTAAATGTCAC
TGACTTGGGTCTGGGCTATGAAACCAAGGAGGAACTGATTTTTAGGTACTGCAGCGGCTCTTGCGATGCAGCTGAGACA
ACGTACGACAAAATATTGAAAAACTTATCCAGAAATAGAAGGCTGGTGAGTGACAAAGTAGGGCAGGCATGTTGCAG
ACCCATCGCCTTTGATGATGACCTGTCGTTTTTAGATGATAACCTGGTTTACCATATTCTAAGAAAGCATTCCGCTAAAA
GGTGTGGATGTATCTGA

>gi|169159808|gb|EU311537.1| Streptococcus pneumoniae strain SP231 pilus islet 2 pitA pseudogene, complete sequence; SipA (sipA), PitB (pitB), and SrtG1 (srtG1) genes, complete cds; and srtG2 pseudogene, complete sequence
AAAAAAACTCCTTTTTTATCATAAAATAGAATACAAAACTTCGGAAAAATATGATTTGAGAAATTTTATGAATTTTGAA
AAAATATAATAATTCAAATTCAGAAGACTTTCAAAAGAAATTTTCGAGCGATTAGAATAAATATATGCATCTTTATTTT

ATGAAATTTTTTACTATTTTTTCCATTTTACCACAGAAAAAGTTGAAAAACAAAGAAGAATCATACTTTTTTGTAATTTA
TAATAACATGTTTTCAAACACATTAATTGACACGAAGTTAATATTATTATACAATTATCTTATGGAAGTCTCAGATTTAT
ATAGTATCAGATTAAAACATATTCTAGCAAGGGAGATAGCATGAACGTTCAATATGATTTTAAGAAGATTCAATATTTT
ACCAGTAGTTTAGTTATCTTTCTCGCTATTCTTTTTTTGTGTGCACCAATTAATTCTTTACGTGCAGATTCAACAACTGAA
CCTCAGACAACTCTGCACAAAACGATTACTCCGATATCAGGGCAAGAAGACCAGTATGAGTTGTCACTGGATATCACA
TCTAAACTGGGAACGGAGACCCAGACGGAGACCCAGTCAGAACCCTTGGATGTAGTCTTGGTTGCCGATTTTTCAGGG
AGTATGGAAGAGCGAGATGTGTGGTCTTACTCTAGTAGACGATACATTAGTAGGATTGAAGCACTAAAACATACACTG
AAAGGTGTGAATGGTCGTCAGGGGCTCATTGATACAATTCTTTCTAATTCCCAAAACCGTCTGTCTATAGTTGGTTTTGC
CGGAAAGATTGATAATCAGTATAATGGCCGTTATTATAATGAATATTATCTGAGTTATCAATATGGAACTTGGCCAAAT
TGAGCTGGTTGGTATTCAAATATCTCTTCATATGATGATGCTAAAACTTTAGTATCTTGGAGCACGGATTCTAATAGCTC
AAAAAATATTGTTAGTTCGTTAACAATTGCTGACTCTAGTCATTCTTATGGTATGGACGCGGGCATTGGCACTGGGACA
AATATAAATGCTGGGTTAACTGAAGCTCAAAGATTGTTGCAAAGTGCAAGGGCTGGGGCAAAAAAAGTAGTTATTCTG
CTGTCAGATGGCGAAGCTAATATGTATTACGAGTCTAATAGTGGGAGAACAATATATAACTATTATTCTAATCCAAATG
TGGGACGTATGATTGATACTCCATATTGGTTTACCTCTGGTTTAGAGAGAGGAATGCTGAATATATCTAGTTTAATAGCT
CCAAAAATAGATGGCTTTTATTCAATCAAATTCAGATATATAGGTTCAAACGATAGTATCACATCTCTTAAAGGATATA
TCAGTGGTTATAATTCTGGAATCCCCAACGAAATATTTTCTGCCAATAATGAAAATGACTTGCAACAAAAATTCAAAGA
AATCACAGATAAAATTCTACCTCTAGGCGTACACCATGTAACTATATCAGATGTCTTGTCCAAGTACGTGCAGCTGTTA
CCTGGTGATGCTTCACACCTTCGTGTCGTCAAAATCAAGGATGGTAACGAGCAAGAACTGAATGACAATCAAGTTACG
ATTGAAACTAAGAAGAACGAACAGGGATTAGTGGAAGTAACAGCCAAGTTTAATCCGAGTTACACTTTGGAGGATGAC
GCCAAGTACGTTCTCAAGTTTACTGTCACCTCTAGCCAAGAGGCATTTGATGCGATTGCGGGTGATAAGACACTTACTA
GTGATGATGCCGAAGAAGCCGATGCTACTAAACTCTACTCCAACAAGGGGGCAAAAGTTGCCTATTCCTATGGTATTGG
GACCTCACGTACCAAAATAAAAGACTATTCTGAGAAGCCCACTTTCAAGCCGTCAGATCCATTGACGGTTCCTGTAGAG
ATTGAGTGGAAAGGTGTGGATGGAAAATCAAATCCATCAGCAAATCGTCCACCTAGTGTCGAATTAAACTTAAACCAA
AAGAAAGATGGAAGTATAAAGGATTCCTATCGAAAGGTCACTAGTCCAGTTCAAACGAATAGTTTTACTGAAAATACT
AGTTTTGCAAAGGTAGCTAAGGGATATGACTACGAACTGAAAGCACCAGACGCTCCGGGATACACAGTCGAAGTTCAA
AAGACAGGTACGAAAGAGAAACCATCCTTCAAAGTTATTTACCGACAGCTTCCAAGTCTCACCGTAAAGAAAATCCTA
GAAGGTGAACAATCACCTAATAAATCTTTCACAATTAATGTTACCCTTTCAGATAAGGATGGCAAGCCGATTAACGGCA
AGTTTGGGAATACAACAGTGACTAACGGGAAAGCACAGATTTCTCTCAAAAATAGTCAGGAAACTGCCCTCAGTTATC
TGCCTCGTGATACCCACTATAAGGTGGAAGAAGTAGAGAACTCTAGAACGGGATATCATGTCACCTATGAAAAACAAG
AGGGGACTTTGTCAGAGGATGTTCAAACAATCGTCACCAACCACAGACTTCCGACACTTTCAGTCACAAAAAAAGTTA
CAGGTGCTTTTGCTAATCTTCTGCAATCCTTTAAGATTACCATTAACGTAAAGGATGCGCAAAATAAACCATTGAATGG
ATCGTATAGTGCAATAGTAAATAATCAAAAAACAACGCTACAATTCACCAATGGTAAGGCGACAGTTGATCTAAAGAA
AGATAAAACCATCAAGATTCTCGACCTTCCTCTAAATGCTCGTTATAGTATCGAAGAAGAAGCAAGTTCGTCTCGTGGG
TATCAGGTGTCCTATGATAAAAAAGAAGGAACTCTTGATGCAAATAAGTCTGCGACAGTCACGAATAATAAAAACAGC
GTACCTGAAACGGGAATTGACTTCTTGAGTAGCACTCTCGTGCTTGGAGTCGTTCTTCCTCTAGGAGGGATCTTCTTTAT
CATCTTACTTGGTCACCTTGTGGTGAATAGGAGGAAGTGATGCTGCTTAAAAAGAAACATAAGAAACCAGTAACACAA
GTCAATCGGGATAAGTCTCCGCCGAGTGTCTGGGGAGATATCCTTTACTTAGTCAGTAAACTTCTGATGGTTGGATTTGT
ACTAGCCACCCTTTACTTTTTCGTCTTTGGATTATTAAGATACAATGACGATGGCATGAAGCCCGCCTTAAAAGATGGC
GACTTGGTCGTCTATTATAGGTTGGATAAACGCTATTCGATTGGTGATTTGCTAGTCTATAGTTATAAAGGTAAGGAAA
GAGTGGCGCGTGTCATAGCAACCGAAGGAAGTACAATCGATATAAACGAAAATGGTCTCATCATCAACGGTTCTCCTC
AACAAGAGCAAGATATCTACAAAGAAACGCTGCTCTATAAGGAAGGGGCAACCTTCTCGATGAAAGTCCCAGCAGGAC
AACTTTTTGTCCTCGGGGACAATCGAACAACGGCTGTAGACAGTCGTGCTTTTGGAACCATCCCTATACAGGATACTCA
AGGAAAAGTTGTAACAGTAATTAGAAGACGAGGCTTTTGATGATCATAATGAAAAAAGAAAATAAAAAAACAAAAGA
AATAATCATGAAAAAAACATTCTTTAAAAAGCTATTCACTGCAAGCATTGCAGCTATAACCGCTTTGTCCGTATTCAGA
GGTGTCCCGACTTTTGCGGATGATAATTCAGCAATAACCAAAGCAAATGGTGAAAATAATGCTGTTGTGAAGATTAATA
AAACGTTGAATATTGCAGAGGGAATAACAACACCAACAGCGACATTTACATTTAAGTTTACAGAAAAAACAGGACAAT
CTTCTAACGGTGCGCCATATCAAACCGGAGTTGCAATTCCAGATAGAAATGTAGAATACAATAAAAATGATCACCCAA
CTGCTGATAAGATTCAAAAAGCAACAGAAGACATTTTTTCGGGAGTTGCTTATGGCCATGCTGGTGAATACGTTTATGA
TGTAGCGGAAGCAAAAACTGGATGGCAGGCGATTACCAAAAATGGTAAAACAATTGATGCCATGAGATACGACAAAC
GTACATATGAAATGCACGTTATTGTTAAGAATAAAGTAAATGGTGGTGTCTATATTTCATCAGTATACTTTAAGGAAAA
TAATAAATCTAACGCCCCTAAAGTAGAACCAAGTGAACAAGGCGTTTATAATTTATTTGATAACACATATACCAAAGAC
GCAAGTAAGGAGCCTAATCCTGATGATCCGAGTCAAGTAGACCCCAATGCGAAAGCATTAACAATTACTAAAAAAGTT
GATGGAGCTTCAGGGGATAAAACAAGAGATTTCCAATTCCATATCAAGATTCAACTTCCAAGTACAAATAAAACAGCA
GAAACCCCTGTTACGAATATTATAGTAAAACATGGATCTAAGTCAGAGGTGTTGGCAGTAGTGACCCCAGCAGATACA
GTTGAGTACAATTTTACTCTTAAAGATGGTGAAACATTTACAGTTGAACAACTACCAGCAGGTTCTAAATATACAGTAA
CTGAAACTGGAGTAGCAGGTTATACAGATTCATCAATTTATACTACAAATGGTGCAGAACAAACATCTCAAGGACAAA
AAAATGTAGATTTTACATTAACAGATATCCTCATAGGTGAAAAGAAAAACGACAACA
AAGTTACTAACAAAATCGACGACGTTACTCCTACTGGTCTCTTGATTGATAACCTTCCATTCATTTTGATGATTGGTCTT
GGTTTGGCTGGATTTGTTGTCTTGTCTAAAAAACGTAGAGAAGCCTAAACGGCTGGGGGCAGATGATGAAAACCAAGC

GTGAGAAACCAAAAAAGAGTCTGTCTAGGCGTCTCGTTCTTGCTGTGGATGGGGTGATCAATCACTTGCTGCTCATTTT
TGCAGCTTTGATCTTTCTCTTTGGTTTCTACGCCCTTTGGGATTCCAACCAAGTCTACTCCTTAGCTTCGTCAAGTGAGTA
CGAAGCTTATCGACCTGTCACGACGCAACAGGATGAGCTGGCCAGTTTTTCAGGCTTCAGCAAACTCCAAGAACTCAAT
CCCGAAGTCCTCGGTTGGATCAATGTCTACGGCACCAATATCGACTATCCCTTAGTCCAAGCCAAGGACAATGAAAAGT
ATCTCAACAAGGACTCCAAAGGTGAGTTTGCAGCGACAGGCGCTATCTTTCTCGATGCACGAAATAATCCTAAGTTCGA
AGACTTTAATACCATTATCTACGGGCACCACGTAGAAAATGGGGTCATGTTTGGTGATGTGGCTAAGTTTGCTGATCAG
GAATTTTTTGACCAGCATCGTTACGGTAGTATATACTACAATGGTGTGGAAAAAGGGCTCGAGATCTTTGAGATGTTGG
AGGTTGATGCCTATGACTTTAACATCTATGATCCAGGAATACAGGGTGAGGACCGCCAGCAGGCCTATCTAGACCACCT
GCTCTCAGTCGCCATGCACAAGCGGGATATCTCACTCTCACCGAGTGATCGTATCATCCTACTCAGTACCTGTTTTCTCG
ATGTGACCAATGGTCGTCATATCGTAGTCGCAAAGATTACAGACACTGTCCCTAAAAATACTTTCCATACAAAAAAATC
AAAACCATTTCCATACAGTGTCTTTGATGACTCGTCTCTTGGACGTTTCCTCTCATCAATCCCACTATGGATTTGGTACC
TTATCTTGTTTGTATTGTTCTTGCTCTTGATTTTCTTACTCCTTGTCCTCTACTTGATCCTACGTCGTAGAAGAGAGTA
AAAAAAATGCAAGAAGCAGACCCTTTTACTGACTAAGGGTGAATAGAAAAGAGGCGGTTTTGTCCCGCCTCCTTTCCTT
TTAAGAAATCTCGCATAGAAAGGAGTGCTGATGACGGTTCAAAAGAGCGCGATTTAAAAACGTATTTCTGGTATTCTTC
TGTGTTTTTGTAGCTCTTTTTAGTTGGCAGAGAGTAGTAGAAGCAAGTGACTATGATCACTATAATCCTATTGAAAAGG
ATGCTTCGAGCACAGGTTTTGAAACCCTACAGCACTTGAACAAAGATGTTTGCGGTTGGATTAGCCTTGATGGGACCAA
GGTAGACTATCCGCTTCTACAAAGTCAGGATAATGTCAAATACCTTGACCGCAATGCCTTTGGCGATTATACGATAACG
GGATCAATTTTTCTCGACTATCGCTTTAATCCCAACTTTACTGATTTTAATACGATCATCTACGGACACTCTATGGCTTC
AGGGGCTATGTTCGGTGAGATTAAGAAATTTGCTGATAAGGAATTCTTCGACCAGCATCGCTACGGTTCTATCTACTAC
AATGGTCGAGAACGTGGTCTTGAAATTTTTGGGATTTTAGAAGTGGATGCCTATGACACGGAGATTTATCGAACCTTGA
GTTCCAAGGATGAGGAACACCAGGCTTACTATCAATATCTGCTAAGTAAAGCCAAGTACAAGCGAGATGTTTCCTTGAC
CTCAACGGACAAGATTGTTTTATTAAGCACCTGTTTCCTTAACATAACTAATGGACGACATATCCTTTTAGCTAAAATAA
CGGATGCACCGGTGAAAGTAGCACAGGATGAAAGTGGGGAAGCAGTAGGAACACGGTATTTCGATCACGGAATTCCC
ACTTCTTGTCCTTATCATCCTAATTATAAGACGAGATAAAAAAACAAAAGAGCAAAAGAAATAGTAGACGTAGGAAAC
GTTTACTATTTTTTTCTTTCATGATAGAATGGATAGAGAAAATAATATCATAAGGATGATAAC

>gi|225007537|ref|NR_027353.1| Homo sapiens CD8a molecule (CD8A), transcript variant 4, transcribed RNA
AGCACCCAAGGGCTGGTCAACCAAGCTGGGGGTTGAATTTCCATCCAGCAATGCAGGCCATGGGAGGCTGCAGCAGTG
ACGCTGTCAGATCCCCTTTGTGAGAATAATAATTTTTATAACAACGTGGCTGGAGGACTGATCAGGAGAGAGACTGGTG
TGAATTGAAGGCTGTTGCAATGGCTCCAAGAAGAGATGAGGCTGTGTGTTTTCAGCTGCCCCCAGTTGCCTGGCCAGGC
TGCCTCGACGGCCCTATTCACGGGCCCCAGCCTCCTCGCCGGGCTGGAAGGCGACAACCGCGAAAAGGAGGGTGACTC
TCCTCGGCGGGGGCTTCGGGTGACATCACATCCTCCAAATGCGAAATCAGGCTCCGGGCCGGCCGAAGGGCGCAACTT
TCCCCCCTCGGCGCCCCACCGGCTCCCGCGCGCCTCCCCTCGCGCCCGAGCTTCGAGCCAAGCAGCGTCCTGGGGAGCG
CGTCATGGCCTTACCAGTGACCGCCTTGCTCCTGCCGCTGGCCTTGCTGCTCCACGCCGCCAGGCCGAGCCAGTTCCGG
GTGTCGCCGCTGGATCGGACCTGGAACCTGGGCGAGACAGTGGAGCTGAAGTGCCAGGTGCTGCTGTCCAACCCGACG
TCGGGCTGCTCGTGGCTCTTCCAGCCGCGCGGCGCCGCCGCCAGTCCCACCTTCCTCCTATACCTCTCCCAAAACAAGC
CCAAGGCGGCCGAGGGGCTGGACACCCAGCGGTTCTCGGGCAAGAGGTTGGGGGACACCTTCGTCCTCACCCTGAGCG
ACTTCCGCCGAGAGAACGAGGGCTACTATTTCTGCTCGGCCCTGAGCAACTCCATCATGTACTTCAGCCACTTCGTGCC
GGTCTTCCTGCCAGCGAAGCCCACCACGACGCCAGCGCCGCGACCACCAACACCGGCGCCCACCATCGCGTCGCAGCC
CCTGTCCCTGCGCCCAGAGGCGTGCCGGCCAGCGGCGGGGGGCGCAGTGCACACGAGGGGGCTGGACTTCGCCTGTGA
TATCTACATCTGGGCGCCCTTGGCCGGGACTTGTGGGGTCCTTCTCCTGTCACTGGTTATCACCCTTTACTGCAACCACA
GGAACCGAAGACGTGTTTGCAAATGTCCCCGGCCTGTGGTCAAATCGGGAGACAAGCCCAGCCTTTCGGCGAGATACG
TCTAACCCTGTGCAACAGCCACTACATTACTTCAAACTGAGATCCTTCCTTTTGAGGGAGCAAGTCCTTCCCTTTCATTT
TTTCCAGTCTTCCTCCCTGTGTATTCATTCTCATGATTATTATTTTAGTGGGGGCGGGGTGGGAAAGATTACTTTTTCTTT
ATGTGTTTGACGGGAAACAAAACTAGGTAAAATCTACAGTACACCACAAGGGTCACAATACTGTTGTGCGCACATCGC
GGTAGGGCGTGGAAAGGGGCAGGCCAGAGCTACCCGCAGAGTTCTCAGAATCATGCTGAGAGAGCTGGAGGCACCCA
TGCCATCTCAACCTCTTCCCCGCCCGTTTTACAAAGGGGGAGGCTAAAGCCCAGAGACAGCTTGATCAAAGGCACACA
GCAAGTCAGGGTTGGAGCAGTAGCTGGAGGGACCTTGTCTCCCAGCTCAGGGCTCTTTCCTCCACACCATTCAGGTCTT
TCTTTCCGAGGCCCCTGTCTCAGGGTGAGGTGCTTGAGTCTCCAACGGCAAGGGAACAAGTACTTCTTGATACCTGGGA
TACTGTGCCCAGAGCCTCGAGGAGGTAATGAATTAAAGAAGAGAACTGCCTTTGGCAGAGTTCTATAATGTAAACAAT
ATCAGACTTTTTTTTTTTATAATCAAGCCTAAAATTGTATAGACCTAAAATAAAATGAAGTGGTGAGCTTAACCCTGGA
AAATGAATCCCTCTATCTCTAAAGAAAATCTCTGTGAAACCCCTATGTGGAGGCGGAATTGCTCTCCCAGCCCTTGCAT
TGCAGAGGGGCCCATGAAAGAGGACAGGCTACCCCTTTACAAATAGAATTTGAGCATCAGTGAGGTTAAACTAAGGCC
CTCTTGAATCTCTGAATTTGAGATACAAACATGTTCCTGGGATCACTGATGACTTTTTATACTTTGTAAAGACAATTGTT
GGAGAGCCCCTCACACAGCCCTGGCCTCTGCTCAACTAGCAGATACAGGGATGAGGCAGACCTGACTCTCTTAAGGAG
GCTGAGAGCCCAAACTGCTGTCCCAAACATGCACTTCCTTGCTTAAGGTATGGTACAAGCAATGCCTGCCCATTGGAGA
GAAAAAACTTAAGTAGATAAGGAAATAAGAACCACTCATAATTCTTCACCTTAGGAATAATCTCCTGTTAATATGGTGT
ACATTCTTCCTGATTATTTTCTACACATACATGTAAAATATGTCTTTCTTTTTTAAATAGGGTTGTACTATGCTGTTATGA

GTGGCTTTAATGAATAAACATTTGTAGCATCCTCTTTAATGGGTAAACAGCATCCGAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAA

>gi|219485608|ref|XM_002239788.1| Branchiostoma floridae hypothetical protein (BRAFLDRAFT_102194) mRNA, complete cds

ATGGCGGCATCAGCAGGTGCAGTGTTCTTACTTCTAGTGTACACGTCATCAGTTCATGGGTATCCAAGACAAAGATTTC
GAGTCTACCTTGACAGTGATGCTTCTGCCGTTGCCGACTGTGCGTTTGAAGGCAACAGCCTGTGTTCCTGGACCCAGGG
ACAGGACGACACCATGGACTGGTCTGTCCTACAGGGATGGGACGCTACAGGCGACACCTACTGGACGACGTTGCCAAG
CACGGACCACACCCTAGCAACAAGTGACGGCTGGTACGTAGCCGTAGAATCTCCCGCCACAGTCACAACCGGTCGGCC
ACTCGCACGCCTGCAGTCCCCAGTAATCCCGCATGGTGAAAAGGCTGAGCAGTGTTTGACGTTTCATTACCACATGACC
GAGGACTACCCCGCTGATCTGGGCTTCGATGCAAAGAACTTGCTGAATGTGTACCTGCTATCGGAGGGCGAGTTGAAG
GGACCGATTTGGACTATGGGAGAACTGCTCTATGGCTACTGGAACCAGGCAGAACTTTCTTTCACCACATATAAGGACT
TTCAGATCGTGCTGGAGGCGGTGGTAGAGTCCCAGAATAGCGGTGACATCGCTCTTGATGACATCAAACTCAGAAATC
GGCCGTGTGCAGGTTGTAATCACATCCTGACGGACCGTACGGGCACGTTCACCAGCCCGCGCTACCCGGACCCCTACCC
GCAGAACACGGACTGCACCTGGACCATCAAGGCGCCCGACGACAAGAGGATCCGCCTGGCCTTCGACCTCATCGACAT
CGTGGAGGACGAGAACTGCGAGATCGACTACGTCGCCGTGAGCGGGTCGGAGAACGGGGACGGCATCAGGTAA

These homologues were BLAST searched in *Leishmania Major* GeneDatabase and looked for homology with any of the sequences in the database. It was found that one of the homologue showed 58% of homology with one of the sequence in *L.major* GeneDatabase. Branchiostoma floridae hypothetical protein mRNA showed 58% of homology with ubiquitin ligase, putative mRNA and this sequence was used in Primer3 to design the primers for the PCR reaction.
The FASTA format of the sequence is mentioned below.

>gi|157865079|ref|XM_001681196.1| Leishmania major ubiquitin ligase, putative (LmjF09.0740) mRNA, complete cds
ATGGGCTACAACCCCGACATCCCGTCGACGTGGAGCATCAAGCAGCTCCTTGAGTACGCGCGGCAGCGCGGAATCGAC
GTACCGAGCGGCTGTGAGAAGTCGGAGCTCGTCACCATCGTCGACGAGTACTTGACAGACGAACTGCTCGCGGAGCAC
ATCCGCAAAGAAAACGCACTGGACACGGACGTCATTGAAGACATCGCAGGCGCTAGCACCGCGTCCATTCTACGGCAT
GCGGCGCACCCCTCCTATACCAGCGTGCTGCGTCGCCATCCAGCAAAACTGCAAGCGATTATTGAAGCCGGCATCGAG
CAGCTCCAATCGCTTCCGTTTGTTGCACGGCTATGCAGTGCGCGAGTGCACAGTCCTCTGCAGGCGTGGCTGGGCACCT
TAGCCTCGGACTGCGGCGATGGCAAGTGTCTGCTGACTGTCCTCACAGCGTCAAACAGCGACGTAGCAGCCTCGCCGG
TGTGTGGGCGCTTCTGTGGAGAGCAAGAAATGGTGGTACGCTGCCTCGACTGTGGCGCCGACTCCACATGCGTCATGTG
CATGGACTGCTTCCGCCACTCTCCGTGCGTCAACCATCGCTACCGTATCACGCAGAGCTCAGGCGGTGGCATGTGCGAC
TGCGGCGATCCGACCGCGTGGAAGCTGGAGTCCTTCTGCAGCCGTCATCGTGCCGCCGCTGCAGCTGCAGCGGCGGGT
GACGGGAACAGTGTGTCCGACCCCCTCGACACCATGGCGGAAGAGGATCGGGTGTGGGCGGTGCCGGTGCTGCGAGGC
ATTATTCAGTTCACTGCTGTTACGCTCCTGCAGCACACGCGTCTGCAGGTGTTGCAGAAGCGGGCGCAGTCACGAGCGA
GAAAGCAGCGAAGGGCGGCAAGCGGCGCCGGGAACAGCACTGCTAGTGGCAGCGGCACTAGCACAGCCGGAACTGCA
GCAACGCCCAGCTTTAGCGACGCGGAAGCCGCGCTGCTGAACGATGTTGCTCGCATCTCGGATAAGTTCGACTGGTTAG
AGGAGTGGATGTCAGAAGTGACGCGGCAGCTGTGGGAGCTCATGCAGGCCAGCGACATCGCCAAGCATCTCGTGGCGC
AGCTCTGGGCAGAGCCGGTGCGAATGGCGGTGTGCGGCGGTCCAGGCGCAGTTGTCACCGCCGCGCACGACAACGAGG
AAGAAAAAATAGGCGGGGAGACCGGCACGGCCGCCTCGATGGCGGACCCCGTCCTCGACGCGTTACCGCCAAACTTCA
CCTGCGTGCACATGGTGTTCCTTCACGAAGCATCGCGGCCCGTGACGGCCCACGAGCAGCCTCTGAATGCCACCAGCA
GATCCGCGCCGTGGCTCAATAACCTCCTTCGGTGCGTCGGCCACTGCTTGAGCAACGCCAAGTTTCGCTTCCCGGTCGG
CGGGCTAATGGCGGCATATGCCGAGTACATCAACGCACAGGAGCGTTCTGAGGACAAGAACGAGTGCTTGTCGCAGTA
CCAGGTGCAGGTGCTAACGAACCCCGACGTGATCCGCCACCTGATGACACCTTCACACCTGCCTTACAAGGCGGCAAC
GAACACAGTGCTGCACAGACAGCTGAGCGCGTTGCTCTACGCGTTATGGCTCGCGCGTGACACCGTCAGTGGCGCCTCT
GCGAACCCGGTGACGGAAACGGCGGCACTCTCGCGCGACGGGCTCCTGTTTGTTGCCCACCAAGCCACGCCCGTCTCG
CTGGCCAGCGTGACCGCGTTACAGTCGTGTCTGAGCGGGTCCCCTGCGGCCTGCTACACCCTTGTGCTGAACCGACTCG
CGTGGCGCGCCTTTGCGCAGGTCGCCGGCAGTCTGGCTATCAGCGGCTACATCTGCAAGGACCCCAATGCACCAACAG
ACTCGCAGCACTACCCCACGCCGGACTTCATTGGTCGCTGCATCTCGCGGCACATGTGGTTGTGGGTGCATGCCATGCG
GGTGGTTCTGTCATGTCTTCTGAGACTGCCTGCAACAGCAGCAACGGATTCATCGGCGTCGATGCCCGCCCTGCCCCAC
GATTCCATCGCAGAGGTGCTGGGGATCACGACGGAGCTCTGGGAAACGTGGCGCGGTGCCACCGCCTTCGATACACCA
CCGGATCGCCTGCGGCAGAAAGGCCCGCTGCTGGAGATGCTGTCGGCATGCAGGGTGGCGAATGGGAAGGACGAAGC
CCTCTTGCTGCTGGAGCACGGCCCCATGGCGGCGGCGCAGTCGCGCGTACTGGGAAAGCAGCCATTGGCGGAGCAGGT
AGGAACGGCAGTGACCGCCAAAACAAGCATCCTGGCGGCCCCTGCGTTGGCGACGGCGATGGGGTACACAATACAGG
CTCTCTATGAGATTAGCCGGACCATGGATGTCGTCTTCGAAAAGCAGCGCGACGGCTTTTGCCGGCGGTGTCGCTCCAT
CGTGCTGGAGCGAAAGACCTTCATGGGTCTGCCGAGAAGCCAACAAGCAGCGAGCCATGGTGGCACTTCAGAAAATGA
GGAGGCGCGCGACGCAGACGGGAAGGCAGGTGCAAGGCTCCCGGCGCCGTCGACGTGCTCGGTGGACATTGTGGAGT

ACACGCTGCTGGAGCCACACGCGCACGCCACTACCCTTTCCAACCACCTCCCTCGCCTATTCGGGGCTGTTCTCACGGC
ATGGGTCATCGAAAGCCAACGCGCCACATCTCGGCTGCAAAGTGCGGCGCCATCGTCCTCGGCCGCCCCGGCGGAGGT
GCTCGCCATGACTCGCATATCTGGCCTTCCAAGGGCGAGCACGAACAGCACCGCAGCAGCAGGGGAGACGTCGGACCT
CCCGGCCACCCCCGCACACGCGCGCCGCCCTCTGCGGTCTCTGCTGGACGCCTTCTTTCAGTTGCGAGCGGAGGTAACA
CACAACGAGCAGGATCGCCTCACCTTTTCTCAGCAGCTCCTCGATTCGTTGGTCATGCCGCATGTGCTGACCGGCCAGG
TCTTCGACGGCTTGTGGCGGCGCGGCGATTACGATGTTTCATCTGCCGTCACGATGTACCTGACCTTCTCGCGCGGCGTG
TCGGCAGAATTCGACATCCTGATGATGCAGGTGCTCACGATGGAGCTGGCACCGGCCGACATGTCGCTGCAGATCCTCC
AGCGCTTCTCTTACGCGAAGCGACGCCTGCGCAACGTGGACACAGAGCATGCGCATCAGGCGCCGCGGCGACGTGCCA
GCGCATTGCACGGGCAGCGACCAGCGGATAAGCTCCCTGAGAAGCCCTTGGACGCCCTCCCAACGATGCTGGCAGCGC
GGCGGCCGTTCTTCCGGCGCTGCCTGAACGGCTACAGGCACTTTTTGCGCCTCATTCTCACGATCGTGACGGACGTGTC
CAAGGCGGCATTCCAGGCCCCCATGTCCTCGCCCGTCATCGACCGCATCGTCGGGGGCCTCCTCGCGCACGGTAAGGTC
TCCCACTCCTCCATCGTCGCCAACGTCAATGGAACCGCCCGTGGCGGCGACTACGGCGACGCCGACGAGGACGAGAAG
GACGCCAGCGGCAACGTGATCAAGTTCTCGAGGCTGATCGACGCCGCGATCCGGAAGCTGGCGGTCGCCGAGAACTCG
GCGCAAGGGAAACAGTTTCAACTGAAGGACGTGGATGTGTGGCGCACACACGTGGGCCTGTACCACATCGGGGTCGTT
GATAGCCATCTCGAGGAGTTTTACAAGACGTACCGAGGCCTCGCCAGCGCAGCAGATGCGGCAAAGCAGAAAGAGGC
AGAGTCGCATGCGCTCGCTGATGGCGACAAGAGCGACAATGCCACGGGCGGTGGCGGGCAGCAGCAGAAACGCATCT
CTCTGCCACCCACGCAACTGTCCGACACGAGCATGTACGCAGAGCTAGTGCCCACCACCCGTGCGTTGCTGCACACGG
ACGCCGTCCTGTTGCCGGCGCTCTACGTACTGCACTTGTACGCCAACTGCCACGTGCCCGCTGCTGTCGCTGCCGCCTCG
GATGCAAAGGCTGCAGTGGCAGACAAGGAGGGTACGGCGTCCACGGCTGCCTTGGCACAGGAGTCGCCGCTCTCTCGG
CACGACAACGCTATGCACGAGAGTGCCGATGATGGAAGGGAAGACACCGAGAGCGGGAGCGAGGATGACCGCCACAG
TGACGACAGGGAAGACCCCAACGTGATCACGCGCGAGGCGTTGCTGCACGCCACCACAACTCTGTACCTGTGCGTGCA
AGACTGCGTCTCCATCACGCGGGCCATGCGGGCTATCGAAGGCTACAGTGGCGAGATGGCGTCCGCTACGGAGAAGGG
CGTCATCTCCTGGGACCTCCTCGAGCTGTACCTGCGTCGATTCGCCATCAACGCGCCCTCCTTCACTCACGCCTCGTGTG
TGCAGTGGCTGCCATTGCCGGAGCTGGTGCCGTGCCAAACCTTGCTCGAGAAGCTGCAGACCCCTGTTCGACTCCACAC
ACACACCTCAAACGGCACGCCCACCGCCATCGCCACTTCTTCCTCCGCTGACATGCTGCACCGTCTGCGTGGCTATCTG
CTCTCGAACAAGGACACCGATCCTTACGGGTGCCTGGAGATGGTGGAGGCGGTGCTCGTGCAGACGGGACTGGCCACC
TTTAGCACAGCCGACTTGCTGGATGAGCAGAAGACGAAGGATGAGGCGGCGCAGAGTCGGAAGAAGGTCATTCAGGA
GCGGCAGGCCTTGTTGATGCAGCGCATGCGCAGTCGCGCGCTGAAGGCATCGGAGAAGATGAAGGTGGCCGCCACCGC
GGCGCAGCAGCAAGAAGGGCAACGCGGTGGCGACGAGGCCTCTGGCCCCGCGGCA
TCGGCGTCGGAGAAGTCCCCTACGGGCGGTGTGATTGGCTCACTTCTTGCGAAGCTGCTGCTGGAGCTGACGACGTTGG
ACTGCTGCGTGTGCCGCTCCACGACGGAGGAGCCACTCTTTCTGCTCTGCCACACCGGCACCTCCGGCATGCTGCCTCA
GCTGGGCGCGCTTTCGCTGCCGGACGGCCGCCATGTGCACAGCCACCTGAGCATGTGCGGCCACGCCGCACACAAGTC
GTGCGTGGAGAAAGTGTTTGTCCGGCTGGCAGTGCTGTGGCAGCGGTGGAACTTCCGGAGTCAGTTCTACCTCGGCCCC
ACGGAGTTCAACTGCCCGGTCTGCACAACCATCATTACAGCGCTGTGCCCGATGCCGGTGCTCTCCTGCGGCGGCAACG
GCGACTGGACGACCGCCACCTCCCGGACAACGCGAATGCTCTCGTCGGCTGCGACTCCCTTTGCGTCTCTCTTCGAGGA
ACTGCAGAACGGCACCGTATCGGCCAGCAACCAGCGCGCCGTCGACGTCGCGGCGGAGTTCCACACAACCTTGGCAAA
CGCCGCCGTCGGCTTCAGCCCGTCCGAAGATGTCCCCGCCATTGTGTCCGCCGAAGACGAGCTCCAGCAGAAGAACGA
GGCCGCCTGGCAGCTGTCCGAGGCCATCCGCGCCTTCTGTTACGCCTGCCACCTGCAGCTGGAGGCCATCAAGGCCGGC
CAAGAACTCAGTCATCGTGACCTCATCGGGCTTCTGTCGGTCCTAGTGAGCATCATGCCCGCTGAACTGGAGCGGCAGC
AAGCCGACCTGCGCTCCAACTACGCACGAAACATGCAGGACAACGAGAGCTTGCTGGTCATGGATGCGCTTCTGCAGC
CCCGCGACGCGAGCAGATTGATCAGTGCCCACGTGCTGACACGGGTGCTGCCCTCTATGCCGGTGAACGTTATAGCAC
GGCTCGTGGCTGTTGCGTGTGCCAGCAACACGGACGACAACGGCGAAGGGTCACATGCCGCCGCCGCCGCCGCCGCGT
CGCTGGAGGAGGAGTTCACGGGCGTGACGGCGGCGCTTTGGCAAACCGTGGGCGTGCTGACACTCCTCAAGGCCCTTG
TCGTCGAAGATGCCCATCATGGCGTCGTCCTCCGCAGCAGCACTTTCACCGTGGCGGGCGCCGTCACCTTTGCGGTGCC
GAGCGTGGCCAGTGTGAGGACTCCGGTATCCCGCTGCACCTCGATTGTGCGCATGCTGCAGTACCTGTTGCCAGTGCCG
CACAAGAGCACGGAGGCGGAGCTGCAGGTGGTTGGCCAAGATATATTGGCCTGCATGCAGGCGGCCACCACACCAGCC
ACCCTCTCCTGCTCCGGACATGGTATGGAAATCGTGAAGTCAGCCGTGCCCTACACCACACCAGAGGAGTGGGTCGCG
CGGATCGCGGAGGAGCTGCTGCACCTGCCCTCCGTGTACACCACCGTTCTAACACGCTTTGCAGAGCACAAGCTGTGCG
CCATATGCCACCAAGAGCCGACGAAGCCAGTGGTGTGCTGCCGATGCGGGAAGCTGATGTGCATGCAGCCGCCAAACT
CGCCACCGGAGCTGTACACGCACACGCGCACCTGCGGCGGCGCTGTTGGCATCTTCCTCGTTGTCCGTACCGCCAACTT
CTACGTGCTGGAGCTGACGTCTGGGCGTGTCTACCACTACCCCTCCAACTACACGGACGAGTACGGCGAGCACGACCG
CAACCTGCGTCGTGGCATTCCGCTCTTCCGCAACACCGAGGAGACCCAGAAACTGGTCTTGACGTGGATGCTGAACAA
GTGGGGTGCCGTATCGTCCATTTTCGGCGTCTCAAACCGCATGGATCTCTCCACCCTGTAG

ACKNOWLEDGEMENT

I would like to express the deepest appreciation to my project supervisor, Dr. Stephen Heath who has been a constant source of knowledge support during the entire course of this work. Without his guidance and persistent help, this dissertation would not have been possible.

I would also like to convey my sincere thanks to my family and friends for their laudable support both morally and technically.

Thiruvarutchelvan S