BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 9(16)

A comprehensive analysis of LACK (Leishmania homologue of receptors for activated C kinase) in the context of Visceral Leishmaniasis

Sukrat Sinha*, Abhay Kumar & Shanthy Sundaram

Centre for Biotechnology, University of Allahabad, Allahabad, U.P-211002, India; Sukrat Sinha – Email: sukratsinha@gmail.com; *Corresponding author

Received September 03, 2013; Accepted September 05, 2013; Published September 23, 2013

Abstract:
The *Leishmania* homologue of activated C kinase (LACK) a known T cell epitope from soluble *Leishmania* antigens (SLA) that confers protection against *Leishmania* challenge. This antigen has been found to be highly conserved among *Leishmania* strains. LACK has been shown to be protective against *L. donovani* challenge. A comprehensive analysis of several LACK sequences was completed. The analysis shows a high level of conservation, lower variability and higher antigenicity in specific portions of the LACK protein. This information provides insights for the potential consideration of LACK as a putative candidate in the context of visceral leishmaniasis vaccine target.

Keywords: Next-generation sequencing, Genome Assembly, Bioinformatics.

Background:
Leishmaniasis is an infectious disease complex caused by several species that are members of the protozoan parasite genus *Leishmania*. In humans, disease manifestation ranges from self-healing cutaneous lesions to life threatening visceral leishmaniasis (VL). This disease complex affects 12 million people, and there are 1.5 million new cases annually **[1, 2]**. Visceral Leishmaniasis known as kala-azar, black fever and dumdum fever and it is the most severe form of leishmaniasis. VL, caused by *Leishmania donovoni, Leishmania infantum* and *Leishmania chagasi*, remains the main agent of morbidity and mortality in leishmaniasis. The parasite has a simple life cycle, and abundant clinical and experimental evidence indicates that of all the parasitic diseases, leishmaniasis in particular should be an appropriate target for effective control through vaccination. There are, however, no vaccines in routine use against any form of the disease **[3, 2]**. Currently available vaccines against a variety of infectious diseases mediate protection by a long-lived humoral response through the production of antibodies. For diseases such as tuberculosis, malaria, human immunodeficiency virus infection, and leishmaniasis, however, the cellular immune response comprising primarily Th1 and CD8 effector T cells has been shown to be critical for mediating protection against the infection **[4]**. Visceral leishmaniasis is fatal if not treated and development of a vaccine with long-term immunity remains a challenge. The attachment of *Leishmania* promastigotes to macrophages, crucial for intracellular parasitism and for the outcome of the infection, has been demonstrated by many investigators to be a specific receptor-mediated event **[5-7]**. The factors that underlie the immunodominance of the LACK epitope in the I-A^d Th response remain unknown, although these are unlikely to reflect the overall abundance of the LACK (Leishmania homologue of receptors for activated C kinase) protein because other *L. major* surface proteins such as GP63 are more highly expressed [**8, 9**]. Whether the focus of the immune response on LACK is detrimental to the parasite remains unknown, but immunization with LACK is highly efficacious in

protection against subsequent infection **[10, 11, 12]**. Emphasis has been placed on the critical role of LACK **[13-16]** and lipophosphoglycan **[17]** that independently mediate parasite attachment to macrophages. Both molecules, when reconstituted into liposomes, mediate protection against cutaneous leishmaniasis and are considered as good vaccine candidates **[18]**. The current study has been undertaken with an idea to determine the regions of identity, similarity and antigenicity in the LACK protein which will help in the development of a vaccine against all forms of VL.

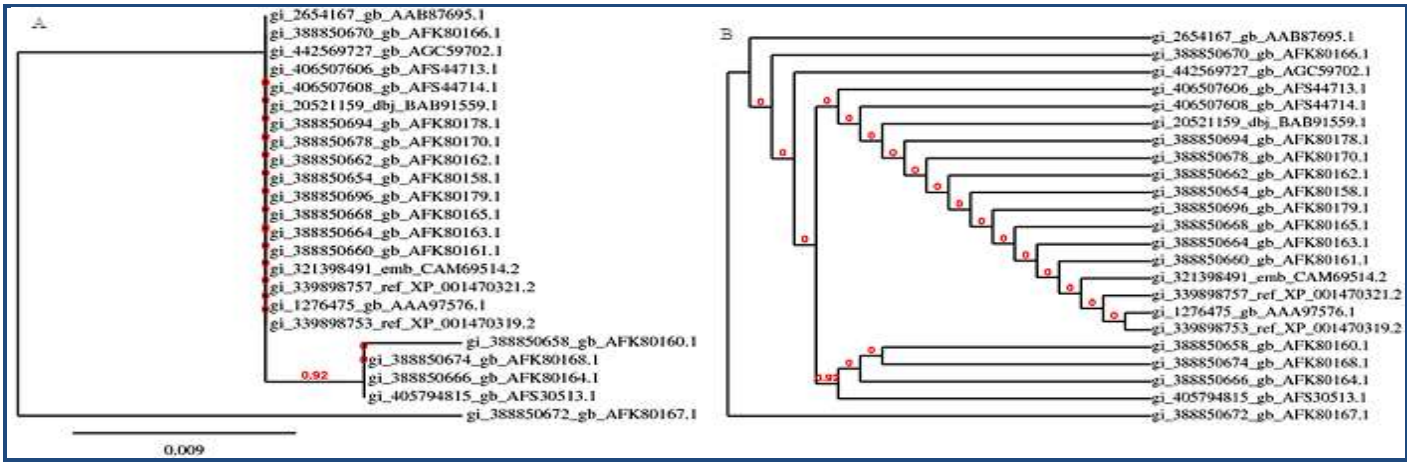

Figure 1: Panel **A)** The Phylogram showing distances in geological time scale of evolutionary relationship in LACK sequences in different species strains of *Leishmania* parasite causing visceral leishmaniasis. As can be interpreted nodes of most of the species strains of *Leishmania* strains overlap indicating homology during evolution and conservation of amino acid residues of LACK. Panel **B)** The Cladogram showing distances in geological time scale of evolutionary relationship in LACK sequences in different species strains of *Leishmania* parasite causing visceral leishmaniasis. The distances in geological time scale comes out to be approximately same in almost all the species strains which indicates lower possibility of variation during evolution.

Methodology:

Dataset Creation

In this study, we aligned the protein sequences of LACK of *Leishmania spp* associated with Visceral Leishmaniasis over with the similar prediction conditions and compared the results obtained. The protein sequences were procured from the National Center for Biotechnology Information (NCBI) through their entrez search.

Multiple Sequence Alignment

By Multiple Sequence alignment using Clustal W and T-COFFEE we arranged the twenty five different sequences of LACK of *Leishmania* species associated with Visceral Leishmaniasis obtained from NCBI to identify regions of homology that may be a consequence of functional, structural, or evolutionary relationships.

Basic Local Alignment Search Tool

Basic Local Alignment Search Tool, or BLAST, has been applied for comparing primary biological sequence information, like the amino-acid sequences of LACK obtained from different species of *Leishmania*. A BLAST search enabled us to compare a query sequence with a library or database of sequences, and identify library sequences that resemble the query sequence above a certain threshold.

Conservation

Changes at a specific position of an amino acid sequence that preserve the physico-chemical properties of the original residue.

Bit score

The value S is derived from the raw alignment score S in which the statistical properties of the scoring system used have been taken into account. Because bit scores have been normalized with respect to the scoring system, they have been used to compare alignment scores from different searches.

H

H is the relative entropy of the target and background residue frequencies. H can be thought of as a measure of the average information (in bits) available per position that distinguishes an alignment from chance. At high values of H, short alignments can be distinguished by chance, whereas at lower H values, a longer alignment may be necessary.

HSP

High-scoring segment pair. HSPs are local alignments with no gaps that achieve one of the top alignment scores in a given search

P Value

The probability of an alignment occurring with the score in question or better. The P value is calculated by relating the observed alignment score, S, to the expected distribution of HSP scores from comparisons of random sequences of the same length and composition as the query to the database. The most highly significant P values are close to 0.

PAM

Percent Accepted Mutation has been used to quantify the amount of evolutionary change in a protein sequence. 1.0 PAM unit is the amount of evolution which will change, on average, 1% of amino acids in a protein sequence. A PAM(x) substitution matrix is a look-up table in which scores for each amino acid substitution have been calculated based on the frequency of that substitution in closely related proteins that have experienced a certain amount (x) of evolutionary divergence.

Identity
The extent to which two (nucleotide or amino acid) sequences are invariant.

Similarity
The extent to which nucleotide or protein sequences are related. The extent of similarity between two sequences can be based on percent sequence identity and/or conservation. In BLAST similarity refers to a positive matrix score.

Expect value
The Expect value (E) describes the number of hits one can "expect" to see by chance when searching a database of a particular size. It decreases exponentially as the Score (S) of the match increases. Essentially, the E value describes the random background noise. The lower the E-value, or the closer it is to zero, the more "significant" the match is.

Gap
A space introduced into an alignment to compensate for insertions and deletions in one sequence relative to another.

Clustal
It is a widely used multiple sequence alignment computer program.

Dendrogram
A dendrogram (from Greek dendron "tree", -gramma "drawing") is a tree diagram used to illustrate the arrangement of the clusters produced by a clustering algorithm.

Phylogeny
A phylogenetic tree or evolutionary tree shows the evolutionary relationships among various biological species or other entities that are believed to have a common ancestor. In a phylogenetic tree, each node with descendants represented the most recent common ancestor of the descendants, and the edge lengths in some trees correspond to time estimates. We referred each node as a taxonomic unit. Internal nodes are generally called hypothetical taxonomic units (HTUs) as they cannot be directly observed.

Protein Variability
We used the following methods to predict the variability in the protein sequences of LACK *Leishmania spp* associated with Visceral Leishmaniasis.

Shannon Entropy
Shannon entropy analysis is possibly the most sensitive tool to estimate the diversity of a system. For a multiple protein sequence alignment, the Shannon entropy (H) for every position can be determined. H ranges from 0 (only one residue in present at that position) to 4.322 (all 20 residues are equally represented in that position). Typically, positions with H >2.0 are considered variable, whereas those with H < 2 are considered conserved. Highly conserved positions are those with H <1.0.

Simpson Diversity Index
The Simpson index is another diversity index calculated from genotype proportions. This index describes the chance that two genotypes sampled at random and with replacement from a community will be from the same species. The value of this index ranges between 0 and 1, the greater the value, the greater the sample diversity.

Wu-kabat Variability coefficient
The Wu-Kabat variability coefficient is a well-established descriptor of the susceptibility of an amino acid position to evolutionary replacements. It highlights stretches of accentuated amino acid variation.

Antigenicity Prediction
The antigenicity/ immunogenicity of the peptides were predicted using antigenicity index software. Several methods based on various physio-chemical properties of experimental determined epitopes (flexibility, hydrophibility, accessibility) have been published for the prediction of antigenic determinants, of which the antigenic index and Preditop are good examples. Perhaps the simplest method for the prediction of antigenic determinants is that of Kolaskar and Tongaonkar (1990), which is based on the occurrence of amino acid residues in experimentally determined epitopes **[19]**.

Discussion:
The scores of Clustal W of sequences of LACK obtained from different species strains of *Leishmania* parasite associated with visceral leishmaniasis have been found to be between 89 to 100 which predict a high level of homology and conservation and negligible percentage of gaps amongst the amino acid residues. Similarly, T-COFFEE **(Figure 2)** results show significant identity, similarity and positives towards the Good (Red) then towards Average (Yellow) and least towards Bad (Green) which indicate high level of conservation amongst residues. The Phylogram **(Figure 1 Panel A)** and Cladogram **(Figure 1 Panel B)** analyses also show tight vicinity among the LACK residues during the process of evolution since nodes are very close to each other. The Cladogram is smaller in length, it has fewer homoplasies and it is more parsimonious. These LACK sequences obtained of one branch from eighteen different species strains of *Leishmania* are more closely related as they arose from gene duplication. Also the four sequences obtained from the other branch are also a product of gene duplication. Only gi_388850672_gb_AFK80167.1 arose from a separate branch which indicates close relationship amongst the other sequences during evolution.

The Shanon Variability coefficient **(Figure 3 Panel A)** is 1 which confirms our results that variability is less and conservation is more among the residues. To confirm our findings we found that Simpson Variability coefficient **(Figure 3 Panel B)** comes out to be 0.46 which again indicates lower variability amongst the residues. The WuKabat variability index **(Figure 3 Panel C)** has a mean value less than 5 which collaborates well with above findings indicating lower tendency for mutations and variations in genotypic and phenotypic level of LACK in the considered species strains of *Leishmania*. The antigenicity plot and immunogenicity prediction **(Figure 3 Panel D)** shows that Threonine, Serine, Phenylalanine, Tryptophan, Phenylalanine, Arginine, Glutamate, Cystine, Valine and Alanine are associated with imparting immunogenicity to the LACK protein.

Figure 2: The MSA shows results of T-Coffee alignment of LACK in different species strains of *Leishmania* parasite causing visceral leishmaniasis. A major chunk of amino acid sequence lies in the red portion of T-Coffee indicating conservation of residues amongst different species and strains.

ISSN 0973-2063 (online) 0973-8894 (print)
Bioinformation 9 (16): 832-837 (2013)

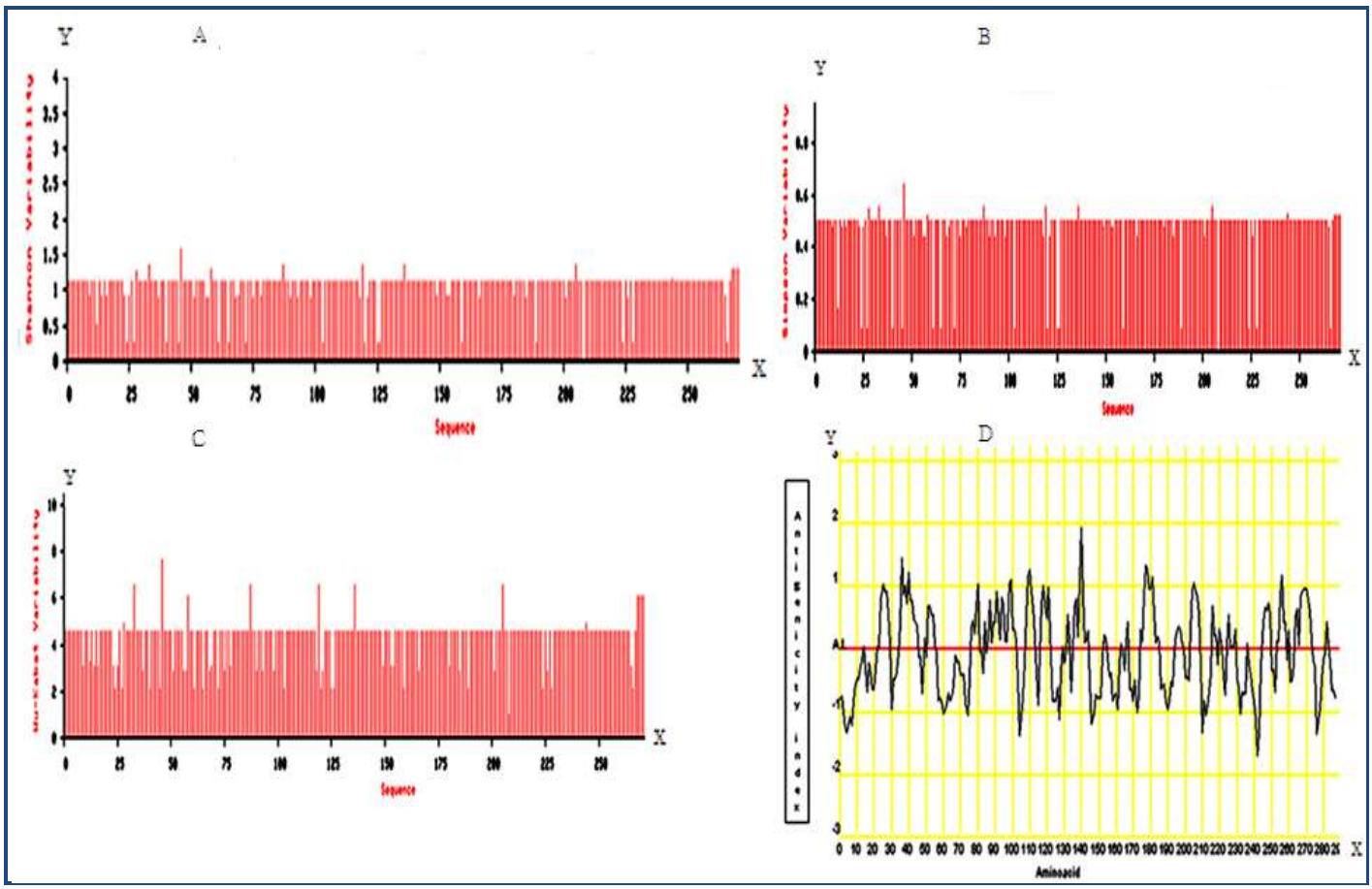

Figure 3: Panel **A)** The graph shows Shanon's variability plot of LACK in different species strains of *Leishmania* parasite causing visceral leishmaniasis. Since H<2 for most of the amino acid residues hence variability comes out be extremely low thereby indicating higher level of identity and similarity amongst the amino acid residues; Panel; **B)** The graph shows Simpson's variability plot in different species strains of *Leishmania* parasite causing visceral leishmaniasis. As can be interpreted from the results the value of variability comes out to be lower than 1, it indicated lower diversity and higher conservation amongst amino acid residues; Panel; **C)** The graph shows Wu-Kabat variability plot in different species strains of *Leishmania* parasite causing visceral leishmaniasis. It reconfirms our observation of lower variability and higher conservation amongst amino acid residues of LACK; Panel **D**) The graph shows (Antigenic index Vs Aminoacid) of LACK in *Leishmania* species in different species strains of *Leishmania* parasite causing visceral leishmaniasis. The average amino acids present at the peak on the graph having antigenic index 1or greater than 1:-25---T, 37---S, 40---S, 55---F, 80---W, 91---F, 100---F, 119---R, 121---W, 139---W, 140---V, 141---S, 179---G, 180---S, 181---Y, 205---W, 265---C, 270---A.

Conclusion:

A comprehensive analysis of LACK sequences available in GenBank is presented in this report. The data reported here show sequence conservation, lower variability and higher antigenicity among known LACK sequences. This provides ample insights for considering LACK as a putative candidate for further validation and analysis in the context of visceral leishmaniasis vaccine target.

Acknowledgement:

Sukrat Sinha is thankful to the financial support provided by the Department of Science and Technology (India) Fast Track scheme for Young Scientists No. SB/FT/LS-345/2012 and Abhay Kumar is thankful for the Department of Biotechnology (India) Postgraduate Fellowship.

References:

[1] Alexander J & Russell DG, *Advan Parasitol*. 1992 **31:** 175 [PMID: 1496927]
[2] HW Murray *et al. Lancet*. 2005 **366:** 9496 [PMID: 16257344]
[3] Khalil EA *et al. Lancet*. 2000 **356**: 9241 [PMID: 11075771]
[4] Seder RA & Hill AV, *Nature*. 2000 **406**: 6797 [PMID: 10963610]
[5] Chang KP & Fong D, *Ciba Found Symp*. 1983 **99**: 113 [PMID: 6357669]
[6] MS Klempner *et al. J J Infect Dis*. 1983 **148**: 3 [PMID: 6619572]
[7] Wilson ME & Pearson RD, *J Immunol*. 1986 **136**: 12 [PMID: 3711662]
[8] Pingel S *et al. J Exp Med*. 1999 **189**: 7 [PMID: 10190902]
[9] Etges *et al. J Biol Chem*. 1986 **261**: 9098 [PMID: 3522584]
[10] Tapia *et al. Microbes Infect*. 2003 **5**: 73 [PMID: 12650765]
[11] Gonzalo *et al. Microbes Infect*. 2001 **3**: 701 [PMID: 11489418]
[12] Gurunathan *et al. J Exp Med*. 1997 **186**: 1137 [PMID: 9314562]
[13] Russell DG & Wilhelm H, *J Immunol*. 1986 **136**: 7 [PMID: 3950420]
[14] Chang CS & Chang KP, *Proc Natl Acad Sci U S A*. 1986 **83**: 1 [PMID: 3079902]

[15] Fong D & Chang KP, *Proc Natl Acad Sci U S A*. 1982 **79**: 23 [PMID: 6961414]
[16] DG Russell & SD Wright, *J Exp Med*. 1988 **168**: 1 [PMID: 3294332]
[17] Handman E & Goding JW, *EMBO J*. 1985 **4**: 2 [PMID: 4018028]
[18] Russell DG & Alexander J, *J Immunol*. 1988 **140**: 4 [PMID: 3257774]
[19] Kolaskar AS & Tongaonkar PC, *FEBS Lett*. 1990 **276**: 1 [PMID: 1702393]

Edited by P Kangueane
Citation: **Sinha *et al*.** Bioinformation 9(16): 832-837 (2013)