Bioinformation Volume 5 *open access*

www.bioinformation.net Issue 8 **Hypothesis**

A gp63 based vaccine candidate against Visceral Leishmaniasis

Sukrat Sinha*, Shanthy Sundaram, Anand Prakash Singh, Ashutosh Tripathi

Center for Biotechnology, University of Allahabad, Allahabad, Uttar Pradesh, India-211002; Sukrat Sinha- Email: sukratsinha@gmail.com; *Corresponding author.

Received November 16, 2010; Accepted November 24, 2010; Published January 22, 2011

Abstract:
Visceral leishmaniasis is a macrophage associated disorder which leads to a profound decrease in the natural immunotherapeutic potential of the infected subjects to combat the disease. The major surface glycoprotein gp63 has been found to be a significant vaccine candidate against visceral leishmaniasis. The current study addresses the levels of similarity and identity in the gp63 obtained from different species of Leishmania viz donovoni, chagasi and infantum linked to the cause of visceral leishmaniasis. The results from BLAST, Phylogram and Cladogram studies indicate significant identity, similarity and conservation of important residues in the protein which lead us to conclude that a common gp63 based vaccine can be used as a therapeutical tool against visceral leishmaniasis caused by different species strains of leishmania.

Keywords: Visceral leishmaniasis (VL), Glycoprotein 63(gp63), Basic Local Alignment Search Tool (BLAST), Cladogram, Phylogram

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]**. 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 **[2, 1]**. 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 **[3]**.

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 **[4-6]**. Emphasis has been placed on the critical role of two abundant surface molecules gp63 **[7-10]** and lipophosphoglycan **[11]** 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 **[12]**. The current study has been undertaken with an idea to determine the regions of identity in the gp63 protein which will help in the development of a vaccine against all forms of VL.

Methodology:
By Sequence alignment we arranged the sequences of protein obtained from NCBI to identify regions of similarity that may be a consequence of functional, structural, or evolutionary relationships between the sequences.

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 gp63 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.

Bioinformation **Volume 5** *open access*

www.bioinformation.net Issue 8 **Hypothesis**

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

ClustalW: command line interface.
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 represents the most recent common ancestor of the descendants, and the edge lengths in some trees correspond to time estimates. Each node is called a taxonomic unit. Internal nodes are generally called hypothetical taxonomic units (HTUs) as they cannot be directly observed.

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.

Discussion:
BLAST **(Table 1 see supplementary material)** shows a high level of similarity and identity almost in the range of more than 90% and negligible percentage of gaps among the amino acid residues of the gp63 molecule. T-COFFEE **(Figure 2)** results show high level of identity, similarity and positives almost towards Good then towards Average and least towards Bad which again strongly advocate the high level of conservation with in aminoacid residues of gp3. CLADOGRAM and PHYLOGRAM **(Figure 1)** analysis also show tight vicinity among the gp63 residues during the process of evolution since the nodes are very close to each other. The Cladogram is lower in length, it has fewer homoplasies and it is more parsimonious. CAB51784.1 and CAB42815.1 are more closely related and both of them are close towards CAB42816.1. Moreover CAB51783.1 is related to all three. In the other branch CAB51797.2 and CAB51794.1 are closely related. They are further close to CAB51793.1. These gp63 proteins obtained from seven different species strains of Leishmania are more closely related amongst themselves as they arose from gene duplication. Moreover the sequence CAB06018.1 has altogether a different branch and it has relatively much less similarity with other gp3 proteins. The scale of Phylogram is 0.3 and it also strongly supports the results of Cladogram analysis.

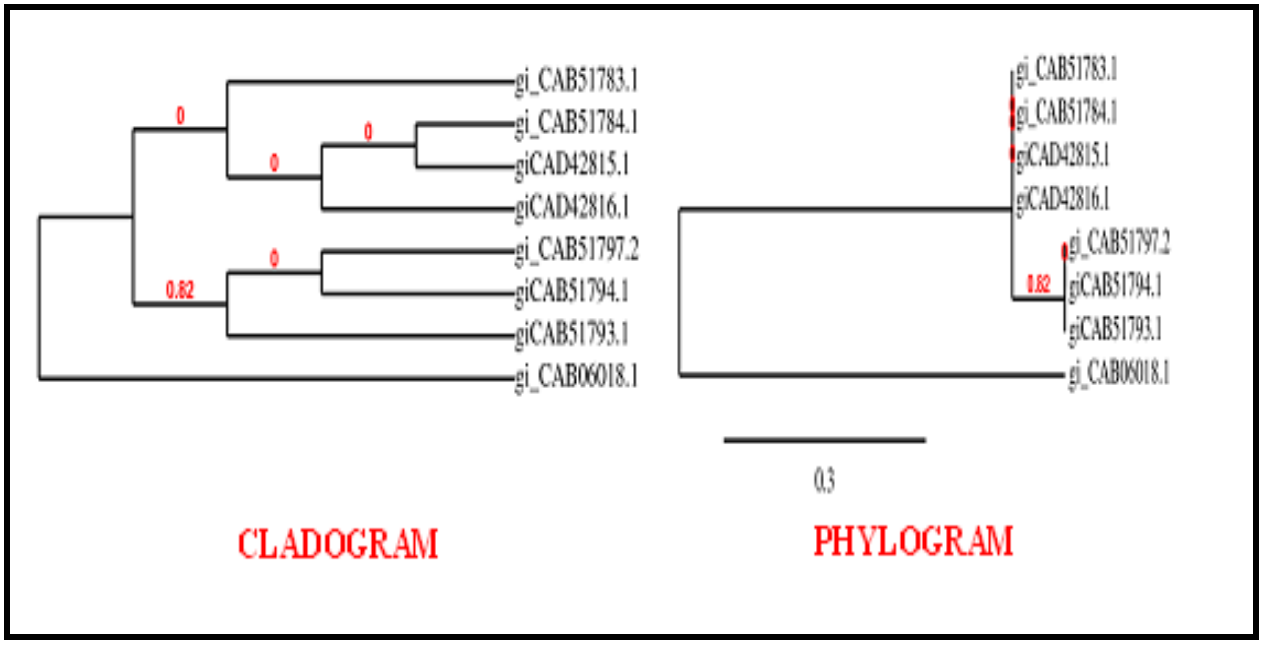
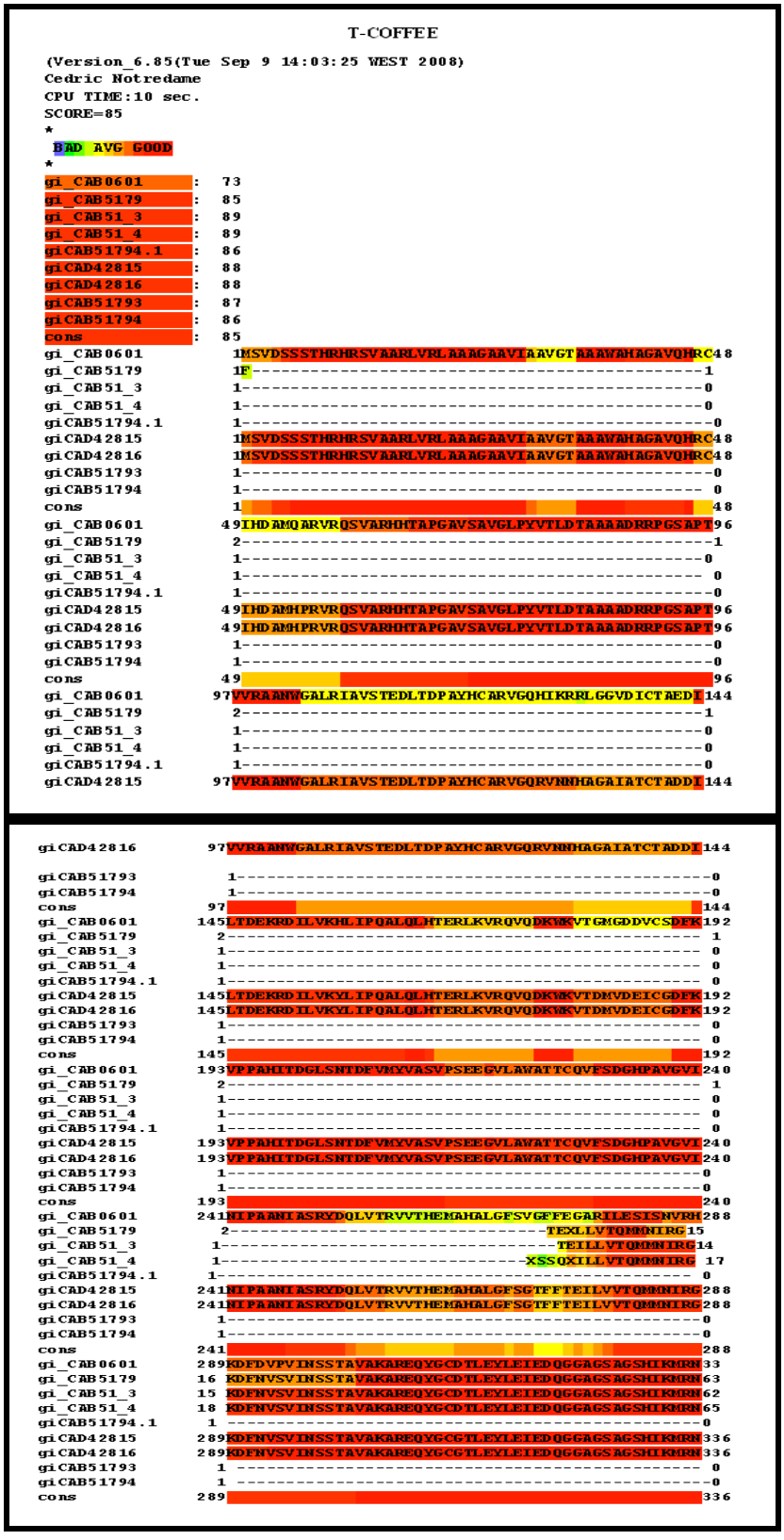


Figure 1: Cladogram and Phylogram analysis of gp63 sequences of different species strains of *Leishmania* associated with Visceral Leishmaniasis

Bioinformation Volume 5 *open access*

www.bioinformation.net Issue 8 **Hypothesis**



Bioinformation Volume 5 *open access*

www.bioinformation.net Issue 8 **Hypothesis**

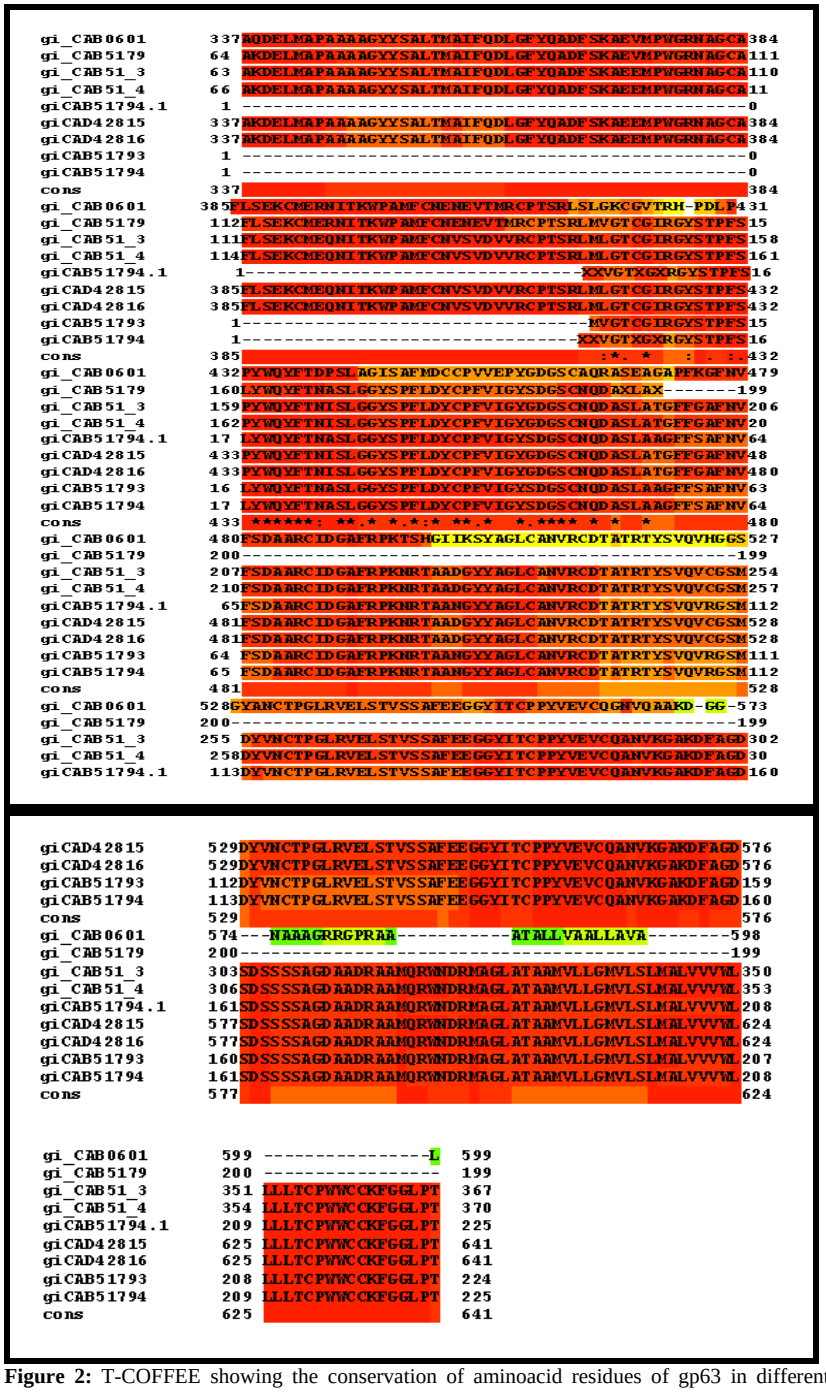


Figure 2: T-COFFEE showing the conservation of aminoacid residues of gp63 in different species strains of *Leishmania* associated with Visceral Leishmaniasis

Conclusion:

Previous findings support the major Leishmania surface glycoprotein (gp63) as a proteinase **[4]** with acidic pH optimum **[6]**. Specifically, evidences also show that it is a metalloenzyme, possibly with Zn in its active site and that it is present on the surface of both leishmanial stages with certain molecular changes during leismanial differentiation. The studies also suggest that its catalytic activity as a proteinase may protect Leishmania to parasitize macrophages successfully **[15]**. Immunization of susceptible BALB/c mice with gp63 entrapped in these vesicles protected virtually all of the mice against progressive nonhealing infections with L. donovoni **[17]**. Moreover data from other studies demonstrated that gp63 encapsulated in stable cationic liposomes provides long-term protection **[17]** in contrast to previous studies of protein-based vaccines **[16]**. Our results of BLAST Cladogram and Phylogenetic tree analysis throw light on the fact that a high level of conservation and identity amongst gp63 residues may help in the designing of a common vaccine against visceral

leishmaniasis caused by different species strains of *Leishmania*. T-COFFEE (Version_6.85) also showed that the level of similarity ranged from average to good in the strains we selected. The results reinstate our claim that a common gp63 based vaccine can be designed against all forms of visceral leishmaniasis.

Acknowledgement:
Sukrat Sinha is thankful to the financial support provided by the University Grants Commission, India under Dr. D.S Kothari Postdoctoral Fellowship.

References:

[1] HW Murray *et al. Lancet* **366**: 9496 (2005) [PMID: 16257344]
[2] EA Khalil *et al. Lancet* **356**: 9241 (2000) [PMID: 11075771]
[3] RA Seder & AV Hill. *Nature* **406**: 6797 (2000) [PMID: 10963610]
[4] KP Chang & D Fong. Ciba Found. Symp. 99 (1983) [PMID: 6357669]
[5] MS Klempner *et al. J. J. Infect. Dis.* **148**: 3 (1983) [PMID: 6619572]
[6] ME Wilson & RD Pearson. *J. Immunol.* **136**: 12 (1986) [PMID: 3711662]
[7] DG Russell & H Wilhelm. *J. Immunol.* **136**: 7 (1986) [PMID: 3950420]
[8] CS Chang & KP Chang. *Proc. Natl. Acad. Sci. U. S. A.* **83**: 1 (1986) [PMID: 3079902]
[9] D Fong & KP Chang. *Proc. Natl. Acad. Sci. U. S. A* **79**: 23 (1982) [PMID: 6961414]
[10] DG Russell & SD Wright. *J. Exp. Med.* **168**: 1 (1988) [PMID: 3294332]
[11] E Handman & JW Goding. *EMBO J.* **4**: 2 (1985) [PMID: 4018028]
[12] DG Russell & J Alexander. *J. Immunol.* **140**: 4 (1988) [PMID: 3257774]
[13] S Karlin & F Altschul. *Proc Natl Acad Sci U S A* **8**: 76 (1990) [PMID: 2315319]
[14] SF Altschul. *J Mol Biol* **219**: 3 (1991) [PMID: 2051488]
[15] G Chaudhuri *et al. JBC* **264**: 13 (1989) [PMID: 2708373]
[16] S Me´ndez *et al. J. Immunol.* **166**: 8 (2001) [PMID: 11290794]
[17] Swati Bhowmick *et al. InfecImmun* **76**: 3 (2008) [PMID: 18195029]

Edited by F Chiappelli
Citation: **Sinha *et al.*** Bioinformation 5(8): 320-325 (2011)

Bioinformation Volume 5 *open access*

www.bioinformation.net Issue 8 **Hypothesis**

Supplementary material:

Table 1: BLAST between gp63 of different species strains of *Leishmania* associated with Visceral Leishmaniasis

SPECIES/SPECIES	ACCESSION NO/ACCESSION NO	SCORE	EXPECT VALUE	IDENTITIES	POSITIVES	GAPS
L.INFANTUM/L.INFANTUM	CAB06018/CAB51797	326 BITS(835)	6e-94	155/198(78%)	166/198(83%)	1/198(0%)
L.INFANTUM/L.INFANTUM	CAB06018/ CAB51794	213 BITS(543)	5e-60	104/139(74%)	107/139(76%)	0/139(0%)
L.INFANTUM/L.INFANTUM	CAB06018/ CAB51793	216 BITS(549)	1e-60	107/154(69%)	113/154(73%)	1/154 (0%)
L.INFANTUM/L.INFANTUM	CAB51797/ CAB51794	108 BITS(270)	8e-29	50/53(94%)	50/53(94%)	0/53(0%)
L.INFANTUM/L.INFANTUM	CAB51797/ CAB51793	116 BITS(291)	2e-31	53/54(98%)	53/54(98%)	0/54(0%)
L.INFANTUM/L.INFANTUM	CAB51793/ CAB51794	450 BITS(1157)	1e-131	221/223(93%)	221/223(93%)	0/223(0%)
L.DONOVANI/L.DONOVANI	CAD42816/CAD42815	1330 BITS(3441)	0.0	641/641(100%)	641/641(100%)	0/641(0%)
L.DONOVANI/L.DONOVANI	CAD42816/CAB51784	762 BITS(1968)	0.0	363/365(99%)	364/365(99%)	0/365(0%)
L.DONOVANI/L.DONOVANI	CAD42816/CAB51783	766BITS(1978)	0.0	365/367(99%)	366/367(99%)	0/367(0%)
L.DONOVANI/L.DONOVANI	CAD42815/CAB51784	762 BITS(1968)	0.0	363/365(99%)	364/365(99%)	0/365(0%)
L.DONOVANI/L.DONOVANI	CAD42815/CAB51783	766 BITS(1978)	0.0	365/367(99%)	366/367(99%)	0/367(0%)
L.DONOVANI/L.DONOVANI	CAB51784/CAB51783	756 BITS(1952)	0.0	365/365(100%)	365/365(100%)	0/365(0%)
L.DONOVANI/ L.INFANTUM	CAD42816/CAB06018	917 BITS(2370)	0.0	482/572(84%)	502/572(87%)	1/572(0%)
L.DONOVANI/ L.INFANTUM	CAD42816/ CAB51797	386 BITS(992)	4e-112	182/198(91%)	187/198(94%)	0/198(0%)
L.DONOVANI/ L.INFANTUM	CAD42816/ CAB51794	390 BITS(1001)	4e-113	213/223(95%)	215/223(96%)	0/223(0%)
L.DONOVANI/ L.INFANTUM	CAD42816/ CAB51793	398 BITS(1022)	2e-115	216/224(96%)	218/224(97%)	0/224(0%)
L.DONOVANI/ L.INFANTUM	CAD42815/CAB06018	917 BITS(2370)	0.0	482/572(84%)	502/572(87%)	1/572(0%)
L.DONOVANI/ L.INFANTUM	CAD42815/ CAB51797	386 BITS(992)	4e-112	182/198(91%)	187/198(94%)	0/198(0%)
L.DONOVANI/ L.INFANTUM	CAD42815/ CAB51794	390BITS(1001)	4e-113	213/223(95%)	215/223(96%)	0/223(0%)
L.DONOVANI/ L.INFANTUM	CAD42815/ CAB51793	398 BITS(1022)	2e-115	216/224(96%)	218/224(82%)	0/224(0%)
L.DONOVANI/ L.INFANTUM	CAB51784/CAB06018	473BITS(1217)	7e-138	230/294(78%)	244/294(82%)	1/294(0%)
L.DONOVANI/ L.INFANTUM	CAB51784/ CAB51797	379BITS(974)	3e-110	181/194(93%)	185/194(95%)	0/194(0%)
L.DONOVANI/ L.INFANTUM	CAB51784/ CAB51794	384BITS(987)	1e-111	213/223(95%)	215/223(96%)	0/223(0%)
L.DONOVANI/ L.INFANTUM	CAB51784/ CAB51793	392BITS(1008)	4e-114	216/224(96%)	218/224(97%)	0/224(0%)
L.DONOVANI/ L.INFANTUM	CAB51783/CAB06018	473BITS(1217)	6e-138	230/294(78%)	244/294(82%)	1/294(0%)
L.DONOVANI/ L.INFANTUM	CAB51783/ CAB51797	382BITS(982)	3e-111	183/197(92%)	187/197(94%)	0/197(0%)
L.DONOVANI/ L.INFANTUM	CAB51783/ CAB51794	384BITS(986)	1e-111	213/223(95%)	215/223(96%)	0/223(0%)
L.DONOVANI/ L.INFANTUM	CAB51783/ CAB51793	392BITS(1007)	4e-114	216/224(96%)	218/224(97%)	0/224(0%)