



**Faculdade
Pernambucana
de Saúde**

**AVALIAÇÃO DA NEK COMO ALVOS FARMACOLÓGICOS
CONTRA LEISHMANIOSES POR MÉTODOS EM BIOINFORMÁTICA**

**Alunos: Crhisllane Rafaela dos Santos Vasconcelos
Geórgia Fernanda de Oliveira
Marconi Rego Barros Jr.**

Recife, 2014

AVALIAÇÃO DAS PROTEÍNAS NEK COMO ALVOS FARMACOLÓGICOS CONTRA LEISHMANIOSES POR MÉTODOS EM BIOINFORMÁTICA

Marconi Rego Barros Jr^{1,4}; Chrisllane Rafaela dos Santos Vasconcelos^{2,4}; Georgia
Fernanda de Oliveira^{3,4}; Rafael de Freitas e Silva^{2,4}

1. Laboratório de Estudos Moleculares e Terapia Experimental;
2. Centro de Pesquisas Aggeu Magalhães – FioCruz (CPQAM);
3. Genomika Diagnósticos;
4. Faculdade Pernambucana de Saúde

ABSTRACT

Leishmaniasis is a neglected disease caused by protozoa from Trypanosomatidae family and in more severe cases can lead an individual to death. It exhibits both deficiencies in its diagnostic property as in its treatment. NEK are an important group of proteins with essential cellular functions, becoming thus potential targets for the development of diagnostic methods, vaccines and more effective therapeutic methodologies. **Objective:** Evaluate diagnostic and therapeutic properties of NEK to leishmaniasis through methods in bioinformatics. **Methodology:** It has been used on-line tools such as BLAST, ORF finder, CD Search, TBLASTn, Wolf PSORT, Mholline, Modeller, Procheck, Chimera, Gene Ontology, SignalP4.1, in addition to MEGA and MrBayes. **Results:** The alignment with Homo sapiens proteins have not exhibited significant values; besides when compared to *Trypanosoma cruzi* it has been observed conserved domains in CD Search and Clustal omega. In addition, gene expression has been up-regulated in blood stage, which is relevant in the search to diagnostic and vaccine/drug targets. The sequence and structure alignment have demonstrated regions where

possible drugs can bind, which can be a promising point to performance of future researches to development of new molecules that can act against several species of trypanosomatids. **Conclusions:** The diagnostic and therapeutic properties have been strengthened in all analyzes done in this work.

Key words: Leishmaniasis, Trypanosomatid, NEK, bioinformatics.

RESUMO

Leishmaniose é uma doença negligenciada causada por protozoários da família Trypanosomatidae, e pode levar o indivíduo a morte nos casos mais severos. Apresenta tanto deficiências em sua propriedade diagnóstica como no tratamento. As NEK constituem um grupo importante de proteínas com funções celulares essenciais, constituindo-se assim, potenciais alvos para o desenvolvimento de métodos diagnósticos, vacinas e de metodologias terapêuticas mais eficazes. **Objetivo:** Avaliar as propriedades diagnóstica e terapêutica das NEK para leishmaniose por métodos em bioinformática. **Métodos:** Foram utilizadas ferramentas disponíveis on-line como o BLAST, ORF finder, CD Search, TBLASTn, Wolf PSORT, Mholline, Modeller, Procheck, Chimera, Gene Ontology, SignalP4.1, além do MEGA e MrBayes. **Resultados:** Os alinhamentos com NEK de Homo sapiens não apresentavam valores significativos; já comparando as NEK de Leishmanias com as de *Trypanosoma cruzi*, foi observado na pesquisa de domínios conservados e no alinhamento com a ferramenta Clustal ômega, áreas bastante conservadas. Além disso, a expressão gênica apresentou-se mais acentuada no estágio sanguíneo, o que é relevante para busca de alvos diagnósticos e vacinais/medicamentosos. Os alinhamentos tanto de sequência quanto de estrutura apresentaram regiões para possível atracamento de fármacos, abrindo assim espaço para pesquisa de fármacos com alvo em mais de uma espécie de tripanosomatídeos. **Conclusões:** As capacidades diagnóstica e terapêutica foram reforçadas em todas análises feitas.

Palavras-chave: Leishmania; Tripanosomatídeos; NEK; bioinformática.

Contato: rafa.freitass86@gmail.com

INTRODUÇÃO

Leishmaniose é uma doença causada por protozoários que pertencem à família Trypanosomatidae e ao gênero *Leishmania*. Esta parasitose se apresenta de três formas, a cutânea, a mucosa e a visceral (Murray et al., 2005), e seus casos mais graves pode levar o indivíduo a morte. É estimado que cerca de 12 milhões de pessoas estejam infectadas por alguma espécie deste gênero, com uma incidência anual de 1 a 2 milhões de pessoas, sendo os países em desenvolvimento os que apresentam a maior prevalência (Hotez et al., 2008). No Brasil, a prevalência também é alta. Além disso, o caro e/ou tóxico tratamento, o qual requer longos períodos de terapia supervisionada, a resistência apresentada pelos parasitas contra as principais drogas utilizadas, e a ausência de vacinas efetivas (Rezende et al., 2012) tornam esta parasitose um importante problema de saúde pública no país, sendo considerada uma doença negligenciada.

NEK (NIMA related kinases) é um conjunto de proteínas quinase, cuja estrutura está relacionada à NIMA. Atualmente é conhecida a existência de mais NEK em tripanosomatídeos (em torno de 20) do que em humanos (em torno de 15) e nestes últimos estas importantes enzimas estão envolvidas na regulação do ciclo celular. Nos parasitas as funções destas proteínas estão ainda obscuras, sendo que algumas delas devem ter funções únicas nos protozoários. Uma função conhecida é por desempenhar papel essencial na duplicação do corpo basal e por estar expressamente mais elevada no sangue do que na forma procíclica (Naula et al., 2005). Em *Trypanosoma brucei* estão envolvidas na diferenciação do parasita e na citocinese e em *Leishmania* sp. estão envolvidas na regulação de processos ciliares, como na montagem, desmontagem e regulação do tamanho destas estruturas (Cao et al., 2009). Acredita-se que possam ser alvos importantes e promissores para formulação de novas terapias, como também para

elaboração de métodos diagnósticos de detecção inicial mais acentuada e eficaz e com predição prognóstica mais precisa.

Atualmente ainda são utilizadas na terapêutica da Leishmaniose drogas desenvolvidas há décadas atrás e que provocam efeitos colaterais graves. As NEK constituem um grupo importante de proteínas com funções celulares essenciais não somente para a sobrevivência do parasita, como também para sua viabilidade e proliferação, constituindo assim potenciais alvos para o desenvolvimento de vacinas e de metodologias terapêuticas mais eficazes. Quanto a capacidade diagnóstica estas proteínas apresentam domínios que as diferenciam das proteínas humanas e cuja expressão indicam a expansão dos parasitas e progressão da infecção.

Por isso, objetivamos aqui avaliar estas proteínas por métodos de bioinformática quanto ao seu possível uso terapêutico e diagnóstico.

OBJETIVOS

Avaliar por métodos de bioinformática a possível capacidade terapêutica e diagnóstica de proteínas NEK.

Objetivos específicos

- Alinhar sequência protéica NEK de *L. braziliensis* contra o banco de sequência de *Leishmania*, *Homo sapiens* e geral;
- Predizer a estrutura 3D da query, dos melhores hits e compará-las;
- Buscar domínios conservados, das sequências protéicas alinhadas e compará-los;

- Prever a localização subcelular;
- Inferir a função por ontologia gênica;
- Construir a árvore filogenética;
- Prever a expressão gênica.

MÉTODOS

Figura 01.

Aquisição de sequência protéica

A sequência protéica foi de uma proteína quinase putativa de *Leishmania braziliensis* que foi coletada no banco de dados do NCBI e cujo gi é 134066129. Esta sequência foi a query utilizada como referência para realização do trabalho e foi o input do método mencionado a seguir.

Alinhamento de sequências protéicas

A sequência input foi alinhada contra o banco de dados total do NCBI, contra o banco de *Leishmania* e contra o banco de *Homo sapiens* pela ferramenta disponibilizada no próprio site, o BLAST.

Busca de ORF

A busca de ORF foi realizada pela ferramenta OrfFind do NCBI.

Predição da estrutura 3D

A predição da estrutura 3D das proteínas foi realizada pelo método de modelagem comparativa utilizando a plataforma online Mhonline, que por sua vez utiliza um conjunto de ferramentas desde de alinhamento até avaliação do modelo, sendo esses: Blast, como ferramenta de alinhamento; Modeller, como ferramenta pra predição da estrutura e Procheck, para avaliação da qualidade estereoquímica dos modelos (Cavalcanti, et al., 2003). Para visualização e alinhamento das estruturas foi utilizado o programa Chimera (Pettersen, et al. 2004).

Busca de domínios conservados

Os domínios protéicos foram encontrados por meio da ferramenta CD SEARCH disponível gratuitamente no NCBI.

Busca de localização subcelular protéica

A localização subcelular foi predita pela ferramenta on-line e gratuita Wolf PSORT. Foram utilizadas a sequência referência de *L. braziliensis* além das sequencias de *L. mexicana*, *L. infantum*, *L. major*, *T. cruzi* e *H. sapiens*.

Ontologia Gênica

A ontologia gênica foi feita pela ferramenta on-line Gene Ontology. A proteínas foram submetidaa e o alinhamento de melhor P value foi escolhido. Foi construído então o gráfico a partir de todas as associações relativas a este melhor hit.

Construção da Árvore Filogenética

A árvore filogenética foi construída pelos programas MEGA e MrBayes. As sequências utilizadas foram os melhores hits do blastp contra o banco não especificado do NCBI. Os melhores hits incluindo espécies diferentes além da própria query foram escolhidos com imput para formulação da árvore filogenética. Antes do software MrBayes ser processado as sequências protéicas foram alinhadas no programa MEGA. Também neste mesmo programa foi feita a árvore fenética. Os parâmetros utilizados foram: método de filogenia foi o Bootstrap e o modelo foi da distância-p. O arquivo gerado foi executado no MrBayes com os seguintes parâmetros: ngen=1000000 e samplefreq=10000. A visualização da árvore foi feita pelo programa FigTree.

Predição de Expressão Gênica

A query e os 10 melhores hits do blastp contra banco de Leishmania foram submetidos ao alinhamento contra banco de ESTs (Expressed sequence tags) do NCBI – tblastn – para predição da expressão gênica.

RESULTADOS

Alinhamento de sequências

Foi realizado o blastp para alinhamento da query contra o banco de dados geral do NCBI. Os melhores hits do blastp apresentaram um e-value de 0.0 com identidade entre 76-100% e cobertura entre 97-100% sendo as melhores pertencentes a outras espécies de Leishmania (*L. mexicana*, *L. infantum*, *L. major*) (figura 02A). Também foi

visto outros bons resultados de identidade com outros integrantes da família dos tripanosomatídeos, como os protozoários simbióticos de bactérias *Angomonas deanei* e *Strigomonas culicis*, além do *Phytomonas sp.* e *Trypanosoma cruzi* e *T. brucei*. Esses apenas confirmaram a similaridade que as proteínas NEK compartilham por entre os integrantes da família dos tripanosomatídeos.

A query também foi alinhada contra o banco de dados de *Homo sapiens* no NCBI. A sequência que apresentou o melhor e-value (apenas $2e^{-48}$) apresentou uma cobertura de 51% e somente 35% de identidade (figura 02B).

Figura 02^a e figura 02B.

Busca de ORF

A busca de ORF foi realizado pelo OrfFind cujo resultado encontra-se na figura a seguir (Figura 03).

Figura 03.

Busca de domínios conservados de proteína

Foi realizado a busca por domínios conservados dos melhores hits do blastp. Foram escolhidas as sequências de maior identidade pertencentes a espécies diferentes de importância clínica: *L. mexicana*, *L. infantum*, *L. major* e *T. cruzi*.

Foi observado que todas as sequências protéicas apresentaram o domínio Pkc like superfamily, indicando que todas são proteínas quinases serina/treonina no intervalo de aminoácidos entre a posição 35 e 301 para as espécies do gênero

Leishmania e 32 a 298 para o *T. cruzi*. Os e-value foram bastante significativos, sendo o maior de $8,01 \times 10^{-102}$ e o menor e mais significativo para o *T. cruzi* ($1,23 \times 10^{-107}$) (figuras 04A, 04B e 04C). Esse domínio catalítico também apresentou identidade com o domínio da quinase humana que catalisa a transferência de grupo gama-fosforil do ATP para resíduos de serina/treonina em substratos proteicos, esse grupo está presente tanto nas NEK humanas quanto nas dos parasitas pesquisados pois todas são proteínas quinases.

Foi encontrado ainda identidade das sequências utilizadas com um trecho multi-domínio (acesso PTZ00283) o qual não foi encontrado quando aplicado as duas sequências de maior identidade no blastp contra banco de *H. sapiens* (imagens 2A). Sugerindo que proteínas quinases que apresentaram identidade de sequência acima de 60% do núcleo do domínio catalítico tem uma maior probabilidade de serem inibidos pelo mesmo grupo de compostos com baixa massa molecular. No caso em estudo, as sequências protéicas apresentaram identidade de 100% do domínio catalítico, o que se pode sugerir este sítio como alvo de drogas inibitórias.

Figuras 04A, B e C.

Análise de localização subcelular protéica

Por meio da ferramenta on-line Wolf PSORT foram preditas a localização subcelular dos melhores hits sequências do alinhamento do blastp. Foi observado que as sequências de *L. braziliensis* e *T. cruzi* apresentaram semelhantes, alinhando-se com proteínas do banco com localização citoplasmática e nuclear. Para *L. mexicana*, *L. infantum* e *L. major* além da localização citoplasmática e nuclear, surgiu também a mitocôndria como outro possível local onde estas proteínas poderiam estar presentes. Já

para as sequências de *H. sapiens* uma apresentou localização citoplasmática e nuclear enquanto a segunda apresentou também uma possível localização extranuclear (figura 05).

Figura 05.

Ontologia Gênica

A query foi submetida na ferramenta *Gene Ontology*. O alinhamento de melhor P value foi com a NEK1. Foi possível a predição de várias funções à sequência submetida o qual pode ser visto na imagem 4. Pode-se observar função de montagem dos cílios, ou seja, sua morfogênese e organização, o que condiz com dados observados na literatura. Além disso, foi possível inferir também participação na mitose, material pericentriolar, atividade de serina/treonina quinase (como observado no *CD Search*) e de ligação de ATP. (Figura 06A, B)

Figura 06A, B.

Predição da estrutura 3D

Dos 10 hits selecionados a partir do blastp contra o banco de leishmania, a modelagem teve valores estereoquímicos aceitáveis para 5 deles, tendo uma mediana de ângulos em regiões não aceitáveis de apenas 0.7%. O alinhamento entre as estruturas mostrou regiões de sobreposição entre os melhores modelos das 5 proteínas (figura 07A, B), considerando que a estrutura das proteínas é mais conservada que a sequência, assim o alinhamento de estruturas confere maior especificidade ao alinhamento de sequências (Verli, 2014); As regiões alinhadas sendo elas tanto alfa-

hélices quanto loops, podem ser avaliadas e testadas como alvos, visando o atracamento de fármacos em várias proteínas com funções importantes do parasito.

Figura 07A, B.

Árvore Fenética e Filogenética

Através da construção fenética e filogenética (figuras 08A, B) pôde-se observar, que nas NEK avaliadas de *Leishmania*, os eventos de duplicação ocorreram de maneiras independentes e não recentes. Uma avaliação com as NEK de *Homo sapiens* podem ressaltar a distância entre elas reforçando-as como possíveis alvos terapêuticos.

Figura 08A, B

Busca de peptídeo-sinal

O software SignalP 4.1 não conseguiu prever nenhum peptídeo-sinal para as sequências submetidas (*L. braziliensis*; *L. mexicana*, *L. infantum*, *L. major*, *T. cruzi* e *H. sapiens*). Abaixo o resultado para a sequência referência *L. braziliensis* (figura 09).

Figura 09.

Predição da Expressão Gênica

Foi realizado o alinhamento contra banco de ESTs da query e dos 10 melhores hits do alinhamento contra banco de *Leishmania*. Os resultados foram expressos na figura a seguir (figura 10). A grande maioria das sequências alinhadas apresentaram expressão gênica mais intensa na fase sanguínea e epimastigota.

O alinhamento com o clustal Omega (material suplementar 1) mostrou muitas regiões altamente conservadas entre as NEK de *Leishmania* e *Trypanosoma cruzi*.

A soma desses resultados indica que a formulação de uma possível droga capaz de interagir com as NEKs de modo geral poderia não somente atuar no tratamento da Leishmaniose como também de outras doenças parasitárias como a Doença de Chagas, apresentando uma gama maior de proteínas e, conseqüentemente, parasitas alvos, como foi visto através das análises geradas. Isso é preferível já que drogas inibitórias altamente específicas tem provocado resistência parasitária, tornando estas substâncias ineficientes (Shah et al., 2004). Além disso, tal droga teria poucas chances de interagir com NEK humanas, pois, como foi demonstrado, as proteínas humanas apresentaram significativa pouca similaridade com as proteínas parasitárias.

CONCLUSÃO

A partir das análises realizadas pode-se concluir que as NEK apresentam potencial para uso no desenvolvimento de métodos diagnóstico e terapêutico. No entanto, as análises realizadas aqui pertencem a um estudo inicial, servindo de base e de estímulo para futuros trabalhos que testem estas proteínas constatando-se tais propriedades.

CONFLITOS DE INTERESSE

Os autores declaram não existir conflito de interesses.

AGRADECIMENTOS

Agradecemos a contribuição da Prof^a Dr^a Ana Benko Isepon pelo apoio no desenvolvimento deste trabalho e ao Prof^o Rafael de Freitas e Silva pela orientação.

REFERÊNCIAS

Murray HW, Berman JD, Davies CR, Saravia NG. Advances in leishmaniasis. *Lancet*. 2005; 366: 17.

Hotez PJ, Bottazzi ME, Franco-Paredes C, Ault SK, Periago MR .The Neglected Tropical Diseases of Latin America and the Caribbean: A Review of Disease Burden and Distribution and a Roadmap for Control and Elimination. *Plos Neglected Tropical Diseases*. 2008; 2: e300.

Rezende AM, Folador EL, Resende DdM, Ruiz JC. Computational Prediction of Protein-Protein Interactions in Leishmania Predicted Proteomes. *PLoS ONE*.2012; 7(12): e51304.

Naula C, Parsons M, Mottram J.C. Protein kinases as drug targets in trypanosomes and Leishmania. *Biochim Biophys Acta*. 2005; 1754(1-2): 151-159.

Cao M, Li G, Pan J. Regulation of cilia assembly, disassembly, and length by protein phosphorylation. *Methods Cell Biol*. 2009; 94:333-46.

Cavalcanti MC, Baia OF, Rossle S, Bisch P, Targino R, Pires PF, Campos MLM, Mattoso, MLQ. Structural Genomic Workflows Supported by Web Services. In: Proc DEXA 2003 workshops, BIDM'03, IEEE CS Press, 2003; 1; 45-49.

Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem.* 2004; Oct;25(13): 1605-12.

Verli H. *Bioinformática da Biologia à flexibilidade molecular.* Porto Alegre. 2014.

Shah NP, Tran C, Lee FY, Chen P, Norris D, Sawyers CL. Overriding imatinib resistance with a novel ABL kinase inhibitor. *Science.* 2004; 305:399–401.

PÁGINA DE ILUSTRAÇÕES

Figura 01.

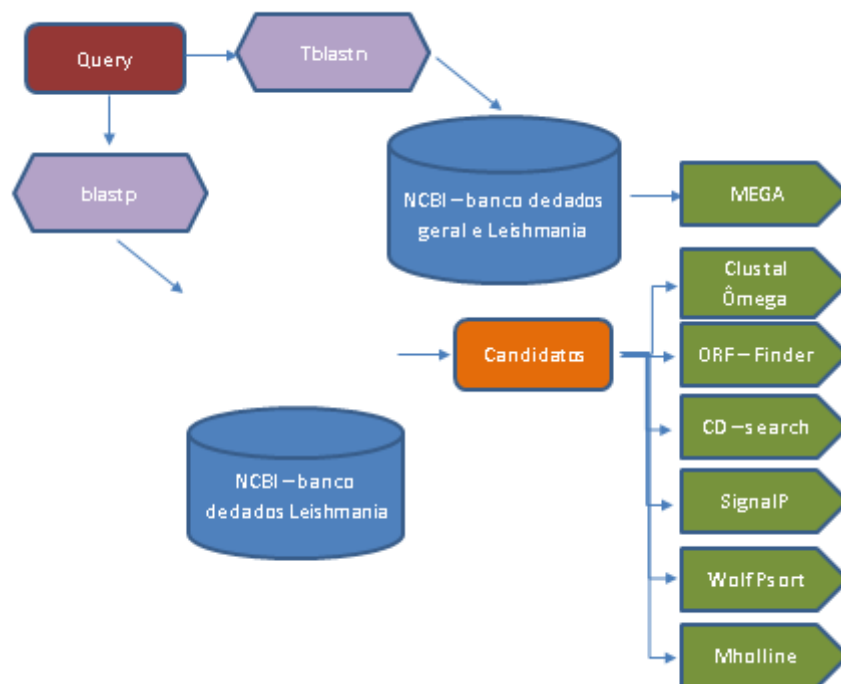


Figura 02A.

Sequences producing significant alignments:

Select: All None Selected: 20

Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input checked="" type="checkbox"/> putative nima-related protein kinase [Leishmania braziliensis MHOM/BR/75M/2904]	1061	1061	100%	0.0	100%	XP_001588787.1
<input checked="" type="checkbox"/> putative hypothetical protein, conserved nima-related protein kinase [Leishmania mexicana MHOM/GT/2001AJ1103]	1033	1033	100%	0.0	97%	XP_003874496.1
<input checked="" type="checkbox"/> conserved hypothetical protein [Leishmania infantum JPCM5]	1032	1032	100%	0.0	97%	XP_001469740.1
<input checked="" type="checkbox"/> putative nima-related protein kinase [Leishmania major strain Friedlin]	1029	1029	100%	0.0	97%	XP_001686724.1
<input checked="" type="checkbox"/> protein kinase [Angomonas deanei]	826	826	97%	0.0	79%	EPY34554.1
<input checked="" type="checkbox"/> protein kinase [Strigomonas culicis]	820	820	97%	0.0	78%	EPY26184.1
<input checked="" type="checkbox"/> protein kinase [Strigomonas culicis]	809	809	96%	0.0	78%	EPY35010.1
<input checked="" type="checkbox"/> protein kinase [Strigomonas culicis]	807	807	97%	0.0	77%	EPY35781.1
<input checked="" type="checkbox"/> protein kinase [Angomonas deanei]	803	803	96%	0.0	78%	EPY33973.1
<input checked="" type="checkbox"/> unnamed protein product [Phytomonas sp. isolate EM1]	794	794	98%	0.0	76%	CCW60021.1
<input checked="" type="checkbox"/> serine/threonine protein kinase [Trypanosoma cruzi strain CL Brener]	765	765	97%	0.0	74%	XP_811722.1
<input checked="" type="checkbox"/> serine/threonine protein kinase [Trypanosoma cruzi Dm28c]	764	764	97%	0.0	74%	ESS68047.1
<input checked="" type="checkbox"/> serine/threonine protein kinase, putative [Trypanosoma cruzi]	764	764	97%	0.0	74%	FKG04005.1
<input checked="" type="checkbox"/> serine/threonine protein kinase [Trypanosoma cruzi strain CL Brener]	763	763	97%	0.0	74%	XP_816171.1
<input checked="" type="checkbox"/> serine/threonine protein kinase [Trypanosoma cruzi strain CL Brener]	761	761	97%	0.0	74%	XP_816170.1

Figura 02B.

Sequences producing significant alignments:

Select: [All](#) [None](#) Selected: 15

Alignments [Download](#) [GenPept](#) [Graphics](#) [Distance tree of results](#) [Multiple alignment](#)

Description	Max score	Total score	Query cover	E value	Ident	Accession
<input checked="" type="checkbox"/> NEK1 protein [Homo sapiens]	177	177	51%	2e-48	35%	AAH15147.1
<input checked="" type="checkbox"/> Chain A, Crystal Structure Of Human Nima-Related Kinase 1 (Nek1) [Homo sapiens]	174	174	51%	7e-48	34%	AAPC_A
<input checked="" type="checkbox"/> NIMA (never in mitosis gene a)-related kinase 3, isoform CPA_d [Homo sapiens]	172	172	51%	9e-48	37%	EAX08908.1
<input checked="" type="checkbox"/> NIMA (never in mitosis gene a)-related kinase 1, isoform CPA_c [Homo sapiens]	175	175	51%	6e-47	35%	EAX04789.1
<input checked="" type="checkbox"/> serine/threonine-protein kinase Nek3 isoform b [Homo sapiens]	172	172	52%	2e-46	36%	NP_001139571.1
<input checked="" type="checkbox"/> NEK4 protein [Homo sapiens]	172	172	54%	3e-46	30%	AAH15515.1
<input checked="" type="checkbox"/> NIMA (never in mitosis gene a)-related kinase 3, isoform CPA_c [Homo sapiens]	172	172	52%	3e-46	36%	EAX08907.1
<input checked="" type="checkbox"/> serine/threonine-protein kinase Nek3 isoform a [Homo sapiens]	172	172	52%	3e-46	36%	NP_002489.1
<input checked="" type="checkbox"/> serine/threonine-protein kinase Nek1 isoform 4 [Homo sapiens]	177	177	51%	3e-46	35%	NP_001186328.1
<input checked="" type="checkbox"/> serine/threonine-protein kinase Nek1 isoform 3 [Homo sapiens]	177	177	51%	4e-46	35%	NP_001186327.1
<input checked="" type="checkbox"/> serine/threonine-protein kinase Nek1 isoform 5 [Homo sapiens]	177	177	51%	4e-46	35%	NP_001186329.1
<input type="checkbox"/> unnamed protein product [Homo sapiens]	172	172	52%	4e-46	36%	BAD18511.1

Figura 03.

NCBI ORF Finder (Open Reading Frame Finder)

PubMed Entrez BLAST OMIM Taxonomy Structure

gi|134066129|emb|CAM43918.1| putative nima-related protein kinase [Leishmania braziliensis MHOM/BR/75/M2904]

View 1 GenBank Redraw 100 SixFrames

Frame	from	to	Length
-1	3..343	342	
-3	1..342	342	
+3	36..191	156	

Figura 04.

NCBI

Conserved Domains

HOME SEARCH GUIDE NewSearch Structure Home 3D Macromolecular Structures Conserved Domains Pubchem BioSystems

Conserved domains on [gi116284336]gb|AAH15147.1] View Concise Results

NEK1 protein, partial [Homo sapiens]

Graphical summary Zoom to residue level show extra options

Query seq. 1 75 150 225 300 375 450

RTP binding site
active site
10e substrate binding site
activation loop (H-loop)

Specific hits
Superfamilies PKC_like superfamily
Multi-domains S_TKc

List of domain hits

#	Name	Accession	Description	Interval	E-value
1	STKc_Nek1	cd08218	Catalytic domain of the Protein Serine/Threonine Kinase, Never In Mitosis gene A (NIMA) ...	3-258	0e+00
2	S_TKc	smart00220	Serine/Threonine protein kinases, catalytic domain; Phosphotransferases. Serine or ...	4-258	4.39e-96

References:

- Marchler-Bauer A et al. (2011), "CDD: a Conserved Domain Database for the functional annotation of proteins.", *Nucleic Acids Res.*39(D)225-9.
- Marchler-Bauer A et al. (2009), "CDD: specific functional annotation with the Conserved Domain Database.", *Nucleic Acids Res.*37(D)205-10.
- Marchler-Bauer A, Bryant SH (2004), "CD-Search: protein domain annotations on the fly.", *Nucleic Acids Res.*32(W)327-331.

Help | Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

NCBI

Conserved Domains

HOME SEARCH GUIDE NewSearch Structure Home 3D Macromolecular Structures Conserved Domains Pubchem BioSystems

Conserved domains on [gi119629313]gb|EAX08908.1] NIMA (never in mitosis gene a)-related kinase 3, isoform CRA_d [Homo sapiens] >[gi119629314]gb|EAX08909.1] View Concise Results

NIMA (never in mitosis gene a)-related kinase 3, isoform CRA_d [Homo sapiens]

Graphical summary Zoom to residue level show extra options

Query seq. 1 50 100 150 200 250 299

RTP binding site
active site
10e substrate binding site
activation loop (H-loop)

Superfamilies PKC_like superfamily
Multi-domains S_TKc

List of domain hits

#	Name	Accession	Description	Interval	E-value
1	STKc_Nek3	cd08219	Catalytic domain of the Protein Serine/Threonine Kinase, Never In Mitosis gene A (NIMA) ...	3-257	0e+00
2	S_TKc	smart00220	Serine/Threonine protein kinases, catalytic domain; Phosphotransferases. Serine or ...	4-252	2.25e-97

References:

- Marchler-Bauer A et al. (2011), "CDD: a Conserved Domain Database for the functional annotation of proteins.", *Nucleic Acids Res.*39(D)225-9.
- Marchler-Bauer A et al. (2009), "CDD: specific functional annotation with the Conserved Domain Database.", *Nucleic Acids Res.*37(D)205-10.
- Marchler-Bauer A, Bryant SH (2004), "CD-Search: protein domain annotations on the fly.", *Nucleic Acids Res.*32(W)327-331.

Help | Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

NCBI

Conserved Domains

HOME SEARCH GUIDE NewSearch Structure Home 3D Macromolecular Structures Conserved Domains Pubchem BioSystems

Conserved domains on [gi407849175]gb|EK.G04005.1] View Concise Results

serine/threonine protein kinase, putative [Trypanosoma cruzi]

Graphical summary Zoom to residue level show extra options

Query seq. 1 75 150 225 300 375 450 495

RTP binding site
active site
10e substrate binding site
activation loop (H-loop)

Superfamilies PKC_like superfamily
Multi-domains PTZ00283

List of domain hits

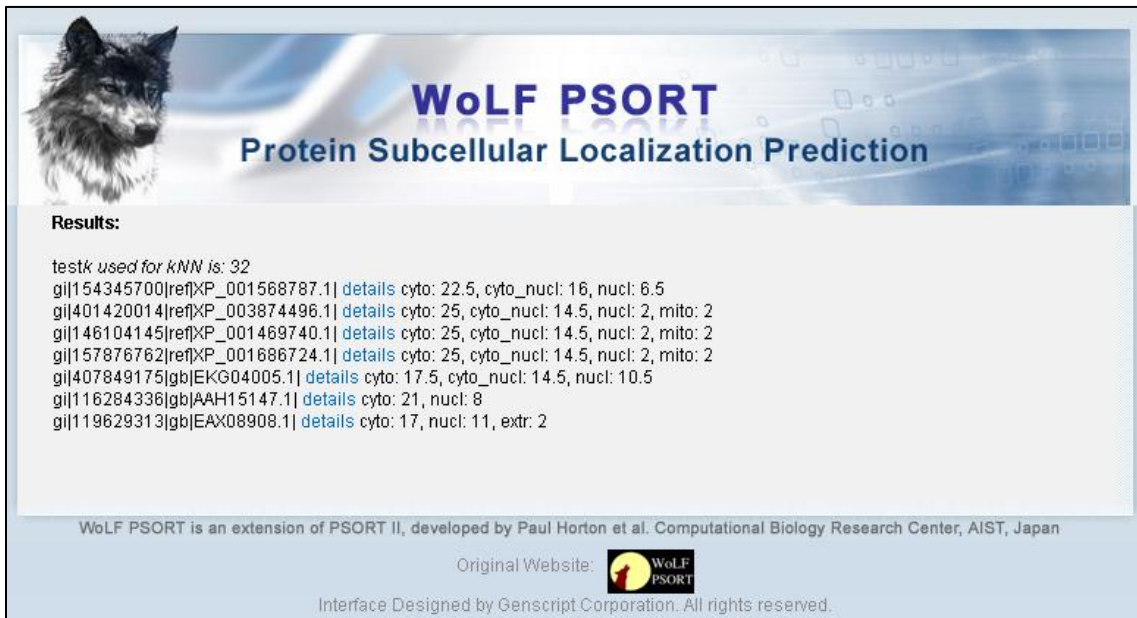
#	Name	Accession	Description	Interval	E-value
1	STKc_Nek	cd08215	Catalytic domain of the Serine/Threonine Kinase, Never In Mitosis gene A (NIMA)-related kinase; ...	32-298	1.23e-107
2	PTZ00283	PTZ00283	serine/threonine protein kinase, Provisional	1-495	0e+00

References:

- Marchler-Bauer A et al. (2011), "CDD: a Conserved Domain Database for the functional annotation of proteins.", *Nucleic Acids Res.*39(D)225-9.
- Marchler-Bauer A et al. (2009), "CDD: specific functional annotation with the Conserved Domain Database.", *Nucleic Acids Res.*37(D)205-10.
- Marchler-Bauer A, Bryant SH (2004), "CD-Search: protein domain annotations on the fly.", *Nucleic Acids Res.*32(W)327-331.

Help | Disclaimer | Write to the Help Desk
NCBI | NLM | NIH

Figura 05.



WoLF PSORT
Protein Subcellular Localization Prediction

Results:

testk used for kNN is: 32

gi|154345700|ref|XP_001568787.1| [details](#) cyto: 22.5, cyto_nucl: 16, nucl: 6.5

gi|401420014|ref|XP_003874496.1| [details](#) cyto: 25, cyto_nucl: 14.5, nucl: 2, mito: 2

gi|146104145|ref|XP_001469740.1| [details](#) cyto: 25, cyto_nucl: 14.5, nucl: 2, mito: 2


gi|157876762|ref|XP_001686724.1| [details](#) cyto: 25, cyto_nucl: 14.5, nucl: 2, mito: 2

gi|407849175|gb|EKG04005.1| [details](#) cyto: 17.5, cyto_nucl: 14.5, nucl: 10.5

gi|116284336|gb|AAH15147.1| [details](#) cyto: 21, nucl: 8

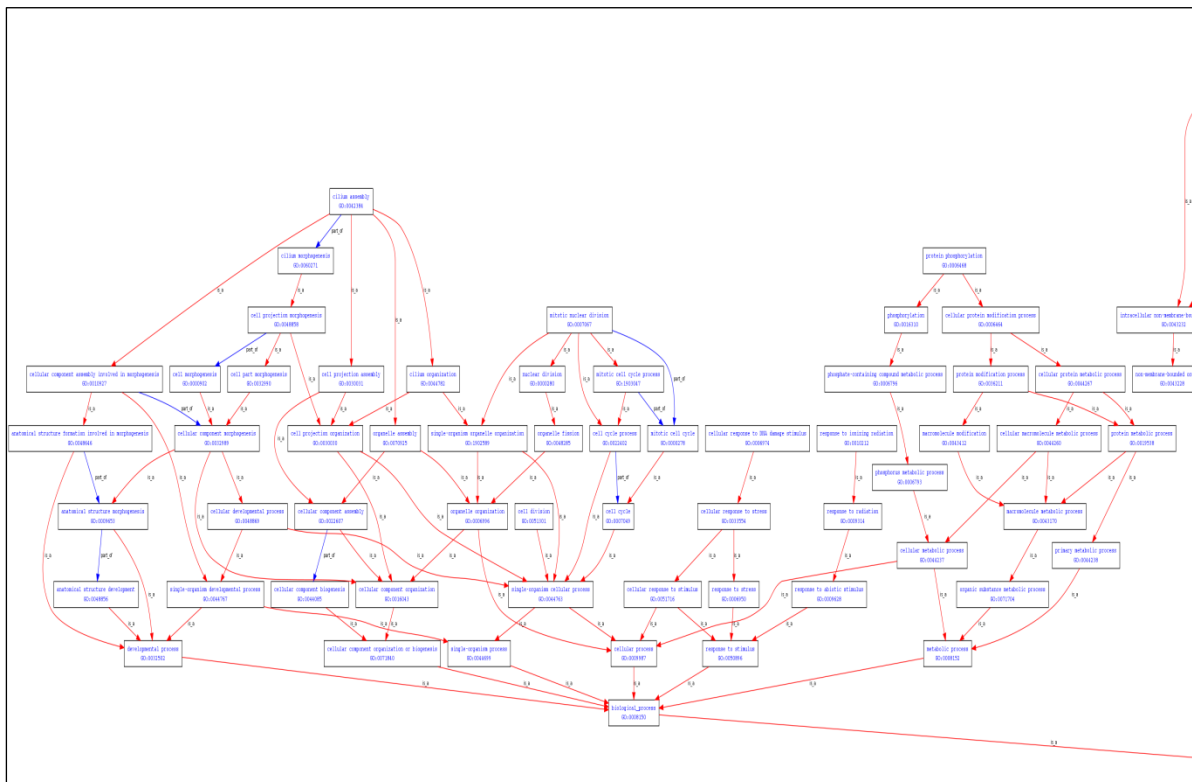
gi|119629313|gb|EAX08908.1| [details](#) cyto: 17, nucl: 11, extr: 2

WoLF PSORT is an extension of PSORT II, developed by Paul Horton et al. Computational Biology Research Center, AIST, Japan

Original Website: 

Interface Designed by Genscript Corporation. All rights reserved.

Figura 06A, B.



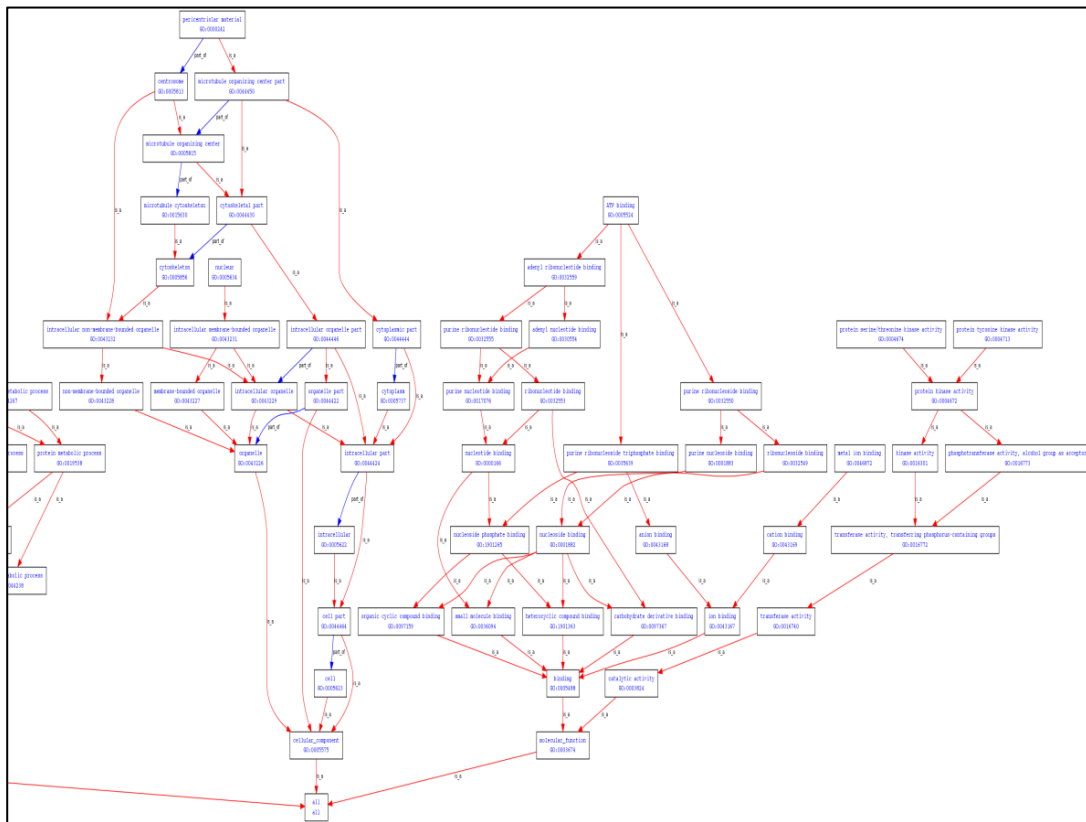


Figura 07A, B.

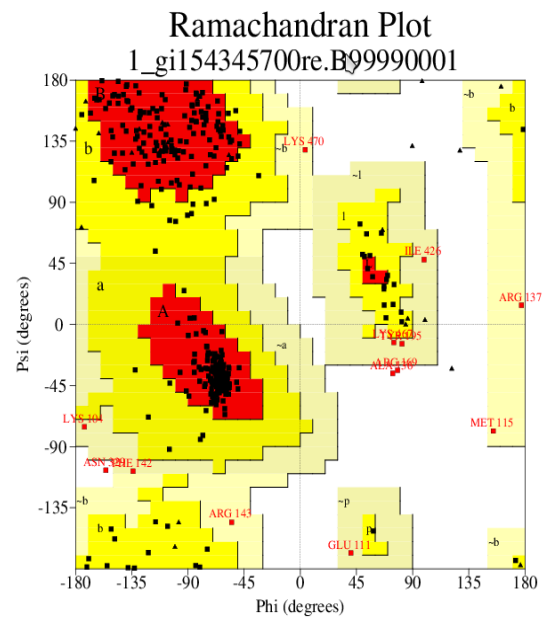
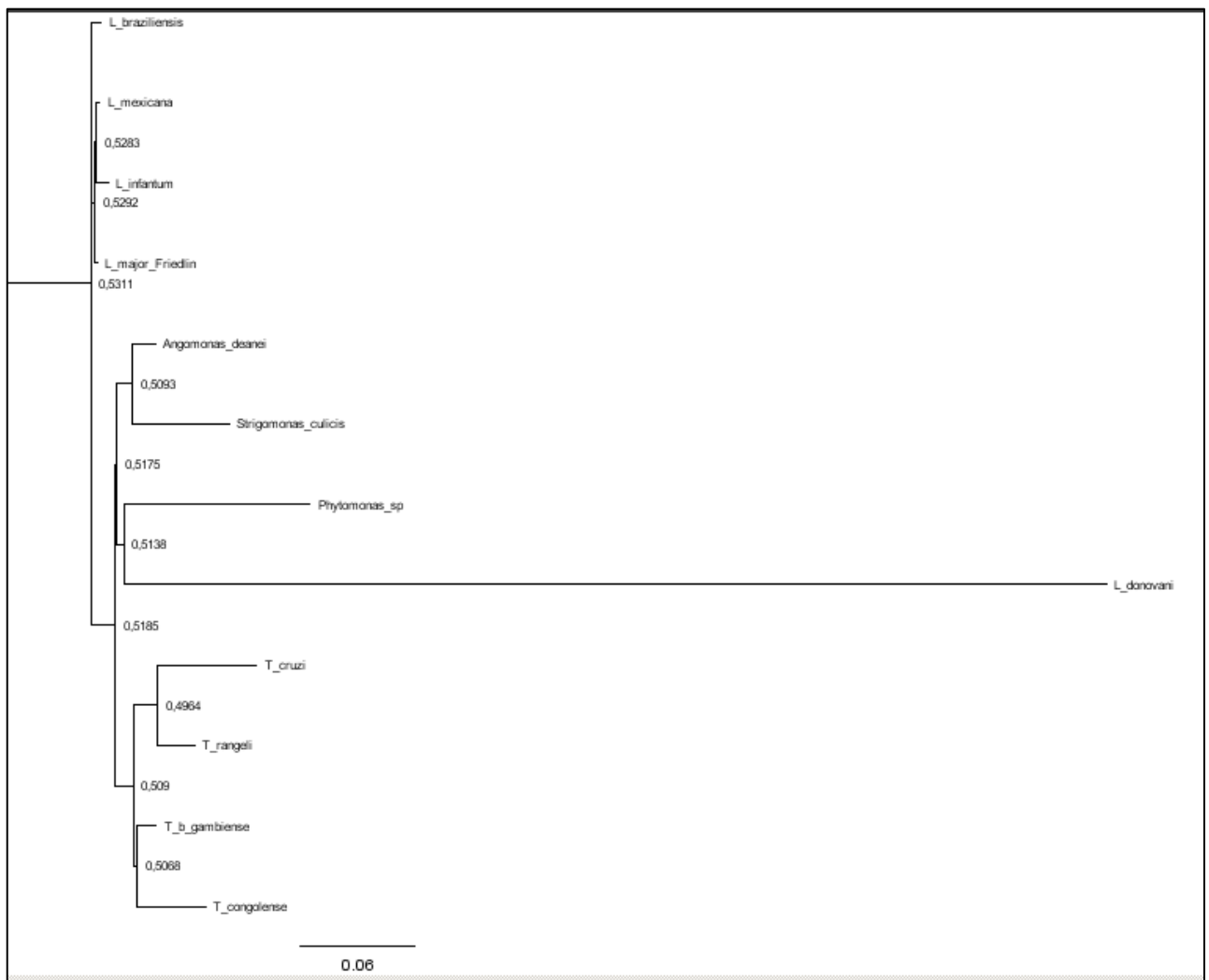
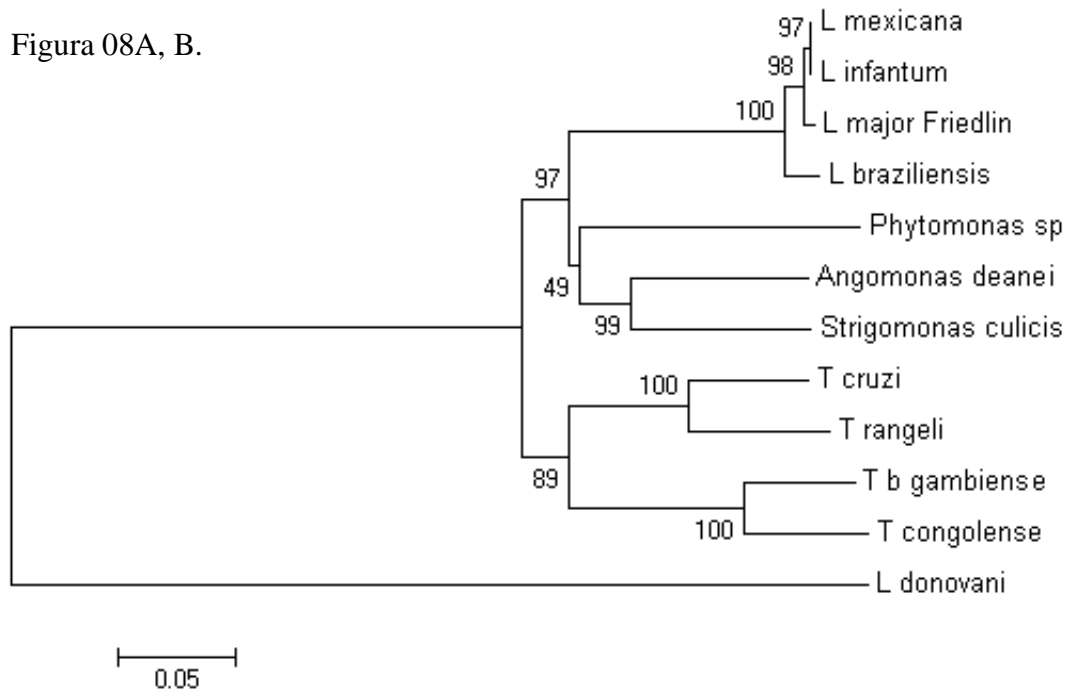


Figura 08A, B.



PÁGINA DE LEGENDAS

Figura 01. Esquema didático da metodologia.

Figura 02A. A foto dos hits do Blastp sem especificar o banco;

Figura 02B. A foto dos hits do Blastp contra o banco de Homo sapiens.

Figura 03. Resultado do ORF Finder.

Figura 04. Pesquisa de domínios conservados no cd-search disponível no site do NCBI. A. apresentado o domínio PKc_like superfamily, e o PTZ00283; B. Apresentando além do domínio PKc_like superfamily, os domínios STKc_Nek1 e S_TKc; C. apresentando além do PKc_like superfamily, o domínio S_TKc.

Figura 05. Resultados da predição de localização subcelular pela ferramenta on-line Wolf Psort.

Figura 06A, B. Ontologia gênica predita pela ferramenta *Gene Ontology*. A. Primeira parte. B. Continuação.

Figura 07. A. alinhamento entre as estruturas 3D preditas pelo servidor Mholline das proteínas: gi154345700 (roxo), gi154345702 (verde), gi154345339 (azul), gi154342748 (marrom) e gi154337543 (vermelho). B. Gráfico de Ramachandram da modelagem que ficou como mediana, apresentando só 0.7% dos ângulos em regiões não aceitáveis.

Figura 08A, B. Arvore fenética e filogenética demonstrando padrão evolutivo das proteínas NEK.

Figura 09. Busca por peptídeo sinal, exemplo que se seguiu em todas as proteínas.

Figura 10. Verde: Positivo; Amarelo: Zero; LCP: Fase exponencial promastigotas de *L. chagasi*; MTS: Estágio Metacíclico; PS: Estágio Procíclico; ES: Estágio epimastigota; MG: *Mimulus guttatus*; BS: Estágio sanguíneo.