

UNIVERSIDADE FEDERAL DO ABC CCNH
CENTRO DE CIÊNCIAS NATURAIS E HUMANAS

**A evolução da Ecotina em Kinetoplastida e a hipótese de
transferência genética horizontal**

Max Mario Fuhlendorf

Santo André - SP
2021

A evolução da Ecotina em Kinetoplastida e a hipótese de transferência genética horizontal

Monografia de conclusão de curso apresentada ao
Curso de Bacharelado em Ciências Biológicas da
Universidade Federal do ABC, como requisito
parcial à conclusão do curso.

Orientadora: Prof.^a Dr.^a Márcia Aparecida Sperança

Santo André - SP
2021

SUMÁRIO

SUMÁRIO	2
RESUMO	3
ABSTRACT	4
1. INTRODUÇÃO	5
2. METODOLOGIA	9
3. RESULTADOS E DISCUSSÃO	11
4. CONSIDERAÇÕES FINAIS	20
REFERÊNCIAS BIBLIOGRÁFICAS	21

RESUMO

A Doença de Chagas, apesar do sucesso de políticas públicas de controle vetorial, ainda tem incidência anual preocupante no Brasil, especialmente nas precárias e crescentes fronteiras da Amazônia legal. Seu agente etiológico, o parasita *Trypanosoma cruzi*, possui gene que codifica uma proteína inibidora de serinopeptidase (ISP), ortóloga da inibidora ecotina encontrada em *E. coli* e em uma gama de outras bactérias gram-negativas. Nestas bactérias, especialmente as que invadem tecidos de artrópodes e vertebrados, uma variedade crescente de estudos indica que a ecotina é um fator fundamental na defesa bacteriana contra as barreiras imunológicas dos organismos infectados. Evidências recentes indicam que as ISPs sintetizadas por *Leishmania major*, que são também ortólogas de ecotina, tem função exógena, facilitando a sobrevivência do parasita em face do sistema imunológico hospedeiro. Traçando um paralelo entre estudos com bactérias portadoras de ecotina e as ISPs de *L. major* — parasita que pertence a clade irmão de *T. cruzi* — fica claro o papel fundamental que estas proteínas inibidoras de serinopeptidases desempenham na modulação da resposta imunológica dos hospedeiros destes parasitas. A seguir, comparamos a posição da ecotina em vários membros do grupo Kinetoplastida para reforçar a hipótese de sua possível origem comum a partir de transferência horizontal.

ABSTRACT

Chagas disease, despite the success of public policies for vector control, still has a worrying annual incidence in Brazil, especially in the precarious and growing frontiers of the legal Amazon. Its etiologic agent, the *Trypanosoma cruzi* parasite, has a gene that encodes a serinepeptidase inhibitor protein (ISP), an ortholog of the ecotin inhibitor found in *Escherichia coli* and a range of other gram-negative bacteria. In these bacteria, especially those that invade arthropod and vertebrate tissue, a growing variety of studies indicates that ecotin is a key factor in bacterial defense against the immunological barriers of infected organisms. Recent evidence indicates that ISPs synthesized by *Leishmania major*, which are also ecotin orthologs, have an exogenous function, facilitating the parasite's survival against the host immune system. Drawing a parallel between studies with ecotin-bearing bacteria and ISPs of *L. major* — a parasite that belongs to a sister group of *T. cruzi* — it is clear the fundamental role that these serinepeptidase inhibitor proteins play in modulating the immune response of these parasite hosts. Next, we compared the position of ecotin in various members of the Kinetoplastida group to reinforce the hypothesis of its possible common origin from horizontal transfer.

1. INTRODUÇÃO

A Doença de Chagas, apesar de todos os avanços médicos e das políticas públicas de controle epidemiológico, afeta ainda um número significativo de brasileiros anualmente. O agente etiológico da Doença de Chagas é o parasita *Trypanosoma cruzi*, transmitido através de picadas de insetos da subfamília Triatominae, através de transfusões sanguíneas ou oralmente pela ingestão de alimentos contaminados com dejetos de insetos vetores. A incidência da doença pela transmissão clássica por triatomíneos coabitantes de residências humanas foi fortemente reduzida graças à efetividade dos programas públicos de controle vetorial e melhoria habitacional. Entretanto, há um crescente número de casos nos estados da Amazônia legal, onde existem triatomíneos silvestres que, por não invadirem habitações humanas, não podem ser facilmente controlados por medidas de contenção tradicionais. O resultado frequente é a contaminação por via oral através de ingestão de produtos contaminados por vetores infectados com *T. cruzi*, como a polpa de açaí (PASSOS et al., 2012), além da expansão da transmissão vetorial por triatomíneos silvestres, cuja incidência só deve crescer conforme prossegue o desmatamento da região (BRUM-SOARES et al., 2010): só entre 2000 e 2011, foram mais de 1200 casos agudos confirmados na Amazônia legal (ARAUJO-JORGE, 2013). A seguir, examinamos a importância da ecotina, enzima inibidora de serinopeptidases expressada em *T. cruzi*, e a relação filogenética dos vários genes ortólogos de ecotina dentro do clado que contém o parasita.

O *T. cruzi* pertence à ordem Trypanosomatida (CHUNG et al., 1983, p. 2), que também abriga o gênero *Leishmania*, outro parasita humano de ampla distribuição que causa doenças de difícil tratamento. Análises genéticas de membros desta ordem revelaram a presença de genes que codificam a produção de inibidores de serinopeptidase (ISPs), nos gêneros *Leishmania* e *Trypanosoma*. Os ISPs observados nestes tripanossomatídeos parecem ser ortólogos do gene que codifica a ecotina, um inibidor encontrado apenas em bactérias gram-negativas, sugerindo uma possível transferência genética lateral em algum ponto da evolução destes táxons. Este fato, bastante interessante por si só, torna-se mais intrigante quando consideramos que o sequenciamento não revelou a presença, no genoma dos parasitas estudados, de serinopeptidases dos tipos que a ecotina inibe, dado que indica que ISPs de tripanossomatídeos tem alvos exógenos (IVENS et al., 2005).

Dada a semelhança estrutural e a provável herança compartilhada da ISP de *T. cruzi* com a ecotina de bactérias, é válida uma contextualização mais aprofundada da ecotina. A ecotina é um inibidor de serinopeptidase isolada originalmente no periplasma em *E. coli*, nomeada pela sua capacidade de inibir a enzima digestiva tripsina (do inglês, **E. coli trypsin inhibitor**) (CHUNG et al., 1983). Hoje é classificada junto com inibidores homólogos — incluindo os ISPs de *Leishmania* e *Trypanosoma* — na família I11 da classificação MEROPS (MEROPS Family I11, 2015). Ortólogos de ecotina foram encontrados em vários gêneros de bactérias gram-negativas, incluindo *Escherichia*, *Yersinia* e *Pseudomonas* (EGGERS et al., 2004). Pesquisas posteriores mostraram que, em procariotos, a ecotina funciona como uma versátil inibidora competitiva de uma variedade de serinopeptidases (BABINE; ABDEL-MEGUID, 2008, p. 171–173).

Há uma infinidade de inibidores de serinopeptidase nos mais diferentes tipos de organismos, algo que é de se esperar em vista da importância metabólica desta classe de peptidases. Em sua maioria são pequenas moléculas com um alto grau de especificidade de substrato, raramente contendo mais de duzentos aminoácidos. A maior distinção da ecotina em relação a outros ISPs é sua ação relativamente genérica, ou seja, sua eficácia inibidora contra uma lista de diferentes enzimas serinopeptidases com funções biológicas tão distintas quanto tripsina, quimotripsina, elastase, colagenase, calicreína, etc (MCGRATH; GILLMOR; FLETTERICK, 1995).

Em *E. coli* o gene da ecotina codifica um monômero de 142 aminoácidos, sendo ambos os resíduos no sítio de clivagem P1 e P1' ocupados por metioninas. Sozinha, a molécula se apresenta como um homodímero, mas complexada com moléculas de serinopeptidases a ecotina forma um complexo quaternário. A estrutura resultante é um heterotetrâmero composto por dois monômeros de ecotina e duas unidades da serino-peptidase, e é bastante estável por ter vários pontos de interação molecular. A primeira unidade da enzima inibida tem dois pontos de contato com o homodímero: um dos monômeros de ecotina faz interface com o sítio ativo da serino-peptidase, enquanto o outro monômero interage com um sítio distal secundário. A segunda unidade de serinopeptidase realiza as mesmas interações com o dímero de ecotina, com sentido invertido no lado oposto do par de proteínas. Cada molécula de ecotina monomérica interage, por sua vez, com o outro monômero de ecotina, com o sítio primário de uma das unidades da enzima alvo e com um sítio secundário na segunda unidade inibida. A ecotina é o único inibidor conhecido de serinopeptidases que faz interfaces em mais de um sítio da superfície

enzimática desse modo, representado na fig. 1 abaixo (MCGRATH; GILLMOR; FLETTERICK, 1995).

Fig. 1 - Interação da ecotina (triângulos vermelho e laranja) com seus substratos, serinopeptidases (quadriláteros verde e azul).



Fonte: MCGRATH; GILLMOR; FLETTERICK, 1995 (fig. 6)

A ecotina parece ser uma potente ferramenta procariota na evasão do sistema imunológico. Um exemplo clássico é o da inibição de NE: em humanos a enzima elastase neutrofílica (abreviada NE, do inglês *neutrophil elastase*), uma serinopeptidase produzida em granulócitos neutrófilos, é uma das principais imuno defesas no combate a invasores patogênicos. Um dos meios de ação da NE é a clivagem da proteína de membrana externa A (OmpA, do inglês *outer membrane protein A*) em bactérias gram-negativas fagocitadas, dificultando sua multiplicação (BELAAOUAJ; KIM; SHAPIRO, 2000; WEINRAUCH et al., 2002). Por esta razão, a ecotina é um alvo importante para pesquisas farmacológicas (BABINE; ABDEL-MEGUID, 2008, p. 171).

Em eucariotos, têm havido uma variedade de pesquisas sobre as funções de ISPs no gênero *Leishmania*. No sequenciamento genético de *Leishmania major*, espécie evolutivamente bastante próxima de *T. cruzi*, foram encontrados três genes codificantes de ortólogos de ecotina, denominados ISP1, ISP2 e ISP3 (ESCHENLAUER et al., 2009). O ISP2 é produzido por *L. major* durante todas as fases de seu ciclo vital, com concentração maior nas fases infectantes promastigotas metacíclicas, e age de maneira análoga à ecotina de *E. coli*, inibindo elastase neutrofílica (NE) nos macrófagos do hospedeiro (FARIA et al., 2011). Morrison et. al, em estudo publicado em 2012, chegaram à conclusão que a ISP1 em *L. major* divergiu durante sua evolução, assumindo função completamente diferente no ciclo de vida do parasito: perdeu boa parte de sua capacidade inibidora de

serino proteases, e passou a ter função endógena, provavelmente relacionada à formação e manutenção do corpo flagelar (MORRISON et al., 2012).

A presença de um gene codificante de ISP em *T. cruzi*, táxon próximo de *L. major*, indica que há probabilidade deste gene ter função exógena importante no ciclo de vida do organismo, protegendo a célula invasora das imuno defesas do hospedeiro de maneira análoga ao gene ISP2 de *L. major* e da ecotina de *E. coli*. Portanto, a caracterização de ISP de *T. cruzi* pode ser um importante passo na construção de futuras terapias e profilaxias para a Doença de Chagas.

2. METODOLOGIA

As sequências de aminoácidos dos ISPs foram obtidas do banco de dados NCBI (*National Center for Biotechnology Information, U.S. National Library of Medicine*) utilizando seu pacote de busca automatizado BLAST, com a ecotina de *E. coli* como alvo de busca (BORATYN et al., 2013). O alinhamento de sequências utilizando o algoritmo *MUSCLE* e a análise filogenética de *maximum likelihood* foram realizados no software *SeaView* (GOUY; GUINDON; GASCUEL, 2010), com a seleção de matrizes de substituição de aminoácidos com melhor ajuste realizado com o software *PROTTEST 3* (DARRIBA et al., 2011). Os arquivos de árvore resultantes foram editados manualmente para padronizar as etiquetas de folhas terminais, e os arquivos de imagem dos cladogramas aqui apresentados foram exportados utilizando a ferramenta online *iTol Web Tree Tool* (LETUNIC; BORK, 2011).

Para realizar entrada de dados no banco de dados do projeto de visualização de loci, sequências gênicas completas para análises de locos gênicos foram obtidas do mesmo banco de dados NCBI utilizando a busca *tBLASTn*, uma ferramenta que toma sequência de aminoácidos como entrada e busca por sequências de nucleotídeos correspondentes, com a busca limitada a genomas anotados no *RefSeq* (BORATYN et al., 2013). Os dados foram inseridos e manipulados utilizando-se software customizado, desenvolvido para a tese de Mestrado do autor (FUHLENDORF, 2018). A lista completa de sequências com os respectivos links e IDs *NCBI GenBank* estão listados na Tabela 1. Após a entrada de dados ser completada o banco de dados foi manipulado manualmente utilizando chamados SQL para identificar genes próximos à esquerda e à direita dos homólogos de ecotina em várias espécies, com auxílio do mapa visual gerado pelo software mostrando que os homólogos de ecotina ocorrem, na maioria das espécies, em dois loci diferentes separados por 50 kbp. Um destes genes é um gene hipotético codificante de catanina, e os outros três são três genes hipoteticamente codificantes de proteínas, que foram etiquetados CHP1, CHP2 e CHP3. Utilizando o genoma de *L. braziliensis* como referência, estas quatro sequências de aminoácidos foram passadas pela ferramenta NCBI *tBLASTn*, com as mesmas configurações utilizadas para os homólogos de ecotina, e os CDSs resultantes foram manualmente etiquetados *CHP1*, *CHP2*, *CHP3* e *katanin-like* no banco de dados do visualizador de loci. Os labels e imagens foram gerados em inglês para facilitar seu uso em eventual publicação dos resultados.

Tabela 1 - Lista de itens *RefSeq* inseridos no banco de dados do gerador de imagens de loci gênicos

Leishmania braziliensis MHOM/BR/75/M2904	NC_009307.2	https://www.ncbi.nlm.nih.gov/nuccore/NC_009307.2
Leishmania braziliensis MHOM/BR/75/M2904	NC_018242.1	https://www.ncbi.nlm.nih.gov/nuccore/NC_018242.1
Leishmania infantum JPCM5	NC_009399.2	https://www.ncbi.nlm.nih.gov/nuccore/NC_009399.2
<i>Leishmania major</i> strain Friedlin	NC_007256.2	https://www.ncbi.nlm.nih.gov/nuccore/NC_007256.2
Leishmania mexicana MHOM/GT/2001/U1103	NC_018319.1	https://www.ncbi.nlm.nih.gov/nuccore/NC_018319.1
Leishmania panamensis	NC_025860.1	https://www.ncbi.nlm.nih.gov/nuccore/NC_025860.1
Leptomonas pyrrocoris	NW_015438382.1	https://www.ncbi.nlm.nih.gov/nuccore/NW_015438382.1
Leptomonas pyrrocoris	NW_015438394.1	https://www.ncbi.nlm.nih.gov/nuccore/NW_015438394.1
Trypanosoma brucei brucei TREU927	NC_007278.1	https://www.ncbi.nlm.nih.gov/nuccore/NC_007278.1
Trypanosoma brucei gambiense DAL972	NC_026738.1	https://www.ncbi.nlm.nih.gov/nuccore/NC_026738.1
<i>Trypanosoma cruzi</i> strain CL Brener	NW_001849489.1	https://www.ncbi.nlm.nih.gov/nuccore/NW_001849489.1
Trypanosoma grayi	NW_008825978.1	https://www.ncbi.nlm.nih.gov/nuccore/NW_008825978.1
Trypanosoma grayi	NW_008826261.1	https://www.ncbi.nlm.nih.gov/nuccore/NW_008826261.1

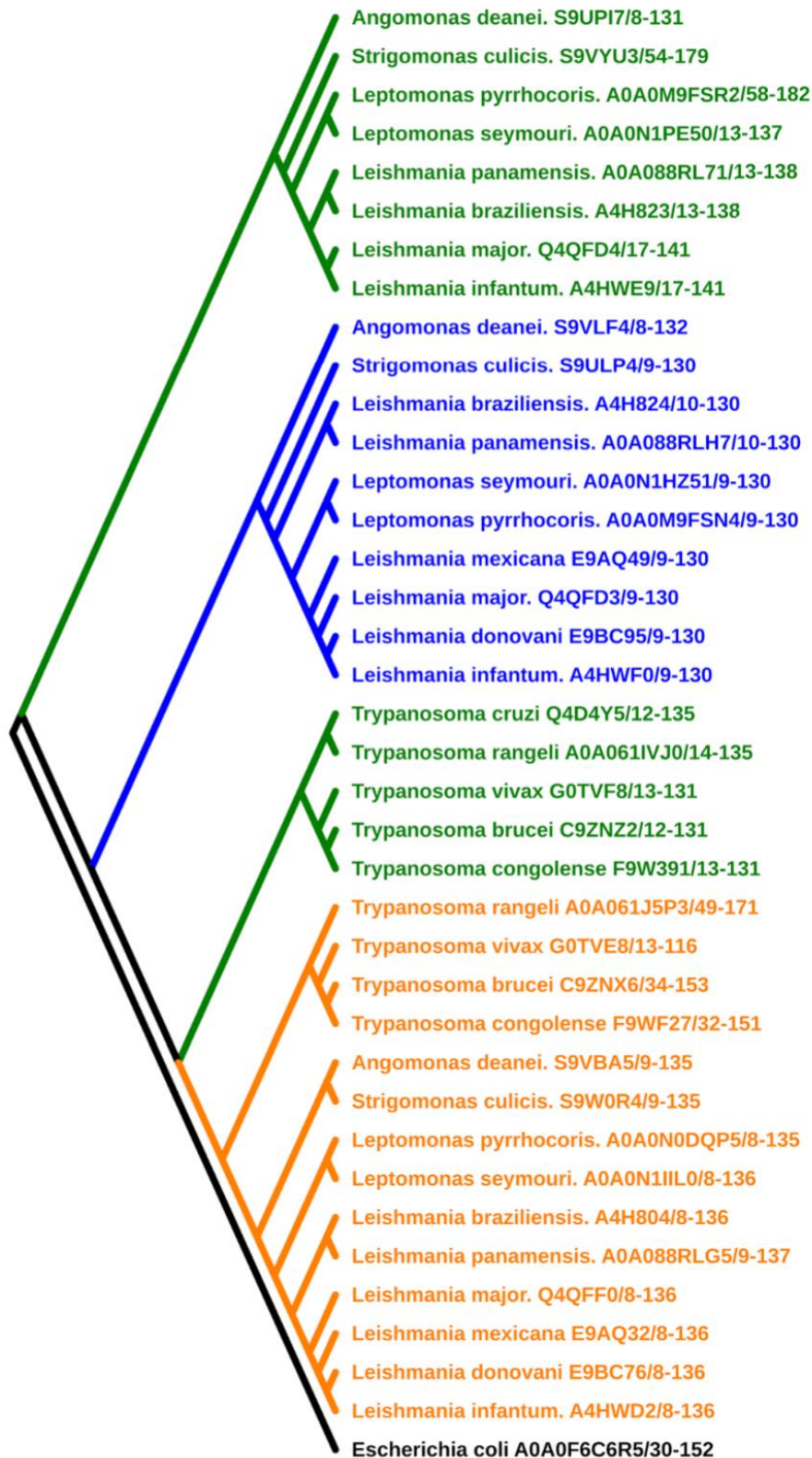
Apesar do foco principal desta pesquisa ser a análise de dados pré-existentes, nós também expressamos um homólogo de ecotina de *T. cruzi* para testar sua atividade *in-vitro* contra uma variedade de serinoproteases. Os testes de atividade enzimática ainda estão ocorrendo. Para esta parte do projeto, realizado sob orientação e no laboratório da Prof.^a Dr.^a Márcia Aparecida Sperança da Universidade Federal do ABC, o gene codificante de ISP2 do *T. cruzi* cepa Y foi extraído por técnica de PCR a partir de DNA purificado, com uso de *primers* desenhados sob medida. O gene foi então clonado e expressado no Laboratório de Agentes Patogênicos utilizando técnicas padronizadas de expressão recombinante de Sambrook e Russel (SAMBROOK; RUSSELL, 2001).

3. RESULTADOS E DISCUSSÃO

O ISP2 de *T. cruzi* foi expresso e purificado com sucesso, e atualmente está passando por testes de reação enzimática no Laboratório de Agentes Patogênicos na Universidade Federal do ABC, em um grupo de estudo separado liderado pela Prof.^a Sperança. Estudos preliminares indicam que esta ISP2 de *T. cruzi* inibe fortemente serinoproteases como a tripsina e a neutrophil elastase, com eficiência ao menos tão forte quanto a da ecotina de *E. coli*. Dados numéricos ainda estão indisponíveis, visto que os testes laboratoriais ainda estão em andamento, mas estes resultados preliminares parecem indicar que a ISP2 de *T. cruzi* muito provavelmente age de maneira similar aos ISP2 das espécies do gênero *Leishmania*, protegendo o parasita através da inibição dos macrófagos do hospedeiro (FARIA et al., 2011).

Análise por inferência filogenética *maximum likelihood* resultou na árvore representada na fig. 2, com a ecotina de *E. coli* como o grupo externo e as ISPs 1, 2 e 3 marcadas com cores, laranja, verde e azul respectivamente. A topologia da árvore indica fortemente que os ISPs homólogos de ecotina diferenciaram-se há muito tempo, sendo provável que os ISP1 e ISP2 ao menos tenham estado presentes no genoma destes organismos desde a diferenciação entre os gêneros *Leishmania* e *Trypanosoma*.

Figura 2 - Cladograma de *maximum likelihood* dos homólogos de ecotina (ISPs), com ISP1 marcado em laranja, ISP2 marcado em verde e ISP3 marcado em azul



A utilização do software de geração de imagens de loci gênicos, descrita em detalhes na tese de Mestrado do autor (FUHLENDORF, 2018), Resultou das imagens das figuras 3 e 4, que mostram os cromossomos inteiros e um recorte de uma área de interesse

respectivamente. A fig. 3 é útil apenas para permitir a comparação entre os dados de *Leishmania* em relação à sua posição geral no cromossomo: os homólogos de ecotina ocorrem no cromossomo 15 destas espécies; os dados das outras espécies ou está incompleto ou foi mal anotado, resultando em grandes sequências contig. A fig. 4 clarifica as posições cromossômicas, mas dados para *T. grayi* e *Leptomonas pyrrhocoris* ainda são fragmentados.

Figura 3 - Resultado do software gerador de imagens de loci, com os cromossomos inteiros obtidos do GenBank anotados com as etiquetas: ISPs em vermelho, e genes circundantes CHP1, CHP2, CHP3 e katanina em verde, amarelo, azul e laranja, respectivamente

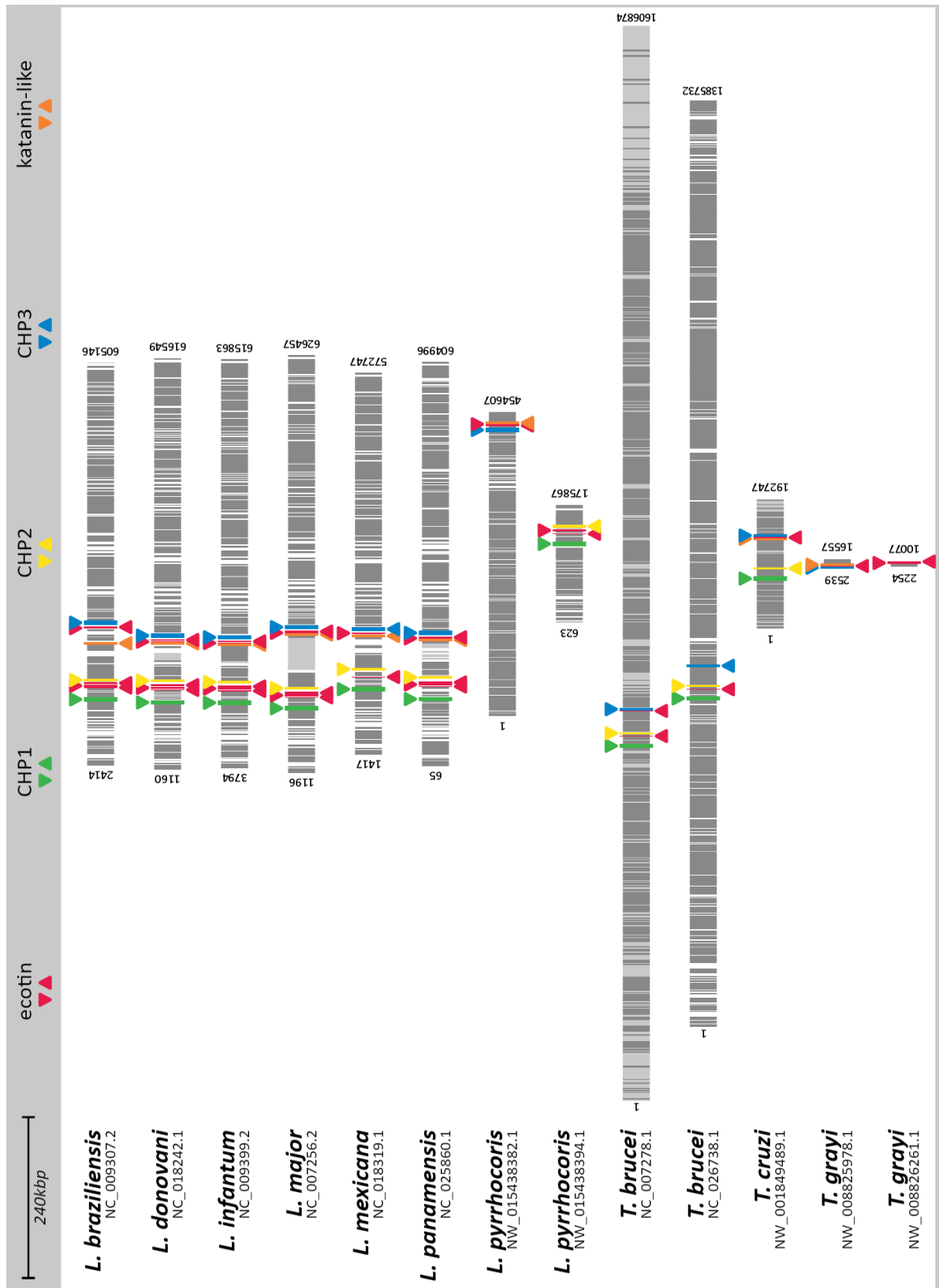
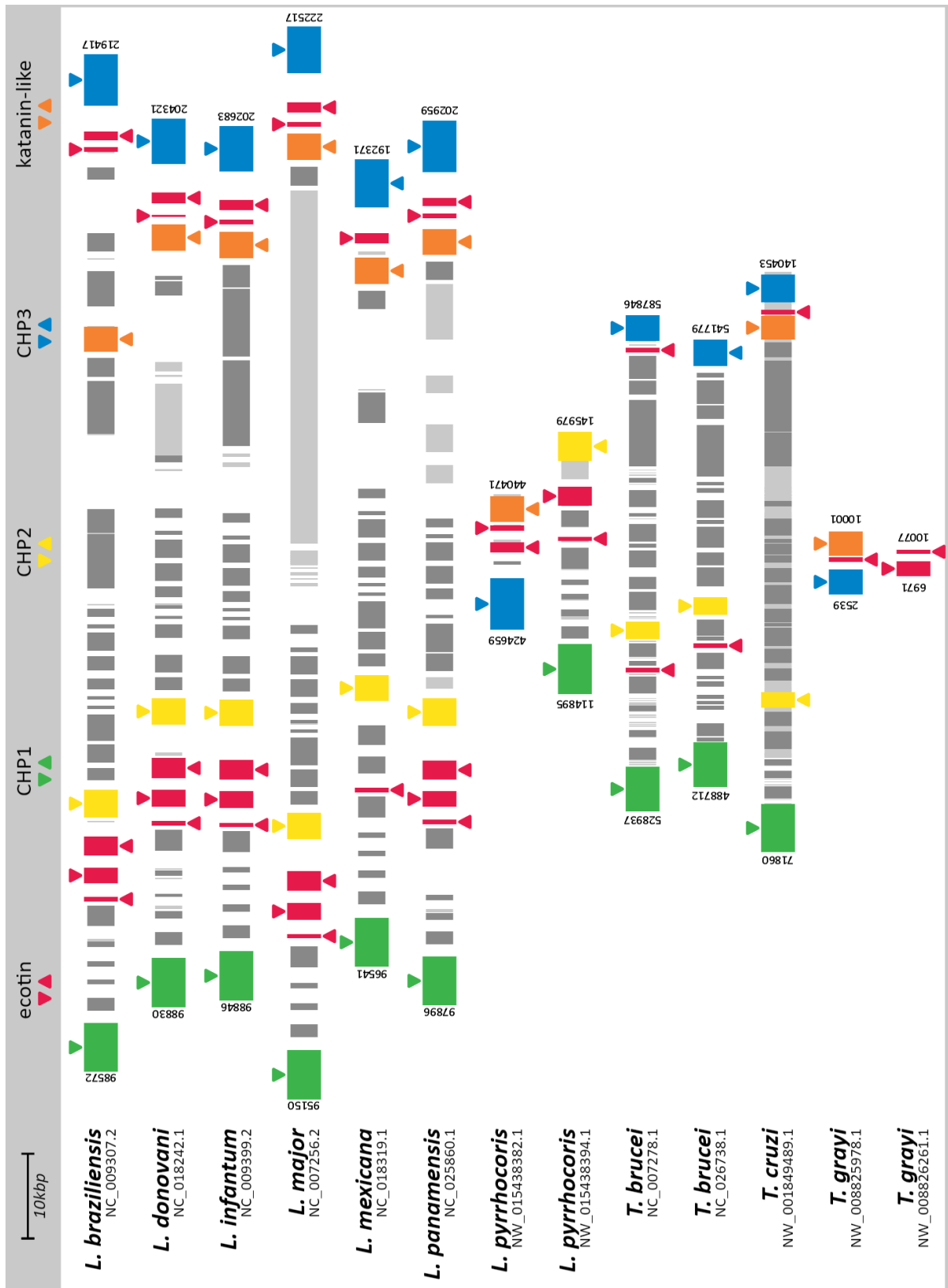
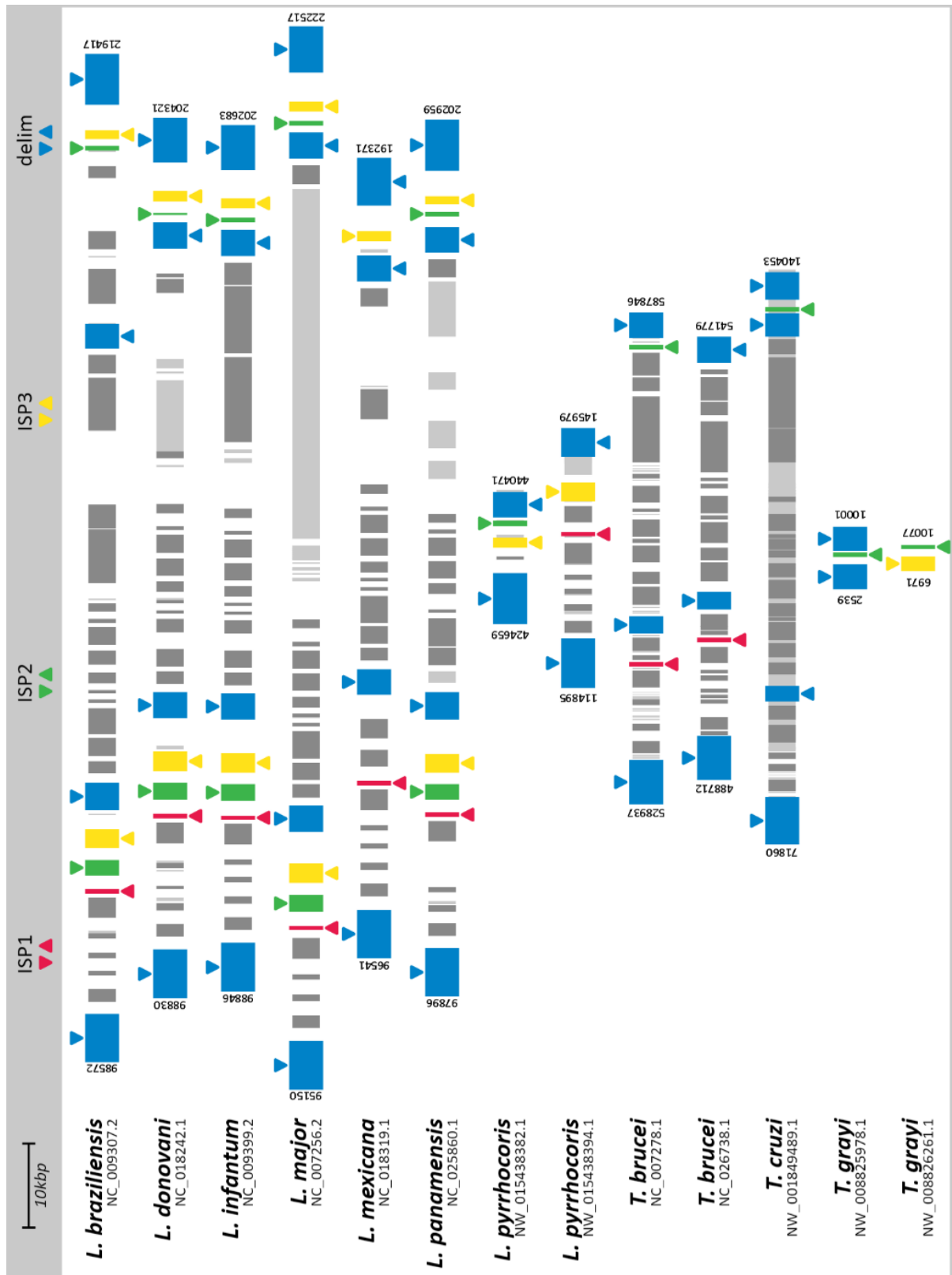


Figura 4 - Resultado do software gerador de imagens de loci: recorte das áreas de interesse delimitadas pelos genes circundantes, com as etiquetas: ISPs em vermelho, e genes circundantes CHP1, CHP2, CHP3 e katanina em verde, amarelo, azul e laranja, respectivamente



As imagens das figuras 3 e 4 mostram até 5 homólogos de ecotina nas várias espécies de *Leishmania*. Essa duplicação não havia sido publicada em artigos anteriores. Análise de dados de sequência e inspeção visual do contexto genômico dado por essas imagens conferem forte suporte à ideia de que homólogos de ecotina sofreram várias duplicações e/ou múltiplos eventos de transferência genética horizontal antes da diferenciação dos gêneros modernos em Trypanosomatida. Examinando-se a fig. 4 com cuidado pode-se notar que a maioria das espécies de *Leishmania* manteve todas as cinco cópias de ISP, enquanto que as espécies de *Trypanosoma* perderam ao menos algumas cópias. Para facilitar esta análise visual, uma segunda imagem com recortes das áreas de interesse foi gerada (fig. 5), através de manipulação manual do banco de dados. Na nova imagem gerada, as etiquetas dos genes circundantes CHP 1-3 e *katanin-like* foram alteradas para “delim” (em azul), e os homólogos prováveis de ISP1, ISP2 e ISP3 foram identificados com cores, vermelho, verde e amarelo respectivamente. Em todas estas imagens, os genes circundantes CHP 1-3 e *katanin-like* (ou “delim” na fig. 5) Servem como etiquetas acessórias, delimitando os loci completos dos homólogos de ecotina e auxiliando na identificação de qual variante o homólogo representa, ISPs 1, 2 ou 3.

Figura 5 - Resultado do software gerador de imagens de loci: esta imagem representa os mesmos dados da fig. 4, mas com as ISPs individuais identificadas em ISP 1, 2 e 3 (vermelho, verde e amarelo respectivamente), e com todas as proteínas circundantes identificadas em azul, e mantidas na imagem apenas para ilustrar os loci gênicos corretamente.



Sabemos que os ISPs têm sua origem em bactérias, e que surgiram no genoma de parasitas Kinetoplastida através de transferência genética horizontal, porque eles não aparecem em nenhum outro eucarioto. A pergunta sem resposta é como. Observando-se as figuras 4 e 5, e mantendo em mente o cladograma da fig. 2, podemos formar uma hipótese de como os tripanossomatídeos adquiriram seus homólogos de ecotina. O próximo parágrafo é especulativo, mas dada a evidência, provavelmente não está muito longe da realidade:

O ancestral de todos os tripanossomatídeos participou de vários eventos de transferência genética horizontal com bactérias possuidoras de ecotina, ou houve um evento de transferência único seguido de várias duplicações gênicas ao longo da evolução dos atuais membros do grupo, ou houve várias transferências e duplicações. Se houve várias transferências genéticas horizontais, elas provavelmente não ocorreram mais do que três vezes, para gerar a ISP1, ISP2 e ISP3, e as cópias adicionais de ISP2 e ISP3 que aparecem nas espécies de *Leishmania* foram provável resultado de duplicações subsequentes. As posições de ISP1 e ISP2 em *T. brucei*, no primeiro e segundo locus de ecotina respectivamente, e a provável homologia da ISP2 com a ISP2 de *T. cruzi*, apontam para uma duplicação ancestral do locus, anterior à divergência dos gêneros *Leishmania* e *Trypanosoma*. Neste cenário as espécies de tripanossomas perderam cópias do gene subsequentemente, que faz sentido visto que os genomas de *Trypanosoma* são muito mais compactos do que os genomas de *Leishmania* nas imagens apresentadas; este fato sugere que, ao longo do tempo, as espécies do primeiro grupo perderam muito mais genes por deleção do que as espécies do segundo grupo, e reforça a probabilidade de que *T. cruzi* e *T. brucei* tenham perdido cópias de ISP ao longo de sua evolução. Uma possível sequência de eventos baseada neste conjunto de dados limitado é a seguinte: o ancestral comum de Trypanosomatida tinha três cópias de ecotina no primeiro locus (na posição aproximada de 120kbp do cromossomo 15), obtidos através de transferências laterais com bactérias possuidoras do gene, ou através de uma simples transferência seguida de duplicações contíguas. Posteriormente, os genes ancestrais dos ISPs 2 e 3 neste locus sofreram um evento de duplicação simultânea, criando o segundo locus em aproximadamente 190kbp no mesmo cromossomo. Subsequentemente, várias das espécies descendentes perderam algumas destas cópias.

A preservação do homólogo ISP2 em várias espécies destes parasitas é um fato interessante, e faz sentido dada a ampla evidência da sua importância contra os sistemas imunológicos dos hospedeiros. Outro fato interessante é que os parasitas *T. brucei*

preservaram o variante ISP1 em todos os casos analisados, enquanto que *T. cruzi* os perderam. Como o ISP1 parece estar envolvido no desenvolvimento flagelar e na mobilidade de promastigotos que habitam o vetor inseto nas espécies de *Leishmania* (MORRISON et al., 2012), esta poderia ser a razão de sua preservação em *T. brucei* e preservação em *T. cruzi*. Estas espécies são membros das seções Salivaria e Stercoraria respectivamente, com ciclos de vida e métodos de transmissão diferentes. Enquanto o *T. cruzi* é transmitido por hemípteros, com a deposição de parasitas infectantes nas fezes sobre a picada na pele do hospedeiro vertebrado, os parasitas infectantes de *T. brucei* vivem na glândula salivar de insetos dípteros, e são injetados nos hospedeiros vertebrados pela probóscide destes insetos, como ocorre nas espécies de *Leishmania*. É possível que a variante ISP1 confira alguma vantagem aos tripanossomatídeos que têm nos insetos dípteros seu hospedeiro intermediário. Esta associação hipotética pode servir como base de pesquisas futuras, com aplicações potenciais em saúde pública.

As especulações apresentadas nos últimos parágrafos são apresentadas apenas para encorajar pesquisas futuras. Por mais tentador que seja afirmar sua validade, nosso conjunto de dados é muito limitado em abrangência e de baixa qualidade em alguns casos para que façamos afirmações assertivas. A anotação genômica automatizada que resulta na maior parte dos nossos dados é limitada, e algumas das sequências utilizadas apresentam erros, omissões ou outros problemas. Analisando os dados de *L. mexicana* na figura 4, por exemplo, é aparente que as sequências codificantes entre a primeira ocorrência de ISP e a etiqueta CHP2 são de fato ISP2 e ISP3, mas na anotação genômica automatizada elas aparecem como “proteínas desconhecidas”. Mesmo assim, com o cada vez mais rápido crescimento no volume de dados genômicos disponíveis, estas especulações poderão ser desenvolvidas no futuro com o surgimento de dados mais precisos e confiáveis.

Este trabalho demonstra sem sombra de dúvida a ubiquidade de grandes volumes de dados genômicos online que ainda não foram estudados. A quantidade de informação que pode ser obtida com baixíssimo investimento financeiro, utilizando-se ferramentas de bioinformática gratuitas, já é enorme; hoje, com o universo de Big Data e com custos de sequenciamento em queda exponencial, este fato se tornará cada vez mais óbvio com a passagem do tempo. A próxima geração de biólogos pode muito bem ser obrigada a aprender a programar antes de aprender os nomes de todas as famílias de plantas e animais.

4. CONSIDERAÇÕES FINAIS

A biologia é a mais interdisciplinar das chamadas “ciências duras”. Um biólogo não pode mais se contentar com um conhecimento enciclopédico de fatos, como foi comum no século passado. Hoje, biólogos aspirantes precisam lembrar que a biologia moderna existe tanto no trato de grandes volumes de dados quanto no conhecimento de fisiologia e nomenclatura. Abraçar ferramentas e métodos de bioinformática é o único caminho à frente para aqueles que não querem ser deixados para trás nas ciências biológicas, especialmente em genética, ecologia e outros campos onde o volume de dados na pesquisa contemporânea sobrepassa tão completamente a capacidade de análise de um indivíduo sem auxílio computacional.

Adiante, nós biólogos precisamos nos adaptar à este admirável novo mundo de Big Data e análise computacional — e empurrar nossos colegas a fazer o mesmo — ou arriscar nos tornarmos rapidamente obsoletos.

REFERÊNCIAS BIBLIOGRÁFICAS

- ARAUJO-JORGE, T. **Doença de Chagas-Fiocruz**. Disponível em: <<http://www.agencia.fiocruz.br/doen%C3%A7a-de-chagas>>. Acesso em: 9 abr. 2015.
- BABINE, R. E.; ABDEL-MEGUID, S. S. **Protein Crystallography in Drug Discovery**. [s.l.] John Wiley & Sons, 2008.
- BELAAOUAJ, A.; KIM, K. S.; SHAPIRO, S. D. Degradation of outer membrane protein A in *Escherichia coli* killing by neutrophil elastase. **Science**, v. 289, n. 5482, p. 1185–1188, ago. 2000.
- BORATYN, G. M. et al. BLAST: a more efficient report with usability improvements. **Nucleic Acids Res.**, v. 41, n. W1, p. W29–W33, jul. 2013.
- BRUM-SOARES, L. M. et al. Morbidity of Chagas disease among autochthonous patients from the Rio Negro microregion, State of Amazonas. **Rev. Soc. Bras. Med. Trop.**, v. 43, n. 2, p. 170–177, 2010.
- CHUNG, C. H. et al. Purification from *Escherichia coli* of a periplasmic protein that is a potent inhibitor of pancreatic proteases. **J. Biol. Chem.**, v. 258, n. 18, p. 11032–11038, set. 1983.
- DARRIBA, D. et al. ProtTest 3: fast selection of best-fit models of protein evolution. **Bioinformatics**, v. 27, n. 8, p. 1164–1165, abr. 2011.
- EGGERS, C. T. et al. The periplasmic serine protease inhibitor ecotin protects bacteria against neutrophil elastase. **Biochem. J.**, v. 379, n. Pt 1, p. 107–118, abr. 2004.
- ESCHENLAUER, S. C. P. et al. Influence of parasite encoded inhibitors of serine peptidases in early infection of macrophages with *Leishmania major*. **Cell. Microbiol.**, v. 11, n. 1, p. 106–120, 2009.
- FARIA, M. S. et al. *Leishmania* inhibitor of serine peptidase 2 prevents TLR4 activation by neutrophil elastase promoting parasite survival in murine macrophages. **J. Immunol.**, v. 186, n. 1, p. 411–422, jan. 2011.
- FUHLENDORF, M. M. **Loci image generator and the evolution of trypanosomatid ecotin : customized software as a tool for evolutionary analysis**. Tese de mestrado—Santo André: UFABC, 2018.
- GOUY, M.; GUINDON, S.; GASCUEL, O. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. **Mol. Biol. Evol.**, v. 27, n. 2, p. 221–224, fev. 2010.
- IVENS, A. C. et al. The genome of the kinetoplastid parasite, *Leishmania major*. **Science**, v. 309, n. 5733, p. 436–442, jul. 2005.
- LETUNIC, I.; BORK, P. Interactive Tree Of Life v2: online annotation and display of phylogenetic trees made easy. **Nucleic Acids Res.**, v. 39, n. Web Server issue, p. W475–8, jul. 2011.
- MCGRATH, M. E.; GILLMOR, S. A.; FLETTERICK, R. J. Ecotin: lessons on survival in a protease-filled world. **Protein Sci.**, v. 4, n. 2, p. 141–148, fev. 1995.
- MEROPS Family I11**. [s.l: s.n.].
- MORRISON, L. S. et al. Ecotin-like serine peptidase inhibitor ISP1 of *Leishmania major* plays a role in flagellar pocket dynamics and promastigote differentiation. **Cell. Microbiol.**, v. 14, n. 8, p. 1271–1286, ago. 2012.
- PASSOS, L. A. C. et al. Sobrevivência e infectividade do *Trypanosoma cruzi* na polpa de açaí: estudo in vitro e in vivo. **Epidemiologia e Serviços de Saúde**, v. 21, n. 2, p. 223–232, 2012.
- SAMBROOK, J.; RUSSELL, D. W. **Molecular cloning: a laboratory manual**. 3rd ed ed. Cold Spring Harbor, N.Y: Cold Spring Harbor Laboratory Press, 2001.
- WEINRAUCH, Y. et al. Neutrophil elastase targets virulence factors of enterobacteria. **Nature**, v. 417, n. 6884, p. 91–94, maio 2002.