



MINISTÉRIO DA EDUCAÇÃO
UNIVERSIDADE FEDERAL DO PIAUÍ
CAMPUS SENADOR HELVÍDIO NUNES DE BARROS
Chefia do Curso de Ciências Biológicas-Modalidade Licenciatura



RAYLLA CAROLINE DE SOUSA

**DESENVOLVIMENTO DE UM MATERIAL DIDÁTICO TEÓRICO-PRÁTICO
AUXILIAR NO ENSINO DE BIOINFORMÁTICA**

Picos
2013

RAYLLA CAROLINE DE SOUSA

**DESENVOLVIMENTO DE UM MATERIAL DIDÁTICO TEÓRICO-PRÁTICO
AUXILIAR NO ENSINO DE BIOINFORMÁTICA**

Trabalho de Conclusão de Curso apresentado à Coordenação do curso de Ciências Biológicas, modalidade Licenciatura, do Campus Senador Helvídio Nunes de Barros, Universidade Federal do Piauí, como pré-requisito para a obtenção do título de Licenciado em Ciências Biológicas.

Orientador:

Prof. Dr. Luís Evêncio da Luz

Picos

2013

Eu, **Raylla Caroline de Sousa**, abaixo identificado (a) como autor (a), autorizo a biblioteca da Universidade Federal do Piauí a divulgar, gratuitamente, sem ressarcimento de direitos autorais, o texto integral da publicação abaixo discriminada, de minha autoria, em seu site, em formato PDF, para fins de leitura e/ou impressão, a partir da data de hoje.

Picos-PI, 07 de Junho de 2013.

FICHA CATALOGRÁFICA

Serviço de Processamento Técnico da Universidade Federal do Piauí
Biblioteca José Albano de Macêdo

S725d Sousa, Raylla Caroline de.
 Desenvolvimento de um material didático teórico-prático auxiliar no ensino de bioinformática / Raylla Caroline de Sousa. – 2013.
 CD-ROM : il. ; 4 ¾ pol. (97 p.)

Monografia (Licenciatura em Ciências Biológicas) – Universidade Federal do Piauí. Picos-PI, 2013.

Orientador (A): Prof. Dr. Luís Evêncio da Luz

1. Ensino. 2. Pesquisa. 3. Material Didático. 4. Bioinformática. I. Título.

CDD 570.280 7

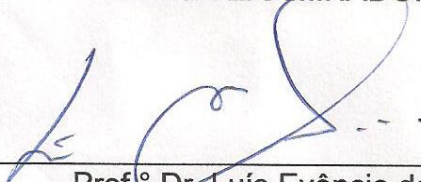
RAYLLA CAROLINE DE SOUSA

**DESENVOLVIMENTO DE UM MATERIAL DIDÁTICO TEÓRICO-PRÁTICO
AUXILIAR NO ENSINO DE BIOINFORMÁTICA**

Monografia apresentada à coordenação do Curso de Ciências Biológicas da Universidade Federal do Piauí, como um dos requisitos para a obtenção do Título de Licenciado em Ciências Biológicas.

Aprovada pela Banca Examinadora em Picos (PI), 10 / 04 / 2013.

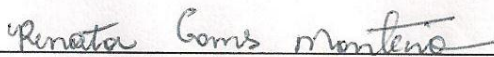
BANCA EXAMINADORA



Prof.º Dr. Luís Evêncio da Luz
Universidade Federal do Piauí-UFPI/CSHNB
Orientador



Prof.º Dr. Daniel Barbosa Liarte
Universidade Federal do Piauí-UFPI/ CMPP
Titular



Prof.ª Me. Renata Gomes Monteiro
Universidade Federal do Piauí-UFPI/ CSHNB
Titular

Prof.ª Me. Nilda Masciel Neiva Gonçalves
Universidade Federal do Piauí-UFPI/ CSHNB
Suplente

Dedico esta vitória a **Deus**, aos meus pais (**Renilda e João**) e aos meus irmãos (**Renilson e Renilcilanne**) por todo amor que me dão e por me fazer perceber a grande beleza da vida.

AGRADECIMENTOS

Á **Deus** por me conceder o dom da vida e da inteligência, estando sempre presente, guiando-me no caminho da verdade transformando sempre a minha vida, me dando força, ressuscitando meus sonhos. Deus estará sempre comigo!

Aos meus pais (**Renilda e João**) que sempre estiveram presentes, me dando muito amor e carinho, me ajudando vencer os obstáculos e alcançar esta conquista fazendo o possível sem medir sacrifícios. Sem vocês não teria conseguido. Amo vocês!

Aos meus irmãos (**Renilson e Renilcilanne**) que estiveram sempre ao meu lado me dando apoio, carinho, incentivo. Esta conquista também pertence a vocês. Amo vocês!

Ao meu namorado, amigo (**Jheycon**) pela cumplicidade, pela ajuda generosa, paciência, incentivo e por entender verdadeiramente o que esta pesquisa significa para mim. Te Amo!

A **minha família**, em geral, que sabiamente contribuiu, fazendo com que eu chegasse até aqui com menos dificuldade. Vocês sempre comprometidos com o bem estar do próximo.

Ao professor **Dr. Daniel Barbosa Liarte**, pela enorme dedicação a este projeto, pela confiança e paciência que teve comigo, pelos brilhantes ensinamentos, você é um excelente profissional e pessoa, nunca havia conhecido alguém tão cheio de sabedoria e honra. Apesar de todas as dificuldades enfrentadas e momentos de superação as conquistas foram superiores!

Ao professor **Dr. Luís Evêncio da Luz**, que me acolheu, a quem muitas vezes recorri nos momentos de dúvida. Permitindo a realização desta conquista. Obrigada pelos ensinamentos e excelentes sugestões!

Aos demais professores membros da banca examinadora (**Me. Renata Monteiro Gomes e Me. Nilda Masciel Neiva Gonçalves**) pelo tempo dispensado a leitura desse trabalho e pelas contribuições que certamente me farão crescer profissionalmente.

A minha amiga (**Danielly Costa**), que mesmo distante me fez despertar o potencial que carrego comigo, me incentivando a lutar pelo o que desejasse você estará sempre no meu coração.

As minhas amigas da universidade (**Rayane Macedo, Any Doésia, Conceição de Maria, Maria Francisca, Maura Moura**) que seguiram juntas comigo nessa caminhada, minhas companheiras de conquista.

Aos meus amigos do peito (**Ailkon e Loisláyne**) pelos momentos de alegrias e sonhos, sempre me estimularam a continuar e nunca desistir. Vocês são maravilhosos!

Aos professores da **Universidade Federal do Piauí – Campus de Picos**, que contribuíram para o meu desenvolvimento acadêmico e pessoal.

Aos graduandos do curso de **Ciências Biológicas** e do curso de **Nutrição** da **Universidade de Federal do Piauí – Campus de Teresina** que se dispusera a colaborar com o desenvolvimento dessa pesquisa.

A turma de **Biologia 2008.2** pelo carinho, amizade e respeito que tiveram comigo, apesar das brigas dividimos muitos momentos de alegrias e vitórias.

A todos que colaboraram para com a minha vida acadêmica minha sincera gratidão.

MUITO OBRIGADA A TODOS!

“Acredite que nenhum de nós
Já nasceu com jeito
Pra super-herói
Nossos sonhos
A gente é quem constrói
É vencendo os limites
Escalando as fortalezas
Conquistando o impossível
Pela fé.
Tantos recordes
Você pode quebrar
As barreiras
Você pode ultrapassar
E vencer
Campeão, vencedor
Deus dá asas, faz teu voo
Essa fé que te faz imbatível
Te mostra o teu valor”
(Jamily)

RESUMO

Bioinformática é o campo da ciência em que a biologia, ciência da computação e tecnologia da informação se fundem para formar uma única disciplina. O objetivo final do campo é permitir a descoberta de novos conhecimentos biológicos, bem como para criar uma perspectiva global de que princípios unificadores da biologia podem ser discernidos. A informática é um meio educacional que vem ganhando cada vez mais importância no cenário mundial, com as inovações tecnológicas acontecendo a todo instante nos sistemas de informação. A escola se vê no momento de se adequar a essas mudanças e de assimilar e incorporá-la ao contexto da educação. A importância do computador e as mudanças que ele provoca no processo de ensino e aprendizagem são hoje conhecidas, contudo, mesmo havendo o reconhecimento da necessidade do uso do computador e de programas computacionais como ferramentas educacionais, estes instrumentos são pouco familiares à maioria dos alunos e educadores brasileiros. O presente trabalho propõe uma análise da situação do ensino e pesquisa de Bioinformática verificando a ocorrência de cursos de pós-graduação disponível no Brasil bem como o ensino e pesquisa desenvolvida na graduação. Além do desenvolvimento de um material didático auxiliar no ensino de Bioinformática, que este conteve na sua organização conteúdo teórico e prático, citando algumas características particulares sobre Noções Básicas de Bioinformática, Banco de Dados, Obtenção e Edição de Sequências Biológicas, Alinhamento Global e Local, Mapa de Restrição. O manual didático visa contribuir como uma estratégia de ensino/aprendizagem na utilização das ferramentas básicas desta nova área promissora e moderna, a bioinformática.

Palavras Chaves: Ensino. Pesquisa. Material Didático. Bioinformática.

ABSTRACT

Bioinformatics is the field of science in which biology, computer science and information technology merge to form a single discipline. The ultimate goal of course is to enable the discovery of new biological knowledge as well as to create a global perspective from which unifying principles in biology can be discerned. The computer is an educational medium that is gaining more and more importance on the world stage, with technological innovations happening all the time in information systems. The school sees fit at the time of these changes and to assimilate and incorporate it into the context of education. The importance of the computer and the changes it causes in the process of teaching and learning are now known, however, even with the acknowledgment of the use of computers and computer programs as educational tools, these tools are unfamiliar to most students and Brazilian educators. This paper proposes a situation analysis of the teaching and research of Bioinformatics checking the occurrence of postgraduate courses available in Brazil as well as teaching and research developed at graduation. Besides the development of educational material aid in the teaching of Bioinformatics, which contained this organization in its theoretical content and practical, citing some particular characteristics of Understanding Bioinformatics, Database, Collection and Biological Sequence Editing, Alignment Global and Local Restriction map. The manual aims to contribute as a didactic strategy teaching / learning in the use of basic tools of this new and promising area modern bioinformatics.

Key Words: Education. Search. Handouts. Bioinformatics.

LISTA DE ILUSTRAÇÕES

Foto 01 - Turma de Ciências Biológicas (CMPP) - Laboratório de Biologia.....	27
Foto 02 - Turma de Nutrição (CMPP) – Sala de Aula.....	27
Foto 03 - Abordagem teórica de noções básicas de bioinformática.....	33
Foto 04 - Alunos realizando a prática: Acesso a diferentes bancos de dados.....	34
Foto 05 - Alunos na prática de edição de dados com o BioEdit.....	35
Foto 06 - Alunos realizando a prática de alinhamento global.....	36
Foto 07 - Alunos efetivando a prática de alinhamento local.....	36
Foto 08 - Alunos desenvolvendo a montagem do mapa de restrição.....	37

LISTA DE QUADROS

Quadro 01 - Cursos de Pós-Graduação em Bioinformática no Brasil.....	28
Quadro 02 - Grades Curriculares dos Cursos e <i>Campi</i> da Universidade Federal do Piauí Pesquisados.....	30

LISTA DE TABELAS

Tabela 01 - Respectivas médias avaliativas da palestra e do Material Didático.....	38
--	----

LISTA DE ABREVIATURAS E SIGLAS

AB³C: Associação Brasileira de Bioinformática e Biologia Computacional

BBNet: BrazilianBioNet

BLAST: Basic Local Alignment Search Tool

CAPES: Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

CGP: Coordenadoria Geral de Pesquisa

ClustalW: Alinhamento de Sequências de Proteínas ou DNA

CMPP: Campus Ministro Petrônio Portella

CMRV: Campus Ministro Reis Velloso

CTS: Ciência, Tecnologia e Sociedade

DNA: Ácido Desoxirribonucleico

IFSC/USP: Instituto de Física de São Carlos / Universidade de São Paulo

ISMB: Intelligent Systems for Molecular Biology

NBI: Núcleo de Bioinformática

NCBI: Centro Nacional para Informação em Biotecnologia

PRPPG: Pró-Reitoria de Pesquisa e Pós-Graduação

RENORBIO: Rede Nordeste de Biotecnologia

RNA: Ácido Ribonucleico

sp: Espécie

TIC's: Tecnologias de Informação e da Comunicação

UAPI: Universidade Aberta do Piauí

UFPI: Universidade Federal do Piauí

SUMÁRIO

1. INTRODUÇÃO.....	17
2. OBJETIVOS.....	19
2.1 Geral.....	19
2.2 Específicos.....	19
3. REFERENCIAL TEÓRICO.....	20
4. MATERIAL E MÉTODOS.....	24
4.1 Público Alvo.....	24
4.2 Levantamento Bibliográfico.....	24
4.3 Levantamento de Dados dos Cursos de Pós-Graduação em Bioinformática.....	24
4.4 Levantamento de Dados da Disciplina Bioinformática Ofertada na Graduação.....	25
4.5 Levantamento de Dados dos Projetos Pesquisa Cadastrados nos <i>Campi</i> da Universidade Federal do Piauí.....	25
4.6 Confecção do Material Didático.....	26
5. RESULTADOS E DISCUSSÕES.....	28
5.1 O Ensino de Bioinformática na Pós-Graduação.....	28
5.2 A Bioinformática na Graduação da Universidade Federal do Piauí.....	29
5.3 A Utilização da Bioinformática nas Pesquisas Realizadas na Universidade Federal do Piauí.....	31
5.4 Material Didático.....	32
5.4.1 Construção do Manual Didático – Ferramentas Básicas de Bioinformática.....	32
5.4.2 Avaliação do Material Didático Auxiliar Teórico-Prático.....	37
6. CONSIDERAÇÕES FINAIS.....	41
REFERÊNCIAS BIBLIOGRÁFICAS.....	42

APÊNDICES.....	45
APÊNDICE A – MATERIAL DIDÁTICO.....	46
APÊNDICE B – QUESTIONÁRIO DE AVALIAÇÃO DO MANUAL DIDÁTICO....	97

1. INTRODUÇÃO

A educação se insere nas novas estratégias de sobrevivência e de existência capazes de orientar as sociedades e culturas no sentido de se habilitarem melhor para interagir e trocar conhecimentos científicos, técnicos e tecnológicos em espaços reais e virtuais. Ganha força a defesa da tese da alfabetização científica e tecnológica, que vem sendo discutida desde os anos 70 e que contém em sua formulação o debate sobre a relação entre Ciência, Tecnologia e Sociedade (CTS) (GOUVÊA; LEAL, 2001).

A biologia celular e molecular teve grande desenvolvimento e, nos últimos anos, vem atravessando a fase denominada genômica, em que os pesquisadores se concentram na descrição do sequenciamento do repertório de genes de seres vivos (genomas), desde vírus e bactérias até o homem, e na identificação de genes responsáveis por características fenotípicas normais ou patológicas, com a perspectiva de decifrar e definir, nos próximos anos, as informações completas dos repertórios de genes típicos de cada espécie (SILVA, 2000).

Amui, 2006 afirma que os avanços científicos vem contribuindo cada vez mais nas áreas biológicas e científicas oferecendo ferramentas computacionais e sistemas específicos que em análise de dados *in silico*, que fornecem resultados rápidos e confiáveis. Essa abordagem *in silico* é representada por metodologias baseadas na criação de algoritmos dentro dessa nova e importante ciência do século 21, a bioinformática (PROSDOCIMI, 2003).

Esta ciência, a bioinformática, introduziu metodologias de análise das macromoléculas biológicas e de suas interações, permitindo a experimentação nas telas de computadores, com enorme economia de tempo e de complexas operações de bancada (SILVA, 2000).

A bioinformática envolve aspectos multidisciplinares, intimamente ligados ao conhecimento biológico, fundamentais no momento atual em que as pesquisas ressaltam o seu valor e utilidade para a Ciência. Suas aplicações são variadas, sejam no diagnóstico de doenças, desenvolvimentos de novos fármacos ou análise de dados originados de genomas sequenciados, a partir da utilização de algoritmos para comparar as sequências desses DNAs (AMUI, 2006).

Pelo fato da bioinformática se caracterizar como uma área recente, ela necessita de diretrizes claras e precisas para utilização de suas ferramentas

computacionais. Nesse sentido, o incentivo para o desenvolvimento deste estudo, foi com o intuito de verificar a situação do ensino e pesquisa de bioinformática e ainda fornecer noções básicas mais relevantes sobre algumas ferramentas de bioinformática, em razão dessa ciência ser indispensável na realização de pesquisas modernas biológicas, caracterizando os avanços da genética atual.

2. OBJETIVOS

2.1 Geral

- Fornecer um material didático que facilite a utilização de ferramentas básicas de bioinformática.

2.2 Específicos

- Verificar a atual situação do ensino e pesquisa de bioinformática no Brasil e na Universidade Federal do Piauí;
- Confeccionar um manual didático auxiliar no ensino de bioinformática;
- Ensinar o manuseio de ferramentas básicas de bioinformática a estudantes do nível superior de ensino;
- Demonstrar a importância da bioinformática para análise de dados biológicos.

3. REFERENCIAL TEÓRICO

Bioinformática é o campo da ciência em que a biologia, ciência da computação e tecnologia da informação se fundem para formar uma única disciplina. O objetivo final do campo é permitir a descoberta de novos conhecimentos biológicos, bem como para criar uma perspectiva global de que princípios unificadores da biologia podem ser discernidos (NCBI, 2006).

Conjuntos de dados sempre crescentes, novas fronteiras na ciência, o desenvolvimento de novas abordagens para análise e apresentação de dados, e a investigação de novas e complexas perguntas são as forças motrizes da bioinformática (DEGRAVE *et al.*, 2007).

Neste século, verifica-se a emergência da era genômica e pós genômica, em que se destaca a Bioinformática como uma ferramenta emergente e essencial. A partir dessas ferramentas disponíveis online, é possível obter, armazenar, classificar e interpretar os volumosos dados gerados por sequenciamento do DNA e proteínas, caracterizando as ômicas, como a genômica e a proteômica (RIBEIRO *et al.*, 2012).

Qualquer pessoa que faça pesquisa envolvendo o estudo de algum sistema de interesse biológico, usando ferramentas computacionais. Essa é a definição dada pelo docente do Instituto de Física de São Carlos (IFSC/USP), Leandro Martínez, ao bioinformata, profissional e pesquisador que, nos últimos anos, tem ocupado papel relevante na ciência, no Brasil e no mundo (MARTÍNEZ, 2012).

O bioinformata utiliza a computação como meio de resolver problemas biológicos (SOARES, 2006). Assim, conforme Arbex *et al.* (2006), esse profissional precisa ter conhecimento sobre:

- Biologia molecular, para, no mínimo, entender o problema;
- Matemática e probabilidade, de forma que fosse capaz de estabelecer um modelo de representação do problema;
- Computação, possibilitando a implementação da solução correta e apropriada, sendo esta inicialmente conhecida ou não;
- Estatística, para que as soluções fossem comprovadamente confiáveis.

É provável que somente após a dramática multiplicação, em tamanho e em número, dos bancos de dados de sequências resultantes dos projetos genoma, que a bioinformática tenha sido amplamente reconhecida como uma disciplina chave

para a nova era da genética, da biotecnologia e da biologia de sistemas (DEGRAVE et al., 2007).

Gerhardt (2001) considera que Neshich, um físico-químico de origem sérvia é um pioneiro da bioinformática nacional. Foi ele o idealizador da BrazilianBioNet (BBNet), uma rede de usuários de bioinformática, formada em 1992, que permitiu o primeiro acesso dos cientistas brasileiros aos programas de análise de sequências de DNA de forma gratuita, por intermédio de um computador (servidor) da Embrapa Recursos Genéticos e Biotecnologia.

Levy (2002) divulga a notícia, através do jornal da Unicamp, que a bioinformática brasileira já conta com uma poderosa ferramenta para análise da estrutura e das funções das proteínas codificadas pelos genes na formação dos seres vivos. Trata-se do Núcleo de Bioinformática (NBI), um sofisticado laboratório que, por meio de simulações em computadores, permitirá realizar os chamados estudos de genoma estrutural e genoma funcional, que vão além do simples mapeamento dos genes.

Além disso, no Brasil existe uma sociedade científica a Associação Brasileira de Bioinformática e Biologia Computacional (AB³C) que consiste em promover a interação de especialistas de várias áreas relevantes, tais como: bioinformática, ciência da computação, biologia, medicina, agronomia, veterinária, ecologia, evolução, biologia de sistema, biologia molecular e celular, bioquímica, física, matemática, estatística, probabilidade, controle automático e processamento de sinais (AB³C, 2004).

A Bioinformática vem se sobressaindo no Brasil, que teve destaque no Congresso Internacional de Bioinformática promovido por o ISMB 2006 (The 14th Annual International Conference on Intelligent Systems for Molecular Biology). Conforme o presidente da comissão organizadora do ISMB 2006 e pesquisador da Embrapa Informática Agropecuária, Goran Neshich, o congresso foi realizado no Brasil devido, principalmente, ao fato de haver pesquisas em bioinformática em todo o país. "Quanto mais produzimos, mais atraímos a atenção estrangeira", ressalta. Para o pesquisador, comunidades científicas estrangeiras consideram dinâmicas as equipes que atuam em bioinformática no Brasil (EMBRAPA, 2006).

Sobre o que envolve bioinformática, no Brasil, já existem grupos fortemente dedicados a estudos, e que fazem o desenvolvimento de softwares. E os maiores desafios, sem dúvidas, consistem em descobrir a cura de graves doenças através do

desenvolvimento de tais modelos e programas. Há também empresas que trabalham em parcerias com universidades (caso do IFSC, inclusive) para o estudo do melhoramento de enzimas para aproveitamento industrial, como, por exemplo, aquelas que compõem o bagaço da cana-de-açúcar, para produção de etanol. Há, também no país, empresas que fazem estudo de genomas de patógenos (pragas agrícolas), conta MARTÍNEZ (2012).

Esta ciência possui várias ferramentas como marcadores moleculares, que são úteis nos testes de paternidade, cultura de tecidos, que é útil na preservação da biodiversidade e engenharia genética, e no desenvolvimento de variedades para produção de alimentos mais saudáveis e nutritivos. A Biotecnologia integra um conjunto de tecnologias na qual a Bioinformática está incluída, sendo esta a ligação mais forte entre a Biotecnologia e a Computação, principalmente no âmbito do desenvolvimento de ferramentas computacionais que suportem ao trabalho cooperativo (BONGIOLO, 2006).

Não só no Brasil, mas no mundo, a bioinformática é carreira nova, que tem crescido muito após a automatização das técnicas de sequenciamento de DNA (MARTÍNEZ, 2012).

A informática é um meio educacional que vem ganhando cada vez mais importância no cenário mundial, com as inovações tecnológicas acontecendo a todo instante nos sistemas de informação. A escola se vê no momento de se adequar a essas mudanças e de assumir e incorporá-la ao contexto da educação (SILVEIRA e BARONE, 1999).

A importância do computador e as mudanças que ele provoca no processo de ensino e aprendizagem são hoje conhecidas, contudo, mesmo havendo o reconhecimento da necessidade do uso do computador e de programas computacionais como ferramentas educacionais, estes instrumentos são pouco familiares à maioria dos alunos e educadores brasileiros (ARAÚJO e GOULART, 2006).

Normalmente, as aulas de Biologia são carregadas de conteúdos e fatos correlacionados ao cotidiano, e muitas vezes possuem características que requerem capacidade de abstração, exigindo dos alunos que não só decorem, mas que compreendam. Por meio de imagens e simulações conseguidas com recursos de informática, o aluno poderá compreender muitos fenômenos biológicos, ampliando assim as possibilidades de aprendizagem (OLIVEIRA e RAZERA, 2007).

Por isso, é importante adequar a metodologias que possam prender a atenção e despertar um maior interesse pelos alunos, como o uso das Tecnologias de Informação e da Comunicação (TIC's), já que essas ferramentas passaram a fazer parte do cotidiano do aluno que nasceu em uma era tecnológica (OLIVEIRA, 2008).

4. MATERIAL E MÉTODOS

Esse trabalho foi desenvolvido entre os anos de 2012 e 2013 na Universidade Federal do Piauí (UFPI), nesse estudo foi realizado um levantamento de dados relacionado à bioinformática em nível de ensino e pesquisa, também foi utilizado o material didático desenvolvido para auxiliar os alunos durante palestras que foram ministradas.

4.1 Público Alvo

Este estudo tem como alvo alunos e professores em geral que queiram conhecer e aprender ou aprimorar seus conhecimentos sobre bioinformática.

4.2 Levantamento Bibliográfico

A coleta de dados iniciou-se primeiramente com o levantamento de dados bibliográficos sobre bioinformática por meio de buscas a internet, a sites, portais, artigos, revistas, teses, manuais, apostilas que tratavam sobre o assunto.

4.3 Levantamento de Dados dos Cursos de Pós-Graduação em Bioinformática

Para a obtenção dos dados sobre a situação de ensino de pós-graduação em bioinformática no Brasil foi realizado uma consulta a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por meio do site <<http://www.capes.gov.br/>>, esse portal registra o quadro atual de todos os cursos reconhecidos de mestrado e doutorado existentes na federação brasileira.

Em relação ao nível estadual foi feito uma busca da estrutura curricular dos cursos de pós-graduação promovidos pela UFPI no qual possuísse áreas afins com bioinformática, pois a UFPI não oferece cursos de pós-graduação em bioinformática. Os cursos analisados foram o de Mestrado em Biotecnologia os dados foram obtidos pelo site <www.ufpi.br/ppgbiotec>. E o outro curso analisado, foi o Doutorado da Rede Nordeste de Biotecnologia – RENORBIO seus dados foram consultados através do site <www.posgraduacao.ufpi.br//renorbio>.

4.4 Levantamento de Dados da Disciplina Bioinformática Ofertada na Graduação

A coleta de dados abrangeu a matriz curricular dos cursos de Ciências Biológicas, Ciências da Computação e algumas áreas de Ciências da Saúde (Biomedicina, Farmácia, Enfermagem, Odontologia, Medicina Veterinária) de todos os *Campi* da UFPI, perfazendo um total de cinco unidades, que são o *Campus* Ministro Petrônio Portella em Teresina, *Campus* Professora Cinobelina Elvas em Bom Jesus, *Campus* Ministro Reis Veloso em Parnaíba, *Campus* Senador Helvídio Nunes de Barros em Picos e o *Campus* Doutor Amílcar Ferreira Sobral na cidade de Floriano. Cada *Campus* dessa instituição de ensino promove cursos de diferentes áreas, então foram levados em consideração os cursos que eram promovidos por cada *Campus* e as suas respectivas grades curriculares que se encontravam disponíveis no site no dia da consulta (21/11/2012). Todas as informações foram obtidas por meio do endereço eletrônico da instituição <www.ufpi.br>.

A Universidade Aberta do Piauí – UAPI, outra instituição de ensino, também foi consultada as suas estruturas curriculares disponíveis no site <www.uapi.edu.br> no dia 21/11/2012 com a análise curricular foi verificada os cursos de Ciências Biológicas e Sistema de Informação, pois esses eram os cursos disponíveis que tinham mais afinidade com o tema bioinformática.

4.5 Levantamento de Dados dos Projetos Pesquisa Cadastrados nos *Campi* da Universidade Federal do Piauí

Os projetos científicos cadastrados dos *Campi* da UFPI são fornecidos pela Pró-Reitoria de Pesquisa e Pós-Graduação (PRPPG) e pela Coordenadoria Geral de Pesquisa (CGP), os dados foram obtidos por meio do acesso ao endereço eletrônico <ufpi.br/cgp/> no qual disponibilizou, no dia do acesso (20 e 21 de novembro de 2012), um total de mil quatrocentos e setenta projetos cadastrados do período de 2006 a 2012. A correlação desse total de projetos com o uso da bioinformática no desenvolvimento da pesquisa foi feita através de uma análise subjetiva promovida pelos títulos de cada projeto.

4.6 Confeção do Material Didático

A busca de uma abordagem facilitadora dos temas biológicos, em especial dos vinculados a biologia molecular, conduziu-nos a uma metodologia embasada na utilização de recursos tecnológicos básicos e todos disponíveis na Internet de acesso gratuito, e a utilização prioritária de alguns instrumentos que possibilitem a visualização esquemática ou fictícia do fenômeno biológico estudado, para melhor elucidação do mesmo (GOUVEA; LEAL, 2001).

A elaboração do material didático seguiu as seguintes etapas:

- I) Pesquisas bibliográficas em artigos, apostilas, acesso a sites, portais relacionados ao contexto do material didático;
- II) Treinamento para manipular as ferramentas básicas de bioinformática e conceitos específicos referente ao tema;
- III) Escolha dos conteúdos para conter no material didático;
- IV) Construção do manual didático;
- V) Realização das palestras de curta duração, com o título “Manual Didático – Ferramentas Básicas de Bioinformática” que foram realizadas no *Campus* Ministro Petrônio Portella (CMPP) em Teresina – PI com a participação dos acadêmicos do curso de Ciências Biológicas e do Curso de Nutrição (Foto 01 e Foto 02);
- VI) A metodologia das palestras consistiu em conteúdos teórico e prático, onde foi apresentado uma breve introdução teórica da atividade que seria proposta e em seguida, foi realizado as práticas pelos participantes através da utilização do manual didático que continha os passos a ser seguido para a execução da atividade prática sugerida;
- VII) Tais atividades foram desenvolvidas nas dependências do Laboratório de Biologia e na sala de aula da UFPI do *Campus* Ministro Petrônio Portella (CMPP) em Teresina – PI, totalizando uma carga horária de oito horas, sendo que foram desenvolvidas duas palestras, cada uma com duração de quatro horas com duração de dois dias;
- VIII) Avaliação - Ao final das palestras os alunos foram incentivado a responder perguntas quantitativa sobre o desempenho da apresentação das palestras, a utilidade do manual didático, a linguagem utilizada neste, se o material continha todos os passos necessários e ainda manifestarem suas opiniões por meio de

perguntas subjetivas em relação aos pontos positivos, negativos, críticas e sugestões sobre o manual didático.



Foto 01 - Turma de Ciências Biológicas (CMPP) - Laboratório de Biologia. Fonte: Arquivo Pessoal, 2013.



Foto 02 - Turma de Nutrição (CMPP) – Sala de Aula. Fonte: Arquivo Pessoal, 2013.

5. RESULTADOS E DISCUSSÕES

5.1 O Ensino de Bioinformática na Pós-Graduação

A bioinformática é uma área das Ciências Biológicas que está em rápido desenvolvimento, com o intuito de verificar a situação do ensino de Bioinformática no Brasil foi efetuada uma avaliação da atual situação dos cursos de pós – graduação no país. Realizada através da consulta de dados disponibilizada no portal da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES <www.capes.gov.br> no dia 20/11/2012. Neste portal são registrados todos os cursos cadastrados nacionalmente. Através desta busca de dados foram encontradas quatro cursos, em duas modalidades de ensino, Mestrado e Doutorado. Estes resultados podem ser visualizados no Quadro 01.

Quadro 01 - Cursos de Pós-Graduação em Bioinformática no Brasil

Nível de Ensino	Instituição de Ensino	Ano de Criação
Mestrado	Universidade de São Paulo – USP	2008
Mestrado	Universidade Federal do Paraná – UFPR	2009
Doutorado	Universidade Federal de Minas Gerais - UFMG	2002
Doutorado	Universidade de São Paulo – USP	2002

Fonte: Site da CAPES <www.capes.gov.br>, acessado em 20/11/2012.

Conforme demonstrado no quadro, verificou-se que existem apenas dois cursos de cada nível com dependências administrativas Federais e Estaduais que se encontram, atualmente, em situação de funcionamento. Com base nesses resultados fica explícita a carência de cursos de pós-graduação em bioinformática no país devido a pouca disponibilidade de cursos oferecidos pelas universidades, sendo necessário um maior investimento educacional nessa área, pois esta ciência possui um caráter essencial para o entendimento de sistemas biológicos mais complexos.

Há apenas três Universidades do Brasil que disponibilizam cursos de pós-graduação em bioinformática, a Universidade de São Paulo, a Universidade Federal do Paraná e a Universidade Federal de Minas Gerais. Portanto, a UFPI não oferece curso de pós – graduação em Bioinformática, com isso foi realizado outro levantamento de dados com o intuito de verificar os cursos de pós-graduação oferecidos pela UFPI que possuísse alguma correlação com a bioinformática e tivessem a disciplina de bioinformática na sua estrutura curricular. Essa busca de informação resultou em dois cursos de pós-graduação disponibilizados na UFPI que foram o Mestrado em Biotecnologia e o Doutorado da Rede Nordeste de Biotecnologia.

Os dados obtidos do Mestrado de Biotecnologia pela consulta no site <www.ufpi.br/ppgbiotec> mostrou que este mestrado disponibiliza uma disciplina de Bioinformática obrigatória com 30h/aulas. E o doutorado da Rede Nordeste de Biotecnologia – RENORBIO de acordo com o site <www.posgraduacao.ufpi.br//renorbio> disponibiliza uma disciplina de caráter obrigatório (Bioinformática) e ainda oferece duas disciplinas do tipo eletivas (Bioinformática Aplicada a Saúde e a Bioinformática na Evolução Molecular). Concluindo que ainda é pouca a quantidade de disciplinas ofertadas nos cursos de pós-graduação no ensino da UFPI, apesar do esforço ainda é necessário um aumento de disciplinas na sua grade curricular voltadas para a bioinformática que hoje é necessária para analisar dados da biologia molecular. Ainda estar no começo, mas sem dúvida as instituições de ensino precisam urgentemente se mobilizar para formar profissionais decisivos para a “medicina” futura.

5.2 A Bioinformática na Graduação da Universidade Federal do Piauí

Como mostrado anteriormente, a oferta de cursos habilitados em bioinformática no Brasil é pouco, favorecendo assim há uma escassez de profissionais capacitados para ensinar como utiliza-la. Foi realizada uma busca de dados no site da UFPI <www.ufpi.br> com a finalidade de consultar as estruturas curriculares dos cursos de graduação da UFPI que pudessem ofertar a disciplina bioinformática. As grades curriculares analisadas foram dos cursos de Ciências Biológicas e da Saúde, além do curso Ciência da Computação, estes foram escolhidos por terem maior afinidade com a área de bioinformática. Apenas algumas

grades curriculares se encontravam disponíveis no site da UFPI até a data da realização da pesquisa (21/11/2012). O Quadro 02 mostra a relação dos *Campi* e seus respectivos cursos que foram analisados.

Quadro 02 - Grades Curriculares dos Cursos e *Campi* da Universidade Federal do Piauí Pesquisados

Universidade Federal do Piauí	Curso de Graduação Analisados
<i>Campus</i> Ministro Petrônio Portella – Teresina	Ciências da Computação; Ciências Biológicas; Farmácia; Odontologia; Enfermagem.
<i>Campus</i> Professora Cinobelina Elvas – Bom Jesus	Medicina Veterinária; Engenharia Agrônômica; Ciências Biológicas.
<i>Campus</i> Ministro Reis Veloso – Parnaíba	Biomedicina; Engenharia Da Pesca; Ciências Biológicas.
<i>Campus</i> Senador Helvídio Nunes de Barros – Picos	Não foram encontradas as grades curriculares
<i>Campus</i> Doutor Amílcar Ferreira Sobral – Floriano	Não foram encontradas as grades curriculares

Fonte: Site UFPI <www.ufpi.br>, acessado em 21/11/2012.

O resultado obtido dessa análise foi que o curso de Biomedicina do *Campus* Ministro Reis Velloso – CMRV na cidade de Parnaíba, que é ofertada a disciplina de “Bioinformática” com 60h/aulas no sexto período na modalidade obrigatória em sua estrutura curricular. E no *Campus* Ministro Petrônio Portella – CMPP em Teresina possui o curso de Licenciatura Plena em Ciências Biológicas com a disciplina optativa “Bioinformática” de 60h/aulas e o curso Bacharelado em Ciência da Computação com uma disciplina de caráter optativo “Tópicos em Bioinformática” de 60h/aulas”.

Além da UFPI foi analisada, também, outra instituição de ensino a Universidade Aberta do Piauí – UAPI, através de uma consulta no site

<www.uapi.edu.br> no dia 21/11/2012. A grade curricular observada foram a dos cursos de Sistema de Informação e Licenciatura Plena em Ciências Biológicas, a pesquisa revelou que somente o curso de Ciências Biológicas continha uma disciplina relacionada diretamente com o campo de bioinformática, e esta graduação disponibiliza uma disciplina “Bioinformática” de 60h/aulas de caráter optativo e/ou eletiva.

Todo esse levantamento de dados dessas duas Universidades que compõe o Estado do Piauí mostra que o curso de graduação disponibilizado tanto pelos *Campi* da UFPI quanto pela a UAPI consiste em um déficit de ofertas de disciplinas durante a graduação. Surge a necessidade de atualização dos conteúdos das disciplinas científicas adequando-as aos currículos e programas atualizados para educar as gerações do século XXI (ABREU, 2008). Hoje em dia os cientistas dispõem de uma série de ferramentas, sabemos que a área de Biologia Molecular e Biotecnologia, assim como a Ciência de uma forma geral, encontra-se em constante avanço, fazendo com que educadores da área de Ciências se preocupem em aproximar a Ciência e a Tecnologia, presentes no cotidiano dos alunos, do universo de sala de aula (CARVALHO & BOSSOLAN, 2009).

5.3 A Utilização da Bioinformática nas Pesquisas Realizadas na Universidade Federal do Piauí

A bioinformática possui um caráter multidisciplinar, ela utiliza várias ferramentas que podem ser aplicadas no desenvolvimento de pesquisas das diferentes áreas de conhecimento. Adicionalmente foi avaliada como a bioinformática e suas ferramentas têm sido utilizadas para o desenvolvimento de pesquisa nos *Campi* da UFPI.

A realização desse estudo foi feito através de uma busca de dados no na Pró-Reitora de Pesquisa e Pós Graduação (PRPPG) e na Coordenação Geral de Pesquisa (CGP) por meio do site <ufpi.br/cgp/>. Nessa consulta foi encontrado um total de mil quatrocentos e setenta projetos cadastrados que estavam disponíveis no site durante o período de 2006 a 2012, esses dados foram obtidos entre os dias 20 e 21 de novembro de 2012 e todos os projetos dos diferentes cursos e áreas dos *Campi* que se encontrados listados foram analisados. Desse total foram escolhidos quarenta e dois projetos por apresentar uma provável correlação com a área de

bioinformática para desenvolverem sua pesquisa científica, levando em conta que essa análise possui um caráter subjetivo, por se basear apenas nos títulos do projeto, em que estes sinalizavam de alguma forma a utilização dessa ciência para desenvolverem seus trabalhos.

Dentre esses quarenta e dois projetos, os resultados apresentaram que poucos projetos desenvolvidos nos *Campi* da UFPI tem relação com a área da bioinformática, apenas dois projetos especificavam em seus títulos, que realmente utilizavam a bioinformática para a obtenção dos resultados. Um desses projetos pertence ao *Campus* Ministro Petrônio Portella – Teresina referente ao ano de 2010 e o outro projeto é do *Campus* Senador Helvídio Nunes de Barros – Picos do ano de 2012. Embora, provavelmente outros projetos pesquisas pudessem utilizar a bioinformática.

Esses dados relatados mostram que essa área é pouco utilizada em pesquisas realizadas pelos *Campi* Universitários Federais do Piauí, a bioinformática é uma área das ciências biológicas que está em rápido crescimento no campo científico e está sendo desenvolvida para atender à necessidade de manipular-se com grandes quantidades de dados genéticos e bioquímicos (QUEIROZ, 2002).

5.4 Material Didático

Esse item aborda a criação e avaliação do manual didático auxiliar teórico-prático nas palestras sobre “Ferramentas Básicas de Bioinformática”.

5.4.1 Construção do Manual Didático – Ferramentas Básicas de Bioinformática

O manual didático foi escrito no período de janeiro a fevereiro do ano de 2013, onde este contém conteúdos teórico e prático, além da utilização de figuras para melhor ilustrar os passos que seriam seguidos. Segue em apêndice uma cópia do Manual Didático – Ferramentas Básicas de Bioinformática.

Antes de iniciar os capítulos sempre era feito uma abordagem teórica do assunto e em seguida, os discentes realizava a prática proposta utilizando como roteiro o manual didático.

Adiante serão descritos os conteúdos que foram escolhidos para compor o material didático.

- Capítulo 1: Noções Básicas de Bioinformática

Este primeiro capítulo versa sobre uma abordagem teórica (Foto 03) sucinta sobre o que é a Bioinformática, seu fator de fortalecimento e as áreas que esta abrange. A bioinformática usa o poder computacional para catalogar, organizar e estruturas estas informações em uma entidade compreensiva e extremamente importante para a Biologia (QUEIROZ, 2002). A bioinformática consiste principalmente na análise computacional de sequências de DNA, RNA e proteínas. Essa nova ciência é interdisciplinar e surgiu na última década devido a uma necessidade urgente pela utilização de ferramentas sofisticadas para a análise de um crescente número de dados que veio a ser produzido em biologia molecular (PROSDOCIMI, 2007).



Foto 03 - Abordagem teórica de noções básicas de bioinformática. Fonte: Arquivo Pessoal, 2013.

- Capítulo 2: Banco de Dados

Em consequência da grande quantidade de informações de sequências de nucleotídeos e de aminoácidos que são produzidas atualmente, principalmente em projetos Genoma, Transcriptoma e Proteoma, o uso dos bancos de dados vem assumindo uma importância crescente na bioinformática (PROSDOCIMI, 2003).

Em Bioinformática existem vários bancos de dados de domínio público que podem ser usados livremente por meio do acesso online. No final da explicação teórica foi proposta uma prática (Foto 04) de acesso a cinco diferentes bancos de dados mais utilizados em bioinformática.



Foto 04 - Alunos realizando a prática: Acesso a diferentes bancos de dados. Fonte: Arquivo Pessoal, 2013.

- Capítulo 3: Obtenção e Edição de Sequências Biológicas

Consiste no processo de obter sequências de DNA, RNA e proteínas nos Bancos de Dados e manusear uma ferramenta de Bioinformática, o programa BioEdit. Nesse programa é possível editar as sequências biológicas, obter a composição de aminoácidos existente nas sequências. A prática (Foto 05) realizada foi a obtenção da sequência nucleotídica de *Leishmania braziliensis* e a edição desta no BioEdit.



Foto 05 - Alunos na prática de edição de dados com o BioEdit. Fonte: Arquivo Pessoal, 2013.

- Capítulo 4: Alinhamento Global e Local

O alinhamento de sequências consiste no processo de comparar duas sequências (de nucleotídeos ou proteínas) de forma a se observar seu nível de identidade. Essa técnica de comparação de sequências é implementada segundo um conceito de desenvolvimento de programas conhecido como um algoritmo guloso e é um dos pilares de toda a bioinformática (PROSDOCIMI, 2007). A prática (Foto 06 e Foto 07) disponibilizada foi a realização de alinhamento global e local entre sequências de *Leishmania sp.* O programa utilizado para o alinhamento local de sequências foi o BLAST (Basic Local Alignment Search Tool) encontrado em <<http://www.ncbi.nlm.nih.gov/BLAST/>>. E o alinhamento global com o programa BioEdit e a ferramenta utilizada foi o ClustalW (Alinhamento de Sequências de Proteínas ou DNA).



Foto 06 - Alunos realizando a prática de alinhamento global. Fonte: Arquivo Pessoal, 2013.



Foto 07 - Alunos efetivando a prática de alinhamento local. Fonte: Arquivo Pessoal, 2013.

- Capítulo 5: Mapa de Restrição

Endonucleases de restrição, também chamadas enzimas de restrição, são proteínas bacterianas que cortam em fragmentos a longa molécula linear de DNA (DNA Goes to School, 2007). Os mapas de restrição é um desenho gráfico dos locais onde essas enzimas de restrição cortam o DNA em regiões específicas. A prática (Foto 08) aplicada foi à construção de um mapa de restrição.



Foto 08 - Montagem do mapa de restrição. Fonte: Arquivo Pessoal, 2013.

5.4.2 Avaliação do Material Didático Auxiliar Teórico-Prático

Para a avaliação do manual foi realizado palestras com os alunos de Ciências Biológicas e de Nutrição do *Campus* Ministro Petrônio Portella - Teresina matriculado na disciplina Genética Humana sendo que continha em torno de vinte alunos por turma onde os mesmos estavam supervisionados pelo Professor Dr. Daniel Barbosa Liarte. As palestras exerceram um cronograma de execução de quatro horas onde foram expostas teorias e atividades práticas para domínio e conhecimento das ferramentas básicas de bioinformática com auxílio do manual didático, internet e ferramentas computacionais.

No final os graduandos voluntários fizeram uma avaliação quantitativa da palestra e do material didático utilizado (segue em apêndice uma cópia do questionário), onde as notas obedeceram a escala de 0 (zero) a 10 (dez), como mostra a Tabela 01 a seguir.

Tabela 01 - Respectivas médias avaliativas da palestra e do Material Didático

	Quantidade de alunos voluntários	Média das notas de 0 a 10	Mediana das notas de 0 a 10
Em relação à apresentação da palestra.	24	8,62	8,75
Em relação à utilidade do material didático	24	9,16	9,5
Em relação à linguagem usada no material didático	24	9,16	9,0

Fonte: Arquivo Pessoal, 2013.

Ao final da palestra, os alunos também foram instigados a manifestarem opiniões subjetivas, ressaltando os pontos positivos, negativos, críticas e sugestões sobre o desempenho geral das atividades práticas apresentadas, conforme demonstram algumas opiniões representadas por letras de (A) a (U).

○ Em relação aos pontos positivos:

(A) *O Material de ótima ajuda - explicativo, principalmente para os principiantes.*

(B) *Linguagem simples de fácil entendimento, a utilização das imagens ajuda bastante.*

(C) *O material é pioneiro e se debruça sobre um campo pouco explorado.*

(D) *O material em si achei que é uma ferramenta muito construtiva, pois ele nos orienta passo a passo de como realizar pesquisas em banco de dados e como utilizar programas.*

(E) *O material aplicado auxilia o usuário na busca por dados no programa, principalmente para as pessoas sem um inglês básico.*

(F) *Ele é realmente útil na utilização dos programas, porque torna mais fácil a manipulação dos sites e do programa BioEdit para os alunos de graduação, que normalmente não tem acesso durante as aulas.*

○ *Em relação aos pontos negativos:*

(G) *Em alguns tópicos não continha todos os passos necessários para a realização da prática.*

(H) *Alguns erros no conteúdo, digitação e organização.*

(I) *Atualizar os dados de acordo com o site.*

(J) *Foi observado que o material continha algumas instruções parcialmente incorretas.*

(K) *Nenhuma.*

○ *Em relação as críticas:*

(L) *O trabalho estava de fácil compreensão, apesar da falta de alguns dados.*

(M) *Explicar os resultados, objetivos.*

(N) *O trabalho facilita muito o trabalho de quem pretende estar nesta área. É um bom recurso e uma boa idéia.*

(O) *O material está adequado para ser aplicado na disciplina de biocomputação.*

(P) *Não tenho.*

- Em relação as sugestões:

(Q) *Deveria ser exposto também o porque dos resultados e como utilizá-los na prática.*

(R) *Olhar as dicas e erros apontados pela turma.*

(S) *Aumentar o tamanho da letra e tirar cópias coloridas.*

(T) *Continuar desenvolvendo o material nesse âmbito da bioinformática, uma vez que essa área é carente de materiais para o início do estudo da bioinformática.*

(U) *Nenhuma.*

Pelo fato de ser a primeira edição do manual este continha inicialmente alguns erros embora, todos os erros apontados pelos alunos no dia da palestra posteriormente foram corrigidos.

Antes de iniciar a atividade prática proposta no manual, foi desenvolvida uma breve explicação oral de como utilizar as ferramentas de bioinformática na prática para realização de uma pesquisa, mas como perspectiva serão acrescentados esse conteúdo explanado no material didático.

Diante dos limites que se restringiu a aplicação do manual, este colaborou para uma compreensão mais ampla do alunado sobre a importância da bioinformática e suas ferramentas computacionais com a finalidade de desenvolver novos conhecimentos e pesquisas científicas que podem ser conduzido por essa área. Com essa era tecnológica, o manual didático de bioinformática se caracteriza essencial como um recurso pedagógico facilitando a aprendizagem do aluno e auxiliando o ensino do professor no cotidiano do contexto educacional.

6. CONSIDERAÇÕES FINAIS

Sabemos que provavelmente o número de disciplinas e projetos pesquisa que envolvem bioinformática deve estar subestimado porque essa área nova muitas vezes é utilizada sem ser mencionada, mas que nossos resultados mostram um panorama confiável da situação da bioinformática no Brasil e na UFPI.

A incorporação da bioinformática no contexto educacional ainda é pouco familiar relacionado à disponibilidade de cursos e conteúdo curricular no processo de aprendizagem nos níveis de ensino de pós-graduação e de graduação. Portanto, visando uma contribuição para minimização desta problemática, este estudo dedica-se a uma proposta dinâmica e interativa para o ensino dos conteúdos de bioinformática, em especial desenvolvendo um recurso pedagógico (manual didático) para auxiliar o processo de ensino-aprendizagem.

A forma didática básica de organizar o processo de ensino não pode ser mais voltada para aquele ensino tradicional onde o professor é o centro do saber que repassa o conhecimento aos alunos e estes decodificam.

O professor sempre deve utilizar em sua prática docente material educativo, pois, os alunos devem ser motivados a estabelecer relações, a compreender e participar dos avanços tecnológicos já que o mundo se caracteriza cada vez mais de ambientes informatizados. Uma das melhores maneiras para desenvolver o conhecimento e habilidades é relacionar a teoria e a prática que permitam aos alunos manipularem objetos que despertem a participação e o prazer de aprender.

A opinião comum dos alunos foi que o material didático é muito útil para o processo de aprendizagem, de fácil utilização, atrativo, sendo que a inclusão de figuras torna o aprendizado mais agradável. A hipótese inicial, que foi confirmada com o desenvolvimento da palestra era a de que os estudantes iam interagir com a máquina e com os programas, gerando um ambiente de aprendizagem e resultando assim em uma aula mais estimulante e de fácil compreensão.

Diante disso, o nosso trabalho pôde promover aos aprendizes de Biologia e de Nutrição uma orientação, sem muitas dificuldades, para aquisição de conhecimento sobre a bioinformática, estimulando uma possibilidade de manipulação de dados biológicos, utilizando essa ciência em prováveis pesquisas científicas desenvolvidas por estes graduandos, através dessa inserção no seu meio educacional.

REFERÊNCIAS BIBLIOGRÁFICAS

AB³C, **Associação Brasileira de Bioinformática e Biologia Computacional**. Incentivo a Compreensão da importância da Biologia Computacional no Brasil e na América Latina. 2004. Disponível em: <<http://www.ab3c.org/>>. Acessado em: 30 jan. 2013.

ABREU, R. G. **Tecnologia e ensino de ciências: recontextualização no “Novo Ensino Médio”**. Rio de Janeiro, 2008.

AMUI, Saulo França. **Do laboratório ao campo virtual: Desenvolvimento de um banco de dados de venenos de serpentes Brasileiras e análise computacional de estruturas primárias de fosfolipases A2**. Dissertação de Mestrado, apresentado à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/ USP - Área de concentração: Toxicologia, 2006.

ARAÚJO, K. T. e GOULART, E, E. **Um estudo sobre a utilização das TIC's na região de Santo André**. Revista Iberoamericana de Educação, n. 38/3; Santo André, 25 mar. 2006.

ARBEX, W; COSTA, V. M. M. S; SILVA, M. V. G. B. **A Bioinformática como Ferramenta nas Pesquisas Atuais**. Viçosa, agosto de 2006.

BONGIOLO, Elisângela. **Análise Panorâmica da Bioinformática no Brasil: Propostas da Gestão de Pessoas para os Laboratórios de Pesquisa**. Criciúma, dezembro de 2006.

CARVALHO, J. C. Q; BOSSOLAN, N. R. S. **Algumas concepções de alunos do ensino médio a respeito das proteínas**. Encontro Nacional de Pesquisa em Educação em Ciência / VII Enpec. Florianópolis, 2009.

DEGRAVE, Wim; MACHADO, C. J. S. **Bioinformática e Saúde: Uma Visão Geral**. RECIIS – R. Eletr. de Com. Inf. Inov. Saúde. Rio de Janeiro, v.1, n.2, Sup.1, p.Sup286-Sup287, jul.-dez. 2007.

DNA Goes to School, Inc. **Análise de Fragmentos de Restrição**. Apostila da Oficina Prática de Genética, Genoma e Biotecnologia. Segundo módulo, 2007. Disponível em: <<http://www.odnavaiaescola.com.br/modulo2.pdf>>. Acessado em 26 fev. 2013.

EMBRAPA. **Brasil se destaca em congresso internacional de bioinformática**. 2006. Disponível em: <http://www.cnptia.embrapa.br/files/agrinforma04_06.pdf>. Acessado em: 30 jan. 2013.

GERHARDT, Isabel. **"Ferrão" ajuda a achar função de proteína**. 2001. Disponível em: <<http://www.cbi.cnptia.embrapa.br/press/fsp.html>>. Acessado em 30 jan. 2013.

GOUVEA, G; LEAL, M. C. **Uma visão comparada do ensino de Ciência, Tecnologia e Sociedade na escola e em um Museu de Ciências.** Revista Ciência e Educação, v.7, n.1, p. 67-84, 2001.

LEVY, Clayton. **Instalado laboratório 'sem paredes' no campus. Embrapa usa bioinformática na produção de novos produtos e na descoberta da cura de doenças grave.** 2002. Disponível em: <http://www.unicamp.br/unicamp/unicamp_hoje/ju/outubro2002/unihoje_ju196pag9b.html>. Acessado em: 30 jan. 2013.

MARTÍNEZ, Leandro. **O mundo, o Brasil e a Bioinformática.** 2012. Disponível em: <http://www.ifsc.usp.br/index.php?option=com_content&view=article&id=1022:omundo-o-brasil-e-a-bioinformatica&catid=7:noticias&Itemid=224>. Acessado em 30 jan. 2013.

NCBI, **National Center for Biotechnology Information. Bioinformatics.** 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/About/primer/bioinformatics.html>>. Acesso em 30 jan. 2013.

OLIVEIRA, M. de F. A. **O uso das tecnologias da informação e da comunicação no ensino aprendizagem de Biologia.** Paraná, 2008.

OLIVEIRA, M. N. V e RAZERA, J. C.C. **A Informática educativa em escolas públicas no interior do Brasil: Argumentos e Perspectivas.** Travessias. 04ed. ISSN 1982-5935, 2007.

PROSDOCIMI, Francisco. **Curso online Introdução a Bioinformática.** Biotecnologia Ciência & Desenvolvimento, 2007. Disponível em: <http://www2.bioqmed.ufrj.br/prosdocimi/FProsdocimi07_CursoBioinfo.pdf>. Acessado em 26 fev. 2013.

PROSDOCIMI, F. *et al.* **Bioinformática: Manual do Usuário.** Revista de Biotecnologia Ciência e Desenvolvimento – n°29, 2003.

QUEIROZ, Alexandre. **Apostila de Introdução a Bioinformática.** Universidade Federal do Rio Grande do Norte, Centro de Biociências, Departamento de Biofísica e Farmacologia, Disciplina de Bioinformática, 2002.

RIBEIRO, Júnior. *et al.* **Bioinformática como recurso pedagógico para o curso de Ciências Biológicas na Universidade Estadual do Ceará – UECE – Fortaleza, Estado do Ceará.** Acta Scientiarum. Education. Maringá, v. 34, n. 1, p. 129-140. jan.- june, 2012.

SILVA, Luiz Hildebrando Pereira. **Ciências Biológicas e Biotecnologia: Realidades e Virtualidades.** São Paulo em perspectiva, 2000.

SILVEIRA, S. R. e BARONE, D. A. C. **Estudo e Construção de uma Ferramenta de Autoria Multimídia para a Elaboração de Jogos Educativos.** Dissertação de Mestrado em Ciências da Computação). Instituto de Informática, Universidade Federal do Rio Grande do Sul, Porto Alegre, 1999.

SOARES, Edson. **Profissão do futuro: Bioinformata vive entre bits e células.** 2006. Disponível em: <<http://www.mqservicos.com.br/secao.asp?i=14&c=801>>. Acessado em: 21 fev. 2013.

APÊNDICES

APÊNDICE A – MATERIAL DIDÁTICO



UNIVERSIDADE FEDERAL DO PIAUÍ
CAMPUS SENADOR HELVÍDIO NUNES DE BARROS

Manual Didático

Ferramentas Básicas de Bioinformática

Raylla Caroline de Sousa

Picos

2013

APRESENTAÇÃO

Esse manual didático concede instruções explícitas e precisas sobre a utilização de ferramentas básicas de bioinformática. Ele será utilizado como principal consulta e referência na realização de nossas atividades durante as Palestras.

Com o objetivo de proporcionar o contato com a bioinformática através do manuseio de algumas ferramentas básicas, para facilitar a compreensão, esse manual possui conteúdos teóricos e práticos, que estão demonstradas através de uma breve introdução e por figuras que se encontram todas enumeradas aparecendo abaixo de cada procedimento enfatizando todos os campos que serão analisados.

Espero que aproveitem bastante as palestras.

Abraço!

LISTA DE FIGURAS

Figura 01- Página inicial do NCBI.....	55
Figura 02- Visualização de outros bancos de dados incluso no portal.....	55
Figura 03- Página Inicial do banco de Dado GenBank.....	56
Figura 04 - Banco de Dados incluso no GenBank.....	56
Figura 05 - Página inicial do banco de dado GeneDB.....	57
Figura 06 - Uma das maneiras de acessar os organismos existentes no GeneDB...	57
Figura 07 - Outra maneira de acessar os organismos existentes no GeneDB.....	58
Figura 08 - Página inicial do portal ExPASy.....	58
Figura 09 - Acesso a outros bancos de dados inclusos no portal.....	59
Figura 10 - Pagina inicial do banco de dados Swiss-Prot.....	59
Figura 11- Página inicial do banco de dado PDB.....	60
Figura 12 - Página Inicial do site para baixar o programa BioEdit.....	61
Figura 13 - Acesso ao Banco de Dados NCBI.....	62
Figura 14 - Resultado de busca da sequência nucleotídica de <i>Leishmania braziliensis</i>	62
Figura 15 - Informações da sequência biológica <i>L.braziliensis</i>	63
Figura 16 - Salvando a sequência <i>L. braziliensis</i> no banco de dados.....	63
Figura 17 - Página Inicial do BioEdit, como abrir uma sequência.....	64
Figura 18 - Abrindo a sequência de <i>L. braziliensis</i> no formato FASTA.....	64
Figura 19 - Sequência de <i>L. braziliensis</i> aberta no BioEdit.....	65
Figura 20 - Copiando uma sequência.....	65
Figura 21- Colando uma sequência.....	66
Figura 22 - Sequência de DNA duplicada.....	66
Figura 23 - Procedimentos para transformar DNA em RNA.....	67
Figura 24 - Sequência editada de DNA para RNA.....	67
Figura 25 - Sequência de RNA duplicada.....	68
Figura 26 - Procedimento de reversão da sequência de RNA para DNA.....	68
Figura 27 - Sequência transformada em DNA.....	68
Figura 28 - Transformação de DNA em aminoácido.....	69
Figura 29 - Sequência de Aminoácido.....	69
Figura 30 - Procedimento para verificar a composição de aminoácidos.....	70
Figura 31 - Resultado da composição de aminoácido.....	70
Figura 32 - Procedimento para salvar um arquivo.....	71
Figura 33 - Salvando o arquivo no formato FASTA.....	71
Figura 34 - Busca a sequência nucleotídica de <i>L. amazonensis</i> no NCBI.....	73
Figura 35 - Resultados de sequências de <i>L. amazonensis</i> RNA helicase.....	73
Figura 36 - Informações da sequência de <i>L. amazonensis</i> RNA helicase.....	74
Figura 37 - Salvando a sequência de <i>L. amazonensis</i> RNA helicase.....	74
Figura 38 - Download da sequência de <i>L. amazonensis</i> RNA helicase.....	75

Figura 39 - Abrindo a sequência <i>L. amazonensis</i> no BioEdit.....	75
Figura 40 - Sequência de <i>L. amazonensis</i>	76
Figura 41- Funções necessárias para importar sequências.....	76
Figura 42 - Importando uma nova sequência.....	77
Figura 43 - Sequência de <i>L.braziliensis</i> importada.....	77
Figura 44 - Várias sequências em um único arquivo.....	77
Figura 45 - Funções para realizar o alinhamento global.....	78
Figura 46 - Janela do ClustalW Options.....	78
Figura 47 - Etapa para realizar o alinhamento global.....	79
Figura 48 - Processando o alinhamento global.....	79
Figura 49 - Alinhamento Global das sequências de <i>Leishmania sp.</i>	80
Figura 50 - Analisando o alinhamento global.....	80
Figura 51 - Opções para salvar o arquivo.....	81
Figura 52 - Salvando o arquivo no formato FASTA.....	81
Figura 53 - Página inicial do BLAST Assembled RefSeq Genomes.....	82
Figura 54 - Abrindo a sequência <i>Leishmania amazonensis.fasta</i> no BioEdit.....	83
Figura 55 - Mostrando o local do nome da sequência.....	83
Figura 56 - Nova janela mostrando toda a sequência.....	83
Figura 57 - Selecionando a sequência completa.....	84
Figura 58 - Mostrando as opções de BLAST.....	84
Figura 59 - Local de inserção da sequência.....	85
Figura 60 - Opções do BLAST.....	85
Figura 61- Mais opções para realizar o BLAST.....	86
Figura 62 - Processado o alinhamento local.....	86
Figura 63 - Resultado do Alinhamento local.....	87
Figura 64 - Gráfico do alinhamento local.....	87
Figura 65 - Lista descritiva do alinhamento local.....	88
Figura 66 - Sequências alinhadas.....	88
Figura 67 - Mostrando os Gap do alinhamento local.....	88
Figura 68 - Sequência <i>L. infantum</i> aberta no BioEdit.....	90
Figura 69 - Procedimento para um Mapa de Restrição.....	90
Figura 70 - Menus para realizar um Mapa de Restrição.....	91
Figura 71 - Funções do menu Create Restriction Map.....	91
Figura 72 - Outras funções do menu Create Restriction Map.....	92
Figura 73 - Mapa de Restrição da sequência <i>L.infantum</i>	92
Figura 74 - Resultado da tabela que mostra a frequência e posição do nucleotídeo.....	93
Figura 75 - Resultado da tabela que mostra as enzimas que cortam até cinco vezes.....	93
Figura 76 - Resultado da posição nucleotídica e as enzimas.....	94
Figura 77 - Resultados das enzimas de restrição que não cortam a sequência.....	94

LISTA DE ABREVIATURAS E SIGLAS

BLAST – Basic Local Alignment Search Tool.

ClustalW – Alinhamento de sequência de proteínas ou DNA.

DNA – Ácido Desoxirribonucleico.

FASTA – Scan a Protein or DNA Sequence Library for Similar Sequence.

L. sp – *Leishmania espécie*.

NCBI – National Center for Biotechnology Information.

RNA – Ácido Ribonucleico.

SUMÁRIO

1. NOÇÕES BÁSICAS DE BIOINFORMÁTICA.....	52
2. 1ª PRÁTICA: BANCO DE DADOS.....	54
2.1 Introdução.....	54
2.2 Acesso a Diferentes Bancos de Dados.....	55
3. 2ª PRÁTICA: OBTENÇÃO E EDIÇÃO DE SEQUÊNCIAS BIOLÓGICAS.....	61
3.1 Introdução.....	61
3.2 Obtenção e Edição de Sequências Biológicas.....	61
4. 3ª PRÁTICA: ALINHAMENTO GLOBAL E LOCAL.....	72
4.1 Introdução.....	72
4.2 Realização de um Alinhamento Global.....	72
4.3 Realização de um Alinhamento Local.....	82
5. 4ª PRÁTICA: MAPA DE RESTRIÇÃO.....	89
5.1 Introdução.....	89
5.2 Montagem de um Mapa de Restrição.....	90
REFERÊNCIAS BIBLIOGRÁFICAS.....	95

1. NOÇÕES BÁSICAS DE BIOINFORMÁTICA

Bioinformática é o campo da ciência em que a biologia, ciência da computação e tecnologia da informação se fundem para formar uma única disciplina. O objetivo final do campo é permitir a descoberta de novos conhecimentos biológicos, bem como para criar uma perspectiva global de que princípios unificadores da biologia podem ser discernidos (NCBI, 2006). O Bioinformata falando de forma resumida utiliza a computação como meio de resolver problemas biológicos (SOARES, 2006).

Desde os anos 1980, a Bioinformática tem adquirido uma importância crescente nas ciências biológicas. Devido à complexidade de ferramentas tanto de *hardware* quanto de *software*, e a necessidade de ter-se um gerenciamento especializado de infraestruturas de rede, processamento e armazenamento de dados, muitas instituições de pesquisa organizaram um núcleo de bioinformática (DEGRAVE *et al.*, 2007).

A bioinformática teve como fator para o seu fortalecimento o projeto Genoma, visto que a necessidade de um sistema de computação que armazenasse e analisasse as enormes quantidades de dados gerados por esse projeto se torna imprescindível e exige cada vez mais da área da computação (ARBEX, 2006). Essa nova ciência envolve diversas áreas do conhecimento, tais como a engenharia de softwares, a matemática, a estatística, a ciência da computação, biologia molecular, etc (PROSDOCIMI *et al.*, 2002).

A bioinformática é um campo das ciências biológicas que está em rápido crescimento e está sendo desenvolvida para atender à necessidade de manipular-se com grandes quantidades de dados genéticos e bioquímicos (QUEIROZ, 2002).

A Bioinformática representa, hoje, um dos grandes desafios para se tentar decifrar o genoma, pois, ao mesmo tempo em que é uma forma de se conseguir informações imediatas para os dados do genoma que vem sendo descobertos, também é a base para um sucesso científico futuro (LENGAUER, 2001 *apud* BONGIOLO, 2006).

Hoje em dia os cientistas dispõem de uma série de ferramentas, sabemos que a área de Biologia Molecular e Biotecnologia, assim como a Ciência de uma forma geral, encontra-se em constante avanço, fazendo com que educadores da área de

Ciências se preocupem em aproximar a Ciência e a Tecnologia, presentes no cotidiano dos alunos, do universo de sala de aula (CARVALHO & BOSSOLAN, 2009). A seguir serão apresentadas atividades práticas relacionadas com a área de bioinformática.

2. 1ª PRÁTICA: BANCO DE DADOS

2.1 Introdução

Um banco de dados pode ser considerado uma coleção de dados interrelacionados, projetado para suprir as necessidades de um grupo específico de aplicações e usuários. Um banco de dados organiza e estrutura as informações de modo a facilitar consultas, atualizações e deleções de dados (PROSDOCIMI *et al.*, 2003).

A construção de bancos de dados para armazenamento de informações de sequências de DNA e genomas inteiros, proteínas e suas estruturas tridimensionais, bem como vários outros produtos da era genômica, tem sido um grande desafio, mas simultaneamente extremamente importante (SANTOS & QUEIROGA, 2003). O investimento contínuo na construção de bancos de dados públicos é um dos grandes motivos do sucesso dos projetos genoma e, em especial, do Projeto Genoma Humano. Devido à magnitude do conjunto de dados produzidos torna-se fundamental a organização desses dados em bancos que permitam acesso on-line (PROSDOCIMI *et al.*, 2003).

Os bancos de dados de genoma representam hoje uma das principais ferramentas de suporte para os biólogos moleculares e geneticistas. É fundamental para a pesquisa nesta área realizar cadastros de sequências e de algumas anotações relacionadas, e realizar consultas nestes bancos a fim de levantar dados para análises biológicas. Entre estas análises é possível destacar a comparação de sequências e o descobrimento de novos genes, funções e características de uma nova sequência (SEIBEL *et al.*, 2000).

A finalidade dessa prática consiste em acessar e conhecer determinados bancos de dados que são mais utilizados em bioinformática, para isso é necessário está apenas conectado à internet. Todos os bancos de dados que serão apresentados estão disponíveis em domínio público e podem ser utilizados livremente.

2.2 Acesso a Diferentes Bancos de Dados

1º PASSO: Acesso ao portal NCBI

O NCBI é um Centro Nacional de Informação Biotecnológica, no qual fornece informações biomédicas e genéticas, e essa base de dados se encontra disponíveis na internet através do acesso ao portal que agrega vários outros bancos de dados. Iniciaremos esta prática com o acesso a página inicial do portal NCBI - National Center for Biotechnology Information, através do endereço eletrônico: <<http://www.ncbi.nlm.nih.gov/>>.

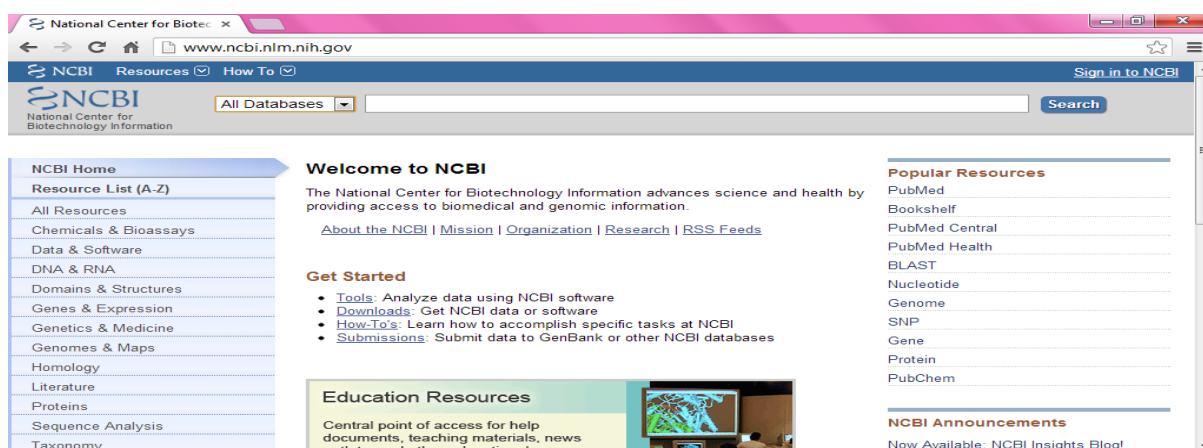


Figura 01. Página inicial do NCBI. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov/>>, acessado em 21/02/13.

Para visualizar a existência de outros bancos de dados que estão incluso nesse portal do NCBI, clique na opção “All Databases”. Dentro do portal NCBI encontra-se o banco de dados GenBank, que este será acessado a seguir.

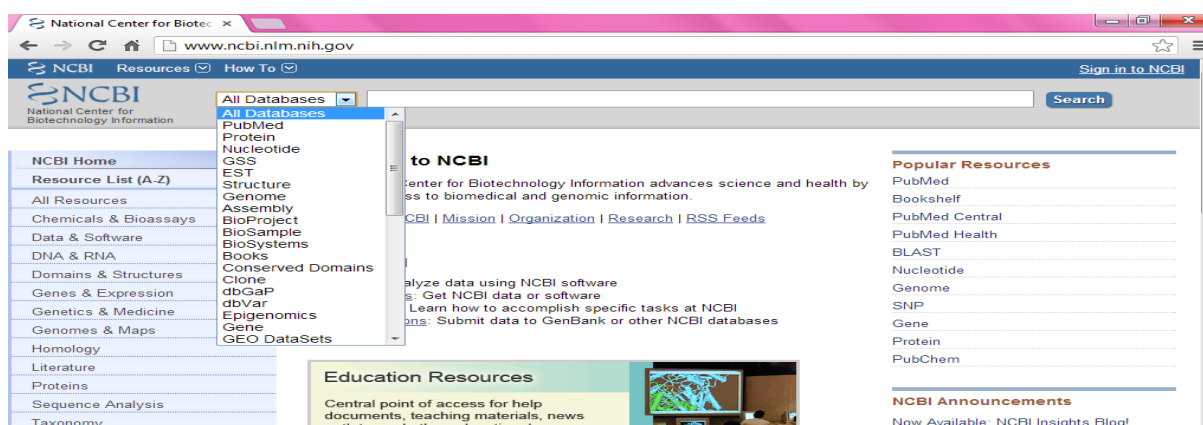


Figura 02. Visualização de outros bancos de dados incluso no portal NCBI. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov/>>, acessado em 21/02/13.

2º PASSO: Acesso ao Banco de Dados Genbank

O Genbank é outro tipo de banco de dados que armazena sequências de nucleotídeos e proteínas, além de informações biológicas relevantes sobre cada sequência, que são, por exemplo, o nome científico e a taxonomia do organismo de origem, um conjunto de anotações que especificam regiões codificantes na sequência (função, estrutura, etc.), e também outras regiões de relevância biológica (SEIBEL *et al.*, 2000).

Vamos realizar neste momento uma busca a página do banco de dados Genbank, por meio do endereço eletrônico: <<http://www.ncbi.nlm.nih.gov/genbank/>>.

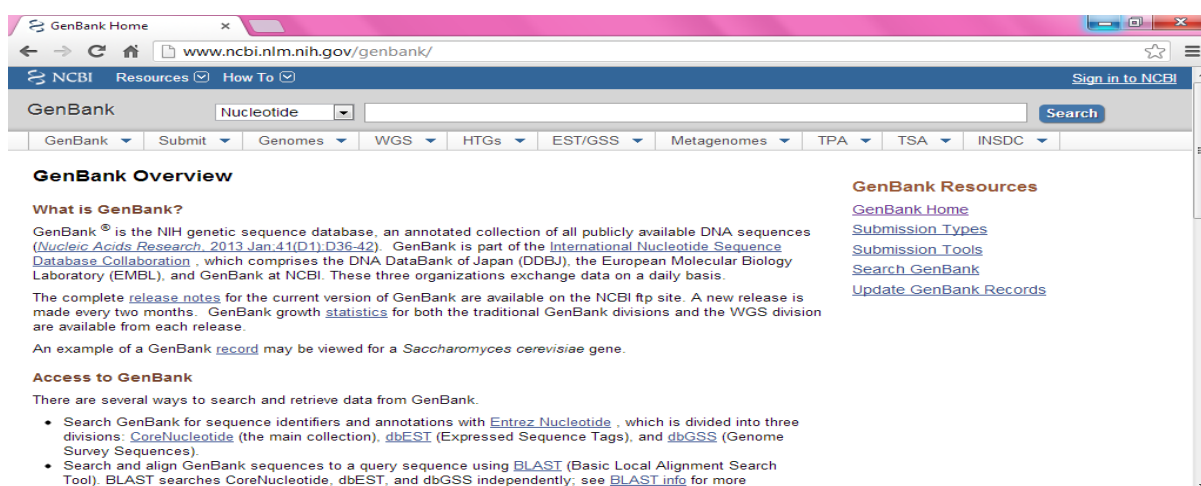


Figura 03. Página Inicial do banco de Dado GenBank. Fonte: Site GenBank <<http://www.ncbi.nlm.nih.gov/genbank/>>, acessado em 21/02/13.

Visualize que o banco de dados Genbank também possui outros bancos de dados inclusos.

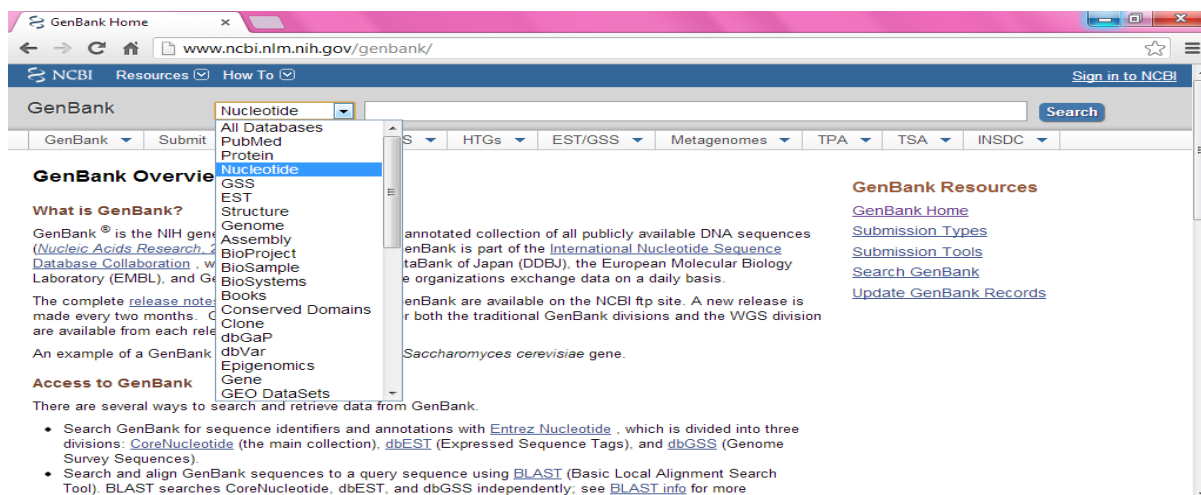


Figura 04. Banco de Dados incluso no GenBank. Fonte: Site GenBank <<http://www.ncbi.nlm.nih.gov/genbank/>>, acessado em 21/02/13.

3º PASSO: Acesso ao Banco de Dados GeneDB

O banco de dados GeneDB contém informações do genoma de organismos patogênicos. Vamos conhecê-lo, realizando uma consulta ao banco de dados o GeneDB, por meio do site <<http://www.genedb.org>>.

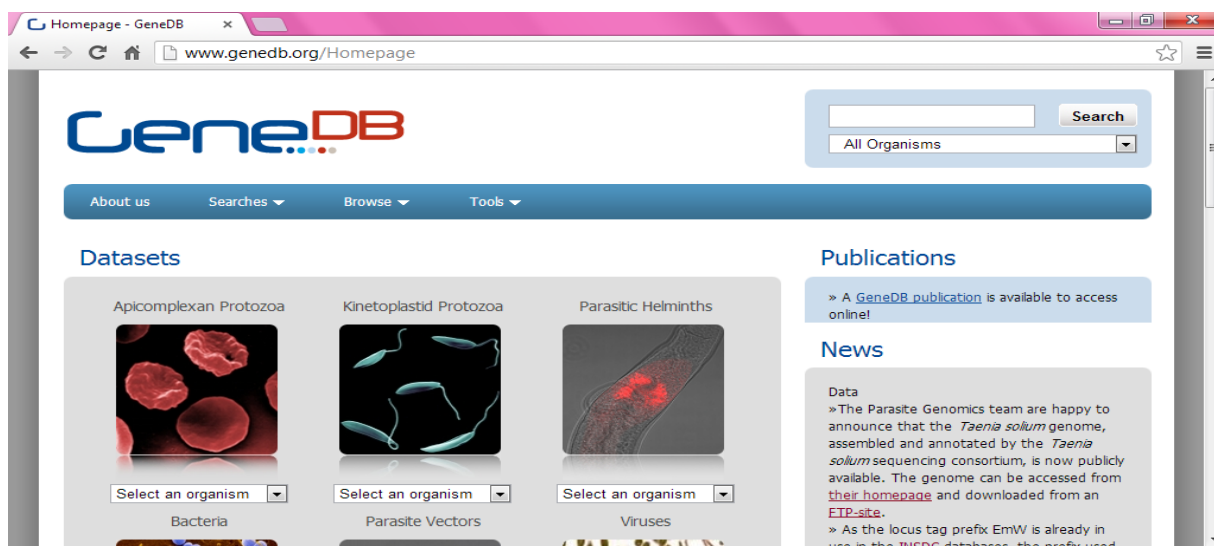


Figura 05. Página inicial do banco de dados Gene DB. Fonte: Site GeneDB <<http://www.genedb.org>>, acessado em 21/02/13.

Nesta página há duas maneiras de se obter a lista de todas as espécies biológicas existentes no banco de dados GeneDB.

A primeira maneira de busca é através da opção “All Organisms”, clique sobre ela e mostrará uma lista com todos os organismos existentes no GeneDB.

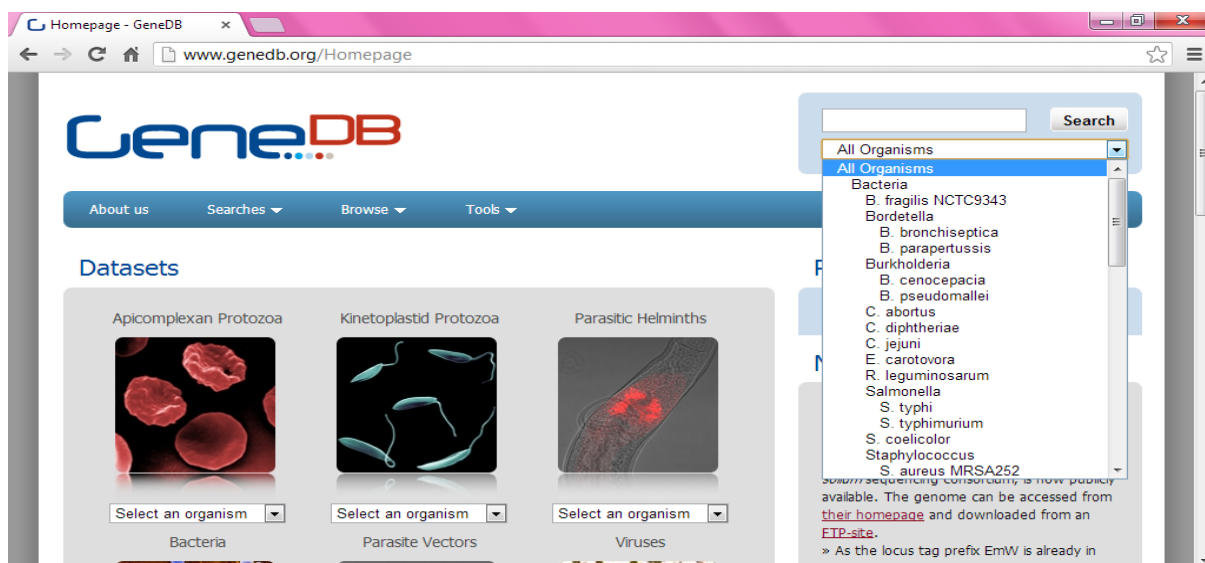


Figura 06. Uma das maneiras de acessar os organismos existentes no GeneDB. Fonte: Site GeneDB <<http://www.genedb.org>>, acessado em 21/02/13.

Ou então clique em “Select an organism”, que fica localizado abaixo das imagens, para buscar o organismo desejado através da sua categoria taxonômica.

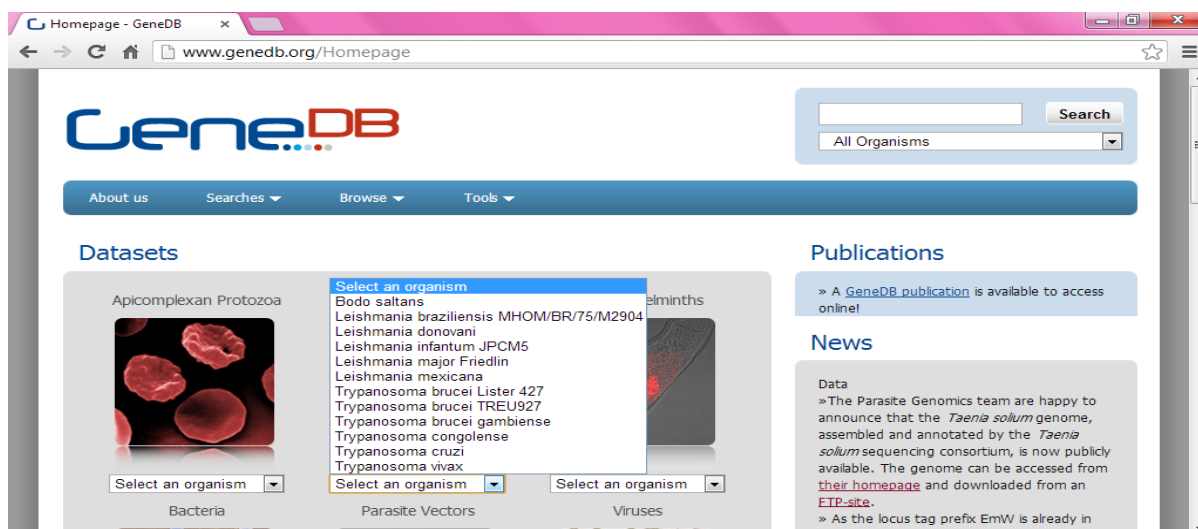


Figura 07. Outra maneira de acessar os organismos existentes no GeneDB . Fonte: Site GeneDB <<http://www.genedb.org>>, acessado em 21/02/13.

4º PASSO: Acesso ao Banco de Dados Swiss-Prot

O portal ExpASY disponibiliza o acesso ao banco de dados Swiss-Prot, que este banco armazena sequências de proteínas e suas respectivas características moleculares, anotado manualmente por uma equipe de especialistas (PROSDOCIMI *et al.*,2003).

Realize um acesso ao portal ExpASY com o seguinte endereço eletrônico: <<http://www.expasy.org>>.



Figura 08. Pagina inicial do portal ExpASY. Fonte: Site ExpASY <<http://www.expasy.org>>, acessado em 21/02/13.

O portal ExPASy também possui vários outros bancos de dados, inclusive o banco de dados Swiss-Prot. Clique na opção “Query all database” e verifique.

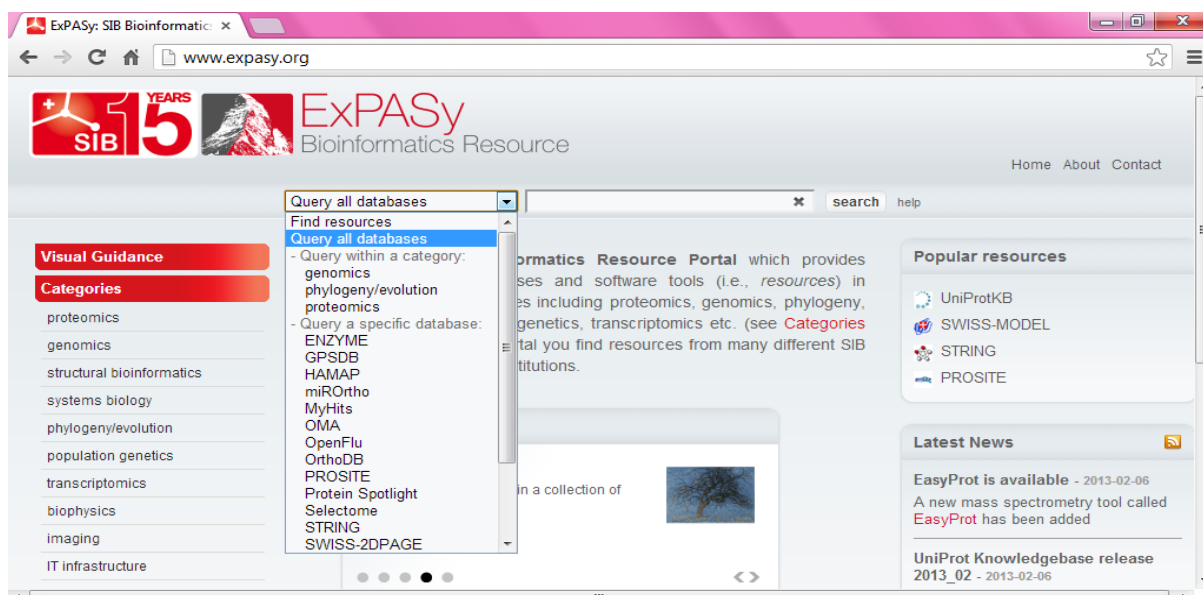


Figura 09. Acesso a outros bancos de dados inclusos no portal. Fonte: Site ExPASy <<http://www.expasy.org>>, acessado em 21/02/13.

Você deve ter notado que ao procurar o banco de dados Swiss-Prot, na lista ele não aparece dentre as opções, mas isso é devido ao fato que o banco de dados Swiss-Prot se uniu com outro banco de dados o UniProtKB. Então, você pode clicar na opção “UniProtKB” ou também pode acessar o banco de dados Swiss-Prot diretamente pelo site <http://web.expasy.org/docs/swiss-prot_guideline.html>.

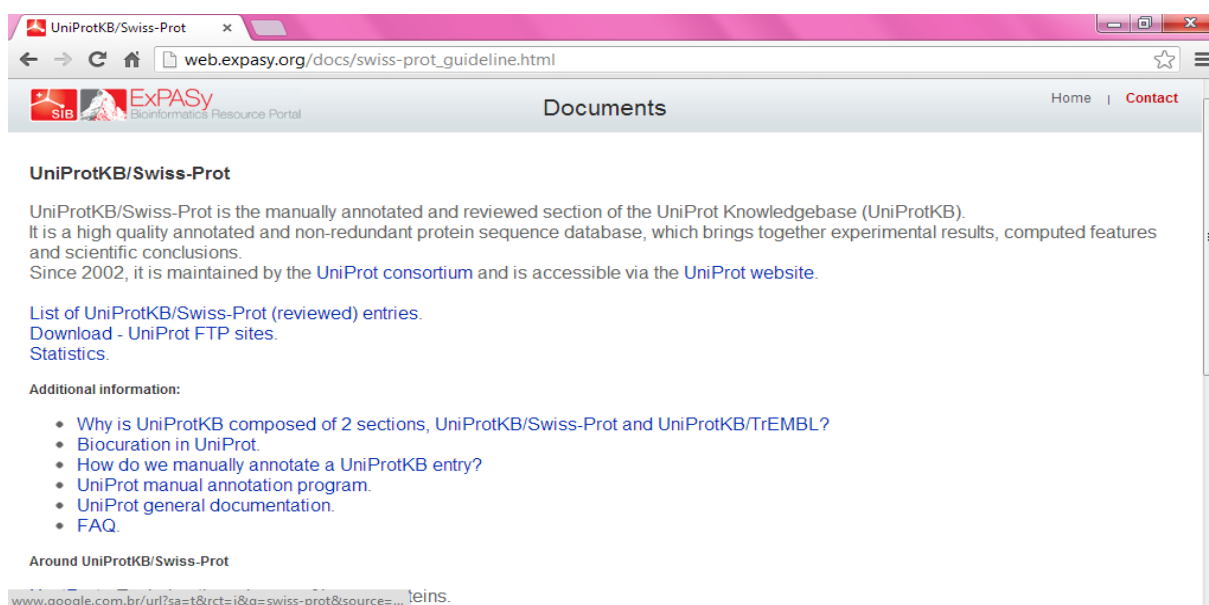
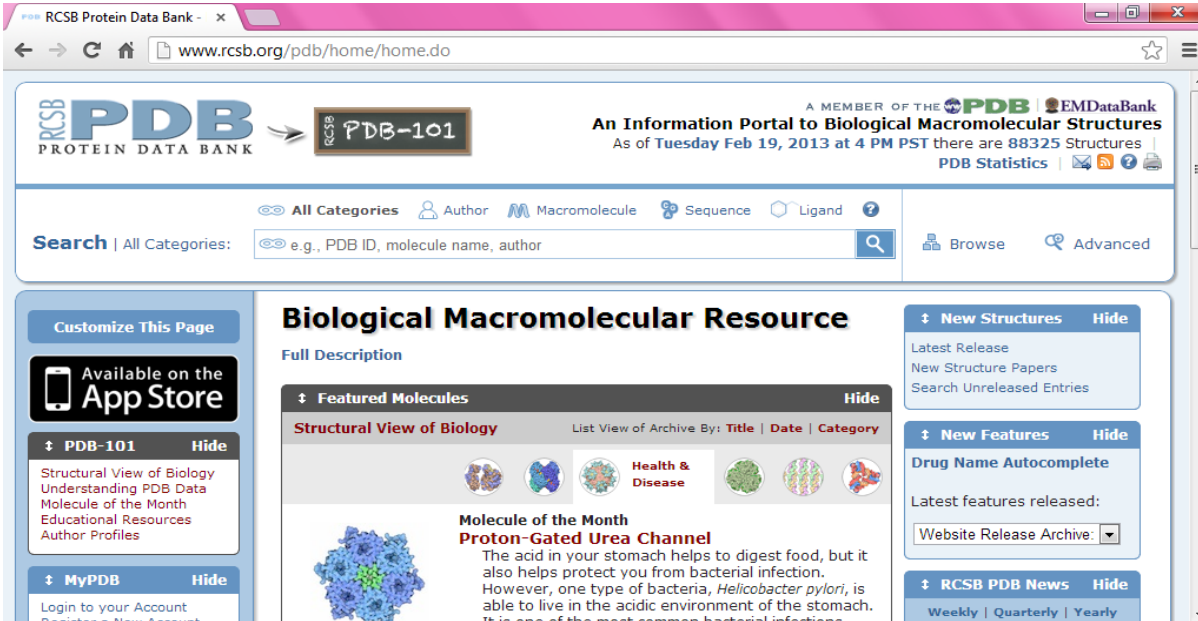


Figura 10. Página inicial do banco de dados Swiss-Prot. Fonte: Site Swiss-Prot <http://web.expasy.org/docs/swiss-prot_guideline.html>, acessado em 21/02/13.

5° PASSO: Acesso ao Banco de Dados PDB

O banco de dados PDB difere por conter estruturas tridimensionais de proteínas e ácidos nucleicos. Conheça esse banco de dados PDB acessando o endereço eletrônico: <<http://www.rcsb.org>>.



The screenshot shows the RCSB Protein Data Bank homepage. At the top, there is a navigation bar with the PDB logo and a search bar. Below the search bar, there are several sections: 'Customize This Page' with links to 'PDB-101' and 'MyPDB'; 'Biological Macromolecular Resource' with a 'Full Description' section; and a 'Molecule of the Month' section for 'Proton-Gated Urea Channel'. The page also features a 'New Structures' section and a 'New Features' section with a 'Drug Name Autocomplete' feature. The footer includes 'RCSB PDB News' with options for 'Weekly', 'Quarterly', and 'Yearly' updates.

Figura 11. Página inicial do banco de dado PDB. Fonte: Site PDB <<http://www.rcsb.org>>, acessado em 21/02/13.

3. 2ª PRÁTICA: OBTENÇÃO E EDIÇÃO DE SEQUÊNCIAS BIOLÓGICAS

3.1 Introdução

O objetivo dessa prática é aprender passo a passo de como buscar as sequências biológicas de DNA, RNA e Proteínas por meio do acesso on-line a um banco de dados e a utilização de outra ferramenta básica da bioinformática, o programa BioEdit que irá manipular as sequências biológicas obtidas. Caso você não tenha esse programa BioEdit o download pode ser obtido através do site: <<http://www.mbio.ncsu.edu/bioedit/page2.html>>.

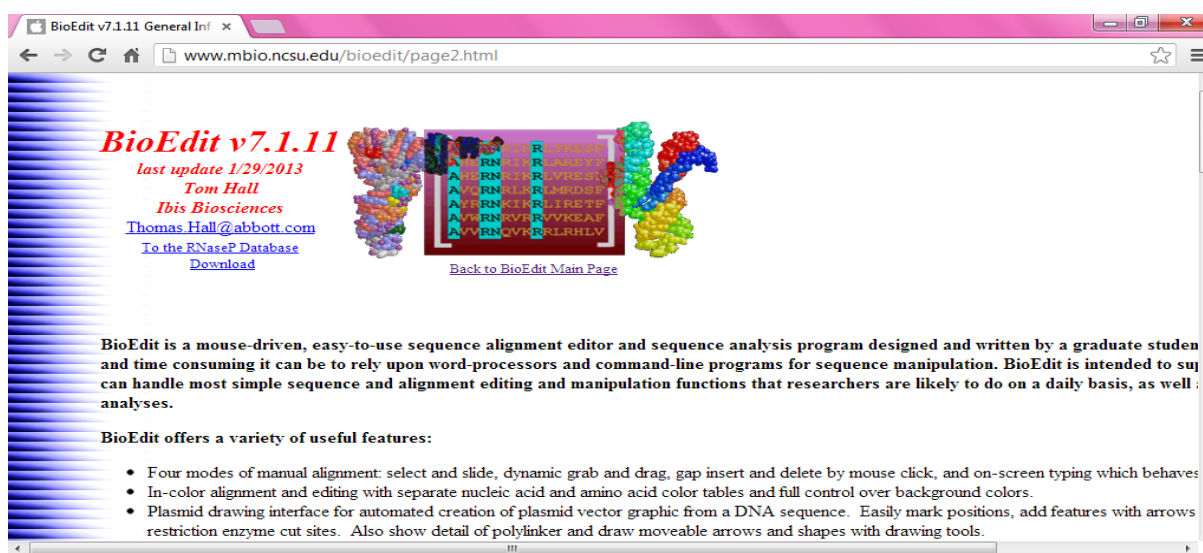


Figura 12. Página Inicial do site para baixar o programa BioEdit, Fonte: Site MBIO <<http://www.mbio.ncsu.edu/bioedit/page2.html>>, acessado em 24/02/13.

3.2 Obtenção e Edição de Sequências Biológicas

1º PASSO: Obtenção de Sequências Biológicas

Para a realização desta prática é necessário à obtenção de uma sequência biológica de um determinado organismo e para isso é preciso realizar uma consulta a um banco de dados. O organismo escolhido para ser utilizado nessa prática pertencerá ao Reino Protista, nesse caso temos duas opções de bancos de dados, que pode ser através do acesso ao banco de dados GeneDB ou NCBI, conforme ensinado na prática anterior, pois estes bancos armazenam dados biológicos dessa

categoria taxonômica. Vamos começar! O banco de dados escolhido será o NCBI, acesse-o por meio do endereço eletrônico <<http://www.ncbi.nlm.nih.gov/>>. Ao exibir a página inicial selecione a ferramenta de busca “Nucleotide” ao lado digite o nome do organismo que deseja obter a sequência, nesse caso utilizaremos como exemplo a espécie *Leishmania braziliensis* RNA helicase.



Figura 13. Acesso ao Banco de Dados NCBI. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov/>>, acessado em 22/02/13.

Ao clicar na opção “Search” (pesquisar) uma nova página surgirá contendo os resultados de busca de sequências nucleotídicas de *Leishmania braziliensis* RNA helicase existentes no portal NCBI, a opção “*Leishmania braziliensis* strain ET-ATP dependent RNA helicase gene, partial cds” será utilizada para dar continuidade à prática.

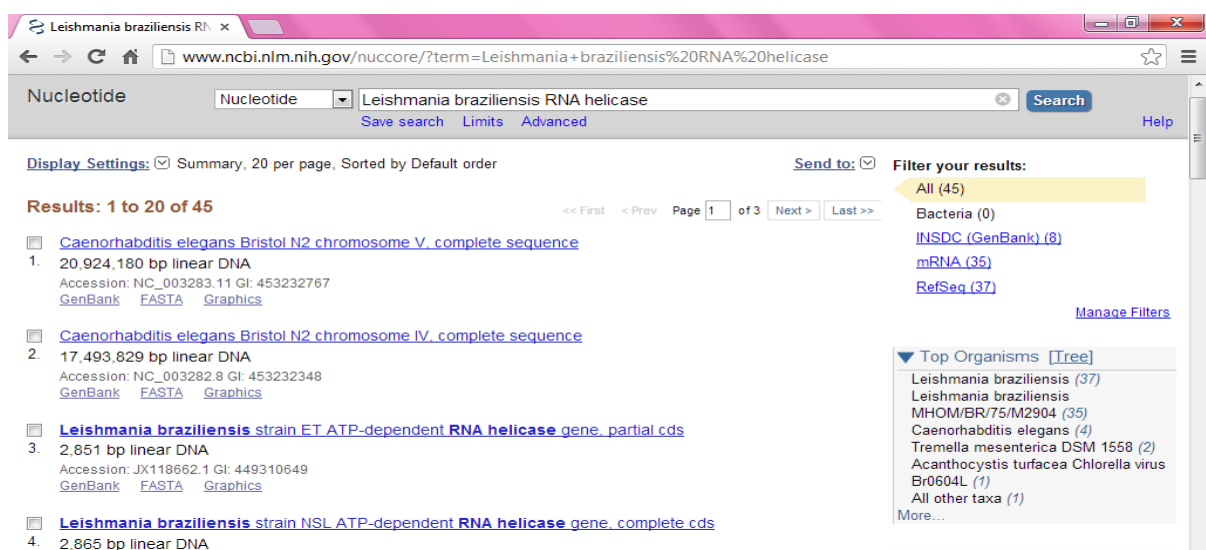


Figura 14. Resultado de busca da sequência nucleotídica de *Leishmania braziliensis*. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov/>>, acessado em 22/02/13.

Após selecionar a opção desejada, uma nova página surgirá mostrando todas as informações da sequência biológica. Clique agora na opção “Send” para salvar sua sequência, marque as opções “Complete Record” e “File”, lembrando que o arquivo deve estar no formato FASTA e em seguida clique em “Create File” para fazer o download da sequência.

O formato FASTA é compatível com outro programa, no caso o BioEdit que adiante será ensinado como manuseá-lo para abrir e editar as sequências obtidas.



Leishmania braziliensis strain ET ATP-dependent RNA helicase gene, partial cds

GenBank: JX118662.1

FASTA Graphics

Go to:

LOCUS JX118662 2851 bp DNA linear INV 16-FEB-2013

DEFINITION Leishmania braziliensis strain ET ATP-dependent RNA helicase gene, partial cds.

ACCESSION JX118662

VERSION JX118662.1 GI:449310649

KEYWORDS .

SOURCE Leishmania braziliensis

ORGANISM [Leishmania braziliensis](#)

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmaniinae; Leishmania; Leishmania braziliensis species complex.

REFERENCE 1 (bases 1 to 2851)

AUTHORS Santos,R.F., Amaral,M.S., Maciel,T.E.F., Vidigal,P.M.P.,

Send: Complete Record Coding Sequences

Choose Destination

File Clipboard

Collections Analysis Tool

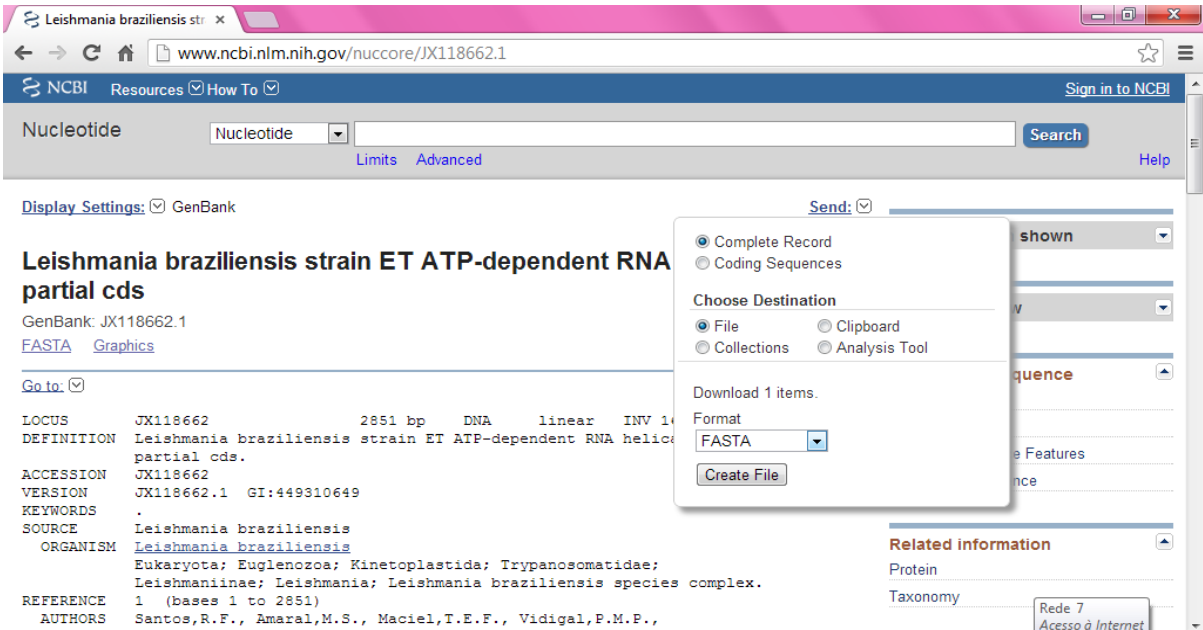
Download 1 items.

Format

FASTA

Create File

Figura 15. Informações da sequência biológica *L.braziliensis*. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov>>, acessado em 22/02/13.



Leishmania braziliensis strain ET ATP-dependent RNA helicase gene, partial cds

GenBank: JX118662.1

FASTA Graphics

Go to:

LOCUS JX118662 2851 bp DNA linear INV 16-FEB-2013

DEFINITION Leishmania braziliensis strain ET ATP-dependent RNA helicase gene, partial cds.

ACCESSION JX118662

VERSION JX118662.1 GI:449310649

KEYWORDS .

SOURCE Leishmania braziliensis

ORGANISM [Leishmania braziliensis](#)

Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmaniinae; Leishmania; Leishmania braziliensis species complex.

REFERENCE 1 (bases 1 to 2851)

AUTHORS Santos,R.F., Amaral,M.S., Maciel,T.E.F., Vidigal,P.M.P.,

Send: Complete Record Coding Sequences

Choose Destination

File Clipboard

Collections Analysis Tool

Download 1 items.

Format

FASTA

Create File

Figura 16. Salvando a sequência *L. braziliensis* no banco de dados. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov>>, acessado em 22/02/13.

2° PASSO: Abrir Sequência no BioEdit

Antes de iniciarmos os procedimentos do segundo passo, precisaremos agora de outra ferramenta básica da bioinformática o programa o BioEdit, que terá a funcionalidade de abrir e editar as sequências extraídas do banco de dado NCBI. O BioEdit é um programa de edição de sequências biológicas com um manuseio relativamente fácil, as operações são realizadas apenas com alguns cliques. E agora iremos dar continuidade com a nossa atividade prática, abra o programa Bioedit, posteriormente clique nas opções “File”, “Open” e selecione a sequência que foi salva anteriormente como mostra à figura 17 e 18 a seguir.

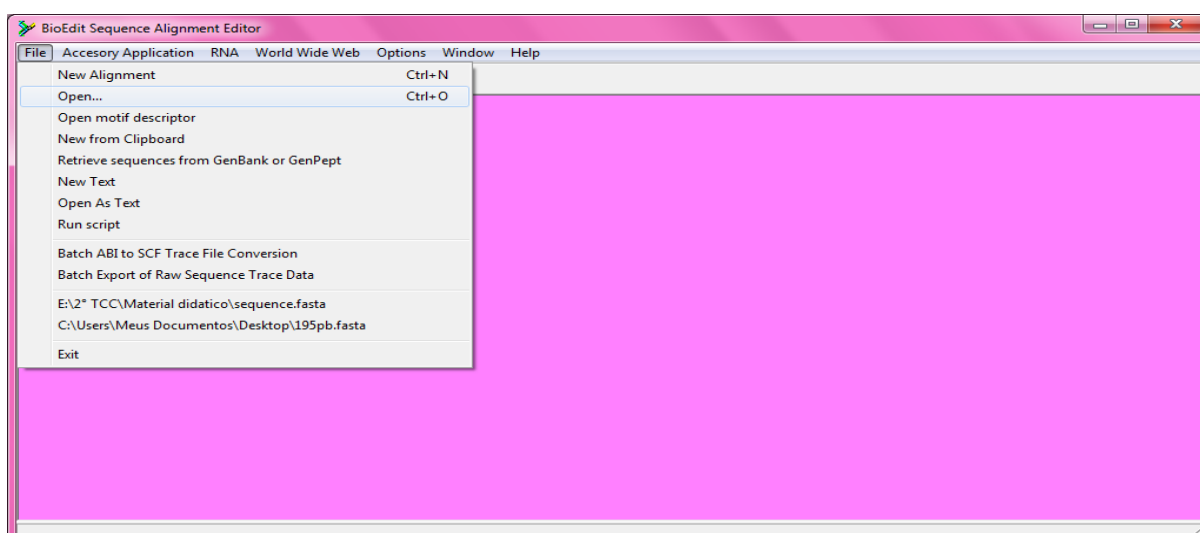


Figura 17. Página Inicial do BioEdit, como abrir uma sequência. Fonte: Própria Autora, 2013.

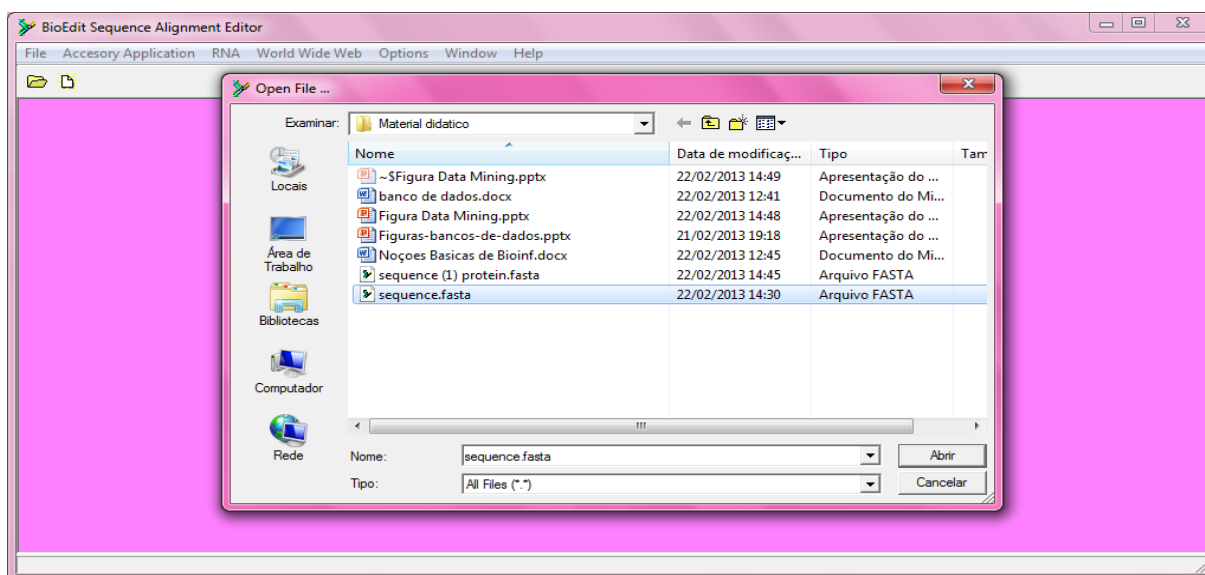


Figura 18. Abrindo a sequência de *L. braziliensis* no formato FASTA. Fonte: Própria Autora, 2013.

Ao abrir o arquivo FASTA, no BioEdit, percebe-se que a sequência está sendo representada pelas bases nitrogenadas, que caracteriza a composição nucleotídica de uma molécula de DNA (Ácido Desoxirribonucleico).

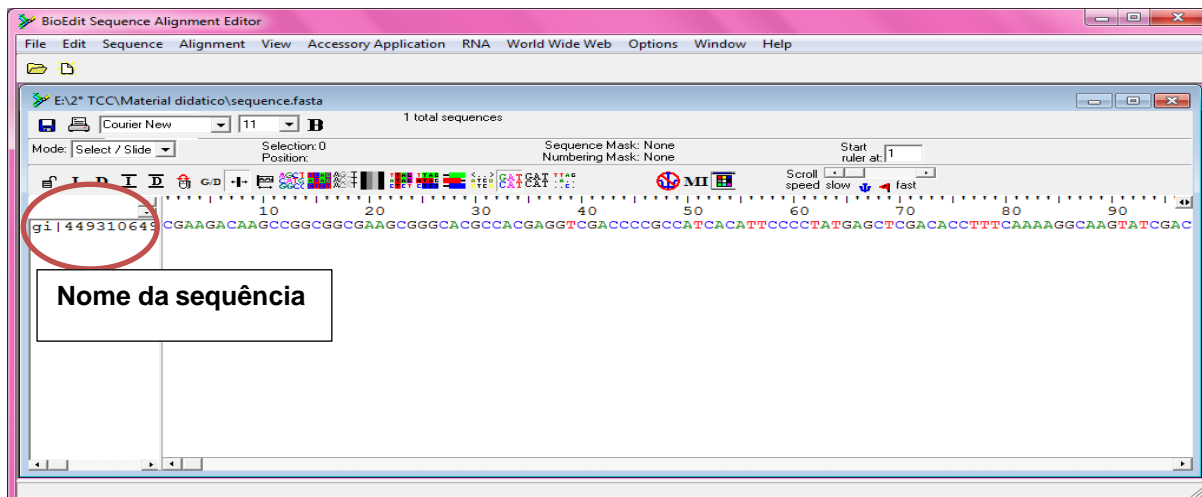


Figura 19. Sequência de *L. braziliensis* aberta no BioEdit. Fonte: Própria Autora, 2013.

3º PASSO: Edição de Sequência no BioEdit

Agora a proposta para esse passo é a realizar a transformação de uma sequência nucleotídica de DNA em uma sequência nucleotídica de RNA (Ácido Ribonucleico). Para isto, é necessário duplicar a sequência de DNA, vamos realizar essa etapa selecionando a sequência clique sob o nome da sequência, a figura 19 mostra o local do nome da sequência, depois clique no menu “Edit”, e nas as opções “Copy Sequence(s)” e “Paste Sequence(s)” que correspondem a função, respectivamente, copiar e colar.

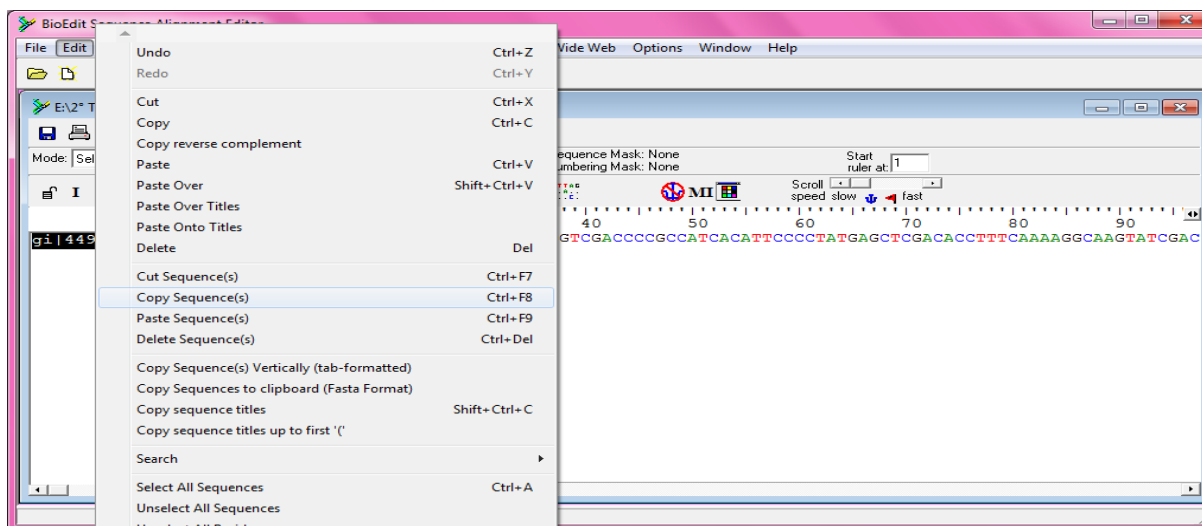


Figura 20. Copiando uma sequência. Fonte: Própria Autora, 2013.

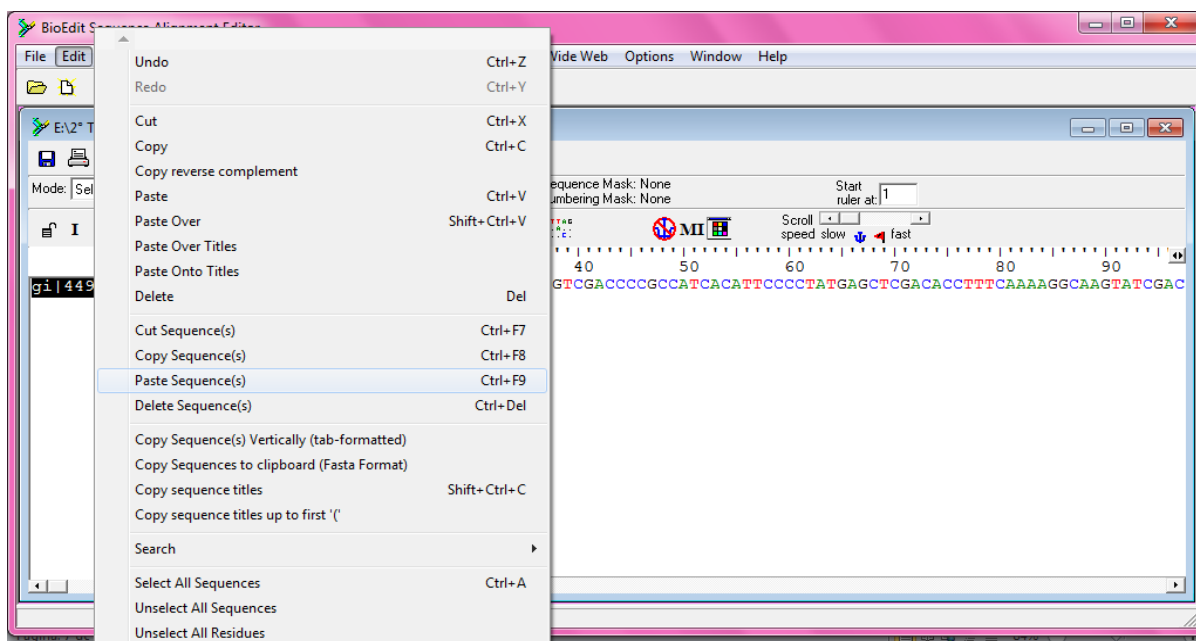


Figura 21. Colando uma sequência. Fonte: Própria Autora, 2013.

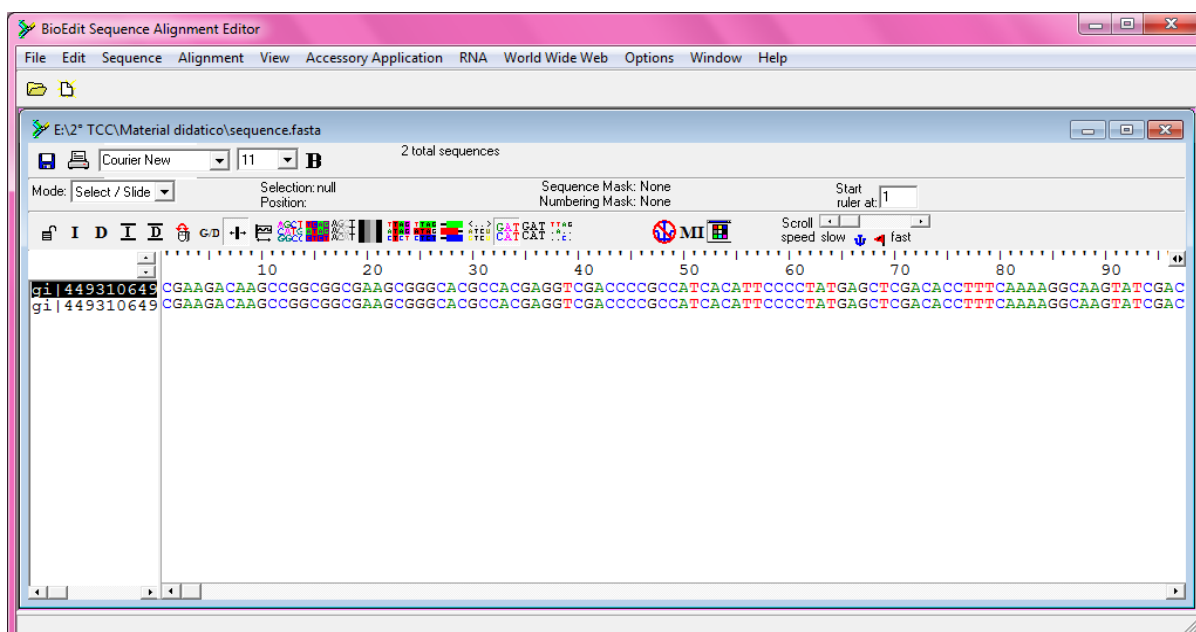


Figura 22. Sequência de DNA duplicada. Fonte: Própria Autora, 2013.

Com a sua sequência duplicada se torna mais fácil agora realizar a transformação das moléculas genéticas. Inicie selecionando a segunda sequência, a que foi copiada, clique no menu “Sequence”, em seguida na função “Nucleic Acid” e no item “DNA → RNA” e pronto sua sequência foi alterada, como mostra a figura 24.

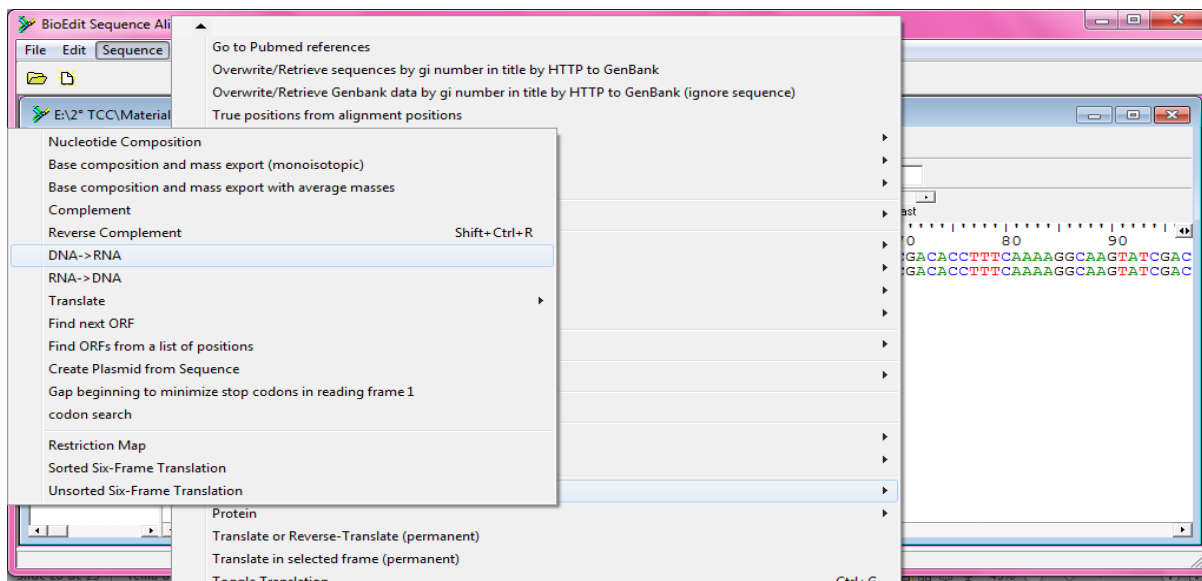


Figura 23. Procedimentos para transformar DNA em RNA. Fonte: Própria Autora, 2013.

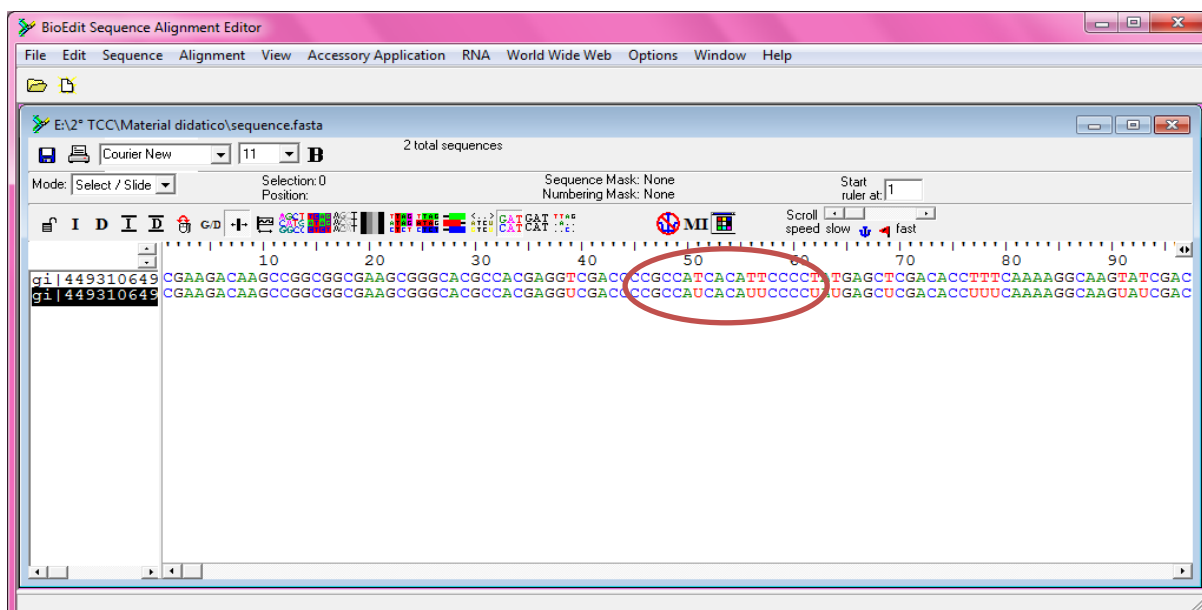


Figura 24. Sequência editada de DNA para RNA. Fonte: Própria Autora, 2013.

4º PASSO: Composição de aminoácido

Vamos explorar mais uma função do programa, que relaciona a sequência nucleotídica de DNA com a composição de aminoácidos. Primeiramente copie e cole a sequência de RNA, como já foi ensinado, agora iremos reverter a sequência de RNA colada para DNA, vá ao menu “Sequence”, opções “Nucleid Acid” e “RNA→DNA”, como ilustra as figuras 26 e 27 a seguir.



Figura 25. Sequência de RNA duplicada. Fonte: Própria Autora, 2013.

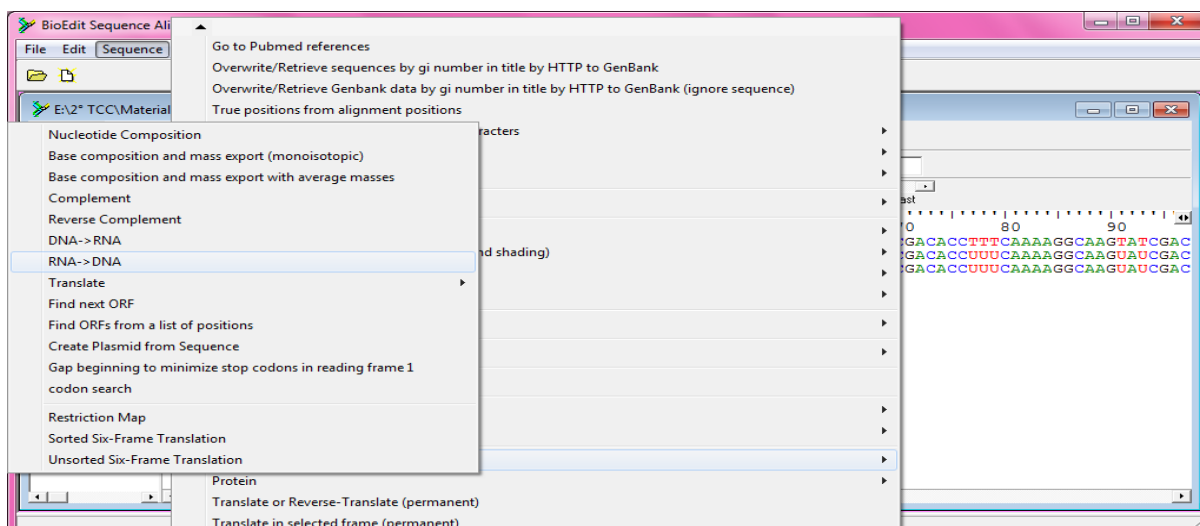


Figura 26. Procedimento de reversão da sequência de RNA para DNA. Fonte: Própria Autora, 2013.

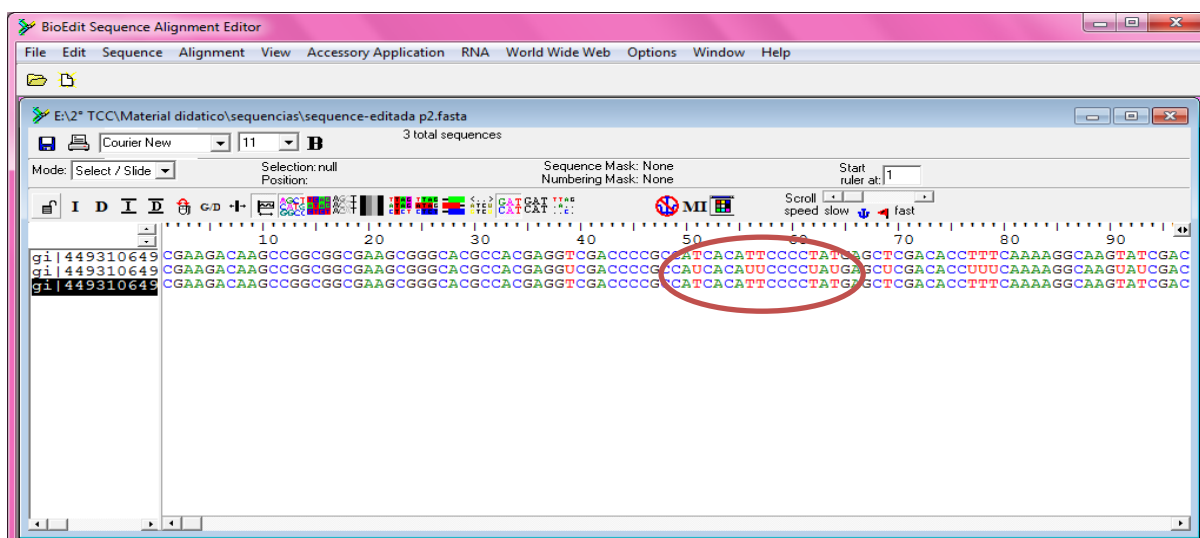


Figura 27. Sequência transformada em DNA. Fonte: Própria Autora, 2013.

Para verificar a composição de aminoácidos é necessário transformar a sequência nucleotídica de DNA em uma sequência de aminoácido. Clique no menu “Sequence” e na função “Translate or Reverse-Translate (permanent)” e a sua sequência é modificada numa sequência de aminoácidos. Perceba que a sequência possui o símbolo (*), essa é a representação dos códon de terminação, quando há a presença desse símbolos implica dizer que esta sequência não é uma proteína e sim uma sequência de aminoácido.

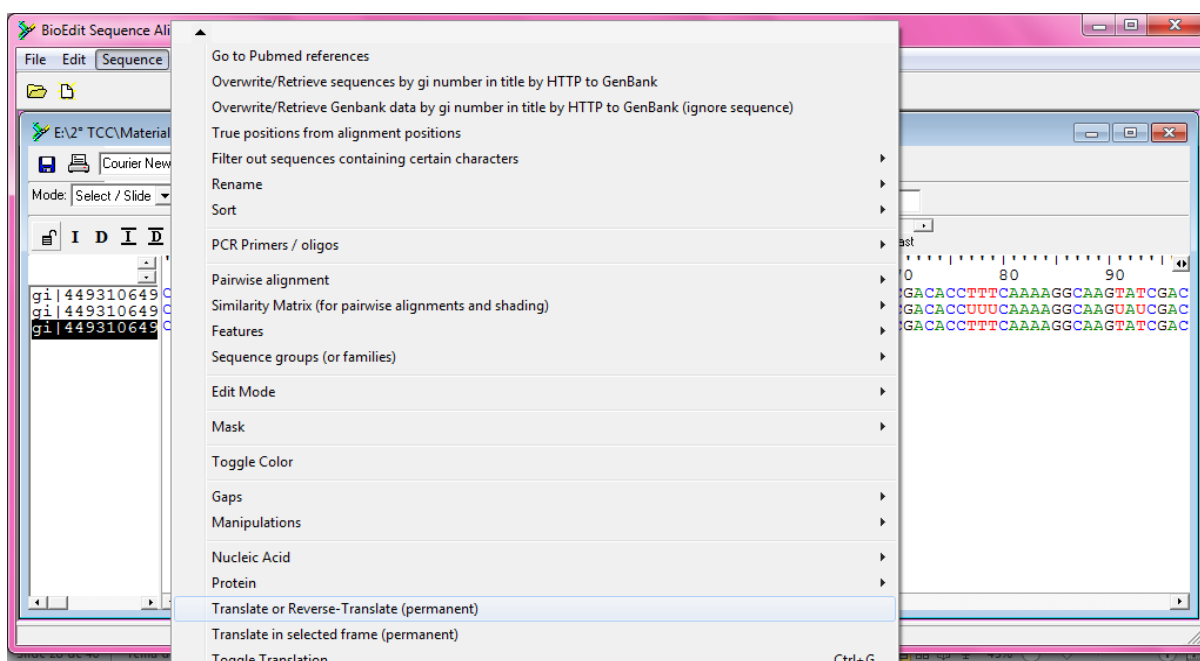


Figura 28. Transformação de DNA em aminoácido. Fonte: Própria Autora, 2013.

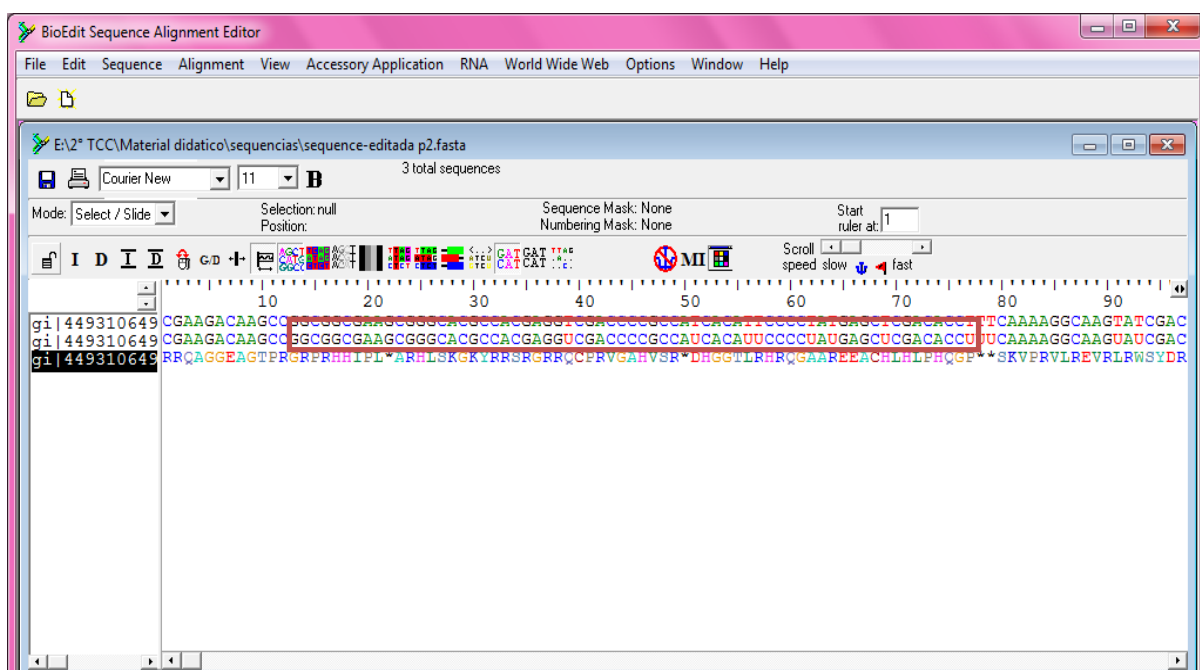


Figura 29. Sequência de Aminoácido. Fonte: Própria Autora, 2013.

E por fim, vamos visualizar a quantidade de aminoácidos existente na sequência de aminoácido de *Leishmania braziliensis*, basta apenas clicar em “Sequence”, “Protein” e em “Amino Acid Composition” que surgirá duas novas janelas, uma representada por um gráfico e a outra uma lista com os resultados obtidos da pesquisa composicional de aminoácidos.

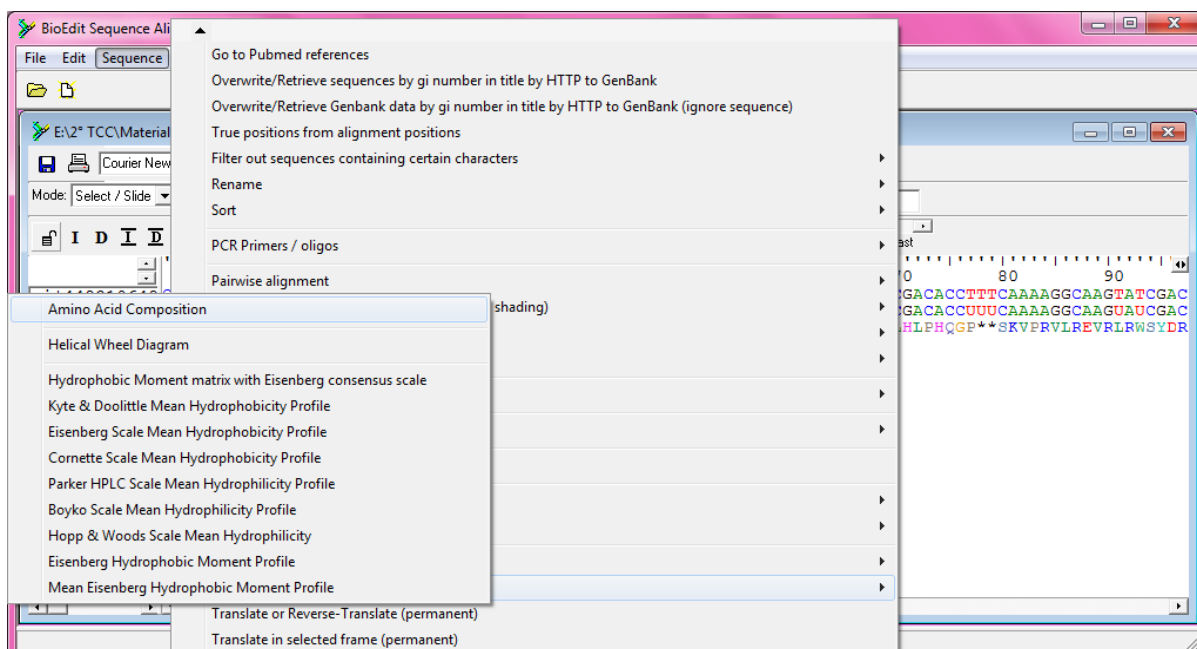


Figura 30. Procedimento para verificar a composição de aminoácidos. Fonte: Própria Autora, 2013.

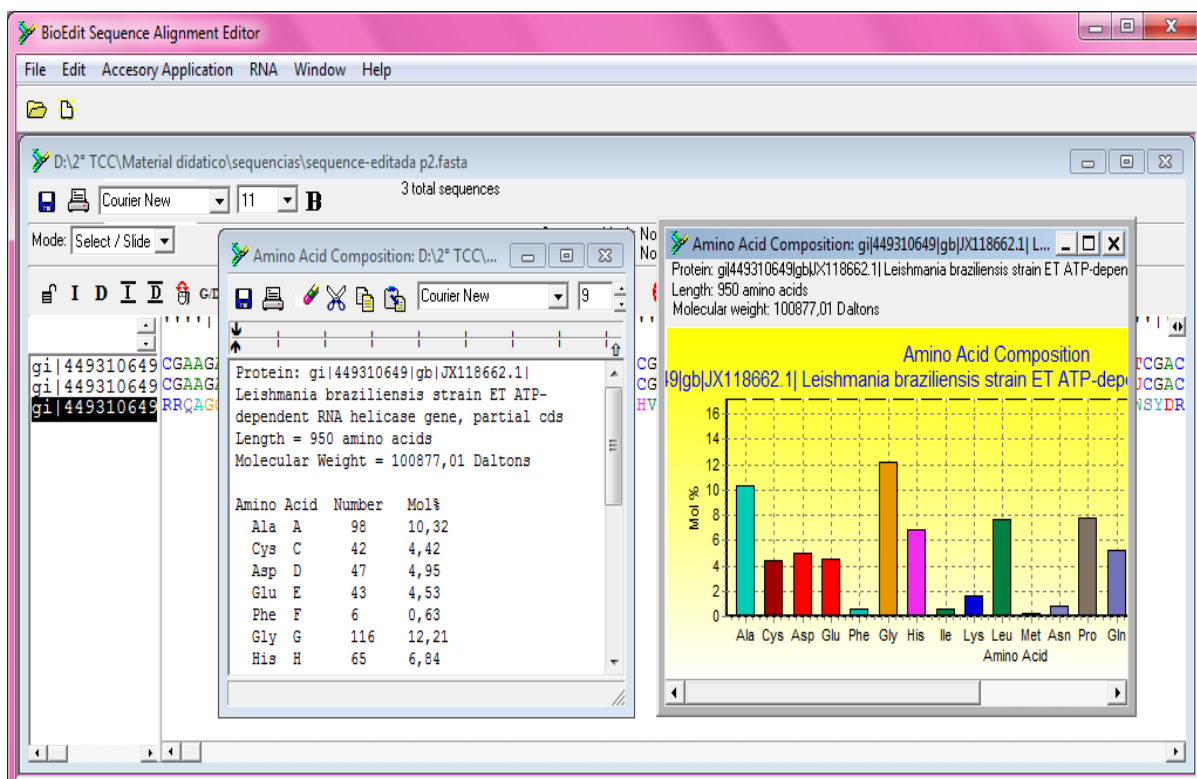


Figura 31. Resultado da composição de aminoácido. Fonte: Própria Autora, 2013.

5º PASSO: Salvar o arquivo

Esse passo é apenas para ensinar como salvar suas sequências produzidas, no qual sempre é necessário fazer ao final de seu trabalho. O passo inicial que deve ser feito é minimizar a lista e o gráfico, depois clique no menu “File” na função “Save As”, e uma nova janela abrirá, escolha o local desejado para armazenar, lembrando que deve selecionar o formato FASTA e pronto suas sequências foram salvas.

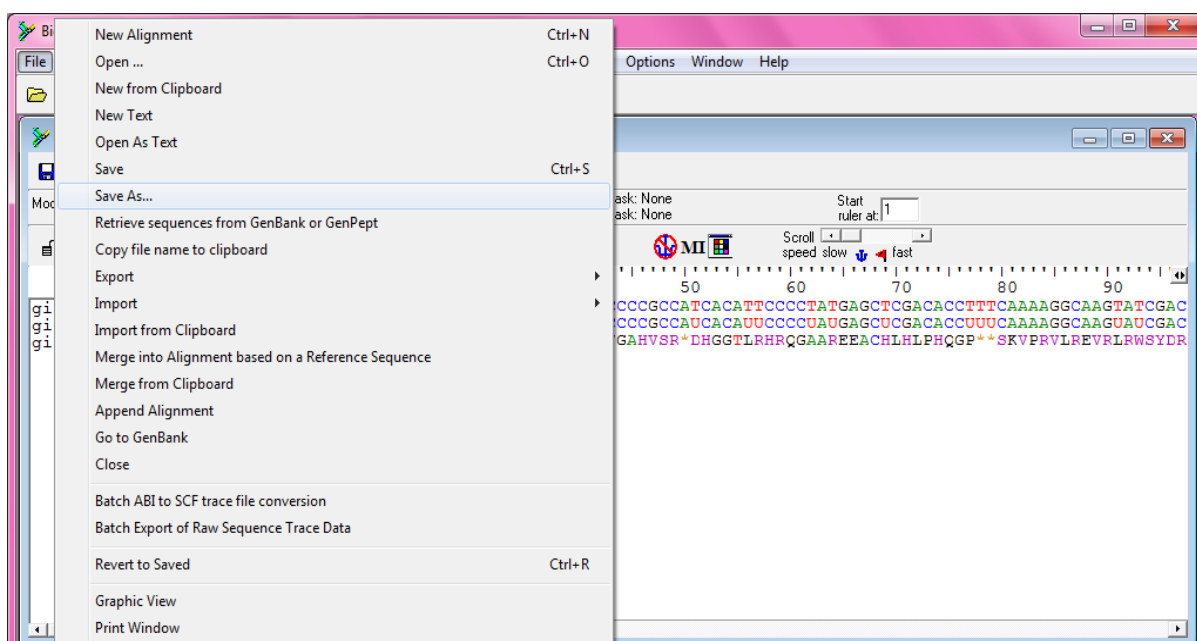


Figura 32. Procedimento para salvar um arquivo. Fonte: Própria Autora, 2013.

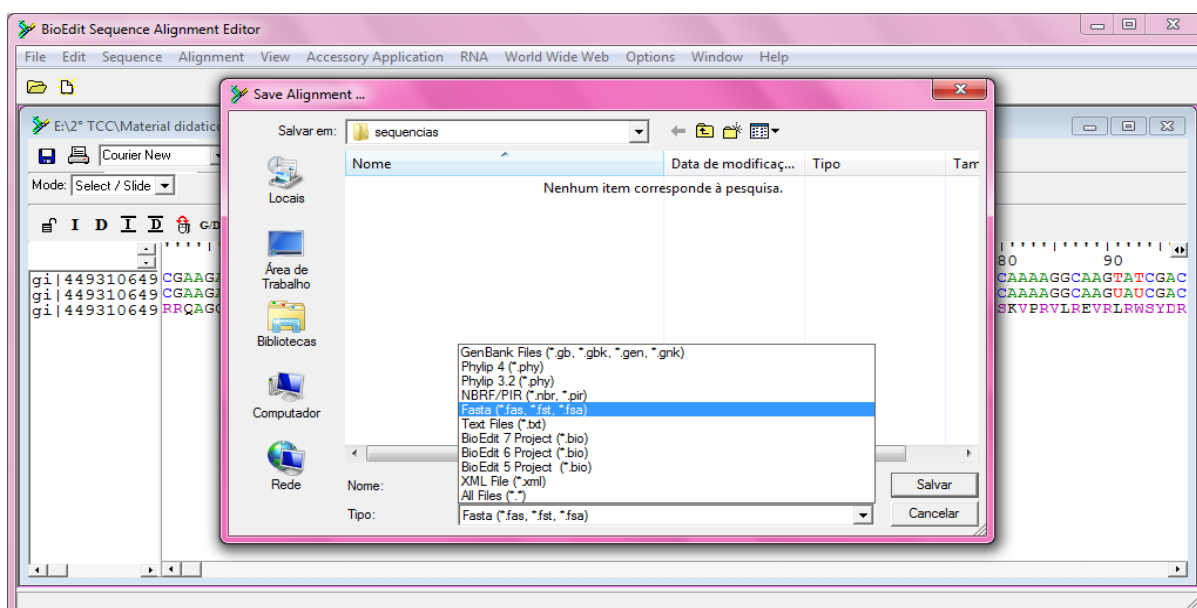


Figura 33. Salvando o arquivo no formato FASTA. Fonte: Própria Autora, 2013.

4. 3ª PRÁTICA: ALINHAMENTO GLOBAL E LOCAL

4.1 Introdução

O alinhamento de sequências biológicas, resumidamente falando, se baseia no processo de comparar duas ou mais sequências de forma a se observar seu nível de semelhança. A similaridade das sequências biológicas pode ser analisada por completa ou apenas entre fragmentos desta sequência. Esse processo de alinhamento pode ser feito de forma global ou local, entre sequências de nucleotídeos ou de proteínas.

Existem centenas de aplicações do alinhamento de sequências, tanto na identificação de genes e proteínas desconhecidas, quanto na comparação da ordem de genes em genomas de organismos proximamente relacionados (sintenia), no mapeamento de sequências expressas dentro de um genoma para identificação de genes, na montagem de genomas e em diversas outras aplicações (PROSDOCIMI, 2007).

Vamos ver agora nessa prática como é feito os dois tipos de alinhamento, o global e o local, para isso os materiais necessários são: o Programa BioEdit e está conectado a internet para realizar alguns procedimentos.

4.2 Realização de um Alinhamento Global

1º PASSO: Obtenção de sequência no banco de dados NCBI

O alinhamento global compara duas ou mais sequências em toda a sua extensão, maximizando as regiões de semelhança que estão sendo comparadas. Para realizar esse tipo de alinhamento execute os procedimentos que serão ensinados abaixo.

Como exemplo, iremos utilizar quatro diferentes espécies de *Leishmania*, que serão *L. amazonensis*, *L. braziliensis*, *L. infantum* e *L. mexicana*, as sequências de RNA helicase. Para obtê-las, acesse o Banco de Dados NCBI <www.ncbi.nlm.nih.gov>, selecione a opção “Nucleotide”, digite o nome do organismo desejado (*Leishmania amazonensis* RNA helicase) e clique em “Search”.

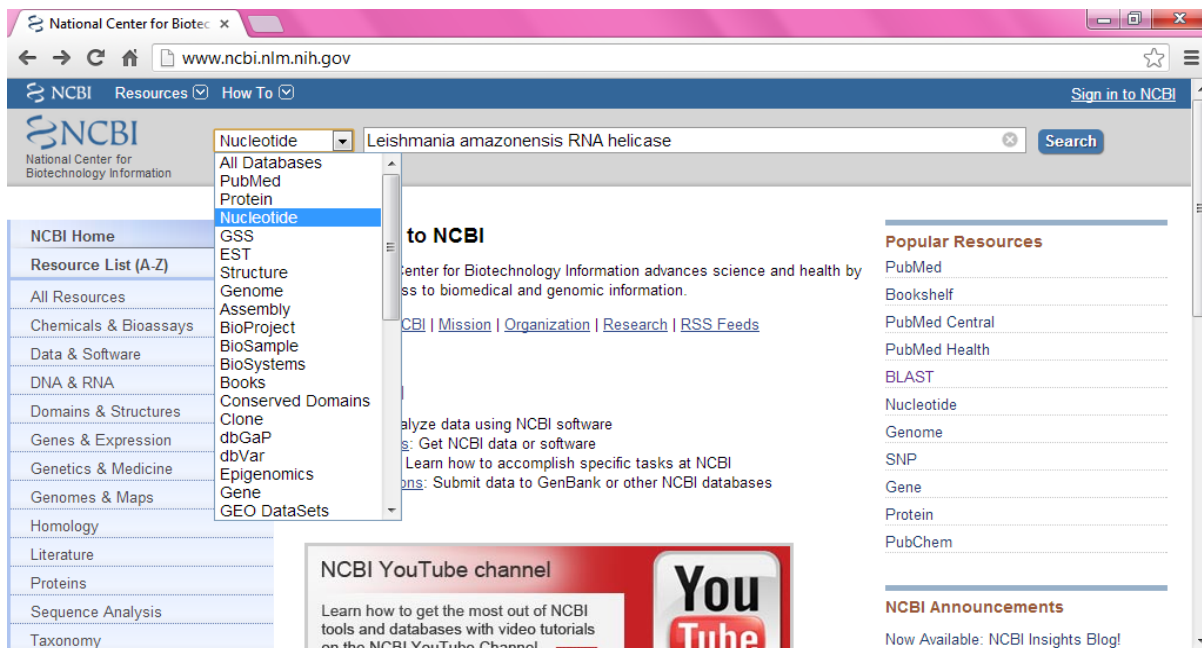


Figura 34. Busca a sequência nucleotídica de *L. amazonensis* no NCBI. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov>>, acessado em 26/02/13.

Após clicar na opção “Search” uma nova página abrirá mostrando os resultados de busca, clique na opção “*Leishmania amazonensis* ATP-dependent RNA helicase (hel 100) mRNA complete cds” e uma nova janela irá expor as informações do organismo, como ilustrado nas figuras 35 e 36 abaixo.

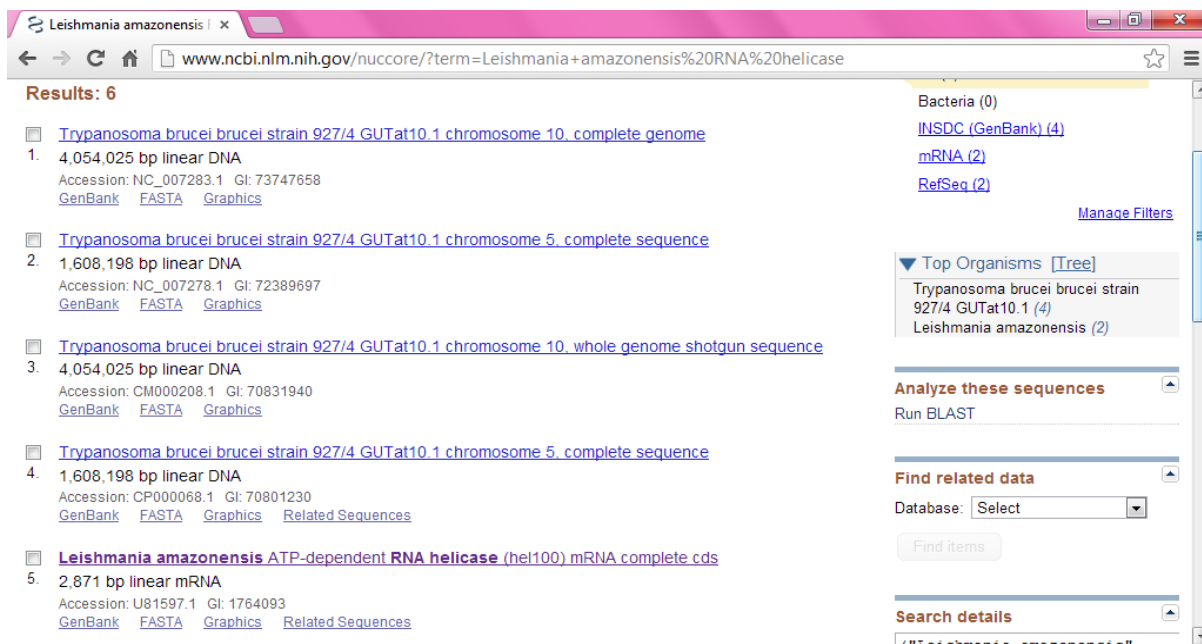


Figura 35. Resultados de seqüências de *L. amazonensis* RNA helicase. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov>>, acessado em 26/02/13.



Leishmania amazonensis ATP-dependent RNA helicase (hel100) mRNA complete cds

GenBank: U81597.1

FASTA Graphics

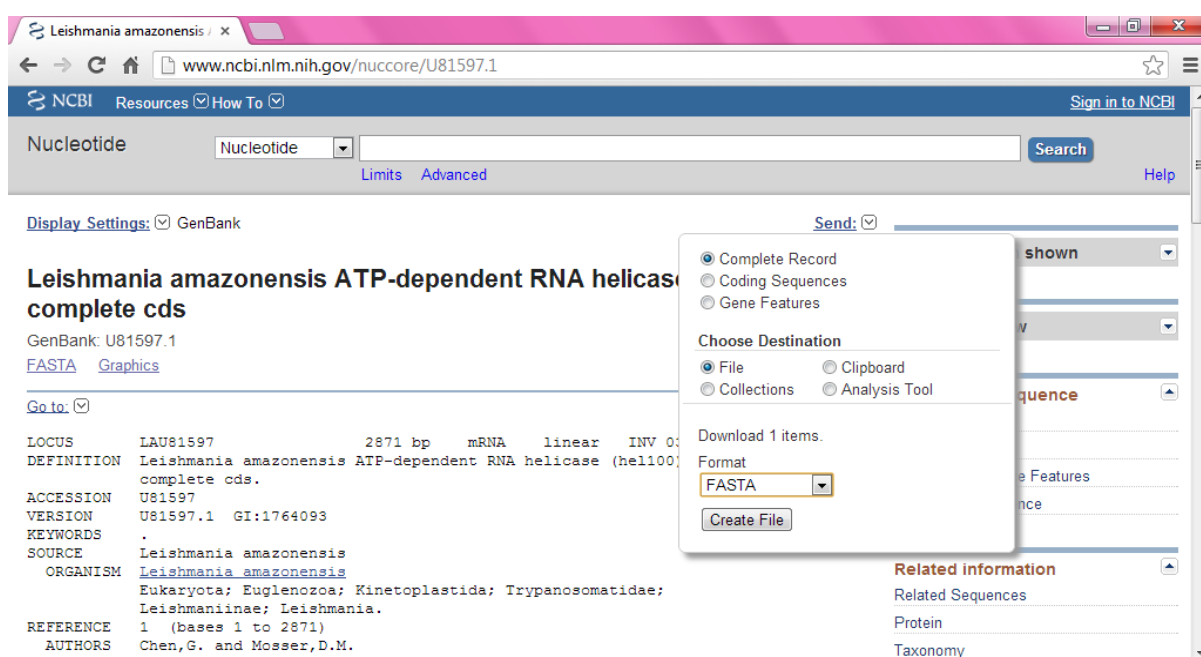
Go to:

LOCUS	LAU81597	2871 bp	mRNA	linear	INV 03-JAN-1997
DEFINITION	Leishmania amazonensis ATP-dependent RNA helicase (hel100) mRNA complete cds.				
ACCESSION	U81597				
VERSION	U81597.1	GI:1764093			
KEYWORDS	.				
SOURCE	Leishmania amazonensis				
ORGANISM	Leishmania amazonensis				
	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmaniinae; Leishmania.				
REFERENCE	1 (bases 1 to 2871)				
AUTHORS	Chen, G. and Mosser, D.M.				

Figura 36. Informações da sequência de *L. amazonensis* RNA helicase. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov>>, acessado em 26/02/13.

Agora iremos salvar a sequência biológica RNA helicase de *L. amazonensis*, basta apenas clicar na opção “Send”, marcar as opções “Complete Record” e “File”, selecione o formato FASTA e clique no botão “Create File” que em seguida iniciará o download da sequência.

Após cada download renomeie as sequências colocando o nome do organismo, por exemplo, sequence L. a. fasta.



Leishmania amazonensis ATP-dependent RNA helicase (hel100) mRNA complete cds

GenBank: U81597.1

FASTA Graphics

Go to:

LOCUS	LAU81597	2871 bp	mRNA	linear	INV 03-JAN-1997
DEFINITION	Leishmania amazonensis ATP-dependent RNA helicase (hel100) mRNA complete cds.				
ACCESSION	U81597				
VERSION	U81597.1	GI:1764093			
KEYWORDS	.				
SOURCE	Leishmania amazonensis				
ORGANISM	Leishmania amazonensis				
	Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmaniinae; Leishmania.				
REFERENCE	1 (bases 1 to 2871)				
AUTHORS	Chen, G. and Mosser, D.M.				

Figura 37. Salvando a sequência de *L. amazonensis* RNA helicase. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov>>, acessado em 26/02/13.

The screenshot shows the NCBI website interface for the sequence U81597.1. The title is "Leishmania amazonensis ATP-dependent RNA helicase (hel100) mRNA complete cds". The sequence is 2871 bp long. The download button for "sequence (1).fasta" is highlighted with a red arrow.

Figura 38. Download da sequência de *L. amazonensis* RNA helicase. Fonte: Site NCBI <<http://www.ncbi.nlm.nih.gov>>, acessado em 26/02/13.

Repita todos os procedimentos anteriores substituindo apenas o nome do organismo pelos outros tipos de *Leishmania sp*, citadas no início do primeiro passo. Escolha as opções "*L. infantum* (LINJ_24_0240)" e "*L. mexicana* (LMXM_33_2620)" e "*L. braziliensis* (a mesma sequência obtida na prática anterior)".

2º PASSO: Abrir várias sequências em um único arquivo

Após obter as quatro diferentes sequências de *Leishmania sp* no banco de dados NCBI iremos abrir todas essas sequências em um único arquivo no BioEdit para realizarmos o alinhamento global. Abra o programa BioEdit, clique no menu "File" e em "Open", selecione uma das sequência de *Leishmania sp*, depois clique em "Abrir" e a sequência será aberta.

The screenshot shows the BioEdit Sequence Alignment Editor interface. The 'Open File...' dialog box is open, displaying a list of files in the 'sequencias' folder. The file 'sequence L.a.fasta' is selected. The 'Abrir' button is visible at the bottom right of the dialog box.

Nome	Data de modificaç...	Tipo	Tarr
sequence (1).protein.fasta	22/02/2013 14:45	Arquivo FASTA	
sequence L.a.fasta	25/02/2013 00:45	Arquivo FASTA	
sequence L.b.fasta	22/02/2013 14:30	Arquivo FASTA	
sequence L.i.fasta	25/02/2013 00:40	Arquivo FASTA	
sequence L.m.fasta	25/02/2013 00:36	Arquivo FASTA	
sequence-editada p2.fasta	22/02/2013 15:12	Arquivo FASTA	

Figura 39. Abrindo a sequência *L. amazonensis* no BioEdit. Fonte: Própria Autora, 2013.

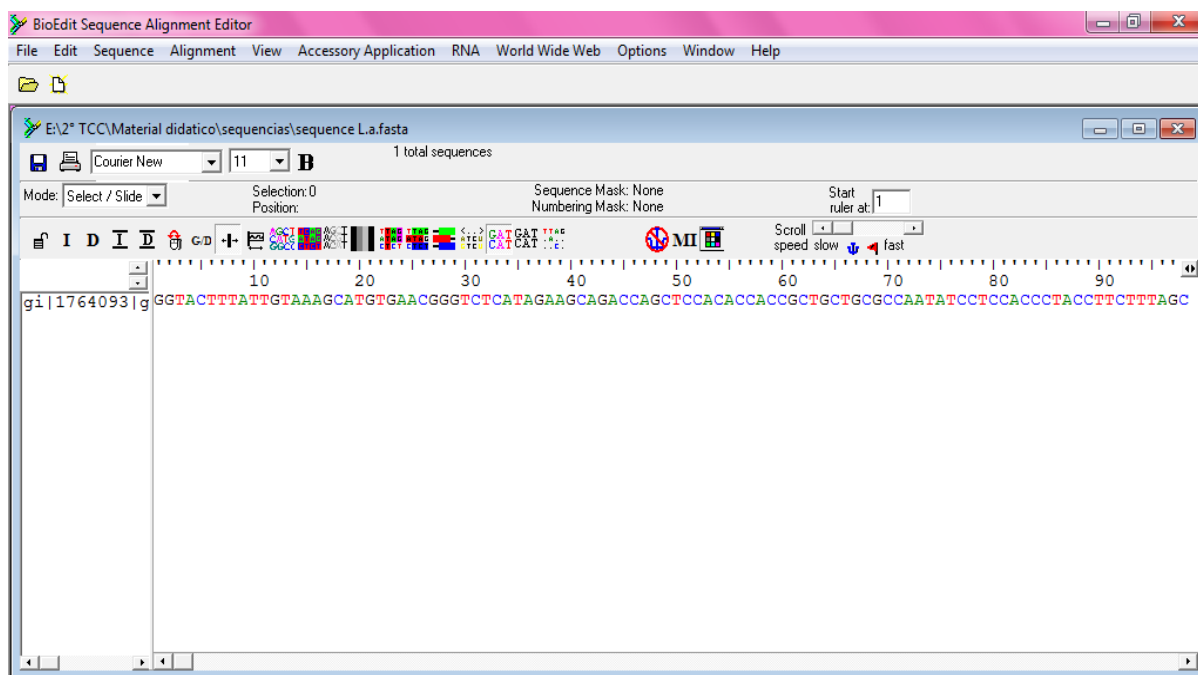


Figura 40. Sequência de *L. amazonensis*. Fonte: Própria Autora, 2013.

Com a sua sequência aberta, vá ao menu “File”, depois na função “Import” e “Sequence alignment file”. Uma nova janela aparecerá para que você selecione outra sequência de *Leshimania sp*, a qual ficara localizada abaixo da primeira sequência existente.

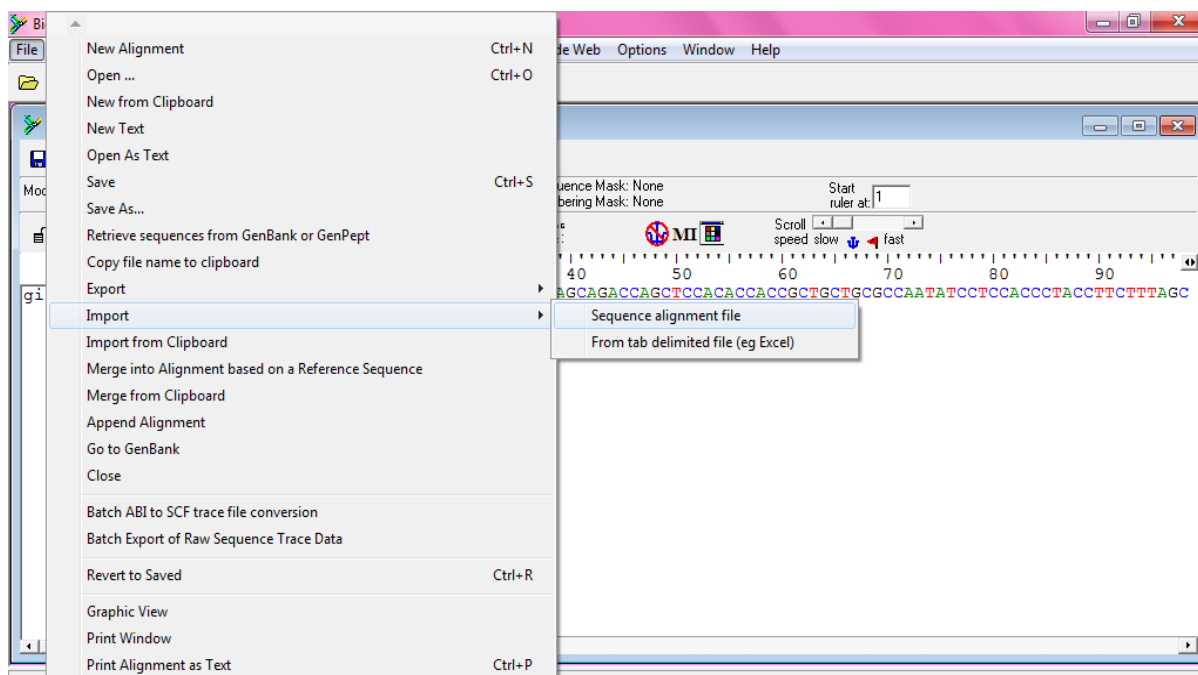


Figura 41. Funções necessárias para importar sequências. Fonte: Própria Autora, 2013.

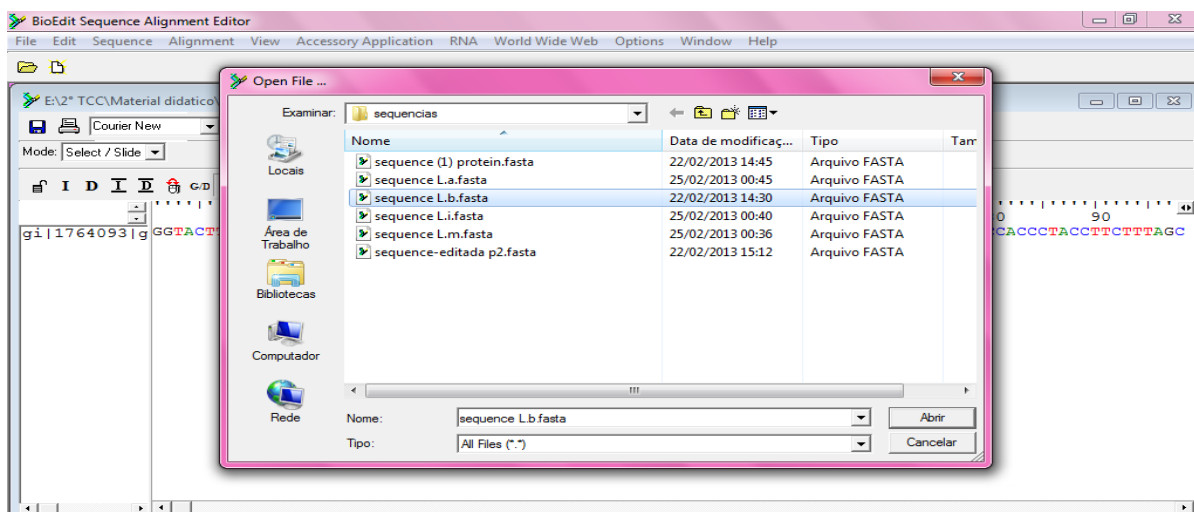


Figura 42. Importando uma nova sequência. Fonte: Própria Autora, 2013.

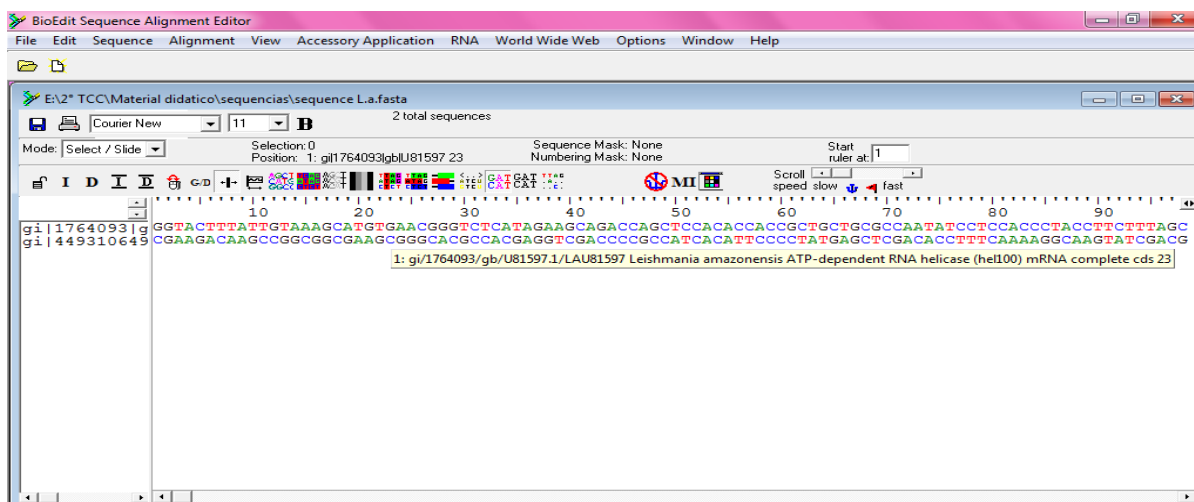


Figura 43. Sequência de *L.braziliensis* importada. Fonte: Própria Autora, 2013.

Repita o item anterior com as outras duas sequências restantes, para deixar as quatro sequências de *Leishmania sp* em um único arquivo.

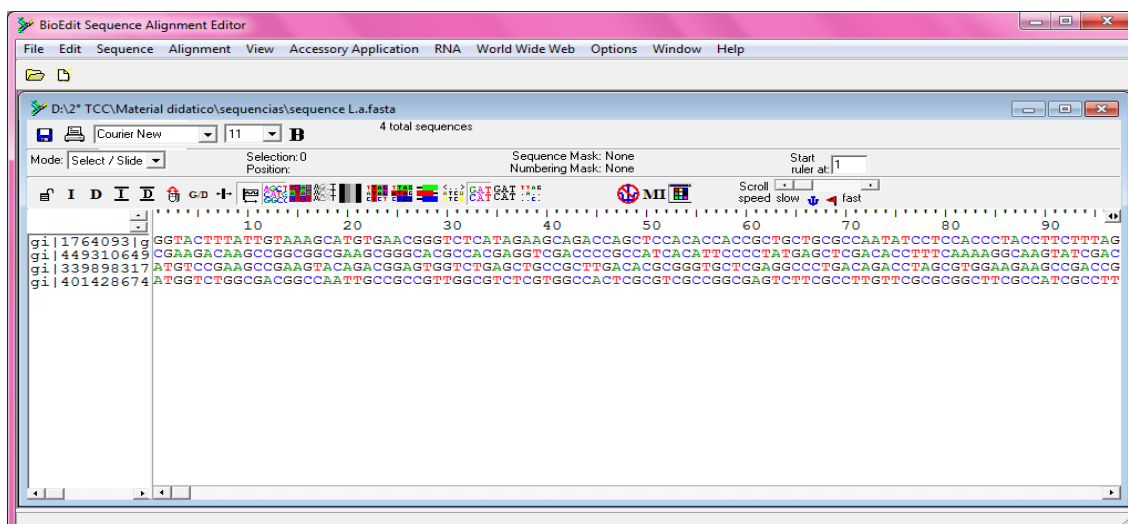


Figura 44. Várias sequências em um único arquivo. Fonte: Própria Autora, 2013.

3º PASSO: Realização de um Alinhamento Global

Selecione as quatro sequências, vá ao menu “Accessory Application”, escolha a função “ClustalW Multiple alignment” que abrirá uma nova janela, clique em Run ClustalW e depois em “OK” que iniciara o processamento para alinhar as sequências.

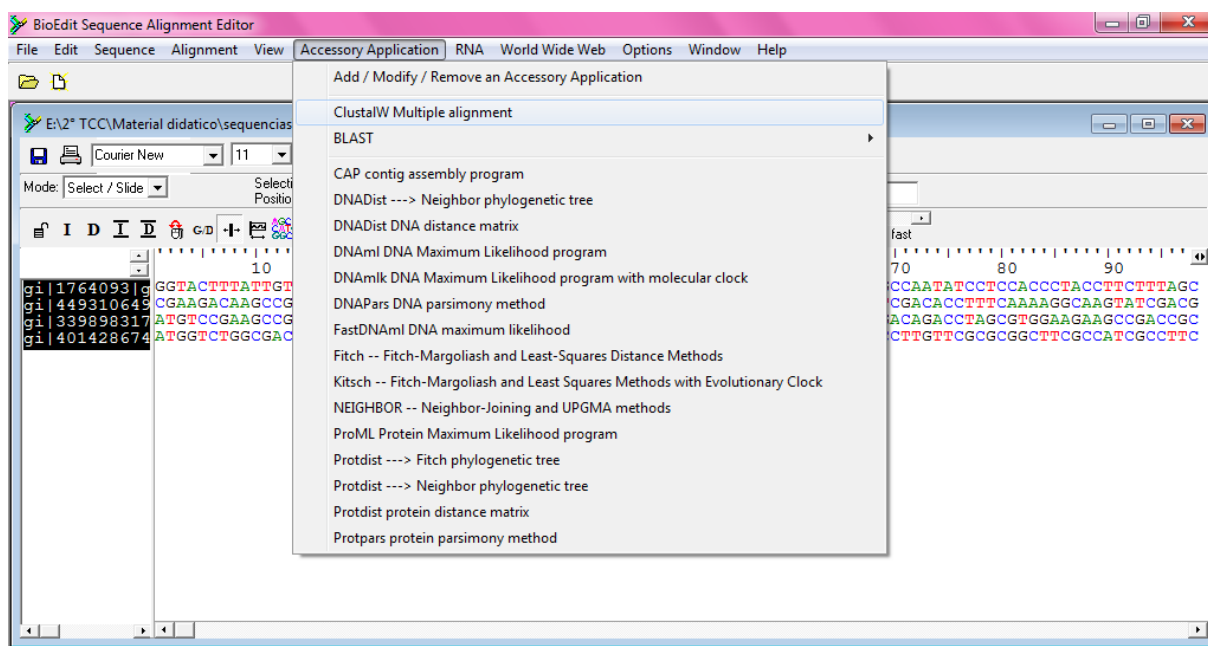


Figura 45. Funções para realizar o alinhamento global. Fonte: Própria Autora, 2013.

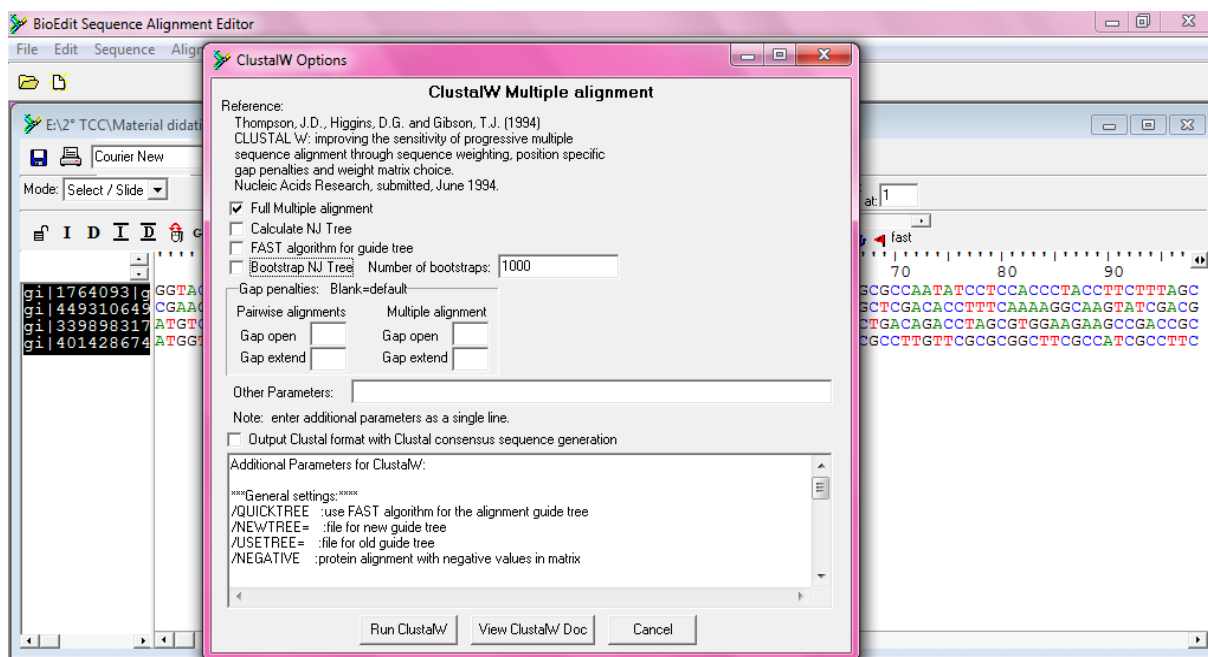


Figura 46. Janela do ClustalW Options. Fonte: Própria Autora, 2013.

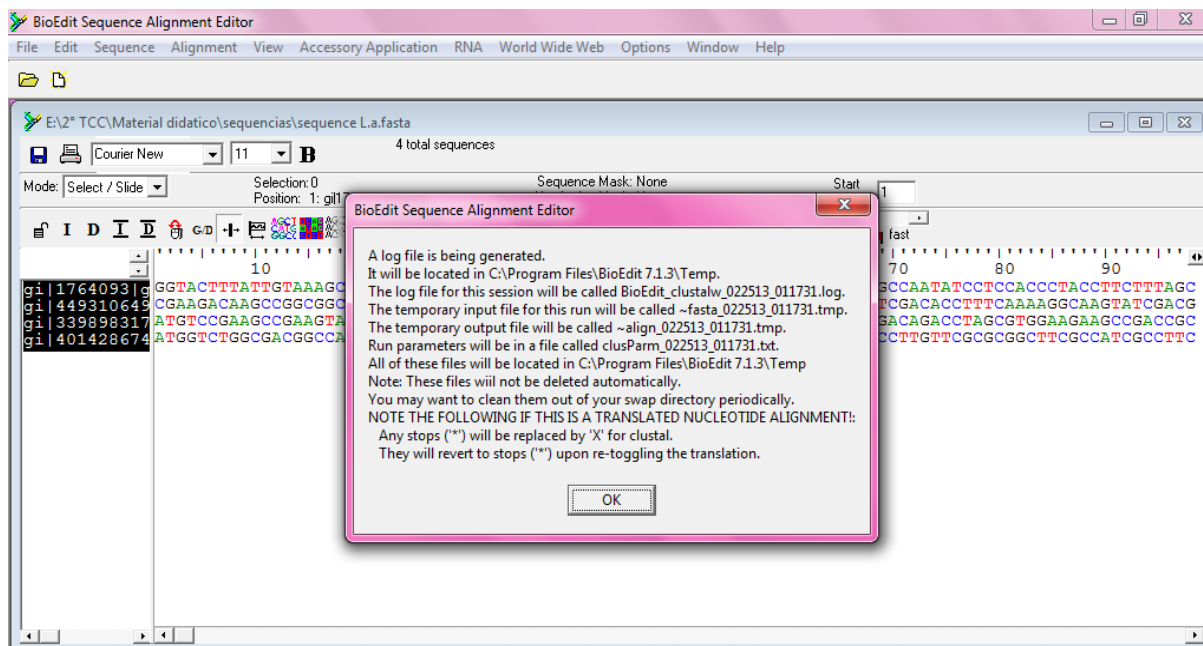


Figura 47. Etapa para realizar o alinhamento global. Fonte: Própria Autora, 2013.

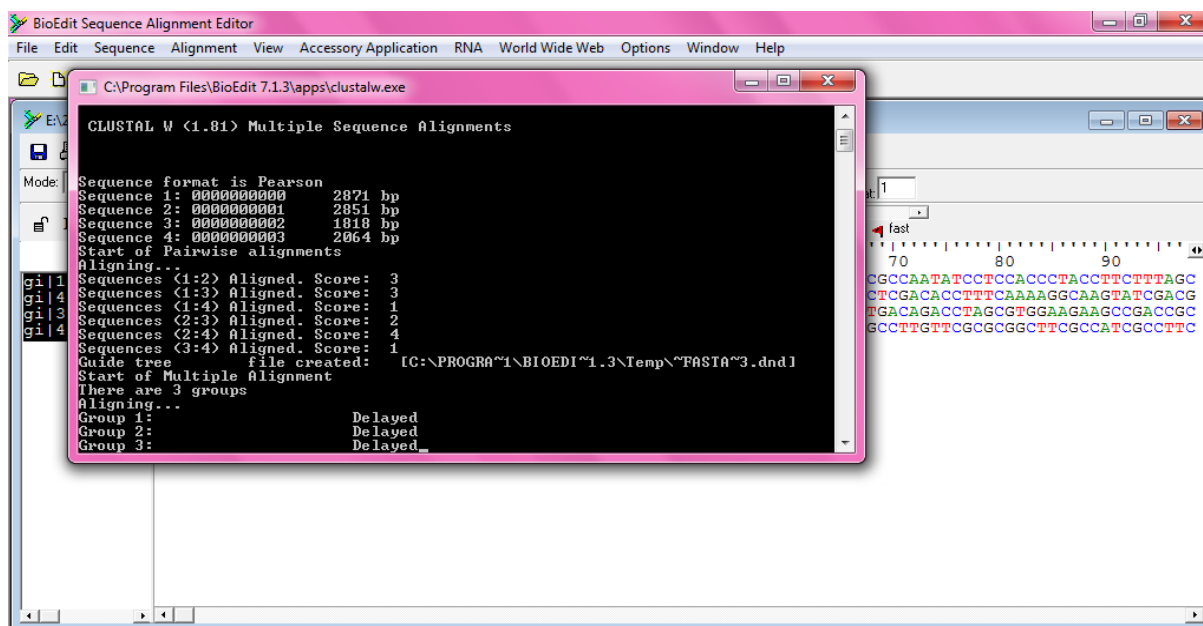


Figura 48. Processando o alinhamento global. Fonte: Própria Autora, 2013.

Após processar os dados, seu alinhamento global será concluído. Note que, para as sequências serem alinhadas em toda a sua extensão é bem simples, basta apenas alguns cliques. Você deve ter percebido que a nova janela com seu alinhamento global apresentam uns “traços” nas sequências. Você sabe o que eles significam? Não. São chamados de *gaps* (falhas), que são símbolos introduzidos pelo programa para representar os locais que não são similares entre as sequências.

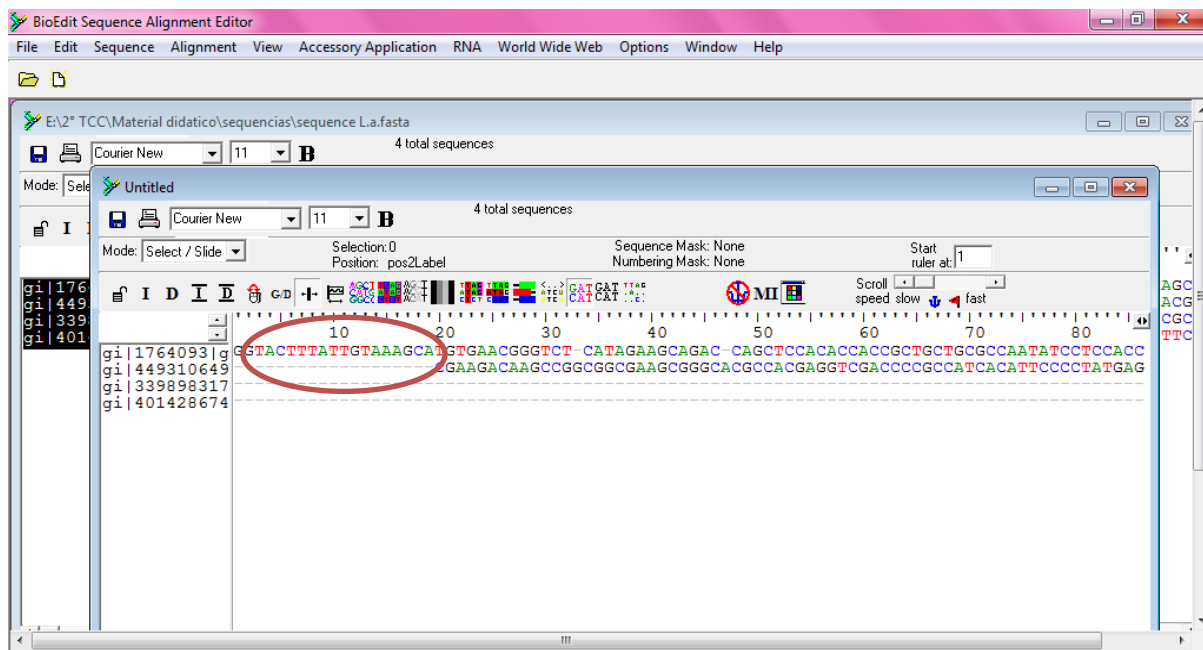


Figura 49. Alinhamento Global das sequências de *Leishmania* sp. Fonte: Própria Autora, 2013.

Mova a barra de rolagem, localizada abaixo das sequências, e veja em toda a extensão da sequência que ela irá ter locais em que são muito semelhantes, embora apresente alguns gaps, pois alinhar sequências é colocar uma sobre a outra de modo que todo base faça correspondência com outra ou com um gap. Os gaps representam inserções ou deleções entre as sequências.

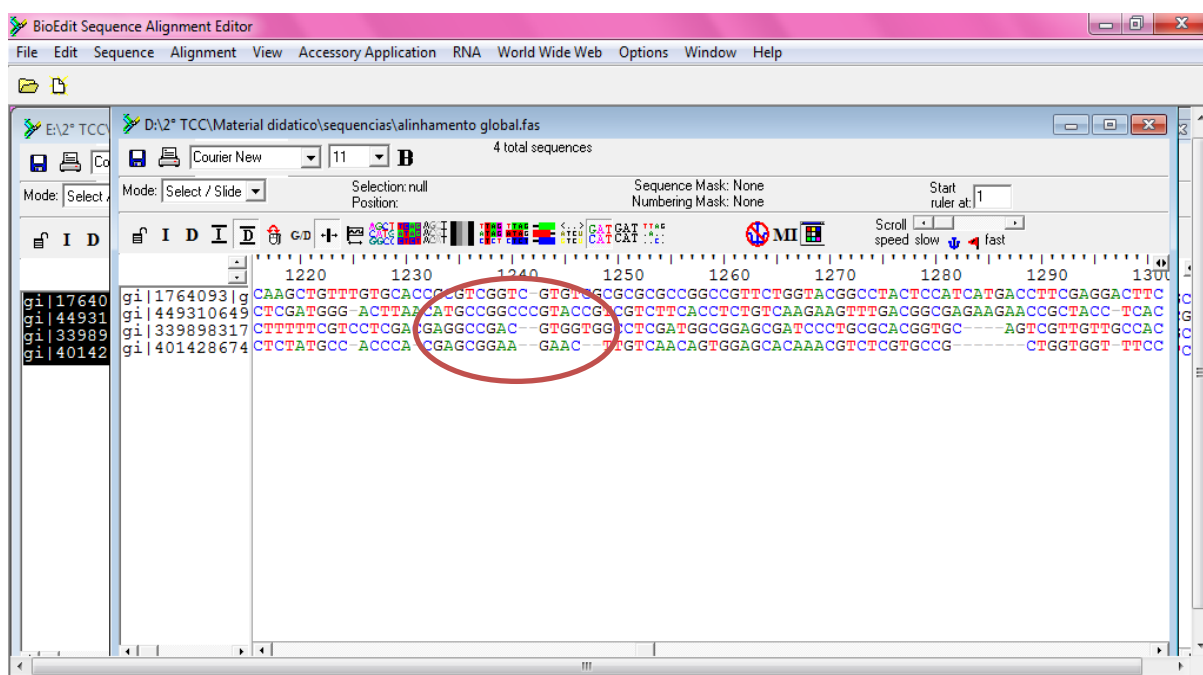


Figura 50. Analisando o alinhamento global. Fonte: Própria Autora, 2013.

Salve sua atividade clicando nas opções seguintes “File”, “Save As”, dê nome ao arquivo e selecione o formato FASTA.

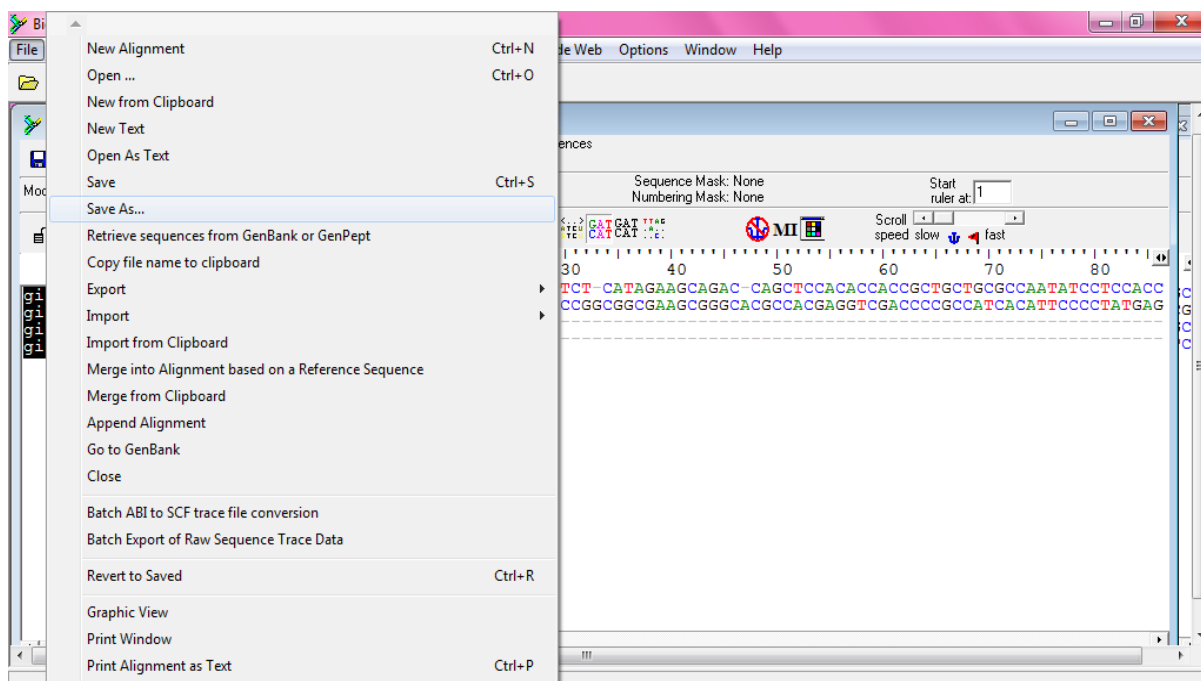


Figura 51. Opções para salvar o arquivo. Fonte: Própria Autora, 2013.

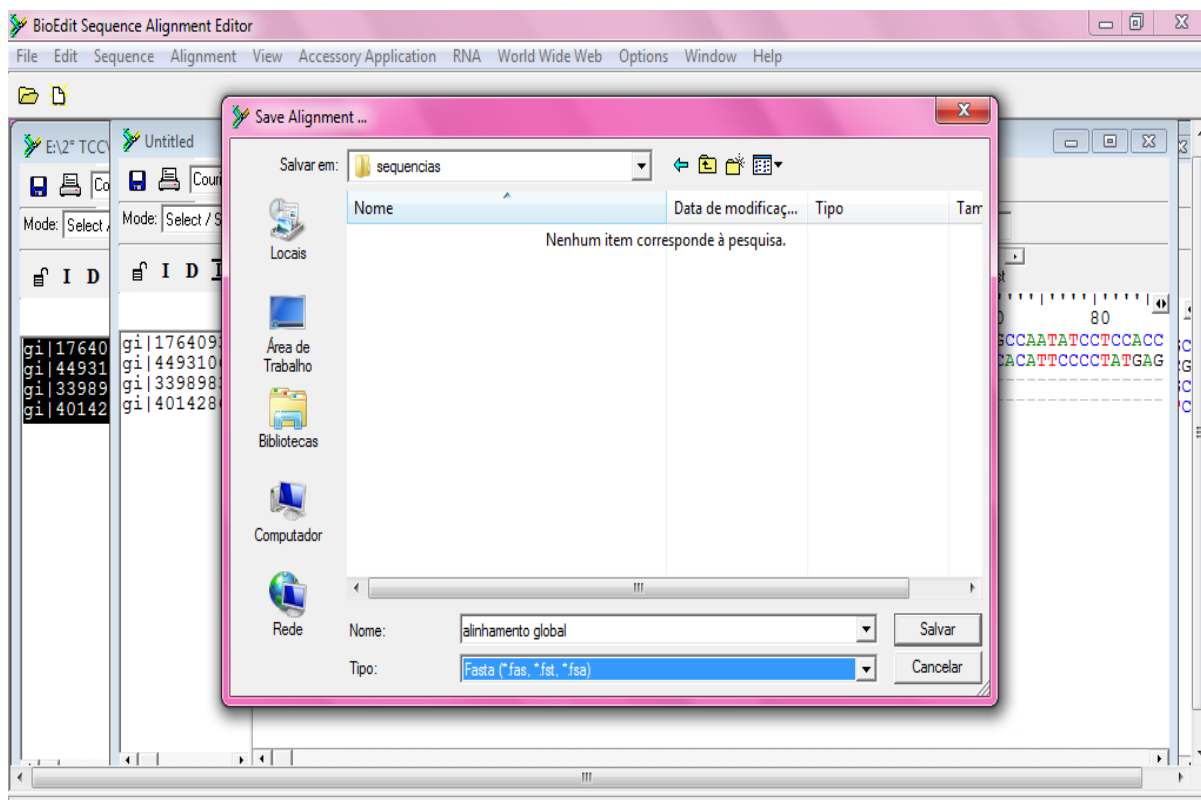


Figura 52. Salvando o arquivo no formato FASTA. Fonte: Própria Autora, 2013.

4.3 Realização de um Alinhamento Local

1° PASSO: Abrir sequência no programa BioEdit

O alinhamento local é utilizado para comparar duas ou mais sequências de tamanhos diferentes e são analisadas trechos conservados. Nessa prática iremos comparar e alinhar uma sequência de um gene com um universo de genes armazenados em um banco de dados. A ferramenta mais utilizada para esse tipo de alinhamento é o BLAST (Basic Local Alignment Search Tool) que está disponível no Banco de Dados do portal NCBI por meio do endereço <<http://blast.ncbi.nlm.nih.gov/>>.

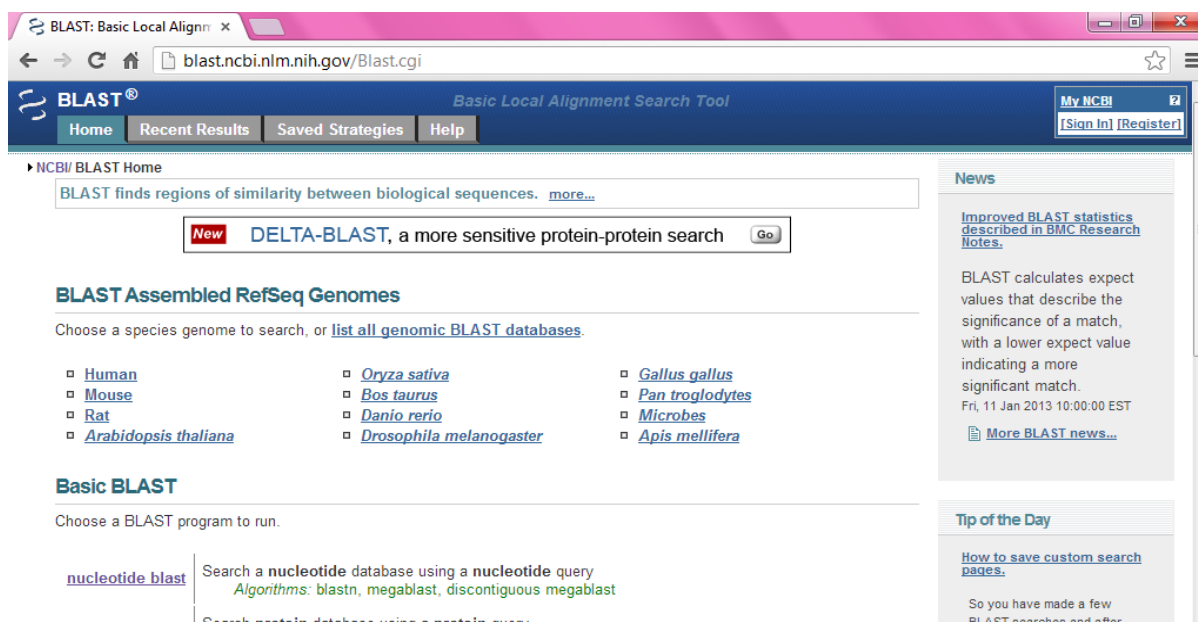


Figura 53. Página inicial do BLAST Assembled RefSeq Genomes. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

Antes de acessarmos esse site precisamos obter uma sequência gênica, para isso abriremos o programa BioEdit, localize o arquivo da sequência de “*Leishmania amazonensis.fasta*” e clique em abrir. Ao abrir a sequência dê um duplo clique sobre o nome desta, como mostra a figura 55, e será disponibilizado uma nova janela contendo toda a sequência, basta seleciona-la por completa e copia-la apertando as teclas “ctrl+c”.

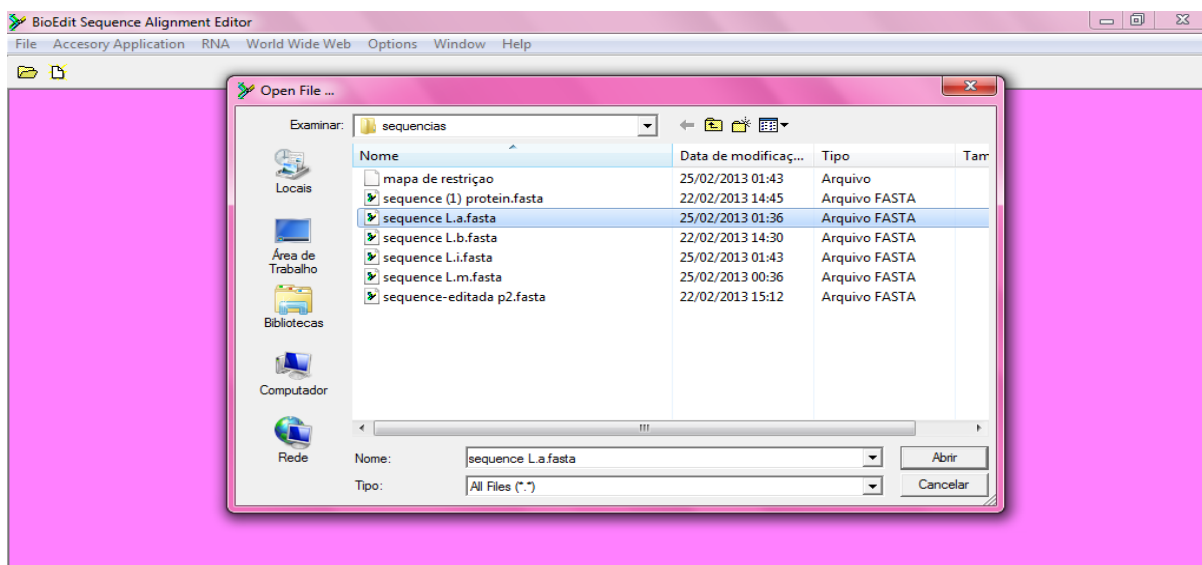


Figura 54. Abrindo a sequência *Leishmania amazonensis.fasta* no BioEdit. Fonte: Própria Autora, 2013.

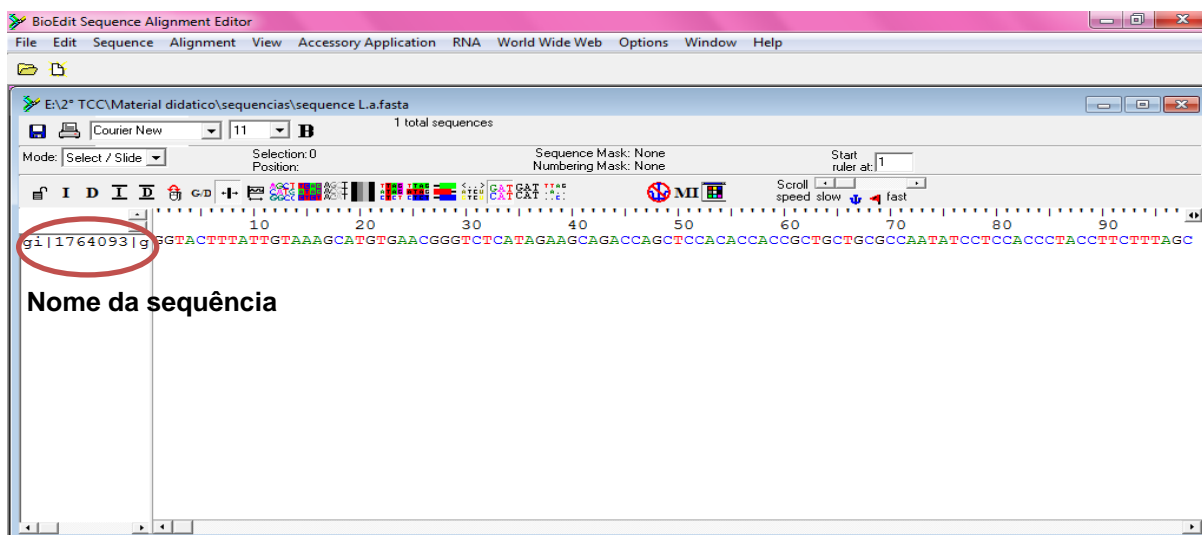


Figura 55. Mostrando o local do nome da sequência. Fonte: Própria Autora, 2013.

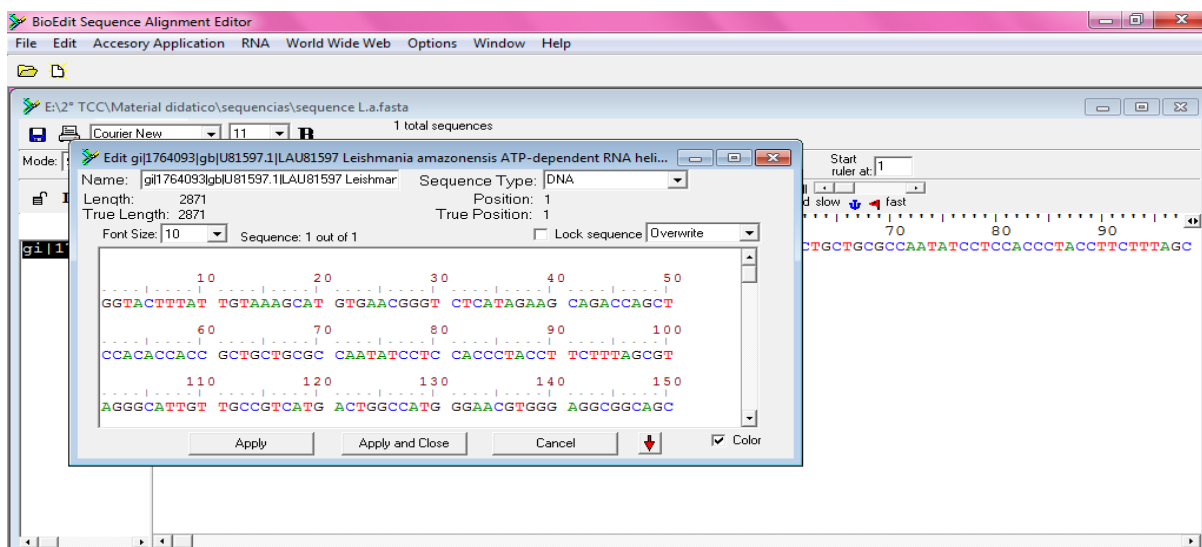


Figura 56. Nova janela mostrando toda a sequência. Fonte: Própria Autora, 2013.

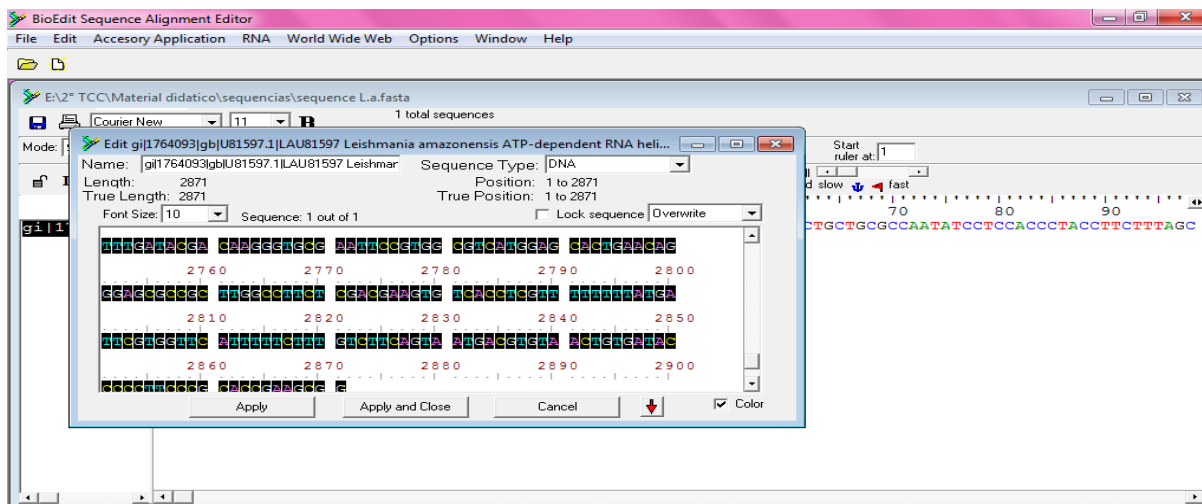


Figura 57. Selecionando a sequência completa. Fonte: Própria Autora, 2013.

2º PASSO: Realização do Alinhamento Local

Após obter a sequência desejada, vá agora para o site <<http://blast.ncbi.nlm.nih.gov/>> para utilizarmos o programa BLAST que será executado. Observe que na página inicial do site temos a função de escolher o genoma de espécies específicas que se localiza sob a opção “BLAST Assembled RefSeq Genomes” e temos ainda a opção “Basic BLAST” que mostra o tipos de BLAST existentes como ilustrado na figura 58. Vamos selecionar o tipo de BLAST o qual procure um banco de dados de nucleotídeo utilizando uma sequência de nucleotídeo, pois lembrem a sequência que será alinhada é a *L. amazonensis* que possui uma composição de nucleotídeo, então você deve clicar na opção “nucleotide blast”.

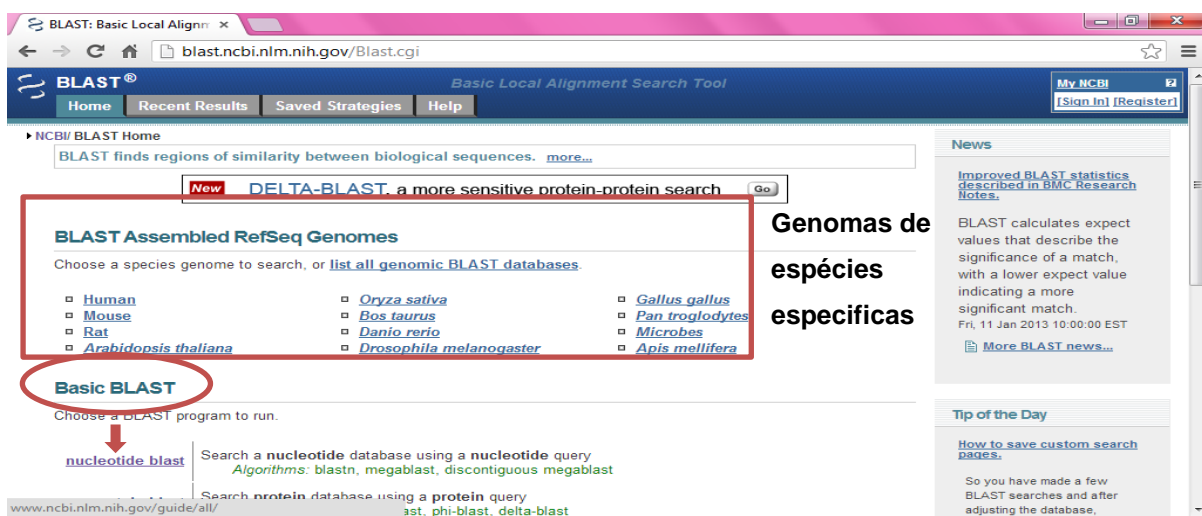


Figura 58. Mostrando as opções de BLAST. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

Ao escolher o tipo de BLAST uma nova página abrirá, na opção “Enter Query Sequence” há um quadro em branco onde você deve digitar a sua sequência, recorde que já havíamos copiado a sequência nucleotídica, basta apenas colar neste espaço apertando as teclas “ctrl+v”. E depois escolha as opções para sua pesquisa, clique em “Others” e selecione a opção “Nucleotide collection (nr/nt)”, no final da página selecione as opções “Show results in a new window” e “BLAST” que iniciará o processamento de busca de dados para ocorrer o alinhamento local (As figuras 59 até a 62, mostram todas essas opções).

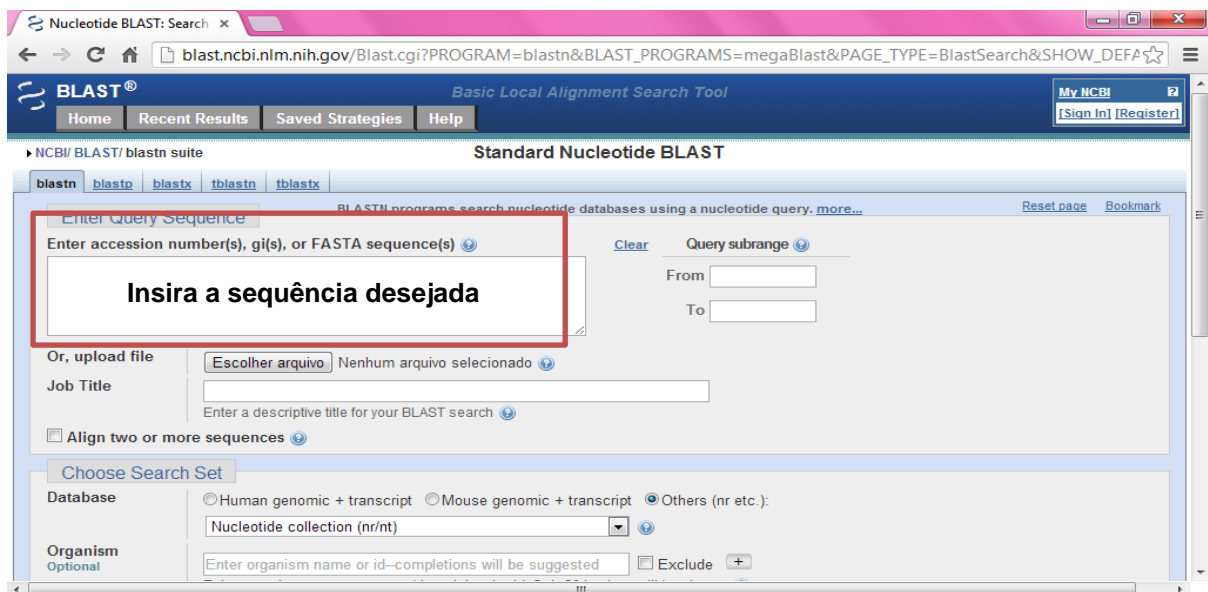


Figura 59. Local de inserção da sequência. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

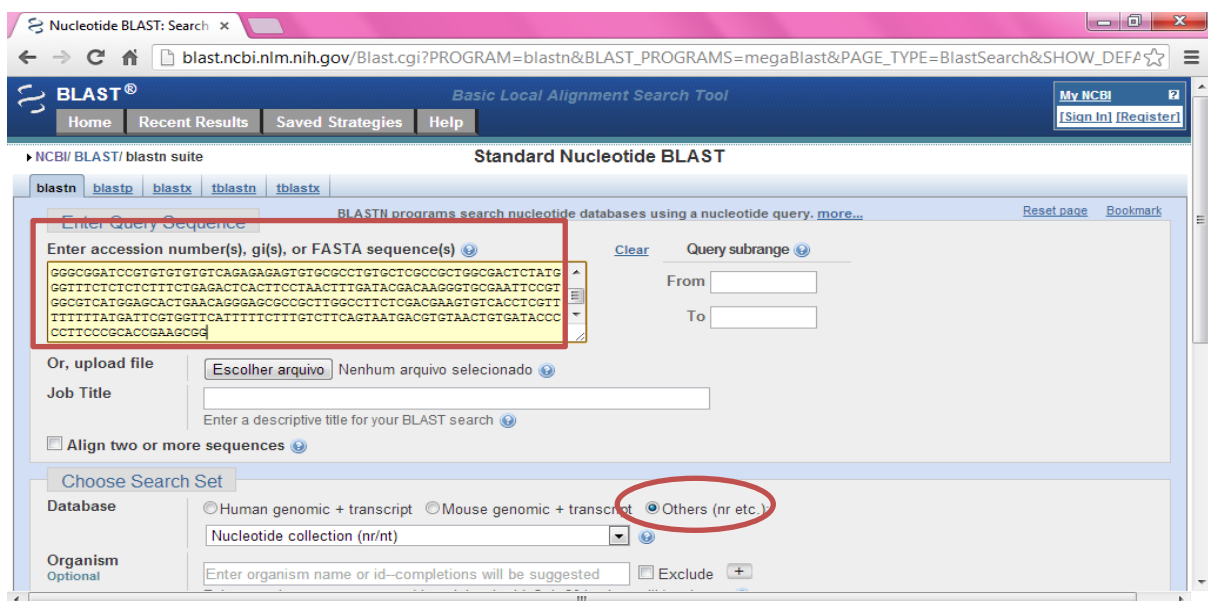


Figura 60. Opções do BLAST. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

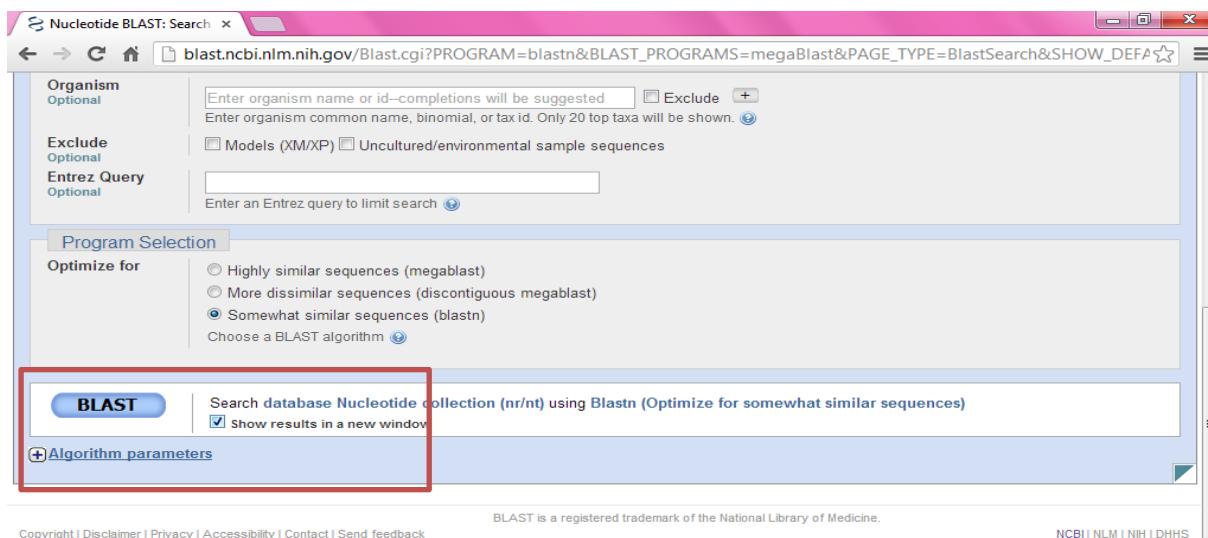


Figura 61. Mais opções para realizar o BLAST. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

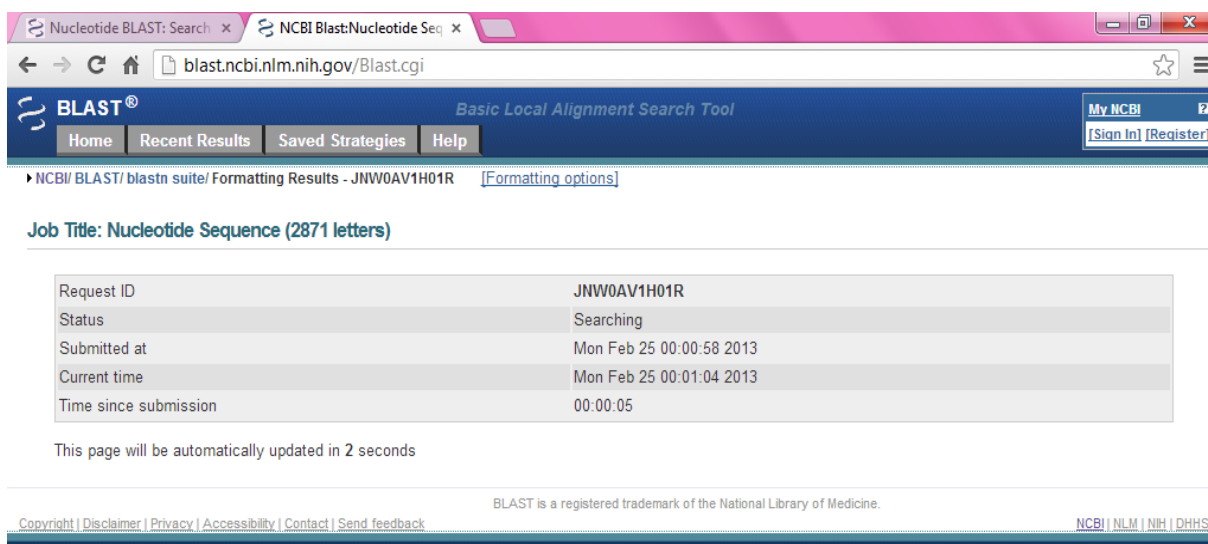


Figura 62. Processado o alinhamento local. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

3º PASSO: Resultados do Alinhamento Local

Após o processamento de busca, o alinhamento local é concluído. Vamos verificar as três opções de resultados promovidos por esse alinhamento, que são o gráfico, uma lista descritiva e as sequências alinhadas, para visualizá-los mova a barra de rolagem. Analisemos primeiro o gráfico, nesse gráfico temos duas escalas uma de cor e outra a régua. Cada cor representa a quantidade de nucleotídeos que são semelhantes entre as sequências e a extensão da barra mostra a posição de bases que são semelhantes nas sequências analisadas.

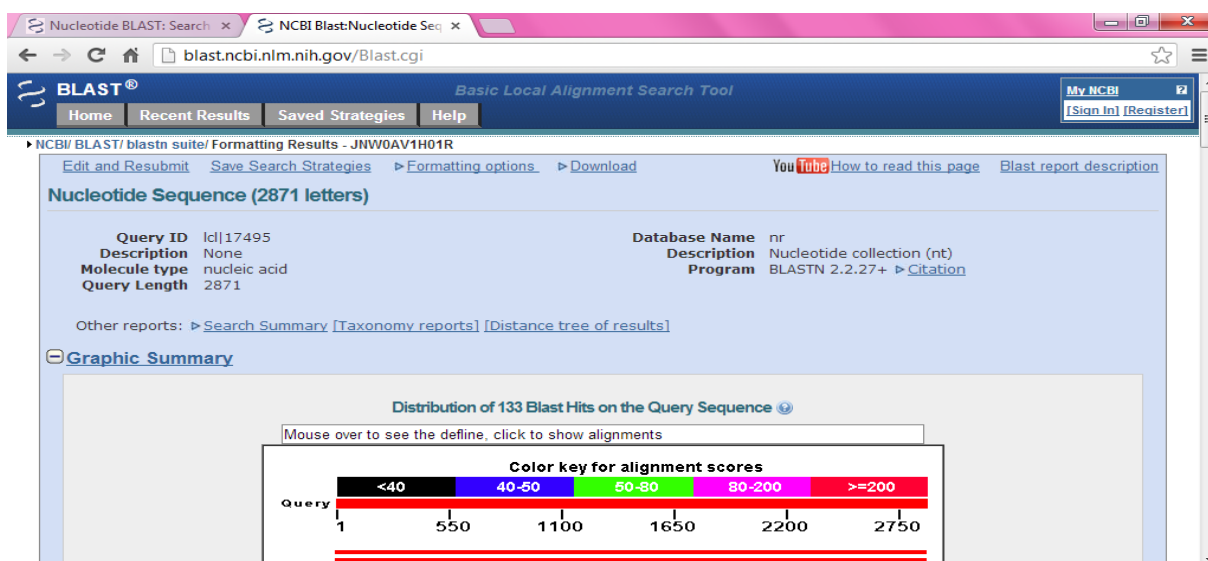


Figura 63. Resultado do Alinhamento local. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

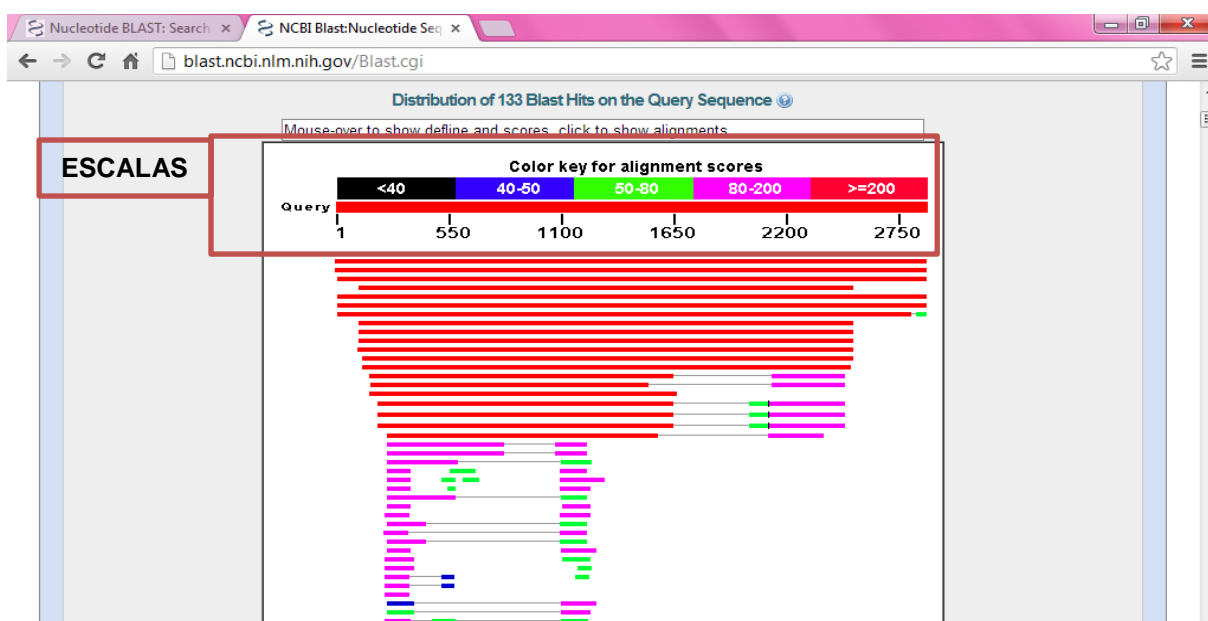
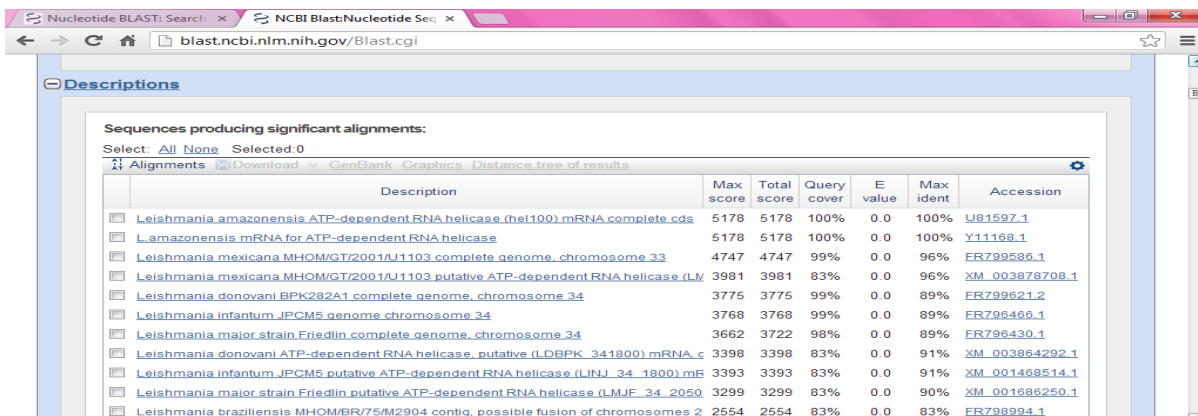


Figura 64. Gráfico do alinhamento local. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

Além do gráfico, tem outra maneira de mostrar o resultado do alinhamento local uma lista descritiva da sequência, que revela o nome da sequência que foi comparada com a sequência de *L. amazonensis*, o item “Max e Total Score” mostram que quanto maior for o score mais as sequências são parecidas, “Query cover” exibe em porcentagem a extensão de semelhança entre as sequências, “E value” expõe a possibilidade das sequências serem parecidas ao acaso, “Max ident” mostra o percentual do grau de semelhança entre as sequências e “Accession” é o nome de acesso a sequência.



Sequences producing significant alignments:

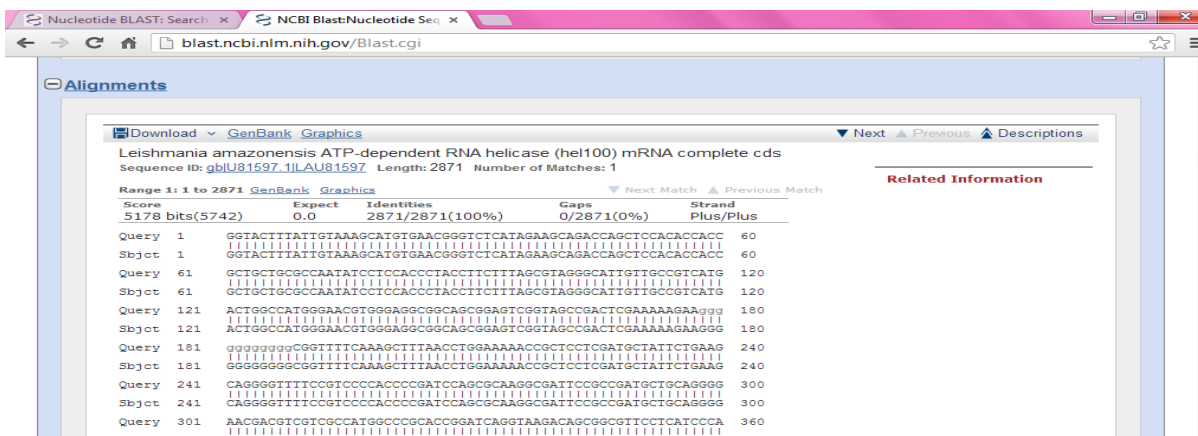
Select: All None Selected: 0

Alignments Download GenBank Graphics Distance trees of results

Description	Max score	Total score	Query cover	E value	Max ident	Accession
Leishmania amazonensis ATP-dependent RNA helicase (hel100) mRNA complete cds	5178	5178	100%	0.0	100%	U81597.1
L.amazonensis mRNA for ATP-dependent RNA helicase	5178	5178	100%	0.0	100%	Y11168.1
Leishmania mexicana MHOM/GT/2001/UJ1103 complete genome_chromosome_33	4747	4747	99%	0.0	96%	FR799586.1
Leishmania mexicana MHOM/GT/2001/UJ1103 putative ATP-dependent RNA helicase (LN	3981	3981	83%	0.0	96%	XM_003878708.1
Leishmania donovani BPK282A1 complete genome_chromosome_34	3775	3775	99%	0.0	89%	FR799521.2
Leishmania infantum JPCMS qenome.chromosome.34	3768	3768	99%	0.0	89%	FR796466.1
Leishmania major strain Friedlin complete genome_chromosome_34	3662	3722	98%	0.0	89%	FR796430.1
Leishmania donovani ATP-dependent RNA helicase putative (LDBPK_341800) mRNA c	3398	3398	83%	0.0	91%	XM_003864292.1
Leishmania infantum JPCMS putative ATP-dependent RNA helicase (LINJ_34_1800) mR	3393	3393	83%	0.0	91%	XM_001468514.1
Leishmania major strain Friedlin putative ATP-dependent RNA helicase (LMJF_34_2050	3299	3299	83%	0.0	90%	XM_001686250.1
Leishmania braziliensis MHOM/BR/75/M2904 contig_possible fusion of chromosomes 2	2554	2554	83%	0.0	83%	FR798994.1

Figura 65. Lista descritiva do alinhamento local. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

A outra maneira de resultado é expor as sequências nucleotídicas alinhadas uma sobre a outra, isto é, a sequência do banco de dados alinhada com a sequência de *L.amazonensis*. Ao mover a barra de rolagem você irá perceber que existirão falhas entre as sequências são os gap, que é inserido quando não há similaridade entre as mesmas, como mostra a figura 67.



Leishmania amazonensis ATP-dependent RNA helicase (hel100) mRNA complete cds
Sequence ID: gblU81597.1|LAU81597 Length: 2871 Number of Matches: 1

Range 1: 1 to 2871 GenBank Graphics

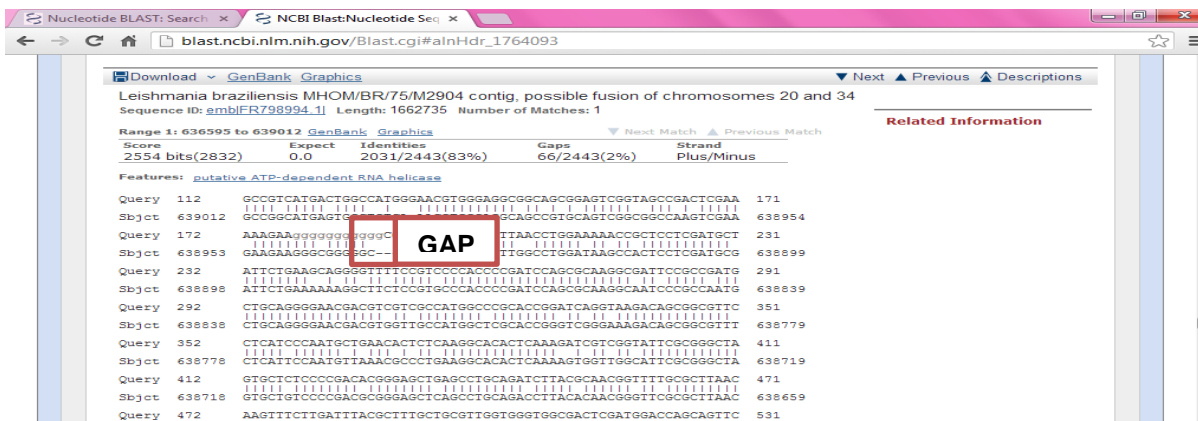
Score	Expect	Identities	Gaps	Strand
5178 bits(5742)	0.0	2871/2871(100%)	0/2871(0%)	Plus/Plus

```

Query 1      GGTACTTATTGTAAGCATGTGAACGGCTCTCATAGAAGCAGACCAGTCCACACCACC
Sbjct 1      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 61     GCTGCTGGCCAAATCTCCACCCCTACCTCTTTAGCGTAGGGCAITTTGCCGTCATG
Sbjct 61     | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 121    ACTGGCCATGGGAACGTGGGAGGCGGCAGCGGATCCGTAGCCGACTCGAAAAAGAGGG
Sbjct 121    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 181    ggggggggCGGTTTTCAAGCTTTAACTGGAAAAACCGCTCCTCGATGCTATTCTGAAG
Sbjct 181    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 241    CAGGGTTTTCCGTCGCCACCCCGATCCAGCGCAAGGCGATTCGCCGATGCTCAGGGG
Sbjct 241    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 301    AACGAGCTCGTCCGATGGCCGACCGGATCAGGTAAGACAGCGCGGTTCCCTCATCCCA
Sbjct 301    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Figura 66. Sequências alinhadas. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acesso em 26/02/13.



Leishmania braziliensis MHOM/BR/75/M2904 contig, possible fusion of chromosomes 20 and 34
Sequence ID: embIFR798994.1 Length: 1662735 Number of Matches: 1

Range 1: 636995 to 639012 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
2554 bits(2832)	0.0	2031/2443(83%)	66/2443(2%)	Plus/Minus

Features: putative ATP-dependent RNA helicase

```

Query 112    GCGGTCATGACTGGCCATGGGAACGTGGGAGGCGGCAGCGGAGTCCGATGCGGACTCGAA
Sbjct 639012  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 172    AAAGAAgggggggggggC
Sbjct 638953  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 232    ATTCTGAAGCAGGSGTTTTCCGTCGCCACCCCGATCCAGCGCAAGGCGATTCGCCGATG
Sbjct 638898  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 292    CTCAGGGAAACGACTCTCCCATGSCCCGACCGGATCAGSTAAACACACGCGHTC
Sbjct 638838  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 352    CTCATCCCAATGCTGAACACTCTCAAGGCACACTCAAAGATGTCGGTATTTCGGGGGCTA
Sbjct 638778  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 412    GTGGCTCCCCGACACGGGAGCTGAGGCTGCAGATCTTACGCCAACGGTTTTGCGCTTAAC
Sbjct 638718  | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Query 472    AAGTTCTTGATTACGCTTTGCTGCGTTGGTGGGTGGCGACTCGATGACACAGCAGTTC
Sbjct 531      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Figura 67. Mostrando os Gap do alinhamento local. Fonte: Site NCBI BLAST <<http://blast.ncbi.nlm.nih.gov/>>, acessado em 26/02/13.

5. 4ª PRÁTICA: MAPA DE RESTRIÇÃO

5.1 Introdução

Enzimas de restrição ou endonucleases de restrição são enzimas bacterianas que cortam pontos específicos da molécula de DNA quebrando as ligações fosfodiéster que ligam um nucleotídeo a outro. As enzimas de restrição são extremamente específicas, pois cada enzima de restrição reconhece e corta apenas uma determinada sequência de nucleotídeos, por exemplo, vamos supor que a enzima de restrição α reconhece a sequência nucleotídica TAGCGT, então ela só irá cortar o DNA nas regiões onde esta combinação de nucleotídeos ocorrerem. As enzimas de restrição constituem a principal ferramenta da tecnologia do DNA recombinante. Graças a elas podemos isolar um gene específico que nos interessa para colocá-lo dentro de outro organismo (DNA Goes to School, 2007).

A cada fragmento obtido após a digestão enzimática, é realizada uma técnica que envolve a separação de fragmentos por eletroforese, a comparação do tamanho dos fragmentos resultantes permite fazer a localização aproximada dos locais de corte, ou seja, elaborar um mapa de restrição. Os mapas de restrição são uma representação dos locais de cortes realizados pelas enzimas de restrição, normalmente esses mapas são utilizados laboratorialmente para clonar uma região de interesse.

Nessa prática iremos construir um mapa de restrição utilizando como exemplo a sequência nucleotídica de *Leishmania infantum*, obtida na prática anterior, além da utilização do programa BioEdit.

5.2 Montagem de um Mapa de Restrição

1º PASSO: Realização de um Mapa de Restrição

Para iniciar a construção do mapa de restrição, primeiro abra a sequência “*Leishmania infantum*.fasta” no programa BioEdit.

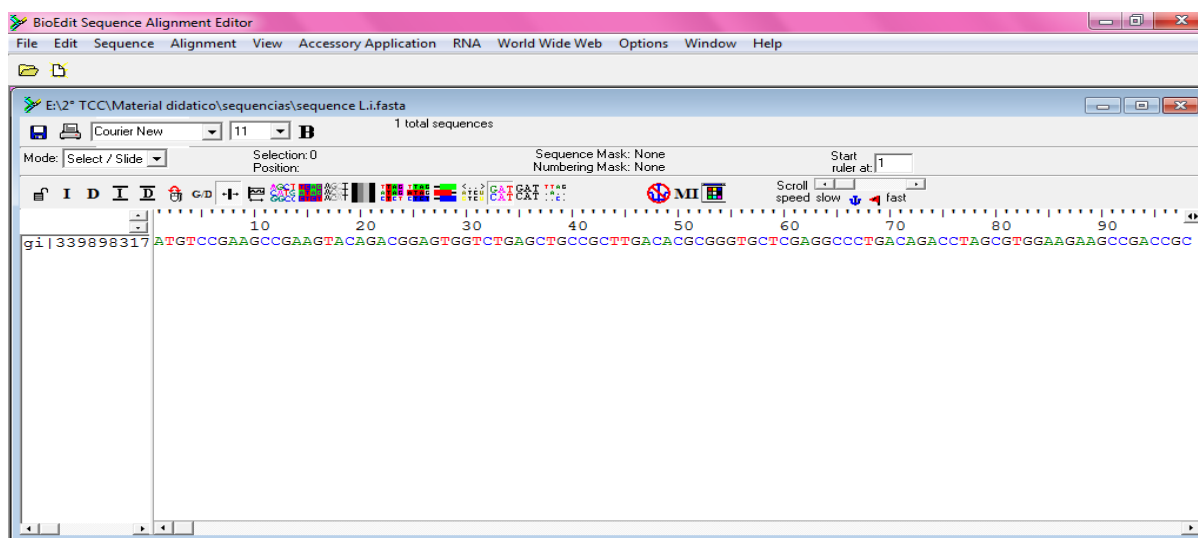


Figura 68. Sequência *L. infantum* aberta no BioEdit. Fonte: Própria Autora, 2013.

Selecione a sequência, em seguida clique no menu “Sequence”, escolha as funções Nucleic Acid e “Restriction Map”. Ao clicar nessas opções abrirá uma nova tela “Create Restriction Map” onde você pode notar que há diferentes funções, iremos verificar os menus que estão demonstrados, através das setas, na figura 70.

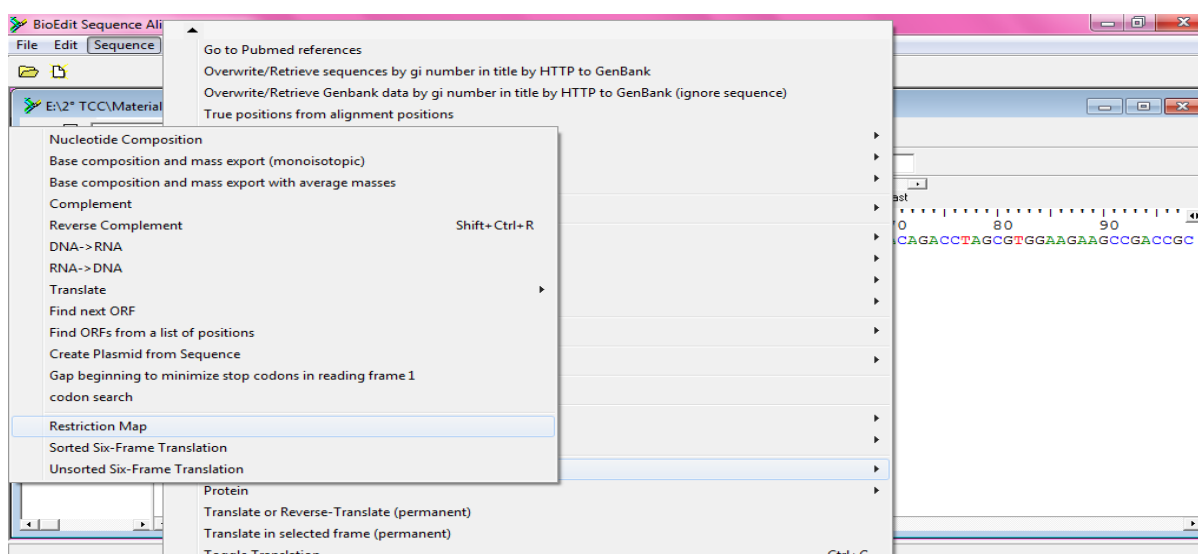


Figura 69. Procedimento para um Mapa de Restrição. Fonte: Própria Autora, 2013.

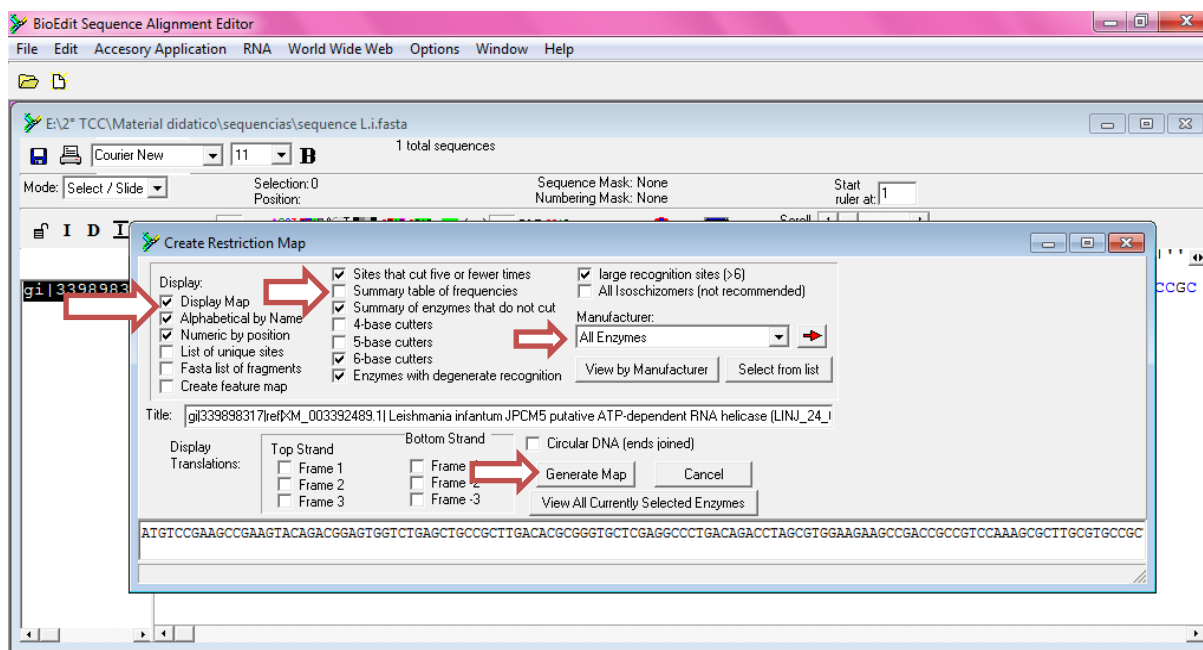


Figura 70. Menus para realizar um Mapa de Restrição. Fonte: Própria Autora, 2013.

2º PASSO: Resultados do Mapa de Restrição

Vamos conferir, então, cada um desse menu e suas respectivas funções.

A função do menu “Display Map” exibe o mapa, as opções “Alphabetical by Name” e “Numeric By position” mostra o nome das enzimas por ordem alfabética e por posição numérica, respectivamente.

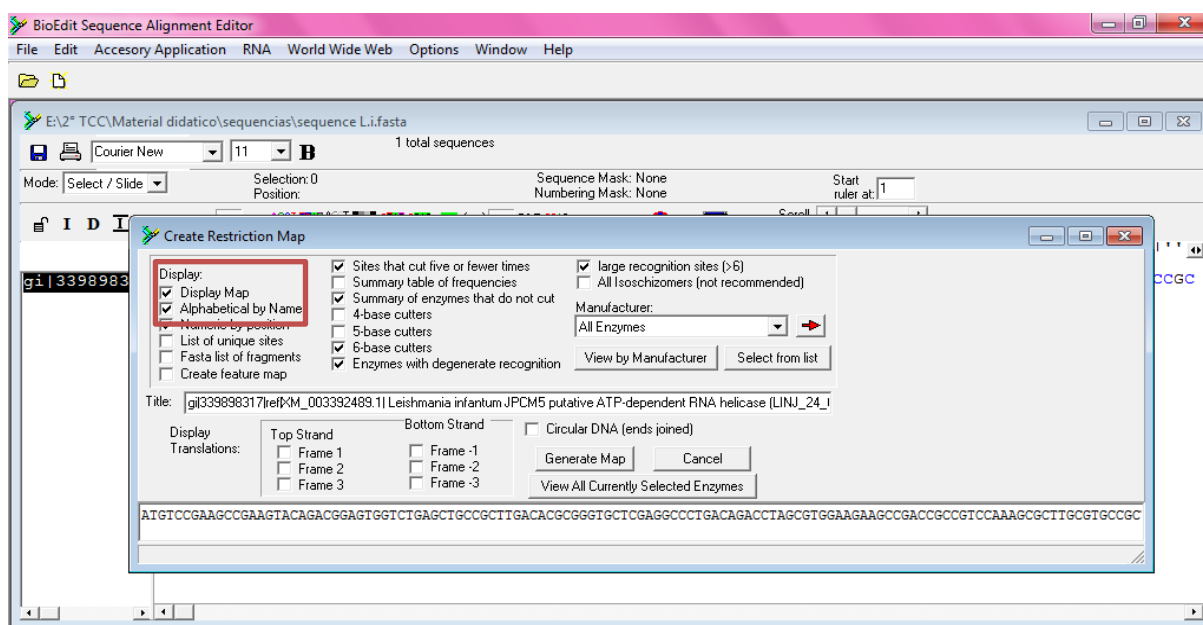


Figura 71. Funções do menu Create Restriction Map. Fonte: Própria Autora, 2013.

E o botão “Sites That cut five or Fewer times” realiza a função de sítios que cortam cinco vezes ou menos um pedaço de DNA, há também a opção “Summary of enzymes that do not cut” que mostra as enzimas que não cortam a sequência. A opção “Manufacturer” permite escolher o fabricante da enzima, selecionaremos a opção “All Enzymes”, ou seja, todas as enzimas que existe no Banco de Dados do programa BioEdit.

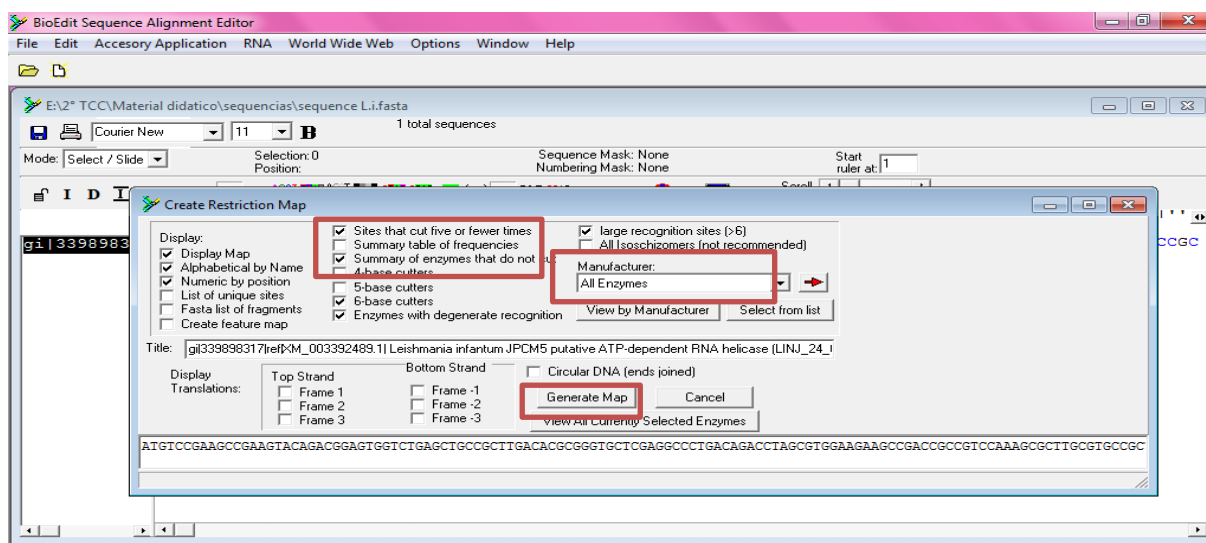


Figura 72. Outras funções do menu Create Restriction Map. Fonte: Própria Autora, 2013.

E a opção “Generate Map” ao clicá-la irá gerar o mapa de restrição, que mostrará os resultados que foram solicitados. Agora clique nesta opção para gerar o mapa de restrição. Vários resultados serão exibidos basta mover a barra de rolagem para visualizá-los.

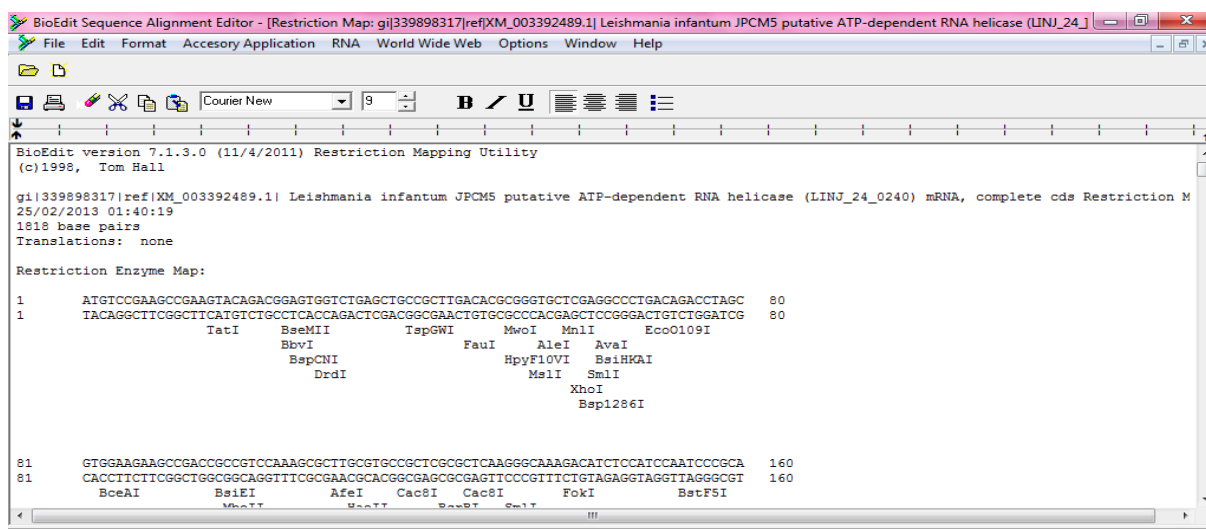
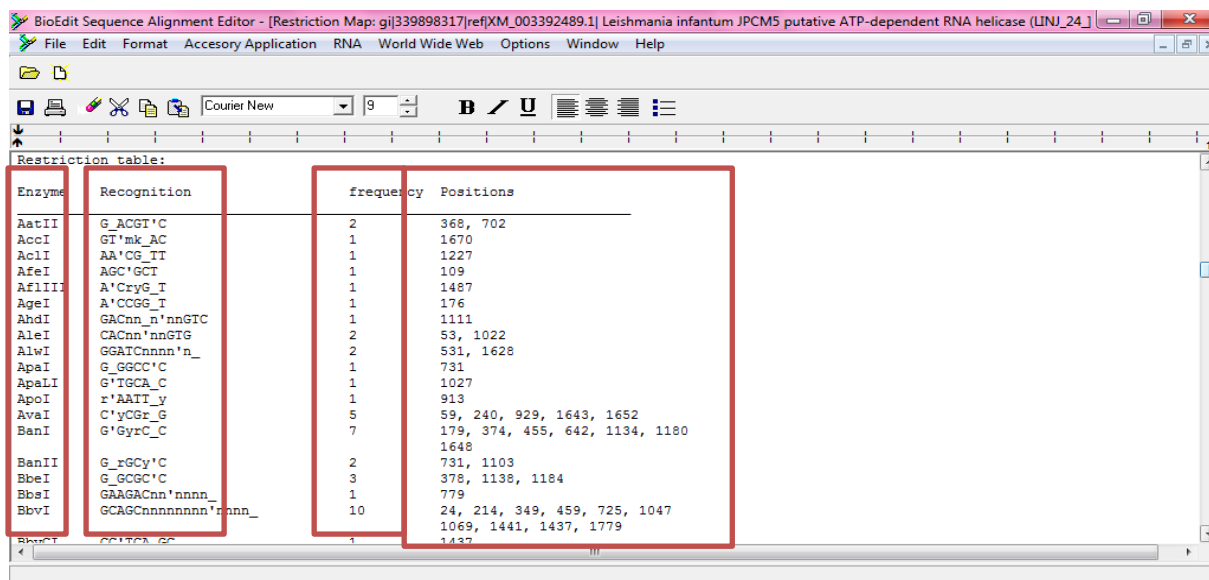


Figura 73. Mapa de Restrição da sequência *L.infantum*. Fonte: Própria Autora, 2013.

O primeiro resultado a ser exibido é uma tabela que relaciona todas as enzimas escolhidas com as bases que elas reconhecem, a frequência que estas seqüências aparecem e a posição que cortam o DNA.

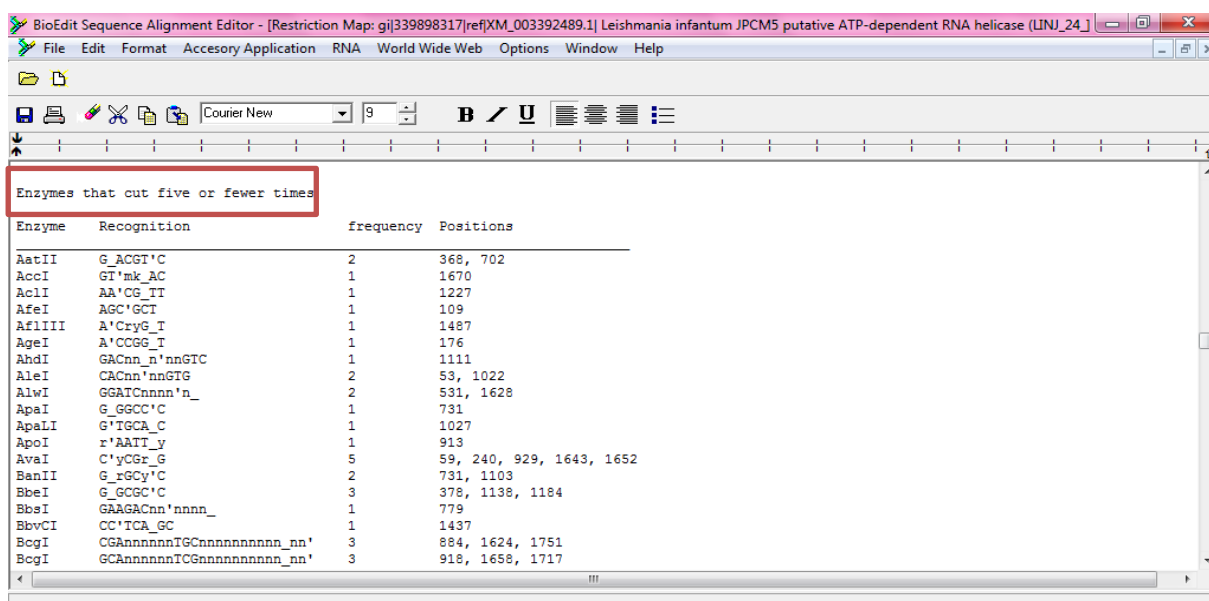


Restriction table:

Enzyme	Recognition	frequency	Positions
AatII	G_ACGT^C	2	368, 702
AccI	GT^mk_AC	1	1670
AclI	AA^CG_TT	1	1227
AfeI	AGC^GCT	1	109
AflIII	A^CryG_I	1	1487
AgeI	A^CCGG_I	1	176
AhdI	GACnn_n^nnGTC	1	1111
AleI	CACnn^nnGTC	2	53, 1022
AlwI	GGATCnnnn^nn_	2	531, 1628
ApaI	G_GGCC^C	1	731
ApaLI	G^TGCA_C	1	1027
ApoI	r^AATT_y	1	913
AvaI	C^yCGr_G	5	59, 240, 929, 1643, 1652
BanI	G^GyrC_C	7	179, 374, 455, 642, 1134, 1180, 1648
BanII	G_rGCy^C	2	731, 1103
BbeI	G_GCGC^C	3	378, 1138, 1184
BbsI	GAGACnn^nnnn_	1	779
BbvI	GCAGCnnnnnnn^nnnn_	10	24, 214, 349, 459, 725, 1047, 1069, 1441, 1437, 1779

Figura 74. Resultado da tabela que mostra a frequência e posição do nucleotídeo. Fonte: Própria Autora, 2013.

Os outros resultados exibidos são uma tabela que mostra as enzimas de restrição que corta até cinco vezes a seqüência nucleotídica. Outra tabela com a posição do nucleotídeo e a respectiva enzima que corta nesse local. E uma lista com as enzimas que não cortam a seqüência.



Enzymes that cut five or fewer times

Enzyme	Recognition	frequency	Positions
AatII	G_ACGT^C	2	368, 702
AccI	GT^mk_AC	1	1670
AclI	AA^CG_TT	1	1227
AfeI	AGC^GCT	1	109
AflIII	A^CryG_I	1	1487
AgeI	A^CCGG_I	1	176
AhdI	GACnn_n^nnGTC	1	1111
AleI	CACnn^nnGTC	2	53, 1022
AlwI	GGATCnnnn^nn_	2	531, 1628
ApaI	G_GGCC^C	1	731
ApaLI	G^TGCA_C	1	1027
ApoI	r^AATT_y	1	913
AvaI	C^yCGr_G	5	59, 240, 929, 1643, 1652
BanII	G_rGCy^C	2	731, 1103
BbeI	G_GCGC^C	3	378, 1138, 1184
BbsI	GAGACnn^nnnn_	1	779
BbvCI	CC^TCA_gc	1	1437
BcgI	CGAnnnnnnIGCnnnnnnnnn^nn'	3	884, 1624, 1751
BcgI	GCAnnnnnnICGnnnnnnnnn^nn'	3	918, 1658, 1717

Figura 75. Resultado da tabela que mostra as enzimas que cortam até cinco vezes. Fonte: Própria Autora, 2013.

Position	Enzyme (s)
16	TatI w'GTAC_w
24	BseMII CTCAGnnnnnnnn'_nn'
24	BbvI GCAGCnnnnnnnn'_nnnnn_
25	BspCNI CTCAGnnnnnnnn'_nn'
28	DrdI GACnn'_nn'nnGTC
38	TspGWI ACGGAnnnnnnnnn'_nn'
45	FauI CCGCnnnn'_nn_
49	MwoI GCnn'_nnn'nnGC
50	HpyF10VI GCn'_nnnnn'nGC
53	AleI CACnn'_nnGTC
53	MslI CAynn'_nnrTG
55	MnlI CCTCnnnnnn'_n'
59	AvaI C'yCGr'_G
59	SmlI C'TyrA'_G
59	XhoI C'TCGA'_G
60	BsiHKAI G_wGCw'C
60	Bsp1286I G_dGCh'C
64	EcoO109I rG'GnC_Cy
83	BceAI ACGGCnnnnnnnnnnnn'_nn_
96	BsiEI CG_ry'CG

Figura 76. Resultado da posição nucleotídica e as enzimas. Fonte: Própria Autora, 2013.

Position	Enzyme (s)
1717	BcuI GCAnnnnnnnnCGnnnnnnnnnn'_nn'
1723	Tth111I GACn'_n_nGTC
1729	HgaI GACGCnnnnnn'_nnnnn_
1739	FspI TGC'GCA
1751	BcgI CGAnnnnnnnTGCnnnnnnnnnn'_nn'
1771	MspA1I CmG'CKG
1774	BslI CCnn'_nnn'nnGG
1776	FauI CCGCnnnn'_nn_
1779	BbvI GCAGCnnnnnnnn'_nnnnn_
1789	BseRI GAGGAGnnnnnnnn'_nn'
1793	EcoS7I CTGAGnnnnnnnnnnnnnn'_nn'
1793	EcoS7MI CTGrAGnnnnnnnnnnnnnn'_nn'
1810	MnlI CCTCnnnnnn'_n'

Enzymes that do not cut:

AarI, Acc65I, AflII, A1oI, A1oI, AlwNI, AscI, AseI, AsiSI, AvrII, BaeI, BaeI
 BamHI, BciVI, BglII, B1pI, EmrI, EmtI, BsaBI, BspEI, BspHI, BsrGI, BssHII, BstBI
 BstEII, BstXI, BstYI, BstZ17I, Bsu36I, ClaI, DraI, EarI, EcoICRI, EcoRI, EcoRV
 Fall, FseI, FspAI, HpaI, KpnI, MfeI, MmeI, NcoI, NdeI, NheI, NotI, PacI, PciI
 PflMI, PmeI, PpiI, PpiI, PstI, PstI, PvuI, RsrII, SacI, SapI, SbfI, SfiI
 SgrAI, SmaI, SnaBI, SpeI, SrfI, SspI, StuI, SwaI, TaqII, TaqII, TspDII, XbaI, XcmI
 XmaI, XmnI

Figura 77. Resultados das enzimas de restrição que não cortam a sequência. Fonte: Própria Autora, 2013.

REFERÊNCIAS BIBLIOGRÁFICAS

ARBEX, W; COSTA, V. M. M. S; SILVA, M. V. G. B. **A Bioinformática como Ferramenta nas Pesquisas Atuais**. Viçosa, agosto 2006.

CARVALHO, J. C. Q; BOSSOLAN, N. R. S. **Algumas concepções de alunos do ensino médio a respeito das proteínas**. Encontro Nacional de Pesquisa em Educação em Ciência / VII Enpec. Florianópolis, 2009.

DEGRAVE, Wim; MACHADO, C. J. S. **Bioinformática e Saúde: Uma Visão Geral**. RECIIS – R. Eletr. de Com. Inf. Inov. Saúde. Rio de Janeiro, v.1, n.2, Sup.1, p.Sup286-Sup287, jul.-dez. 2007.

DNA Goes to School, Inc. **Análise de Fragmentos de Restrição**. Apostila da Oficina Prática de Genética, Genoma e Biotecnologia. Segundo módulo, 2007. Disponível em: <<http://www.odnavaiaescola.com.br/modulo2.pdf>>. Acessado em: 26 fev. 2013.

LENGAUER, T. **Computational Biology at the Beginning of the Post-genomic Era**. Lecture Notes for Computer Science Volume 2000. Informatics: 10 Years Back - 10 Years Ahead. Reinhar Wilhelm (Ed.), Springer. Berlin, p. 341-355 2000. Berlin: University of Bonn, 2001. p. 341-355. *apud* BONGIOLO, Elisângela. **Análise Panorâmica da Bioinformática no Brasil: Propostas da Gestão de Pessoas para os Laboratórios de Pesquisa**. Monografia apresentada à Diretoria de Pós-graduação da Universidade do Extremo Sul Catarinense- UNESC, para a obtenção do título de especialista em MBA Gestão de Pessoas, dezembro de 2006.

NCBI, **National Center for Biotechnology Information. Bioinformatics**. 2006. Disponível em: <<http://www.ncbi.nlm.nih.gov/About/primer/bioinformatics.html>>. Acesso em: 21 fev. 2013.

PROSDOCIMI, Francisco. **Curso online Introdução a Bioinformática**. Biotecnologia Ciência & Desenvolvimento, 2007. Disponível em: <http://www2.bioqmed.ufrj.br/prosdocimi/FProsdocimi07_CursoBioinfo.pdf>. Acessado em: 26 fev. 2013.

PROSDOCIMI, F. *et al.* **Bioinformática: Manual do Usuário**. Revista de Biotecnologia Ciência e Desenvolvimento – n°29, 2003.

QUEIROZ, Alexandre. **Apostila de Introdução a Bioinformática**. Universidade Federal do Rio Grande do Norte, Centro de Biociências, Departamento de Biofísica e Farmacologia, Disciplina de Bioinformática, 2002.

SANTOS, Fabrício R. & QUEIROGA, José Miguel. **Bioinformática aplicada à Genômica**. Manuscritos para capítulo do Biowork IV. Departamento de Biologia Geral, ICB, UFMG. Belo Horizonte, MG, 2003.

SEIBEL, L. F. B, Lemos M., Lifschitz S. **Bancos de Dados de Genoma**. Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro, 2000.

SOARES, Edson. **Profissão do futuro: Bioinformata vive entre bits e células**. 2006. Disponível em: <<http://www.mqservicos.com.br/secao.asp?i=14&c=801>>. Acessado em: 21 fev. 2013.

APÊNDICE B – QUESTIONÁRIO PARA AVALIAÇÃO DO MATERIAL DIDÁTICO**AVALIAÇÃO DO MANUAL DIDÁTICO**

1. Qual nota você daria de 0 (zero) a 10 (dez) para a Apresentação das Palestras?

2. Qual nota você daria de 0 (zero) a 10 (dez) para a Utilidade do Material Didático na utilização das Ferramentas Básicas de Bioinformática?

3. Qual nota você daria de 0 (zero) a 10 (dez) para a Linguagem utilizada no Material Didático?

4. O Material Didático continha todos os passos necessários para realizar as atividades práticas propostas?

() Sim () Não

5. Dê sua opinião em relação ao Material Didático:

Pontos Positivos-

Pontos Negativos-

Críticas-

Sugestões-