

A BIOINFORMÁTICA COMO FERRAMENTA DE FORMAÇÃO DE RECURSOS HUMANOS NO IFRN**A. Q. P. Farias, P. F. S. Chacon, N. R. R. da Silva**Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte – Campus Natal Zona Norte
neyvan.rodrigues@ifrn.edu.br - aqpf@hotmail.com - pablofchacon@hotmail.com

Artigo submetido em outubro/2011 e aceito em janeiro/2013

RESUMO

Experimentos bem sucedidos e amplamente conhecidos, como o Projeto Genoma Humano, confirmam a potencialidade da bioinformática como novo ramo científico, os quais antes seriam inviáveis sem o desenvolvimento de tecnologias para o sequenciamento e análise dos dados obtidos. Nesse preâmbulo, esse trabalho tem como principal objetivo apresentar uma proposta de capacitação dos alunos do IFRN através da proposta de um curso de extensão apresentando ferramentas introdutórias de biologia computacional. A Bioinformática como ramo da ciência tem se mostrado bastante promissora, revelando-se uma nova ferramenta na agregação do conhecimento para a biologia molecular atrelado a informática, essa

integração tem como intuito, decifrar as características mais intrínsecas das moléculas responsáveis pelo regimento da vida na Terra. Em consequência disso, surge à necessidade de um profissional que possua uma formação multidisciplinar, sendo capaz de dominar técnicas de biologia molecular, matemática, estatística, engenharia de software e ciências da computação. Para o alcance do objetivo proposto, este trabalho foi desenvolvido através de aulas teórico-práticas de ferramentas de biologia computacional, aproveitando as capacidades tecno-científicas desenvolvidas no Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Norte, para apuração dos conceitos e características da temática relacionada.

PALAVRAS-CHAVE: bioinformática, recursos humanos, sequenciamento, curso.**BIOINFORMATICS AS A TOOL FOR HUMAN RESOURCE TRAINING IN IFRN****ABSTRACT**

Successful and widely known experiments such as the Human Genome Project, have confirmed the potential of bioinformatics as a new branch of science, achievements that would not be feasible without the development of technologies for sequencing and data analysis. In this preamble, this work has as main objective to present a proposal to form bioinformaticians through an extension course in introductory Biology Computational tools. This science now has been very promising, revealing a new tool in the aggregation of knowledge for molecular biology in order to decipher the intrinsic characteristics of the

molecules responsible for the regiment of life on Earth. As a result, there has been the need for a professional who has a multidisciplinary background, being able to master techniques of molecular biology, mathematics, statistics, software engineering and computer science. To achieve the proposed objective, this work was based on action research, through an extension course, taking advantage of techno-scientific skills developed at the Federal Institute of Education, Science and Technology of Rio Grande do Norte, for investigating the concepts and characteristic of related topics.

KEY-WORDS: bioinformatics, formation, professional, lack, course.

A BIOINFORMÁTICA COMO FERRAMENTA DE FORMAÇÃO DE RECURSOS HUMANOS NO IFRN

INTRODUÇÃO

Desde a descoberta do DNA como a substância relacionada pelo armazenamento da informação genética e elucidação da sua estrutura molecular na década de 50, tal molécula passou a constituir o principal foco de estudo da biologia molecular. O processo de sequenciamento iniciou-se na década de 70 através de um processo manual e rudimentar sendo sequenciada de 10 a 100 bases por vez. A partir da década de 90 vem ocorrendo um desenvolvimento de novas tecnologias de sequenciamento de DNA, contribuindo para o aumento do volume de dados biológicos a serem analisados utilizando ferramentas de bioinformática.

A biologia molecular então progrediu exponencialmente, devido ao melhoramento da automação na geração de dados de sequências de genótipos para fenótipos, tornando-se muito mais uma ciência de informação. Entretanto, de acordo com FILHO *et al.* (2002), esse aumento na quantidade de sequências genéticas a serem armazenadas passou a exigir algoritmos computacionais eficientes que promovessem o compartilhamento, análise e armazenamento desses dados. Segundo EDWARDS *et al.* (2009), o mais importante para os cientistas ao redor do mundo é o fato da análise desses dados estar atrelada a qualidade das sequências e suas anotações mantidas nos bancos de dados públicos. Nesse âmbito surge a bioinformática como uma nova área de estudos e trabalho, que promoverá um meio de conexão entre os dados biológicos e as hipóteses científicas levantadas nas pesquisas ligadas, por exemplo, ao fluxo da informação gênica.

Outra abordagem, segundo POLANSKI *et al.* (2007), sugerem que uma próxima etapa da bioinformática seria o aperfeiçoamento das tecnologias na área da proteômica, permitindo a avaliação da abundância e atividades de milhares de famílias protéicas de uma forma mais eficiente. FILHO *et al.* (2002) definem essa nova etapa da pesquisa biológica como “Era Pós-Genômica”, consistindo na mudança do foco de atenção dos ácidos nucleicos para as proteínas, a fim de adicionar significados às sequências obtidas com os projetos genoma, visando colher os frutos desses projetos através do desenvolvimento e/ou desenho de fármacos e vacinas.

Tendo em vista a abrangência de conhecimentos necessários para o desenvolvimento das pesquisas em bioinformática, essa nova área tem característica de ser multidisciplinar, necessitando da integração de diferentes profissionais, como relatado pela AB3C (Associação Brasileira de Bioinformática e Biologia Computacional). Segundo FILHO *et al.* (2002) a bioinformática envolve as áreas de engenharia de softwares, a matemática, a estatística, a ciência da computação e a biologia molecular (ver Figura 1). Há a necessidade de sistemas computacionais de grande porte bem como profissionais especializados para a manipulação dos mesmos. Com isso surge um novo profissional: o bioinformata. De uma forma resumida, esse profissional deve ser capaz de identificar os problemas biológicos e solucioná-los através do uso de ferramentas computacionais (SOARES, 2006).

inserção desses profissionais ao mercado de trabalho.

2 - METODOLOGIA

Considerando-se o potencial tecno-científico desenvolvido pelos Institutos Federais no que diz respeito à difusão de conhecimentos relacionados principalmente às tecnologias, essa proposta teve como foco a realização no IFRN - Campus Natal Zona Norte de um curso de extensão de introdução a ferramentas de Bioinformática para os alunos do Campus e comunidade acadêmica de outras instituições de ensino superior do estado. A ideia inicial do curso configurou-se na apresentação dos conteúdos e realização de aulas teórico-práticas referentes à bioinformática e no final do curso verificou-se através de questionários de avaliação, o nível de aceitação e aprendizado no que diz respeito às ferramentas de apresentadas no curso. A equipe executora da proposta foi constituída de um professor orientador, doutor na área de ciências biológicas e dois alunos bolsistas de extensão do último ano do curso técnico integrado regular de informática. O intuito ao compor tal equipe é aliar membros ligados à biologia e à informática, visando à permuta de conhecimentos na obtenção do sucesso do projeto.

Para formatação do curso e discussão das principais ferramentas foi proposto um aprofundamento dos alunos bolsistas com relação aos conceitos biológicos relacionados com o conteúdo que seria trabalhado no curso. Sendo assim, realizou-se um processo de treinamento que consistiu em reuniões periódicas entre a equipe, além de pesquisa na literatura relacionada às temáticas apresentadas. Além disso, para o desenvolvimento das aulas teóricas-práticas foi preparado a infra-estrutura necessária, através da instalação do software livre MEGA (KUMAR *et al.*, 2004) nos computadores do laboratório de informática do Campus Natal - Zona Norte. Os algoritmos de alinhamento tais como: CLUSTAL W (THOMPSON *et al.*, 1994), T-Coffee (NOTREDAME *et al.*, 2000), MAFFT (KATOHI *et al.*, 2002) e linguagem Perl (NG *et al.*, 2004) foram utilizados nas aulas de alinhamento de sequências e programação através da utilização de servidores online, não havendo necessidade de instalação desses pacotes de bioinformática.

A divisão dos conteúdos ministrados no curso foi determinada durante a fase de formatação das aulas do curso que procurou privilegiar a homogeneidade dos assuntos a serem tratados abordando-os de forma igualitária, sem focar em nenhum tópico específico. Vislumbrou-se a fidelidade dos mesmos com a temática apresentada e a complementaridade deles, garantindo que haja uma fluência na apresentação das aulas, abordando o tema da forma clara, concisa e objetiva. Desse modo, privilegiando-se os pontos anteriormente descritos, optou-se por dividir a temática do curso como demonstrado (Tabela 1):

Tabela 1: Conteúdo Programático.

TEMA	ASSUNTOS	HORAS/AULA
Palestra do Prof. Valdir Balbino – UFPE	Formação de Recursos Humanos na área de Bioinformática para o mercado de trabalho.	5
Introdução à bioinformática	Definição de bioinformática, como surgiu, áreas de atuação e perspectivas de desenvolvimento.	5
Noções de biologia molecular	Fluxo Gênico, ômicas, PCR, eletroforese, clonagem de genes, enzimas de restrição, sequenciamento de DNA e barcoding.	10
Elaboração de scripts biológicos com a linguagem PERL	Conceitos básicos acerca da aplicação da linguagem para problemas biológicos, tipos de variáveis, estruturas de controle de fluxo e de iteração, sub-rotina, funções da linguagem, expressões regulares e desenvolvimento de scripts aplicados.	15
Utilização de bancos de dados biológicos	Noções gerais sobre bancos de dados biológicos (conceito, tipos e contexto de surgimento), NCBI (GenBank e Blast), Uniprot/Swiss-Prot, KEGG e PDB.	10
Noções de montagem e anotação gênica	Análise de cromatogramas (PHRED), softwares para montagem (PHRAP e CAP3), análise de contigs (Consed, VecScreen, ORF Finder) e anotação a nível de nucleotídeos, proteínas e processos biológicos.	5
Alinhamento de sequências gênicas e protéicas	Definição de alinhamento, porque alinhar, tipos de alinhamento e algoritmos (CLUSTAL, POY, T-COFFEE e MAFFT) e softwares de alinhamento (MEGA, BioEdit, Mesquite).	10

O cronograma de aulas definido procurou atender a todos os interessados, de forma que o curso teve duração de duas semanas totalizando uma carga horária de 60 horas de aula.

Ao fim do curso, foi aplicado um questionário para os alunos do curso, no intuito, de avaliar o desempenho da equipe com relação aos conteúdos ministrados, juntamente com o grau de satisfação, aprendizado e interesse em seguir na área de bioinformática.

3 - ANÁLISE E INTERPRETAÇÃO DOS DADOS

Uma vez concluído o período de aulas do curso e aplicado o questionário avaliativo para todos os alunos inscritos, foi possível determinar os pontos relevantes observados e passíveis de análise. Primeiramente verificou-se no ato das inscrições – nas quais houve uma sondagem para verificação da origem dos candidatos – a presença de alunos tanto ligados às ciências biológicas, graduados ou frequentando algum curso nessa área nas instituições de ensino superior como a UFRN; quanto de discentes do IFRN em cursos técnicos ou de licenciatura voltados à informática e eletrônica. Tal fato evidencia a heterogeneidade do público inscrito, comprovando os dados relatados anteriormente no que diz respeito à abrangência da bioinformática.

Por se tratar de um campo recente e até agora pouco explorado, percebeu-se logo na primeira aula que a maioria dos alunos ainda não tinha conhecimento sobre a área de atuação da bioinformática, na resolução dos problemas biológicos com os quais a ciência depara-se nas pesquisas biotecnológicas. A partir dos diálogos em sala de aula, percebeu-se que para muitos alunos do curso tiveram um primeiro contato com a área, até então desconhecida, e esse fato despertou o interesse em seguir na área. Tal fato comprova a importância de ministrar um curso com essa temática, principalmente em um cenário onde ela ainda é pouco divulgada e concomitantemente enfrenta a escassez de profissionais especializados.

Como era esperado, percebeu-se uma maior facilidade dos alunos em assimilar os assuntos da bioinformática ligados as suas respectivas áreas de estudo e/ou atuação. Alunos de informática do IFRN identificavam detalhes acerca do funcionamento dos scripts biológicos apresentados nas aulas práticas em linguagem Perl, enquanto os alunos de biologia pertencentes a instituições de ensino superior encontravam aplicações práticas em seus campos de atuação através dos softwares de bioinformática. Foi evidenciado a partir das respostas ao questionário aplicado ao fim do curso, o grau de satisfação dos alunos com os tópicos apresentados no curso e o desejo de seguir carreira na área ou de ao menos utilizarem as ferramentas disponibilizadas em suas práticas profissionais (Figura 2).

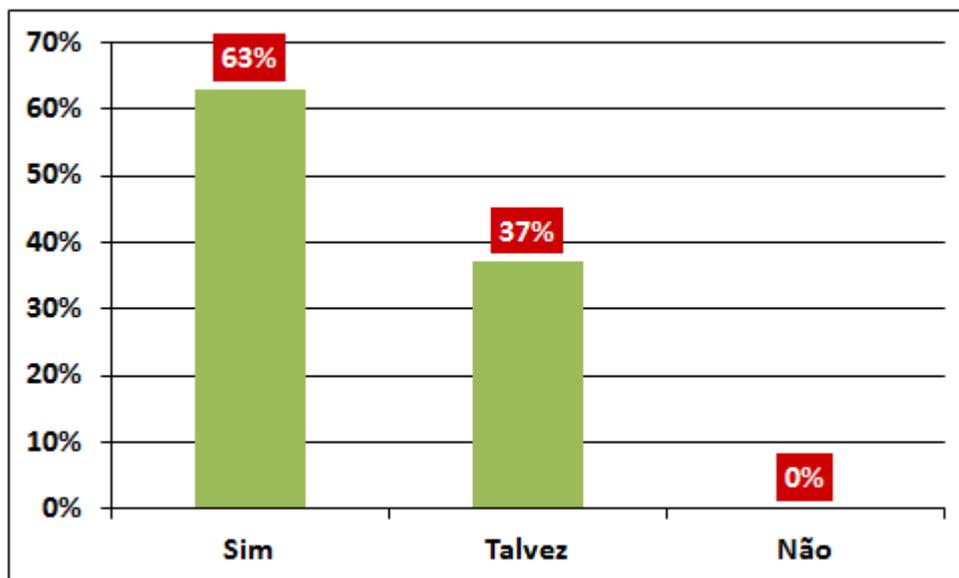


Figura 2 - Gráfico mostrando a pretensão dos alunos do curso em seguir na bioinformática ou de utilizar ferramentas da área no campo profissional.

A partir do contato com os participantes do curso durante as aulas já lecionadas e através do questionário aplicado, percebe-se o foco de utilização dos conhecimentos adquiridos de cada um dos dois grupos de alunos: “biólogos” e “informatas”. Observa-se nos discentes ligados à biologia o interesse em aplicar a experiência obtida no manuseio de ferramentas computacionais na elucidação de problemas biológicos com os quais se deparam no seu cotidiano, enquanto os estudantes de informática passaram a vislumbrar um novo nicho de mercado no qual poderão utilizar os seus conhecimentos e habilidades computacionais e desenvolvimento e implementação de códigos e softwares de caráter biológico.

No que tange as características do curso oferecido, pode-se evidenciar com base nas informações obtidas após a análise dos questionários a satisfação da maioria dos discentes com relação à estrutura dos tópicos e aulas práticas-teóricas, à equipe organizadora, aos conteúdos e práticas apresentados (Figuras 3, 4 e 5).

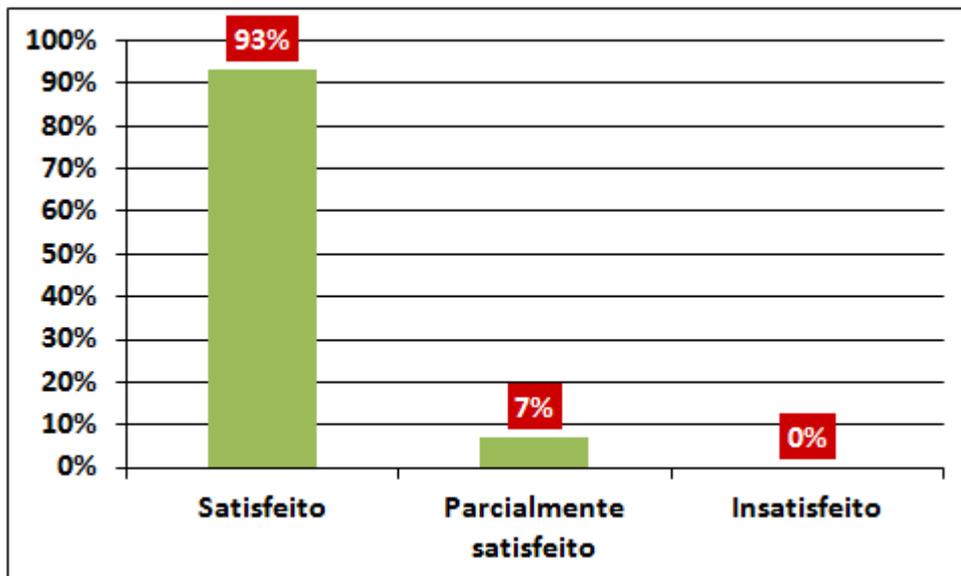


Figura 3 - Gráfico mostrando o nível de satisfação dos alunos quanto à organização do curso.

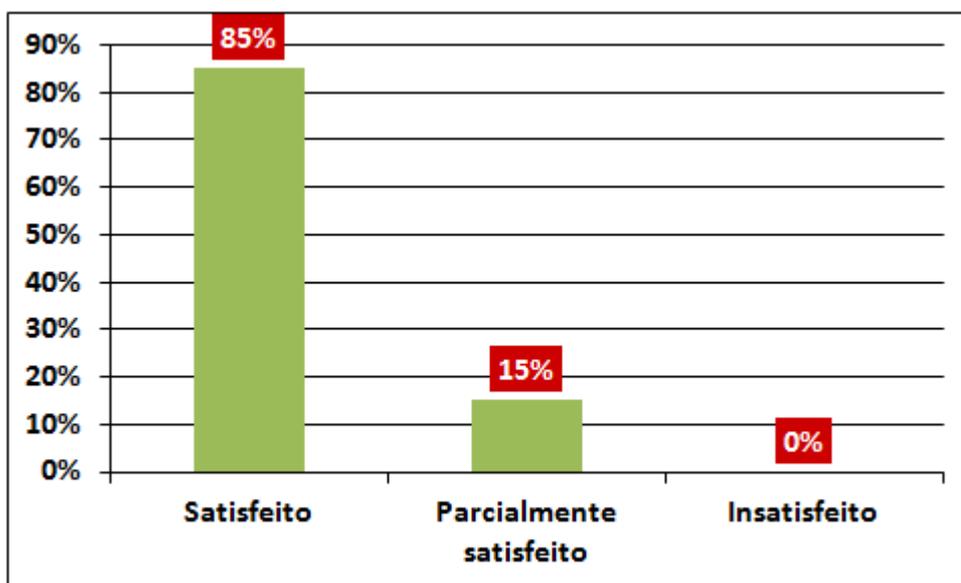


Figura 4 - Gráfico mostrando o nível de satisfação dos alunos quanto à abordagem dos conteúdos apresentados no curso.

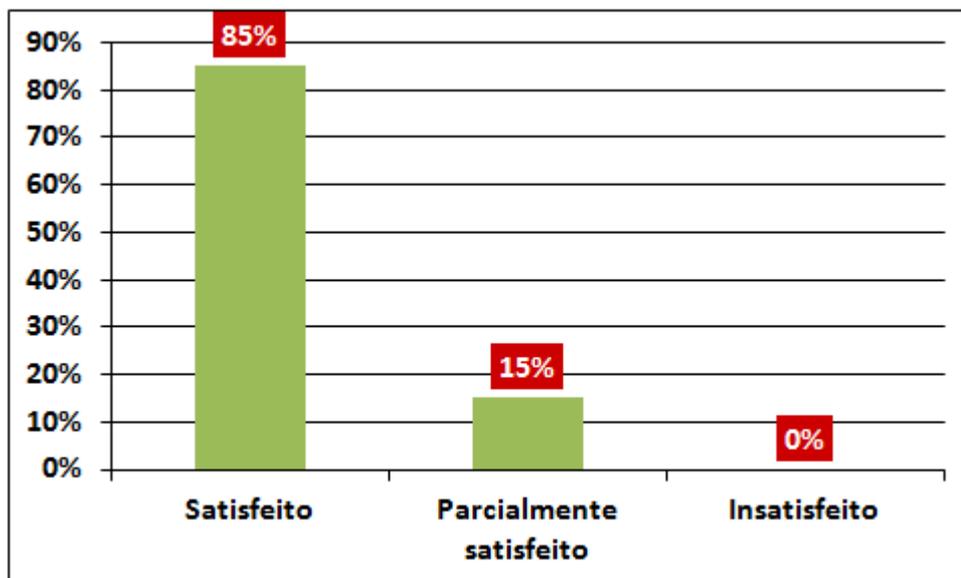


Figura 5 - Gráfico mostrando o nível de satisfação dos alunos quanto às práticas do curso.

Fica evidente dessa forma o sucesso do curso desenvolvido observando-se a boa aceitação e assimilação do conteúdo lecionado. Com isso, pode-se perceber que o projeto se mostrou favorável à ampliação dos conhecimentos e experiências dos alunos, contribuindo para que se tornem profissionais capacitados em cada uma de suas áreas e, provavelmente, futuros bioinformatas.

4 - CONSIDERAÇÕES FINAIS

Tomando-se como base os dados aqui gerados nesse projeto e a discussão da equipe organizadora com os alunos do curso, nota-se claramente o entusiasmo dos discentes tanto ligados à biologia quanto à informática de adentrarem nesse novo ramo da pesquisa científica. Assim, ressalta-se a importância da divulgação de ferramentas em bioinformática na formação de profissionais que possibilitarão um novo caminho para comprovação de hipóteses biológicas ou mesmo a aplicação de algoritmos e técnicas computacionais voltados às mais diferentes disciplinas como a biologia molecular, medicina clínica, farmacologia, biotecnologia, ciência forense e antropologia (LESK, 2008).

Dessa forma, a realização do I Curso de Introdução a Ferramentas de Bioinformática possibilitou a obtenção dos objetivos pretendidos, por meio da disponibilização de um canal eficiente de difusão de conhecimentos na temática abordada, o qual espera-se que seja apenas um estopim na formação de profissionais gradativamente mais completos e especializados nessa nova área de estudo. Faz-se relevante destacar a importância dessa temática de trabalho para o IFRN, o qual beneficia-se através do incentivo à propagação de novas habilidades para os discentes, que associem o conhecimento trabalhado no ensino médio com as competências tecnológicas moldadas principalmente nos cursos voltados à informática e eletrônica.

É imprescindível que iniciativas como essa sejam comumente adotadas pelos institutos federais como por outras instituições de ensino, visando à especialização profissional no país, sobretudo em áreas onde a mão de obra é escassa e necessária como a bioinformática. Nesse cenário, vale salientar iniciativas semelhantes à que foi apresentada como a do professor Dr. Francisco Prosdocini, um dos pioneiros na divulgação no país da temática abordada através da elaboração de um curso *on line* ministrado em 2007. PROSDOCIMI (2007) afirma que sua ideia representou uma iniciativa produtiva e proveitosa, tendo os alunos adquiridos conteúdo e aprendido a compreender muito sobre a lógica e o pensamento em bioinformática atrelada as questões ou problemáticas biológicas no meio acadêmico.

Através do trabalho realizado fica clara a importância do incentivo e manutenção das práticas de pesquisa em bioinformática no instituto, inicializadas através do curso oferecido, capazes de despertar nos alunos o interesse na complementação dos conhecimentos adquiridos no nosso curso de extensão. Sendo assim, acredita-se na promissora possibilidade futura de ampliação dessa iniciativa através da consolidação de um núcleo de pesquisas em bioinformática no campus do IFRN Natal Zona Norte, sustentada através de uma cooperação entre o IFRN e a Universidade Federal do Rio Grande do Norte. Desse modo, espera-se expandir gradativamente essa ramificação da pesquisa científica, contribuindo para divulgar os conhecimentos nessa área inserindo raízes de propagação do meio acadêmico através da ampliação de outros âmbitos do desenvolvimento humano.

REFERÊNCIAS BIBLIOGRÁFICAS

1. ASSOCIAÇÃO BRASILEIRA DE BIOINFORMÁTICA E BIOLOGIA COMPUTACIONAL. **Incentivo a Compreensão da importância da Biologia Computacional no Brasil e na América Latina**. Disponível em: <<http://www.ab3c.org/>>. Acesso em: 05 set. 2011, 17:00.
2. BENTON, D. **Bioinformatics**: Principles and potential of a new multidisciplinary tool. [S.l.: s.n.], 1996, p. 261-272. (Trends Biotechnol, v. 14).
3. EDWARDS, D., STAJICH, J., HASEN, D. **Bioinformatics**: Tools and Applications. New York: Springer, 2009.
4. FILHO, F.C. *et al.* Bioinformática: Manual do Usuário. **Biotecnologia Ciência e Desenvolvimento**, Brasília, v. 5, n. 29, p. 12-25, dez./jan. 2002. Disponível em: <<http://www.biotecnologia.com.br/revista/bio29/bioinf.pdf>>. Acesso em: 04 set. 2011, 15:00.
5. KATO, K. *et al.* MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. **Nucleic Acids Research**, v. 30, n. 14, p. 3059-3066, 2002. Disponível em: <<http://nar.oxfordjournals.org/content/30/14/3059.full.pdf>>. Acesso em: 24 set. 2011, 01:00.
6. KUMAR, S., TAMURA, K., NEI, M. MEGA3: Integrated Software for Molecular Evolutionary Analysis and Sequence Alignment. **Briefings in Bioinformatics**, v. 5, n. 2, p. 150-163, jun. 2004. Disponível em: <http://www.kumarlab.net/pdf_new/KumarTamura03.pdf>. Acesso em: 23 set. 2011, 23:30.

7. LESK, A.M. **Introdução à Bioinformática**. 2. ed. São Paulo: Artmed, 2008.
8. NG, G., SARATHY, G., ING-SIMMONS, N. **Perl Documentation**. Dez. 2004. Disponível em: <<http://perl.enstimac.fr/perl-all-en-5.8.5.pdf>>. Acesso em: 24 set. 2011, 02:00.
9. NOTREDAME, C., HIGGINS, D.G., HERINGA, J. T-Coffee: A Novel Method for Fast and Accurate Multiple Sequence Alignment. **J. Mol. Biol**, p. 205-217, 2000. Disponível em: <<http://www.tcoffee.org/Publications/Pdf/tcoffee.pdf>>. Acesso em: 24 set. 2011, 00:30.
10. POLANSKI, A., KIMMEL, M. **Bioinformatics**. New York: Springer, 2007.
11. PRODOCIMI, F. **Curso On Line: Introdução à Bioinformática**. [S.l.]: Portal Biotecnologia, 2007. Disponível em: <http://www2.bioqmed.ufrj.br/prosdocimi/FProsdocimi07_CursoBioinfo.pdf>. Acesso em: 09 set. 2011, 16:00.
12. SOARES, E. **Profissão do futuro: bioinformata vive entre bits e células**. [S.l.]: Portal IDGNow, 2006. Disponível em: <<http://idgnow.uol.com.br/carreira/2006/05/02/idgnoticia.2006-04-28.1060287223/>>. Acesso em: 05 set. 2011, 17:30.
13. THOMPSON, J.D., HIGGINS, D.G., GIBSON, T.J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. **Nucleic Acids Research**, v. 22, n. 22, p. 4673-4680, 1994. Disponível em: <<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC308517/pdf/nar00046-0131.pdf>>. Acesso em: 24 set. 2011, 01:30.