

**UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE BIOCÊNCIAS
GRADUAÇÃO EM CIÊNCIAS BIOLÓGICAS
DEPARTAMENTO DE GENÉTICA E BIOLOGIA EVOLUTIVA**

**BIO0307
BIOLOGIA MOLECULAR PARA BACHARELADO**

<p>ROTEIRO DE AULAS PRÁTICAS</p> <p>ANÁLISES COMPUTACIONAIS DE SEQUÊNCIAS DE ÁCIDOS NUCLEÍCOS</p>

4ª Edição

PROFESSORES RESPONSÁVEIS:

Profa. Dra. Luciana Amaral Haddad

Prof. Dr. Eduardo Gorab

Profa. Dra. Regina Célia Mingroni Netto

**SÃO PAULO
2017**

Aula Prática 1 – Análise computacional de sequências de ácidos nucleicos

Objetivos: Revisão sobre a reação em cadeia da polimerase (PCR). Desenho de iniciadores no NCBI para reação em cadeia da polimerase (PCR) e conhecimento dos bancos de dados Herança Mendeliana no Homem *Online* (OMIM) e *Medical Genetics* (MedGen).

Estratégia: Desenhar iniciadores para identificação pela PCR e sequenciamento da mutação $\Delta F508$ no gene *CFTR* (do inglês, *Cystic Fibrosis Transmembrane conductance Regulator*).

- (1) A fibrose cística é uma doença de herança autossômica recessiva, causada por mutações no gene *CFTR*.
- (2) A forma mais simples e imediata para obter informações concisas sobre uma doença genética é hoje no banco de dados *Medical Genetics* (MedGen) do NCBI (*National Center for Biotechnology Information*, Bethesda, MD). O NCBI fornece acesso a informações genômicas e biomédicas.
- (3) Entrando no site MedGen (<http://www.ncbi.nlm.nih.gov/medgen>), uma busca por 'cystic fibrosis' recupera no primeiro resultado um resumo sobre a doença. Clique neste resultado. Informações mais aprofundadas podem ser encontradas nos bancos Gene Reviews e OMIM.
- (4) No NCBI, o site OMIM (do inglês, *Online Mendelian Inheritance in Man*) cita brevemente algumas informações disponíveis sobre todas as doenças genéticas humanas conhecidas, recuperadas da literatura científica no *pubmed*. Uma vez que o OMIM foi construído a partir do *pubmed* de forma manual e não compreensiva, o seu texto não tem o objetivo de ser completo embora tenda a conter as primeiras informações de associação de um gene a um fenótipo.
- (5) Verifique algumas informações sobre a fibrose cística no OMIM, para o qual há um *link* nesta mesma página: OMIM 219700.
- (6) Cerca de 70% dos pacientes com fibrose cística, em populações Caucásicas, tem a mutação $\Delta F508$, uma deleção de três bases levando à deleção de um resíduo de fenilalanina na posição 508 da proteína CFTR. Esses dados podem ser observados se voltarmos a página e mudarmos para a *database* 'Clinical Variants – ClinVar' no mesmo menu. **ClinVar** é o banco de dados do NCBI que agrega informações sobre variantes de sequência do genoma humano e sua relação com a saúde humana. Faça uma busca por 'cystic fibrosis deltaF508'. Clique em 'Search'.
- (7) Clique sobre a mutação p.Phe508DelPhe e observe na nova página alguns aspectos importantes:
 - a. 'Clinical Significance': pathogenic
 - b. (Clinical) 'Condition': cystic fibrosis
 - c. 'Variant type': deletion
 - d. 'Genomic location': 7q31.2
 - e. Protein change: F508DelF

- f. *Links* para OMIM (descrito acima), dbSNP (banco de dados de polimorfismos de base única), 1000genomes (aula P6).
- (8) Tendo conhecido algumas portas de entrada no NCBI para obter informação sobre a doença, iremos agora desenhar um par de iniciadores para amplificar a sequência de DNA do **éxon 10 do CFTR** que contém a mutação que estudaremos.
- (9) A descrição **p.Phe508delPhe** é a denominação correta desta mutação em relação à **sequência proteica (p.)** da CFTR, definindo a **deleção da fenilalanina em sua posição 508**. Em relação à **sequência de código (c.) do gene**, define-se corretamente a mutação como **c.1521-1523delCTT**, i.e, a **deleção de três bases, CTT**, nas posições **1521 a 1523** da sequência de código do gene *CFTR*.
- (10) Para amplificação pela reação em cadeia da polimerase (PCR), os iniciadores deverão hibridar com as sequências intrônicas adjacentes ao éxon 10, isto é, os íntrons 9 e 10.
- (11) Abra o arquivo *Word* enviado por email, contendo a sequência do éxon 10 do *CFTR* e regiões intrônicas adjacentes. Copie a sequência incluindo seu cabeçalho.
- (12) Para o desenho dos iniciadores, utilizaremos o *primer3*, um programa de computador que busca identificar em uma sequência de DNA curtos trechos de fita simples com características que os definam como bons iniciadores para a PCR. Assim, os resultados vêm como pares de iniciadores em ordem decrescente de acordo com a adequação aos critérios adotados. Estes critérios compreendem:
- temperatura de fusão (do inglês, *melting*) em torno de 60°C (tamanho do oligonucleotídeo entre 18 e 25 bases);
 - produto de PCR entre 200 e 1000 pb;
 - inexistência de auto-complementariedade;
 - inexistência de complementariedade com seu par;
 - presença preferencial de C ou G nas duas últimas bases a 3';
 - 40 a 60% da sequência de cada iniciador com conteúdo de [C+G]; e
 - sequência a ser amplificada (amplicom) com 40 a 60% de [C+G].
- (13) O *site* do NCBI usa o *Primer3* para o desenho de iniciadores e então os submete a análise pelo BLAST com um banco de sequências definido pelo usuário. O objetivo do BLAST após o desenho dos iniciadores é verificar se estes podem identificar por complementariedade mais de um alvo no genoma do organismo selecionado, o que não é desejável, pois queremos especificidade à PCR.
- (14) Para acessar o *site* Primer-blast, entre em <http://www.ncbi.nlm.nih.gov/tools/primer-blast/index.cgi>
- (15) O *site* aconselha usar uma RefSeq (aula P2) como molde para a PCR sempre que possível, porque auxilia o programa a identificar sequências inespecíficas, tendo-a como padrão. No entanto, hoje usaremos a nossa sequência do arquivo *Word*. Cole-a no local indicado.

- (16) Determine o tamanho desejado para sequência a ser amplificada (amplicom, '*PCR product size*') que, para nós, será entre 300 e 700 pb.
- (17) A temperatura de fusão (T_m) será 60°C (opt), variando entre 57°C e 63°C. O máximo de diferença de T_m entre os dois iniciadores será de 3°C.
- (18) Escolha um banco de sequências não-redundantes, no nosso caso, '*Genome (reference)*' e '*Homo sapiens*' como organismo.
- (19) Clique em '*Get primers*'. Aguarde pelos resultados.

(20) Uma observação importante é que há **486 pb** de sequência do **íntron 9** e **1.323 pb** do **íntron 10** no arquivo. **O tamanho do éxon 10 é de 192 pb.**

- (21) Observe os resultados. O que você identifica? Escolha um par de iniciadores.
- (22) Com a sua escolha, uma vez sintetizados os oligonucleotídeos e padronizada a reação da PCR, como você poderá observar se há a mutação **c.1521-1523delCTT** ($\Delta F508$)?
- (23) Para restringir as alternativas, você pode definir que o primeiro iniciador deverá estar entre a base 1 e a 486 (íntron 9). O segundo iniciador deverá ser no íntron 10 (da base 679 à 2001). Mantenha os parâmetros anteriores e clique em '*Get primers*'. Aguarde.
- (24) Um banco de dados de mutações do gene *CFTR*, relacionadas à fibrose cística, está no *site* do *Clinical and Functional Translation of CFTR* (CFTR2) e tem o objetivo de facilitar a comunicação entre pesquisadores, acelerar o processo de triagem de mutações no gene e uniformizar o formato para os relatos de mutações e sua localização precisa. Para acessá-lo, entre em <http://www.cftr2.org>. Você deverá concordar com os termos de uso do *site*.
- (25) NO menu, haverá '*Variant List History*'. Na página que se abriu, selecione '*CFTR2_8August2016.xlsx*'. No arquivo excel que se abre, observe a diversidade de mutações causadoras de fibrose cística. Busque pela mutação do nosso estudo digitando F508. Compare sua frequência na população (coluna F) à de outras mutações.

Curiosidades

- (1) Como você acha que a fibrose cística é triada pelo teste do pezinho no recém-nascido?
- (2) Uma gestante deseja realizar diagnóstico pré-natal para fibrose cística. Ela e o marido são primos e têm, em sua família, vários afetados pela doença e alta taxa de consanguinidade. Em testes genéticos realizados, não se encontrou mutação no gene *CFTR* dos familiares afetados, mas o diagnóstico clínico é evidente. Como você poderia fazer um diagnóstico pré-natal e aconselhamento genético para o casal?

Referências bibliográficas

Steve Rozen and Helen J. Skaletsky (2000) Primer3 on the WWW for general users and for biologist programmers. In: Krawetz S, Misener S (eds) *Bioinformatics Methods and Protocols: Methods in Molecular Biology*. Humana Press, Totowa, NJ, pp 365-386