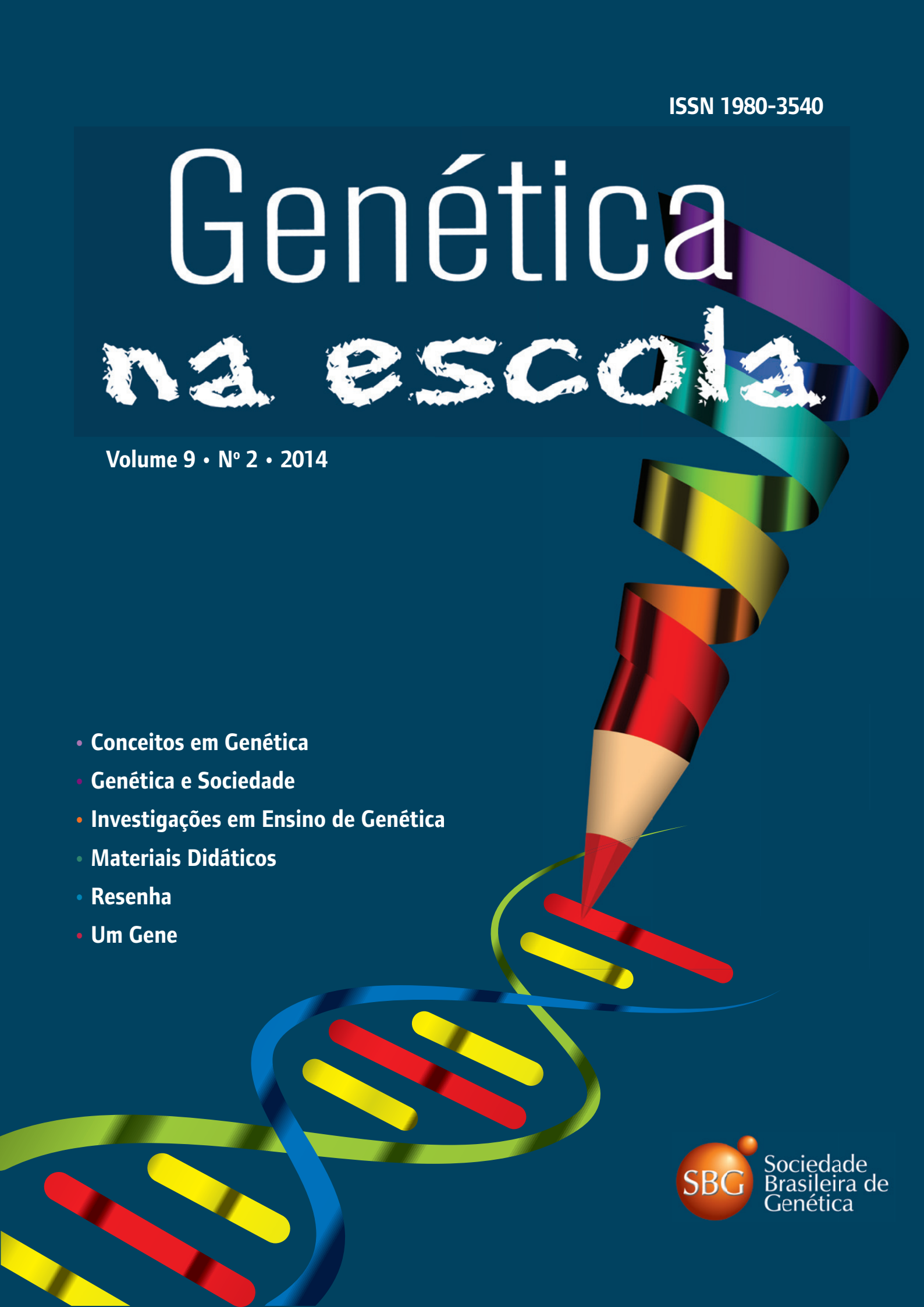
ISSN 1980-3540

Genética na escola

Volume 9 • Nº 2 • 2014

- Conceitos em Genética
- Genética e Sociedade
- Investigações em Ensino de Genética
- Materiais Didáticos
- Resenha
- Um Gene



SBG Sociedade Brasileira de Genética

Índice

Conceitos em Genética

Herança monogênica: além de Mendel, além do DNA 80

Genética e Sociedade

A Genética e as invasões biológicas: dois estudos de caso de bivalves invasores do Brasil 86

Investigações em Ensino de Genética

Generalizações que distanciam os conhecimentos dos livros didáticos das referências em Genética 92

Materiais Didáticos

Avaliação da depressão endogâmica em populações de araticunzeiro: apoio pedagógico ao Ensino de Genética 104

Tirinhas no ensino da estrutura, função e conceito de gene 118

A Genética contra os crimes ambientais:
identificação de madeira ilegal provenientes de unidades de conservação utilizando marcador molecular 124

Como desvendar enigmas genéticos a partir da comparação de sequências 136

Resenha

Mitos da Genética Humana 146

Um Gene

O gene da intolerância à lactose 148

Genética na Escola – ISSN: 1980-3540

Genética na escola

Sociedade Brasileira de Genética

Herança monogênica: além de Mendel, além do DNA

Maria de Nazaré Klautau-Guimarães[1], Sabrina Guimarães Paiva[2], Silviene Fabiana de Oliveira[1]

[1] Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, campus Darcy Ribeiro, Universidade de Brasília.
[2] Instituto Federal de Educação, Ciência e Tecnologia do Tocantins (IFTO) e doutoranda no Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, campus Darcy Ribeiro, Universidade de Brasília.

Autor para correspondência: nklautau@unb.br

O artigo propõe uma reflexão sobre o ensino da herança monogênica, apontando as limitações decorrentes da sua simplificação e sugerindo a apresentação do quadro atual dos conceitos envolvidos.É importante transmitir ao aluno de graduação a diversidade de novos conhecimentos que se contrapõem à ideia simples da herança monogênica, que foi inicialmente proposta e ainda persiste no ensino. O artigo não pretende apresentar uma fórmula pronta e sim uma ideia a ser construída por professores e demais graduados.

O ENSINO DA HERANÇA MONOGÊNICA

O modelo clássico de investigação genética, principalmente da espécie humana, baseia-se em estudos familiares e na transmissão de um fenótipo, que é tradicionalmente utilizado como exemplo em sala de aula e nos livros didáticos. Porém, de fato, apenas poucas características apresentam reconhecidamente herança monogênica, que conceitualmente representa a transmissão de uma característica, de uma geração a outra, cuja expressão do fenótipo depende de somente um gene. E ainda, o número de características e doenças cuja variação fenotípica pode ser explicada por mutações em um único lócus está se reduzindo significativamente ao longo das décadas, devido ao avanço do conhecimento na área da genômica.

Rotineiramente, para o ensino dos padrões de herança em humanos utiliza-se a observação da transmissão de características fenotípicas de fácil visualização, como cor dos olhos, lóbulos da orelha, cor de cabelos, entre outros. Porém, comumente o fenótipo dessas características é simplificado para duas classes distintas derivadas do produto da expressão de um único gene e dois alelos com relação de dominância completa entre eles, isso é, um de expressão dominante e o outro, de expressão recessiva. Grande parte dos livros didáticos apresenta a variação e a expressão fenotípica desse modo simplificado, apontando uma relação genótipo-fenótipo direta e associada a um padrão de herança mendeliana, como apresentado no artigo *Mitos da Genética Humana* (MCDONALD, 2011). Considerando a literatura atual e o banco de dados *Online Mendelian Inheritance in Man* (OMIN), grande parte das características humanas de fato não se encaixa nesse modelo genético simples.

HISTÓRICO

O interesse humano pela compreensão da variação genética é descrita desde a cultura grega até a contemporânea. Porém, o primeiro grande marco nesta área ocorreu em meados do século XIX com os trabalhos de Gregor Mendel, que introduziram os princípios da herança biológica. Em seus trabalhos demonstrou que as variações detectadas para algumas características da ervilha eram transmitidas de geração a geração, em um determinado padrão. Após a redescoberta dos trabalhos de Mendel em 1900, o interesse nesta área envolveu estudos em diferentes organismos modelos, os quais confirmaram, de forma geral, a aplicação dos princípios da herança mendeliana. Esses princípios referem-se aos pares de fatores (genes), aos conceitos de dominância/recessividade e à segregação aleatória dos alelos para os gametas. Assim, a variação fenotípica está relacionada à presença de dois alelos para um determinado gene nas características analisadas que apresentem o padrão de segregação mendeliana.

Após o desenvolvimento da microscopia, foi possível associar os conhecimentos gerados na área da herança com os da área da citologia. Desta forma foi observado que os fatores de Mendel (genes) e os cromossomos tinham comportamento similar durante a meiose, o que levou a proposta da **Teoria Cromossômica da Herança**. Essa proposta estabelece que as características hereditárias são controladas por genes que se localizam nos cromossomos, que são transmitidos por meio dos gametas, mantendo a continuidade genética de geração em geração. Esse foi o inicio do conhecimento sobre as bases físicas da herança.

No século XX, com o desenvolvimento da genética, começaram a ser observadas variações ao padrão de segregação mendeliana. Com isso, novos conceitos foram cunhados na tentativa de registrar e explicar esses eventos. A partir da observação dessas variações, o conhecimento sobre a variação genética já se encontrava **além de Mendel**.

Ainda no século XX, o segundo grande marco na área da Genética foram as descobertas das bases química, molecular e funcional da herança com a definição da estrutura dos ácidos nucléicos e do **Dogma Central da Biologia Molecular**, conceito de que o fluxo da informação genética ocorre do DNA ao RNA e do RNA para as proteínas. Contudo, uma série de descobertas sobre os genes e os processos de expressão gênica dificultou a interpretação clássica do gene como unida-

de hereditária que possui estrutura, função e localização definidas, afetando diretamente a compreensão da herança monogênica. Diante da complexidade do genoma e da maquinaria celular, a proposta de uma relação de 1:1:1 entre um gene, um produto proteico e uma função apresenta-se simplificada. Considerando a crise do conceito de gene, ressalta-se que sua expressão não ocorre de forma isolada e sim na dependência da estrutura do DNA, do ambiente celular e externo, o que demonstra a definição de padrão de herança muito mais complexa (JOAQUIM; EL HANI, 2010).

Já no século XXI, o terceiro grande marco, foram os conhecimentos sobre a estrutura e função dos genes e genomas advindos das novas tecnologias. Nos últimos anos, a área da epigenética está esclarecendo as bases da variação fenotípica herdada por meio de processamento de sinais externos. As interações do genoma com uma diversidade de fatores ambientais geram diferentes padrões de expressão por meio de modificações químicas reversíveis do DNA e da estrutura da cromatina (KLUG *et al*. 2013).

Realmente, não é possível predizer completamente o quadro fenotípico com base na alteração genética para a maioria das características e/ou doenças ditas monogênicas, devido a efeitos adicionais que a transformam em características complexas de padrões de herança atípicos. Assim, o padrão de herança monogênica já se encontra **além de Mendel e além do DNA**.

REFLEXÕES SOBRE O ENSINO ATUAL DA HERANÇA MONOGÊNICA NA GRADUAÇÃO

Os novos conhecimentos nos obrigam a fazer uma reflexão sobre o ensino da herança monogênica, principalmente diante da crise do conceito de gene. Destacamos alguns pontos importantes a abordar como: a) a ampliação da visão sobre hereditariedade para além do genecentrismo; b) a amplitude da variação fenotípica esperada para uma característica monogênica; c) e a diversidade de fatores envolvidos na relação genótipo x fenótipo. Se considerarmos os vários modos de interação alélica e o número de alelos, essa relação com certeza será de um gene: muitos fenótipos. Se adicionarmos a interação com outros genes e com o meio ambiente, poderá advir uma maior variação fenotípica.

Por outro lado, a compreensão da etiologia molecular de doenças genéticas está alterando nossa percepção da transmissão da doença. Foi evidenciado um maior nível de complexidade da variação fenotípica em doenças de herança mendeliana e isso nos leva a ver que a nossa classificação em herança monogênica e multifatorial é uma simplificação. Cada vez mais se observa que grande parte das doenças genéticas apresenta variação fenotípica como resultado de ação de alelos de vários genes, além de outros fatores extragênicos (BADANO; KATSANIS, 2002).

Os clássicos exemplos de herança monogênica, Fibrose cística (FC) e Fenilcetonúria (PKU), apresentam complexidade genética que envolve genes modificadores na extensa variação fenotípica, mesmo na presença de genótipos idênticos para os genes principais, CFTR e PAH, respectivamente. A herança do gene específico segue o padrão de segregação mendeliana, entretanto o conhecimento da mutação não pode predizer exatamente o

fenótipo do paciente, indicando que outros fatores genéticos e ambientais são importantes moduladores da expressão fenotípica (BADANO; KATSANIS, 2002).

A dificuldade no estabelecimento da correlação genótipo x fenótipo (PKU e CF), indica que o modelo mendeliano é útil para a identificação primária da causa genética das doenças familiais e também para cálculos de riscos de recorrência, mas pode ser um modelo incompleto para entender a natureza do defeito celular e fisiológico. Provavelmente, ao menos parte dessas doenças genéticas hoje consideradas monogênicas poderiam ser melhor descritas como de **herança oligogênica**, isso é, a transmissão de características determinadas primariamente por etiologia genética que requerem a ação sinérgica de alelos de um pequeno número de lócus. Embora as bases moleculares dos fenômenos descritos para as doenças de herança oligogênica permaneçam ainda pouco entendidas, sabe-se que a etiologia de determinados casos são devidos à estrutura dos produtos gênicos envolvidos, como no caso de produtos poliméricos, complexos protéicos, relação ligante-receptor e produtos que participam de uma mesma via bioquímica. Assim, a maioria das mutações provavelmente apresenta efeito quantitativo no fenótipo, o que faz parte da expansão do conceito monogênico (BADANO; KATSANIS, 2002).

Nesse sentido, Shawky (2009) apresenta os diversos aspectos a serem considerados sobre a extensão da variação que pode ser encontrada para uma característica de herança monogênica, os quais envolvem a base molecular das mutações e outros mecanismos. São eles: 1) um gene, uma mutação, muitos fenótipos; 2) um gene, muitas mutações, muitos fenótipos; 3) outros mecanismos como dosagem gênica, regulação gênica, interação gênica, genes modificadores, processamento (*splicing*) alternativo, *imprinting* genômico, mecanismos epigenéticos e interação com o ambiente.

Novas perspectivas de olhar a hereditariedade considerando a importância do ambiente materno na transmissão de características hereditárias foram colocadas por Maurel e Kanellopoulos-Langevin (2008). As autoras enfatizam que a transmissão hereditária envolvem, no mínimo, três processos: a herança mendeliana (nuclear), a herança citoplasmática, que envolve o mtDNA e os componentes citoplasmáticos do ovócito, e o microquimerismo. Esse terceiro processo envolve a troca de células entre o feto e a mãe (nos dois sentidos) que podem permanecer ativas na vida adulta. A interação materno/fetal durante a gestação e no aleitamento envolve a troca de macromoléculas, organelas subcelulares e células vivas que envolvem processos epigenéticos na transmissão de características. Esses processos apresentam efeito no desenvolvimento do embrião, no processo da regulação da função imune da criança e pode permanecer por várias gerações (MAUREL; KANELLOPOULOS-LANGEVIN, 2008).

CONSIDERAÇÕES FINAIS

Os estudos dos mecanismos envolvidos nas características de herança monogênica permitiram o esclarecimento da fisiopatologia de diferentes doenças e, consequentemente, alterou a percepção de padrão de transmissão, revelando a complexidade da determinação fenotípica. Embora muitas características tenham um componente genético e mostrem recorrência familiar, não é observado um padrão de herança mendeliana.

Torna-se necessário apontar ainda que, além do componente genético, há diversos fatores envolvidos na determinação de uma característica ou doença, que ampliam a variação fenotípica. É importante que o estudante tenha uma visão global e dinâmica dos possíveis níveis de organização relacionados à expressão fenotípica: o molecular (genes), o micro (cromossomos e células), o macro (organismo e ambiente) e suas interações.

No texto procuramos propor uma reflexão sobre o ensino da herança monogênica, apontando as limitações decorrentes da sua simplificação e sugerindo a apresentação do quadro atual dos conceitos envolvidos. A revisão de conceitos anteriormente mal compreendidos na formação dos professores e profissionais da saúde pode permitir a melhor compreensão dos novos conhecimentos sobre a herança biológica e romper o preceito de que a maioria das características fenotípicas apresenta herança monogênica.

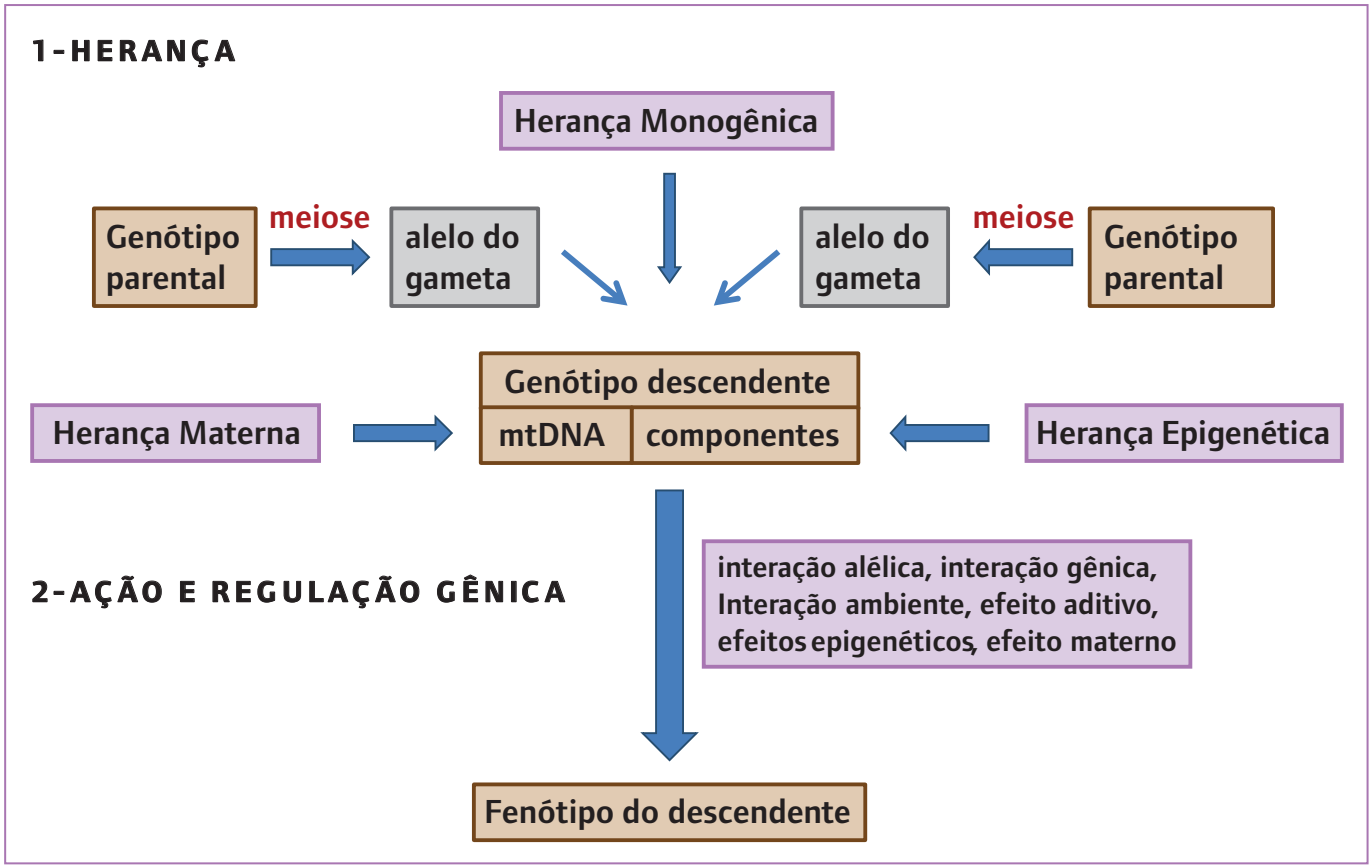


Figura 1.
Esquema sugerindo que a percepção da transmissão de um fenótipo possa ser abordada em duas etapas: 1. transmissão do material genético; 2. ação e controle da expressão do gene.

Na tentativa de facilitar a visualização dos diversos mecanismos envolvidos no ensino da herança monogênica levantados nessa reflexão, a Figura 1 sugere a apresentação da transmissão de um fenótipo em duas etapas: 1. transmissão do material genético (herança) e 2. ação e controle da expressão do gene. Essa figura pode ser utilizada pelos docentes no sentido de nortear a discussão e apresentação de cada etapa e/ou processo.

REFERÊNCIAS

BADANO, J.L.; ANDKATSANIS, N. Beyond Mendel: a evolving view of human genetic disease transmission, *Nature Reviews*, v.2, p.770-789, 2002.

JOAQUIM, L.M.; EL-HANI,C.N. A genética em transformação: crise e revisão do conceito de gene, *Scientla Studia*, v.8, n.1, p.93-128, 2010.

KLUG, W.S; CUMMINGS, M.R.; SPENCER, C.A.; PALLADINO, M. A. *Conceptos de Genética*, 10a edición, Madrid: Pearson Educación, S.A, 2013.

MACDONALD, J. (2011). Myths of Human Genetics. University of Delaware. Disponível em: <http://udel.edu/~mcdonald/MythsHumanGenetics.pdf>, Acessado em 03/02/2014.

MAUREL, M.C.; KANELLOPOULOS-LANGEVIN, C . Heredity—Venturing Beyond Genetics. *Biology of Reproduction*, v.79, p.2–8, 2008.

SHAWKY, R. M . One gene, many phenotypes. *Egyptian Journal of Medical Human Genetics*, v. 10, p. 1- 12, 2009.

A Genética e as invasões biológicas: dois estudos de caso de bivalves invasores do Brasil

Ramon Pereira Lopes[1], Michelle Rezende Duarte[2], Edson Pereira da Silva[3]

[1] Mestrando em Biologia Marinha, Universidade Federal Fluminense
[2] Doutoranda em Biologia Marinha, Universidade Federal Fluminense
[3] Laboratório de Genética Marinha e Evolução, Departamento de Biologia Marinha, Universidade Fluminense, Niterói, RJ

Autor para correspondência: gbmedson@vm.uff.br

As bioinvasões têm causado prejuízos econômicos e ambientais em todo mundo e já começam a trazer os mesmos danos ao Brasil. Uma característica interessante das espécies invasoras é que elas não compartilham um passado evolutivo com a comunidade invadida. Quando são introduzidas em um novo ambiente, as espécies invasoras ficam sujeitas a mudanças evolutivas que podem nos trazer informações importantes sobre a genética das populações.

BIOINVASÃO E SEUS IMPACTOS

Invasão biológica, ou simplesmente bioinvasão, é o termo utilizado para definir a presença de determinada espécie em uma região onde ela não foi registrada anteriormente. Existem dois tipos de invasões: expansões e introduções. Enquanto as expansões consistem na dispersão de organismos por mecanismo natural, as introduções ocorrem quando as espécies são transportadas por atividades humanas, intencionalmente ou não, para uma área onde não ocorriam.

O veículo ou atividade pela qual uma espécie é transportada e introduzida em um novo *hábitat* é denominado vetor. No meio aquático, podem ser identificados àqueles que resultam em introduções consideradas intencionais, como a aquicultura e a aquariofilia. Vetores considerados acidentais ou não intencionais incluem a água e o sedimento dos tanques de lastro, equipamentos de recreação, as incrustações nos cascos de navios, plataformas de petróleo, boias de navegação, flutuantes e os detritos sólidos flutuantes. Também merecem atenção organismos epibiontes, simbiontes ou parasitas que podem estar associados à introdução via aquicultura e aquariofilia, bem como os canais (de nível, comportas, irrigação).

De certa forma é surpreendente que espécies invasoras se estabeleçam em novos *hábitat*. Isto porque, para que uma introdução seja bem sucedida, a espécie invasora deve ser compatível com o novo ambiente. Por exemplo, pequenas mudanças de temperatura, hora do dia ou número de indivíduos introduzidos podem representar a diferença entre o sucesso ou o fracasso de uma bioinvasão. Ou seja, há um conjunto de variáveis que fazem o fenômeno da bioinvasão um sistema complexo de difícil previsão. A despeito disso, é possível identificar algumas características que tornam maiores as chances de sucesso desse jogo. Espécies generalistas, ou seja, com uma maior amplitude de tolerância a fatores ambientais (temperatura, salinidade, luminosidade etc.) apresentam vantagens em relação àquelas que apresentam menor amplitude. O caranguejo verde europeu *Carcinus maenas*, por exemplo, é capaz de viver e se reproduzir numa ampla gama de temperaturas, salinidades e *hábitat*, bem como apresenta hábito alimentar onívoro, alimentando-se de diversas plantas e animais. A espécie apresenta, ainda, grande fecundidade e é extremamente agressiva na competição com outras espécies de crustáceos e bivalves do ambiente. Por outro lado, o sucesso de uma espécie como invasora também depende do ambiente e da comunidade invadida, que pode apresentar poucos competidores ou predadores ou, ainda, recursos disponíveis para serem explorados. Este foi o caso, por exemplo, da alga do gênero *Spartina* que invadiu bancos de lama na Baía de Willapa, Estados Unidos.

As invasões biológicas são desafios complexos e de grande extensão ao bem–estar das populações humanas. Espécies invasoras podem diminuir o rendimento dos cultivos no setor de maricultura, trazer prejuízos operacionais e financeiros para indústrias,

usinas e transportes marítimos em função de incrustações em larga escala, provocar a perda do potencial pesqueiro nos ambientes em que foram introduzidas e, até, desencadear epidemias ou ondas de intoxicação capazes de matar ou incapacitar milhões de pessoas a cada ano. Se alguma ameaça ambiental pode ser considerada irremediável esta é a bioinvasão. Ao contrário da poluição física ou química, agentes biológicos se reproduzem e se espalham autonomamente por muito tempo e longas distâncias. Apenas uma pequena fração das invasões biológicas é passível de erradicação e, mesmo assim, com custos imensos. A despeito das suas consequências adversas, existem muitas dificuldades políticas em se lidar com o problema devido ao fato de que os seus custos econômicos, ambientais e sociais se contrapõem aos benefícios sociais e econômicos advindos do livre comércio.

MUDANÇAS EVOLUTIVAS EM ESPÉCIES INVASORAS

As espécies invasoras não compartilham um passado evolutivo com a comunidade invadida e esse fato faz das invasões biológicas um modelo interessante para compreensão dos processos que atuam na evolução dessas espécies e suas populações. As espécies invasoras estão sujeitas a cinco tipos de mudanças evolutivas:

1. **Hibridação –** Cruzamento entre espécies de bioinvasores com espécies nativas ou com outras espécies invasoras.
2. **Rearranjos genômicos –** Fenômeno que pode determinar a rápida adaptação das populações invasoras. A poliploidia (duplicação do genoma) e a alopoliploidia (hibridização seguida de duplicação do genoma) são processos de reconhecida importância na evolução das plantas.
3. **Modificação do genoma induzida pelo estresse –** A exposição às condições bióticas e abióticas do novo ambiente pode causar uma instabilidade no genoma, nesse caso, mediada pelo estresse ambiental e tais modificações facilitam a adaptação.

Essas três primeiras mudanças, de um modo geral, trazem aumento da variação genética populacional.

4. **Efeito de determinados genes –** De forma semelhante, o efeito de determinados genes e não obrigatoriamente o seu número, pode auxiliar a habilidade de colonização das espécies invasoras.
5. ***Bottleneck* –** O efeito gargalo (em português), por outro lado, é um caso extremo de deriva genética (quando o acaso tem papel importante na determinação de quais genes estarão presentes na próxima geração) que ocorre pela redução drástica do tamanho populacional (ver Figura 1), tendo como consequência a redução dos níveis de variação genética da população. A chegada acidental de um ou poucos indivíduos de uma espécie em um novo ambiente, como se dá no caso das bioinvasões, é um exemplo de *bottleneck* (conhecido, nesse caso, como efeito fundador). Nesses casos, os baixos níveis de variação genética podem afetar a capacidade das populações de se adaptarem ao novo ambiente.

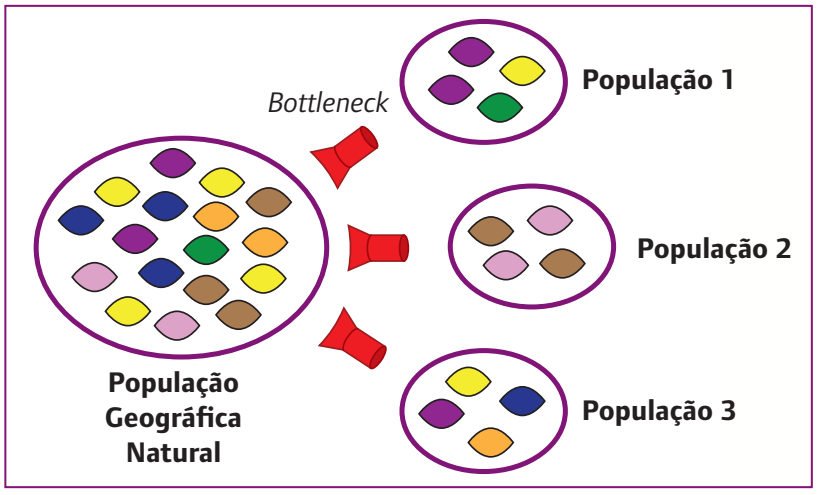


Figura 1.
Representação gráfica do efeito *bottleneck* ou gargalo no qual uma população geográfica natural dá origem a várias populações menores, cada uma com apenas parte da variação genética original.

Outro fator importante para o sucesso das espécies invasoras é a qualidade do ambiente invadido: sabe-se, por exemplo, que ambientes poluídos podem facilitar o crescimento de algumas espécies invasoras, provavelmente, pela diminuição da competição.

BIVALVES INVASORES NO BRASIL

Os bivalves são organismos exclusivamente aquáticos pertencentes ao filo dos moluscos. Eles formam um grupo extremamente bem sucedido e diversificado, podendo ocorrer em ambientes de salinidade diversa como água salgada, doce ou salobra.

Existem registros de bivalves exóticos introduzidos em várias partes do mundo, sendo que no Brasil, até o momento, foram identificadas nove espécies: *Myoforceps aristatus*; *Mytilopsis leucophaeta*; *Perna perna*; *Corbicula fluminea*; *Corbicula manilensis*; *Corbicula largillierti*; *Limonperna fortunei*; *Crassostrea gigas* e *Isognomon bicolor*.

Os bivalves *Myoforceps aristatus* e *Mytilopsis leucophaeta* são classificados como **espécies detectadas**: classificação dada para as espécies que tiveram registros isolados no ambiente natural, não tendo registros posteriores do aumento de abundância ou dispersão. Caso a espécie detectada volte a aparecer, sendo encontrada de forma recorrente e apresentando indícios de aumento populacional, ela passa a ser classificada como **espécie estabelecida ou naturalizada**, porque já mantém interações com as demais espécies. Essa é a situação em que se enquadra o bivalve *Perna perna*, que foi introduzido entre os séculos XVIII e XIX e é encontrado hoje em grande parte do litoral brasileiro.

Quando as espécies exóticas vencem as pressões ecológicas e conseguem se dispersar para outras regiões são chamadas de **espécies invasoras**, atual classificação de *Corbicula fluminea*, *Corbicula manilensis*, *Corbicula largillierti*, *Crassostrea gigas*, *Limonperna fortunei* e *Isognomon bicolor*.

A ABORDAGEM GENÉTICA

É possível afirmar que as invasões biológicas são, geralmente, constituídas de rápidos eventos evolutivos, resultando em populações geneticamente dinâmicas, tanto no espaço, quanto no tempo. Assim, a abordagem genética constitui-se numa importante alternativa no esforço para entender o processo de estabelecimento dessas espécies nos ecossistemas aquáticos do Brasil.

Como já foi discutido antes, o sucesso no estabelecimento de uma população bioinvasora pode estar relacionado com o seu nível de variação genética, pois ela oferece às populações invasoras uma maior flexibilidade nas respostas às pressões do novo ambiente. Uma técnica molecular muito utilizada para mensuração da variação genética e que já teve grande impacto nos estudos populacionais é a eletroforese de aloenzimas. De uma forma bem resumida, essa técnica pode ser definida como a migração de partículas sob a ação de uma corrente elétrica. Ela se baseia nas características físico-químicas das proteínas, ou seja, suas diferentes formas, tamanhos, mas, principalmente, suas cargas elétricas. Proteínas diferentes apresentam cargas elétricas também diferentes, fato que se manifestará em diferentes mobilidades eletroforéticas.

A aplicação da técnica de eletroforese de aloenzimas para o estudo de espécies invasoras apresenta como vantagens a fácil preparação de extratos, o preço relativamente barato e sua objetividade. No entanto, seu uso também apresenta algumas limitações, como por exemplo, a subestimativa da variação genética total. Mas, talvez, a maior limitação do método esteja no fato de que a variação amostrada por eletroforese seja, na realidade, uma variação fenotípica, ou seja, dos produtos diretos dos genes e não dos genes ou DNA em si.

A eletroforese de aloenzimas já auxiliou com sucesso alguns trabalhos com bioinvasores no Brasil e em outros países, por exemplo, bivalves do gênero *Mytilus* no Chile, o mexilhão zebra *Dreissena polymorpha* no Canadá, a lagosta *Orconectes limosus* na França e, no Brasil, a mosca *Zaprionus indianus*. Em todos estes casos, a eletroforese de aloenzimas se mostrou uma boa aliada no estudo das bioinvasões. A seguir, são apresentados dois estudos de caso com os bivalves invasores *Isognomon* bicolor e *Limnoperna fortunei* que demonstram como a abordagem genética pode ajudar a entender o processo de invasão biológica em meio aquático.

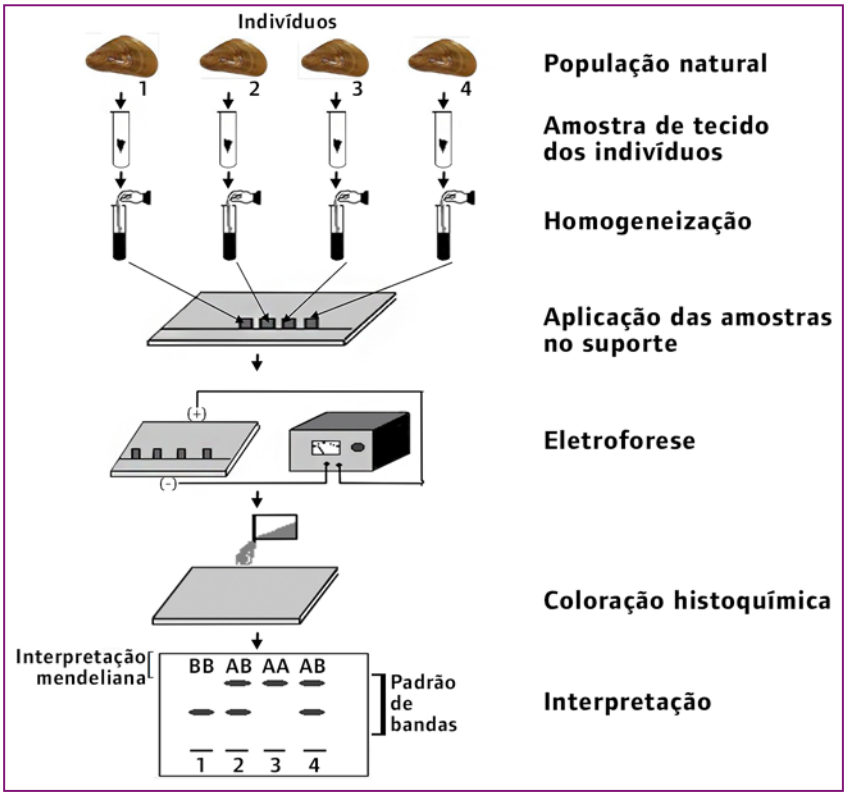


Figura 2.
Resumo gráfico do método de eletroforese de aloenzimas empregada para análise da variação genética em populações.

Isognomon bicolor

Bivalve marinho pertencente à família Isognomonidae, endêmico do mar do Caribe. Possui rápido crescimento, sendo encontrado em altas densidades nos costões rochosos, ocorrendo desde a faixa superior do médio litoral até sete metros de profundidade. Suas conchas adotam as mais diversas formas e isso permite que este bivalve possa se expandir entre e sobre os demais organismos incrustantes do costão rochoso, limitando a habilidade que esses organismos poderiam ter para obtenção de alimento. *I. bicolor* está presente no litoral brasileiro desde meados da década de 1980, quando foi registrado em Atol das Rocas (Natal, RN). Atualmente, este bivalve está presente em grande parte da costa brasileira, incluindo as regiões Nordeste (RN, PE e BA), Sudeste (SP e RJ) e Sul (PR e SC).

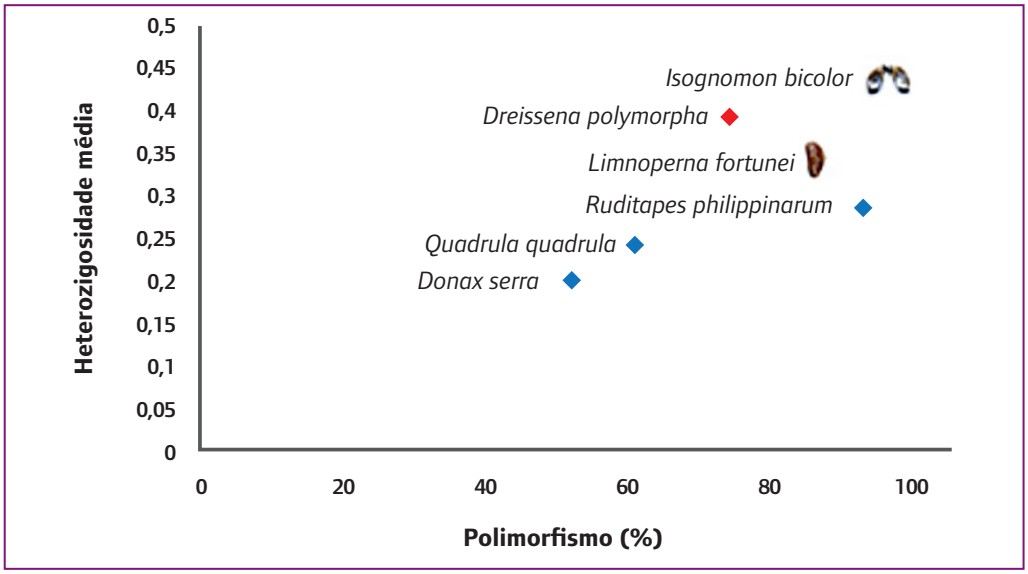
Investigando os padrões de variação genética do *Isognomon bicolor* na Praia de Itaipu (Niterói, RJ), em três momentos no tempo (2005, 2009 e 2013), foram encontrados altos níveis de variação genética, bem como sua estabilidade nos três anos estudados. De certa forma, o resultado é surpreendente, uma vez que se esperava que essas populações invasoras estivessem sob o efeito do *bottleneck* (gargalo) e, portanto, apresentassem níveis de variação genética (polimorfismo, número de alelos, heterozigosidade etc.) baixos. Mas não é caso de *Isognomon bicolor* quando comparada com outras espécies de bivalves invasores (ver Figura 3).

Altos níveis de variação genética em bioinvasores podem ser explicados como consequência de múltiplas invasões, uma hipótese plausível considerando que embarcações de grande porte, como as que fazem transportes intercontinentais, são muito frequentes em nosso litoral.

Limnoperna fortunei

Mais conhecido como mexilhão dourado, é um bivalve de água doce pertencente à família Mytilidae, endêmico do sul da Ásia. É capaz de se fixar em uma grande variedade de substratos e possui grande capacidade adaptativa, atingindo densidades de até 150.000 indivíduos/m^2. Com 30 dias e aproximadamente 0,5 cm de comprimento já começam a se reproduzir. No Brasil, esta espécie foi registrada em 1998 no Pantanal Mato-Grossense e no sistema da Lagoa dos Pa-

Figura 3.
Níveis de variação genética em populações de I. bicolor em comparação com outras espécies invasoras (*Dreissena polymorpha*, *Limnoperna fortunei*) e Bivalves (*Donax serra*, *Quadrula quadrula*, *Ruditapes philippinarum*).



tos. Tornou-se uma praga nas bacias dos rios Paraná, Paraguai, Uruguai e Jacuí/Patos. Os impactos ambientais causados pelo mexilhão dourado envolvem danos ao abastecimento público de água, destruição de hábitats e a diminuição da fauna nativa.

As populações invasoras de *Limnoperna fortunei*, tanto no Brasil (Lago Guaíba, RS; Porto Esperança, MS e Usina de Itaipu, PR, Praia de Itaipu, RJ) quanto na Argentina (Balneário Bagliardi e Usina de Yaciretá), apresentaram altos níveis de variação genética. Este padrão é semelhante àquele encontrado para *Isognomon bicolor* e, do mesmo modo, indica que o efeito *bottleneck* parece não ter atuado em nenhuma das cinco populações estudadas.

Embora surpreendente, a variação genética alta em espécies invasoras não é uma exceção. As bioinvasões mediadas por água de lastro, como é, provavelmente, o caso de *Limnoperna fortunei*, podem ser o resultado de um único evento de bioinvasão a partir de tamanhos populacionais muito grandes dos invasores ou, ainda, de múltiplas invasões.

Os estudos com o *Isognomon bicolor* e *Limnoperna fortunei* indicam que vários episódios de introdução podem ter facilitado a instalação e a permanência desses bivalves no novo hábitat. Caso nenhuma medida seja tomada para prevenir episódios de introdução de espécies exóticas nas águas brasileiras, a bioinvasão pode se tornar uma ameaça importante à biodiversidade local pela possibilidade de extinção de espécies nativas.

PARA SABER MAIS

FERNADES, F. C.; SOUZA, R. C. C. L.; SILVA, E. P. In: FERNANDEZ, M. A.; SANTOS, S. B.; PIMENTA, A.; THIENGO, S. C. (ORGS). *Bivalves marinhos introduzidos no Brasil. Tópicos em Malacologia: Ecos do XIX Encontro Brasileiro de Malacologia*, Sociedade Brasileira de Malacologia – SBMa. Rio de Janeiro: Technical Books Editora, 2011. p. 300-307.

MACHADO, C. J. S.; OLIVEIRA, A. E. S.; MATOS, D. M. S.; PIVELLO, V.; CHAME, M.; SOUZA, R. C. C. L.; CALAZANS, S. H.; SILVA, E. P. Recomendações para elaboração e consolidação de uma estratégia nacional de prevenção e controle das espécies exóticas no Brasil. *Ciência e Cultura*, v. 61, n.1, p. 42-45, 2009.

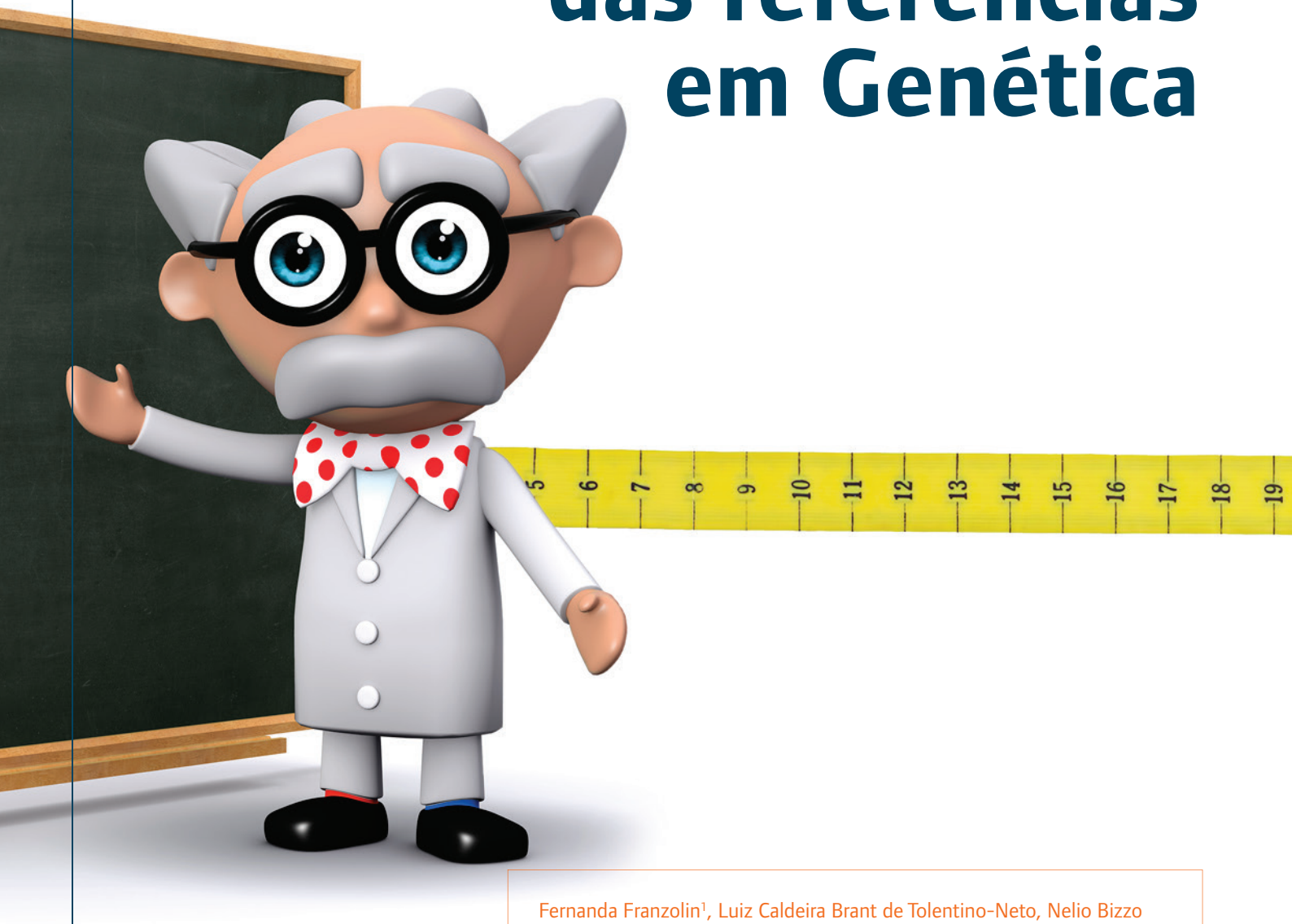
SILVA, E. P. Marcadores moleculares no rastreamento da bioinvasão. In: SILVA, J. S. V.; SOUZA, R. C. C. L. (ORGS). *Água de Lastro e Bioinvasão*, Rio de Janeiro: Editora Interciência, 2004. p. 191-203.

SILVA, E. P. Genética Marinha. In: SOARES-GOMES, A.; PEREIRA, R. C. (EDS). *Biologia Marinha*, 2ª edição. Rio de Janeiro: Editora Interciência, 2009. p. 333-351.

SILVA, E. P. Genética da bioinvasão do mexilhão-dourado. In: MANSUR, M. C. D.; SANTOS, C. P.; PEREIRA, D.; PAZ, I. C. P.; ZURITA, M. L. L.; RODRIGUEZ, M. T. R; NEHRKE, M. V.; BERGONCI, P. E. A. (ORGS). *Moluscos Límnicos Invasores do Brasil: Biologia, Prevenção, Controle*. Porto Alegre: Redes Editora, 2012. p. 51-57.

SOUZA, R. C. C. L.; CALAZANS, S. H.; SILVA, E. P. Impactos das espécies invasoras no ambiente aquático. *Ciência e Cultura*, v. 61, n.1, p. 35-41, 2009.

Generalizações que distanciam os conhecimentos dos livros didáticos das referências em Genética



Fernanda Franzolin[1], Luiz Caldeira Brant de Tolentino-Neto, Nelio Bizzo

[1] Centro de Ciências Naturais e Humanas, Universidade Federal do ABC, São Paulo
[2] Centro de Educação, Universidade Federal de Santa Maria, Rio Grande do Sul
[3] Faculdade de Educação, Universidade de São Paulo, São Paulo

Autor para correspondência - fernanda.franzolin@ufabc.edu.br

Como os conhecimentos de Genética presentes nos livros didáticos se aproximam ou se distanciam do conhecimento de referência utilizado na formação dos professores? Uma pesquisa em livros didáticos brasileiros e estadunidenses traz resultados referentes a um tipo de distanciamento frequentemente encontrado: as generalizações, que consistem em atribuir a um conhecimento particular um caráter geral. Ambos os materiais apresentam generalizações que os distanciam do conhecimento veiculado nos livros de referência utilizados na graduação, sendo que nos livros brasileiros há maior incidência de generalizações quando tratam da meiose e, nos EUA, a maior frequência está no conteúdo expressão gênica. Durante a exposição dos resultados deste trabalho são apresentados alguns exemplos de como alguns livros evitam essas generalizações, bem como, as possibilidades de encurtar tais distanciamentos seja com mudanças na formação inicial, seja com o trabalho em sala de aula.

INTRODUÇÃO

Selecionar e adequar o conhecimento a ser ensinado estão entre as tarefas mais importantes do professor e é fundamental que sejam abordadas nos cursos de formação inicial de professores. Geralmente, nos cursos de licenciatura, os conceitos científicos são abordados com grande grau de aprofundamento, mas não é incomum estudantes concluintes da graduação afirmarem que sabem muito sobre o conteúdo específico de sua área, mas temem ir para a sala de aula por não saber como transmiti-los. Cursos que possuem uma separação entre a formação científica e a formação pedagógica do professor são questionados pela literatura e contrariam as normas legais que regem a matéria, em nível nacional, pelo menos desde 2002 (BRASIL, 2002)

Deve-se salientar que o conhecimento gerado pela ciência e aprendido pelo futuro professor em sua graduação não consiste no mesmo conhecimento a ser ensinado na educação básica. Primeiramente, é preciso considerar que o saber escolar não é composto apenas pelo saber científico, mas também

pelo saber cotidiano (LOPES, 1999), o que não elimina a importância de aproximar o aluno do Ensino Médio do conhecimento produzido pela Ciência. Apesar de não ser papel da escola necessariamente substituir as concepções dos alunos por ideias científicas, apresentando-as como superiores a outros saberes, é preciso permitir que os estudantes desenvolvam uma visão de mundo compatível com a ciência (COBERN, 1996), ou seja, fazê-los ser capaz de compreender o conhecimento produzido e que, em certos contextos, pode ajudar em nossa vida cotidiana. Isso seria importante não por se considerar esse conhecimento como melhor do que os outros, ou uma verdade inquestionável, mas por se considerar que esse conhecimento é *confiável*, pois foi gerado no *consenso* de uma comunidade científica, tendo sido avaliado por diversos pesquisadores de uma determinada área de conhecimento e por eles aceito (ZIMAN, 1985). Pode-se também considerar que os indivíduos desenvolvem imagens científicas do mundo das quais obtêm diversas deduções. Essas opiniões sobre a realidade científica servem de base para muitas ações sociais, o que também justifica a importância do conhecimento científico ser ensinado no ambiente escolar (ZIMAN, 1985).

Porém, não é possível ensinar esse saber científico nas escolas utilizando a mesma linguagem com a qual os cientistas se comunicam. Chevallard (1991), ao trazer para a Didática da Matemática o conceito de transposição didática, criado por Verret em 1975, serve hoje de referência também para pesquisadores de outras áreas. Este autor argumenta que o saber gerado pela ciência, o saber erudito (ASTOLFI, 2000), não é o mesmo que é ensinado para os alunos. A transposição didática seria a transformação do objeto de saber em objeto de ensino.

Os cursos de graduação utilizam tanto artigos científicos, os quais são produtos diretos da produção de conhecimento científico, quanto livros-texto, que já passaram por um certo grau de transposição didática, mantendo grande aproximação com o saber produzido na academia. Todavia, a linguagem dessas fontes não foi elaborada para alunos de Ensino Fundamental ou Ensino Médio. É evidente que a transposição didática precisa ser adequada para cada nível, para que, como afirma Chevallard, possa se tornar compreensível para o estudante. Aprender a realizar transformações nesse estágio do saber, transformando-o em um saber apropriado ao aluno, torna-se então um desafio para os futuros professores.

Se, por um lado, é importante o licenciando estar preparado para realizar a transposição didática, outra necessidade é que ele também seja capaz de fazer o que Chevallard denomina como vigilância epistemológica, que consiste numa espécie de verificação da pertinência dos resultados do processo. Para este autor, apesar de necessário, o processo de transposição didática pode causar "disfunções inadequadas".

Em sala de aula, os professores utilizam os livros didáticos tanto para preparar aulas quanto para ministrá-las. Esses são materiais em que o processo de transposição didática pode ser evidenciado (FORQUIN, 1992) e, consequentemente, podem revelar disfunções inadequadas. Para identificá-las o professor precisa ter, além de atenção, domínio do conhecimento produzido pela ciência. Vários pesquisadores têm apontado elementos que merecem ser melhorados nesses materiais, inclusive na área de Genética (ex: EL HANI *et al.* 2007; ESCRIBANO; SAHELICES, 2004; GERICKE; HAGBERG, 2007).

Neste texto, relatamos parte dos resultados de uma pesquisa que investigou como os conhecimentos de Genética presentes nos livros didáticos se aproximam e se distanciam do conhecimento de referência utilizado na formação dos professores. Este estudo procurou ampliar o quadro de investigações apresentado pelas pesquisas acima citadas, propondo um novo olhar para os distanciamentos encontrados entre o livro didático e a referência. Procuraremos especificamente relatar os resultados referentes a um tipo de distanciamento frequentemente encontrado nesta investigação: as generalizações que consistem em atribuir a um conhecimento particular um caráter geral (por exemplo, quando certa característica de um grupo de organismos específico é atribuída a todos os organismos).

METODOLOGIA

Nesta pesquisa foram analisados conteúdos relacionados à área de Genética presentes em uma amostra de seis obras didáticas. Três delas eram livros brasileiros de Biologia para o Ensino Médio, aprovados pela avaliação do Programa Nacional do Livro Didático para o Ensino Médio (PNLEM) de 2007. Pensando em verificar se os seus resultados eram específicos de nossa realidade ou poderiam ser encontrados em outro contexto, também foram investigados mais três livros didáticos utilizados por professores nos Estados Unidos.

Nessas seis obras foram analisados três grupos de conteúdos relacionados ao ensino de Genética: *Meiose, Leis de Mendel e Expressão gênica*. Os textos desses livros didáticos sobre esses conteúdos foram comparados com uma literatura de referência amplamente utilizada para a formação de licenciados em Ciências Biológicas (GRIFFITHS *et al.* 2008). Quando um distanciamento entre o conteúdo presente na principal obra de referência e no livro didático foi identificado, ou quando essa principal obra de referência não continha informações suficientes para a comparação de um determinado assunto, outras fontes, igualmente consagradas, foram consultadas (ex. ALBERTS *et al.* 2008, KREBS; GOLDSTEIN; KILPATRICK, 2008, LODISH *et al.* 2007, WATSON *et al.* 2008). Ao longo do manuscrito, por brevidade, essas obras serão chamadas somente de *referência*.

Para a análise, refinou-se uma metodologia desenvolvida colaborativamente em pesquisas anteriores (DEL CARLO, 2007; FRANZOLIN, 2007; NARCISO-JUNIOR, 2008) que considera que o conhecimento escolar está sujeito a duas fontes de influência: O *laxismo*, que tende a distanciar o conhecimento a ser ensinado do conhecimento científico, visando torná-lo compreensível ao aluno; e o *rigorismo*, que tende a se opor a essa tendência, aproximando o conhecimento a ser ensinado do conhecimento científico, buscando sua correção e compromisso com os cânones científicos. Ambas as influências são essenciais na constituição do conhecimento escolar e atuam concomitantemente, tendendo a deslocar o conhecimento em sentidos opostos. Se, por um lado, a ausência de uma delas poderia gerar conhecimentos inacessíveis para quem aprende, por outro lado poderia gerar descompromisso com o conhecimento aceito pela comunidade científica. Sofrendo a ação dessas duas forças, o conhecimento ensinado na escola distancia-se do conhecimento de referência.

O esquema apresentado na figura 1 ilustra como esta pesquisa considera os distanciamentos presentes nos livros didáticos. Esta imagem não tem a pretensão de ser uma ferramenta gráfica na qual se plote precisamente cada afastamento, mas sim a intenção de representar esquematicamente as ideias que norteiam a metodologia utilizada.

Na imagem, cada faixa horizontal representa um nível de conhecimento, representando o saber científico em um primeiro nível e, subsequentemente, os saberes que transitam no ensino superior e nos diferentes níveis da educação básica. Um conhecimento poderia se distanciar da referência em um eixo *vertical* devido à necessidade de se adequar o conhecimento à faixa etária do aluno por meio da *pedagogização*, e também por um eixo horizontal, devido à necessidade de *didatização*, ambos importantes para a ocorrência do aprendizado. A *pedagogização* é aqui considerada como a apropriação do conhecimento para a compreensão do aluno em seu nível de ensino. Já a *didatização* leva em conta o uso de recursos variados como metáforas e analogias para facilitar a compreensão dos conhecimentos. Os pontos representam conhecimentos hipotéticos que sofreram distanciamentos horizontais e verticais em relação ao conhecimento científico.

Entretanto, consideramos que existe um limite de aceitabilidade sobre o quanto um conhecimento a ser ensinado pode ou não se afastar do conhecimento de referência. Este limite foi representado no esquema pelo contorno do cone.

O cone tem a base mais larga, pois quanto mais jovem é o aluno, maior é a necessidade de didatização (ocorrendo um maior distanciamento horizontal). O diâmetro do cone é diminuído ao se aproximar do conhecimento científico, devido à necessidade do aumento de rigor ocorrer na medida em que é aumen-

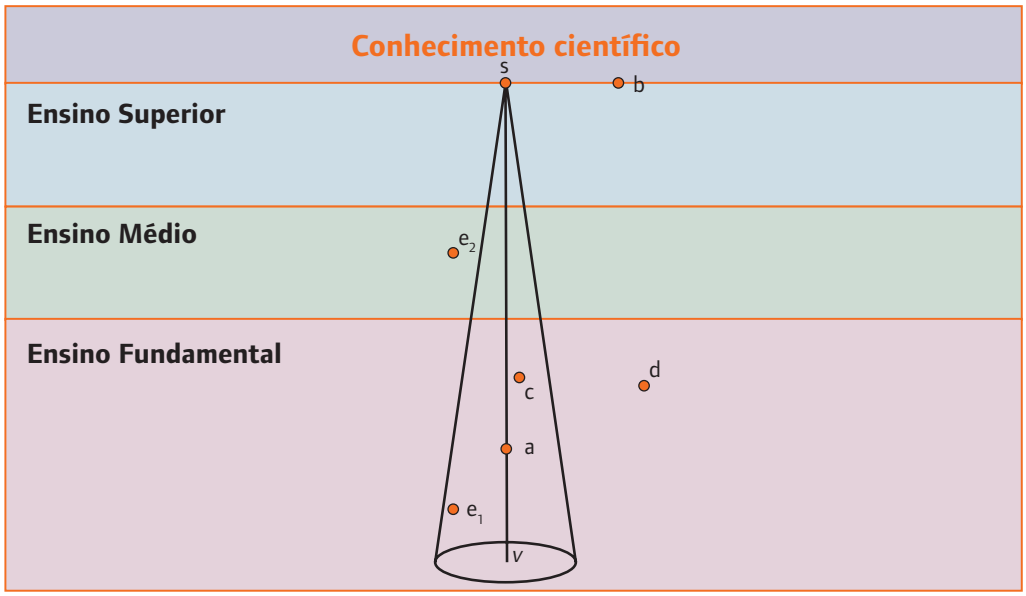


Figura 1.
Representação das ideias que norteiam a metodologia de pesquisa utilizada, ilustrando os possíveis distanciamentos encontrados entre os conhecimentos ensinados nos diferentes níveis de ensino e aqueles apresentados pela referência. O ponto *a* refere-se a um conhecimento que se distancia *verticalmente* da ciência de referência (ponto *s*, no vértice do cone), devido ao processo de *pedagogização*. O eixo *v* refere-se ao eixo de maior rigor com relação ao conhecimento científico. O afastamento dos conhecimentos em relação a este eixo decorre da necessidade de *didatização*, portanto, caracteriza-se como distanciamento horizontal. Desta forma, o ponto *b* refere-se a um conhecimento que se distancia *horizontalmente* da referência. Os pontos *c* e *d* referem-se a conhecimentos ensinados no Ensino Fundamental que se distanciam verticalmente e horizontalmente da referência. Entretanto, *c*, por ser considerado um conhecimento que sofre um *distanciamento aceitável*, encontra-se dentro do cone, e *d*, ao ser caracterizado por ter um *distanciamento não aceitável*, encontra-se fora do cone. Já o ponto e_1 representa um conhecimento que, ao ser ensinado no Ensino Fundamental, é também considerado como um conhecimento que possui um *distanciamento aceitável*, pois é proveniente de uma transposição didática necessária ao nível de ensino correspondente. Entretanto, devido ao maior rigorismo no Ensino Médio, este mesmo conhecimento, representado também pelo ponto e_2, é representado fora do cone, pois seu distanciamento não é mais considerado aceitável neste nível de ensino.

tado o nível de ensino. Portanto, os limites de quão aceitáveis são os distanciamentos são mais amplos na base do sistema educativo e se restringem progressivamente.

Os conhecimentos representados dentro do cone (pontos a, c e e1) são aqueles cujos distanciamentos em relação ao conhecimento científico (ponto s) de referência são considerados aceitáveis, pois são importantes ou necessários para que a aprendizagem ocorra. Já os situados fora do cone (pontos b, d e e2) estariam fora dos limites desta aceitabilidade, pois além de não serem necessários para a faixa etária do aluno, podem dificultar o aprendizado, na medida em que se torna muitas vezes incompatível com os novos conhecimentos transmitidos.

A metodologia aqui utilizada diferencia-se de outras utilizadas para analisar livros didáticos pois não considera como erros conceituais os distanciamentos entre o conhecimento científico e o livro didático. Reconhecemos tanto a necessidade da transposição didática, como a necessidade de se impor limites aos distanciamentos, e ainda, que os distanciamentos podem ser de diferentes naturezas. Um caso particular de distanciamento, aprofundado neste trabalho, são as generalizações, que consistem na atribuição de um caráter geral a um conhecimento particular.

Para quantificar os distanciamentos e aproximações e, consequentemente, as generalizações, foi preciso estabelecer *unidades de registro*. Estas, segundo Bardin (2007), são unidades de significação, as quais se caracterizam como "segmentos da mensagem" que farão parte da "contagem frequencial". Portanto, procurou-se analisar todos os conhecimentos referentes a cada conteúdo analisado (*Meiose, Leis de Mendel e Expressão gênica*), fracionando cada frase do texto em informações a serem comparadas com a bibliografia de referência, as quais foram consideradas como *unidades de registro*. Por exemplo, na frase "[...] No processo de produção do RNA, denominado transcrição gênica, as duas cadeias do DNA se separam e uma delas serve de molde ao RNA; [...]" (LDA, v. 1, p. 250), foram identificadas três unidades de registro: 1) "O processo de síntese de RNA é denominado transcrição"; 2)"Para a transcrição, é preciso que as duas cadeias de DNA se separem"; 3) "O RNA é produzido a partir de um molde de DNA". É também importante deixar claro que a mesma informação foi contabilizada como unidade de registro apenas uma vez no mesmo livro, independentemente de ser repetida pelo autor. Porém, para cada livro em que ela aparecia, um registro diferente era realizado.

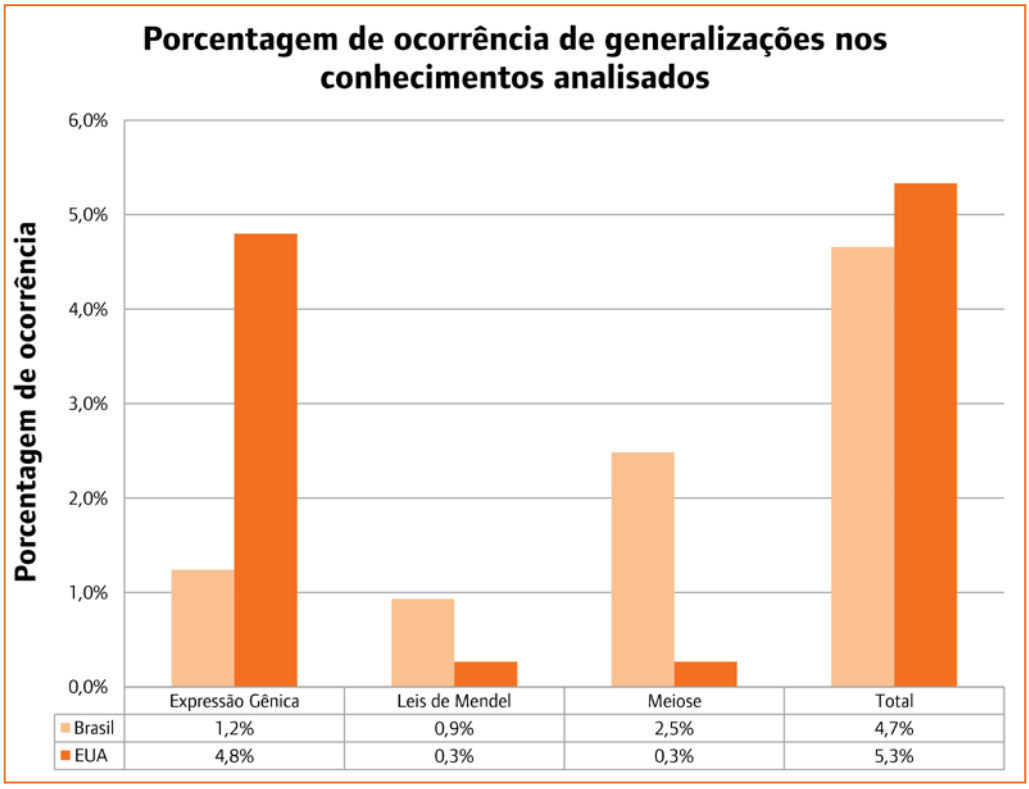
Neste artigo, as obras analisadas serão mencionadas por códigos, sendo LDA, LDB e LDC a designação dos livros brasileiros e LDJ, LDK, LDL, a dos livros estadunidenses. Ao final, para fins de consulta e esclarecimentos, inserimos uma chave de identificação dos livros analisados.

RESULTADOS

Considerando que o número de conhecimentos analisados (*unidades de registro*) nos livros do Brasil foi 322 e nos dos EUA foi de 375, e que o número de generalizações encontradas foi de 15 no primeiro contexto e 20 no segundo; ao se calcular e comparar a porcentagem de conhecimentos que se caracterizam como generalizações em cada contexto pode-se perceber que, de forma geral, sem se ater a conteúdos específicos do ensino de Genética, a porcentagem de generalizações nos livros didáticos brasileiros é próxima à dos estadunidenses (Figura 2).

Contudo, existe uma distribuição desigual de generalizações de acordo com o conteúdo analisado, sendo que os livros brasileiros têm maior incidência de generalizações no conteúdo 'meiose', e nos EUA a maior frequência está no conteúdo 'expressão gênica'.

Figura 2.
Gráfico representando a porcentagem de conhecimentos analisados que foram classificados como distanciamentos do tipo generalização nos livros didáticos brasileiros e estadunidenses (número de conhecimentos analisados: Brasil = 322, EUA =375; número de generalizações encontradas: Brasil = 15, EUA =20).



EXPRESSÃO GÊNICA: GENERALIZAÇÕES AO ENSINAR OS PROCESSOS DE TRANSCRIÇÃO E TRADUÇÃO

A maioria das generalizações foi encontrada durante a análise dos conhecimentos sobre expressão gênica.

Uma primeira generalização foi evidenciada em dois livros estadunidenses (LDK e LDL), que afirmam que a expressão gênica é a expressão do ácido desoxirribonucleico (DNA) para o ácido ribonucleico (RNA) e do RNA para proteína, sem deixar claro, como faz a referência, que a expressão gênica pode ter como produto final simplesmente um RNA funcional (ALBERTS *et al.* 2008, KREBS; GOLDSTEIN; KILPATRICK, 2008). Apesar de os autores dos livros didáticos se referirem às diferentes moléculas de RNA existentes, acabam generalizando ao não deixar claro que nem sempre o processo de expressão gênica resultará em uma proteína.

Já outros livros, como LDA, LDB, LDC e LDJ aproximam-se da referência ao deixar mais claro que os genes, ou o DNA, podem dar origem a diferentes tipos de RNAs, e não apenas ao mRNA que será utilizado na produção de uma proteína: "As moléculas de RNA transcritas a partir do DNA podem ser de três tipos principais: RNA mensageiro (RNAm), RNA ribossômico (RNAr) e RNA transportador (RNAt)" (LDA, v. 1, p. 251).

Todavia, a maioria das generalizações que aparece em relação a esse conteúdo ocorre quando os livros detalham mecanismos da expressão gênica sem mencionar que estão descrevendo aqueles que ocorrem apenas em eucariontes, ou procariontes, ou àqueles que não ocorrem em todos os organismos. Em decorrência, existe a possibilidade de que o aluno considere que os processos descritos são universais.

Isso foi evidenciado quando os livros didáticos informam que a síntese de RNA ocorre no núcleo (LDK e LDL) e o RNA sintetizado *migra* para o citoplasma (LDL). Como organismos procariontes não possuem núcleo, essa descrição não procederia para esses seres. A referência (ALBERTS *et al.* 2008, KREBS; GOLDSTEIN; KILPATRICK (2008), LODISH *et al.* 2007) atesta que em procariontes, tanto a transcrição, como a tradução ocorrem no mesmo compartimento celular.

Outros livros, contudo, conseguem evitar essa generalização, dizendo que a síntese de RNA em procariontes ocorre no nucleóide ("[...] Nos organismos procarióticos, que não apresentam núcleo, a síntese desses tipos de RNA ocorre no nucleóide, a região da célula onde se localiza o cromossomo desses organismos. [...]" (LDA, v. 1, p. 251)) ou, de forma menos específica, ao dizer que esse processo ocorre no citoplasma ("'Transcription takes place in the nucleus of eucayotic cells and in the DNA-containg region in the cytoplasm of prokaryotic cells") (LDJ, p. 206) ou simplesmente deixando claro que estão descrevendo o processo que ocorre em eucariontes (LDB e LDJ).

Outra generalização similar ocorreu no tratamento do processo de tradução. A referência diz que após sair do sítio P, o tRNA, que carrega o primeiro tRNA, vai para o sítio do ribossomo em procariontes (GRIFFITHS *et al.* 2008, KREBS; GOLDSTEIN; KILPATRICK, 2008), e dirige-se diretamente para o citoplasma em eucariontes (KREBS; GOLDSTEIN; KILPATRICK, 2008). Entretanto, tanto o LDK como os LDA, LDJ e LDL distanciam-se da referência ao não especificarem que a passagem do tRNA pelo sítio E, ou sua saída diretamente para o citoplasma, depende do tipo de organismo em que a tradução está ocorrendo, generalizando a informação apresentada como se ela ocorresse em qualquer tipo de célula.

Já um dos livros estadunidenses (LDK), distancia-se da referência ao dizer que o código começa a ser traduzido desde a ponta 5' do mRNA, desconsiderando que, entre essa ponta e o código de início, existe um trecho que não é traduzido, conforme menciona a referência (GRIFFITHS *et al.* 2008b, p. 336).

Mais um distanciamento presente em um livro brasileiro (LDB) refere-se a como ocorre a conexão entre um aminoácido e o tRNA. No livro diz-se apenas que o tRNA captura os aminoácidos, enquanto a referência diz que a ligação entre um aminoácido e um tRNA é realizada por uma enzima denominada aminoacil-tRNA sintetase (GRIFFITHS *et al.* 2008, KREBS; GOLDSTEIN; KILPATRICK, 2008). Todavia, há livros didáticos que se aproximam da referência ao dizer que há enzimas que realizam a ligação entre o aminoácido e o tRNA: "Enzymes first attach a specific amino acid to one end of each tRNA according to the genetic code" (LDJ, p. 208) e [...] The tRNA is folded into a cloverleaf shape and is activated by an enzyme that attaches a specific amino acid to the 3'end. [...] (LDK, p. 338).

Um livro brasileiro (LDB) e outro estadunidense (LDJ) também generalizam dizendo que o que confere a especificidade entre o tRNA e o aminoácido é o anticódon, enquanto a referência diz que o anticódon *pode* determinar qual será o aminoácido a ser ligado ao tRNA. Entretanto, dependendo da sintetase, outros mecanismos podem ser determinantes, como a leitura pela sintetase de outros nucleotídeos presentes em diferentes posições do tRNA. Neste caso, a sequência do anticódon acaba não sendo necessariamente requerida para o reconhecimento da tRNA sintetase (ALBERTS *et al.* 2008, KREBS; GOLDSTEIN; KILPATRICK, 2008).

Outra generalização encontrada em um dos livros brasileiros (LDB) e nos três estadunidenses ocorre quando informam que o anticódon pareia-se ao códon complementar do mRNA. Para a referência, o anticódon do tRNA *pode* parear-se com o mRNA complementar, mas há pareamentos onde nem

todas as bases do anticódon são complementares ao do códon. Muitas células possuem menos do que 61 tRNAs e, portanto, há casos em que um mesmo tRNA pareia com mais de um códon para o mesmo aminoácido por meio de um pareamento não padrão, em que geralmente os dois primeiros pares de bases são complementares mas o terceiro, não (LODISH *et al*. 2007).

MEIOSE: GENERALIZAÇÕES AO ENSINAR COMO OS GENES SE DISTRIBUEM NOS GAMETAS

A maioria das generalizações referentes ao processo de meiose ocorre quando o livro didático também trata como universal etapas que não ocorrem em todos os organismos.

Dois livros didáticos, um brasileiro (LDC) e outro estadunidense (LDL) afirmam que durante a meiose os centríolos dirigem-se para os polos opostos da célula. Entretanto, não especificam que estão falando estritamente da divisão celular em células animais. Segundo a referência (ALBERTS *et al*. 2008), os centríolos são estruturas encontradas em células desse grupo de organismos, mas há células de outros seres como plantas superiores e até mesmo muitos oócitos animais, que não as possuem. Dessa forma, o aluno que lê tal descrição de meiose pode pensar que todas as células possuem essas estruturas e se dividem da mesma forma.

O mesmo tipo de distanciamento ocorre em outros livros brasileiros quando falam sobre a finalização da primeira divisão da meiose. Eles descrevem simplesmente que ocorrem processos como descondensação dos cromossomos homólogos (LDA, LDB, LDC), reconstituição da membrana nuclear (LDB, LDC) e a interfase (LDB). Todavia, não deixam claro, como faz a referência (GRIFFITHS *et al*. 2008b, LODISH *et al*, 2007), que isso ocorre apenas em determinados grupos de organismos.

Esta generalização foi facilmente evitada utilizando-se algumas expressões como "em algumas espécies" (LDJ e LDK) que a indicam como uma possibilidade (LDC e LDK). Por exemplo: "In some species, the chromosomes uncoil, the nuclear membrane reappers, and nuclei re-form during telophase I" (LDK, p. 274); "Entre o final da primeira divisão e o início da segunda, pode haver um pequeno intervalo chamado **intercinese**, em que não ocorre duplicação do DNA (LDC, p. 106. Grifo no original.).

Uma última generalização relacionada à meiose foi identificada em um dos livros brasileiros (LDB) que diz que os cromossomos homólogos, que migram para os polos *opostos* da célula, durante a primeira divisão da meiose, são *idênticos*, exceto nos lugares onde ocorreu permuta. Todavia, ao resumir o processo de meiose, a referência (GRIFFITHS *et al*., 2008b) deixa claro que cada um dos cromossomos do par de homólogos pode carregar alelos diferentes desde o início do processo de meiose. Isso evidencia que os homólogos podem não ser idênticos, simplesmente pelo motivo de terem alelos diferentes, e não somente devido à recombinação. Portanto, durante a primeira divisão da meiose, um homólogo que migra para um dos polos da célula é diferente do seu par que migra para o outro.

LEIS DE MENDEL: GENERALIZAÇÕES AO FALAR SOBRE LINHAGENS PURAS E SEGREGAÇÃO INDEPENDENTE

Apenas duas generalizações foram encontradas durante a análise do tópico Leis de Mendel. Uma dessas foi identificada quando alguns livros se referem às linhagens puras.

A referência, assim como parte dos livros didáticos analisados, explica essas leis apoiando-se nas explicações de Mendel sobre seus experimentos e resultados. Como esses experimentos envolvem o cruzamento de organismos de linhagens puras, a referência define que uma linhagem pura com relação a um fenótipo específico é aquela capaz de produzir, quando cruzada consigo mesma, uma prole em que todos os indivíduos possuam esse mesmo fenótipo (GRIFFITHS *et al.* 2008b). Já dois livros, um brasileiro (LDA) e outro estadunidense (LDL) distanciam-se da referência ao considerar que uma linhagem pura é aquela que, quando cruzada consigo mesma, produz indivíduos iguais a si, podendo o leitor interpretar que todas as características dessa linhagem são herdadas por sua prole, e não apenas um fenótipo específico que está sendo considerado, mesmo que o exemplo se refira apenas a uma característica.

Todavia, outros livros se aproximam da referência. O LDC diz que **linhagens puras** são aquelas que, por autofecundação, produzem apenas descendentes que em *determinadas* características são iguais à geração progenitora: "Mendel iniciou seus trabalhos obtendo linhagens puras, isto é, plantas que, por autofecundação, só originavam descendentes iguais a elas mesmas, em relação a determinadas características [...]" (LDC, p. 120-121. Grifo no original.). O LDJ também se aproxima da referência ao dizer que uma linhagem pura para uma determinada característica é aquela que, na autopolinização, produz apenas indivíduos com essa mesma característica: "[...] Plants that are **true-breeding**, or pure, for trait always produce offspring with that trait when they self-pollinate. [...]" (LDJ, p. 175. Grifo no original.).

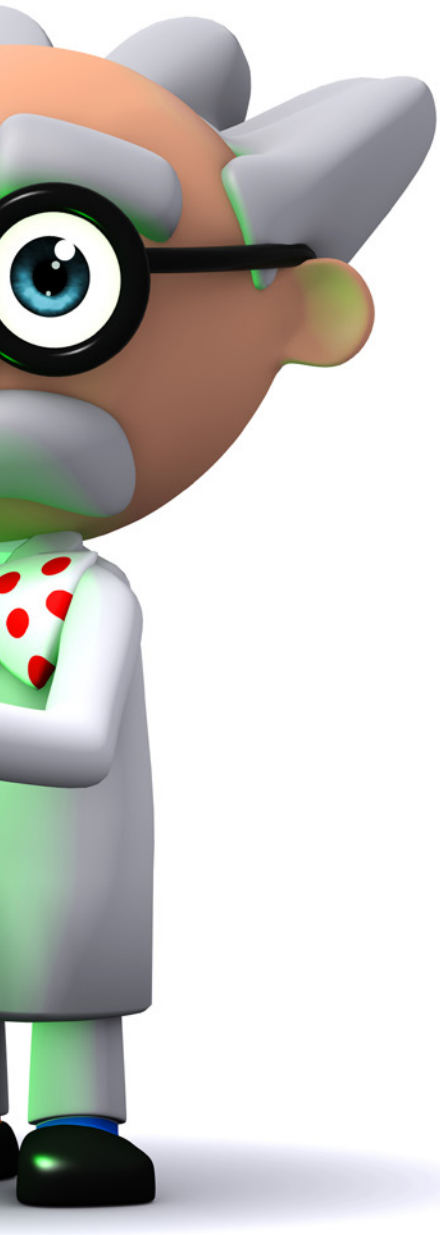
A outra generalização foi identificada quando dois livros consideram como segregação independente as que ocorrem em genes localizados em diferentes pares de cromossomos. Não especificam, como faz a referência (GRIFFITHS *et al.*, 2008b, WATSON *et al.* 2007), que elas podem também ocorrer em genes localizados no mesmo cromossomo, se houver uma longa distância entre eles (LDA e LDC), devido ao processo de permuta. Apesar de apresentarem separadamente capítulos que falam sobre a permuta, os dois livros didáticos não as mencionam como possibilidade de segregação independente.

CONCLUSÕES E CONSIDERAÇÕES FINAIS

Os dados pesquisados mostram que as generalizações estão presentes nos conteúdos de Genética nos livros didáticos brasileiros, porém não são exclusivas do nosso contexto. Foi também possível perceber, na amostra analisada, que a frequência de generalizações muda de acordo com o conteúdo para cada localidade, sendo nos livros brasileiros mais presentes no tópico meiose e nos livros estadunidenses bem concentradas nos conhecimentos sobre expressão gênica.

Algumas pesquisas já trazem dados sobre problemas na apresentação dos conteúdos de Genética encontrados nos livros didáticos (ex. EL HANI *et al.* 2007; FERREIRA; JUSTI 2004; GERICKE; HAGBERG, 2007). Todavia, a presente pesquisa traz um novo elemento a esse quadro de investigação, que consiste em não apenas tratar os distanciamentos como erros e sim explorar a diversidade de naturezas que eles possam ter. Neste artigo, exploramos uma dessas possibilidades: as generalizações.

Optamos por não considerar essas generalizações simplesmente como erros conceituais, pois, para ser considerado um erro conceitual, é preciso que o conhecimento tenha necessariamente quatro aspectos, os quais vêm sendo utilizados durante vários anos em processos oficiais de avaliação de livros didáticos: "[...] estar baseado em premissas que não são aceitas pela comunidade científica; constituir uma forma de conceber e explicar relações (modelo mental); estar fundamentado em conhecimentos adaptativos no contexto escolar; demonstrar-se falso e não adaptativo em situações da vida real [...]"



(BIZZO, 2009, p. 45). As generalizações aqui encontradas não atendem a todos esses aspectos conjuntamente, principalmente quanto ao segundo mencionado.

Por exemplo, o livro (LDL) diz que a síntese de RNA ocorre no núcleo e o RNA sintetizado *migra* para o citoplasma, apesar de não dar maiores detalhes sobre o fato de estar se referindo apenas a eucariontes; por outro lado, fica claro em outros momentos que não faz parte do modelo mental dos autores a existência de apenas um modelo de célula. Por exemplo, cerca de cem páginas antes do tratamento da transcrição gênica, ao falar sobre procariontes e eucariontes, eles afirmam que os organismos procariontes não possuem material genético organizado na forma de núcleo (LDL, p. 173). Apesar de tais generalizações não serem caracterizadas como erros conceituais, consideramos que é preciso refletir sobre a pertinência do uso dessas generalizações. Durante a exposição dos resultados procuramos apresentar alguns exemplos de como alguns livros evitam essas generalizações. A simples referência sobre o grupo de organismos ao qual o processo se refere (eucariontes ou procariontes), ou o uso de termos que relativizem as informações (como "geralmente"), de verbos que indiquem possibilidade e podem ajudar a evitá-las sem, entretanto, precisar entrar em detalhes desnecessários aos estudantes para os quais esses conteúdos estão sendo introduzidos.

Acreditamos que reconhecer essas generalizações é importante para os professores, para que possam relativizá-las para os alunos. Nesse sentido, o ensino do conteúdo específico da área de conhecimento é essencial durante a formação do professor. Entretanto, a resolução brasileira que institui as Diretrizes Curriculares Nacionais para a educação básica diz que "os conteúdos a serem ensinados na escolaridade básica devem ser tratados de modo articulado com suas didáticas específicas" (BRASIL, 2002, p. 2). É preciso também que o futuro docente aprenda a ensinar esses conhecimentos, sabendo como utilizar estratégias metodológicas, conhecer possibilidades de contextualização, e ainda aprender a realizar a transposição didática e a vigilância epistemológica mencionadas por Chevallard (1991), sendo capaz de avaliar a pertinência tanto dos distanciamentos quanto das aproximações realizadas entre o conhecimento de referência e os conteúdos ensinados pelos livros didáticos e por ele próprio. As generalizações aqui apontadas podem da mesma forma estar presentes na fala do próprio professor. É também por isso que, durante sua formação, é essencial que ele aprenda a refletir sobre a sua própria pertinência e o seu próprio objetivo.

Diante dessa necessidade, não é possível pensarmos em um curso de licenciatura que acredita que a qualidade da formação do professor esteja apenas atrelada ao conhecimento do saber erudito. É preciso que, juntamente com esses conhecimentos, o futuro docente exercite a reflexão responsável e cuidadosa sobre como ensiná-los. Esse processo pode ser iniciado durante sua formação, desde os primeiros momentos, durante as aulas de conhecimento específico de Genética e em outras disciplinas específicas de conhecimentos da Biologia, nas quais pode tanto realizar exercícios de transposição didática, quanto analisar as referentes aos materiais didáticos.

É importante que os professores reconheçam essas – e muitas outras – generalizações a fim de que possam relativizá-las para os alunos quando as julgarem necessárias. Uma vez reconhecida a generalização, o professor pode, desde seu planejamento até a avaliação do conhecimento, explorar as peculiaridades do ensino da Ciência que envolvem, entre outros tantos aspectos, a transposição didática. Ou seja, nem sempre se aprende na escola o conhecimento tal qual ele é praticado, há um intervalo – de tempo e de profundidade – entre a sua produção na academia e a sua difusão nos bancos escolares. Conforme afirma Chevallard (1991), o conhecimento ensinado pode envelhecer e desacordos podem ser gerados pelo progresso na produção do conhecimento científico.

Uma das formas de se evitar as generalizações é a constante procura por atualização por parte do professor. Como o livro didático é uma produção editorial com um longo percurso entre sua concepção e publicação, os novos conhecimentos gerados na academia demandam anos para figurarem em seus textos. Os livros avaliados pelo Programa Nacional do Livro Didático (PNLD), por

exemplo, são apresentados ao MEC dois anos antes de serem entregues aos alunos, com os quais permanecem por três anos. Portanto, de saída, eles não podem incluir conhecimentos que tenham sido gerados até os cinco últimos anos.

Todavia, nem todas as generalizações envolvem a ausência de conhecimentos novos da Ciência. Muitas incidem sobre conhecimentos já consensuais. O exemplo de que alguns livros afirmam simplesmente que a síntese de RNA ocorre no núcleo, generalizando, sem levar em conta a existência desses processos em organismos procariontes, ilustra essa situação. Quando os autores escreveram seus textos, o conhecimento sobre a existência de organismos procariontes já era um consenso. Apenas para se ter uma ideia, uma edição da referência (ALBERTS *et al.* 1997), publicada cerca de dez anos antes da publicação do referido livro didático, já trazia diferenciações entre os processos de transcrição de procariontes e eucariontes, o que já não era novidade na época dessa publicação. Outros livros didáticos do mesmo período aqui analisados, todos em que foi encontrado tal distanciamento, também apresentavam essa informação.

Pode-se cogitar a hipótese do autor, ou de um professor em sua sala de aula, ter a intenção de omitir essa informação pelo fato de ainda não ter trabalhado conhecimentos prévios com os alunos. Todavia, não é o caso, pois, no próprio livro, conhecimentos sobre eucariontes e procariontes são apresentados previamente. Um distanciamento pode ser realizado na sala de aula, quando o professor sente que ele é necessário para a compreensão do aluno, já que, segundo Chevallard (1991), a transposição didática é importante para adequar o conhecimento a ser ensinado em um conhecimento compreensível para o aluno. No entanto, a vigilância epistemológica seria importante para se considerar até quando os distanciamentos gerados são positivos e necessários para a compreensão do aluno.

Cabe ao professor, atento às inovações da ciência, às novas demandas vindas dos alunos, e ao consenso já gerado na academia, não apenas analisar e agilizar essa transposição, mas também ser cuidadoso durante esse processo, procurando meios de relativizar as generalizações. O professor não precisa necessariamente descartar o livro didático para evitá-las. Uma leitura critica e atenta com os alunos, apontando-as, e o uso de expressões simples como *"geralmente"*, *"pode"*, *"na maioria das vezes"*, *"em organismos eucariontes"*, podem ser a solução em muitos casos.

REFERÊNCIAS

ALBERTS, B.; JOHNSON, A.; LEWIS, J; RAFF, M.; ROBERTS, K.; WALTER, P. *Molecular Biology of the Cell*. 5.th ed.. New York e London: Garland Science, 2008.

ALBERTS, B.; BRAY, D.; LEWIS, J; RAFF, M.; ROBERTS, K.; WATSON, J. D. *Biologia Molecular da Célula*. 3. ed. Porto Alegre: Artmed, 1997.

ASTOLFI, J.P., E. DAROT, Y. DRINSBURGER-VOGEL E J. TOUSSAINT. As palavras-chave da didática das ciências. (tradução de Maria Ludovina Figueiredo), Lisboa: Instituto Piaget, 2000

BARDIN, L. *Análise de conteúdo*. Lisboa: Edições 70, 2007.

BIZZO, N. *Mais Ciência no Ensino Fundamental*: metodologia de ensino em foco. São Paulo: Editora do Brasil, 2009.

BRASIL, Resolução CNE/CP 01/2002. Brasília: Conselho Nacional de Educação, 2002.

CHEVALLARD, Y. *La Transposición Didáctica*: Del saber sabio al saber enseñado. Buenos Aires: Aique, 1991.

COBERN, W. W. Worldview Theory and Conceptual Change in Science Education. *Science Education*, v. 80, n. 5, p. 579-610, 1996.

DEL CARLO, S. *Conceitos de Física na educação básica e na Academia: aproximações e distanciamentos*. 2007.97 p. Tese (Doutorado). Faculdade de Educação, Universidade de São Paulo, São Paulo, 2007.

EL-HANI, C. N.; ROQUE, N.; VANZELA, A. L. L.; SOUZA, A. F. L.; MARQUES, A. C.; VIANA, B. F.; KAWASAKI, C. S.; LEME, C. L. D.; FARIA, D.; MEYER, D.; OMENA, E.; OLIVEIRA, E. S.; ASSIS, J. G. A.; FREGONEZE, J.; QUEIROZ, L. P.; CARVALHO, L. M.; NAPOLI, M.; CARDOSO, M. Z.; SILVEIRA, N. A.; HORTA, P. A.; SANO, P. T.; ZUCOLOTO, R. B.; TIDON, R.; SILVA, S. A. H.; ROSA, V. L.; ROCHA, P. L. B. Brazilian High School Biology Textbooks: Main

Conceptual Problems in Genetics and Cell & Molecular Biology. In: INTERNATIONAL MEETING ON CRITICAL ANALYSIS OF SCHOOL SCIENCE TEXTBOOK, 2007, Hammamet, Tunisie. *Proceedings*... Hammamet, Tunisie: International Organisation for Science and Technology Education, 2007.

ESCRIBANO, D. D.; SAHELICES, C. C. Imágenes externas de gen y cromosoma en materiales instruccionales para la enseñanza de la Biología en el sistema educativo venezolano. *Revista Brasileira de Pesquisa em Educação em Ciências*, v. 4, n. 2, p. 74-86, 2004.

FORQUIN, J. C. Saberes escolares, imperativos didáticos e dinâmicas sociais. *Teoria e Educação*, Porto Alegre, n. 5, p. 28-49, 1992.

FRANZOLIN, F. *Conceitos de Biologia na educação básica e na Academia: aproximações e distanciamentos*. 2007. 162 p. Dissertação (Mestrado). Faculdade de Educação, Universidade de São Paulo, São Paulo, 2007.

GERICKE, N. M.; HAGBERG, M. The phenomenon of gene function as described in textbooks for upper secondary school in Sweden - a comparative analysis with historical models of gene function. In: INTERNATIONAL MEETING ON CRITICAL ANALYSIS OF SCHOOL SCIENCE TEXTBOOK, 2007, Hammamet, Tunisie. *Proceedings*... Hammamet, Tunisie: International Organisation for Science and Technology Education, 2007.

GRIFFITHS, A. J. F.; WESSLER, S. R; LEWONTIN, R. C.; CARROLL, S. B. *Introduction to Genetic Analysis*. 9.th ed. New York: Feeman and Company, 2008b.

KREBS, J; GOLDSTEIN, E. S.; KILPATRICK, S.T. *Lewin's Genes X*. 10.th ed. Boston, Toronto, London, Singapore: Jones & Barlet Pub, 2008.

LODISH, H. ; BERK, A.; KAISER, C. A.; KRIEGER, M.; SCOTT M. P.; BRETSCHER, A. PLOEGH, H. *Molecular Cell Biology*. 6.th ed. New York: WH Freeman & Co., 2007. 1150 p.

LOPES, A. R. C. *Conhecimento escolar: ciência e cotidiano*. Rio de Janeiro: EdUERJ, 1999.

NARCISO-JUNIOR, J. *Conceitos de Química na educação básica e na Academia: aproximações e distanciamentos*. 2008. 87 p. Dissertação (Mestrado). Programa Interunidades de Ensino de Ciências, Universidade de São Paulo, São Paulo, 2008.

WATSON, J. D.; BAKER, T. A.; BELL, S.P.; GANN, A; LEVINE, M.; LOSICK, R. *Molecular Biology of the Gene*. 6.th ed. San Francisco: Pearson Education, 2006.

ZIMAN, J. *Enseñanza y aprendizaje sobre la ciencia y la sociedad*. México: Fondo de Cultura Económica, 1985.

APÊNDICE A

Chave de identificação dos livros didáticos utilizados como amostra de livros para análise nesta pesquisa

Código na amostra	Editora	Autores	Título	Edição
LDA	Moderna	José Mariano Amabis e Gilberto Rodrigues Martho	Biologia	2ª Edição – 2005
LDB	IBEP	Augusto Adolfo, Marcos Crozetta e Samuel Lago	Biologia – Coleção Vitória Regia	2ª edição – 2005
LDC	Moderna	José Arnaldo Favaretto e Clarinda Mercadante	Biologia	1ª edição – 2005
LDJ	Holt, Rinehart and Winston	John H. Postlethwait e Janet L. Hopson	Modern Biology	2009
LDK	Glecoe	Alton Biggs *et al*.	Biology	2009
LDL	Prentice Hall	Kenneth R. Miller e Joseph Levine	Biology	2008

Avaliação da depressão endogâmica em populações de araticunzeiro: apoio pedagógico ao Ensino de Genética*

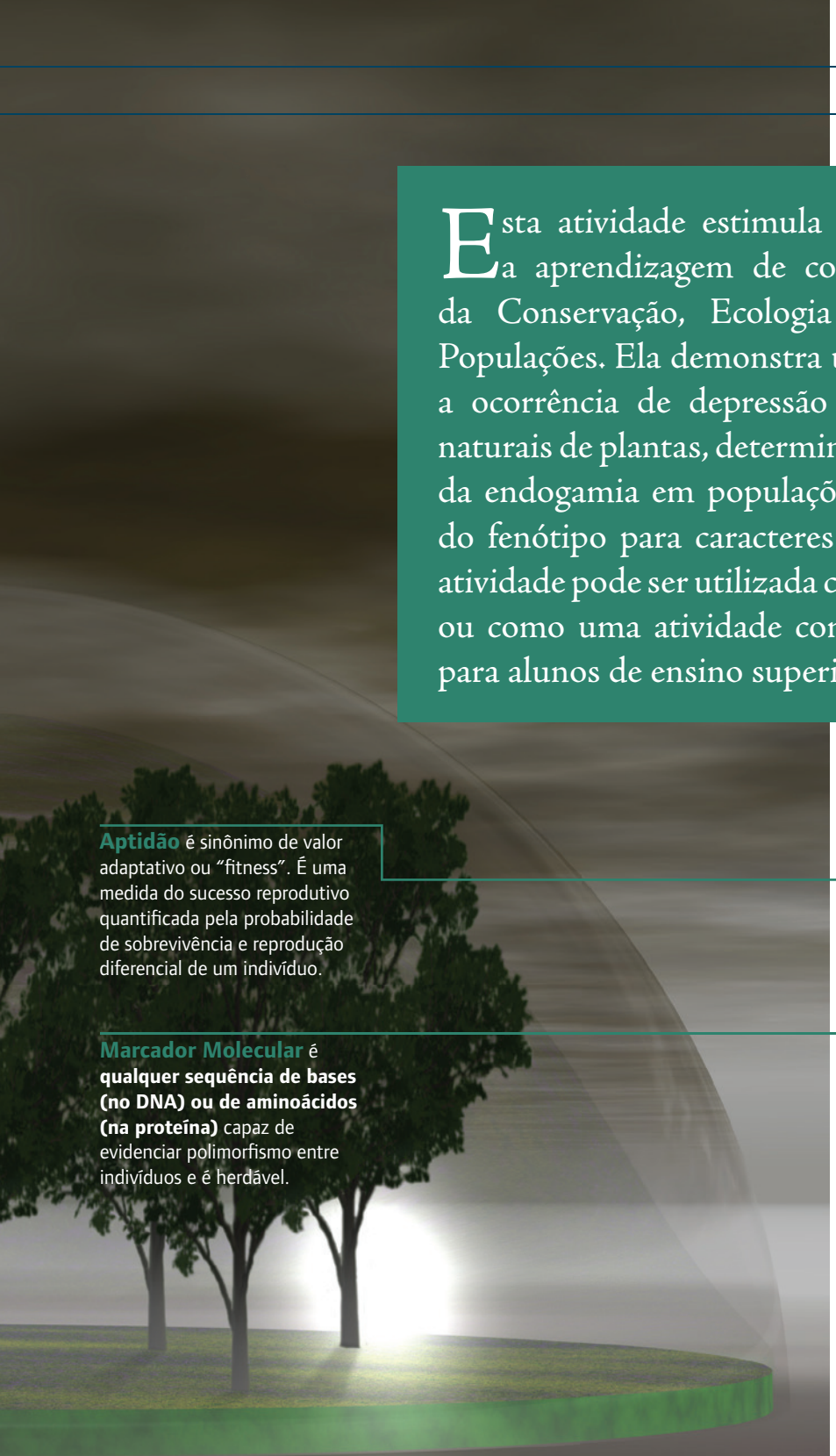
Edivaldo Barbosa de Almeida Júnior[1,3]; Núbia Esther de Oliveira Miranda[1,2]; Rosane Garcia Collevatti[1,2]

[1] Laboratório de Genética & Biodiversidade, ICB, Universidade Federal de Goiás, Goiânia, GO.
[2] Programa de Pós-Graduação em Ecologia e Evolução, ICB, Universidade Federal de Goiás (UFG), Goiânia, GO.
[3] Programa de Pós-Graduação em Genética e Melhoramento de Plantas, Escola de Agronomia, Universidade Federal de Goiás, Goiânia, GO.

E-mail para correspondência: rosanegc68@hotmail.com.

*Material didático desenvolvido na disciplina de Genética da Conservação, coordenado pela Profa. Rosane Garcia Collevatti, do curso de graduação em Ecologia e Análise Ambiental do Departamento de Ecologia, Instituto de Ciências Biológicas, Universidade Federal de Goiás, como uma atividade do Estágio Docência (bolsistas CAPES/UFG) dos discentes dos Programas de Pós-graduação em Ecologia e Evolução e Genética e Melhoramento de Plantas.

Esta atividade estimula o raciocínio lógico e auxilia a aprendizagem de conceitos básicos de Genética da Conservação, Ecologia Molecular e Genética de Populações. Ela demonstra uma metodologia para avaliar a ocorrência de depressão endogâmica em populações naturais de plantas, determinada a partir do conhecimento da endogamia em populações naturais e do valor médio do fenótipo para caracteres quantitativos adaptativos. A atividade pode ser utilizada como simulação de aula prática ou como uma atividade complementar à aula expositiva para alunos de ensino superior.



Aptidão é sinônimo de valor adaptativo ou "fitness". É uma medida do sucesso reprodutivo quantificada pela probabilidade de sobrevivência e reprodução diferencial de um indivíduo.

Marcador Molecular é **qualquer sequência de bases (no DNA) ou de aminoácidos (na proteína)** capaz de evidenciar polimorfismo entre indivíduos e é herdável.

Microssatélites são curtas sequências de 1 a 4 nucleotídeos de comprimento que são repetidas sequencialmente no genoma. Ex: $(CA)_{30}$ representa uma sequência de 30 repetições dos nucleotídeos citosina (C) e adenina (A). Essas regiões têm altas taxas de mutação, por isso são muito polimórficas o que as tornam adequadas para discriminar indivíduos geneticamente.

FUNÇÃO PEDAGÓGICA

A atividade procura mostrar os efeitos do acasalamento entre indivíduos aparentados na redução da **aptidão** da prole, por meio da avaliação de caracteres adaptativos, ou seja, caracteres ligados ao vigor das progênies. Nesse sentido, a atividade proposta procura utilizar dados provenientes da análise genética de **marcadores moleculares** do tipo **microssatélites**, para estimar os coeficientes de endogamia, e dados de caracteres quantitativos adaptativos para medir a aptidão da prole. Dessa forma, são trabalhados conceitos básicos de Genética de Populações e Genética Quantitativa, além de métodos de Ecologia Molecular.

OBJETIVO

O objetivo da atividade é permitir a compreensão da depressão endogâmica, ou seja, a redução da aptidão da prole devida ao cruzamento entre indivíduos aparentados. A partir de dados simulados (fictícios) do genótipo da prole oriunda de diversas plantas de diferentes populações e de caracteres quantitativos adaptativos, o estudante poderá determinar o nível de endogamia entre os indivíduos de uma prole, bem como os valores médios de caráter adaptativo. Este procedimento permite identificar se há depressão endogâmica para o caráter adaptativo estudado em proles oriundas de cruzamento entre indivíduos mais aparentados.

PROBLEMA PROPOSTO

Em populações tipicamente **exogâmicas** muitos **lócus** gênicos estão em **heterozigose** e pode ocorrer dominância entre **alelos**. A consequência desse fato é que alelos recessivos deletérios podem permanecer na população, ao longo das gerações, mascarados pela **dominância** porque esses alelos não são expostos à seleção natural. Quando ocorre acasalamento entre parentes (endogamia) é maior o número de alelos compartilhados, herdados de ancestral comum, do que quando o cruzamento ocorre entre indivíduos não relacionados da população. Por isso, alelos recessivos deletérios presentes na população podem expressar seu fenótipo, em uma proporção dos descendentes, pois estarão em **homozigose**. A implicação imediata da expressão fenotípica desses alelos é a redução no vigor da prole. Esse efeito é conhecido por depressão endogâmica.

A endogamia em populações naturais pode ocorrer naturalmente, por restrições intrínsecas da espécie à dispersão. Assim, os indivíduos de uma mesma população cruzam entre si com frequência maior que entre indivíduos de populações diferentes. Entretanto, a restrição ao **fluxo gênico** pode ser resultado da fragmentação, redução e isolamento dos hábitats causados especialmente pela atividade humana (efeitos antrópicos). A fragmentação leva à restrição do fluxo gênico e potencialmente à depressão endogâmica, comprometendo a conservação das espécies. Por isso, a depressão endogâmica causada por efeitos antrópicos deve ser minimizada nas populações naturais, em programas de manejo e conservação, com o intuito de manter o potencial evolutivo dessas populações e a sua persistência ao longo das gerações (FRANKHAM *et al.* 2003)

Nesta atividade foi proposta a avaliação da ocorrência de depressão endogâmica em populações naturais, a partir de teste de **progênie** no araticunzeiro (*Annona crassiflora* Mart., Annonaceae). Esta espécie é uma fonte de alimento e renda para pequenos agricultores no Centro-Oeste que utilizam os frutos *in natura* ou para produção de doces, sorvetes e sucos. O araticunzeiro é uma árvore autocompatível, ou seja, pode ocorrer autopolinização. A polinização é feita por besouros da espécie *Cyclocephala octopunctata* Burmeister (CAVALCANTE *et al.* 2009). A atividade foi desenvolvida com dados simulados para os genótipos e para os caracteres quantitativos. Assim, foram amostradas quatro populações naturais de *A. crassiflora* (populações 01, 02, 03 e 04). De cada população foram coletadas

Exogamia é o cruzamento entre indivíduos não aparentados ou pouco aparentados geneticamente.

Lócus é a região do cromossomo onde se localiza um gene, ou marcador molecular.

Heterozigose é o estado em que um lócus apresenta alelos diferentes nos cromossomos homólogos.

Alelos são as diferentes formas alternativas do mesmo gene.

Dominância é a interação alélica em que o alelo dominante se manifesta no fenótipo.

Progênie é a descendência ou prole de um indivíduo. É o conjunto de indivíduos que apresentam pelo menos um dos genitores em comum.

Homozigose é o estado em que um lócus apresenta o mesmo alelo nos cromossomos homólogos.

Fluxo gênico é a migração de genes entre populações. Em plantas, pode ocorrer em dois estágios – via pólen (migração de gametas) e via semente (migração de indivíduos).

sementes de 13 plantas matrizes. Foi semeada uma semente por matriz em **delineamento experimental inteiramente casualizado**. Conhecendo o parentesco dos indivíduos (quanto maior a proximidade genética, maior a semelhança fenotípica) pode-se usar um delineamento experimental inteiramente casualizado, que permita ao pesquisador separar os efeitos ambientais e genéticos na composição do fenótipo. Quatorze anos após o plantio, para avaliar a aptidão dos indivíduos, as 13 plantas já em idade reprodutiva, tiveram o lócus microssatélite Acr 12 caracterizados, para determinação do nível de endogamia das populações. Foi medido o diâmetro maior dos frutos (*DM*) de quatro árvores (A, B, C e D) por população (populações 01, 02, 03 e 04). Esse caráter quantitativo está ligado à aptidão dos indivíduos porque frutos maiores possuem mais sementes e atraem mais dispersores. Em caso de ocorrência da depressão endogâmica, espera-se que as populações com maiores valores de endogamia, estimada pelo parâmetro *f* (coeficiente de endogamia), apresentem as menores médias fenotípicas para o caráter avaliado.

Delineamento experimental inteiramente casualizado (DIC) é um delineamento experimental que se refere à forma como são organizadas as unidades experimentais. O modelo mais simples de delineamento é o inteiramente casualizado. Neste delineamento, as unidades experimentais são destinadas a cada tratamento por meio de sorteio de forma casual e considera os princípios de repetição e casualização, ou seja, os tratamentos do experimento são divididos em parcelas de forma inteiramente casual (por sorteio). Para se aplicar este tipo de delineamento experimental é necessário que o material experimental seja semelhante e que as condições de estudo sejam completamente uniformes (por exemplo, mesmo tipo de solo ou de adubação).

Para a avaliação da ocorrência de depressão endogâmica nas populações é preciso:

1. Obter os **genótipos** das 13 plantas matrizes, que foram plantadas no experimento e das quais foram amostradas os frutos, e calcular os coeficientes de endogamia (*f*) das populações.
2. Obter os dados quantitativos, ou seja, o diâmetro maior dos frutos (DM), das plantas em delineamento experimental e condições controladas, para obter as médias fenotípicas da prole de cada árvore (médias de diâmetro maior dos frutos).
3. Comparar os valores dos coeficientes de endogamia com as médias fenotípicas das plantas de cada população e interpretar os resultados observados.

Genótipo é a constituição genética de uma célula ou indivíduo para um ou mais lócus.

INSTRUÇÕES PARA O PROFESSOR

1. Esta atividade poderá ser realizada individualmente, ou em grupos de alunos de, no máximo, três pessoas.
2. Cada grupo deverá receber o problema proposto, uma cópia do procedimento para realizar a atividade, uma cópia de cada painel (1.1, 1.2, 1.3, 2.1, 2.2 e 3) e das questões para serem discutidas.
3. É recomendável que o professor aplique esta atividade em turmas que já tiveram contato prévio com os conceitos de estrutura e organização do material genético, marcadores moleculares e segregação mendeliana.

PROCEDIMENTO PARA OS ESTUDANTES

1. Ler com atenção o problema proposto.
2. Analisar o Painel 1.1 que apresenta os resultados da genotipagem do lócus microssatélite Acr 12 de treze plantas pertencentes a quatro populações diferentes.
   a. Observar os genótipos das plantas adultas e anotar os genótipos das plantas no Painel 1.2.
   b. Comparar os alelos de cada indivíduo (banda preta) com os da escada alélica "M" bandas vermelhas, para obtenção dos genótipos.
   c. Denominar os alelos dos indivíduos de forma que correspondam aos nomes dos alelos da escada alélica.
3. A partir dos genótipos obtidos, calcular as frequências alélicas e estimar a heterosigosidade esperada (*He*) sob as condições de Equilíbrio de Hardy-Weinberg (EHW) e a heterozigosidade observada (*Ho*) para estimar a endogamia (*f*) da população sendo:
   - Frequência alélica *f(Ai)*: ***f(Ai) = ni / nt***, onde
     - ni – número de alelos i presentes nos 13 indivíduos
     - nt – número total de alelos em todos os 13 indivíduos
   - Heterozigosidade esperada pelo equilíbrio de Hardy-Weinberg *He*: $H_e = 2pq + 2pr + 2qr$, sendo que
     - *p*, *q* e *r* representam as frequências de cada um dos três alelos (A1, A2

e A3, respectivamente) presentes no lócus Acr 12.

- Heterozigosidade observada (*Ho*):

$$Ho = \frac{Hobs}{N}$$, sendo

  - *Hobs* número de genótipos heterozigotos observados nos 13 indivíduos
  - *N* número total de genótipos avaliados.

- Endogamia (*f*):

$$f = \frac{He - Ho}{He};$$

Anotar no Painel 1.3 as estimativas *He*, *Ho* e endogamia (*f*) das populações 01, 02, 03 e 04.

4. Analisar o Painel 2.1 que apresenta os frutos coletados de quatro indivíduos (A, B, C e D) adultos de *Annona crassiflora*, provenientes de sementes coletadas nas populações naturais 01, 02, 03 e 04, que foram mantidos em condições controladas de solo, irrigação e adubação, em delineamento experimental inteiramente casualizado.

   a. Medir o diâmetro maior do fruto, com o auxílio de régua comum, respeitando a escala 1:6 cm.

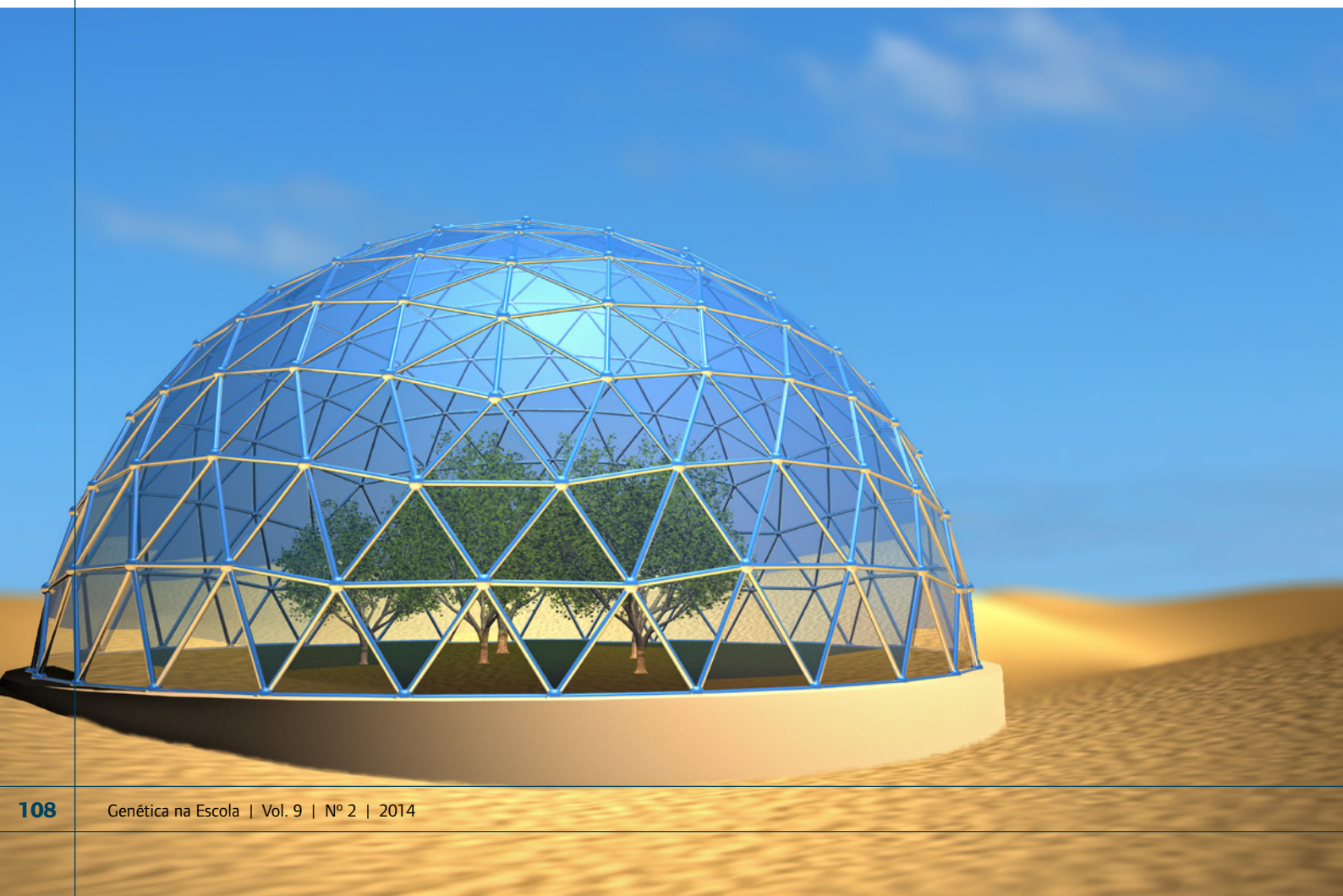
   b. Anotar os valores de diâmetro no Painel 2.2

5. Calcular a média de diâmetro maior do fruto por população e anotar no Painel 3. Transcrever para o Painel 3 os valores do coeficiente de endogamia de cada população do Painel 1.3.

6. Calcular o coeficiente de depressão endogâmica δ (LANDE; SCHEMSKE, 1985) para a(s) população(ões) endogâmica(s) para o caráter adaptativo diâmetro do fruto e transcrever para o Painel 3. As populações não endogâmicas tem valor de δ = 0,0

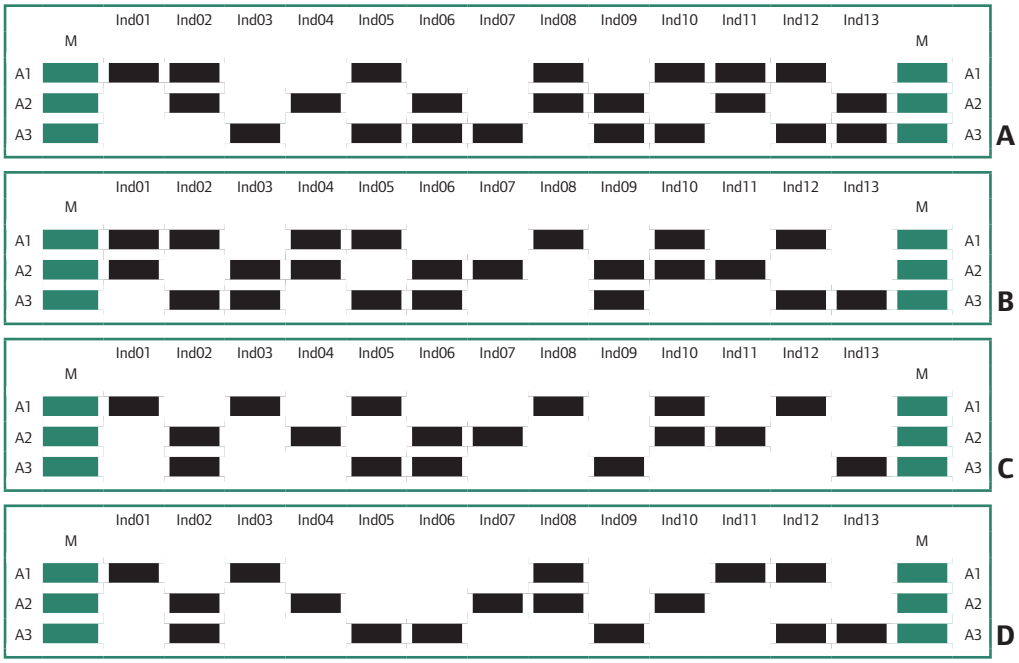
$$\delta = 1 - \left(\frac{X_1}{X_0}\right)$$, sendo

  - x_1, valor do caráter adaptativo da população endogâmica
  - x_0, valor do caráter adaptativo da população não endogâmica com o maior valor do caractere entre as populações analisadas.



PAINÉIS

Painel 1.1.
Representação esquemática do gel de poliacrilamida, corado com nitrato de prata com os genótipos para o lócus microssatélite Acr 12 de treze plantas pertencentes às populações 01; 02; 03 e 04. "M" é a escada alélica para determinação dos alelos.



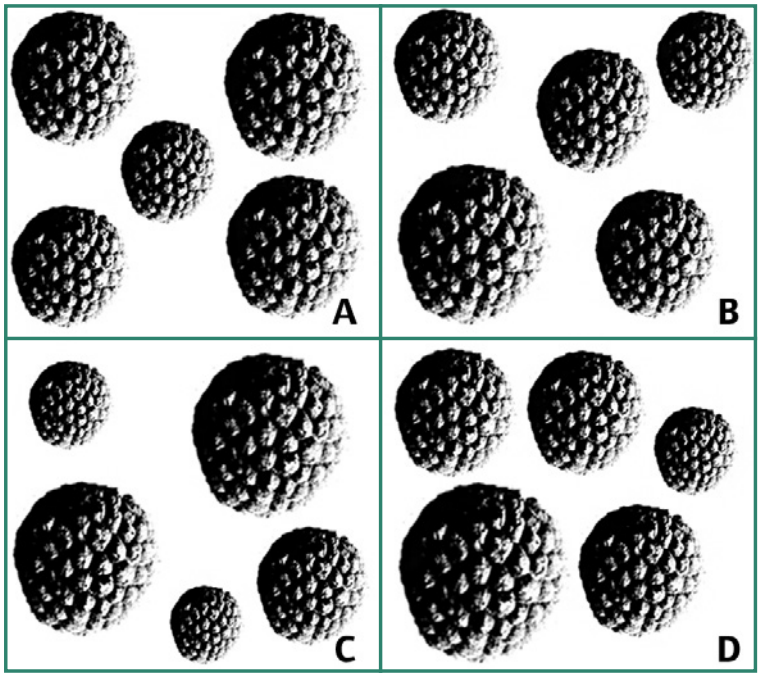
Painel 1.2.
Genótipos de treze indivíduos pertencentes às populações 01, 02, 03 e 04 para o lócus microssatélite Acr 12 da espécie *Annona crassiflora* Mart.

População 01		População 02		População 03		População 04	
Indivíduos	**Genótipos**	**Indivíduos**	**Genótipos**	**Indivíduos**	**Genótipos**	**Indivíduos**	**Genótipos**
Ind01		Ind01		Ind01		Ind01	
Ind02		Ind02		Ind02		Ind02	
Ind03		Ind03		Ind03		Ind03	
Ind04		Ind04		Ind04		Ind04	
Ind05		Ind05		Ind05		Ind05	
Ind06		Ind06		Ind06		Ind06	
Ind07		Ind07		Ind07		Ind07	
Ind08		Ind08		Ind08		Ind08	
Ind09		Ind09		Ind09		Ind09	
Ind10		Ind10		Ind10		Ind10	
Ind11		Ind11		Ind11		Ind11	
Ind12		Ind12		Ind12		Ind12	
Ind13		Ind13		Ind13		Ind13	

Índices	Populações			
	01	02	03	04
H_e				
H_o				
f				

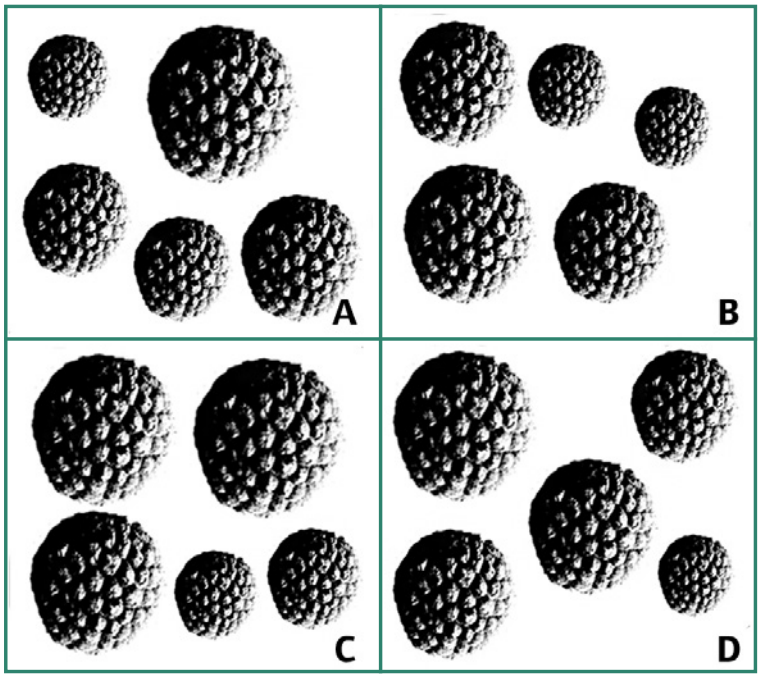
Painel 1.3.
Estimativas de diversidade genética (heterozigosidade esperada sob equilíbrio de Hardy-Weinberg: H_e), heterozigosidade observada (H_o) e endogamia (f) das população 01, 02, 03 e 04 referente ao lócus microssatélite Acr 12 da espécie *Annona crassiflora* Mart.

População 01

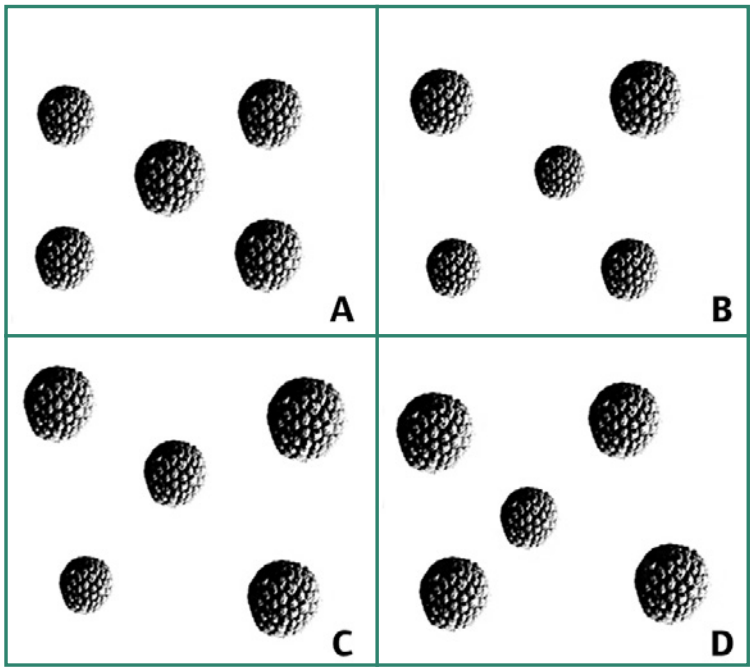


Painel 2.1.
Frutos coletados de quatro indivíduos (A, B, C e D) adultos de *Annona crassiflora*, provenientes de sementes coletadas nas populações naturais 01, 02, 03 e 04, e mantidos em condições controladas de solo, irrigação e adubação, em delineamento experimental inteiramente casualizado.

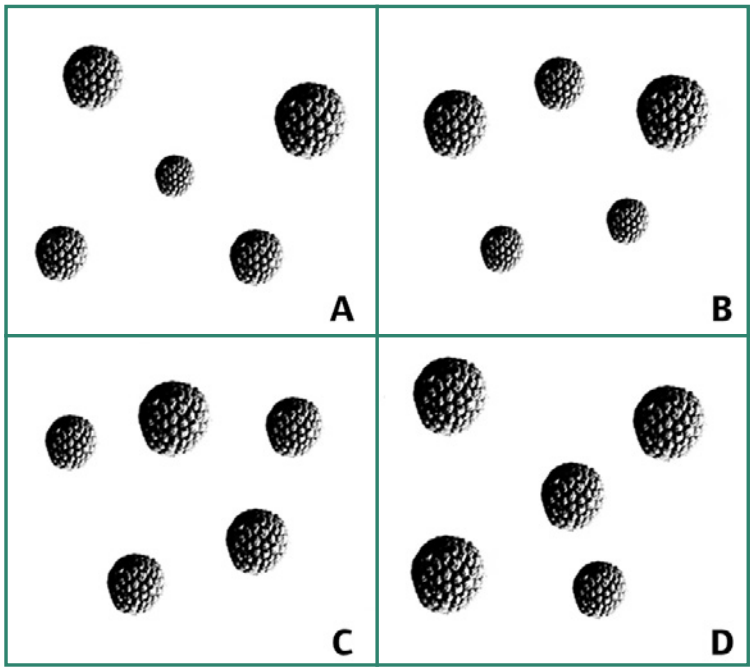
População 02



População 03

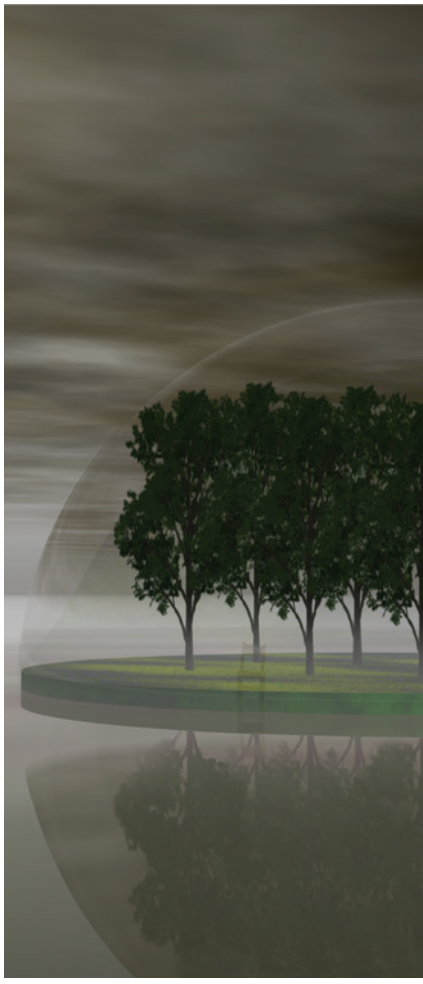


População 04



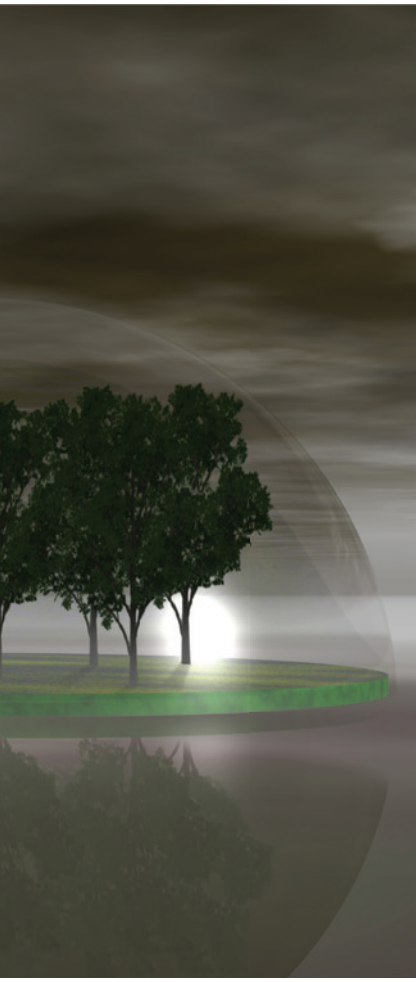
DM					
Indivíduos	**Frutos**	**População 01**	**População 02**	**População 03**	**População 04**
A	01				
	02				
	03				
	04				
	05				
B	01				
	02				
	03				
	04				
	05				
C	01				
	02				
	03				
	04				
	05				
D	01				
	02				
	03				
	04				
	05				

Painel 2.2.
Diâmetro maior (*DM*) de frutos de quatro indivíduos (A, B, C e D) adultos de *Annona crassiflora*, provenientes de sementes coletadas nas populações naturais 01, 02, 03 e 04 e mantidos em condições controladas de solo, irrigação e adubação, em delineamento experimental inteiramente casualizado.



Painel 3.
Estimativas dos valores do coeficiente de endogamia (*f*) e médias fenotípicas para diâmetro maior do fruto (*DM*) de indivíduos adultos de *Annona crassiflora*, provenientes de sementes coletadas em populações naturais e mantidas em condições controladas de solo, irrigação e adubação, em delineamento experimental inteiramente casualizado.

Índices	Populações			
	01	02	03	04
f				
Média de DM				
δ_{DM}				



ENTENDENDO A ATIVIDADE

Questão 1. Existe relação entre a endogamia e a redução da diversidade genética, que é representada pela heterozigosidade esperada pelo EHW (H_e)?

Questão 2. O que a relação entre os coeficientes de endogamia (*f*) das populações e as médias fenotípicas de diâmetro dos frutos permite concluir?

Questão 3. É possível inferir que as populações naturais de araticunzeiro – *A. crassiflora* – possuem depressão endogâmica?

Questão 4. Qual o papel dos marcadores moleculares microssatélites na avaliação da ocorrência de depressão endogâmica?

Questão 5. Qual a importância da avaliação da depressão endogâmica para as estratégias de manejo e conservação de populações ameaçadas?

Questão 6. Para qual população você aconselharia um plano de manejo devido ao risco de depressão endogâmica?

RESPOSTAS

a. Preenchimento do painel 1.2

Indivíduos	Genótipo	Indivíduos	Genótipo	Indivíduos	Genótipo	Indivíduos	Genótipo
Ind01	A1A1	Ind01	A1A2	Ind01	A1A1	Ind01	A1A1
Ind02	A1A2	Ind02	A1A3	Ind02	A2A3	Ind02	A2A3
Ind03	A3A3	Ind03	A2A3	Ind03	A1A1	Ind03	A1A1
Ind04	A2A2	Ind04	A1A2	Ind04	A2A2	Ind04	A2A2
Ind05	A1A3	Ind05	A1A3	Ind05	A1A3	Ind05	A3A3
Ind06	A2A3	Ind06	A2A3	Ind06	A2A3	Ind06	A3A3
Ind07	A3A3	Ind07	A2A2	Ind07	A2A2	Ind07	A2A2
Ind08	A1A2	Ind08	A1A1	Ind08	A1A1	Ind08	A1A2
Ind09	A2A3	Ind09	A2A3	Ind09	A3A3	Ind09	A3A3
Ind10	A1A3	Ind10	A1A2	Ind10	A1A2	Ind10	A2A2
Ind11	A1A2	Ind11	A2A2	Ind11	A2A2	Ind11	A1A1
Ind12	A1A3	Ind12	A1A3	Ind12	A1A1	Ind12	A1A3
Ind13	A2A3	Ind13	A3A3	Ind13	A3A3	Ind13	A3A3

Painel 1.2. Genótipos de treze indivíduos pertencentes às populações 01, 02, 03 e 04 baseados na amplificação via **PCR** do lócus microssatélite Acr 12 da espécie *Annona crassiflora* Mart.

PCR ou reação em cadeia da polimerase (em inglês *Polymerase Chain Reaction*) é um método de amplificação *in vitro* que permite criar milhões de cópias de um fragmento específico de DNA, sem o uso de um organismo vivo. A técnica permite determinar o genótipo de um indivíduo, para um dado lócus gênico selecionado.

b. Cálculo das frequências alélicas, da heterozigosidade esperada (H_e) sob EHW, heterozigosidade observada (H_o) e coeficientes de endogamia (f)

b.1. População 01:

- Cálculo das frequências alélicas

$n(A1) = 8$

$n(A2) = 8$

$n(A3) = 10$

$nt = 8 + 8 + 10 = 26$

$f(A1) = n(A1) / nt = 8 / 26 = 0,3$

$f(A2) = n(A2) / nt = 8 / 26 = 0,3$

$f(A3) = n(A3) / nt = 10 / 26 = 0,4$

- Cálculo da heterozigosidade esperada sob equilíbrio de Hardy-Weinberg (H_e):

$H_e = 2pq + 2pr + 2qr$

$H_e = 2(0,3 \times 0,3) + 2(0,3 \times 0,4) + 2(0,3 \times 0,4)$

$H_e = 0,660$

- Cálculo da heterozigosidade observada (H_o):

$H_o = f(AiAj) / f(AiAi)$

$H_o = 9/13$

$H_o = 0,692$

- Cálculo do coeficiente de endogamia (*f*):

$$f = \frac{He - Ho}{He} \rightarrow f = \frac{0{,}660 - 0{,}692}{0{,}660} \rightarrow f = -0{,}048$$

b.2. População 02:

- Cálculo das frequências alélicas

$n(A1) = 8$

$n(A2) = 10$

$n(A3) = 8$

$nt = 8 + 10 + 8 = 26$

$f(A1) = n(A1) / nt = 8 / 26 = 0{,}3$

$f(A2) = n(A2) / nt = 10 / 26 = 0{,}4$

$f(A3) = n(A3) / nt = 8 / 26 = 0{,}3$

- Cálculo da heterozigosidade esperada sob equilíbrio de Hardy-Weinberg (He):

$He = 2pq + 2pr + 2qr$

$He = 2(0{,}3 \times 0{,}4) + 2(0{,}4 \times 0{,}3) + 2(0{,}3 \times 0{,}3)$

$He = 0{,}660$

- Cálculo da heterozigosidade observada (Ho):

$Ho = f(AiAj) \,/\, f(AiAi)$

$Ho = 9/13$

$Ho = 0{,}692$

- Cálculo do coeficiente de endogamia (*f*):

$$f = \frac{He - Ho}{He} \rightarrow f = \frac{0{,}660 - 0{,}692}{0{,}660} \rightarrow f = -0{,}048$$

b.3. População 03:

- Cálculo das frequências alélicas

$n(A1) = 8$

$n(A2) = 10$

$n(A3) = 8$

$nt = 8 + 10 + 8 = 26$

$f(A1) = n(A1) / nt = 8 / 26 = 0{,}3$

$f(A2) = n(A2) / nt = 10 / 26 = 0{,}4$

$f(A3) = n(A3) / nt = 8 / 26 = 0{,}3$

- Cálculo da heterozigosidade esperada sob equilíbrio de Hardy-Weinberg (He):

$He = 2pq + 2pr + 2qr$

$He = 2(0{,}3 \times 0{,}4) + 2(0{,}4 \times 0{,}3) + 2(0{,}3 \times 0{,}3)$

$He = 0{,}660$

- Cálculo da heterozigosidade observada (Ho):

$Ho = f(AiAj) \,/\, f(AiAi)$

$Ho = 4/13$

$Ho = 0{,}307$

- Cálculo do coeficiente de endogamia (*f*):

$$f = \frac{He - Ho}{He} \rightarrow f = \frac{0{,}660 - 0{,}307}{0{,}660} \rightarrow f = -0{,}535$$

b.4. População 04:

- Cálculo das frequências alélicas

$n(A1) = 8$

$n(A2) = 10$

$n(A3) = 8$

$nt = 8 + 8 + 10 = 26$

$f(A1) = n(A1) / nt = 8 / 26 = 0{,}3$

$f(A2) = n(A2) / nt = 8 / 26 = 0{,}3$

$f(A3) = n(A3) / nt = 10 / 26 = 0{,}4$

- Cálculo da heterozigosidade esperada sob equilíbrio de Hardy-Weinberg (He):

$He = 2pq + 2pr + 2qr$

$He = 2(0{,}3 \times 0{,}3) + 2(0{,}3 \times 0{,}4) + 2(0{,}3 \times 0{,}4)$

$He = 0{,}660$

- Cálculo da heterozigosidade observada (Ho):

$Ho = f(AiAj) \,/\, f(AiAi)$

$Ho = 3/13$

$Ho = 0{,}231$

- Cálculo do coeficiente de endogamia (*f*):

$$f = \frac{He - Ho}{He} \rightarrow f = \frac{0{,}660 - 0{,}231}{0{,}660} \rightarrow f = -0{,}650$$

c. Preenchimento do painel 1.3

Índices	Populações			
	01	02	03	04
H_e	0,660	0,660	0,660	0,660
H_o	0,692	0,692	0,307	0,231
f	-0,048	-0,048	0,535	0,650

d. Preenchimento do Painel 2.2. Diâmetro dos frutos

DM					
Indivíduos	Frutos	População 01	População 02	População 03	População 04
A	01	16,8	1,8	1,3	1,4
	02	18	3,3	1,5	1,7
	03	12,6	2,4	1,6	0,9
	04	15	2,2	1,3	1,2
	05	18	2,6	1,5	1,1
B	01	14,4	2,6	1,4	1,4
	02	15,6	1,8	1,6	1,1
	03	12,6	1,7	1,1	1,6
	04	19,8	2,9	1,2	1
	05	16,2	2,5	1,3	0,9
C	01	10,8	3,1	1,6	1,2
	02	20,4	3,4	1,4	1,6
	03	19,2	3	1,8	1,3
	04	9,6	1,9	1,2	1,2
	05	14,4	2,1	1,7	1,4
D	01	16,2	3,1	1,7	1,7
	02	15	2,4	1,6	1,6
	03	10,8	2,9	1,4	1,4
	04	22,2	2,7	1,5	1,7
	05	16,2	1,7	1,7	1,2

e. Preenchimento do Painel 3

Índices	Populações			
	01	02	03	04
f	-0,048	-0,048	0,535	0,650
Média de DM	15,7 cm	15,0 cm	8,8 cm	8,0 cm
δ_{DM}	0,0	0,0	0,44	0,49

Questão 1.
A endogamia não é responsável pela redução da diversidade genética, medida por meio da heterozigosidade esperada sob equilíbrio de Hardy-Weinberg (*He*). O acasalamento entre indivíduos aparentados não leva à perda de alelos, mas à fixação deles em genótipos homozigotos. Assim, a diversidade genética é deslocada do componente "dentro de indivíduos", para o componente "entre indivíduos". A endogamia não altera as frequências alélicas, apenas as genotípicas, motivo pelo qual o parâmetro *He* não se modifica, mesmo nas populações endogâmicas.

Questão 2.
Populações endogâmicas (03 e 04) apresentaram valores médios de diâmetro de frutos menores que as populações não endogâmicas (01 e 02).

Questão 3.
Sim, das quatro populações de araticunzeiro avaliadas, duas (03 e 04) apresentaram depressão endogâmica. A população 04 apresentou maior redução de valor adaptativo, com perda de valor fenotípico, ou seja, perda do valor médio do diâmetro de fruto, de 49%.

Questão 4.
Os marcadores moleculares microssatélites apresentam herança mendeliana e por isso são adequados para estimar os índices de endogamia (*f*) em populações naturais. Com essa tecnologia é possível detectar a redução da heterozigosidade observada nas populações (*Ho*) que ocorre em decorrência do acasalamento entre indivíduos aparentados. Ao estimar os índices de endogamia com os marcadores microssatélites, é possível verificar se as populações com os maiores valores de *f* são as que apresentam as menores médias fenotípicas, em relação aos caracteres quantitativos adaptativos.

Questão 5.
A redução na aptidão pode comprometer seriamente a persistência das populações de uma determinada espécie. Portanto, a endogamia pode ameaçar a viabilidade populacional. Nesse sentido, é importante avaliar populações naturais em programas de manejo e conservação de espécies ameaçadas, de modo a detectar a ocorrência de depressão endogâmica e delinear estratégias que visem minimizá-la ao máximo.

Questão 6.
As duas populações endogâmicas (03 e 04) necessitam de manejo devido à depressão endogâmica.

REFERÊNCIAS

CAVALCANTE, T. R. M.; NAVES R. V.; FRANCESCHINELLI, E.V.; SILVA, R. D. Polinização e formação de frutos em araticum. *Bragantia*, v. 68, p. 13-21, 2009.

FRANKHAM, R.; BALLOU, J. D.; BRISCOE, D. A. *Introduction to Conservation Genetics*. Cambridge: Cambridge University Press, 2002, 617 p.

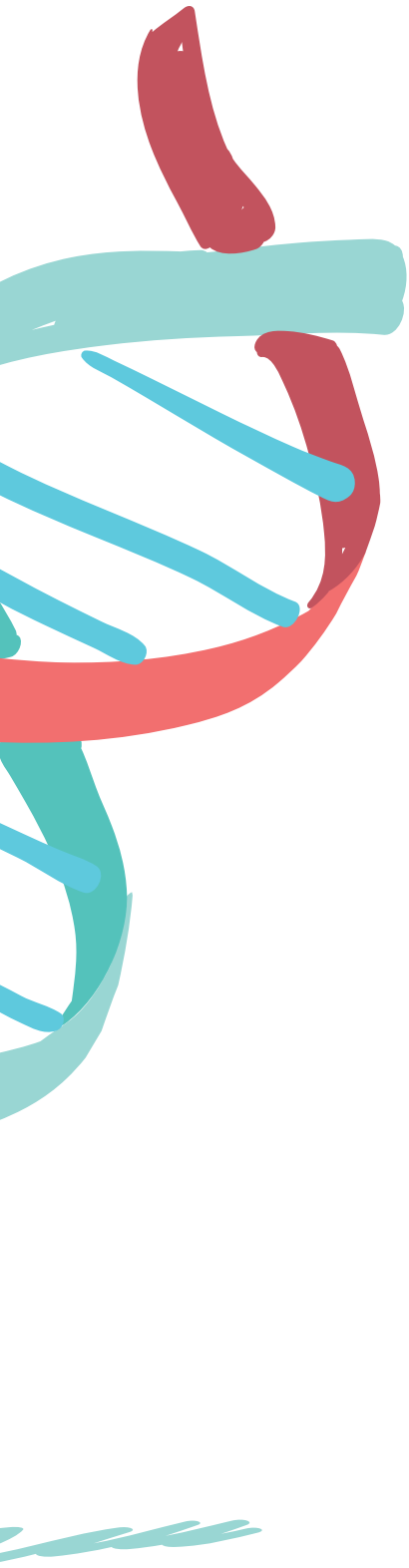
LANDE, R.; SCHEMSKE, D. W. The evolution of self fertilization and inbreeding depression in plants. I. Genetic models. *Evolution*, v. 39, p.24-4, 1985.

Tirinhas no ensino da estrutura, função e conceito de gene

Maria de Nazaré Klautau-Guimarães[1], Mariana Marzullo Pedreira[2], Silviene Fabiana de Oliveira[1]

[1] Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, campus Darcy Ribeiro, Universidade de Brasília.
[2] Mestranda no PPG Ciências da Saúde, Departamento de Genética e Morfologia, Instituto de Ciências Biológicas, campus Darcy Ribeiro, Universidade de Brasília.

Autor para correspondência: nklautau@unb.br ou nazaklautau@gmail.com



Essa proposta didática apresenta a interação entre a arte e a ciência na construção do conhecimento científico. Como arte, explora o potencial pedagógico das tirinhas que permite uma leitura rápida, dinâmica, crítica e reflexiva. Na ciência, aborda a crise do conceito molecular clássico do gene. Assim, espera-se que os estudantes sejam estimulados a organizar e comunicar as novidades sobre a estrutura e função do gene, de maneira criativa e para além do livro didático.

CONCEITO DE GENE

A introdução do conceito de gene, termo empregado pela primeira vez por Wilhelm L. Johannsen em 1909, foi um dos marcos da história da ciência no século XX. O conceito original de gene como unidade indivisível da herança mendeliana foi sendo modificado com os avanços do conhecimento do material genético, sendo que hoje o mais abordado no ensino é o conceito molecular: uma sequência de DNA responsável pela produção de uma molécula de RNA ou de um polipeptídio que desempenha uma função específica. Observa-se uma relação de 1:1:1 entre genes, produtos gênicos e função gênica, que perpassa uma noção de unidade. Este conceito está amplamente difundido tanto no nível escolar médio quanto no superior (EL-HANI, 2007).

Com o aprofundamento nos conhecimentos sobre estrutura e funcionamento dos genomas, o conceito de gene passou a ser contestado, instalando-se assim uma crise que sugere que o mesmo passe por uma revisão. Determinados autores afirmam que o que está por trás da crise do modelo é justamente o tratamento dos genes como unidades (EL-HANI, 2007; MEYER, 2010; MEYER *et al.*, 2013).

Apesar da discussão na área da Filosofia e História da Ciência, observa-se dificuldade na compreensão deste conceito por parte de todos os atores no processo de ensino-aprendizagem da genética. Nos livros didáticos do ensino médio e superior detecta-se uma proliferação de significados para o termo gene, o que pode levar à compreensão equivocada por parte dos estudantes (SANTOS; EL-HANI, 2009; PITOMBO *et al.* 2008). O conceito predominante para estudantes de graduação das áreas biológicas é o molecular clássico, com grande heterogeneidade dos conhecimentos prévios e dificuldades de organização dos conhecimentos adquiridos (MEYER, 2010; PEDREIRA *et. al.*, 2013). Nas principais revistas de divulgação científica observa-se a utilização de metáforas intimamente relacionadas ao determinismo genético (GOLDBACH; EL-HANI, 2008).

PROPOSTA PEDAGÓGICA

O ensino atual apresenta estratégias teóricas tradicionais que se apresentam inadequadas na construção dos conhecimentos científicos e como consequência, há grande perda de interesse por parte dos estudantes. Diante do quadro atual, faz-se necessário que os estudantes sejam estimulados a desenvolver sua curiosidade natural e o seu potencial criativo para muito além do livro didático. Nesse contexto, Araújo-Jorge (2007) apresenta relações entre ciência, arte e educação, onde lembra que "Grandes cientistas, como Galileu Galilei e Leonardo da Vinci, transitaram pelas vias de conexão entre a ciência e a arte, ao desenvolver o conhecimento e o comunicar das mais diferentes formas".

Caruso e Silveira (2009) argumentam que através da produção artística apresenta-se uma pedagogia que contempla articulações entre ensino-aprendizagem e conhecimento-sociedade integrando os conteúdos disciplinares. Estes autores apresentam novos horizontes das potencialidades das histórias em quadrinhos (HQs) de contribuir de forma diferenciada para a construção da cidadania, popularização da ciência e no ensino de ciência em todos os níveis. As tirinhas permitem fazer uma contextualização da aprendizagem do conceito que se pretende transmitir, como também, permitem uma leitura muito rápida e dinâmica o que estimula os estudantes a reler os conceitos de forma crítica e reflexiva.

Os estudos acadêmicos na área de Ensino de Ciências, a nível nacional e internacional, apresentam a utilização das HQs como recurso relevante para os diversos níveis de ensino. E todos apontam a importância do professor na compreensão crítica do enredo das histórias, na seleção do material e no planejamento das atividades (PIZARRO, 2009).

OBJETIVOS DA ATIVIDADE

O objetivo deste material é introduzir tirinhas como recurso didático no ensino de genética, buscando enquadrar os desafios, estimular a discussão e o pensamento crítico relacionado ao conceito, estrutura e função do gene. Pretende-se contribuir para desmistificação de conceitos errôneos e de senso comum entre os estudantes a respeito do conceito molecular clássico do gene, através da linguagem e imagem dos quadrinhos.

Este material tem o seu foco voltado para o aprimoramento do conceito de gene para o ensino superior e formação de professores, preparando-os para estar sempre abertos aos conhecimentos novos e suas aplicações na sociedade. Os estudantes e professores terão a oportunidade de construir o seu próprio conhecimento diante de um desafio, as tirinhas, que requer pesquisa e elaboração de pequenos textos.

CARACTERÍSTICAS ESPECÍFICAS DO MATERIAL

Com o intuito de conectar os estudantes ao novo material foi desenvolvido um personagem principal, que representa um estudante de graduação cursando a disciplina de Genética na universidade. O personagem foi inserido em três contextos comuns aos estudantes: 1 – estudar para uma prova, 2 – participar de palestras e 3 – bater papo em horas de diversão. A ideia central é que os estudantes, ao serem apresentados às tirinhas, identifiquem-se com o personagem, ficando confortáveis e estimulados a terem questionamentos ao longo da disciplina.

A tirinha **Estudando para a prova** (Fig. 1) busca dinamizar o questionamento do conceito de gene. Nesse contexto, o estudante é apresentado estudando em diferentes fontes e pensando em vários conceitos genéticos. O grande desafio seria em como fazer a relação completa entre o conhecimento atual na área de Biologia Molecular, buscando uma definição atual de gene que leve em consideração as recentes descobertas sobre a estrutura, a organização e a regulação do gene. A tirinha **Genética no bar** (Fig. 2) foi desenvolvida para priorizar o caráter argumentativo dos estudantes, enquanto questiona a relação "um gene: uma proteína" após os resultados do Projeto Genoma Humano. Essa atividade proporciona ao estudante a organização dos seus conhecimentos para depois explicá-los e justificá-los, mesmo que para uma plateia fora do meio acadêmico. A tirinha **Na palestra** (Fig. 3) retrata a participação de estudantes em eventos como seminários, palestras ou congressos e objetiva abordar a relação entre a estrutura física do gene e a sequência de aminoácidos de uma proteína, de uma maneira inversa ao que é tradicionalmente ensinado no Dogma Central da Biologia Molecular.

Figura 1.
Estudando para a prova.

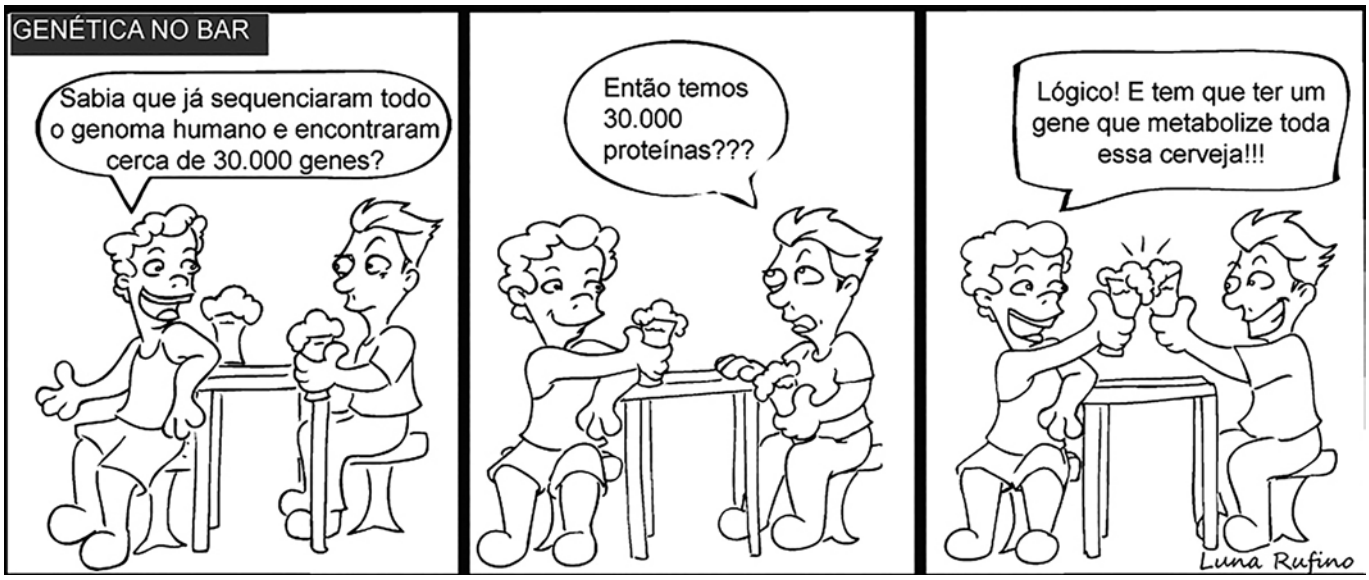


Figura 2.
Genética no bar.

PROCEDIMENTO DE APLICAÇÃO DO MATERIAL

Aqui são apresentadas três propostas de aplicação do material didático de acordo com os objetivos almejados. Isso não impede que outras formas de aplicação sejam sugeridas. Como o objetivo é de construção do conhecimento, o estudante não recebe um material pronto. Dessa forma, não se espera que existam respostas exatas e sim, que as mesmas sejam elaboradas de forma dinâmica com consultas, discussões, trocas de ideias entre os estudantes e a orientação do docente responsável.

(Proposta A) Avaliação dos conhecimentos prévios. Nessa proposta as tirinhas são apresentadas individualmente aos estudantes no início da disciplina e é solicitado que apresentem respostas sem consulta, utilizando apenas o conhecimento prévio ao início

Figura 3.
Na palestra.

da disciplina. Com esta avaliação, o docente poderá estruturar melhor as abordagens desse tema em sua disciplina. Após uma apresentação teórica do conteúdo, os estudantes podem receber suas respostas e rever criticamente as mesmas.

(Proposta B) Construção do conhecimento sobre a estrutura, função e conceito do gene, coerente com os conhecimentos científicos atuais. Para esta proposta é sugerido que as tirinhas sejam aplicadas aos estudantes, em duplas, promovendo a busca de informações em fontes adequadas. Para este objetivo sugerimos a utilização do texto elaborado por Meyer (2010), ou que seja estimulada a busca em outras fontes ou que se utilizem os textos de revisão citados.

(Proposta C) Estimular o estudante para a leitura, o estudo e a contextualização do conceito. Nesta proposta os estudantes são incentivados a construir tirinhas autorais a partir de suas indagações, pesquisas na mídia e discussão de artigos de revisão em sala de aula. Após a elaboração das tirinhas em duplas ou grupos, sugere-se a apresentação à turma, com enfoque no contexto e nas possíveis leituras que a tirinha permita.

AGRADECIMENTO

Agradecimento para Luna Rufino pela ilustração das tirinhas.

REFERÊNCIAS

ARAUJO-JORGE, T. C. Relações entre ciência, arte e educação: relevância e inovação. *Revista E*, São Paulo Sesc, p. 12, 2007.

CARUSO, F. e SILVEIRA, C. Quadrinhos para a cidadania. História, Ciência e Saúde – Manguinhos, Rio de Janeiro, v.16, n.1, p.217-236, 2009.

EL-HANI, C. N. Between the Cross and the Sword: The Crisis of the Gene Concept. *Genetics and Molecular Biology* (Impresso). v. 30, p. 297-307, 2007.

GOLDBACH, T.; EL-HANI, C. N. Entre Receitas, Programas e Códigos: Metáforas e Ideias Sobre Genes na Divulgação Científica e no Contexto Escolar. *Revista de Educação em Ciência e Tecnologia*. Vol 1, p. 153-189, 2008.

MEYER, L. M. N.; BOMFIM, G. C.; EL-HANI, C. N. How to Understand the Gene in the Twenty-First Century? *Science & Education* (Dordrecht). Vol 22, p. 345-374, 2013.

MEYER, L. M. N. *Como ensinar a estudantes universitários de Ciências Biológicas e Ciências da saúde sobre a crise do conceito de gene?* Universidade Federal da Bahia. Dissertação de Mestrado. 2010. Disponível em: < http://twiki.ufba.br/twiki/pub/PPGEFHC/DissertacoesPpgefhc/Lia_Midori_Meyer_Nascimento_2010.pdf >. Disponível em: 17/02/2014.

PEDREIRA, M. M.; RESENDE, T. A.; OLIVEIRA, S. F.; KLAUTAU-GUIMARÃES, M. N. Tirinhas no ensino de genética: potencial para avaliação dos conhecimentos prévios sobre os genes. *Enseñanza de las Ciencias*. Vol extra, p. 2695-2700, 2013.

PITOMBO, M. A.; ALMEIDA, A. M. R.; EL-HANI, C. N. Conceitos de gene e ideias sobre função gênica em livros didáticos de biologia celular e molecular do ensino superior. *Contexto & Educação*. Vol. 77, p. 81-110, 2008.

PIZARRO, M. V. As histórias em quadrinhos como linguagem e recurso didático no ensino de ciências. *VII Enpec*. Florianópolis. 2009.

SANTOS, V. C.; EL-HANI, C. N. Ideias sobre genes em livros didáticos de biologia do ensino médio publicados no Brasil. *Revista Brasileira de Pesquisa em Educação em Ciências*. Vol. 9, n.1, 2009.

A Genética contra os crimes ambientais: identificação de madeira ilegal provenientes de unidades de conservação utilizando marcador molecular*

Núbia Esther de Oliveira Miranda[1,2]; Edivaldo Barbosa de Almeida Júnior[1,3]; Rosane Garcia Collevatti[1,2]

[1] Laboratório de Genética & Biodiversidade, ICB, Universidade Federal de Goiás (UFG), Goiânia
[2] Programa de Pós-Graduação em Ecologia e Evolução, ICB, Universidade Federal de Goiás (UFG), Goiânia
[3] Programa de Pós-Graduação em Genética e Melhoramento de Plantas, Escola de Agronomia, Universidade Federal de Goiás (UFG), Goiânia.

Autor para correspondência: rosanegc68@hotmail.com.

*Material didático desenvolvido na disciplina de Genética da Conservação, coordenado pela Profa. Rosane Garcia Collevatti, do curso de graduação em Ecologia e Análise Ambiental do Departamento de Ecologia, Instituto de Ciências Biológicas, Universidade Federal de Goiás, como uma atividade do Estágio Docência (bolsistas CAPES/UFG) dos discentes dos Programas de Pós-graduação em Ecologia e Evolução, Genética e Melhoramento de Plantas.

A perda acelerada da cobertura vegetal coloca em risco os recursos naturais e a diversidade biológica, além de afetar a produtividade econômica. Atualmente, a despeito da fiscalização para coibir a prática, grande parte dos desmatamentos são feitos de forma ilegal. Para auxiliar o ensino de Genética da Conservação e Ecologia Molecular é proposta uma atividade que demonstra como o uso de marcador molecular, baseado no sequenciamento de DNA de cloroplasto, pode auxiliar na identificação de amostras apreendidas por órgãos ambientais para resolver e estabelecer provas em crimes ambientais envolvendo desflorestamentos. A atividade pode ser utilizada como simulação de aula prática ou como uma atividade complementar à aula expositiva para alunos de ensino superior.

O objetivo da atividade é mostrar como o uso de **marcadores moleculares**, baseados no **sequenciamento** de DNA de cloroplasto (cpDNA), pode ser empregado para determinar a origem da madeira extraída de florestas naturais a partir de amostras apreendidas pelos órgãos de fiscalização ambiental. Para essa finalidade, é preciso comparar as sequências de DNA de amostras apreendidas com as sequências de DNA de um banco de dados pré-existente contendo sequências de amostras de populações na área de ocorrência da espécie, especialmente nas unidades de conservação e nos locais destinados ao extrativismo, com a finalidade de determinar a origem da madeira e se a madeira foi removida de localidades em que o corte é ilegal.

A metodologia inclui a coleta de material apreendido, o sequenciamento e a comparação do mesmo com um banco de dados de sequências visando gerar provas sobre remoção ilegal de madeira, de modo que o relatório possa ser utilizado em processos de crime ambiental. É importante frisar que o pressuposto fundamental do uso desta técnica para identificar o local de origem é a existência de diferenças genéticas detectáveis entre as áreas. Desta forma, para que o local de origem seja determinado, é necessário ter um banco de dados de sequências de DNA abrangendo diversas áreas, tanto Unidades de Conservação quanto áreas de extrativismo. Além disso, muitas vezes somente um marcador molecular não é suficiente para diferenciar amostras de diferentes locais. Dessa forma, pode ser necessário, para determinar o local de origem definitivo, o sequenciamento de mais de um marcador molecular, embora a atividade, aqui apresentada, utilize somente um marcador molecular.

PROBLEMA PROPOSTO

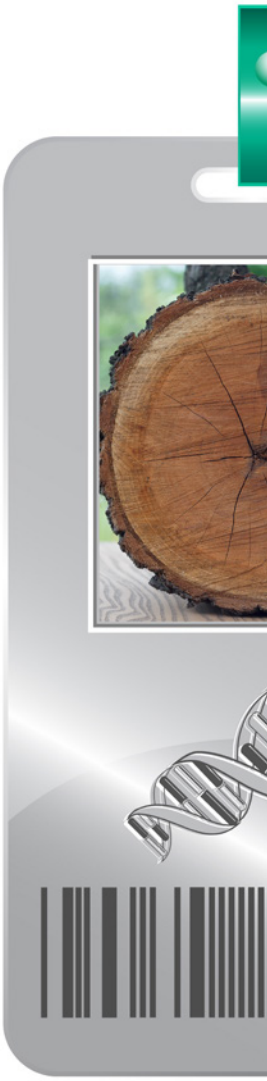
O Brasil possuiu uma das maiores áreas de cobertura vegetal e de diversidade florística do planeta. Além da grande importância dessas áreas para a conservação, elas também constituem uma importante fonte de desenvolvimento econômico para o País, o que acarreta a substituição de áreas nativas por sistemas produtivos como pastagens, lavouras, regiões para extrativismo, reflorestamento e mineração. Essa perda massificada e acelerada da cobertura vegetal, além de afetar a produtividade econômica, também coloca em risco os recursos naturais e a diversidade biológica. Nesse sentido, o Código Florestal, bem como a ação fiscalizatória das agências reguladoras ambientais (IBAMA, ICMBio, órgãos estaduais e municipais) são ferramentas utilizadas para controlar o uso das áreas destinadas ao uso antrópico e proteger os remanescentes florestais. A despeito disso, a taxa de desmatamento no Brasil continua aumentando e boa parte dessa prática está baseada na ilegalidade.

A utilização de técnicas genéticas em investigações forenses cada vez mais vem sendo utilizadas como fontes de evidências em investigações criminais. Na resolução e no estabelecimento de provas em crimes ambientais, esta ainda é uma ferramenta pouco utilizada no Brasil, mas que tem apresentado resultados positivos tanto na identificação de espécies quanto no rastreamento da origem de produtos obtidos por extrativismo ou caça (GRAVENA *et al.* 2008; SANCHES *et al.* 2011). Esses resultados positivos, aliados às técnicas que permitem a obtenção de material genético a partir de madeira já processada, abrem um importante campo de investigação para auxiliar o combate de crimes de desmatamentos florestais (DEGUILLOUX *et al.* 2003).

A aroeira, *Myracrodruon urundeuva* (sin. *Astronium urundeuva*) pertence à família Anacardiaceae e possui outros nomes comuns como aroeira-vermelha, aroeira-mansa, aroeira-branca, aroeira-da-praia, aroeira-do-sertão, aroeira-do-paraná. Outras espécies desta família que também são conhecidas por aroeira são: *Schinus molle*, *Astronium fraxinifolium* e *Lithraea brasiliensis*. De grande plasticidade ecológica, a aroeira ocorre nas Florestas Estacionais Semideciduais, desde o Nordeste, passando pelo Centro-Oeste do Brasil, chegando ao Rio Grande do Sul e estendendo-se à Argentina e ao Paraguai (LORENZI, 1992). Devido ao extrativismo da madeira, *M. urundeuva* encontra-se atualmente ameaçada de extinção (MMA, 2008), sendo seu corte res-

Marcador Molecular é qualquer dado molecular herdável capaz de evidenciar polimorfismo entre indivíduos. Pode ser um fragmento de DNA, codificante ou não, ou uma proteína.

Sequenciamento pode ser definido como a série de métodos empregados para determinar a ordem específica das bases nitrogenadas adenina (A), guanina (G), citosina (C) e timina (T) em uma molécula de DNA (veja o artigo Sequenciamento de DNA: decifrando o manual de instruções dos seres vivos, Genética na Escola 03.03, 45-52, 2009).



tringido e regulamentado pela Portaria N.º 83-N de 26 de Setembro de 1991 do Instituto Brasileiro do Meio Ambiente e dos Recursos Naturais Renováveis – Ibama:

> "Art. 1º. Fica proibido o corte e a exploração da Aroeira legítima ou Aroeira do sertão (*Astronium urundeuva*), das Baraúnas (*Melanoxylon brauna* e *Schinopsis brasiliensis*), do Gonçalo Alves (*Astronium fraxinifolium*) em Floresta Primária.
>
> Art. 2º. A exploração da Aroeira ou Aroeira do Sertão (*Astronium urundeuva*) das Baraúnas (*Melanoxylon brauna* e *Schinopsis brasiliensis*), do Gonçalo Alves (*Astronium fraxinifolium*) em Floresta Secundária, só poderá ser efetivada através de Plano de Manejo Florestal de Rendimento Sustentado, dependendo de projeto previamente aprovado pelo Ibama."

Nesta atividade foi proposta a determinação da origem de madeiras cortadas de aroeira com a utilização de um marcador molecular cloroplastidial (cpDNA) como evidência para uma investigação criminal de exploração madeireira em áreas ilegais.

O Instituto Chico Mendes (ICMBio) recebeu uma denúncia anônima de que quatro lojas moveleiras, de uma cidade localizada no sudeste do estado de Goiás, estariam comercializando produtos provenientes de madeiras de uso não certificado, i.e., oriundas de áreas onde a extração é proibida. Nas quatro lojas denunciadas, no entanto, os proprietários apresentaram documentos certificando que a origem da madeira era de áreas de extrativismo certificado, **Áreas A1** e **A2**. Para avaliar a procedência do material, os fiscais do ICMBio coletaram um lote de amostras de cada loja. Este lote era composto por câmbio da madeira de cinco pranchas por loja, totalizando 20 amostras. Os fragmentos de tecidos do câmbio de cada amostra foram acondicionados em saquinhos esterilizados individuais e conduzidos a um laboratório de análise genética. O DNA das amostras de câmbio foi extraído e a região cloroplastidial trnS-trnG foi sequenciada por pesquisadores deste laboratório. As sequências obtidas foram comparadas com bancos de dados do laboratório, contendo informações da composição genética das duas áreas de extrativismo certificado (**Áreas A1** e **A2**) e de duas populações de aroeira de unidades de conservação onde a extração da espécie é proibida por lei (**Áreas B1** e **B2**).

Para determinar a origem da madeira é preciso:

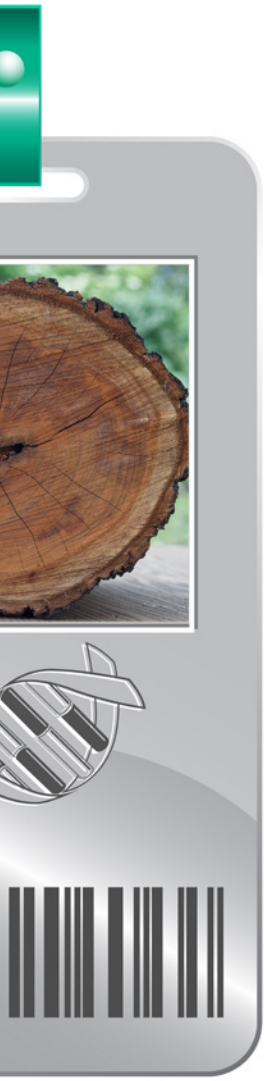
1. Observar as sequências de DNA do banco de dados das áreas de plantio (A1 e A2) e das unidades de conservação (B1 e B2).
2. Determinar as sequências de DNA de cada amostra dos lotes apreendidos nas quatro lojas moveleiras.
3. Comparar as sequências de DNA de cada amostra com as do banco de dados das sequências de DNA de cada área (A1, A2, B1 e B2).
4. A partir da comparação das sequências de DNA, determinar se a madeira é oriunda das áreas de extrativismo A1 ou A2 e, portanto, é realmente certificada, ou se a madeira é oriunda das unidades de conservação, portanto de origem ilegal.

INSTRUÇÕES PARA O PROFESSOR

1. Esta atividade poderá ser realizada individualmente, ou em grupos de alunos de, no máximo, quatro pessoas.
2. Cada grupo deverá receber o problema proposto, uma cópia do procedimento para realizar a atividade, uma cópia de cada painel, de cada tabela e das questões para serem discutidas.
3. É recomendável que o professor aplique esta atividade em turmas que já tiveram contato prévio com os conceitos de genética mendeliana e de populações.

Os recursos didáticos que deverão ser utilizados consistem em três painéis, conforme descrito a seguir:

1. **Painel 1** – Banco de dados contendo as sequências de DNA para o marcador molecular cloroplastidial trnS-trnG das áreas de extrativismo certificado (A1 e A2) e de unidades de conservação (B1 e B2) no estado de Goiás.



Área A1:

```
TCCTAATTGAAAAAAAAGACAATTACTATGTTACATTACAC
TCCTAATTGAAAAAAAAGACAATTACTATGTTACATTACAC
TCCTAACTGAAAAAAAAGACAAATACTATGTCACATTACAC
TCCTAATTGAAAAAAAACACAATTACTATGTTACATTACAC
TCCTAACCGAAAAAAAAGACAAATACTATGTCACATTACAC
TCCTAATTGAAAAAAAACACAATTACTATGTTACATTACAC
```

Área A2:

```
TCCTAATTGAAAAAAAAGACAATTACTATGTTACATTACAC
TCCTAATTGAAAAAAAACACAATTACTATGTTACATTACAC
TCCTAACCGAAAAAAAAGACAATTACTATGTTACATTACAC
TCCTAACTGAAAAAAAAGACAAATACTATGTCACATTACAC
TCCTAATTGAAAAAAAACACAATTACTATGTTACATTACAC
TCCTAACCGAAAAAAAAGACAAATACTATGTCACATTACAC
```

Área B1:

```
TCCTAATTGAGAAAAAAGACAATTACTATGTTACATTACAC
TCCTAATTGAGAAAAACACAATTACTATGTTACATTACAC
TCCTAACCGAAAAAAAAGACAATTACTACCCTACATTACAC
TCCTAACTGAAAAAAAAGACAAATCCTATGTCACATTACAC
TCCTAACCGAAAAAAAACCCAATTACTATGTTACATTACAC
TCCTAACCGAAAAAAAAGACAAATACTATGTCACATTACAC
```

Área B2:

```
TCCTAATTGAGAAAAAAGACAATTACTATGTTACATTACAC
TCCTAATTGAGAAAAACACAATTACTATGTTACATTACAC
TCCTAACCGAAAAAAAAGACAATTACTACCCTACATTACAC
TCCTAACTGAAAAAAAAGACAAATCCTATGTCACATTACAC
TCCTAACCGAAAAAAAACACAATTACTATGTTACATTACAC
TCCTAACCGAAAAAAAAGACAAATACTATGTCACATTACAC
```

Painel 1.
Banco de dados contendo as sequências de DNA para o marcador molecular cloroplastidial trnS-trnG das áreas de extrativismo certificado (A1 e A2) e das unidades de conservação (B1 e B2) no estado de Goiás.

2. **Painel 2** – Localização geográfica das áreas de extrativismo certificado (A1 e A2) e das unidades de conservação (B1 e B2) no estado de Goiás. Os pontos indicam os locais de ocorrência das populações naturais de aroeira e, os quadrados, as áreas para as quais o laboratório de análise genética possui banco de dados.
3. **Tabela 1** – Sequências de DNA de aroeira encontradas nas áreas de extrativismo certificado (A1 e A2) e nas unidades de conservação (B1 e B2).
4. **Painel 3** – Sequências de DNA das amostras dos lotes apreendidos nas quatro lojas moveleiras.
5. **Tabela 2** – Sequências de DNA de aroeira encontradas nos lotes de madeira amostrados nas quatro lojas moveleiras e origem das amostras.

Painel 2.
Localização geográfica das áreas de extrativismo certificado (A1 e A2) e das unidades de conservação (B1 e B2) no estado de Goiás. Os pontos indicam os locais de ocorrência das populações naturais de aroeira e, os quadrados, as áreas para as quais o laboratório de análise genética possui banco de dados.

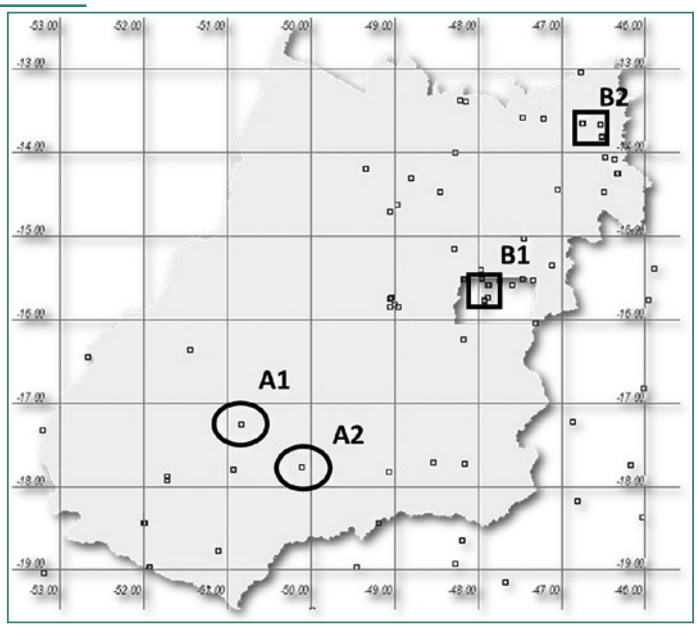


Tabela 1.
Sequências de DNA de aroeira encontradas nas áreas de extrativismo certificado (A1 e A2) e das unidades de conservação (B1 e B2).

Área	Sequências de DNA
A1	
A2	
B1	
B2	

Painel 3.
Sequências de DNA das amostras dos lotes apreendidos nas quatro lojas moveleiras.

Lote 1:

AMOSTRA 1:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC
AMOSTRA 2:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC
AMOSTRA 3:	TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC
AMOSTRA 4:	TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC
AMOSTRA 5:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC

Lote 2:

AMOSTRA 1:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC
AMOSTRA 2:	TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC
AMOSTRA 3:	TCCTAATTGAAAAAAACACAATTACTATGTTACATTACAC
AMOSTRA 4:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC
AMOSTRA 5:	TCCTAACCGAAAAAAAGACAATTACTATGTTACATTACAC

Lote 3:

AMOSTRA 1:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC
AMOSTRA 2:	TCCTAATTGAGAAAAACACAATTACTATGTTACATTACAC
AMOSTRA 3:	TCCTAACCGAAAAAAAGACAATTACTACCCTACATTACAC
AMOSTRA 4:	TCCTAACTGAAAAAAAGACAAATCCTATGTCACATTACAC
AMOSTRA 5:	TCCTAACCGAAAAAAACCCAATTACTATGTTACATTACAC

Lote 4:

AMOSTRA 1:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC
AMOSTRA 2:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC
AMOSTRA 3:	TCCTAACCGAAAAAAACCCAATTACTACCCTACATTACAC
AMOSTRA 4:	TCCTAACCGAAAAAAAGACAAATACTATGTTACATTACAC
AMOSTRA 5:	TCCTAACCGAAAAAAAGACAAATACTACCCTACATTACAC

Lote	Amostras	Sequências de DNA	Área
1	1		
	2		
	3		
	4		
	5		
2	1		
	2		
	3		
	4		
	5		
3	1		
	2		
	3		
	4		
	5		
4	1		
	2		
	3		
	4		
	5		

Tabela 2.
Sequências de DNA de aroeira encontradas nos lotes de madeira amostrados nas quatro lojas moveleiras e origem das amostras.

PROCEDIMENTO PARA OS ESTUDANTES

1. Ler atentamente o problema proposto.
2. Analisar o Painel 1 e Painel 2 – Sequências de DNA cloroplastidial do marcador molecular trnS-trnG para amostras das áreas de extrativismo certificado, A1 e A2, e para as unidades de conservação, B1 e B2, e localização geográfica das áreas:

a) analisar as sequências e determinar quais são presentes em cada área;

b) preencher a Tabela 1 com os resultados, atribuindo nomes diferentes para as sequências diferentes, por exemplo,

S1, S2, S3 etc. As sequências estão alinhadas de forma que cada base nitrogenada em uma posição no DNA em um indivíduo corresponde à mesma posição no indivíduo abaixo;

c) para analisar as sequências de DNA e determinar se são iguais ou diferentes, deve-se comparar cada base nitrogenada (total de 40) e verificar se são iguais ou diferentes;

d) lembrar-se de que sequências diferentes devem possuir nomes diferentes.

3. Analisar o Painel 3 - A partir das sequências de DNA de cada uma das amostras dos quatro lotes apreendidos nas lojas moveleiras:

a) determinar as sequências de DNA de cada amostra;

b) comparar as sequências de DNA das amostras com as do banco de dados das diferentes áreas (A1, A2, B1 e B2), conforme descrito no procedimento 2;

c) atribuir nomes às sequências de DNA das amostras da seguinte forma: se a sequência da amostra for igual à sequência de alguma das áreas do banco de dados, dar o nome igual às sequências de DNA do banco de dados; se as sequências de DNA forem diferentes, atribuir um novo nome; sequências iguais sempre devem ter o mesmo nome;

d) preencher a Tabela 2 com os resultados das sequências de DNA de cada amostra, para cada lote.

4. Analisar os resultados da Tabela 2:

a) determinar quais amostras são realmente oriundas de áreas de extrativismo (Áreas 1 e 2) e quais não são oriundas;

b) preencher os resultados na Tabela 2.

ENTENDENDO A ATIVIDADE

Questão 1: Em relação ao banco de dados das áreas A1, A2, B1 e B2. Quantas e quais sequências de DNA são exclusivas de cada área? Quantas e quais são compartilhadas? Quais áreas são mais semelhantes entre si? Discutir sobre a semelhança entre as áreas.

Questão 2: Foi possível encontrar a origem de todas as madeiras dos quatro lotes de aroeira? O banco de dados utilizado só possuía sequências de DNA de duas áreas de extrativismo e duas unidades de conservação? Este fato causou problemas para a atribuição dos locais de origem das amostras? Comentar sobre os problemas envolvidos na atribuição dos locais de origem e apontar soluções para eles.

Questão 3: Como a genética contribuiu para a elucidação do crime ambiental?

Questão 4: Quais as principais dificuldades na utilização deste método baseado em marcador molecular?

RESPOSTAS

Conjunto de sequências de DNA da Área A1:

```
TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC  [1]
TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC  [1]
TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC  [2]
TCCTAATTGAAAAAAACACAATTACTATGTTACATTACAC  [3]
TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC  [4]
TCCTAATTGAAAAAAACACAATTACTATGTTACATTACAC  [3]
Sequências (S1, S2, S3 e S4)
```

Conjunto de sequências de DNA da Área A2:

```
TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC  [1]
TCCTAATTGAAAAAAACACAATTACTATGTTACATTACAC  [3]
TCCTAACCGAAAAAAAGACAATTACTATGTTACATTACAC  [5]
TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC  [2]
TCCTAATTGAAAAAAACACAATTACTATGTTACATTACAC  [3]
TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC  [4]
Sequências (S1, S2, S3, S4 e S5)
```

Conjunto de sequências de DNA da Área B1:

```
TCCTAATTGAGAAAAAGACAATTACTATGTTACATTACAC  [6]
TCCTAATTGAGAAAAACACAATTACTATGTTACATTACAC  [7]
TCCTAACCGAAAAAAAGACAATTACTACCCTACATTACAC  [8]
TCCTAACTGAAAAAAAGACAAATCCTATGTCACATTACAC  [9]
TCCTAACCGAAAAAAACCCAATTACTATGTTACATTACAC  [10]
TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC  [4]
Sequências (S4, S6, S7, S8, S9 e S10)
```

Conjunto de sequências de DNA da Área B2:

```
TCCTAATTGAGAAAAAGACAATTACTATGTTACATTACAC  [6]
TCCTAATTGAGAAAAACACAATTACTATGTTACATTACAC  [7]
TCCTAACCGAAAAAAAGACAATTACTACCCTACATTACAC  [8]
TCCTAACTGAAAAAAAGACAAATCCTATGTCACATTACAC  [9]
TCCTAACCGAAAAAAACACAATTACTATGTTACATTACAC  [11]
TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC  [4]
Sequências (S4, S6, S7, S8, S9 e S11)
```

Painel 1.
Banco de dados contendo as sequências de DNA para o marcador molecular cloroplastidial trnS-trnG das áreas de extrativismo certificado (A1 e A2) e das unidades de conservação (B1 e B2) no estado de Goiás.

Área	Sequências de DNA
A1	S1, S2, S3 e S4
A2	S1, S2, S3, S4 e S5
B1	S4, S6, S7, S8, S9 e S10
B2	S4, S6, S7, S8, S9 e S11

Tabela 1.
Sequências de DNA de aroeira encontradas nas áreas de extrativismo certificado (A1 e A2) e das unidades de conservação (B1 e B2).

Painel 3.
Sequências de DNA das amostras dos lotes apreendidos nas quatro lojas moveleiras.

Sequências de DNA do Lote 1:

AMOSTRA 1:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC	[1]
AMOSTRA 2:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC	[1]
AMOSTRA 3:	TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC	[2]
AMOSTRA 4:	TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC	[2]
AMOSTRA 5:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC	[4]

Sequências (S1, S2, e S4)

Sequências de DNA do Lote 2:

AMOSTRA 1:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC	[1]
AMOSTRA 2:	TCCTAACTGAAAAAAAGACAAATACTATGTCACATTACAC	[2]
AMOSTRA 3:	TCCTAATTGAAAAAAACACAATTACTATGTTACATTACAC	[3]
AMOSTRA 4:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC	[4]
AMOSTRA 5:	TCCTAACCGAAAAAAAGACAATTACTATGTTACATTACAC	[5]

Sequências (S1, S2, S3, S4 e S5)

Sequências de DNA do Lote 3:

AMOSTRA 1:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC	[4]
AMOSTRA 2:	TCCTAATTGAGAAAAACACAATTACTATGTTACATTACAC	[7]
AMOSTRA 3:	TCCTAACCGAAAAAAAGACAATTACTACCCTACATTACAC	[8]
AMOSTRA 4:	TCCTAACTGAAAAAAAGACAAATCCTATGTCACATTACAC	[9]
AMOSTRA 5:	TCCTAACCGAAAAAAACCCAATTACTATGTTACATTACAC	[10]

Sequências (S4, S7, S8, S9 e S11)

Sequências de DNA do Lote 4:

AMOSTRA 1:	TCCTAATTGAAAAAAAGACAATTACTATGTTACATTACAC	[1]
AMOSTRA 2:	TCCTAACCGAAAAAAAGACAAATACTATGTCACATTACAC	[4]
AMOSTRA 3:	TCCTAACCGAAAAAAACCCAATTACTACCCTACATTACAC	[12]
AMOSTRA 4:	TCCTAACCGAAAAAAAGACAAATACTATGTTACATTACAC	[13]
AMOSTRA 5:	TCCTAACCGAAAAAAAGACAAATACTACCCTACATTACAC	[14]

Sequências (S1, S4, S12, S13 e S14)

Questão 1: Em relação ao banco de dados das áreas A1, A2, B1 e B2, quantas e quais sequências de DNA são exclusivos de cada área? Quantas e quais são compartilhadas? Quais áreas são mais semelhantes entre si? Discutir a semelhança entre as áreas.

A Área A1 não apresenta nenhuma sequência exclusiva. A Área A2 apresenta todas as quatro sequências presentes em A1 mais a sequência S5, que é exclusiva. A Área B1 tem uma sequência exclusiva (S10), compartilha a sequência S4 com todas as outras áreas e as sequências S6, S7, S8 e S9 com a área B2. Esta última tem uma sequência exclusiva (S11). As áreas A1 e A2 são mais semelhantes entre si, bem como as Áreas B1 e B2. Provavelmente, a maior semelhança entre as áreas A1 e A2 e B1 e B2 se deve à proximidade entre as áreas.

Lote	Amostras	Sequências de DNA	Área
1	1	S1	A1 ou A2
	2	S1	A1 ou A2
	3	S2	A1 ou A2
	4	S2	A1 ou A2
	5	S4	A1 ou A2 ou B1 ou B2
2	1	S1	A1 ou A2
	2	S2	A1 ou A2
	3	S3	A1 ou A2
	4	S4	A1 ou A2
	5	S5	A2
3	1	S4	A1 ou A2 ou B1 ou B2
	2	S7	B1 ou B2
	3	S8	B1 ou B2
	4	S9	B1 ou B2
	5	S10	B1
4	1	S1	A1 ou A2
	2	S4	A1 ou A2 ou B1 ou B2
	3	S12	?
	4	S13	?
	5	S14	?

Tabela 2.
Sequências de DNA de aroeira encontradas nos lotes de madeira amostrados nas quatro lojas moveleiras e origem das amostras.

Questão 2: Foi possível encontrar a origem de todas as madeiras dos quatro lotes de aroeira? O banco de dados utilizado só possuía sequências de DNA de duas áreas de extrativismo e duas unidades de conservação. Este fato causou problemas para a atribuição dos locais de origem das amostras? Comentar os problemas envolvidos na atribuição dos locais de origem e apontar soluções para eles.

O rastreamento da origem do material baseia-se nas variações observadas em cada uma das áreas e nas diferenças entre as sequências de DNA de cada área. Seguem as análises:

Lote 1: *Não conseguimos determinar a área exata das amostras, mas conseguimos determinar que 4 amostras (1 a 4) são oriundas de áreas de extrativismo. Já a origem da amostra 5 não pode ser determinada*

porque possui as sequências de DNA S4, compartilhada por todas as áreas.

Lote 2: *Também conseguimos determinar que 4 amostras são de áreas de extrativismo, mas não exatamente qual é a área. Porém, para a amostra 5, conseguimos determinar que é oriunda da Área 2 porque esta amostra possui a sequência S5, exclusiva desta Área.*

Lote 3: *Não conseguimos determinar a origem da amostra 1 porque esta amostra possui as sequências de DNA S4, compartilhada por todos as áreas. Já para as amostras 2, 3, 4, conseguimos determinar apenas que elas são oriundas de unidades de conservação (B1 ou B2), porque estas amostras têm sequências de DNA compartilhadas entre as áreas. Para a amostra 5, determinamos que é oriunda da Área B1. Desta forma, por comercializar madeiras de plantios, a loja moveleira correspondente ao lote 3 pode ser autuada por crime ambiental.*

Lote 4: *A amostra 1 é proveniente de área de extrativismo (A1 ou A2), mas a amostra 2 pode ser oriunda de qualquer uma das quatro áreas. Entretanto, as amostras 3, 4 e 5 não puderam ser rastreadas, ou seja, não podem ter origem determinada. Só sabemos que se trata de material ilegal porque podemos excluir as Áreas A1 e A2 como fontes de origem. As sequências de DNA do lote 4 são muito diferentes das presentes nestas duas áreas. Apesar disso, não conseguimos rastrear sua origem porque o material apreendido apresenta sequências que não estão presentes nos bancos de dados e este problema só poderia ser contornado se contássemos com um banco de dados maior, que englobasse grande parte das populações naturais da aroeira, ao longo de toda sua distribuição.*

Questão 3: Como a genética contribuiu para a elucidação do crime ambiental?

As análises genéticas, além de permitirem a identificação de espécies, também permitiram rastrear a origem do material com base na comparação de sequências obtidas a partir da análise de uma região do DNA. Isto só foi possível porque o marcador molecular utilizado foi capaz de evidenciar o polimorfismo entre indivíduos (amostras) e apresentar variação ao longo da distribuição da espécie investigada. Assim, fomos capazes de rastrear essa variabilidade e identificar sua provável origem. Desta forma, a análise genética é fundamental em problemas onde a identificação por características morfológicas não pode contribuir, por exemplo, para a identificação de origem.

Questão 4: Quais as principais dificuldades na utilização deste método baseado em marcador molecular?

A principal dificuldade é estabelecer bancos de dados robustos e abrangentes, ou seja, com muitos marcadores moleculares e com sequências de indivíduos de muitas áreas, para a maioria das espécies que se encontram ameaçadas por atividades exploratórias ilegais. Isso envolveria a capacitação de profissionais para lidar com a geração de dados genéticos e a designação de verbas específicas para a implementação desse banco de dados. A partir disso, a utilização do método seria facilitada e poderia ser realizada rapidamente, fornecendo indícios robustos para solucionar casos de crimes ambientais.

REFERÊNCIAS

DEGUILLOUX, M.F; PEMONGE, M. H; PETIT, R. J. DNA-based control of oak wood geographic origin in the context of the cooperage industry. *Ann. For. Sci.*, v. 61, p. 97-104, 2004.

GRAVENA, W.; HRBEK, T.; SILVA, V. M. F.; FARIAS, I. P. Amazon river dolphin love fetishes: from folklore to molecular forensics. *Mamm. Sci.*, v. 24, p. 969-978, 2008.

LORENZI, H. Árvores brasileiras: manual de identificação e cultivo de plantas arbóreas nativas do Brasil. São Paulo: *Plantarum*, v.1, p. 368, 1992.

MMA. Instrução Normativa MMA nº 6, de 23 de setembro de 2008, que reconhece as espécies da flora brasileira ameaçadas de extinção. Disponível em: http://www.ibama.gov.br/documentos/lista-de-especies-ameacadas-de-extincao.

SANCHES, A.; PEREZ WAM; FIGUEIREDO, M. G.; ROSSINI, B. C.; CERVINI, M.; GALETTI JR, P.M.; GALETTI, M. Wildlife forensic DNA and lowland tapir (Tapirus terrestres) poaching. *Conservation Genet. Resour*, v. 3, p. 189-193, 2011.

Como desvendar enigmas genéticos a partir da comparação de sequências*

Adriana Maria Antunes[1], Lays Karolina Soares da Cruz[1], Mariana Pires de Campos Telles[2]

[1] Bolsista CAPES no Programa de Pós-Graduação em Genética e Biologia Molecular, ICB, Universidade Federal de Goiás, Goiânia, Campus II, ICB I.
[2] Bolsista PQ1D-CNPq, Laboratório de Genética & Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Campus II, ICB I.

Autor para correspondência: adrianaantunesbio@gmail.com

*Material didático desenvolvido como parte da atividade de Estágio de Docência na disciplina Genética Molecular, coordenado pela Profa. Mariana Pires de Campos Telles, no curso de graduação em Ciências Biológicas, Instituto de Ciências Biológicas, Universidade Federal de Goiás. Apoiado pelo projeto "Desenvolvimento de marcadores, genotipagem e caracterização genômica de espécies do Cerrado (GENPAC 02)" - CNPq 563839/2010-4.

A principal função da atividade é demonstrar como funciona uma das etapas para a análise dos dados de sequenciamento de primeira e/ou segunda geração, utilizando uma ferramenta da Bioinformática conhecida como "BLAST" (*Basic Local Alignment Search Tool*). Ao resolver os enigmas propostos, que simulam situações reais, graduandos ou pós-graduandos manipularão as ferramentas em busca de respostas e consolidarão seus conhecimentos sobre a análise de dados de sequenciamento e as múltiplas possibilidades de uso e aplicações em diferentes áreas.

A principal função da atividade proposta é mostrar uma das etapas para a análise dos dados de sequenciamento, utilizando uma ferramenta da Bioinformática conhecida como "BLAST" (*Basic Local Alignment Search Tool*). As metodologias de sequenciamento geram uma grande quantidade de dados que podem ser organizadas e avaliadas pelas ferramentas analíticas da Bioinformática, uma ciência que usa a matemática e a computação para armazenar, comparar e analisar sequências do DNA, RNA ou de proteínas. Por meio do BLAST é possível inferir relações funcionais e evolutivas entre sequências. A atividade proposta apresenta três sequências de nucleotídeos, denominadas enigma 1, enigma 3 e enigma 5 e duas sequências de aminoácidos, denominadas enigma 2 e enigma 4 que deverão ser analisadas usando a ferramenta "BLAST" para analisar a similaridade entre as sequências que estão sendo analisadas, em comparação com outras sequências armazenadas nos bancos de dados públicos. A atividade permitirá aos estudantes consolidar o conhecimento sobre a análise de dados de sequenciamento e, assim, ampliar a visão e as possibilidades de análises no nível molecular, uma área importante da genética que tem interface com as diferentes áreas das ciências.

PROBLEMA PROPOSTO

O DNA (ácido desoxirribonucleico) e o RNA (ácido ribonucleico) armazenam e transmitem a informação genética nos organismos. Tanto o DNA quanto o RNA são compostos por unidades químicas (monômeros) chamadas de nucleotídeos, unidos em longas cadeias não ramificadas. Por outro lado, a estrutura de uma proteína é comumente definida em termos de quatro níveis hierárquicos. A estrutura primária diz respeito à sequência de resíduos de aminoácidos. A estrutura secundária inclui resíduos de aminoácidos arranjados em conformações particularmente estabilizadas por pontes de hidrogênio, que originam padrões regulares e repetitivos. A estrutura terciária inclui todos os aspectos do padrão de enovelamento tridimensional de uma proteína. Proteínas que apresentam duas ou mais subunidades possuem ainda a estrutura quaternária, que descreve como as várias subunidades estão dispostas no espaço. As metodologias e o equipamento disponíveis atualmente nas áreas de Genética e Biologia Molecular possibilitam sequenciar o DNA, o RNA e as proteínas, permitindo investigar questões biológicas importantes em diversas áreas das ciências.

As metodologias de sequenciamento permitem identificar a ordem das bases nitrogenadas no DNA ou de aminoácidos nas proteínas. Com o passar dos anos, essas metodologias têm aprimorado e possibilitado a geração de um volume de dados cada vez maior. O **método de Sanger** foi proposto primeiramente para analisar sequências de nucleotídeos e durante vários anos foi o único método disponível. Na última década surgiram novas tecnologias de sequenciamento de DNA, denominados de "**segunda geração**" ou, como conhecido em inglês, *Next Generation Sequencing* (NGS), que têm permitido sequenciar genomas inteiros e gerar um volume maior de sequências, em um curtíssimo espaço de tempo, e com custo reduzido. O sequenciamento de proteínas também utiliza novas tecnologias, tais como a **espectrometria de massa**, que também tem permitido uma aceleração na disponibilização de sequências de aminoácidos de proteínas descobertas por grupos de pesquisas em todo o mundo. Um importante banco de dados de sequências é o GenBank (http://www.ncbi.nlm.nih.gov/genbank/) que armazena um grande conjunto de dados que é disponibilizado de forma gratuita ao público.

Uma das maneiras de comparar sequências de nucleotídeos ou aminoácidos é a partir da utilização da ferramenta "BLAST", um algoritmo que analisa comparativamente a sequência de interesse com outras sequências conhecidas e armazenadas nos banco de dados, de diversos organismos. Existem diferentes tipos de BLAST (Tabela 1) que podem ser acessados a partir do *web site* do NCBI (*National Center for Biotechnology Information*) (http://BLAST.ncbi.nlm.nih.gov/BLAST.cgi?CMD=Web&PAGE_TYPE=BLASTHome).

Método de Sanger, também conhecido como método dideoxi, requer a síntese enzimática de uma fita de DNA complementar à fita analisada usando dideoxinucleotídeos, nucleotídeos que quando inseridos interrompem a síntese do DNA, pois não possuem o grupo hidroxila no carbono 3' que é necessário para formação da próxima ligação fosfodiéster. São gerados fragmentos de diferentes tamanhos que são analisados em gel de eletroforese. A análise conjunta dos fragmentos no gel permite ler a sequência de nucleotídeos do fragmento.

Existem várias tecnologias de **sequenciamento de segunda geração** (NGS) comercialmente disponíveis, entre elas a Roche454, a Illumina Solexa, a Life Technologies (SOLiD) e a Helicos BioSciences. Cada uma delas possui características particulares, mas no geral usam métodos baseados em amplificação ou de molécula única na preparação de reações. Para o sequenciamento existem vários métodos como ciclo de terminação reversível, sequenciamento por ligação ou piro sequenciamento.

Para a identificação por **espectrometria de massa**, as proteínas são separadas por eletroforese ou cromatografia e em seguida analisadas por ionização a laser, técnica através da qual são gerados íons e estes analisados através da relação massa-carga.

Tipo	Análise realizada
BLASTn	Procura uma sequência de nucleotídeos em um banco de dados de sequências de DNA
BLASTp	Procura uma sequência de aminoácidos em um banco de dados de proteínas
BLASTx	Procura uma sequência de nucleotídeos em um banco de dados de proteínas
tBLASTn	Procura uma sequência de aminoácidos em um banco de dados de nucleotídeo
tBLASTx	Procura uma sequência de nucleotídeos traduzidos em um banco de dados de nucleotídeos traduzidos

Tabela 1.
Análises realizadas com cada tipo de BLAST.

Alinhamento de sequência – O BLAST faz o alinhamento simples e local. O alinhamento simples é aquele realizado entre duas sequências. O alinhamento local é aquele no qual apenas pedaços ou trechos das sequências são comparados.

No BLAST, a análise de similaridade entre sequências é realizada por meio do **alinhamento da sequência** submetida ao programa (chamada de "*Query*"), com sequências do banco de dados (denominada "*Subject*"). Alinhar duas sequências consiste em comparar a correspondência de nucleotídeos entre elas. Dessa forma, o alinhamento organiza e compara sequências visando identificar regiões de similaridade. No alinhamento, os nucleotídeos nos quais as bases são correspondentes ou iguais são chamados de "match" e nas bases em que não são correspondentes de "*mismatch*". No alinhamento também podem existir regiões com inserção e/ou deleção de um ou mais nucleotídeos na sequência (*indel*), onde se pode inserir um hífen, indicando um "***gap***" (Figura 1).

Gap são espaços introduzidos entre os nucleotídeos ou aminoácidos a fim de obter o melhor alinhamento possível.

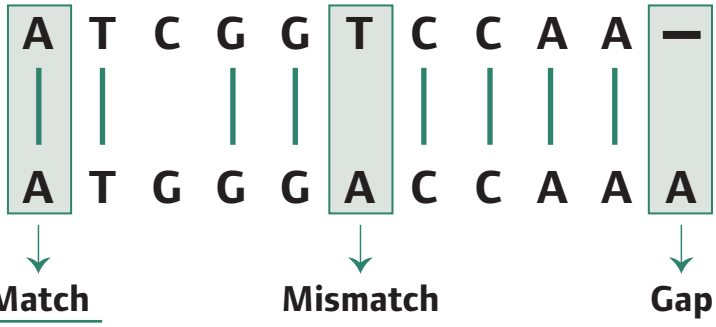


Figura 1.
Exemplo de alinhamento de sequência de nucleotídeos do DNA.

Quando todas as bases do *Query* pareiam com as bases da sequência *Subject*, o alinhamento recebe o nome de "Full-Lenght". No entanto, nem sempre as sequências são alinhadas por completo. Em muitos casos, apenas trechos conservados entre as sequências são correspondentes no alinhamento. Nesse sentido, durante a análise no BLAST a sequência de interesse (*Query*) é alinhada com várias sequências do banco de dados (*Subject*) e pode ser escolhido o melhor alinhamento com base em alguns parâmetros como, por exemplo, o valor de *score* e o *E-value*.

O valor de *score* para cada alinhamento é calculado por meio da atribuição de valores negativos de penalidade para os gaps e *mismatchs*. O melhor alinhamento é aquele que possui maior valore de *score*. O *E-value* também é um parâmetro de confiança para análise de alinhamentos e corresponde ao número esperado de sequências com similaridade tão alta quanto a encontrada por acaso. Quanto maior o *E-value*, menor a confiança de que as correspondências sejam reais, e também maior a chance da similaridade entre as sequências serem devido ao acaso. Geralmente um alinhamento só é considerado real se o *E-value* for menor que 10^{-20}, valor utilizado como ponto de corte.

Quando uma sequência é analisada no BLAST, é importante verificar o valor de *score* e o *E-value* para decidir qual o melhor alinhamento. O melhor resultado da análise no BLAST é chamado de "*Best-Hit*" e, na maioria das vezes, ele é o primeiro resultado da tabela de resultados gerada no BLAST. Se a sequência em análise não tiver similaridade significativa com nenhuma sequencia do banco de dados, tem-se um "*No hit*". Neste caso o *E-value* possui um valor alto.

Nesta atividade está sendo proposta uma análise por meio do BLAST de cinco sequências denominadas "enigma". As sequências do enigma 1, do enigma 3 e do enigma 5 são compostas de nucleotídeos, enquanto as sequências do enigma 2 e do enigma 4 são compostas de aminoácidos. Para a análise das sequências "enigmas" serão usados os cinco tipos de BLAST. O primeiro desafio da atividade é escolher o tipo de BLAST adequado para solucionar o questionamento de cada enigma. Ao submeter as sequências para análise no BLAST, é preciso anotar os valores de *score* e *E-value* e definir o alinhamento *Best-Hit* para responder às questões complementares.

Sequência enigma 1

- enigma 1

TGCAAAAGCAAAATTGATTACTCTGTTTTGGTTTTGGATCAGATCAACCAAAGGAAGCGAAGAAGAAGAAGACTTTATACAC
TCTCGACGCTATTGATTTCGTCTGTGCAGTTTTGTTTCTCTATCAGTTAGTTTTAAAGATGAGTTCTTCAGAGAGTGTGGAAAAC
GAGTGCATGTGTTGGGCTGCAAGAGATCCATCTGGTCTTCTTTCTCCTCATACTATCACTCGCAGGTCTGTTACAACTGACGATG
TTTCACTCACAATCACTCATTGTGGAGTGTGTTACGCTGATGTTATCTGGAGTAGAAACCAACATGGAGACTCCAAGTACCCTT
TGGTTCCTGGGCATGAGATTGCTGGAATAGTGACTAAGGTTGGGCCTAATGTTCAACGATTCAAAGTTGGAGACCATGTTGGTG
TTGGAACGTATGTTAACTCCTGCAGGGAGTGTGAATATTGTAATGAAGGACAAGAAGTTAATTGTGCGAAAGGAGTTTTTACT
TTCAATGGCATTGATCATGATGGCTCTGTTACTAAAGGAGGCTACTCTAGTCACATTGTTGTTCATGAAAGGTACTGCTACAAGA
TACCTGTGGACTATCCCTTGGAATCAGCTGCACCATTACTCTGTGCTGGAATCACGGTTTATGCTCCTATGATGCGTCACAATAT
GAATCAACCTGGTAAATCTCTTGGGGTGATCGGGCTAGGTGGTCTTGGACACATGGCGGTTAAGTTTGGCAAGGCTTTTGGACT
TAGTGTTACGGTTTTTAGCACCAGCATTTCCAAGAAAGAAGAAGCTTTGAATCTGCTAGGAGCTGAGAATTTCGTTATCTCATCT
GACCATGACCAGATGAAGGCACTAGAGAAATCTCTAGACTTTCTAGTTGACACAGCATCTGGTGATCACGCGTTTGATCCTTACA
TGTCTCTCTTGAAGATTGCTGGAACTTATGTATTGGTTGGTTTCCCAAGTGAAATTAAAATCAGTCCTGCCAATCTCAATCTTGG
TATGAGAATGCTCGCTGGAAGTGTAACCGGGGGGACCAAAATAACACAGCAAATGTTAGATTTCTGTGCAGCTCATAAGATTTA
TCCAAACATAGAGGTGATTCCCATTCAAAAGATAAACGAAGCTCTCGAAAGAGTGGTGAAGAAGGACATCAAGTACCGTTTCG
TGATTGACATCAAGAACTCCCTCAAATAGATGTTGCTCAAAGGAAGGAATAATGGAGTCTGTAATAAGAGAATAATACTCACTGC
TACAATCTTTATTACGTATTTTCTCGTTTTTCATTAGTAAAGCAAATAAATTAAGACT

Sequência enigma 2

- enigma 2

VSSCSGVSPGVAGPRQGSQGWVELGFSLGGSGTNPLVLVGKPRPVPARCVLWEERGLGALGWSAGQQRSPEWERPSRAVSPS
SLPQGTEVHDRRGEAVRALYQELLCSGRPASPVSPRPAPPLGLRRAWSRPSPLSHRPGCLDRSHPHLPQCFSKQKFGSLALA
GPVSAV

Sequência enigma 3

- enigma 3

GGTTCCTCTTTTTCTATCTTCCTTCTAGCCCTTCTCTCTTGCTTGACCGTGCCCGCGTCAGCCTACCAAGTACGCAACTCCTCGGG
CCTCTACCATGTCACCAATGATTGCCCTAACTCGAGTATTGTGTACGAGACGGCCGATACCATACTACACTCTCCGGGGTGTG
TCCCTTGCGTTCGCGAGGGTAACAACTCGAGGTGTTGGGTGGCGGTGGCCCCCACAGTCGCCACCAGGGACGGCAAACTCCC
CACAACGCAGCTTCGACGTCATATCGATCTGCTTGTCGGGAGCGCCACCCTTTGCTCGGCCCTCTATGTGGGGGACTTGTG
CGGGTCTGTCTTTCTTGTCGGCCAACTGTTCACCTTCTCCCCCAGGCGCCACTGGACAACGCAAGACTGCAACTGCTCTATC
TACCCCGGCCATATAACGGGTCACCGAATGGCATGGGATATGATGATGAACTGGTCCCCTACAACAGCGCTGGTAGTAGCTCAGC
TGCTCAGGGTCCCGCAAGCCATCGTGGATATGATCGCTGGTGCCCACTGGGGAGTCCTAGCGGGCATAGCGTATTTCTCCATGG
TAGGGAACTGGGCGAAGGTCCTGGTAGTGCTGTTGCTGTTTGCCGGCGTCGATGCCGAGACCTACACCACTGGGGGGAGTGC
TGCCAGGATCACGACCGGACTCGTCAGTCTTTTCAGTCCGGGCGCCAAGCAGAATATCCAGCTGATGAACACCAACGGCAGT
TGGCACATCAATCGCACGGCCCTGAACTGTAATGCGAGCCTCGACACCGGCTGGGTGGCGGGGCTCTTCTACCACCACAAATT
CAACTCCTCGGGCTGCCCCGAGAGGATGGCCAGCTGTAGACCCCTTGCCGATTTTGACCAGGGCTGGAGCCCTATCACCCACGC
CAACGGAAGTGGCCCCGAACATCGCCCCTACTGCTGGCACTACCCCCCAAAGCCCTGTGGTATCGTGCCAGCACGGAACGTGTG
TGGCCCAGTGTATTGTTTCACTCCTAGCCCCGTGGTGGTGGGAACGACCGACGTGCTGGGCGTGCCTACCTACACCTGGGGTGG
TAATGATACGGACTTCTTCGTCCTTAACAACACCAGGCCACCGTTGGGCAATTGGTTTGGTTGCACCTGGATGAACTCGTCTG
GATTTACCAAAGTGTGCGGAGCGCCTCCTTGCGTCATCGGAGGGGTGGGCAACAACACCTTGCACTGCCCTACTGACTGTTTC
CGCAAGCATCCAGAAGCCACATACTCTCGGTGTGGCTCCGGTCCCTGGATCACGCCCAGGTGCCTGGTCCACTATCCTTATAGG
CTTTGGCACTACCCTTGCACCGTCAACTACACCCTGTTCAAGGTCAGGATGTACGTAGGAGGGGTCGAGCACAGGCTGGAGGT
TGCTTGCAACTGGACGCGGGGCGAGCGTTGTGATCTGGACGACAGGGACAGGTCCGAGCTCAGCCCGCTGCTGCTGTCCACCA
CGCAGTGGCAGGTCCTTCCGTGCTCCTTCACGACCTTGCCAGCCTTGACCACTGGCCTCATCCATCTCCACCGGAACATCGTGGA
CGTGCAATATTTGTACGGGGTGGGGTCAAGCATTGTGTCCTGGGCCATCAAGTGGGAATACGCCATTCTCTTATTTCTCCTGCT
TGCAGACGCGCGCATCTGCTCTTGCTTGTGGATGATGTTACTCATATCCCAAGCGGAGGCG

Sequência enigma 4

- enigma 4

```
MVPQTETKAGAGFKAGVKDYRLTYYTPDYMVKDTDILAAFRMTPQPGVPPEECGAAVAAESSTGTWTTVWTDGLTSLDRYK
GRCYDIEPVAGEDNQYIAYVAYPIDLFEEGSVTNLFTSIVGNVFGFKALRALRLEDLRIPPAYAKTFQGPPHGIQVERDKINKY
GRGLLGCTIKPKLGLSAKNYGRAVYECLRGGLDFTKDDENVNSQPFMRWRDRFTFVAEAIYKSQAETGEIKGHYLNVTAATSE
EMMKRAECAKDLGVPIIMHDYLTAGLTANTSLAHYCRDTGLLLHIHRAMHAVIDRQRNHGIHFRVLAKALRLSGGDHLHSG
TVVGKLEGEREVTLGFVDLMRDDYIEKDRSRGIYFTQDWCSLPGVMPVASGGIHVWHMPALVEIFGDDACLQFGGGTLGHPWG
NAPGAAANRVALEACTQARNAGVDLARKGGDVIRAACKWSPELAAACEVWKEIKFEFETIDKL
```

Sequência enigma 5

- enigma 5

```
GGAACCATTATGCACTCTTCAATAGTTTTGGCCACCGTGCTCTTTGTAGCGATTGCTTCAGCATCAAAAACGCGAGAGCTATGCA
TGAAATCGCTCGAGCATGCCAAGGTTGGCACCAGCAAGGAGGCGAAGCAGGACGGCATCGACCTCTACAAACATATGTTCGAG
CACTATCCAGCAATGAAGAAATACTTCAAGCATCGTGAAAATTATACACCGGCCGATGTCCAAAAGGATCCCTTCTTTATTAAACA
AGGTCAAAATATCTTGCTCGCCTGTCACGTTTTGTGCGCCACATACGACGATCGTGAGACATTCGACGCGTACGTTGGTGAGCT
GATGGCACGACACGAGCGGGACCATGTTAAAGTACCGAATGATGTTTGGAATCACTTCTGGGAACATTTCATCGAGTTTCTGG
GAAGTAAGACCACGTTGGACGAGCCAACCAAGCACGCATGGCAAGAGATCGGTAAAGAATTCTCACATGAAATCAGCCACCA
CGGTCGACATTCGGTTCGCGACCATTGCATGAACTCGTTGGAGTATATCGCGATCGGCGATAAGGAACATCAAAAGCAGAATGG
CATTGACCTTTACAAGCATATGTTCGAGCATTATCCACATATGAGAAAGGCATTCAAGGGACGCGAAAACTTCACGAAAGAAGA
CGTTCAAAAGGACGCATTCTTCGTTAACAAGGACACAAGATTCTGTTGGCCCTTCGTATGCTGTGACTCCTCATACGATGACGAG
CCAACATTCGACTATTTTGTTGATGCCCTAATGGATCGTCATATCAAAGATGATATTCATCTACCTCAGGAACAATGGCATGAGTTC
TGGAAATTGTTTGCCGAATATTTGAACGAAAAGAGTCACCAACATTTGACAGAAGCCGAGAAACATGCATGGAGTACAATAGGT
GAGGACTTCGCGCATGAGGCCGATAAGCATGCAAAGGCCGAAAAAGACCATCATGAAGGAGAGCACAAAGAGGAACACCAC
TGAACCAACCCGTCGTCGTTCAACTTAAGCCTTCAGCTTAAGCTCGAGCTAAAGCCTCAGCTTGAGCTCAATCTTATGTCCTCAGG
CCTAAACTTGAATTTTAAAAGCATTTTGTTGAAGCAGTGCTAGCCAATCTCTTATCTTATCGGTGCTATTATCAATTTACTCTATGC
CACCCCCCCCCCCCTCTCTCTGTTCTCTATTTGATATTCTGTTCTTTTAGTGCCAGATGTTAGTACCAGATGTTATTTTCTGCATAAT
TTTCTTCTCTTTACTTCGTTATTTTTTCGTTCTTCTATTTTTATGGCCAATTTTTGTGATGTCGAAGTCAATAAAACCATTTTT
```

O **formato FASTA** permite representar sequências de nucleotídeos e aminoácidos. Na primeira linha do arquivo há o símbolo maior (>) seguido pelo título da sequência e, nas linhas seguintes, a sequência representada pelo código de única letra.

INSTRUÇÕES PARA O PROFESSOR

1. É recomendável que o professor aplique esta atividade em turmas que já possuem conhecimentos sobre genética molecular, assim o público alvo mais adequado são alunos de Graduação ou Pós Graduação na área de Biológicas. Nesse sentido, é importante o conhecimento prévio acerca da estrutura e função do DNA e das proteínas. Além disso, é preciso que os alunos tenham conhecimentos das tecnologias de análise das macromoléculas incluindo métodos de sequenciamento, como o Sanger, e espectrometria de massa, metodologias que geram os dados das sequências que serão analisados a partir do BLAST.
2. A atividade será realizada individualmente por cada estudante que receberá do professor os problemas propostos. Cada estudante deverá executar a análise no BLAST e ao final entregar um relatório. A atividade pode ser realizada como aula prática em laboratório de informática ou como exercício de casa.
3. O professor deverá fornecer aos estudantes, por meio eletrônico, dois arquivos. O primeiro arquivo em **formato FASTA** contendo as cinco sequências a serem consultadas e as questões enigmas. O segundo arquivo pode ser enviado em PDF contendo as orientações para os estudantes e as questões adicionais (A a D) que devem ser respondidas.

PROCEDIMENTOS PARA OS ESTUDANTES

1. Acessar o site do NCBI (http://www.ncbi.nlm.nih.gov/). No menu lateral direito, acessar a ferramenta BLAST. Em "Basic BLAST" acessar o tipo de BLAST que considerar mais adequado para responder cada um dos cinco enigmas. Na caixa de texto copiar a sequência enigma em formato FASTA, ou ainda, escolher o arquivo a ser submetido. Escolher a base de dados a ser consultada. Dar o comando de execução do BLAST e resolver os 5 enigmas abaixo descritos.

- **Enigma 1.** Um pesquisador que trabalha com um determinado gene utilizou um par de ***primers*** que amplifica a região do genoma que corresponde a esse gene específico. Após obter o sequenciamento desta região do genoma, o pesquisador precisa confirmar se houve a amplificação e o sequenciamento da região alvo. Para isso, ele decidiu utilizar o BLAST como ferramenta de bioinformática. Qual versão do BLAST ele precisará utilizar e por quê? Qual o nome do gene com o qual o pesquisador trabalha?
- **Enigma 2.** Foi isolado o material biológico de um indivíduo do qual foi extraída uma dada proteína. Após o sequenciamento, a sequência de aminoácidos está disponível, mas não se sabe de que proteína se trata. Qual BLAST você usaria para identificar a proteína e por quê? Qual proteína foi extraída?
- **Enigma 3.** Você trabalha com sequências de cDNA e quer saber se uma delas codifica ou não alguma proteína no organismo do qual você isolou a amostra. A ferramenta BLAST pode ser usada para responder esta pergunta. Que tipo de BLAST você usaria e por quê? A sequência de cDNA do enigma 3 codifica alguma proteína? Se sim, qual?
- **Enigma 4.** Para investigar uma dada proteína depois de isolá-la e sequenciá-la, você pode usar o BLAST para chegar até a sequência correspondente no genoma? Que BLAST você usaria e por quê?
- **Enigma 5.** Você trabalha com amostras de pacientes com deficiências hematológicas e ainda não sabe exatamente o que causa a anomalia. O sequenciamento dos genes das vias hemolíticas destes pacientes é realizado e você decide usar o BLAST como uma primeira ferramenta para tentar identificar a similaridade dessas sequências com outras sequências nos bancos de dados. Que BLAST você usaria? Houve similaridade com algum gene? Qual? Isso faz sentido com a sua pergunta biológica?

2. Interprete os resultados gerados em cada questão enigma e responda também às seguintes perguntas:

- **Questão A.** Em "Graphic Summary" em que cor as sequências inseridas aparecem e o que isso significa na interpretação dos resultados?
- **Questão B.** Em "Descriptions", se houver, qual foi o resultado do *Best-Hit*, seu *E-value* e *Score* total para as sequências. E o que significa "Complete cds"?
- **Questão C.** Com base nos resultados obtidos no BLAST para as sequências, responda se foi possível identificar a que organismo pertence e também qual é a sua função.
- **Questão D.** Qual a importância da utilização do BLAST no estudo de espécies não modelo, ou que não possuem genoma descrito e de espécies que já têm informações genéticas bem conhecidas?

Primer é uma sequência curta de nucleotídeos que serve como iniciador para a replicação do DNA.

RESPOSTAS PARA OS ENIGMAS

Enigma 1. *Nucleotide BLAST*, pois trata-se de uma sequência de nucleotídeos que será alinhada contra sequências de banco de dados de nucleotídeos. O BLAST indicará em que região do genoma a sequência foi alinhada com maior similaridade, que deve ser exatamente o gene alvo do estudo em questão. A sequência enigma 1 corresponde ao gene "cinnamyl-alcohol dehydrogenase mRNA".

Enigma 2. *Protein BLAST*, pois trata-se de uma sequência de aminoácidos que será alinhada contra sequências de um banco de dados de aminoácidos. O BLAST mostrará em seu *Best Hit* a qual proteína e de qual organismo a sequência *Query* é mais similar, possibilitando descobrir qual a proteína putativa proveniente da amostra em estudo. A sequência enigma 2 corresponde à proteína "Cytochrome b-245, alpha polypeptide, isoform CRA".

Enigma 3. *BLASTx*, a sequência *Query* é composta por nucleotídeos (DNA codificado a partir de RNA) e o resultado deste BLAST será em aminoácidos caso se trate de uma região codificadora de proteína. Também será possível descobrir a proteína putativa e o organismo com maior similaridade para aquela região que contenha sequência depositada no GeneBank. A sequência do enigma 3 corresponde ao Gene do envelope do vírus da hepatite C.

Enigma 4. *T BLASTn*, a sequência *Query* é composta por aminoácidos e o resultado retornaria uma sequência de nucleotídeos, o que permitiria acesso a uma provável sequência genômica para aquela proteína.

Enigma 5. TBLASTx, trata-se de uma sequência de nucleotídeos que é traduzida sendo alinhada contra um banco de dados de nucleotídeos traduzidos. Sim, houve similaridade com um gene mutado da hemoglobina. Isso faz sentido de acordo com a pergunta biológica e pode ajudar a explicar o quadro clínico dos pacientes.

Questão A. O resultado gerado pela comparação de sequências (Figura 2) mostra primeiramente um gráfico sumário, contendo os *scores* associados à correspondência encontrada entre as sequências e quanto maiores os *scores*, mais alta a correspondência. Essa informação também é evidenciada nos *scores* com que as sequências inseridas aparecem. No resultado da sequência enigma 1, o gráfico exibe o alinhamento em rosa e vermelho (*score* de 80 a >=200), que indica alta similaridade. Para a sequência enigma 2, o gráfico exibe *score* de <40 a >=200. Para a sequência enigma 3, o gráfico exibe *score* de >=200. Para a sequência enigma 4, o gráfico exibe *score* >=200. Para a sequência enigma 5, o gráfico exibe *score* de <40 a >=200.

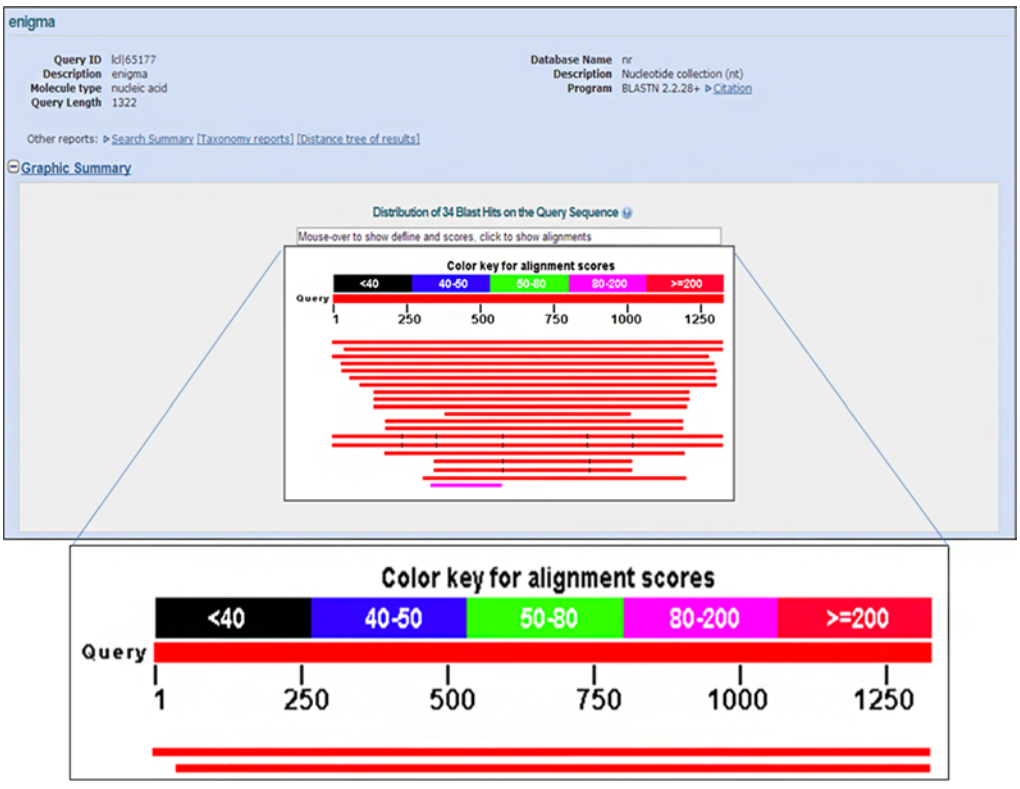
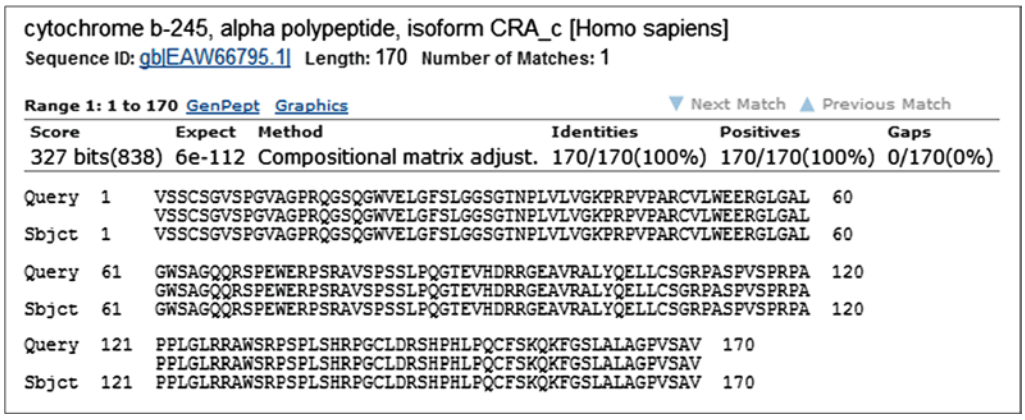


Figura 2.
Resultado gráfico do BLAST para a sequência enigma 1 evidenciando a chave de cores para *scores* de alinhamento.

Questão B. Em geral, o melhor resultado é o "*Best-Hit*", ou seja, o primeiro resultado da tabela de resultados do BLAST. Acessando esse resultado com um clique, aparecem as sequências comparadas e parâmetros onde *Query* é a sequência fornecida para a busca e *Subject* é a sequência conhecida. Foram encontrados *Best-Hits* para as cinco sequências fornecidas (Tabela 2). Quando se tem um "Complete cds", significa que todas as bases informadas no *Query* pareiam com as bases da sequência *Subject*.

Enigma	Best-hit	Espécie	Scores	E-value
1	cinnamyl-alcohol dehydrogenase mRNA, complete cds	*Arabidopsis thaliana*	2442	0,0
2	Cytochrome b-245, alpha polypeptide, isoform CRA_c	*Homo sapiens*	327	6^{e-112}
3	Hepatitis C virus isolate 1A14.32/MN-2/C5/22.05.1995 envelope gene	Virus hepatite c	1157	0,0
4	Large subunit of Rubisco	*Stigeoclonium helveticum*	949	0,0
5	Pseudoterranova decipiens (frameshift mutated) hemoglobin mRNA, complete cds	*Homo sapiens*	797	0,0

Tabela 2.
Best-hit para cada uma das cinco sequências enigmas.


cytochrome b-245, alpha polypeptide, isoform CRA_c [Homo sapiens]
Sequence ID: gb|EAW66795.1| Length: 170 Number of Matches: 1

Range 1: 1 to 170 GenPept Graphics ▼ Next Match ▲ Previous Match

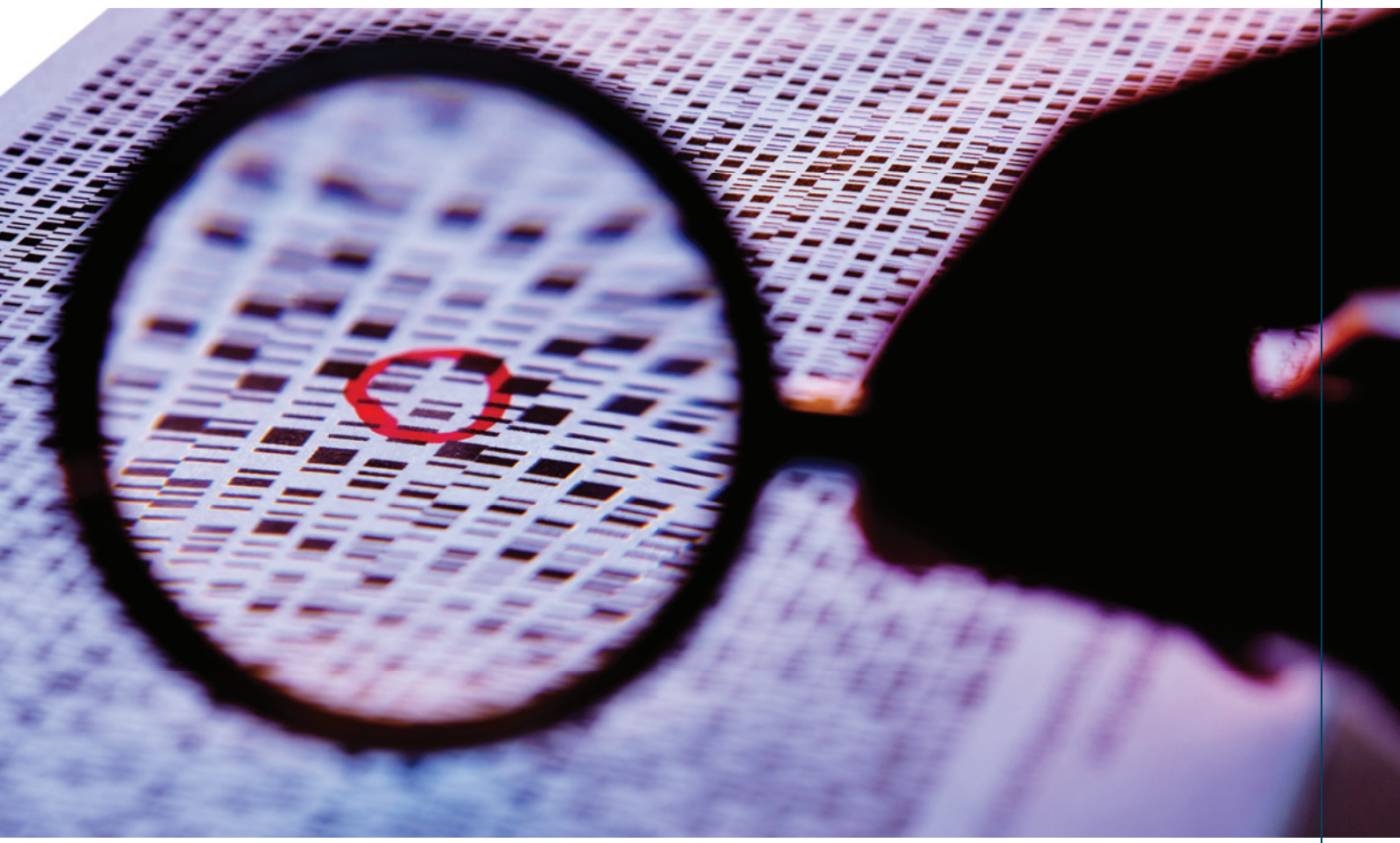
Score	Expect	Method	Identities	Positives	Gaps
327 bits(838)	6e-112	Compositional matrix adjust.	170/170(100%)	170/170(100%)	0/170(0%)

```
Query  1    VSSCSGVSPGVAGPRQGSQGWVELGFSLGGSGTNPLVLVGKPRPVPARCVLWEERGLGAL  60
            VSSCSGVSPGVAGPRQGSQGWVELGFSLGGSGTNPLVLVGKPRPVPARCVLWEERGLGAL
Sbjct  1    VSSCSGVSPGVAGPRQGSQGWVELGFSLGGSGTNPLVLVGKPRPVPARCVLWEERGLGAL  60

Query  61   GWSAGQQRSPEWERPSRAVSPSSLPQGTEVHDRRGEAVRALYQELLCSGRPASPVSPRPA  120
            GWSAGQQRSPEWERPSRAVSPSSLPQGTEVHDRRGEAVRALYQELLCSGRPASPVSPRPA
Sbjct  61   GWSAGQQRSPEWERPSRAVSPSSLPQGTEVHDRRGEAVRALYQELLCSGRPASPVSPRPA  120

Query  121  PPLGLRRAWSRPSPLSHRPGCLDRSHPHLPQCFSKQKFGSLALAGPVSAV  170
            PPLGLRRAWSRPSPLSHRPGCLDRSHPHLPQCFSKQKFGSLALAGPVSAV
Sbjct  121  PPLGLRRAWSRPSPLSHRPGCLDRSHPHLPQCFSKQKFGSLALAGPVSAV  170
```

Figura 3.
Resultado do alinhamento da sequência enigma 2 com a sequência *subject* do banco de dados.



Questão C. Sim, foi possível identificar o organismo e a função de cada uma das sequências informadas, assim como descrito abaixo:

- **Sequência enigma 1:** É o mRNA de uma enzima que catalisa a reação química, cinamil álcool-desidrogenase, de *Arabidopsis thaliana*.
- **Sequência enigma 2:** É um Alpha Citocromo B-245, isoforma CRA de *Homo sapiens*.
- **Sequência enigma 3:** É a sequência de uma proteína "e", de envelope celular de um isolado de Vírus de Hepatite C.
- **Sequência enigma 4:** É um genoma parcial de cloroplasto, gene da proteína Rubisco da alga *Stigeoclonium helveticum*.
- **Sequência enigma 5:** É um RNA ribossomal parcial de gene, região 16 mitocondrial de *Hynobius leechii*.

Questão D. Uma pesquisa no BLAST permite a um investigador comparar uma dada sequência fornecida em uma consulta com a maior biblioteca de base de Dados de Sequências do mundo e identificar sequências que se assemelham à investigada (*Query*) e que estejam acima de um nível mínimo de semelhança. Quando se trabalha com espécies das quais se tem pouco conhecimento de informações genéticas, o BLAST pode ser uma ferramenta muito importante, pois pode ser utilizado, entre outros, para descobrir se as sequências novas em estudo contêm algum gene homólogo já descrito para outra espécie ou ainda se uma destas novas sequências, se sabidamente genes, já foram descritos para outros organismos. Trabalhando com organismos para os quais já se tem informações genéticas disponíveis, por exemplo, organismos modelo, é possível utilizar novos dados adquiridos para esses organismos e compará-los aos dados já publicados no banco de dados do NCBI, podendo inclusive ser verificado o nível de similaridade entre as duas sequências, o que pode fornecer importante informações.

Mitos da Genética Humana

Maria de Nazaré Klautau Guimarães

Universidade de Brasília, Departamento de Genética e Morfologia, Brasília-DF

A página http://udel.edu/~mcdonald/mythintro.html apresenta uma revisão da literatura sobre os mitos da Genética Humana, com acesso gratuito. Foi elaborada por John MacDonald, em 2011, professor e pesquisador do Department of Biological Sciences, University of Delaware. É possível obter o formato pdf no link http://udel.edu/~mcdonald/MythsHumanGenetics.pdf com excelentes ilustrações.

O autor apresenta ampla revisão dos estudos realizados com análises da segregação em famílias e similaridades entre gêmeos para 19 características facilmente identificáveis em humanos como cor dos olhos, lóbulos da orelha, cor do cabelo, sensibilidade ao PTC, capacidade de enrolar a língua etc. Os dados são apresentados em gráficos e tabelas, com discussões e conclusões. São também apresentados, quando existentes, dados sobre a variação dos caracteres entre diferentes populações e são discutidos os aspectos evolutivos. Tais dados levam à reflexão sobre a interpretação do modo de herança dessas características, tornando-se um material excelente de apoio ao professor, como também, de estudo para o aluno de graduação.

A maioria dos trabalhos sobre a base genética dessas características foram realizados nas décadas de 30 a 80 e apresentaram simplificações que levaram a interpretações inadequadas quanto ao modo de herança. Na maioria das vezes, foram interpretados como apresentando herança mendeliana. Tais interpretações inadequadas se apresentam até hoje na maioria dos livros didáticos, tendo se tornado mitos. Apresentam estudos mais recentes sobre algumas dessas características que exploram as bases moleculares da variação, que também são discutidas pelo autor. Apesar das contribuições das técnicas moleculares, ressaltam-se ainda dificuldades na conclusão de alguns estudos, pois algumas características apresentam limitações metodológicas na identificação da variação fenotípica.

Considerando que tais características são de grande interesse por parte dos estudantes, é importante que sejam utilizadas no ensino de genética, de modo que os estimulem a analisar e interpretar a variação fenotípica de maneira adequada. Na maioria das vezes, o autor conclui que a característica não deve ser utilizada como exemplo de herança monogênica. Geralmente, o erro de interpretação é devido à classificação dicotômica da variação fenotípica, desprezando as variações intermediárias.

Por todos esses motivos, é que devemos discutir esses estudos com nossos estudantes, pois a simplificação desses exemplos leva a ideias deterministas sobre a genética humana. À luz dos novos conhecimentos sobre a expressão gênica faz–se necessária uma reflexão sobre a adequação da utilização desses exemplos no ensino para que se possa desmistificar e,ao mesmo tempo, ilustrar como o conhecimento científico evolui ao longo do tempo.

O gene da intolerância à lactose

Renata Palacios[1] e Marcos Edgar Herkenhoff[2]

[1] Departamento de Bioquímica, Centro de Ciências Agroveterinárias, Universidade do Estado de Santa Catarina, Lages
[2] Departamento de Genética, Instituto de Biociências de Botucatu, Universidade Estadual Paulista "Júlio de Mesquita Filho", Botucatu

Autor para correspondência: palaciosrenata2012@gmail.com

A intolerância à lactose acomete em torno de 75% da população mundial e é caracterizada pela ineficiência ou redução na atividade enzimática da lactase, denominada hipolactasia, durante a quebra e digestão da lactose. A hipolactasia pode ser ocasionada por fatores primários, secundários e congênitos. O fator primário, mais frequente, é de natureza hereditária. Esta alteração no gene da lactase surgiu entre 5.000 e 10.000 anos e tornou-se benéfica, garantindo uma vantagem para os seres humanos que tinham o leite como alternativa nutricional.

O GENE DA INTOLERÂNCIA À LACTOSE

A enzima lactase, também conhecida como lactase-florizina hidrolase ou β–-D-galactosidase, está situada na borda em escova dos enterócitos, células responsáveis pela absorção de nutrientes, no intestino delgado. A lactase é responsável por hidrolisar, ou seja, quebrar em partículas menores, a lactose, açúcar não absorvível encontrado no leite, em moléculas menores como a glicose e galactose, para que sejam absorvidas pela mucosa intestinal e liberadas na corrente sanguínea (Figura 1).

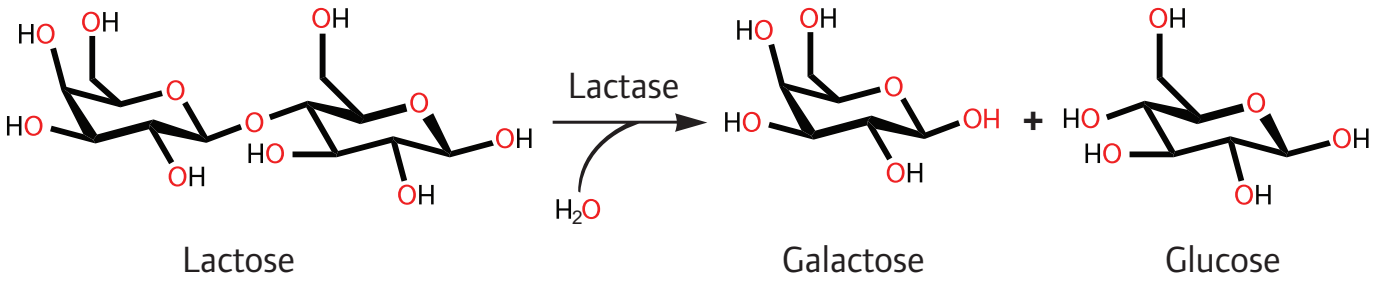


Figura 1.
Quebra da lactose em glicose e galactose em uma reação que envolve a enzima lactase e uma molécula de água. O "O" representa o átomo de oxigênio e, o "H", de hidrogênio. (Fonte: http://commons.wikimedia.org/wiki/File:Lactase.png)

A diminuição da atividade enzimática ou a ausência da enzima é denominada de hipolactasia, ou também conhecida popularmente como lactase não persistente. O declínio na capacidade de quebrar lactose em moléculas menores faz com que esta molécula passe rapidamente para o cólon, onde ocorrerá a sua fermentação pelos microrganismos intestinais. As bactérias localizadas no componente intestinal utilizam a lactose como uma boa fonte de energia e acabam gerando gás metano (CH_4), gás hidrogênio (H), e ácido láctico, causando desconforto por distensão abdominal, flatulência e diarreia. O surgimento dessa sintomatologia gerada pela má absorção intestinal da lactose caracteriza a intolerância a esta molécula.

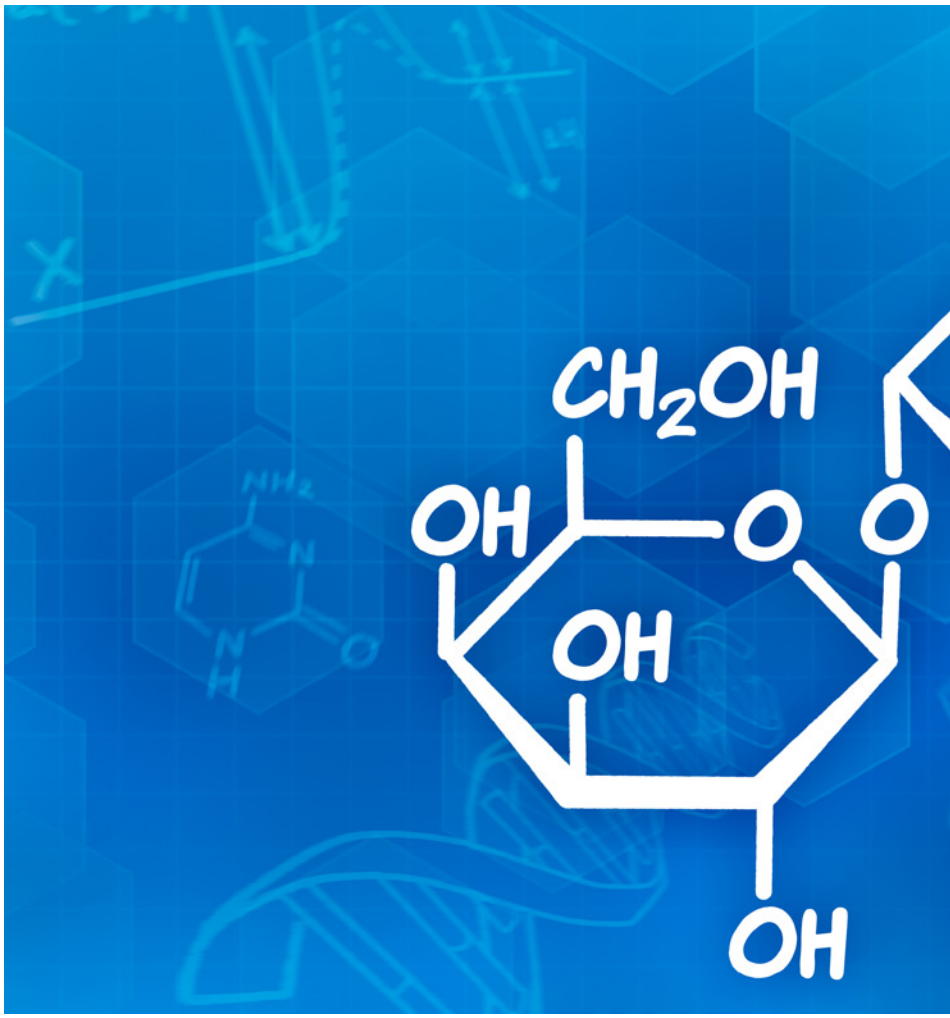
CAUSAS DA INTOLERÂNCIA À LACTOSE

A intolerância à lactose pode ser desencadeada por uma hipolactasia primária, secundária ou intolerância à lactose congênita. A hipolactasia primária ou hipolactasia tipo-adulto é a redução irreversível e geneticamente programada da atividade da enzima lactase e ocorre na maioria da população mundial após o desmame, podendo chegar a um nível de função fisiológica abaixo de 10% durante o envelhecimento. O declínio da enzima também pode ser chamado de lactase não persistente e ocorre em indivíduos que são homozigotos para o alelo recessivo, ou seja, que tenham herdado uma cópia do alelo recessivo do pai e a outra cópia do alelo recessivo da mãe. Dessa forma, o indivíduo precisa ter os dois alelos recessivos para que ele possua a característica, ou expresse o fenótipo, de "lactose não persistente". Portanto, quando os pais são heterozigotos, eles possuem 25% de chances de terem um filho que seja "lactose não persistente".

A hipolactasia secundária e reversível é ocasionada por doenças que causam dano à mucosa intestinal como: giardíase, esquistossomose, doenças inflamatórias intestinais, doença de Crohn, retocolite e enterites infecciosas.

A intolerância à lactose congênita tem um padrão de herança autossômico recessivo, ou seja, o gene em questão está localizado nos cromossomos autossômicos e não nos sexuais. Assim, tanto o homem quanto a mulher possuem duas cópias do gene. No caso da forma autossômica recessiva, é necessário que o indivíduo possua duas cópias do mesmo gene, e que cada cópia do gene tenha o alelo recessivo para gerar o fenótipo. Esta condição é relativamente rara na população e potencialmente letal se não tratada corretamente. A criança que nasce com a falta desta enzima sofre com episódios de diarreias consecutivas que acabam gerando desidratação, uma condição que acarreta um risco grave para um recém-nascido.

A diferença da hipolactasia primária e a congênita está em nível molecular; na hipolactasia primária tipo-adulto a enzima apresenta sua estrutura normal e apenas sofre declínio de sua atividade durante o envelhecimento; na intolerância à lactose congênita, a enzima é ausente ou apresenta mudanças em sua estrutura que impossibilitam que a hidrólise aconteça. Em relação às causas da intolerância à lactose, a hipolactasia primária é a mais frequente.

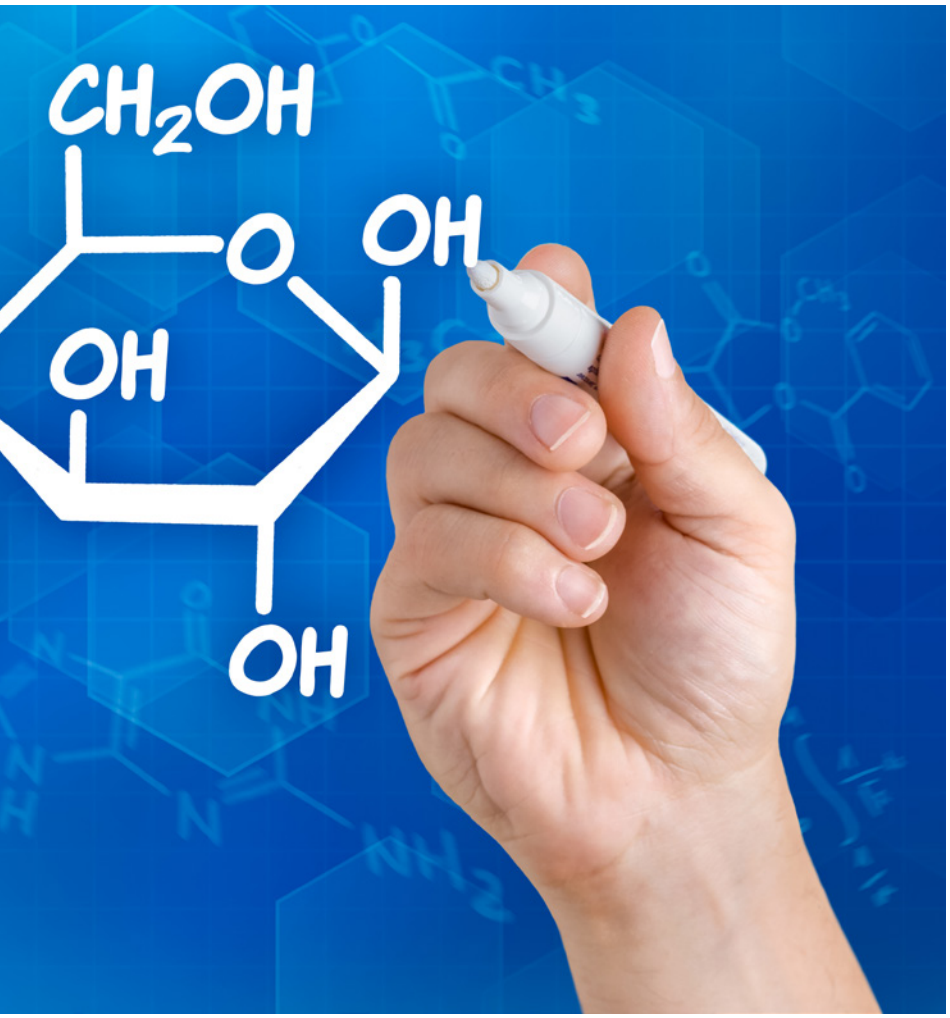


HIPOLACTASIA PRIMÁRIA TIPO-ADULTO E INDIVÍDUOS "LACTASES PERSISTENTES"

A hipolactasia primária acomete em torno de 75% da população mundial, podendo variar conforme as diferentes expressões genéticas, recessivas ou dominantes, nos grupos étnicos. Em uma minoria populacional, essa atividade enzimática pode continuar nor-

malmente, como são os casos encontrados principalmente na Europa, em países próximos ao Mar do Norte, e são chamados de "lactases persistentes". A persistência à lactose é uma característica e também um fenótipo, de natureza hereditária autossômica dominante. Assim sendo, o portador de apenas um alelo que confere a persistência à lactose não apresenta declínio da maturação na expressão da enzima lactase. Esses indivíduos são capazes de degradar o açúcar encontrado no leite em uma mesma proporção durante o envelhecimento, não sofrendo a redução da função fisiológica da enzima e diminuindo as possibilidades de desenvolvimento da intolerância à lactose.

Cada **kb** se refere a 1.000 pb; a sigla pb vem do inglês "pair base" que significa pares de bases.



Éxon é a região do gene que permanece em seu transcrito após o processo de splicing ou excisão dos íntrons do RNA. Os éxons permanecem no RNA maduro e a contêm a informação da sequência de aminoácidos da proteína codificada pelo gene.

Alguns autores discutem que, há milhares de anos, todas as pessoas tinham essa hipolactasia primária. Há mais ou menos 10.000 anos ocorreu a domesticação de animais que produziam leite, como bovinos, caprinos e ovinos. Portanto, naquele período da história o consumo de leite proveniente dessas espécies de animais fora iniciado, e os indivíduos que passavam da infância, continuavam a consumi-los. Dessa forma, mutações que ocorressem no gene da lactase, permitindo que o organismo continuasse a produzir esta enzima, tornar-se-iam vantajosas para os indivíduos que a possuíssem, em relação aos indivíduos que não a apresentassem, pois tal mutação permite que o indivíduo possa usufruir do leite como um alimento rico em nutrientes. Assim sendo, quando surgiu o alelo com a mutação no gene lactase, os indivíduos que possuíam uma cópia deste alelo e puderam usufruir do leite como alimento por mais tempo, consequentemente foram favorecidos pela seleção natural. Tal mutação foi vantajosa para a sobrevivência desses indivíduos nas regiões mais frias, como o norte da Europa, onde a agricultura não era favorável e a inclusão do leite na dieta tornou-se uma estratégia de sobrevivência. As pessoas com a persistência desta enzima eram mais saudáveis e tinham mais filhos; dessa maneira, a frequência deste alelo aumentou.

O GENE RESPONSÁVEL PELA PRODUÇÃO DE LACTASE

O gene responsável pela produção de lactase, denominado LCT, está localizado no cromossomo 2 nos seres humanos. Normalmente esta enzima sofre um declínio na sua capacidade fisiológica, tendo a sua produção inibida após o desmame e resultando na diminuição da expressão de lactase no intestino delgado. Este gene possui quase 49,34 **kb**, tem 17 **éxons** (Figura 2) e seu transcrito (RNAm) possui 6.241 bases. O RNA mensageiro é traduzido em uma proteína que apresenta 1.927 aminoácidos em sua composição.

Alguns segmentos populacionais humanos exibem uma lactase persistente e essa mutação surgiu entre 5000 a 10.000 anos, coincidindo com a domesticação dos animais que produzem leite. Estudos relacionaram a ocorrência dessa persistência da lactase a

duas mutações na **extremidade 5' do gene** da LCT, e estas mutações surgiram de forma independente em populações distintas. A extremidade 5' é responsável pela regulação deste gene, ou seja, por sua inibição na fase adulta. Ambas as mutações, uma sendo uma troca de uma citosina por uma timina no nucleotídeo número 13.910 e outra uma troca de uma guanina por uma adenina no nucleotídeo número 22.018, foram associadas à persistência da lactase. O conhecimento a respeito do gene e de sua sequência de DNA são importantes informações para se utilizar na hipolactasia primária.

A extremidade 5' do gene antecede o ponto de início de transcrição do gene e contém sua região promotora.

Figura 2. Representação esquemática do gene LCT, cujo produto é a proteína lactase. Cada coluna representa um éxon, região do gene que codifica a sequência de aminoácidos que compõe a proteína. O gene LCT possui 17 éxons. (Fonte: http://www.ensembl.org/Homo_sapiens/Location/View?db=core;g=ENSG00000115850;r=2:136545081-136595080)

A IMPORTÂNCIA NO DIAGNÓSTICO DA INTOLERÂNCIA À LACTOSE

Por apresentarem sintomas semelhantes que envolvem diarreia, náuseas, distensão abdominal e eczema, antes de ser diagnosticada no paciente, essa intolerância primária pode ser confundida com outra patologia, a alergia ao leite, no entanto, ambas são fisiologicamente distintas.

As alergias alimentares caracterizam um processo de hipersensibilidade, que é uma alteração exagerada do organismo em uma resposta imune frente a uma substância que ele considera estranha. Neste caso, o organismo do indivíduo começa a produzir anticorpos contra um determinado alimento, e reage como se fosse um patógeno em potencial. Essa reação provoca a liberação de histamina por **mastócitos**. O diagnóstico da patologia que acomete o paciente é realizado através da suspeita clínica com associação da sintomatologia e da comprovação com exames laboratoriais.

O princípio do teste mais difundido em laboratórios de análises clínicas que detecta a intolerância à lactose ou a sua função fisiológica diminuída é o teste de tolerância com sobrecarga oral de lactose, baseado na coleta de glicemia em jejum, seguida da ingestão de 50g de lactose e de coletas de sangue periférico sucessivas entre 0, 15, 60 e 90 minutos. O objetivo é avaliar a curva glicêmica após a ingestão de lactose. O indivíduo com intolerância não consegue quebrar a lactose em glicose e galactose, pela deficiência enzimática e acaba impedindo que os açúcares entrem na corrente sanguínea e elevem a glicose sérica, ou seja, a glicose presente no sangue.

Por avaliar a função fisiológica e esta poder variar em relação à capacidade habitual de digerir certa concentração tolerável pelo paciente e sofrer interferência direta dos diferentes grupos étnicos, que podem compor a ascendência do indivíduo, este teste pode se apresentar inespecífico. Portanto, testes para a detecção das mutações LCT 13910 C>T e LCT 13910 G>A tornam-se ferramentas muito mais efetivas para o diagnóstico desta enfermidade. Atualmente estão disponíveis várias ferramentas que permitem a detecção destes alelos com a mutação que confere a persistência ou não à lactase.

Mastócitos são células do tecido conjuntivo que possuem em seu interior grânulos de substâncias que são mediadores químicos como serotonima e histamina , que são liberadas nas reações alérgicas, e também de heparina, um anticoagulante.

PARA SABER MAIS

BARBOSA, C. R.; ANDREAZZI, M. A. Intolerância à lactose e suas consequências no metabolismo do cálcio. *Revista Saúde e Pesquisa* v. 4, n. 1, p. 81-86, 2011.

MATTAR, R.; MAZO, F. C. Intolerância à lactose: Mudança de paradigmas com a biologia molecular. *Rev. Associação de Medicina Brasileira* v. 56, n. 2, p. 230, 2010.

PRETTO, F. M. et al. Má absorção de lactose em crianças e adolescentes: Diagnóstico através do teste de hidrogênio expirado com o leite de vaca como substrato. *Jornal de Pediatria* v. 78, n. 3, p. 213-218, 2010.

SAHI, T. Hypolactasia and lactase persistence: Historical review and the terminology. *Scand J. Gastroenterol Suppl* v.1, n. 6, p. 202, 1994.

WANG, T. et al. The genetically programmed down-regulation of lactase in children. *Gastroenterology* v. 114, n. 6, p.1230-6, 1998.