Genética na Escola – ISSN: 1980-3540

Código de Barras de DNA: uma atividade para entender a identificação de espécies

Camila Penha Abreu Souza[1], Ana Paula Sampaio Amorim[2], Joana de Jesus Alves[3], Ligia Tchaicka[4]

[1] Pós-graduanda do Programa de Pós-Graduação em Biodiversidade e Conservação, Universidade Federal do Maranhão, São Luis, MA
[2] Pós-graduanda do Programa de Pós-Graduação em Ciência Animal, Universidade Estadual do Maranhão, São Luís, MA
[3] Graduada em Ciências Biológicas Licenciatura, Universidade Estadual do Maranhão, São Luís, MA
[4] Departamento de Química e Biologia, Universidade Estadual do Maranhão São Luís, MA

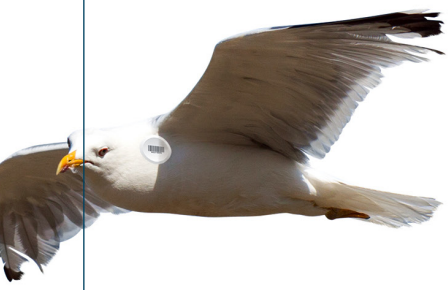
Autor para correspondência: ltchaicka@yahoo.com.br

Palavras-chave: DNA barcoding, taxonomia molecular, citocromo oxidase I, ensino de ciências, bold systems, evolução

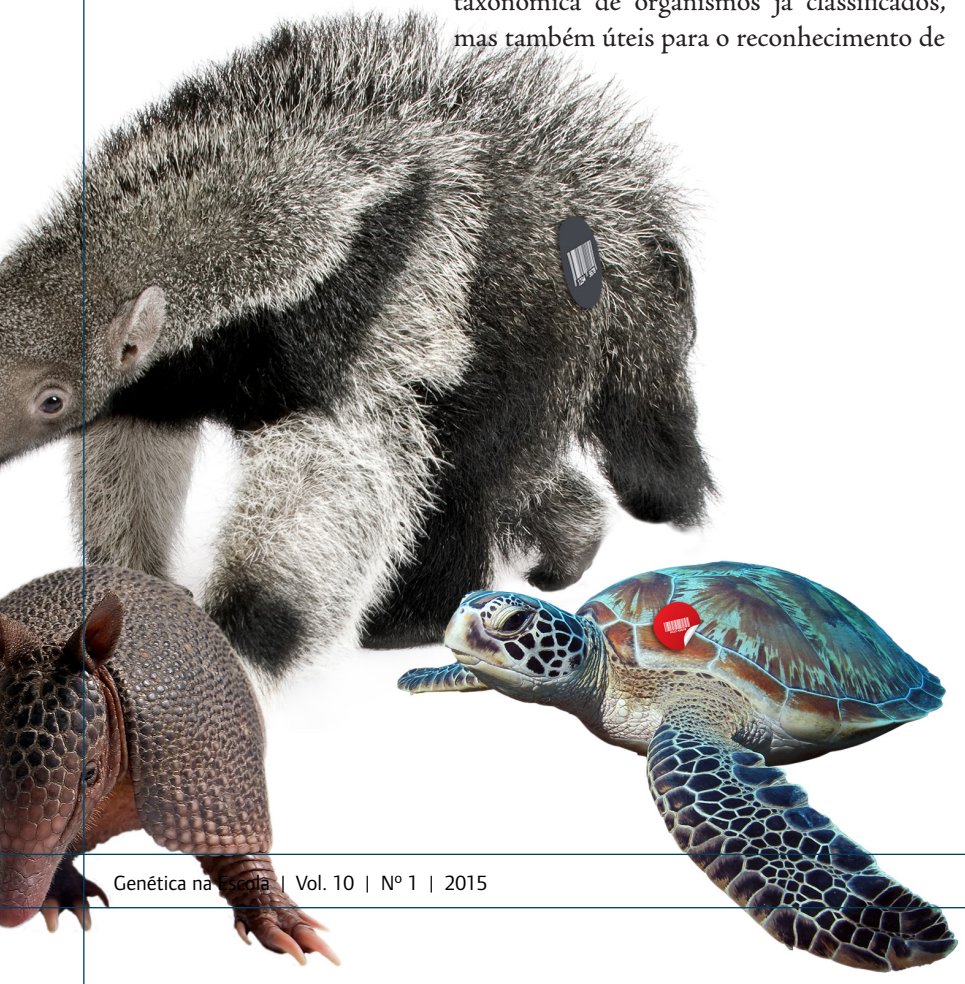


Genética na Escola – ISSN: 1980-3540

Grande parte da biodiversidade que existe sequer foi catalogada pelos cientistas e, em muitos casos, a identificação das espécies é muito trabalhosa ou simplesmente impossível. Uma técnica proposta para agilizar esse processo é o Código de Barras de DNA. A proposta deste trabalho foi apresentar essa técnica a professores em um curso de formação continuada, bem como discutir com eles a pertinência de abordá-la no ensino básico.







A diversidade biológica tem grande importância ambiental, evolutiva e até econômica, mas está ameaçada pela utilização desordenada dos recursos naturais. Acredita-se que muitas espécies poderão ser extintas antes de serem conhecidas e existe, assim, a necessidade de ampliar os esforços para o estudo de organismos das mais diversas regiões de nosso planeta. Uma das principais dificuldades encontradas em pesquisas dessa natureza é a correta identificação de um espécime.

O avanço da genética nas últimas décadas possibilitou o desenvolvimento de ferramentas moleculares aplicáveis não só à identificação taxonômica de organismos já classificados, mas também úteis para o reconhecimento de novas espécies. Uma dessas ferramentas é a técnica de identificação por Código de Barras de DNA (DNA *Barcoding*).

Este trabalho relata uma atividade cujo objetivo foi apresentar a professores do ensino básico essa ferramenta, bem como discutir uma proposta que poderia ser desenvolvida pelos docentes em aulas.

CONTEXTO

O Grupo de Estudos em Genética e Conservação da Universidade Estadual do Maranhão (CGC-UEMA) realizou o curso "*Atualização em Genética: trazendo a Biodiversidade e a Biotecnologia para a sala de aula*", destinado a professores do ensino básico, que atuam em Biologia e Ciências da Natureza. O objetivo foi apresentar e discutir assuntos atuais em genética, bem como desenvolver formas alternativas para tornar o ensino mais interessante e próximo do aluno.

O curso contou com uma fase ministrada à distância e com uma fase posterior com aulas presenciais. Participaram 20 professores, de várias faixas etárias, atuantes no Ensino Fundamental e no Médio, tanto da rede pública como privada, de escolas da capital e do interior do Maranhão.

Código de Barras de DNA: identificando espécies foi uma das várias atividades desse curso, buscando simular uma pesquisa e, para tanto, os participantes assumiram o papel de geneticistas interessados na identificação de espécimes provenientes de um museu. Essa atividade foi totalmente desenvolvida em ambiente virtual, por meio de um fórum de discussão que esteve ativo por 20 dias.

Genética na Escola – ISSN: 1980-3540

DNA *Barcoding* ou Código de Barras de DNA é um método proposto em 2003 pelo grupo de pesquisa de Paul Hebert da Universidade de Guelph, Canadá, que tem por finalidade a identificação de diferentes espécies por meio de uma região gênica padrão. Esse método pode ser comparado a um scanner, como os que existem em supermercados, que utilizam as listras pretas do código de barras para identificar os produtos. A região gênica utilizada deve ser universal e variável, ou seja, deve estar presente em todos os organismos vivos e apresentar diferenciação suficiente entre espécies para distingui-las. Para os animais e a maioria dos eucariotos a região utilizada é um gene codificador de proteínas do DNA mitocondrial conhecido como COI- Citocromo *c* Oxidase Subunidade I (ver Box 2). Para plantas terrestres pode-se utilizar genes do cloroplasto, como o matK e o rbcL. Esse tipo de método tem a vantagem de poder ser utilizado em espécimes em qualquer fase da vida (ex.: ovos, larvas, adultos) e em fragmentos do ser vivo em questão (folhas, pelos, sangue etc.).

Box 1.
DNA *Barcoding*

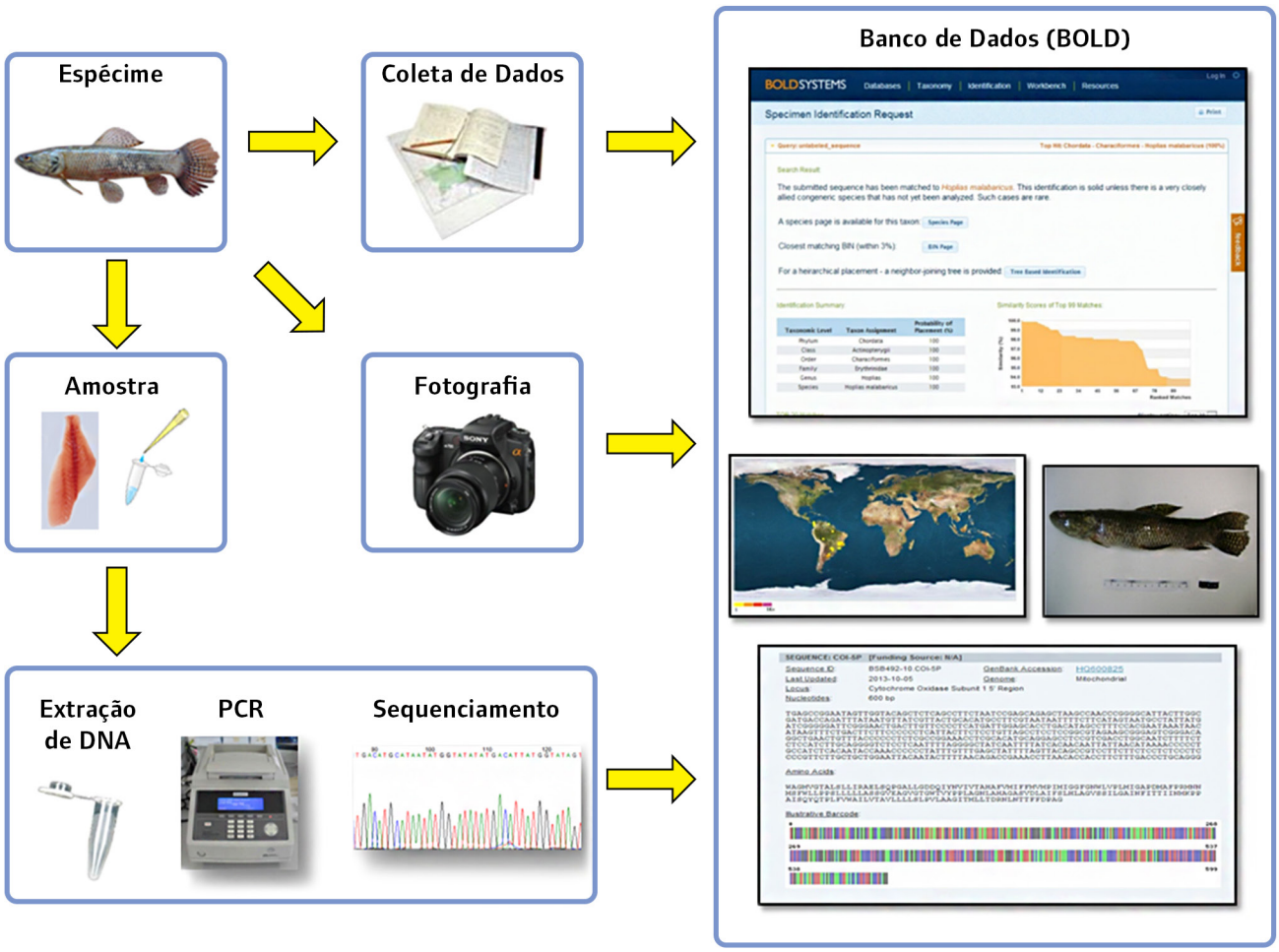


Figura 1.
Esquema do método de DNA *barcoding*

O gene da Subunidade I da Citocromo *c* Oxidase ou COI, é um gene do DNA mitocondrial com aproximadamente 600 pares de bases que codifica uma enzima envolvida no transporte de elétrons na respiração celular. Ele é utilizado na identificação via DNA *barcoding* por sua taxa de mutação ser adequada para distinguir espécies próximas.

Box 2.
Gene COI

DESENVOLVIMENTO DA ATIVIDADE

Por meio de textos, foram apresentados os conceitos de biodiversidade, classificação taxonômica, código de barra de DNA, diversidade genética e marcadores moleculares. É importante citar que as técnicas básicas utilizadas em biologia molecular, como eletroforese, extração de ácidos nucleicos e sequenciamento de DNA já haviam sido abordadas anteriormente.

Os participantes foram então estimulados a discutir sobre o código de barras de DNA e sua aplicação, pesquisar como o método foi desenvolvido e com qual finalidade; em relação às técnicas de biologia molecular, quais foram utilizadas e por que foram escolhidas as regiões do DNA empregadas; finalmente, os grupos de pesquisas foram formados para estudar diferentes espécies.

Em seguida, foi proposto um problema hipotético: cada um dos participantes recebeu tecidos biológicos de cinco espécimes depositados em um museu que necessitava saber quais espécies o material pertencia. Assim, realizou a extração de DNA dos tecidos, PCR (Reação em Cadeia da Polimerase) e sequenciamento do gene COI. Vale lembrar que essas etapas não foram realizadas pelos participantes do curso; o que de fato receberam para análise foram cinco sequências de espécies diferentes (retiradas do banco de dados do GenBank), que representavam os produtos desses sequenciamentos. Receberam também um roteiro com as indicações de como acessar o Bold Systems (ver Box 3) para que pudessem identificar as espécies à qual pertencia cada sequência (veja esse roteiro em anexo).

Os participantes então tiveram a tarefa de acessar o banco de dados e obter dados sobre a espécie que possui a sequência de nucleotídeos analisada. Além da identificação, obtiveram informações sobre a ecologia do animal, imagens e outras informações que estão disponíveis na página. Em seguida, todos foram estimulados a relatar seus resultados, comunicando também as inferências que poderiam fazer a partir de seus dados, como especulações sobre a área geográfica de origem da amostra, se a espécie em questão é rara ou ameaçada, entre outras. Ao final, avaliaram a atividade e a possibilidade de utilizá-la para o ensino na educação básica.

O BOLD (*Barcode of Life Database* – www.boldsystems.org) é um banco de dados público internacional, onde estão depositadas as sequências do gene COI de espécies que já foram investigadas pelo método do código de barras de DNA (atualmente contém mais de 3 milhões de sequências). Esse banco possui dados moleculares, morfológicos e de distribuição das espécies. No Brasil existe um consórcio de instituições para gerar dados sobre a identificação molecular da biodiversidade brasileira, o BrBol (*Brazilian Barcode of Life* - www.brbol.org/pt-br) que deposita as informações que obtém no Bold Systems.

Box 3.
Bold Systems

AVALIAÇÃO DA EXPERIÊNCIA

Todos os participantes foram capazes de realizar a atividade utilizando o roteiro proposto, sendo que alguns relataram dificuldades com a língua inglesa ao utilizar o site do BOLD Systems. Durante o processo, foi possível discutir alguns conceitos e temas além dos que foram propostos inicialmente, como: evolução biológica e origem comum das espécies, a mutação como geradora de variabilidade, a comparação entre taxonomia clássica e molecular, a importância das coleções biológicas, e ainda dados da biologia das espécies pesquisadas. Para além dos conceitos, a discussão englobou as políticas públicas de conservação da biodiversidade, a legislação ambiental e questões econômicas relacionadas à biodiversidade.

O uso do ambiente virtual mostrou-se bastante eficaz para o trabalho neste curso. Os participantes postaram no fórum virtual não apenas as tarefas solicitadas, mas também informações pesquisadas em artigos científicos e sites especializados, enriquecendo a discussão. Alguns conteúdos do fórum foram sistematizados e são apresentados a seguir, relatando as ideias expressas pelos participantes (na transcrição das postagens os participantes são identificados pela letra P, seguida de um número).

É possível afirmar que os professores participantes do curso compreenderam o método de identificação via Código de Barras de DNA. Uma evidência disso é que termos específicos associados ao método (tais como amostragem, análise laboratorial, comparação com banco de dados biológicos) foram mencionados de maneira adequada nas postagens que fizeram no fórum virtual.

O BOLD Systems foi o primeiro banco de dados biológicos com o qual os participantes tiveram contato. Alguns relataram entusiasmo pela descoberta da existência de "*um banco contendo tantas sequências de DNA dos mais diversos organismos e a possibilidade do acesso pela internet*" (P1). A descoberta da existência desses repositórios e as diversas ferramentas da biologia molecular estimularam a procura por outros bancos de dados que permitissem o livre acesso às sequências de nucleotídeos de diversos organismos, e que pudessem ser utilizados pelos professores participantes com seus alunos. Os bancos citados pelos participantes foram: GenBank (do National Center for Biotechnology Information -http://www.ncbi.nlm.nih.gov/Genbank.); DDBJ - DNA Data Bank of Japan (do Japan National Institute of Genetics -http://www.ddbj.nig.ac.jp); EMBL – Bank (do European Molecular Biology Laboratory - http://www.ebi.ac.uk/embl); PDB - Protein Data Bank (repositório de dados sobre estruturas 3D de moléculas de proteínas pertencente ao Research Collaboratory for Structural Bioinformatics – http://www.rcsb.org/pdb/home/home.do).

Possibilidades de aplicação da técnica de Código de Barras de DNA foram bastante discutidas: foi lembrada a identificação de espécies crípticas, de híbridos, de partes de animais que são comercializados ilegalmente, de composição de produtos que são adulterados e vendidos como produtos originais. Uma reportagem que relata uma atividade de Código de Barras de DNA desenvolvida com alunos de ensino médio dos Estados Unidos foi adicionada ao fórum pelos participantes, que se mostraram interessados em desenvolvê-la em suas escolas (Box 4).

Esse é o título de um artigo da revista Ciência Hoje que apresenta uma reportagem sobre um grupo de cientistas dos EUA que desenvolveu uma pesquisa com alunos do ensino médio para analisar o DNA de alguns alimentos por meio da técnica de Código de Barras de DNA. Eles montaram um pequeno laboratório de genética e investigaram a composição dos alimentos servidos por alguns restaurantes. A pesquisa revelou que 25% dos restaurantes analisados enganavam seus clientes servindo alimentos adulterados. Para ter acesso ao artigo completo e a uma entrevista com um dos cientistas que desenvolveu o trabalho segue o link: http://cienciahoje.uol.com.br/alo-professor/intervalo/2012/01/o-atum-que-era-tilapia

Box 4.
O atum que era tilápia

Genética na Escola – ISSN: 1980-3540



Relataram ainda que por meio dessa experiência adquiriram ou aprofundaram o conhecimento sobre técnicas de genética e biologia molecular, e que muitos deles não conheceram durante a graduação. Também revelaram ter compreendido o motivo da escolha do gene COI para a identificação molecular dos animais: *"por apresentar variação nas sequências de pares de bases suficientes para diferenciação de espécies, no entanto não é um bom marcador molecular para as plantas"*(P2).

Em suma, os professores participantes consideraram a atividade interessante e viram nela possibilidades de utilização para o ensino na educação básica, *"pois permitiu aliar a pesquisa com a descoberta"*(P3).

POSSIBILIDADES DE USO DA ATIVIDADE

A atividade permitiu que os participantes associassem os conceitos de genética e biologia molecular com aplicações no estudo da biodiversidade. Despertou o interesse em usar metodologias diferenciadas no âmbito da educação básica e também inserir a pesquisa científica no cotidiano da sala de aula. Uma importante observação é a de que a bioinformática pode ser especialmente interessante para o ensino de ciências na modalidade à distância: os participantes podem realizar todas as tarefas em um computador que tenha acesso à internet, sem a necessidade de um laboratório específico, equipamentos sofisticados ou reagentes muito caros.

A atividade pode ser realizada com alunos de graduação, pós-graduação e em cursos de formação continuada de professores. De acordo com a avaliação dos professores participantes, poderia também ser desenvolvida na educação básica e, especialmente, no ensino médio. Entretanto, é necessário frisar que a compreensão plena de como funciona o método depende do entendimento de como funcionam algumas técnicas de biologia molecular e alguns conceitos sobre diversidade biológica. Além disso, a discussão precisa ser adequada ao nível de ensino no qual ela seja conduzida.

Genética na Escola – ISSN: 1980-3540

ANEXO I – ROTEIRO UTILIZADO NA ATIVIDADE

Código de Barras de DNA: uma atividade para identificação de espécies

Objetivo: Essa atividade apresenta uma simulação de pesquisa através da qual você será um geneticista e tem como desafio identificar e caracterizar espécies.

Problema Proposto: O grupo de pesquisa, ao qual você pertence, recebeu de um museu cinco amostras para que sejam identificadas. Você decidiu utilizar a técnica de Código de Barras de DNA. Os procedimentos realizados em laboratório foram: extração de DNA, amplificação do gene COI pela técnica de PCR e leitura em sequenciador automático. Agora, de posse das sequências de cada espécime, você irá identificá-las no site *Barcode of Life Data Systems*.

Siga os passos:

1. Acesse o **BOLD** (Barcodeof Life Data Systems): http://www.boldsystems.org/
2. Selecione a opção **Identification** no menu da página inicial do BOLD.
3. Será aberta uma nova janela. No menu **Search Databases**, selecionar a opção **Species Level Barcode Records**.
4. Copie e cole no espaço **Enter sequences in fasta format** a sequência que está dentro do retângulo a seguir. Depois de colar, clique em **Submit**.
5. Após algum tempo de processamento, aparecerá o resultado da busca (**Search Result**). Caso a sequência inserida seja encontrada no banco de dados, surgirá um ou mais nomes científicos, que são a identificação da amostra. Clique sobre esses nomes para saber mais sobre a espécie.
6. Repetir o procedimento desde o item 3 com cada sequência a seguir, separadamente:

Sequência 1

CCTATATCTAATCTTCGGAGCATGAGCTGGCATAGTTGGAACCGCCCTCAGCCTCCTCATCCGCGCAGAACTTGGA
CAACCTGGGACTCTACTAGGAGATGATCAAATTTATAATGTAATCGTCACTGCCCATGCCTTCGTAATAATTTTCTT
TATAGTAATACCAATCATGATCGGTGGTTTCGGAAATTGACTAGTCCCACTCATAATCGGCGCCCCCGACATAGCAT
TCCCCCGTATAAACAACATAAGCTTCTGACTACTTCCCCCATCATTTCTTTTACTTTTAGCATCTTCCACAGTAGAAG
CTGGGGCCGGAACAGGGTGAACAGTTTATCCCCCCCTCGCTGGTAACCTAGCCCATGCCGGTGCCTCAGTAGACC
TAGCCATCTTCTCCCTTCACTTAGCAGGTATTTCCTCTATCCTAGGTGCTATCAACTTCATTACAACCGCCATCAACA
TAAAACCCCCAGCCCTATCCCAATACCAAACCCCACTATTCGTATGATCAGTACTTATCACTGCCGTCCTTCTCCTAC
TCTCTCTCCCAGTCCTCGCTGCCGGCATTACTATACTACTAACAGACCGAAACTTAAATACTACATTCTTTGATCCAG
CTGGAGGAGGAGACCCAATCCTGTACCAACACCTCTTCTGATTTTTCGGTCATCCAGAAGTCTATATCCTCATTCTA

Sequência 2

ACTCTTTACCTTCTATTTGGTGCCTGAGCCGGTATGGTAGGAACTGCTCTTAGTCTCCTAATCCGGGCCGAACTAGGTCA
ACCTGGCACACTACTAGGAGATGATCAAATTTATAATGTGGTCGTTACTGCCCATGCTTTTGTGATGATTTTCTTCATAGTAA
TACCTATTATGATTGGAGGGTTTGGTAACTGATTGGTCCCATTAATAATTGGAGCCCCTGACATAGCATTCCCCCGAATGAA
TAACATGAGCTTCTGACTTCTTCCTCCATCTTTTTTACTTCTACTTGCTTCATCTATGGTGGAGGCCGGAGCAGGGACTGGA
TGAACAGTATATCCACCCTTAGCCGGTAATCTGGCTCATGCGGGAGCATCCGTAGATCTAACCATTTTCTCACTCCACCTAG
CAGGTGTCTCTTCGATCTTGGGTGCTATTAATTTTATCACCACTATTATTAATATAAAACCTCCTGCCATATCTCAATACCAA
ACACCCCTTTTTGTATGATCAGTTTTAATCACTGCAGTCCTATTACTCCTATCACTCCCAGTCCTAGCAGCAGGAATTACTA
TGCTATTAACAGATCGAAACCTAAATACCACATTCTTTGATCCTGCCGGAGGAGGAGATCCTATCTTATACCAACACTTATTC

Sequência 3

TCCACCAACCATAAAGACATCGGCACCCTATACCTAATCTTCGGGGCATGAGCAGGAATAGTAGGCACAGCACTCAGCC
TATTAATCCGCGCAGAACTAAGCCAACCAGGTACTCTTCTAGGAGATGACCAAATTTACAACGTTATCGTCACAGCCCACG
CTTTCATCATAATCTTCTTTATAGTTATACCAATCATAATCGGTGGCTTCGGAAACTGACTTGTTCCATTAATAATTGGAG
CACCAGATATAGCATTCCCACGTATAAACAACATAAGCTTTTGACTCTTACCTCCTTCATTATTACTACTTCTAGCATCATCA
GGAATTGAAGCAGGCGCAGGCACAGGCTGAACAGTGTACCCCCCATTAGCTGGAAACCTAGCCCACGCCGGTGCTTCTG
TAGATCTAACCATCTTCTCACTCCACCTAGCTGGTGTGTCTTCAATTTTAGGTGCTATCAACTTCATTACCACAGCAATCAA
CATAAAATCCCCTGCCATATCACAATACCAAACACCCTTATTCGTATGATCCGTACTAATTACAGCTGTCCTATTACTACTTTC
TCTACCAGTACTCGCTGCAGGTATTACCATACTACTCACAGACCGAAATCTAAATACAACCTTCTTTGATCCTTCAGGAGGA
GGAGATCCAATCCTATATCAACACCTATTCTGATTCTTTGGACACCCTGAAGTATACATCCTAATCCTCCCAGGATTTGGCA
TAATCTCCCACATTGTCACCTATTATGCCGGCAAAAAAGAACCATTCGGCTACATAGGAATAGTTTGAGCAATAATATCCAT

Sequência 4

GGTGCTAGTTGACGGTATAGTAGGAACAGCTCTAAGCCTTCTAATTCGCGCTGAATTAGGCCAACCCGGAACTCTGCTCGGA
GACGACCAAATCTACAACGTAGTTGTAACCGCACACGCATTTGTAATAATCTTCTTCATAGTAATACCAATCATAATTGGAG
GATTCGGTAACTGACTTGTTCCCCTAATAATTGGTGCTCCCGATATAGCATTTCCCCGAATAAATAATATAAGCTTCTGACTCC
TCCCTCCCTCATTCCTACTACTCCTCGCATCCTCTATAGTTGAAGCTGGGGCAGGAACAGGCTGAACCGTGTACCCTCCCT
TAGCAGGCAACCTAGCCCATGCAGGAGCTTCAGTAGATCTAACCATTTTCTCTTTACACTTAGCAGGAGTTTCCTCAATTTTA
GGAGCCATCAACTTCATTACAACAATTATCAACATAAAGCCCCCCGCAATGTCACAATACCAAACCCCTCTGTTCGTATGATC
CGTAATAATTACCGCCGTACTACTACTACTCTCGCTCCCTGTATTAGCAGCCGGCATCACAATGCTATTAACAGACCGGAACC
TAAATACAACCTTCTTCGACCCGGCAGGAGGAGGAGACCCTATTCTATATCAACACTTATTCTGATTTTTTGGTCACCCGAA
AGTTTAGGCCCGCCCCCTGCGCGCCTGCCTCTGCGTGAGGGCACTGCCATGCCCCCCCTTCGGACTGAAGAATAAAAACAA
AAATCCGCGCTGCTTTGCTTGCTATTTTGGTTGGCCCCCCGATAGAGAGAGAGAGGAAGGAGAAGATAGTTGTCGGGCTGG

Sequência 5

ATGAATTTTTTCTACAAATCATAAAGATATTGGAACATTATATTTTATTTTTGGTATCTGAGCAGGAATAGTAGGAACATC
TCTTAGTTTATTAATTCGTACTGAACTAGGAAATCCTGGATCATTAATTGGTGATGATCAAATTTATAATACTATTGTTACAG
CTCATGCTTTTATTATAATTTTTTTTATAGTTATACCAATTATAATTGGAGGTTTCGGAAATTGATTAGTTCCATTAATATTAG
GAGCTCCTGATATAGCATTTCCTCGAATAAACAATATAAGATTTTGATTATTACCCCCATCTTTAATTTTATTAATTTCTAGAA
GAATTGTAGAAAATGGAGCAGGAACAGGATGAACAGTGTACCCCCCTCTGTCATCTAATATTGCTCATGGAGGATCTTCTG
TAGATTTAGCCATTTTTTCTTTACATTTAGCTGGAATTTCATCAATTTTAGGAGCAATTAATTTTATTACTACTATTATTAACA
TACGAATTAATAATATAAAATTTGATCAAATACCTTTATTTGTATGAGCAGTAGGTATTACAGCACTTTTATTATTGCTTTCAT
TACCTGTACTTGCAGGTGCAATTACTATACTTTTAACTGATCGTAATATTAATACATCATTTTTTGACCCAGCAGGAGGTGG
TGATCCTATTCTTTATCAACATTTATTTTGATTTTTTGGACATCCTGAAGTTTATATTTTAATTTTACCAGGATTTGGTATAAT
TTCCCATATTATTTCCCAAGAAAGAGGAAAAAAGGAAACATTTGGTTCTTTAGGAATAATTTATGCTATAATAGCAATTGG
TTTATTAGGATTTATTGTTTGAGCTCATCATATATTTACTGTAGGAATAGATATTGATACACGAGCATATTTTACTTCAGCAA
CAATAATTATTGCCGTACCAACAGGTATTAAAATTTTTAGTTGATTAGCAACATTACATGGAACTCAAATTAATTATAGTCCA
TCAATATTATGAAGATTAGGATTTATTTTTTTATTTACTGTTGGAGGTTTAACAGGTGTTGTATTAGCTAATTCATCAATTGA
TATTACTCTCCATGATACATATTATGTTGTTGCTCATTTTCATTATGTTTTATCTATAGGAGCAGTATTTGCTATTTTAGGAG
GATTTGTTCATTGATACCCCTTATTTACAGGTCTAACATTAAATAATTATTTATTAAAAATTCAATTTTTATCAATATTTATTG
GAGTTAATTTAACATTTTTTCCACAACATTTTTTAGGTTTAGCAGGTATACCTCGACGATACTCAGATTATCCCGATATTTTT
TTATCTTGAAATATTATTTCTTCTTTTGGTTCTTATATTTCTTTTTTTTCCATAATTATAATTATTATTATCATTTGAGAATCTA
TAATTAATCAACGTATTATTTTATTTTCATTAAATATACCTTCTTCTATTGAATGATATCAAAATACTCCTCCCGCAGATCATTC

Questões para discussão

A. Qual(is) espécie(s) corresponde(m) a cada sequência analisada? Caracterize cada amostra de acordo com as informações disponíveis no site BOLD (Gênero, Espécie, Família, Descrição do Táxon, Local de Ocorrência/Coleta).

B. Por que foram utilizadas sequências do gene COI para a identificação das espécies abordadas na atividade?

C. A técnica de Código de Barras permite a identificação molecular de espécies a partir de uma pequena amostra de DNA. Com base na atividade e em seus conhecimentos da técnica, como essa técnica poderia ser usada para estudar a biodiversidade?