

BIOINFORMÁTICA:

UMA JANELA PARA A
BIODIVERSIDADE

Organizadoras

Julia Gabriele Dani
Fernanda Pessi de Abreu
Scheila de Ávila e Silva

BIOINFORMÁTICA:

UMA JANELA PARA A
BIODIVERSIDADE

Fundação Universidade de Caxias do Sul

Presidente:
Dom José Gislón

Universidade de Caxias do Sul

Reitor:
Gelson Leonardo Rech

Vice-Reitor:
Asdrubal Falavigna

Pró-Reitor de Pesquisa e Pós-Graduação:
Everaldo Cescon

Pró-Reitora de Graduação:
Terciane Ângela Luchese

*Pró-Reitora de Inovação e
Desenvolvimento Tecnológico:*
Neide Pessin

Chefe de Gabinete:
Givanildo Garlet

Coordenadora da EDUCS:
Simone Côte Real Barbieri

Conselho Editorial da EDUCS

André Felipe Streck
Alexandre Cortez Fernandes
Cleide Calgaro – Presidente do Conselho
Everaldo Cescon
Flávia Brocchetto Ramos
Francisco Catelli
Gelson Leonardo Rech
Guilherme Brambatti Guzzo
Karen Mello de Mattos Margutti
Márcio Miranda Alves
Simone Côte Real Barbieri – Secretária
Suzana Maria de Conto
Terciane Ângela Luchese

Comitê Editorial

Alberto Barausse
Università degli Studi del Molise/Itália

Alejandro González-Varas Ibáñez
Universidad de Zaragoza/Espanha

Alexandra Aragão
Universidade de Coimbra/Portugal

Joaquim Pintassilgo
Universidade de Lisboa/Portugal

Jorge Isaac Torres Manrique
*Escuela Interdisciplinar de Derechos
Fundamentales Praeeminentia Iustitia/
Peru*

Juan Emmerich
*Universidad Nacional de La Plata/
Argentina*

Ludmilson Abritta Mendes
Universidade Federal de Sergipe/Brasil

Margarita Sgró
*Universidad Nacional del Centro/
Argentina*

Nathália Cristine Vieceli
Chalmers University of Technology/Suécia

Tristan McCowan
University of London/Inglaterra





BIOINFORMÁTICA:

UMA JANELA PARA A BIODIVERSIDADE

Organizadoras

Julia Gabriele Dani
Fernanda Pessi de Abreu
Scheila de Ávila e Silva

© das organizadoras

1ª edição: 2025

Preparação de texto: Giovana Letícia Reolon

Leitura de prova: Helena Vitória Klein

Revisão técnica: Aline Zanetti, Emanuel Scherdien da Rosa, Jean Lucas Benvenuti,

Joelmir Mazzoni, Johnatan Vilasboa, Letícia Osório e Mariane Dorneles

Editoração: Ana Carolina Marques Ramos

Capa: Julia Gabriele Dani

Dados Internacionais de Catalogação na Publicação (CIP)

Universidade de Caxias do Sul

UCS – BICE – Processamento Técnico

B615 Bioinformática [recurso eletrônico] : uma janela para a biodiversidade /
organização Julia Gabriele Dani, Fernanda Pessi de Abreu, Scheila de
Ávila e Silva. – Caxias do Sul : Educs, 2025.
Dados eletrônicos (1 arquivo).

Apresenta bibliografia.

Vários autores.

Modo de acesso: World Wide Web.

ISBN 978-65-5807-510-3

1. Bioinformática. 2. Biodiversidade. I. Dani, Julia Gabriele. II.
Abreu, Fernanda Pessi de. III. Silva, Scheila de Ávila e.

CDU 2. ed.: 577:004

Índice para o catálogo sistemático:

1. Bioinformática

577:004

2. Biodiversidade

574.1

Catalogação na fonte elaborada pela bibliotecária

Carolina Machado Quadros – CRB 10/2236

Direitos reservados a:



EDUCS – Editora da Universidade de Caxias do Sul

Rua Francisco Getúlio Vargas, 1130 – Bairro Petrópolis – CEP 95070-560 – Caxias do
Sul – RS – Brasil

Ou: Caixa Postal 1352 – CEP 95020-972 – Caxias do Sul – RS – Brasil

Telefone/Telefax: (54) 3218 2100 – Ramais: 2197 e 2281 – DDR (54) 3218 2197

Home Page: www.ucs.br – E-mail: educs@ucs.br

Sumário

Apresentação

6

ECOLOGIA

Interações biológicas em redes: da observação
na natureza ao uso de ferramentas de análises

9

Daniel A. Carvalho, Carolina V. C. Silva & Gabriel G. Barbosa

O Triângulo de Grime como ferramenta
didática em práticas de ecologia vegetal

25

Francieli Peter-da-Silveira & Ana Boeira Porto

CONSERVAÇÃO

Conservação em foco: análises de dados
georreferenciados para proteger a
biodiversidade

45

Fernando Fernandes & Taciane Schröder

Genômica aplicada à conservação de espécies:
decifrando enigmas taxonômicos

58

Bruna Boizonave Andriola, Gabriele Zenato Lazzari, Yago Beux
& Amanda Kessler

Coleções biológicas: os bancos de dados da
biodiversidade

79

Taciane Schröder, Fernando Fernandes & Julia Gabriele Dani

EVOLUÇÃO

Olhando para o presente e desvendando o
passado com macroevolução

89

Patrícia Sperotto & Sandra Reinales

O código de barras da vida: reconhecendo
espécies de fungos a partir do DNA

106

Nicolas do Carmo Regio & Mauro Carpes Westphalen

Apresentação

A bioinformática tem desempenhado um papel crucial na exploração da biodiversidade, ao integrar métodos computacionais avançados com o estudo da vida em sua forma mais fundamental. Por meio da análise de grandes volumes de dados genômicos, transcriptômicos e proteômicos, torna-se possível investigar a diversidade de organismos com uma precisão e escala antes inimagináveis. Com isso, cientistas podem mapear a variabilidade genética de populações, identificar novas espécies e compreender suas interações em diferentes ambientes.

Para decifrar a biodiversidade, é necessária a capacidade da bioinformática de cruzar informações biológicas com dados ambientais, comportamentais e ecológicos. Por meio de algoritmos, é possível traçar as relações evolutivas entre espécies, revelando a complexidade das adaptações e processos evolutivos ao longo do tempo. A bioinformática não apenas facilita a identificação de padrões em grandes conjuntos de dados como também permite previsões sobre a resiliência de espécies e ecossistemas diante de mudanças ambientais e climáticas.

Além disso, essa área da ciência desempenha um papel importante na conservação da biodiversidade. Ao fornecer dados robustos e modelos preditivos, ela apoia a formulação de estratégias de conservação. Isso se torna essencial em um cenário global de perda acelerada de espécies e habitats. O uso dessas ferramentas possibilita a priorização de áreas de preservação, o monitoramento de populações em risco e a avaliação de impactos ambientais de atividades humanas, garantindo uma abordagem mais integrada e eficiente na proteção da biodiversidade.

Entretanto, muitas vezes essas ferramentas não são abordadas em disciplinas de graduação devido ao seu nível de complexidade. O objetivo deste livro foi reunir, de maneira didática e acessível, algumas ferramentas de bioinformática para discentes

da graduação. Cada capítulo traz uma introdução em linguagem didática juntamente com um roteiro para aula prática. O livro dispõe de um repositório no GitHub (<https://github.com/EbookBioinformaticaBiodiversidade>), onde os materiais para realização de cada aula estão organizados por capítulo.

Os capítulos aqui apresentados foram propostos para auxiliar os docentes e discentes a utilizar essas ferramentas que muitas vezes parecem um grande mistério. A bioinformática pode ser vista como uma janela na qual nos apoiamos para compreender melhor a biodiversidade que nos cerca. Esperamos poder ajudar discentes e docentes nesse processo por meio desta obra.

A comissão organizadora

ECOLOGIA



Interações biológicas em redes: da observação na natureza ao uso de ferramentas de análises



Daniel A. Carvalho¹, Carolina V. C. Silva² &
Gabriel G. Barbosa³

¹Programa de Pós-Graduação em Biologia Animal, Universidade Federal do Rio Grande do Sul, Porto Alegre (Brasil)

²Programa de Pós-Graduação em Ecologia, Instituto Nacional de Pesquisas da Amazônia, Manaus (Brasil)

³Programa de Pós-Graduação em Ecologia, Universidade Federal do Rio Grande do Sul, Porto Alegre (Brasil)

1. Introdução

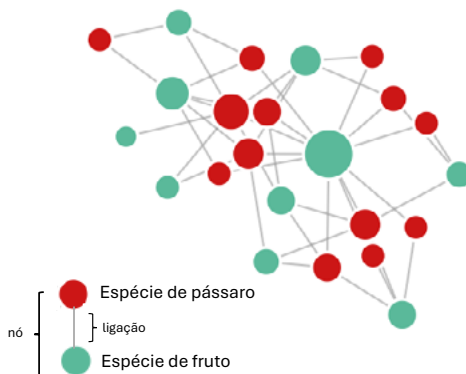
A biodiversidade existente no planeta Terra é influenciada por muitos fatores. Fatores abióticos, por exemplo (condições ambientais), exercem pressões nas quais diferentes espécies podem ser resistentes e resilientes. Nesse contexto, as condições ambientais agem como filtros, “filtrando” quais espécies podem estar presentes, ou não, em determinados ambientes. Entretanto, os fatores bióticos (efeitos dos organismos em um ambiente) também estruturam a biodiversidade nas distintas regiões do globo (Barracough, 2015; Guimarães Júnior *et al.*, 2017; Abdala-Roberts *et al.*, 2019; Cosmo *et al.*, 2023). As interações entre espécies representam bem os fatores bióticos (Margulis; Fester, 1991).

Em ecologia, uma interação biológica é o efeito que um par de organismos (da mesma ou de diferente espécie) tem sobre uma comunidade (conjunto de espécies que ocorrem no mesmo lugar e ao mesmo tempo). As interações ocorrem de diferentes maneiras (Post; Palkovacs, 2009), variando no tempo e no espaço (Poisot *et al.*, 2012; Dáttilo *et al.*, 2013), e ainda podem ocorrer de forma direta (envolvendo contato físico entre os indivíduos; *e.g.* polinização inseto-planta) ou indireta (não envolvendo contato físico, a interação ocorre através de recursos compartilhados, toxinas ou inibidores de

crescimento; *e.g.* quando uma planta secreta uma substância que inibe o crescimento de outras plantas próximas).

Para compreender melhor a complexidade dessas interações e seus impactos ecológicos, é útil visualizar essas relações como uma rede de interações de espécies. Em suma, uma rede de interações é uma representação gráfica, matemática e conceitual. No contexto ecológico, essas redes demonstram como diferentes espécies se relacionam dentro de uma comunidade e como tais interações se estabelecem (Poisot; Stouffer; Gravel, 2015). As redes são formadas graficamente por nós e arestas (ou ligações), em que os nós representam as espécies e as arestas representam as interações entre elas (*e.g.* predação, polinização, competição ou mutualismo) (Figura 1).

Figura 1: Rede de interação entre dispersores de sementes (pássaros) e plantas (frutos) na Inglaterra. Os círculos (nós) vermelhos correspondem às espécies de pássaros; os verdes, às de plantas. As interações são representadas por linhas conectando os pássaros e as plantas. Esta figura foi retirada do site <https://www.web-of-life.es/>, que disponibiliza redes de interação de pesquisadores do mundo todo. Para mais detalhes: Sorensen (1981).

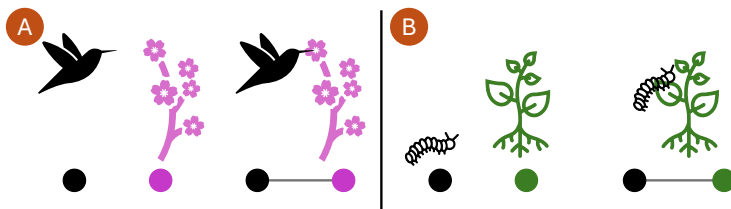


Fonte: imagem gerada pelos autores.

Cada tipo de interação possui suas próprias características e influências sobre a dinâmica e estabilidade do ecossistema (Diamond, 1975; Toju *et al.* 2017). Por exemplo, em uma rede de polinização beija-flor-planta, as arestas podem indicar a frequência

com que o beija-flor visita a planta (Figura 2), enquanto em uma rede de predação lagarta-planta as arestas podem representar a taxa de predação entre a lagarta (predadora) e a planta (presa).

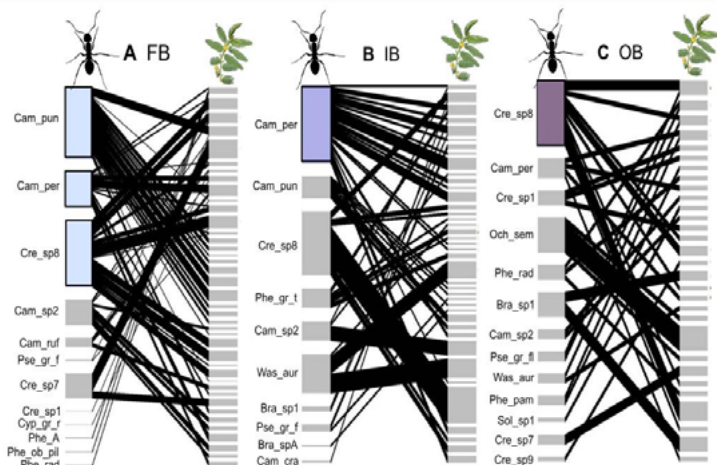
Figura 2: Interação de polinização beija-flor-planta (A) e de predação lagarta-planta (B). Os círculos indicam as espécies; a ligação, a interação entre as espécies.



Fonte: imagem gerada pelos autores.

A análise das redes permite que os(as) pesquisadores(as) visualizem a estrutura das interações. O estudo da estrutura da rede é importante, pois ela influencia na sua estabilidade, ou seja, no quanto é resistente a perturbações (*e.g.* fogo, desmatamento) e extinção de espécies, por exemplo (Lewinsohn *et al.*, 2006). Dessa forma, a estrutura da rede nos ajuda a entender como as espécies estão interconectadas e como mudanças no comportamento de uma única espécie ou do ambiente podem afetar a rede como um todo. Silva *et al.* (2020), por exemplo, investigaram o efeito do fogo na rede de interação formiga-planta (*Chamaecrista repens*) no bioma Pampa. Os autores observaram que, em campos com um mês após a queimada, a quantidade de nectários extraflorais – glândulas secretoras de néctar que atraem formigas, as quais geralmente protegem a planta contra herbivoria (Rico-Gray; Oliveira, 2007) – aumentou, e, em campos há mais de 2 anos sem queimar, a quantidade desse recurso foi bem reduzida. O aumento do recurso permitiu mais espécies de formigas interagindo mais fortemente com a planta, o que pode conferir uma maior efetividade na proteção para a planta. Ou seja, o fogo alterou o habitat, aumentando o recurso disponível, permitindo mais espécies de formiga protegendo a planta.

Figura 3: Redes de interação formiga-planta de indivíduos de *Chamaecrista repens* (direita) e espécies de formigas (esquerda) em campos um mês pós fogo (A), um ano pós fogo (B) e dois anos ou mais sem fogo (C). As linhas representam as interações entre os indivíduos de *C. repens* e as espécies de formigas. Os retângulos representam as espécies de formigas do núcleo central da rede (espécies que interagem fortemente com a planta). Cada cor representa um período pós fogo do campo. Na rede A são três espécies nesse núcleo; ao longo do tempo, reduzem-se a uma. Mais detalhes em Silva *et al.* (2020.)



Fonte: imagem gerada pelos autores.

Para criarmos as redes de interações, podemos utilizar a bioinformática, uma ciência que combina Biologia, Matemática, Ciência da Computação e Estatística, para desenvolver métodos de armazenamento, recuperação e análise de dados biológicos (Hogeweg, 2011). No campo das interações biológicas, essa ciência auxilia na utilização de diferentes softwares e estruturas analíticas para a avaliação do grande volume de dados obtidos a partir de experimentos e observações de campo (Tucker; Duplisea, 2012). Porém, a aplicação da bioinformática nos cursos de Ciências Biológicas ainda é pouco explorada, mesmo que os(as) alunos(as) relatem como essencial o estudo da bioinformática para o avanço da biologia (Marcorin; Pereira, 2022).

Portanto, devido à importância que as interações biológicas têm sobre a dinâmica da biodiversidade, e relação a como é funda-

mental entendermos como as redes de interação estão estruturadas e à baixa aplicação da bioinformática nos cursos de graduação, nossos objetivos neste capítulo são (1) abordar os diferentes tipos de interações (competição, mutualismo, amensalismo, neutralismo, comensalismo e parasitismo) assim como o seu impacto no funcionamento das comunidades e ecossistemas, (2) sugerir uma atividade prática de campo observando as interações biológicas e (3) fornecer ferramentas para a análise de dados de redes de interações.

2. Interações biológicas e sua influência sobre os ecossistemas e a dinâmica de comunidades

Conforme contextualizamos, as interações biológicas são fundamentais para entendermos como os organismos coexistem e influenciam uns aos outros, especificamente no sucesso de uma espécie para utilizar os recursos disponíveis e contribuir para o seu desenvolvimento e reprodução (Townsend; Begon; Harper, 2010). Para medir o efeito das interações sobre as espécies, é necessário compreendermos a intensidade e a frequência com que as interações ocorrem. A intensidade representa a força da interação, enquanto a frequência indica o número de interações em um determinado período. Essa caracterização das interações define os tipos de interações observados a longo prazo, apresentadas a seguir.

2.1. Competição (–/–)

A competição é uma interação ecológica antagônica entre organismos ou espécies, em que o sucesso de um é comprometido pela presença do outro. Isso acontece quando dois ou mais organismos (da mesma ou de diferentes espécies) disputam os mesmos recursos, como espaço e alimento, e esses recursos são insuficientes para atender às necessidades de todos. As espécies que interagem devem ter ganhos negativos sobre o seu sucesso reprodutivo, adaptativo, e na utilização de um recurso. Por exemplo, quando duas espécies competem por um determinado alimento, ambas estarão

gastando energia e tempo para competir, enquanto poderiam obter ganhos superiores caso a outra espécie não estivesse presente.

Este tipo de interação é uma das principais estruturadoras das comunidades biológicas. Sabe-se, por exemplo, que a competição influencia a distribuição das espécies. Espécies que competem diretamente pelos mesmos recursos podem ser excluídas umas das outras, levando à separação espacial ou temporal. Um exemplo clássico é o dos tentilhões de Darwin (Grant; Grant, 2007), aves que apresentam diferentes tamanhos e formas de bico, adaptados para diferentes tipos de alimentação. Inicialmente esses organismos estavam presentes em uma única ilha e compartilhavam de um mesmo tipo de recurso alimentar. Devido à pressão competitiva, muitas espécies se distribuíram ao longo das ilhas e se adaptaram a outras fontes de recursos alimentares. Essa adaptação, para diminuir a sobreposição de recursos alimentares com outras espécies, se deu em parte pela variação das características morfológicas distintas em seus bicos.

2.2. Mutualismo (+/+)

Mutualismo é uma interação ecológica entre duas ou mais espécies em que ambas obtêm ganhos positivos de suas associações. Pode ser diferenciado em dois principais tipos. O mutualismo obrigatório é caracterizado por organismos que vivem em estreita proximidade, geralmente envolvendo trocas essenciais para o seu desenvolvimento e sobrevivência. Um exemplo são os líquens (associações simbióticas entre fungos e algas ou cianobactérias). O mutualismo facultativo, por outro lado, é uma relação benéfica, mas não essencial para a sobrevivência das espécies envolvidas. Um exemplo é a polinização realizada pela grande maioria de espécies de abelhas.

O mutualismo frequentemente promove a coevolução entre espécies, levando a uma maior diversidade genética e ecológica. Na polinização, por exemplo, diferentes espécies de flores e polinizadores, ao longo do tempo evolutivo, desenvolveram adaptações

específicas umas às outras, o que aumentou a variedade de formas, interações mutualísticas (inclusive obrigatórias) e comportamentos no ecossistema atual (Olesen *et al.*, 2007).

2.3. Amensalismo (-/0)

O amensalismo ocorre quando um organismo causa um efeito negativo em outro sem sofrer nenhum efeito em troca. Em outras palavras, um organismo é prejudicado pela presença de outro sem que este seja beneficiado ou prejudicado. Essa interação acontece geralmente através da secreção de substâncias pela espécie inibidora que acaba afetando negativamente a outra espécie, denominada amensal. O exemplo mais conhecido e de grande importância médica é a produção de antibióticos, como a penicilina, que é secretada por fungos do gênero *Penicillium* e inibe a proliferação de bactérias.

2.4. Neutralismo (0/0)

O neutralismo ocorre quando, numa comunidade, a presença de uma espécie não altera ou não tem efeito em outra espécie, embora interajam entre si (Odum; Barrett, 2007). Portanto, as espécies são capazes de coexistir. Pássaros que se alimentam de sementes ou frutos de árvores podem coexistir com herbívoros terrestres (como veados ou coelhos) sem afetar diretamente suas populações. Ambos compartilham o mesmo espaço e recursos alimentares, mas suas interações são limitadas, resultando em um estado de neutralismo.

2.5. Comensalismo (0/+)

Comensalismo é um tipo de interação entre dois organismos na qual uma espécie é beneficiada sem afetar a outra. Um exemplo de comensalismo é a forésia, em que um organismo é transportado por outro, sem nenhuma perda nutritiva ou de desenvolvimento para o hospedeiro (Veiga, 2016). Esse tipo de interação é comum em cupins, cujas espécies fixam-se nas pernas de algumas espécies de besouros para migrar para demais lugares (Błoszyk *et al.*, 2013).

2.6. Parasitismo (+/-)

O parasitismo é um tipo de interação em que um organismo, denominado parasita, vive às custas de outro organismo, chamado hospedeiro. Essa relação é geralmente desigual, beneficiando o parasita enquanto prejudica o hospedeiro. A intensidade dessa interação é determinada pelo impacto no hospedeiro e pelas adaptações do parasita, enquanto a frequência é influenciada pela ocorrência temporal e pela dinâmica populacional dos organismos envolvidos.

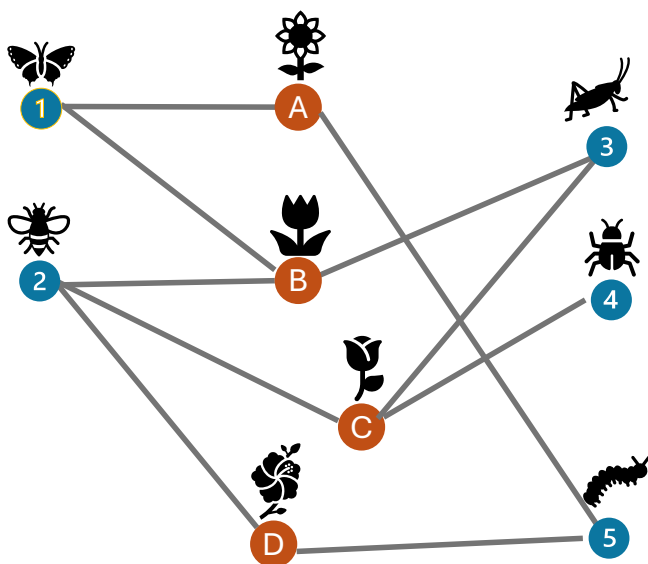
O parasita pode ser classificado como obrigatório, quando depende inteiramente de seu hospedeiro para completar seu ciclo de vida, ou facultativo, quando não possui essa dependência direta. Também podem ser classificados como endoparasitas, quando vivem dentro do corpo de seu hospedeiro (*e.g.* vermes nematódeos e platelmintos) e ectoparasitas quando vivem fixados em alguma parte externa do corpo dos hospedeiros (*e.g.* piolhos, pulgas e carrapatos).

3. Análise de redes de interação e sua importância

A principal forma de representar redes de interações ecológicas é através de grafos, pois nos permitem visualizar a configuração espacial das unidades das interações e os fluxos entre elas. Nesta seção, abordaremos somente o grafo bipartido e as ferramentas de bioinformática que podem ser utilizadas para representar as redes de interações, porém existem distintas maneiras de representação (Bascompte; Jordano, 2014).

No grafo bipartido, temos dois grupos de organismos de cada lado, e as interações ocorrem sempre em diferentes níveis tróficos. Em uma rede bipartida entre insetos que utilizam recursos das plantas (Figura 4), por exemplo, os insetos (1-5) só interagem com as plantas (A-D), não havendo interações planta-planta e inseto-inseto.

Figura 4: Grafo bipartido representando as interações entre insetos polinizadores (1-5) e plantas (A-D). As interações são representadas por arestas que ligam os pares de espécies desses dois grupos.

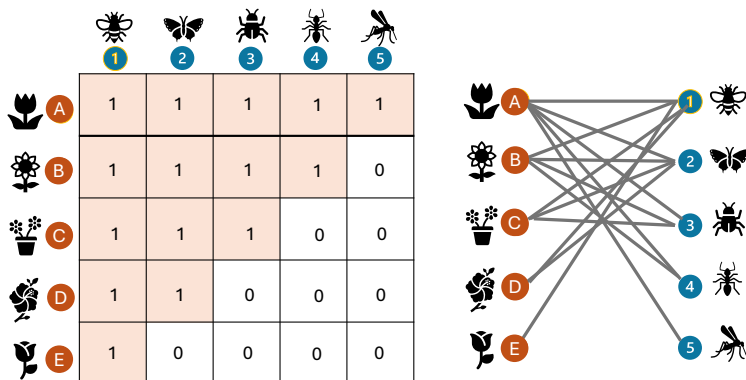


Fonte: imagem gerada pelos autores.

Para além de representarmos as interações através dos grafos bipartidos, também podemos analisar como as redes de interações estão estruturadas. As duas estruturas mais comuns de serem detectadas são o aninhamento e a modularidade (Olesen *et al.*, 2007). O aninhamento ocorre quando espécies especialistas tendem a interagir com subconjuntos das espécies que interagem com espécies mais generalistas (Bascompte *et al.*, 2003). Em outras palavras, as espécies menos conectadas interagem com um subconjunto das espécies com que as mais conectadas interagem (Figura 5). Nesse contexto, ocorre uma alta sobreposição da utilização de recurso que é consumido por diferentes espécies. Isso confere, de acordo com a resistência e resiliência da interação, uma maior robustez da rede, já que espécies generalistas podem manter a funcionalidade

da rede, mesmo que espécies especialistas sejam removidas (Bascompte *et al.*, 2003).

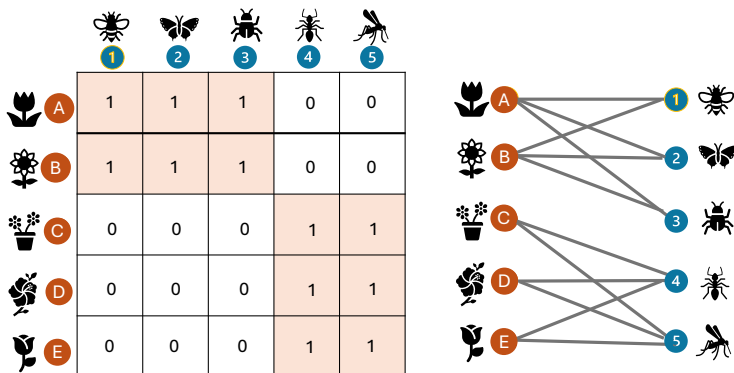
Figura 5: Estrutura aninhada de interações entre plantas (A-E) e insetos visitantes florais (1-5), em que as ligações feitas por espécies com menos parceiros de interação são um subconjunto das ligações feitas por espécies com mais parceiros. Na matriz de dados (esquerda), 1 indica a interação inseto-plantas, e 0 aponta quando não há interação. Na rede (direita), a interação é representada pela linha de ligação entre nós.



Fonte: imagem gerada pelos autores.

Já na modularidade, há a formação de subgrupos de espécies que interagem mais com as espécies do mesmo subgrupo do que com outras espécies da mesma rede (Figura 6), formando módulos (Lewinsohn *et al.*, 2006; Olesen *et al.*, 2007). Em uma rede modular, a sobreposição de interações é grande entre espécies de um mesmo módulo, mas pequena entre espécies de módulos diferentes. Esses subgrupos podem indicar que há uma especialização das espécies e seus recursos (Olesen *et al.*, 2007). Por exemplo, um módulo pode consistir em plantas com flores tubulares e seus polinizadores de língua longa, enquanto outro módulo pode consistir em plantas com flores abertas e polinizadores generalistas.

Figura 6: Estrutura modular das interações entre visitantes florais (1-5) e plantas (A-E). Na matriz (esquerda), as espécies de planta A e B interagem com os insetos 1, 2 e 3. Nas espécies de planta C, D e E, interagem somente com os insetos 4 e 5. O número 1 indica a interação inseto-planta, e 0 quando não há interação. Na rede (direita) é possível visualizar dois grupos separados. A interação representa a ligação entre nós.



Fonte: imagem gerada pelos autores.

A compreensão dos conceitos de modularidade e aninhamento é crucial para entendermos a complexidade das interações ecológicas e suas implicações para a resiliência, estabilidade e funcionalidade dos ecossistemas. Porém, isso também pode ser aplicado a práticas em conservação, gestão de recursos naturais e sustentabilidade (Memmott *et al.*, 2007), pois podemos presumir o efeito no ambiente de uma espécie que possa ser extinta, por exemplo.

4. Atividade prática: explorando as redes de interação inseto-planta

Após aprendermos alguns conceitos de interações biológicas e redes de interação, veremos esses conceitos na prática. Dessa forma, propomos uma atividade cujo objetivo é incentivar os(as) alunos(as) a construir sua própria rede de interação, a partir de observações de interações insetos-planta em uma área específica. Para isso, precisaremos de: caderneta de campo e lápis, câmera fotográfica ou celular (para fotografar as interações), fita métrica

ou barbante (para delimitação das áreas de observação), computador com Excel e o software R instalado. Para instalar o R, acesse: <https://www.r-project.org/>.

4.1. Procedimento metodológico

Selecione um local com diversidade de flores e insetos como o jardim de uma escola, parque, praça ou jardim botânico da sua cidade. Depois, divida a turma em grupos, e cada grupo ficará responsável por um quadrante. No quadrante, faça o levantamento das espécies de plantas com flores e anote. O levantamento pode ser a nível de espécie, mas, caso não seja possível, tire fotos das diferentes plantas e atribua uma identificação (e.g. flor amarela 1).

Durante um período determinado (e.g. uma manhã), os(as) alunos(as) devem observar e registrar as interações dos insetos visitantes florais. Essas observações e registros devem ser realizados para 3-5 indivíduos de planta (de cada espécie escolhida), por um período de 15 minutos. Considere uma visita floral o inseto que entra em contato com alguma estrutura da flor. O registro deve conter: (a) identificação da planta; (b) identificação do inseto (espécie ou ordem); (c) horário da observação; (d) número de vezes que a interação ocorreu.

Depois de ir a campo, passe os dados para o Excel e salve o arquivo no formato “.txt”. A matriz de dados deve conter as espécies de plantas nas linhas, os visitantes florais nas colunas e as células contendo o número de visitas de cada inseto para cada espécie de planta.

Com a matriz pronta, é possível testar a estrutura da sua rede de interação. Para isso, acesse nosso passo a passo (*script*). Guiando-se pelo *script*, será possível observar como as interações estão organizadas (através do gráfico da rede), a sua estrutura (aninhada ou modular) e outras métricas de rede (conectância e especialização). Se você quiser testar as análises antes de ir a campo, disponibilizamos uma matriz de dados de um estudo de po-

linização (Small, 1976), disponível em: <https://drive.google.com/file/d/1vOdil35Uh6TkjXheLEpX66xKvPJJdtbl/view?usp=sharing>.

Após realizar as análises e representar graficamente as redes, debata com os(as) alunos(as) resultados: (1) Qual foi a estrutura da sua rede de interação? (2) Quais espécies de plantas e visitantes florais interagiram com o maior número de espécies? (3) Que importância elas têm para a rede? (4) Na área em que coletou, você esperava encontrar mais ou menos espécies de plantas e insetos? (5) Qual a importância que essas interações inseto-planta que você observou têm no sistema estudado? Essas e outras questões pertinentes ao estudo são muito bem-vindas.

5. Considerações finais

Neste capítulo, exploramos aspectos das interações biológicas através das lentes das redes ecológicas, da observação na natureza a partir de um experimento em campo e do uso de ferramentas da bioinformática para análise de dados biológicos. Ao abordar os diversos tipos de interações entre espécies e seus impactos nos ecossistemas e na dinâmica das comunidades, demonstramos como a bioinformática pode ser uma ferramenta poderosa para entender e visualizar essas complexas relações.

Com a atividade prática proposta neste capítulo, esperamos proporcionar uma oportunidade valiosa para os(as) estudantes aplicarem conceitos teóricos na coleta e análise de dados reais, promovendo uma compreensão mais profunda das interações biológicas. Ao construir suas próprias redes de interação inseto-planta, os(as) estudantes poderão observar diretamente como as espécies interagem e como essas interações se estruturam. Nesse contexto, esperamos que a defasagem, anteriormente citada, acerca da falta da implementação da bioinformática dentro dos cursos de graduação em Ciências Biológicas possa ser, com isso, inicialmente suprida.

Referências

- ABDALA-ROBERTS, L.; PUENTES, A.; FINKE, D. L.; MARQUIS, R. J.; MONTSERRAT, M.; POELMAN, E. H.; BJÖRKMAN, C. Tri-trophic interactions: bridging species, communities and ecosystems. *Ecology Letters*, v. 22, n. 12, p. 2151-2167, 2019.
- BASCOMPTE, J.; JORDANO, P.; MELIÁN, C. J.; OLESEN, J. M. The nested assembly of plant-animal mutualistic networks. *Proceedings of the National Academy of Sciences*, v. 100, n. 16, p. 9.383-9.387, 2003.
- BASCOMPTE, J.; JORDANO, P. *Mutualistic networks*. Princeton: Princeton University Press, 2014. (Monographs in Population Biology, n. 53).
- BASCOMPTE, J.; JORDANO, P. Plant-animal mutualistic networks: the architecture of biodiversity. *Annual Review of Ecology, Evolution, and Systematics*, v. 38, p. 567-593, 2002.
- BARRACLOUGH, T. G. How do species interactions affect evolutionary dynamics across whole communities? *Annual Review of Ecology, Evolution, and Systematics*, v. 46, p. 25-48, 2015.
- BŁOSZYK, J.; GUTOWSKI, J.; GWIAZDOWICZ, D.; MADRA, A. Phoresy of Trichouropoda shcherbakae Hirschmann, 1972 (Acari: Mesostigmata) on beetles of the genus Tetropium Kirby, 1837 (Coleoptera: Cerambycidae) in Białowieża Forest, Poland. *Open Life Sciences*, v. 8, n. 10, p. 986-992, 2013.
- COSMO, L. G.; ASSIS, A. P. A.; AGUIAR, M. A.; PIRES, M. M.; VALIDO, A.; JORDANO, P.; GUIMARÃES JR, P. R. Indirect effects shape species fitness in coevolved mutualistic networks. *Nature*, v. 619, n. 7971, p. 788-792, 2023.
- DÁTTILO, W.; GUIMARAES JR, P. R.; IZZO, T. J. Spatial structure of ant-plant mutualistic networks. *Oikos*, v. 122, n. 11, p. 1.643-1.648, 2013.
- DIAMOND, J. M. Assembly of species communities. In: CODY, M.; DIAMOND, J. M. (eds.). *Ecology and Evolution of communities*. Cambridge: Harvard University Press, 1975. 342-444.
- GRANT, P. R.; GRANT, B. R. *How and why species multiply: the radiation of Darwin's finches*. Princeton University Press, 2007.
- GAUSE, G. F.; NASTUKOVA, O. K.; ALPATOV, W. W. The Influence of Biologically Conditioned Media on the Growth of a Mixed Population of *Paramecium caudatum* and *P. aurelia*. *Journal of Animal Ecology*, v. 3, n. 2, p. 222-230, 1934.

- GUIMARAES JÚNIOR, P. R.; PIRES, M. M.; JORDANO, P.; BASCOMPTE, J.; THOMPSON, J. N. Indirect effects drive coevolution in mutualistic networks. *Nature*, v. 550, n. 7677, p. 511-514, 2017.
- HOGEWEG, P. The roots of bioinformatics in theoretical biology. *PLoS computational biology*, v. 7, n. 3, e1002021, 2011.
- INGS, T. C. *et al.* Ecological networks – beyond food webs. *J. Anim. Ecol.*, v. 78, p. 253-269, 2009.
- LEWINSOHN, T. M.; INÁCIO PRADO, P.; JORDANO, P.; BASCOMPTE, J.; M. OLESEN, J. Structure in plant-animal interaction assemblages. *Oikos*, v. 113, n. 1, p. 174-184, 2006.
- MARCORIN, L.; PEREIRA, M. A. A. Mapeamento de competências em bioinformática: panorama geral dos profissionais das áreas ômicas no Brasil. In: WORKTEC – WORKSHOP DE TECNOLOGIA DA FATEC, 5., 2022. Ribeirão Preto. *Anais [...].* Ribeirão Preto: FATEC, 2022.
- MARGULIS, Lynn; FESTER, René (ed.). *Symbiosis as a source of evolutionary innovation: speciation and morphogenesis*. New York: MIT press, 1991.
- MAZGAJSKI, T. D. Competition for nest sites between the starling *Sturnus vulgaris* and other cavity-nesters – study in forest park. *Acta Ornithologica*, Warsaw, v. 35, p. 103-107, 2000.
- MEMMOTT, J. *et al.* The conservation of ecological interactions. *Insect conservation biology*, p. 226-244, 2007.
- ODUM, E. P.; BARRETT G. W. *Fundamentos de Ecologia*. 5. ed. Tradução: Pégasus Sistemas e Soluções. São Paulo: Thomson Learning, 2007. 282-335.
- OLESEN, J. M.; BASCOMPTE, J.; DUPONT, Y. L.; JORDANO, P. The modularity of pollination networks. *Proc. Natl. Acad. Sci.*, v. 104, p. 1.9891-1.9896, 2007.
- PETERS, R. S. *et al.* Evolutionary history of the Hymenoptera. *Current Biology*, v. 27, n. 7, p. 1.013-1.018, 2017.
- POISOT, T. *et al.* The dissimilarity of species interaction networks. *Ecol. Lett.*, v. 15, p. 1.353-1.361, 2012.
- POISOT, T.; STOUFFER, D. B.; GRAVEL, D. Beyond species: why ecological interaction networks vary through space and time. *Oikos*, v. 124, n. 3, p. 243-251, 2015.
- POST, D. M.; PALKOVACS, E. P. Eco-evolutionary feedbacks in community and ecosystem ecology: interactions between the ecological theatre and the evolutionary play. *Philosophical*

Transactions of the Royal Society B: Biological Sciences, v. 364, n. 1.523, p. 1.629-1.640, 2009.

RICO-GRAY, V.; OLIVEIRA P. S. *The ecology and evolution of ant-plant interactions*. 1. ed. Chicago: University of Chicago Press, 2007.

SILVA, C. V. C. *et al.* Effects of time-since-fire on ant-plant interactions in southern Brazilian grasslands. *Ecol. Indic.*, v. 112, 106094, 2020.

SMALL, E. Insect pollinators of the Mer Bleue peat bog of Ottawa. *Canadian Field Naturalist*, v. 90, p. 22-28, 1976.

TOJU, H. *et al.* Species-rich networks and eco-evolutionary synthesis at the metacommunity level. *Nat. Ecol. Evol.*, v. 1, 0024, 2017.

TOWNSEND, C. R.; BEGON, M.; HARPER, J. L. *Fundamentos em Ecologia*. 3. ed. Porto Alegre: Artmed, 2010. pp. 211-212.

TUCKER, A.; DUPLISEA, D. Bioinformatics tools in predictive ecology: applications to fisheries. *Philosophical Transactions of the Royal Society B: Biological Sciences*, v. 367, n. 1.586, p. 279-290, 2012.

VAN DER KOOL, C. J.; OLLERTON, J. The origins of flowering plants and pollinators. *Science*, v. 368, n. 6.497, p. 1.306-1.308, 2020.

VEIGA, J. P. Commensalism, Amensalism, and Synnecrosis. In: KLIMAN, R. M. (ed.), *Encyclopedia of Evolutionary Biology*. Oxford: Academic Press, 2016. vol. 1, pp. 322-328.

O Triângulo de Grime como ferramenta didática em práticas de ecologia vegetal



Francieli Peter-da-Silveira¹ & Ana Boeira Porto¹

¹Programa de Pós-graduação em Botânica, Universidade Federal do Rio Grande do Sul, Porto Alegre (Brasil)

1. Contextualização

Ao analisar e correlacionar a dinâmica das relações das plantas entre si, a disponibilidade de recursos, a presença de herbívoros e/ou patógenos no ambiente, os efeitos climáticos extremos e a interferência humana, o ecólogo inglês John Philip Grime propôs, em 1974, um sistema de classificação de estratégias ecológicas ainda hoje amplamente utilizado, denominado Triângulo CSR, ou Triângulo de Grime. Esse modelo estabelece três estratégias primárias que representam combinações de características viáveis selecionadas naturalmente sob condições de competição, limitação abiótica ao crescimento (estresse) ou destruição periódica da biomassa (distúrbio) (Pierce *et al.*, 2017).

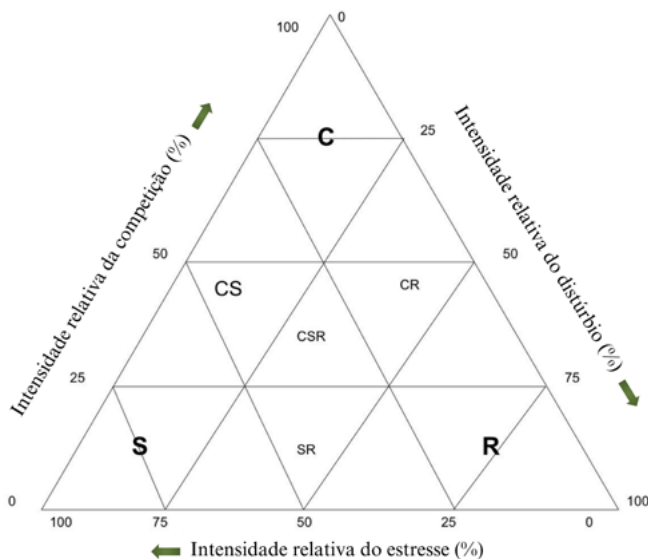
O modelo define como Competidoras (C) plantas que sobrevivem em habitats relativamente estáveis, com baixo nível de estresse e distúrbio, em que a ampla disponibilidade de recursos favorece o crescimento vegetativo e a dominância na comunidade. Em contrapartida, plantas que prosperam em ambientes estressantes, com baixa disponibilidade de luz e nutrientes, por exemplo, tendem a apresentar estratégia conservativa dos recursos com crescimento lento, sendo denominadas Tolerantes ao Estresse (S). Já as plantas Ruderais (R) habitam ambientes onde a disponibilidade de recursos é alta e os distúrbios são frequentes, têm vida curta e investem na produção de propágulos que possam regenerar a população, dado

que a destruição da biomassa é iminente (Grime, 2001; Meindl; Natt; Velardi, 2021).

Originalmente, Grime elaborou um modelo triangular no qual os ângulos do triângulo indicam valores absolutos para cada estratégia CSR. Esse modelo original, posteriormente, foi incrementado observando-se a sobreposição de estratégias e níveis intermediários entre competição, estresse e distúrbio – CR, CSR, CS, SR (Figura 1). O Triângulo de Grime passou a ser empregado em pesquisas de ecologia vegetal, e Pierce *et al.* (2017) demonstraram que as estratégias CSR podem ser avaliadas utilizando apenas três atributos foliares: LA (área foliar, em inglês *leaf area*); SLA (área foliar específica, em inglês *specific leaf area*); e LDMC (teor de matéria seca foliar, em inglês *leaf dry matter content*). O atributo LA, definido como a área unilateral (mm^2) de uma folha individual, é relacionado com a capacidade das plantas de interceptar luz e, assim, apresentar maior produção de biomassa e aptidão competitiva. A SLA é a área unilateral de uma folha fresca (mm^2) dividida por sua massa seca (mg), geralmente positivamente relacionada com a taxa de crescimento relativo potencial e negativamente com a longevidade da folha. E LDMC compreende a massa seca (mg) de uma folha dividida por sua massa fresca (g), estando inversamente relacionado ao SLA, e indica a densidade da folha e sua capacidade de retenção de água e nutrientes. Folhas com alto LDMC são geralmente mais espessas, duráveis e tolerantes a estresse, como seca ou baixa luminosidade, e tendem a ter uma vida útil mais longa e menor taxa de crescimento (Gaudichet *et al.*, 2024; Pérez-Harguindeguy *et al.*, 2013; Reich, 2014); enquanto valores baixos de LDMC estão associados a uma alta taxa de crescimento típica da estratégia ruderal, e valores baixos de SLA são indicativos de adaptação ao estresse. O equilíbrio SLA/LDMC representa o espectro econômico foliar, que é um *trade-off* entre o funcionamento foliar aquisitivo e conservador. Esses dois aspectos permitem posicionar uma espécie dentro de um espaço multivariado em relação a outras espécies e, ao mesmo tempo, atribuir-lhe classes de estratégias ecológicas

distintas (Pierce *et al.*, 2013). Em geral, plantas que otimizam a captura de recursos como luz, água e nutrientes, apresentando folhas com alta SLA e baixa LDMC, facilitando o rápido crescimento. Plantas S adaptam-se a ambientes com recursos escassos e condições adversas, como seca ou solos pobres, apresentando folhas com baixa SLA e alta LDMC, tornando-as mais duráveis e eficientes na retenção de água e nutrientes. Plantas R colonizam rapidamente ambientes perturbados, apresentando características que permitem rápida reprodução e crescimento, com variações nas características foliares (SLA e LDMC) (Pierce *et al.*, 2013).

Figura 1: Esquema ilustrando o Triângulo de Grime e a relação entre a intensidade relativa da competição, do distúrbio e do estresse com as estratégias ecológicas CSR.



Fonte: adaptada de Meindl, Natt e Velardi (2021).

Nos materiais suplementares do artigo de Pierce *et al.* (2017), os autores disponibilizam uma planilha editável, denominada *Strategy*, que habilita pesquisadoras e pesquisadores a avaliarem, de forma prática, estratégias ecológicas no modelo CSR em qualquer

comunidade vegetal. Algumas bases de dados *online*, tais como TRY Plant, Trait Database ([2024]), reúnem e disponibilizam uma ampla gama de atributos funcionais incluindo LA, SLA e LDMC. Dessa forma, utilizando dados obtidos a partir da literatura e a planilha *Stratefy*, é possível implementar como prática em ecologia vegetal a construção e análise do Triângulo de Grime. O método permite reconhecer um *continuum* de estratégias apresentadas pelas espécies, e a partir disso pode-se investigar seus processos evolutivos, como elas se relacionam com a estrutura e a composição das comunidades, como alocam seus recursos para sobrevivência e perpetuação nestes ecossistemas e, finalmente, como compõem e influenciam os processos ecossistêmicos (Grime, 2001). Zanzottera *et al.* (2020), por exemplo, analisaram a nível da comunidade como as estratégias ecológicas podem moldar a estrutura do habitat ao longo de gradientes de sucessão. Já Pierce *et al.* (2022) avaliaram como as estratégias ecológicas das plantas estão associadas às suas compensações anatômicas foliares. E Guo *et al.* (2018) utilizaram o método CSR para explicar os mecanismos de naturalização de espécies em todo o mundo.

Com base no exposto, o presente capítulo apresenta uma proposta de atividade que inclui desde a coleta e a medição de atributos funcionais até a construção do Triângulo de Grime. Esquematizamos as etapas a serem desenvolvidas com base no protocolo proposto por Pérez-Harguindeguy *et al.* (2013) e na planilha *Stratefy* de Pierce *et al.* (2017). O conjunto de dados utilizados aqui para exemplificação é um recorte de uma pesquisa realizada pelas autoras com dados ainda não publicados. A pesquisa em questão tinha como objetivo descrever as estratégias ecológicas apresentadas por plantas aquáticas ao longo de um gradiente de inundação em um complexo de áreas úmidas no Sul do Brasil. Ao final, são apresentados alguns resultados encontrados e uma breve interpretação destes.

2. Objetivo

A atividade visa ampliar o repertório dos docentes para a elaboração de aulas de ecologia vegetal que integrem teoria e prática. Para os discentes, a proposta busca proporcionar uma experiência botânica completa, combinando a coleta de dados em campo com análises subsequentes. Esta prática incentiva a análise de estratégias ecológicas, conforme o modelo CSR (Grime, 2001), observadas em comunidades vegetais.

3. Descrição da atividade

A presente atividade tem como objetivo classificar espécies vegetais quanto às suas estratégias ecológicas (CSR) a partir de atributos foliares (LA, SLA e LDMC). Está dividida em quatro momentos: seleção de espécies e indivíduos; coleta de estruturas foliares; medição dos atributos foliares; e análise de dados.

3.1. Seleção de espécies e indivíduos

O primeiro passo para avaliação das estratégias ecológicas em uma comunidade é escolher quais espécies terão seus atributos foliares mensurados, uma vez que avaliar todas as espécies inventariadas pode demandar muito tempo e esforço. A partir dos dados obtidos no levantamento fitossociológico, seja ele realizado em formação campestre (valores de cobertura, como estimativas de cobertura) ou florestal (valores de contagem, como número de indivíduos), selecione para avaliação as espécies que coletivamente compõem cerca de 80% da abundância relativa cumulativa. Caso esse valor abranja um número muito alto de espécies, alternativas podem ser adotadas, como análise de espécies indicadoras. A seleção das espécies pode também ser direcionada dependendo do escopo da pesquisa e de perguntas específicas que você busque responder.

Para a seleção de indivíduos de cada espécie eleita, alguns critérios devem ser observados. Os indivíduos devem estar: reprodutivamente maduros; aparentemente saudáveis (e.g., sem sinais

de herbivoria e patógenos, a menos que seja parte do objetivo do estudo); e livres de sombreamento. Importante destacar que, tratando-se de plantas clonais, com indivíduos modulares, pode não ser simples distinguir limites individuais. Nesse caso, considere um *ramet* como a unidade básica, isto é, um pequeno ramo acima do solo que está enraizado de forma individual. Escolha indivíduos ou *ramets* de forma aleatória, por meio de transectos ou, ainda, por parcelas. Isso evita que os indivíduos sejam coletados em um único ponto e possibilita coletas mais representativas da comunidade. Tratando-se dos atributos foliares LA, SLA e LDMC, o número de indivíduos selecionados deve ser no mínimo cinco e preferencialmente dez.

3.2. Coleta de estruturas foliares

Caso você tenha optado por amostrar cinco indivíduos por espécie, o número ideal de folhas a serem avaliadas será cinco por indivíduo. Caso tenham sido amostrados dez indivíduos por espécie, quatro folhas para cada indivíduo serão suficientes.

Para evitar que ocorra dessecação, as folhas não devem ser coletadas isoladamente. O recomendado é que elas sejam coletadas junto aos seus ramos e que o destaque seja feito apenas no momento do processamento. Após a coleta, os ramos devem ser embrulhados em papel úmido e guardados em sacos plásticos lacrados. Uma dica é expirar no interior do saco antes de lacrá-lo para aumentar a concentração de CO₂ e umidade e, com isso, diminuir a perda de água pela transpiração. As amostras devem ser armazenadas em uma caixa térmica ou geladeira, permanecendo resfriadas e no escuro. Devem ser processadas o quanto antes possível, de preferência dentro de 24 horas. Mesmo com esses cuidados, no momento do processamento você poderá se deparar com algumas amostras danificadas. Assim, o ideal é coletar amostras extras para evitar que se tenha que retornar ao local de coleta para obter novas.

3.3. Medição dos atributos foliares

Nesta etapa serão necessários os itens listados a seguir: balança de precisão; placa de Petri; pinça; sacos de papel; estufa; fita crepe; computador; *scanner* ou impressora com a função de digitalização. Adicionalmente, para alguns casos, pode ser pertinente realizar a digitalização utilizando uma prancha gabaritada em que seja possível ver a escala em milímetros ou centímetros para assim relacioná-la com a escala real no ImageJ.

3.2.1. Pesagem da massa das estruturas foliares

Vamos começar com as folhas coletadas: escolha qual espécie você começará a analisar e selecione um dos indivíduos coletados. Acomode as folhas em uma placa de Petri para a pesagem das folhas frescas. É imprescindível que a numeração dos indivíduos de cada espécie continue a mesma até o final de todos os passos de pesagem e medição, ou seja, o indivíduo 1 da espécie x será sempre o mesmo. Importante ressaltar que a balança de precisão deve estar calibrada e o peso da placa deve ser desconsiderado antes da pesagem das folhas. Anote o valor do peso fresco, que vai ser utilizado posteriormente para o cálculo de LDMC. Após a pesagem, coloque as folhas no saco de papel, com a identificação numérica do indivíduo, e reserve. Repita esse passo para cada um dos indivíduos das espécies coletadas.

3.2.2 Digitalização

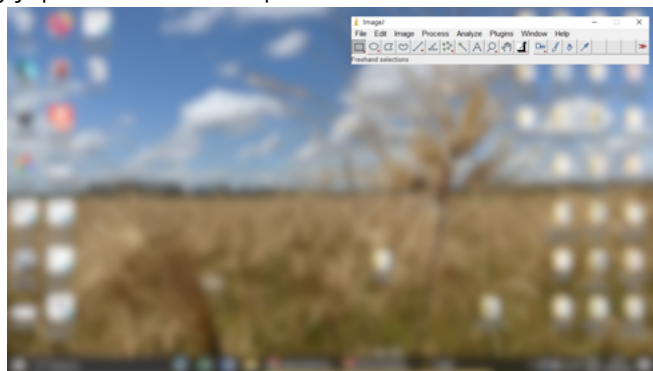
Para esta etapa é necessária a utilização de *scanner* (ou impressora com a função de digitalização), computador e fita crepe. Após a pesagem de todas as folhas de todos os indivíduos, selecione uma espécie para iniciar essa parte do processo. Com o aplicativo do *scanner* aberto no computador, encontre a função de digitalização. Corte pequenos pedaços de fita crepe e os utilize, se necessário, para acomodar as folhas na prancha de digitalização do *scanner*. A maioria das folhas não é perfeitamente plana, e a fita garante que elas fiquem bem dispostas no momento da

digitalização. Importante: as folhas não podem ficar sobrepostas! A sobreposição irá mascarar o valor final da área foliar. Não há a necessidade de preencher os espaços vazios do interior das folhas, o que pode ocorrer quando houver a presença de pequenos buracos causados por herbívoros, por exemplo. Antes da digitalização, confira nas configurações de sua máquina a resolução da imagem que será gerada, ela deverá ter, no mínimo, 150 dpi (pontos por polegadas, do inglês *dots per inch*). Escaneie a imagem em escala de cinza, pois isso reduz a necessidade de convertê-la antes de medir. O documento gerado pela digitalização deverá estar em formato de imagem (PNG, TIFF).

3.2.3 ImageJ

O ImageJ é um aplicativo gratuito, sua instalação é simples e o *download* pode ser feito através do *link* disponível na seção “Para saber mais” deste capítulo. Antes do *download*, você deve escolher o sistema operacional do computador que você realizará as análises (Windows, Linux ou Mac OS X). Após a escolha, será baixada uma pasta zipada que contém alguns arquivos, entre eles, o aplicativo intitulado ImageJ.exe. Com um clique duplo em cima do aplicativo, você abrirá uma pequena aba em sua janela de trabalho, conforme Figura 2.

Figura 2: Exemplo de uma área de trabalho como a janela do aplicativo ImageJ após instalado e aberto para uso.

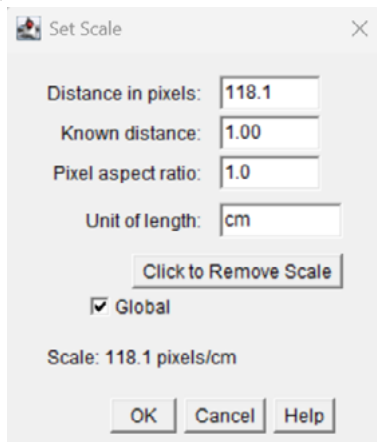


Fonte: imagem gerada pelas autoras.

Antes de iniciar a análise da imagem com o ImageJ, vamos configurar a escala. Contextualizando: dpi é a densidade de pixels por polegada (*inch*) em uma única dimensão. Assim, cada resolução de imagem (150, 300 dpi) possui um número diferente de pixels por unidade de medida. Como a unidade de interesse é em centímetros, o ajuste deve ser feito da seguinte forma:

- 1) divida a resolução da sua imagem por 2,54 (1 inch = 2,54 cm), ou seja, o dpi que você tem por 2,54 (ex.: $300 \div 2,54 = 118,11$);
- 2) no ImageJ, carregue a imagem que você vai analisar clicando em *File* → *Open* e selecione a imagem que será analisada;
- 3) depois, clique em *Analyze* → *Set scale* e preencha conforme a Figura 3.
- 4) a opção *Global* deve estar selecionada para que em a cada imagem aberta a escala seja mantida, cada vez que você fechar o programa esse passo deverá ser repetido (quando uma segunda imagem for aberta, uma caixa de diálogo aparecerá a respeito do *Global scale*, desmarque a primeira opção e marque a segunda, assim a escala será mantida e não aparecerá mais a caixa cada vez que uma nova imagem for aberta);

Figura 3: Demonstrativo de como deve ser feito o preenchimento da janela *Set scale*.

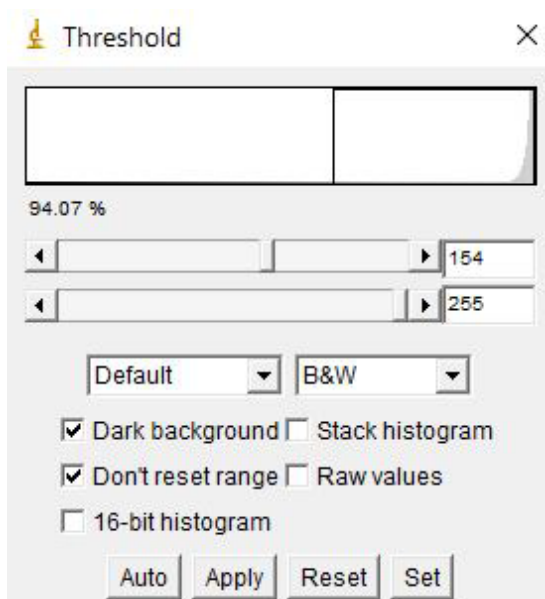


Fonte: imagem gerada pelas autoras.

Com a escala configurada, vamos iniciar nossas análises!

- 1) Clique em *Image* → *Adjust* → *Threshold* → *B&W* → *Apply*, isso informa ao aplicativo que a imagem está em preto e branco (black and white – B&W). Feche a janela (Figura 4).

Figura 4: Demonstrativo da janela *Threshold* para a configuração inicial do ImageJ.



Fonte: imagem gerada pelas autoras.

- 2) No menu *Analyze* → *Set Measurements* estão as funções disponíveis para análise. Como exemplo, selecionamos *Area*, mas você pode selecionar o que for de interesse para a sua análise. Explicações sobre cada função estão disponíveis no *Guia de uso do aplicativo*, a partir da página 136, cujo link está disponível na seção “Para saber mais” deste capítulo. Após selecionar as funções, clique em OK. Uma vez selecionadas, as funções se mantêm para futuras análises com outras imagens, até o programa ser reaberto ou você alterar.

- 3) Clique novamente no menu *Analyze*, depois em *Tools ROI Manager* (ROI = *Regions Of Interest*), mantenha aberta a janela que surgiu.
- 4) Selecione a ferramenta *Wand (trancing) tool* localizada na barra de ferramentas do ImageJ logo abaixo do menu *Analyze* (Figura 5).

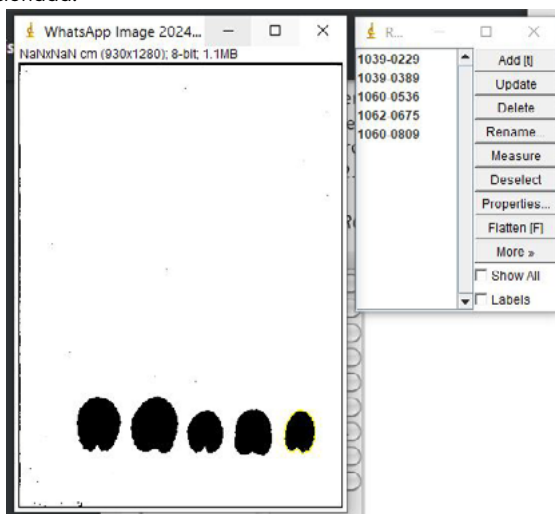
Figura 5: Demonstrativo do ícone da ferramenta *Wand (trancing) tool* na barra de ferramentas do ImageJ.



Fonte: imagem gerada pelas autoras.

- 5) Clique em uma folha e ela será selecionada e contornada em amarelo. Na janela *ROI Manager*, clique em *Add* ou pressione a letra T do teclado para adicionar os valores calculados. Repita o processo para cada folha da imagem (Figura 6).

Figura 6: Demonstrativo da janela *ROI Manager* com os valores de cada folha selecionada.



Fonte: imagem gerada pelas autoras.

- 6) Na janela *ROI Manager*, clique em *Measure*.
- 7) Uma janela *Results* aparecerá com os valores de cada folha. Salve os *Results* clicando em *File* → *Save as*. Dê o mesmo nome que você denominou a imagem. O arquivo salvo será uma planilha Excel. Esses dados serão utilizados para a obtenção dos valores de área em cm². Também é possível copiar os dados diretamente dessa janela para uma planilha Excel. Na escolha dessa última opção, é necessário observar o tipo de separador decimal do seu Excel: se ponto ou vírgula. ImageJ utiliza o ponto como separador decimal, diferente dos pacotes brasileiros de Excel, que têm a vírgula como padrão. É preciso atentar a isso, caso contrário poderá haver problemas na leitura correta dos valores.

Após a análise de todas as coletas, deposite as folhas em sacos de papel identificados com o nome da espécie e o numeral correspondente de cada indivíduo. Os sacos contendo as folhas deverão ser colocados em estufa. É recomendado que a secagem seja em 70°C por pelo menos 72 horas, ou a 80°C por 48 horas. Para a pesagem das folhas secas, recomenda-se utilizar uma pinça para remover as folhas de dentro do saco de papel e acomodá-las na balança. O auxílio da pinça é importante para que a umidade e a gordura presentes nos dedos da mão não entrem em contato com a folha seca, evitando, assim, a interferência no peso seco final.

3.3. Análise de dados

Para esta etapa será necessária a instalação do programa R Studio. Também será necessário o *download* da planilha *Stratefy* (CSR), disponível como apêndice do artigo de Pierce *et al.* (2017). Os *links* para *download* estão na seção “Para saber mais” deste capítulo.

Na planilha *Stratefy* serão inseridos os valores médios dos atributos funcionais para cada espécie que estamos avaliando. A partir dos valores obtidos na janela *Results* no ImageJ, vamos calcular a média da área foliar de cada indivíduo das espécies sele-

cionadas. Assim, para cada indivíduo você deverá somar todos os valores de área foliar encontrados pelo ImageJ e dividir pelo número total de folhas coletadas e mensuradas para o indivíduo. Repita esse cálculo para cada indivíduo das espécies que estão sendo analisadas. Ao final você terá cinco valores médios, se coletou cinco indivíduos, por exemplo. Para calcular o valor médio da espécie basta somar essas médias e dividir pelo número de indivíduos coletados – nesse caso, cinco. A mesma lógica deve ser aplicada para os valores de peso seco e fresco.

3.3.1 Cálculos para obtenção dos valores CSR

Nesta etapa você irá inserir os valores médios de cada espécie, calculados anteriormente, para área foliar, peso fresco e peso seco na planilha *Stratefy*, ilustrada na Figura 7. Insira seus dados nas colunas D, E e F. Dados de área foliar serão inseridos na coluna D, “LA (*leaf area*)”, e deverão estar na unidade de medida de milímetros quadrados (mm²). Na coluna E, “LFW (*leaf fresh weight*)”, serão inseridos valores de peso fresco em miligrama (mg). Na coluna F, “LDW (*leaf dry weight*)”, serão inseridos valores de peso seco em miligrama (mg). Atente para a necessidade de conversão de unidades de medida a partir dos dados originais!


A partir dos valores inseridos nas colunas D, E e F, serão gerados os resultados nas colunas seguintes (G em diante). Note que as estratégias ecológicas para cada espécie podem ser observadas na coluna Q (*Strategy class*). Nas colunas M, N e O estão os valores em porcentagem de C, S e R, respectivamente. Esses valores deverão ser copiados para uma tabela à parte, conforme explicado no item seguinte. Nas colunas J, K e L estão os valores dos atributos LA, LDMC e SLA, utilizados para o cálculo das estratégias ecológicas. Caso você já possua esses valores de atributos, obtidos da literatura ou de outra base de dados, é possível colocá-los diretamente na tabela e, dessa forma, obter os valores de CSR a partir deles. Nesse caso, atente para o fato de que os valores de LDMC devem

ser inseridos em porcentagem. Para convertê-los de mg/g para porcentagem, multiplique-os por mil e divida por cem.

Exemplo: LD₅₀ de 281,3 mg/g.

Sabemos que 1 grama = 1.000 miligramas. Então fazemos: $281,3 \text{ mg}/1.000 \text{ mg} = 0,2813$. Agora multiplicamos por 100 para transformar em porcentagem: $0,2813 \times 100 = 28,13\%$. Resultado: LDMC = 28,13%. Ou seja, 28,13% da massa fresca da folha é matéria seca, e o restante (71,87%) é água.

Figura 7: Captura de tela demonstrativa da planilha *Strategy*.

<div> Stratefy <small>Stratefy v1.0 (4/16/2024) (help on how to use Stratefy is available online) it is a strategic management tool.</small></div>			data input			calculated traits			traits for CSR analysis			CSR analysis output					for plotting tree past values from		
HOW TO Stratefy? <small>Stratefy uses the following input data: LA, LFW and LDW and input these into the equations (1), (2), (3) and (4), LDMC, and LA, LFW and LDW and input these into the equations (5), (6), (7) and (8).</small>			Paste or input your trait measurements here (LA, LFW, LDW) (A, B, C)			Paste or input your trait measurements here (LA, LFW, LDW) (A, B, C)			Paste or input your trait measurements here (LA, LFW, LDW) (A, B, C)			CSR analysis output					for plotting tree past values from		
			LA LFW LDW			LSI LFW LMA			LA LDMC SLC			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output					for plotting tree past values from		
			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			leaf area (mm²) leaf length (mm) leaf width (mm)			CSR analysis output							

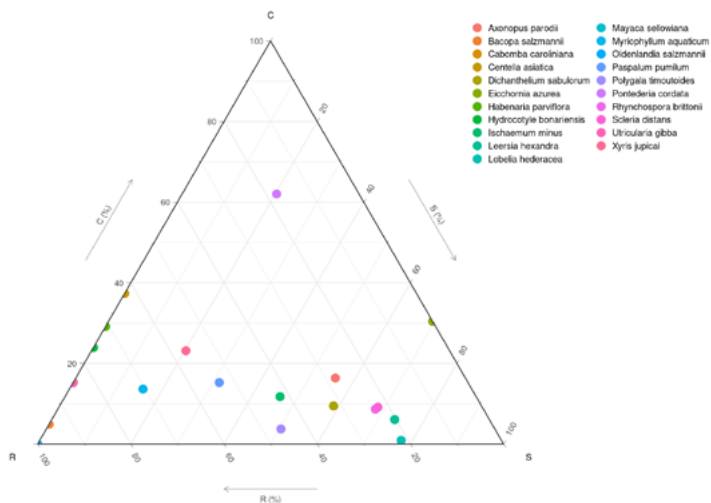
Fonte: Pierce *et al.* (2017).

3.3.2. Construção do gráfico ternário CSR

Com os valores de C, S e R obtidos na planilha *StrateFy* (Figura 7 – colunas M, N e O), você deverá elaborar uma tabela de dados para dar entrada no R. A tabela deve conter quatro colunas, sendo que a primeira deverá indicar o nome das espécies e as demais os valores de CSR calculados a partir da planilha *StrateFy*.

O *script* para construção do gráfico ternário (Triângulo de Grime) demonstrado na Figura 8 está disponível no GitHub (<https://github.com/EbookBioinformaticaBiodiversidade/Cap2>). Nele constam comandos que designam os pacotes utilizados, a abertura dos dados, a confecção e a plotagem do gráfico. O gráfico gerado está em sua forma mais básica e pode ser modificado utilizando outros pacotes e comandos – para alteração de cores e tamanho da fonte, por exemplo.

Figura 8: Triângulo de Grime com a distribuição de estratégias CSR por espécie.



Fonte: imagem gerada pelas autoras.

No gráfico acima, podemos observar que as plantas aquáticas apresentaram um amplo espectro de estratégias ecológicas ao longo do gradiente de inundação estudado. Observamos estratégias primárias – quando os valores de C, S ou R são absolutos – e secundárias – quando os valores de CSR são intermediários. As espécies indicadas diretamente sob o vértice R apresentam valor de R absoluto e se caracterizam exclusivamente como ruderais (*Mayaca sellowiana* e *Oldenlandia salzmännii*). Já as espécies indicadas sob a linha C-R apresentam valor de S nulo e são ruderais competidoras (*Bacopa salzmännii*, *Cabomba caroliniana*, *Centella asiatica*, *Habenaria parviflora*, *Hydrocotyle bonariensis* e *Utricularia gibba*). No lado oposto, a espécie indicada sobre a linha C-S apresenta valor de R igual a zero e é, portanto, competidora tolerante ao estresse (*Eichhornia azurea*). As demais espécies distribuídas no interior do triângulo apresentam valores de C, S e R variados, ou seja, características adaptativas à competição, ao estresse e ao distúrbio em níveis maiores ou menores. Uma única espécie está próxima do vértice C (*Pontede-*

ria cordata), as demais apresentam valores de R e S mais altos. Assim, podemos concluir que o ambiente dinâmico da interface terra-água do gradiente de inundação leva à prevalência de plantas com características ruderais e tolerantes ao estresse, mas mantém uma diversidade de tipos de estratégia.

4. Considerações finais

Esperamos que este material possa enriquecer a prática de docentes em ecologia vegetal e servir como um roteiro de trabalho para estudantes. Ao executar as etapas de pesquisa, cada estudante terá uma vivência da iniciação científica completa, produzindo e interpretando resultados de seu trabalho. É importante que os resultados encontrados sejam amplamente discutidos entre docentes e discentes para que as aplicações do método se conectem à prática. Por exemplo, inda na temática de gradiente de inundação, Albuquerque e Rodrigues-Filho (2019), por exemplo, notaram que a distribuição de plantas aquáticas se dava principalmente entre os eixos R/CR devido à influência positiva que a evaporação exercia sobre as espécies com maior área foliar específica. Assim, concluíram que as condições climáticas influenciaram as estratégias ecológicas e exerceram uma influência específica sobre os traços funcionais de espécies nesses ambientes. Dessa forma, a contextualização do saber aliado a métodos ecológicos que traduzem as relações que enxergamos na natureza contribui para a formação de profissionais despertos para a interconexão existente entre os seres vivos e o ambiente.

Conforme destacado por Pierce *et al.* (2017), entender as relações estratégia-ambiente serão úteis também para prever como as comunidades poderão responder às mudanças no clima, nos ciclos biogeoquímicos e uso da terra, temáticas extremamente pertinentes na atualidade. Ressaltamos a importância da leitura dos artigos de Pierce *et al.* (2017) e Grime (2001), entre outros trabalhos, em que as características de cada uma das estratégias – primárias e se-

cundárias – são dadas em detalhe. Essa base teórica é fundamental para a compreensão e a interpretação dos resultados encontrados.

5. Para saber mais

- Para casos especiais na avaliação de atributos foliares, envolvendo, por exemplo, folhas compostas, heterofilia, plantas sem folhas e inclusão do pecíolo, consulte Pérez-Harguindeguy *et al.* (2013).
- *Download* ImageJ: <https://imagej.net/ij/download.html>
- Guia de uso do ImageJ: <https://imagej.net/ij/docs/guide/146-Part-IV.html>
- *Download* planilha *Stratify*: <https://besjournals.onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2F1365-2435.12722&file=fec12722-sup-0004-TableS2.xls>
- *Download* RStudio: <https://posit.co/downloads/x>
- Script para construção do gráfico ternário: <https://github.com/EbookBioinformaticaBiodiversidade/Cap2>
- Materiais para leitura e interpretação das estratégias ecológicas do Triângulo de Grime:
 - › GRIME, J. P.; PIERCE, S. *The evolutionary strategies that shape ecosystems*. New Jersey: John Wiley & Sons, 2012.
 - › GRIME, J. P. *Plant Strategies, Vegetation Processes and Ecosystem Properties*. 2. ed. Chichester: Wiley, 2001.
 - › PIERCE, S. *et al.* A global method for calculating plant CSR ecological strategies applied across biomes world-wide. *Functional ecology*, v. 31, n. 2, p. 444-457, 2017.

Referências

ALBUQUERQUE, A. C.; RODRIGUES-FILHO, C. A. de S. Influence of climatic variables on CSR strategies of aquatic plants in a semiarid region. *Hydrobiologia*, v. 847, p. 61-74, 2020.

GAUDICHET, C. *et al.* Influence of local environmental gradients on ecological strategies of herbaceous communities in riverine side channels. *Journal of Vegetation Science*, v. 35, n. 1, p. e13227, 2024.

GRIME, J. P. *Plant Strategies, Vegetation Processes and Ecosystem Properties*. 2. ed. Chichester: Wiley, 2001.

GRIME, J. P. Vegetation classification by reference to strategies. *Nature*, v. 250, p. 26-31, 1974.

GUO, W-Y. *et al.* The role of adaptive strategies in plant naturalization. *Ecology letters*, v. 21, n. 9, p. 1.380-1.389.

HAMILTON, N. E.; FERRY, M. Ggtern: Ternary Diagrams Using ggplot2. *Journal of Statistical Software, Code Snippets*, v. 87, n. 3, p. 1-17, 2018. DOI: 10.18637/jss.v087.c03.

IMAGEJ. ImageJ download. Disponível em: <https://imagej.net/ij/download.html>. Acesso em: 28 jun. 2024.

MEINDL, G. A.; NATT, M.-L.; VELARDI, S. (eds). *Humans and the Environment*. [s.l.]: Lumen Learning, 2021. E-book. Disponível em: <https://courses.lumenlearning.com/suny-binghamton-envil01&envi201/>.

PÉREZ-HARGUINDEGUY, N. *et al.* New handbook for standardised measurement of plant functional traits worldwide. *Australian Journal of Botany*, v. 61, p. 167-234, 2013.

PIERCE, S. *et al.* A global method for calculating plant CSR ecological strategies applied across biomes world-wide. *Functional ecology*, v. 31, n. 2, p. 444-457, 2017. DOI: <https://doi.org/10.1111/1365-2435.12722>.

PIERCE, S.; MAFFI, D.; FAORO, F.; CERABOLINI, B. E. L.; SPADA, A. The leaf anatomical trade-offs associated with plant ecological strategy variation. *Plant Ecology*, v. 223, p. 1.233-1.246, 2022.

R CORE TEAM. R: A language and environment for statistical computing. *R Foundation for Statistical Computing*, Vienna, Austria, [2024]. Disponível em: <https://www.R-project.org/>. Acesso em: 28 jun. 2024.

REICH, P. B. The world-wide ‘fast-slow’ plant economics spectrum: a traits manifesto. *Journal of Ecology*, n. 104, p. 275-301, 2014.

TRY Plant Trait Database. [2024]. Disponível em: <https://www.try-db.org/TryWeb/Home.php>. Acesso em: 28 jun. 2024.

WICKHAM, H.; BRYAN, J. Readxl: Read Excel Files. R package version 1.4.3, 2023. Disponível em: <https://github.com/tidyverse/readxl>.

WICKHAM, H. *Ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag, 2016.

ZANZOTTERA, M.; FRATT, M. D.; CACCIANIGA, M.; PIERCE, S.; CERABOLINI, B. E. L. Community-level variation in plant functional traits and ecological strategies shapes habitat structure along succession gradients in alpine environment. *Community Ecology*, v. 21, p. 55-65, 2020.

CONSERVAÇÃO



Conservação em foco: análises de dados georreferenciados para proteger a biodiversidade



Fernando Fernandes¹ & Taciane Schröder¹

¹Programa de Pós-graduação em Botânica, Universidade Federal do Rio Grande do Sul, Porto Alegre (Brasil)

1. Contextualização

Estimativas conservadoras sugerem que a Terra é lar de cerca de 10 milhões de espécies. Mesmo sem considerar a miríade de possibilidades dos dificilmente definidos procariontes (organismos unicelulares sem núcleo definido, como bactérias e arqueas), ainda podemos contar com cerca de 8,7 milhões de eucariontes (organismos cujas células possuem núcleo definido, como protistas, fungos, plantas e animais) (Díaz; Malhi, 2022). Segundo o *Catalogue of Life* (2024), dessas 8,7 milhões estimadas, existem, hoje, cerca de 2 milhões de espécies eucarióticas já descritas, com metade delas sendo insetos e um quinto plantas vasculares. Os fungos representam aproximadamente 7% das espécies conhecidas, enquanto os vertebrados são apenas 4%. Contudo, os números são tão impressionantes que a descrição de todas as espécies eucarióticas pode levar cerca de 1.200 anos, exigindo o trabalho de 303 mil taxonomistas (Díaz; Malhi, 2022).

Todo o potencial da biodiversidade, embora muitas vezes desconhecido, já está seriamente ameaçado. A Terra está enfrentando atualmente seu sexto grande evento de extinção em massa (Cowie; Bouchet; Fontaine, 2022). As atividades humanas são as principais responsáveis por essa crise, pois têm causado degradação dos ambientes naturais e mudanças climáticas que afetam negativamente ecossistemas ao redor do mundo. Esses impactos resultam em uma perda acelerada de biodiversidade (Habibullah *et al.*, 2022). Além

disso, mesmo entre as espécies já descritas, a maioria ainda não foi avaliada quanto ao seu estado de conservação. Isso é especialmente preocupante para grupos historicamente negligenciados, como plantas e fungos (Nic Lughadha *et al.*, 2020). No caso das plantas, somente cerca de 16% das espécies foram oficialmente avaliadas quanto ao estado de conservação (IUCN, 2023), sendo ainda estimado que, dentre as angiospermas, cerca de 45% das espécies enfrentam algum grau de ameaça (Bachman *et al.*, 2024). Já no caso dos fungos, apenas em torno de 0,4% das espécies descritas foram avaliadas (Antonelli *et al.*, 2023; IUCN, [2023]).

A falta de avaliação dessas espécies torna difícil implementar medidas eficazes de conservação, aumentando o risco de extinção para muitas delas. Portanto, a preservação da biodiversidade exige ações urgentes e coordenadas para mitigar os efeitos das atividades humanas e proteger os ecossistemas. Sem essas ações, corremos o risco de perder uma grande parte do patrimônio biológico do planeta, com consequências imprevisíveis para o equilíbrio ambiental e para a própria sobrevivência humana.

Para garantir a eficácia dos esforços de conservação, é essencial inicialmente compreender o nível de ameaça enfrentado por cada espécie. Nesse sentido, a Lista Vermelha da IUCN (International Union for Conservation of Nature) se destaca como um método baseado em dados, que utiliza uma metodologia rigorosa e uniforme para produzir classificações consistentes. Esse sistema é aplicável a todas as espécies, exceto para micro-organismos, proporcionando uma base confiável para iniciativas de conservação em escala global (IUCN, 2022). No Brasil, essa abordagem é refletida na publicação de listas oficiais de espécies ameaçadas, como a atualização da Lista Nacional de Espécies Ameaçadas de Extinção pela Portaria MMA nº 148, de 7 de junho de 2022. Essas listas utilizam os critérios estabelecidos pela IUCN como referência para suas avaliações. Dessa forma, a Lista Vermelha da IUCN se firmou como uma ferramenta essencial para a conservação da biodiversidade, reconhecida por sua eficácia global. É crucial reconhecer a

importância de entender profundamente o estado de conservação das espécies como um passo inicial vital para sua proteção contra a extinção, alinhando-se ao princípio fundamental de “conhecer para preservar”.

A IUCN (2024) utiliza cinco critérios para avaliar o grau de ameaça das espécies: A (redução da população), B (distribuição geográfica restrita e apresentando fragmentação, declínio ou flutuações), C (população pequena e com fragmentação, declínio ou flutuações), D (população muito pequena ou distribuição muito restrita) e E (análise quantitativa de risco de extinção). Esses critérios são aplicados para classificar as espécies em nove categorias de ameaça distintas: segura ou pouco preocupante (LC), quase ameaçada (NT), vulnerável (VU), em perigo (EN), criticamente em perigo (CR), extinta na natureza (EW) e extinta (EX). Destas, as categorias *criticamente em perigo*, *em perigo* e *vulnerável* são consideradas graus de ameaça de extinção.

Na avaliação do estado de conservação de uma espécie, o pesquisador deve submetê-la à análise de todos os critérios, não testando um deles somente se não houver dados disponíveis para aplicá-lo. A categoria de maior ameaça identificada ao final do processo determina a classificação da espécie.

Este capítulo foca na obtenção e no tratamento de dados para avaliar o estado de conservação das espécies seguindo o critério B, o qual trata da distribuição das espécies. Informações geográficas são fundamentais para calcular a extensão de ocorrência (EOO) e a área de ocupação (AOO) da espécie. Para realizar esses cálculos, utilizaremos informações de bancos de dados disponíveis on-line e a ferramenta GeoCAT – Geospatial Conservation Assessment Tool (2024), uma ferramenta geoespacial especializada em análises de conservação, que utiliza como base as categorias e critérios adotados pela IUCN.

2. Objetivo

Compreender os princípios da avaliação do estado de conservação das espécies segundo os critérios da IUCN e aprender a aplicar o critério B.

3. Descrição da atividade “Explorando a biodiversidade na prática”

Primeiramente, vamos entender onde obter os dados de distribuição geográfica. A segunda etapa será o tratamento dos dados, visando garantir sua qualidade e corrigir eventuais erros. Em seguida, realizaremos a análise propriamente dita, calculando a extensão de ocorrência (EOO) e a área de ocupação (AOO) da espécie. Após isso, interpretaremos os resultados e aplicaremos os dados de acordo com as regras do critério B da IUCN.

3.1. *Passo 1: Obtendo os dados de distribuição das espécies*

Os dados de distribuição geográfica das espécies podem ser encontrados em diferentes tipos de bancos de dados, sejam herbários, museus naturais ou plataformas on-line que compilam dados de coleções biológicas, como o Global Biodiversity Information Facility – GBIF (<https://www.gbif.org/>), e o speciesLink (<https://specieslink.net/>) (Para aprender como baixar dados destas plataformas ver capítulo “Coleções biológicas: os bancos de dados da biodiversidade”, deste livro). Para as análises a seguir, precisaremos obter dados de distribuição, que consistem em três colunas: nome da espécie; latitude; e longitude.

Além dessas fontes tradicionais, é importante destacar a existência de aplicativos como o iNaturalist (<https://www.inaturalist.org/>) e o WikiAves (<https://www.wikiaves.com.br/>), que permitem que cidadãos, incluindo estudantes, contribuam ativamente para o registro da biodiversidade. Essas plataformas, além de fornecerem dados de alta qualidade, já são integradas a grandes bancos de dados, como o GBIF, e podem oferecer registros valiosos, especialmente quando revisados com critério. Embora não tenham o mesmo rigor cientí-

fico que as coleções de herbários, com uma verificação cuidadosa os dados provenientes dessas plataformas podem ser bastante acurados. Esse tipo de engajamento pode ser uma excelente forma de envolver os estudantes no processo de coleta e documentação da biodiversidade, complementando os métodos tradicionais.

3.2. Passo 2: Organizando e tratando os dados

Os dados obtidos podem conter erros e imprecisões geográficas, que podem resultar de erros de digitação ou falhas no GPS durante a coleta do material. Idealmente, o pesquisador deve verificar os pontos de ocorrência para assegurar que as populações ainda existem e que os registros não estão incorretamente reportados. Além desses fatores, é comum nos depararmos com outros tipos de erros ao tratar dados de plataformas agregadoras, como o GBIF ou o speciesLink, que devem ser considerados durante a etapa de limpeza dos dados.

Dados duplicados: Muitas vezes, registros de uma mesma coleta podem ser duplicados, devido à presença de duplicatas do material em diferentes museus ou herbários. Agregadores como o GBIF podem tratar esses registros duplicados como coletas distintas, inflando o número de ocorrências. Nesses casos, é necessário identificar e remover as duplicatas.

Coordenadas com centroide do município: Em situações nas quais as coordenadas geográficas do ponto de coleta não estão disponíveis, o centroide do município pode ser usado como uma estimativa do local. No entanto, essa prática pode distorcer os resultados, já que o ponto não reflete a real localização da ocorrência. Para análises de conservação, o uso de coordenadas de centroides municipais deve ser evitado.

Coordenadas ausentes: Em muitos registros, as coordenadas podem estar simplesmente ausentes. Quando não é possível recuperar a informação geográfica correta, o ideal é remover esses registros da análise para evitar imprecisões.

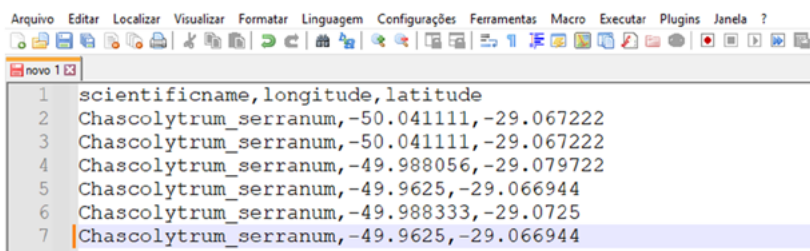
Verificação do local de coleta via Google Earth: Ferramentas como o Google Earth podem ser úteis para verificar se o local de coleta original ainda existe ou foi convertido em áreas urbanas, agrícolas ou de pecuária. A análise de imagens de satélite ao longo dos anos permite observar mudanças no uso da terra e avaliar se o habitat da espécie foi alterado por atividades antrópicas.

Esses erros são comuns e podem comprometer a precisão das análises de distribuição e conservação. Por isso, a verificação e limpeza cuidadosa dos dados são etapas essenciais para garantir resultados mais robustos e confiáveis.

Para tratar os dados obtidos das coleções biológicas, vamos utilizar uma planilha de Microsoft Excel ou o software Notepad++ (<https://notepad-plus-plus.org/>). Para esta prática, utilizaremos o Notepad++.

Abra seus dados no Notepad++ e verifique se todas as colunas estão corretas. Alguns dados baixados podem estar sem a latitude ou a longitude, ou com o nome da espécie escrito de forma diferente. Após isso, separe os dados utilizando apenas vírgulas, mantendo o título de cada coluna. Ao final do tratamento, você deve obter uma organização dos seus dados semelhante à Figura 1.

Figura 1: Dados organizados no software Notepad++.



Fonte: imagem gerada pelos autores.

Agora vamos salvar o arquivo no formato que o GeoCat utiliza. No Notepad++, com seus dados já organizados, clique em "Arquivo", depois em "Salvar como...". Selecione o "Tipo de arquivo" como "All

types” e escreva “.csv” ao final do nome do seu arquivo – por exemplo, “Distribuição_Cserranum.csv”. O arquivo “.csv” também pode ser gerado no software Microsoft Excel. Com o arquivo “.csv” salvo, estamos prontos para efetuar as análises no GeoCat.

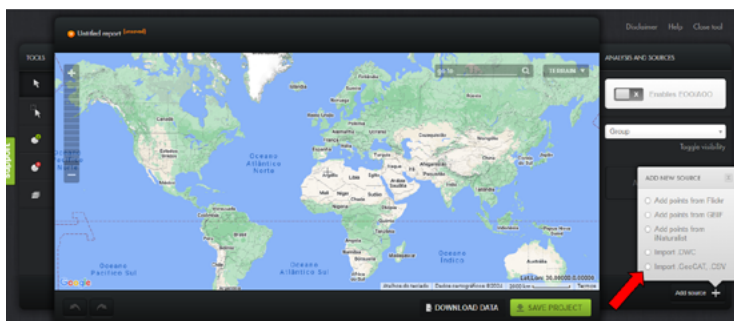
3.3. Passo 3: Utilizando o GeoCat

Ao abrir a ferramenta GeoCAT (<https://geocat.iucnredlist.org/editor>), você se depara com uma tela semelhante à Figura 2. Nesta tela inicial, você irá procurar o botão no canto inferior direito chamado “Add source”. Clicando nele, você verá a opção para fazer upload do seu arquivo “.csv”.

Figura 2: Página inicial da ferramenta GeoCat. A seta vermelha indica onde fazer o upload do arquivo “.csv” para realizar as análises.

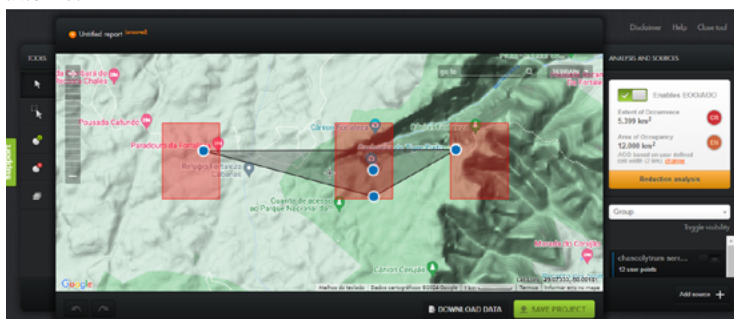
Fonte: IUCN, ([2024]). Disponível no link: <https://geocat.iucnredlist.org/editor>.

Após o upload completo, clique em “import” e a distribuição da sua espécie será carregada no mapa. Ao clicar em “Enable EOO/



AOO”, serão apresentadas a extensão de ocorrência (EOO) e a área de ocupação (AOO), conforme a Figura 3.

Figura 3: Cálculo de extensão de ocorrência (representado em cinza) e área de ocupação (representada em vermelho) para *Chascolytrum serranum* L.N.Silva.



Fonte: IUCN ([2024]). Disponível no link: <https://geocat.iucnredlist.org/editor>.

A extensão de ocorrência (EOO), representada pelo polígono em cinza, é a medida da área total delimitada pelos pontos extremos onde a espécie é encontrada. Essa métrica considera a distribuição geral da espécie. A área de ocupação (AOO), representada em vermelho, refere-se à área onde a espécie realmente ocorre dentro da EOO. A AOO considera apenas as áreas efetivamente ocupadas pela espécie, excluindo regiões onde ela não está presente. É importante mencionar que a Área de Ocupação (AOO), calculada no GeoCat, segue os parâmetros da IUCN, que utiliza uma célula padrão de 2 x 2 km ao redor dos registros de ocorrência. Essa escala visa padronizar a avaliação da área ocupada pelas espécies, permitindo uma comparação mais consistente entre diferentes espécies e regiões. No entanto, essa célula pode ser ajustada conforme necessário, de acordo com a realidade de distribuição de cada espécie. Embora o GeoCat aplique o valor padrão de forma automática, o ajuste manual pode ser considerado para espécies com distribuições muito localizadas ou dispersas, a fim de refletir com mais precisão sua AOO.

Após o cálculo, o GeoCat fornece uma sugestão inicial de nível de ameaça baseada nos valores de extensão de ocorrência (EOO) e área de ocupação (AOO) da espécie. É importante destacar que esses valores funcionam como parâmetros preliminares, auxiliando

no entendimento da distribuição geográfica da espécie. No entanto, a classificação final do estado de conservação exige uma análise mais aprofundada, que segue os passos detalhados do critério B da IUCN. Assim, mesmo que possam ser úteis, as sugestões fornecidas pelo GeoCat não substituem a avaliação oficial, que considera outros fatores, como fragmentação da população, declínios e flutuações, para determinar o nível de ameaça da espécie com maior precisão.

3.4. Passo 4: Avaliando o grau de ameaça que a espécie enfrenta

Para realizar a avaliação conforme o critério B, vamos utilizar o documento “Diretrizes para o Uso das Categorias e Critérios da Lista Vermelha da IUCN”, publicado pela IUCN em 2022, disponível no link: <https://www.iucnredlist.org/resources/redlistguidelines>.

Na página das diretrizes, temos todos os critérios e suas regras de aplicação, mas aqui vamos nos ater ao B, conforme a Figura 4.

Figura 4: Recorte do resumo do critério B usado para avaliar se um táxon pertence a uma categoria de ameaça (criticamente em perigo, em perigo ou vulnerável).

B. Distribuição geográfica restrita e apresentando fragmentação, declínios ou flutuações			
	Criticamente Em Perigo	Em Perigo	Vulnerável
B1 Extensão de ocorrência (EOO)	< 100 km²	< 5.000 km²	< 20.000 km²
B2 Área de ocupação (AOO)	< 10 km²	< 500 km²	< 2.000 km²
E pelo menos duas das seguintes condições:			
(a) População severamente fragmentada, OU número de localizações condicionadas a ameaça	= 1	≤ 5	≤ 10
(b) declínio continuado observado, estimado, inferido ou projetado em: (i) extensão de ocorrência; (ii) área de ocupação; (iii) área, extensão e/ou qualidade do habitat; (iv) número de localizações condicionadas a ameaça ou subpopulações; (v) número de indivíduos maduros.			
(c) flutuações extremas em: (i) extensão de ocorrência; (ii) área de ocupação; (iii) número de localizações condicionadas a ameaça ou subpopulações; (iv) número de indivíduos maduros.			

Fonte: Comitê de Padrões e Petições da IUCN (2022).

Com os cálculos de EOO e AOO em mãos, podemos utilizá-los para aplicar o critério B conforme a Figura 4. O primeiro passo é analisar quantitativamente a EOO e a AOO. De acordo com a tabela, nossa espécie exemplo, *Chascolytrum serranum* L.N.Silva, possui uma extensão de ocorrência inferior a 100 km² (5 km²), o que a categoriza como “criticamente em perigo”. Analisando a área de ocupação, a espécie exemplo possui uma AOO calculada

em 12 km², o que a categoriza como “em perigo”. Além da análise quantitativa, existem subcritérios qualitativos que devem ser avaliados com cuidado pelo pesquisador. Note que, para a avaliação, é necessário satisfazer pelo menos dois dos subcritérios. Conforme os subcritérios forem sendo analisados, o pesquisador vai anotando o código de cada um – por exemplo, o código “bii” indica que a espécie sofreu um declínio (observado, estimado, inferido ou projetado) em sua área de ocupação. Dessa forma, ao final da avaliação o pesquisador terá uma série de códigos que descrevem os tipos de ameaças que a espécie enfrenta.

Como a categoria de maior ameaça define o grau de ameaça da espécie, nossa espécie exemplo foi classificada como “criticamente ameaçada” de extinção, conforme o critério B, com o código Blb(i,ii),c(i,ii). Isso se deve à sua extensão de ocorrência menor que 100 km² e às observações de campo que indicam que uma de suas últimas localidades está ameaçada pelo avanço das atividades humanas, o que também resultou em um declínio na sua distribuição nos últimos anos.

É importante destacar que é essencial fundamentar as inferências realizadas nesta etapa para garantir a qualidade da avaliação. Lembrando que, caso surjam dúvidas, todos os subcritérios do critério B são explicados em detalhes ao longo do manual de diretrizes da IUCN (2022). Outros critérios além do B também se utilizam de dados qualitativos que requerem uma investigação de campo.

4. Considerações finais

Destaca-se que a aplicação cuidadosa de todos os critérios disponíveis é essencial para uma avaliação precisa do estado de conservação das espécies. Cada critério oferece uma perspectiva única e complementar, possibilitando uma análise abrangente que considera fatores tanto quantitativos quanto qualitativos.

Este capítulo concentrou-se em demonstrar como avaliar o critério B, que muitas vezes requer dados detalhados para análises

completas. Reconhecendo que esses dados muitas vezes estão fora do contexto da sala de aula, uma abordagem pedagógica eficaz envolve a criação de cenários hipotéticos durante o aprendizado. Por exemplo, ao simular observações de campo ou fornecer dados fictícios, os estudantes podem praticar a aplicação dos critérios da IUCN em um contexto controlado. Além de possibilitar o desenvolvimento de habilidades técnicas, como a interpretação de mapas de distribuição e o cálculo de EOO e AOO, isso também ensina os futuros conservacionistas a integrarem dados quantitativos e qualitativos de maneira significativa.

Atividade extra: explorando a perda de biodiversidade com MapBiomias

Para aprofundar a discussão sobre a perda de biodiversidade causada pela alteração de habitats, sugerimos uma atividade prática utilizando a plataforma MapBiomias (<https://mapbiomas.org/>), uma ferramenta que permite pesquisar mudanças no uso do solo no Brasil, incluindo o município dos próprios estudantes. Ao visualizar a perda de vegetação natural em suas regiões, os alunos podem ter um entendimento mais concreto e impactante da degradação ambiental ao seu redor.

Atividade sugerida

Explorar o MapBiomias: Oriente os estudantes a acessar a plataforma e investigar a evolução do uso do solo em seu município ou em uma área de interesse.

Analisar a perda de habitat: Peça que identifiquem áreas que já foram cobertas por vegetação nativa e que agora estão ocupadas por atividades humanas, como agricultura, pecuária ou urbanização.

Discussão em sala de aula: Após a pesquisa, promova uma discussão sobre como essa perda de habitat pode impactar a biodiversidade local e globalmente. Incentive reflexões sobre a

importância da conservação e as possíveis formas de engajamento para proteger o meio ambiente.

Essa atividade prática pode ajudar a cativar os estudantes para o tema da conservação, ao conectá-los diretamente com a realidade de suas próprias comunidades, e incentivá-los a se engajar ativamente na proteção da biodiversidade.

Referências

- ANTONELLI *et al.* *State of the World's Plants and Fungi*. London: Royal Botanic Gardens, 2023.
- BACHMAN, S. P.; BROWN, M. J.; NIC LUGHADHA, E.; LEÃO, T. C.; WALKER, B. Extinction risk predictions for the world's flowering plants to support their conservation. *New Phytologist*, p. 01-12, 2024.
- BRASIL. Portaria MMA nº 148, de 7 de junho de 2022. Atualização da Lista Nacional de Espécies Ameaçadas de Extinção. Brasília, DF: Ministério do Meio Ambiente, 2022.
- COWIE, R. H.; BOUCHET, P.; FONTAINE, B. The Sixth Mass Extinction: fact, fiction or speculation?. *Biological Reviews*, v. 97, n. 2, p. 640-663, 2022.
- DÍAZ, S.; MALHI, Y. Biodiversity: Concepts, patterns, trends, and perspectives. *Annual Review of Environment and Resources*, v. 47, p. 31-63, 2022.
- CATALOGUE OF LIFE. Banco de dados on-line. *Catalogue of Life*, [2024]. Disponível em: <https://www.catalogueoflife.org/>. Acesso em: 26 jun. 2024.
- GEOSPATIAL CONSERVATION ASSESSMENT TOOL (GEOCAT). Home Page. [2024]. Disponível em: <http://geocat.kew.org/>. Acesso em: 26 jun. 2024.
- GLOBAL BIODIVERSITY INFORMATION FACILITY (GBIF). Home Page. Disponível em: <http://www.gbif.org>. Acesso em: 27 jun. 2024.
- HABIBULLAH, M. S.; DIN, B. H.; TAN, S. H.; ZAHID, H. Impact of climate change on biodiversity loss: global evidence. *Environmental Science and Pollution Research*, v. 29, n. 1, p. 1073-1086, 2022.
- INTERNATIONAL UNION FOR CONSERVATION OF NATURE (IUCN). Comitê de Padrões e Petições da IUCN. Diretrizes para o Uso das Categorias e Critérios da Lista Vermelha da IUCN. Versão

15.1. 2022. Disponível em: <https://cmsdocs.s3.amazonaws.com/RedListGuidelines.pdf>. Acesso em: 02 ago. 2024.

INTERNATIONAL UNION FOR CONSERVATION OF NATURE (IUCN). Diretrizes para o Uso das Categorias e Critérios da Lista Vermelha da IUCN. Versão 15.1. Preparada pelo Comitê de Padrões e Petições. IUCN, 2022. Disponível em: <https://www.iucnredlist.org/resources/redlistguidelines>.

INTERNATIONAL UNION FOR CONSERVATION OF NATURE (IUCN). Red List version 2023-1. [2023]. Disponível em: https://nc.iucnredlist.org/redlist/content/attachment_files/2023-1_RL_Table_1a.pdf. Acesso em: 31 jul. 2024.

INTERNATIONAL UNION FOR CONSERVATION OF NATURE (IUCN). Banco de dados online: Red List of threatened species. [2024]. Disponível em: <https://www.iucnredlist.org/>. Acesso em: 26 jun. 2024.

NIC LUGHADHA, E. *et al.* Extinction risk and threats to plants and fungi. *Plants, People, Planet*, v. 2, n. 5, p. 389-408, 2020.

SPECIESLINK NETWORK. Home Page. [2024]. Disponível em: www.specieslink.net. Acesso em: 27 jun. 2024.

Genômica aplicada à conservação de espécies: decifrando enigmas taxonômicos



Bruna Boizonave Andriola¹, Gabriele Zenato Lazzari¹, Yago Beux¹ & Amanda Kessler¹

¹Programa de Pós-graduação em Ecologia e Evolução da Biodiversidade, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre (Brasil)

1. Contextualização

1.1. A Era Genômica e os avanços no estudo de populações selvagens

Vivemos em um cenário de acentuado declínio da biodiversidade (ou sexta extinção em massa) e de crise climática, que ameaça a árvore da vida e consequentes serviços ecossistêmicos prestados à humanidade (Ceballos; Ehrlichb, 2023). Nesse contexto, como uma forma de fornecer suporte para medidas legais de proteção de espécies ameaçadas e ações de conservação, é necessário, primeiramente, melhorar a compreensão da biodiversidade global. Para isso, realizar estudos de delimitação de espécies; identificação dos mecanismos genéticos por trás da adaptação local¹; o manejo e o monitoramento de populações em vida livre ou de cativeiro; e programas de reprodução, translocação e reintrodução de espécies na natureza são ações de extrema importância (Hogg, 2023; Segelbacher *et al.*, 2022; Supple; Shapiro, 2018).

O primeiro passo para a conservação é compreendermos quais são as entidades que devem ser conservadas. Quando falamos em entidade, podemos pensar não apenas em espécies, mas também em suas populações. No que diz respeito a espécies, precisamos

¹ Processo por meio do qual as populações apresentam características que conferem maior sobrevivência e reprodução no ambiente local do que em qualquer outro lugar.

realizar a delimitação destas. Há diferentes formas de realizar tal delimitação. Estudos de taxonomia clássicos baseiam-se, principalmente, em dados morfológicos. Com o avanço da tecnologia, a utilização de dados genéticos se tornou mais acessível e indispensável nesses estudos.

Marcadores genéticos², mitocondriais e nucleares tornaram-se muito úteis em estudos que buscam compreender o relacionamento entre as espécies. Contudo, a partir da utilização desses dados, também se pode observar uma maior discordância filogenética, problema obtido devido ao uso de diferentes marcadores e/ou porções do genoma amostrado. A utilização de dados genômicos para compreender a história evolutiva de organismos não-modelo mostra-se extremamente eficiente, desvendando uma biodiversidade desconhecida e seus relacionamentos complexos (Lopes *et al.*, 2023; Musher *et al.*, 2022).

Neste capítulo, quando falamos em abordagens genômicas, fazemos referência a estudos que utilizam genomas completos ou diversos marcadores³ representativos de regiões do genoma. São grandes conjuntos de dados que proporcionam elevado poder analítico, melhorando a precisão de estimativas de diversidade genética, estruturação populacional e história demográfica, bem como permitindo a investigação de questões antes inatingíveis com dados genéticos tradicionais (estes, restritos a poucos *loci* ou pequenas regiões). Ao analisar conjuntos maiores, é possível minimizar os vieses de poucas ou uma região específica, pois a história evolutiva de um *locus* não será, necessariamente, a mesma da espécie. Ao mesmo tempo, dados genômicos também apresentam novos desafios aos pesquisadores, relacionados a custos de geração de dados, complexidade de manuseio destes, disponibilidade de máquinas com capacidade de analisar os dados e interpretação das análises, exigindo o conhecimento de ferramentas de bioinformática.

² Como marcadores microsatélites (ou STRs), sequências repetidas de DNA ao longo do genoma altamente polimórficas. Porções do genoma mitocondrial, como o gene citocromo B e a região controle da mitocôndria.

³ Nesse caso, polimorfismos de nucleotídeo único (ou SNPs) – polimorfismos resultantes da alteração de uma única base, distribuídos ao longo do genoma.

Avanços nas tecnologias de sequenciamento permitiram a obtenção de dados para um número crescente de espécies não-modelo, incluindo o uso de DNA extraído de amostras históricas, amostras não-invasivas (ex. pelos, penas e fezes) e amostras ambientais⁴ (ex. água e solo), comumente obtido em baixas quantidades e com uma qualidade inferior (Hohenlohe; Funk; Rajora, 2021). Além disso, a exemplo da tecnologia de sequenciamento por nanoporos (MinION), atualmente é possível até mesmo sequenciar amostras em campo, com equipamentos portáteis que permitem análises dos resultados em tempo real, rompendo barreiras de acessibilidade em ambientes com recursos limitados ou de difícil acesso (Wasswa *et al.*, 2022).

De forma geral, as opções de sequenciamento diferem principalmente na quantidade de genoma que é sequenciada, com a possibilidade de sequenciar partes específicas do genoma (sequenciamento de representação reduzida – RRS; na ordem de megabytes) ou sequenciar o genoma inteiro de um organismo (sequenciamento de genoma completo – WGS; na ordem de gigabytes). Outros parâmetros importantes envolvem diferentes opções de profundidade de sequenciamento⁵, custo e capacidade de processamento necessária para análise dos dados. Dependendo da abordagem escolhida, os cientistas conseguem desvendar histórias evolutivas complexas, estimar o impacto de cruzamentos consanguíneos, modelar histórias demográficas⁶, identificar mutações deletérias ou associadas ao surgimento de novas espécies, investigar o potencial adaptativo a longo prazo, entre outras possibilidades. A seguir, veremos como a genômica pode ajudar a decifrar enigmas taxonômicos e qual a relevância disso para a conservação de espécies.

1.2. Delimitando espécies e informando ações de conservação

A classificação taxonômica de cada unidade evolutiva deve ser a mais bem resolvida possível, pois é sob ela que se define o *status*

⁴ DNA ambiental (do inglês, *environmental DNA* – eDNA)

⁵ Quantidade média de vezes que cada posição do genoma/sítio é sequenciada.

⁶ Mudanças no tamanho populacional e eventos de migração.

de conservação⁷. Os critérios para a definição do *status* envolvem, além da classificação taxonômica, a quantidade de dados disponíveis, o tamanho populacional da espécie, o grau de fragmentação de suas populações, a distribuição geográfica, entre outros fatores (IUCN, 2012). Uma espécie que apresenta diversas populações isoladas que carecem de estudos ecológicos e genéticos pode estar sob grande ameaça. Essas populações que acreditamos serem compostas de apenas uma espécie podem apresentar um número ainda não descoberto de unidades evolutivamente significativas (ESUs⁸) que precisam urgentemente de medidas de conservação. Para isso, é necessário que estejamos atentos aos conceitos de delimitação de espécies (Frankham; Ballou; Briscoe, 2008).

Entre os cientistas não há um consenso sobre qual conceito é o mais correto. O fenético define uma espécie a partir das similaridades fenotípicas; o ecológico reconhece uma espécie como um grupo que explora e ocupa o mesmo nicho ecológico; e o biológico define que, para um grupo ser considerado uma espécie, os indivíduos devem ser capazes de inter cruzarem entre si e estarem reprodutivamente isolados dos demais (Ridley, 2006). Esse grupo é capaz de trocar material genético dentro da sua população e entre populações adjacentes, produzindo uma prole viável e fértil. Isto geralmente não acontece quando há o cruzamento entre membros de espécies diferentes, as quais podem ser incapazes de reproduzir (isolamento pré-zigótico) ou podem gerar descendentes estéreis (isolamento pós-zigótico). Esse conceito tem sido o mais aceito dentro da genética de populações⁹, biologia evolutiva e biologia da conservação¹⁰ (Frankham; Ballou; Briscoe, 2008).

Apesar de ser o conceito mais usado, estudos recentes mostram que o surgimento de uma espécie é mais complexo do que

⁷ Não avaliada (NE), dados insuficientes (DD), menos preocupante (LC), quase ameaçada (NT), vulnerável (VU), em perigo (EN), criticamente em perigo (CR), extinta na natureza (EW) e extinta (EX).

⁸ Do inglês, *Evolutionarily Significant Units*.

⁹ Área da biologia que estuda a dinâmica populacional das espécies em relação a sua frequência alélica e processos de evolução.

¹⁰ Área que visa compreender os impactos das atividades humanas sobre as espécies, buscando medidas para conservar essas entidades.

apenas a simples separação do seu ancestral, devido ao isolamento reprodutivo. Tal processo pode também envolver eventos de introgressão¹¹ e separação incompleta de linhagens¹², que violam as premissas do conceito biológico, uma vez que uma espécie pode surgir a partir da hibridização de ESUs distintas. É o caso do estudo da história evolutiva dos lobos-marinhos-peruanos (Lopes *et al.*, 2023), que empregou dados genômicos para reconstruir os relacionamentos filogenéticos entre as espécies.

A reconstrução filogenética permite reconhecer a monofilia¹³ de um grupo taxonômico, sua diversidade genética e fluxo gênico entre populações. Os principais métodos de reconstrução filogenética são: a máxima verossimilhança, a inferência bayesiana e o *Neighbor-Joining* (Matioli; Fernandes, 2012).

Em resumo, a aplicação de dados genômicos para estudar a história evolutiva dos organismos nos permite compreender os relacionamentos e a classificação de distintas ESUs, ajudando a identificar e proteger espécies que podem ser morfologicamente similares, mas geneticamente diferentes (Trindade *et al.*, 2021). Uma vez que as identificamos, é possível calcular outras métricas, como a diversidade nucleotídica, o tamanho populacional efetivo¹⁴, o grau de endogamia, entre outras. Esses parâmetros fornecem embasamento para a melhor tomada de decisão no âmbito da conservação da biodiversidade do nosso planeta, permitindo, por exemplo, o planejamento para criação de áreas protegidas e corredores de dispersão.

Os avanços contínuos nas tecnologias de sequenciamento e bioinformática oferecem poderosas ferramentas que prometem auxiliar os profissionais no enfrentamento de desafios complexos relacionados à biodiversidade e à crise climática. A utilização da genômica é uma dessas ferramentas. A aplicação de dados genômi-

¹¹ Inserção de genes de uma espécie no genoma de outra.

¹² Retenção de polimorfismos ancestrais em linhagens próximas evolutivamente.

¹³ Conceito da biologia evolutiva que descreve um agrupamento taxonômico que inclui um ancestral comum e todos os seus descendentes.

¹⁴ Calculado com base no número de indivíduos que contribuem reprodutivamente para a espécie.

cos no estudo de populações e espécies nos auxilia na compreensão dos relacionamentos filogenéticos e na delimitação de entidades evolutivas distintas, importantes para o estabelecimento de unidades prioritárias para conservação. Logo, é importante que os futuros profissionais da área estejam atualizados e preparados para explorar suas potenciais aplicações.

2. Objetivo

Ao final desta atividade, os alunos deverão ser capazes de: (I) aplicar métodos de análise filogenética para delimitação de espécies; (II) comparar as vantagens e limitações de diferentes conjuntos de dados genômicos no contexto de análises filogenéticas; (III) avaliar as filogenias obtidas e a relação entre as espécies; e, por fim, (IV) explorar e avaliar as implicações do resultado para decisões de conservação dos organismos-alvo.

3. Descrição da atividade

3.1. Público-alvo e requisitos para a atividade

Para a execução da atividade, é esperado que os alunos já tenham sido introduzidos à interpretação de filogenias e disponham de computadores com acesso à internet.

3.2. Contextualizando os alunos

As tartarugas-marinhas são um grupo de distribuição global que historicamente tem sofrido reduções populacionais devido à pesca acidental, à caça de ovos e de indivíduos adultos (Jackson, 2001). Esses declínios têm motivado esforços globais de conservação desde 1950, incluindo diversas medidas para proteção de praias, regulação da pesca e estabelecimento de áreas de proteção (Hamann *et al.*, 2010; Wallace *et al.*, 2011). Atualmente, segundo a IUCN¹⁵, quatro espécies são categorizadas como vulneráveis (VU) (*Caretta caretta*, *Chelonia mydas*, *Dermochelys coriacea* e *Lepidochelys olivacea*), duas espécies como criticamente em

¹⁵ União Internacional para a Conservação da Natureza.

perigo (CR) (*Eretmochelys imbricata* e *Lepidochelys kempii*) e uma espécie sem dados suficientes (DD) para categorização (*Nattator depressus*).

A quantidade de dados de sequenciamento disponíveis atualmente traz consigo um desafio: quais tipos de sequências utilizar? A resposta para essa questão vai depender de fatores como o objetivo do trabalho, os tipos de dados que os últimos estudos usaram e a disponibilidade computacional.

Quando pensamos no tipo de pergunta que queremos responder, precisamos também levar em consideração os dados que serão utilizados. Levando em consideração a questão de delimitação de espécies e estrutura de populações. Marcadores genéticos clássicos, como os microssatélites (para a região nuclear) ou genes mitocondriais (ex. D-loop, cyt-b e COI), podem ser úteis como ponto de partida, devido ao baixo custo financeiro e computacional, permitindo a análise de muitos indivíduos e auxiliando a explorar o padrão de agrupamento.

O genoma mitocondrial, uma herança unicamente materna, pode ser utilizado em casos de populações com divergência recente devido à taxa de mutação elevada. Além disso, em termos de esforço computacional, esse tipo de dado é intermediário em relação aos marcadores clássicos e genomas completos, devido ao seu tamanho reduzido. Contudo, ele apresenta propriedades¹⁶ que podem resultar em discordâncias nos resultados quando comparados ao genoma nuclear, como topologias conflitantes, tempos de divergência e diversidade genética.

O genoma nuclear, por sua vez, permite explorar as mesmas questões que o mitocondrial, levando em consideração a recombinação entre os genomas, o que agrega um grau de complexidade na compreensão dos relacionamentos. Esse tipo de dado soluciona as discordâncias filogenéticas encontradas no uso dos marcadores

¹⁶ Sistema de herança matrilinear (transmissão exclusiva da linhagem materna para a prole), taxa de mutação mais elevada que o genoma nuclear e ausência de recombinação.

clássicos, causadas por introgressões/hibridizações e separação incompleta de linhagens. Além disso, também permite fazer estimativas de demografia histórica, heterozigosidade e estimativas populacionais. Quando a localização dos componentes genômicos (regiões codificantes, regiões repetitivas, regiões transponíveis, etc.) está disponível (chamada de anotação genômica), torna possível analisar as funções gênicas, variantes locais ligadas a populações, genes sob seleção e estrutura cromossômica. Entretanto, o custo computacional e financeiro desse tipo de dado é consideravelmente maior em comparação aos mencionados anteriormente.

Dentro das possibilidades para desvendar relacionamentos filogenéticos, o ideal é o uso de genomas nucleares completos, permitindo a realização de análises mais refinadas. Contudo, isso nem sempre é uma realidade devido às limitações computacionais. Sendo assim, na atividade a seguir serão utilizadas sequências compondo dois conjuntos de dados, um com sequências da região controle da mitocôndria (D-loop) e outro com genomas mitocondriais completos. Serão utilizadas sete espécies de tartarugas-marinhas e uma espécie mais distante filogeneticamente (*Chelydra serpentina*) como grupo-externo¹⁷. No primeiro, serão utilizados dois indivíduos por espécie (um indivíduo do Oceano Atlântico e outro da região indo-pacífica) para espécies com distribuição global. No segundo, haverá apenas um indivíduo por espécie.

Dispondo desses dois conjuntos de dados, testaremos se é possível recuperar clados monofiléticos com monofilia recíproca¹⁸ ou se as populações constituem unidades evolutivas distintas. Para estimar o relacionamento evolutivo desse grupo, será realizado o alinhamento das sequências, a estimativa de modelo evolutivo¹⁹ e as estimativas filogenéticas usando dois métodos: máxima verossimilhança e *Neighbor-Joining*.

¹⁷ Espécies distantes filogeneticamente usadas como grupo de referência para determinar as relações evolutivas do grupo a ser analisado.

¹⁸ Quando existe concordância nas topologias estimadas a partir de diferentes dados.

¹⁹ Padrões de substituição nucleotídica.

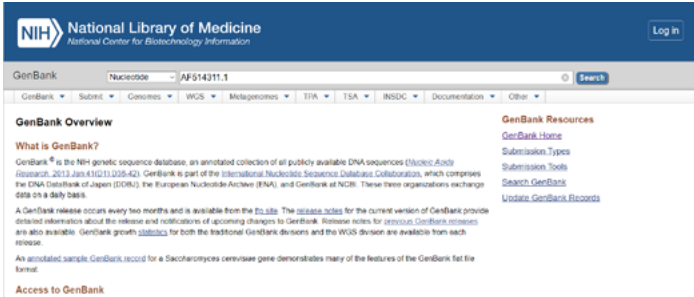
3.3. Obtendo os dados e as ferramentas necessárias

Inicialmente, é preciso baixar as sequências das espécies-alvo. Para isso, acesse a base de dados do NCBI (National Center for Biotechnology Information), disponível em <https://www.ncbi.nlm.nih.gov>.

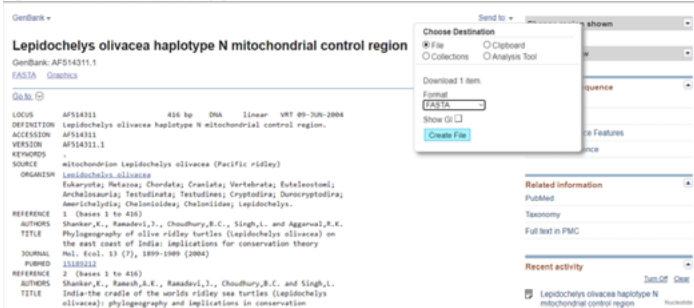
- A) Na barra de pesquisa, selecione a opção “Nucleotide” e digite o acesso “AF514311.1” para buscar a sequência da tartaruga-oliva (*Lepidochelys olivacea*) do Oceano Atlântico (Figura 1A). Baixe a sequência conforme a Figura 1B.

Figura 1: Base de dados de sequências biológicas do NCBI.

(A)



(B)



Fonte: imagens geradas pelos autores.

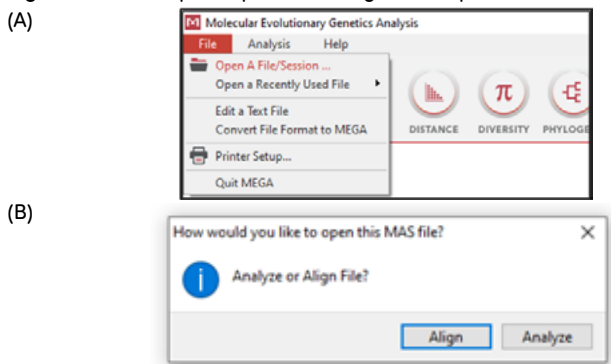
- B) Renomeie o arquivo baixado com o nome da espécie (“*Lepidochelys olivacea_atl*” ou “*Lolivacea_atl*”).
- C) Abra o arquivo e renomeie o cabeçalho (linha que contém o símbolo “>”) com o mesmo nome que foi dado ao arquivo fasta.

- D) As sequências restantes estão disponíveis em <https://github.com/EbookBioinformaticaBiodiversidade/Cap4>.
- E) Agora que já foram baixadas todas as sequências de interesse, é necessário que elas estejam em um único arquivo fasta. Para isso, crie um novo arquivo de texto, adicione a extensão “.fasta” ao nome e copie o cabeçalho e a sequência de cada espécie para esse novo arquivo uma abaixo da outra. Em seguida, instale o programa MEGA 11 (disponível em <https://www.megasoftware.net/home>).

3.4. Alinhamento

- A) Abra o MEGA 11, importe o arquivo com todas as sequências da região controle (Figura 2A) e clique em “align” para abrir as sequências (Figura 2B).

Figura 2: Passo a passo para o carregamento das sequências no MEGA 11.

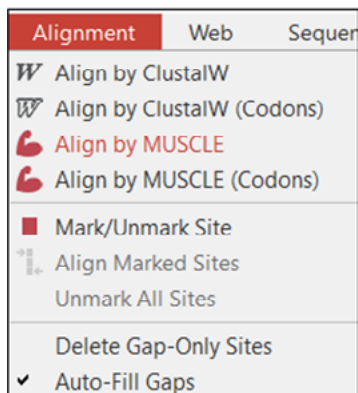


Fonte: imagens geradas pelos autores.

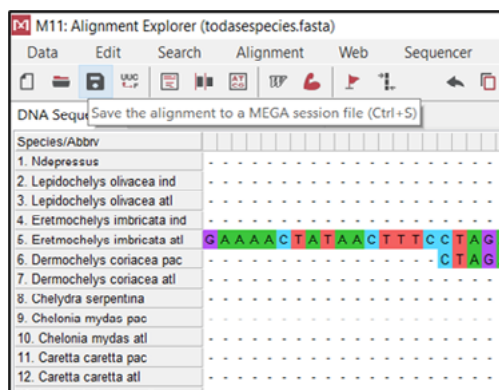
- B) Selecione todas as sequências e clique em *alignment* para selecionar um algoritmo de alinhamento (MUSCLE ou ClustalW) (Figura 3A). Neste exemplo, será usado o algoritmo MUSCLE devido à maior velocidade e precisão com alinhamentos complexos. Salve o alinhamento conforme a Figura 3B.

Figura 3: Passo a passo para o alinhamento de seqüências no MEGA 11.

A)



(B)

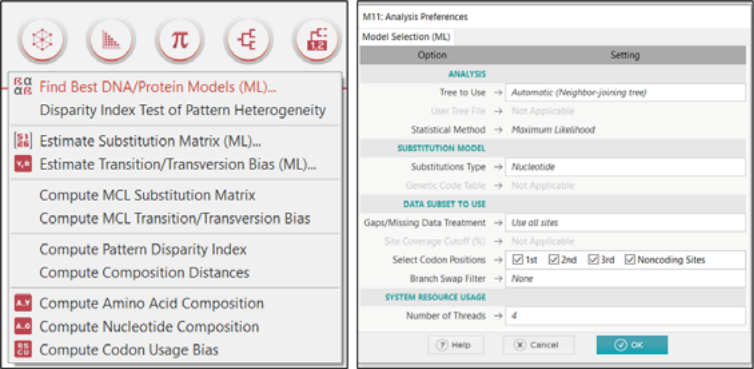


Fonte: imagens geradas pelos autores.

3.5. Análise de dados

- A) Abra o alinhamento da mesma maneira que fez com o arquivo “fasta”, clique em *analyze* e depois em *yes*.
- B) Verifique qual modelo de substituição de nucleotídeo melhor se adequa para o conjunto de dados (Figura 4).

Figura 4: Passo a passo e parâmetros usados na estimativa do modelo de substituição.



Fonte: imagens geradas pelos autores.

C) Ao final do teste será aberta uma janela com a tabela contendo os modelos evolutivos que foram testados (Figura 5). O modelo com o menor valor de BIC (*Bayesian Information Criterion*) é o mais adequado para ser usado na estimativa filogenética, pois garante que o modelo selecionado tem um bom ajuste aos dados, sem possuir parâmetros excessivos que levariam ao sobreajuste (do inglês *overfitting*²⁰).

²⁰ Ocorre quando um modelo estatístico ou de aprendizado de máquina se ajusta tão bem aos dados de treinamento que capta não apenas os padrões reais, mas também o ruído e as variações aleatórias presentes.

Figura 5: Resultado da estimativa de modelo de substituição.

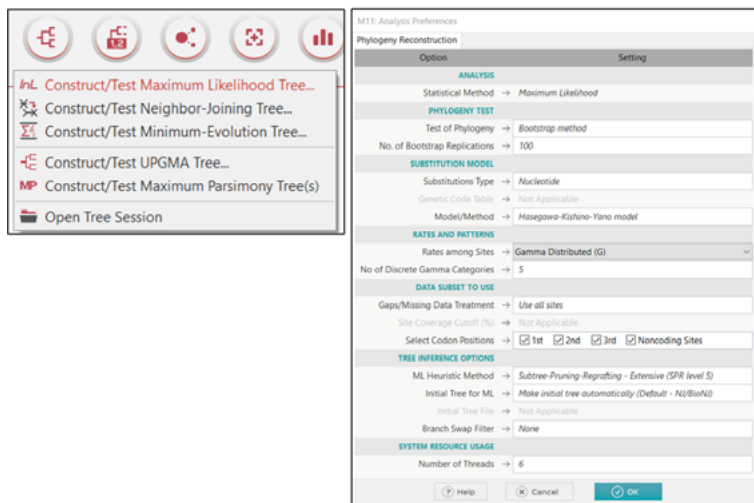
Table. Maximum Likelihood fits of 24 different nucleotide substitution models

Model	Parameters	BIC	AICc	lnL	(+)	(+G)	R	f(A)	f(T)	f(C)	f(G)	r(AT)	r(AC)	r(AG)	r(TA)	r(TC)	r(TG)	r(CA)	r(CT)	r(CG)	r(GA)	r(GT)	r(GC)
HKY+G	28	6787.009	6595.917	-3209.840	n/a	0.52	2.41	0.337	0.338	0.191	0.134	0.045	0.026	0.096	0.045	0.140	0.018	0.045	0.248	0.018	0.246	0.045	0.026
HKY+I	28	6788.762	6597.609	-3270.686	0.45	n/a	2.24	0.337	0.338	0.191	0.134	0.048	0.027	0.096	0.048	0.137	0.019	0.048	0.243	0.019	0.242	0.048	0.027
TN93+G	29	6791.566	6593.595	-3267.670	n/a	0.53	2.40	0.337	0.338	0.191	0.134	0.045	0.025	0.117	0.044	0.117	0.018	0.044	0.207	0.018	0.294	0.045	0.025
TN93+I	29	6793.374	6595.403	-3268.574	0.45	n/a	2.24	0.337	0.338	0.191	0.134	0.047	0.027	0.113	0.047	0.118	0.019	0.047	0.208	0.019	0.284	0.047	0.027
HKY+G+I	29	6793.458	6595.486	-3268.616	0.36	2.23	2.36	0.337	0.338	0.191	0.134	0.046	0.026	0.097	0.046	0.139	0.018	0.046	0.246	0.018	0.245	0.046	0.026
T92+G	26	6795.724	6616.210	-3283.002	n/a	0.54	2.34	0.337	0.337	0.163	0.163	0.046	0.022	0.118	0.046	0.118	0.022	0.046	0.246	0.022	0.246	0.046	0.022
T92+I	26	6797.541	6620.027	-3283.911	0.44	n/a	2.19	0.337	0.337	0.163	0.163	0.048	0.023	0.116	0.048	0.116	0.023	0.048	0.241	0.023	0.241	0.048	0.023
TN93+G+I	30	6797.837	6593.048	-3266.388	0.36	2.20	2.37	0.337	0.338	0.191	0.134	0.045	0.026	0.116	0.045	0.117	0.018	0.045	0.207	0.018	0.293	0.045	0.026
T92+G+I	27	6802.276	6617.942	-3281.861	0.35	2.20	2.31	0.337	0.337	0.163	0.163	0.047	0.022	0.118	0.047	0.118	0.022	0.047	0.244	0.022	0.244	0.047	0.022
GTR+G	32	6807.336	6588.914	-3262.303	n/a	0.52	2.42	0.337	0.338	0.191	0.134	0.041	0.042	0.118	0.041	0.120	0.013	0.073	0.212	0.004	0.297	0.033	0.005
GTR+I	32	6810.312	6591.890	-3263.790	0.45	n/a	2.28	0.337	0.338	0.191	0.134	0.041	0.042	0.114	0.041	0.119	0.016	0.075	0.210	0.005	0.289	0.041	0.007
GTR+G+I	33	6814.086	6588.847	-3261.259	0.35	1.92	2.41	0.337	0.338	0.191	0.134	0.040	0.042	0.118	0.040	0.119	0.014	0.074	0.211	0.004	0.296	0.036	0.006
HKY	27	6892.539	6708.205	-3326.992	n/a	n/a	1.81	0.337	0.338	0.191	0.134	0.055	0.031	0.090	0.055	0.129	0.022	0.055	0.228	0.022	0.227	0.055	0.031
T92	25	6897.055	6726.361	-3338.086	n/a	n/a	1.81	0.337	0.337	0.163	0.163	0.055	0.026	0.110	0.055	0.110	0.026	0.055	0.227	0.026	0.227	0.055	0.026
TN93	28	6897.261	6706.108	-3324.935	n/a	n/a	1.82	0.337	0.338	0.191	0.134	0.055	0.031	0.102	0.054	0.114	0.022	0.054	0.202	0.022	0.258	0.055	0.031
GTR	31	6911.756	6700.150	-3318.930	n/a	n/a	1.82	0.337	0.338	0.191	0.134	0.052	0.046	0.104	0.052	0.115	0.014	0.082	0.203	0.014	0.262	0.036	0.019
K2+G	25	6928.827	6758.132	-3353.971	n/a	0.54	2.20	0.250	0.250	0.250	0.250	0.039	0.039	0.172	0.039	0.172	0.039	0.039	0.172	0.039	0.172	0.039	0.039
K2+I	25	6939.968	6769.274	-3359.542	0.44	n/a	2.06	0.250	0.250	0.250	0.250	0.041	0.041	0.168	0.041	0.168	0.041	0.041	0.168	0.041	0.168	0.041	0.041
K2+G+I	26	6955.158	6777.644	-3362.719	0.00	0.44	3.56	0.250	0.250	0.250	0.250	0.027	0.027	0.195	0.027	0.195	0.027	0.027	0.195	0.027	0.195	0.027	0.027
K2	24	7039.431	6875.558	-3413.691	n/a	n/a	2.40	0.250	0.250	0.250	0.250	0.037	0.037	0.176	0.037	0.176	0.037	0.037	0.176	0.037	0.176	0.037	0.037
JC+G	24	7110.022	6946.148	-3448.967	n/a	0.67	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+I	24	7116.672	6952.799	-3452.312	0.42	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC+G+I	25	7118.770	6948.076	-3448.943	0.15	1.02	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083
JC	23	7190.775	7033.723	-3493.781	n/a	n/a	0.50	0.250	0.250	0.250	0.250	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083	0.083

Fonte: imagem gerada pelos autores.

D) Quando clicar no ícone “*phylogeny*”, verá que existem diversos métodos disponíveis. Neste tutorial serão usados apenas dois: *Neighbor-Joining*, que utiliza uma matriz de distância par a par para determinar os agrupamentos; e máxima verossimilhança, que determina quais combinações de topologia, comprimentos de ramos e modelo evolutivo são mais prováveis de serem observados com base nas sequências fornecidas. Os parâmetros que serão usados estão nas Figuras 6A e 6B.

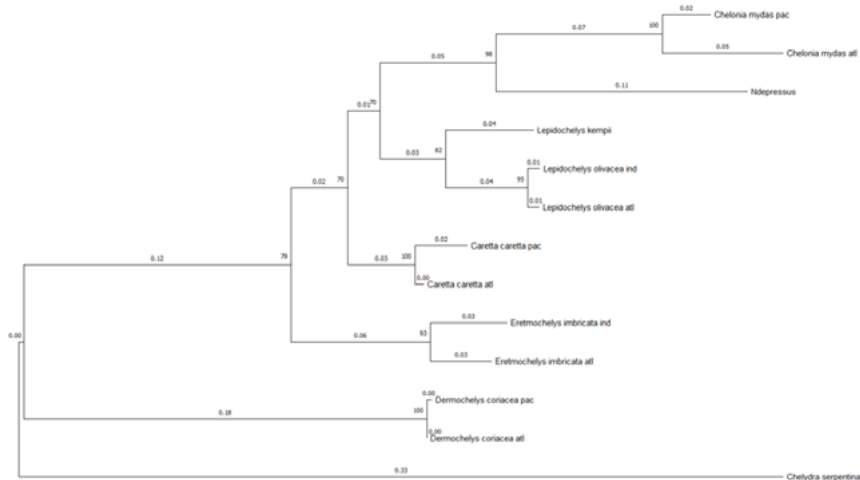
Figura 6: Passo a passo e parâmetros usados na estimativa filogenética.



Fonte: imagens geradas pelos autores.

E) A filogenia esperada é apresentada na Figura 7. Exporte a filogenia em formato *newick* (.nwk) para poder abri-la em outros programas de visualização, como o Figtree (disponível em: <https://github.com/rambaut/figtree/releases>).

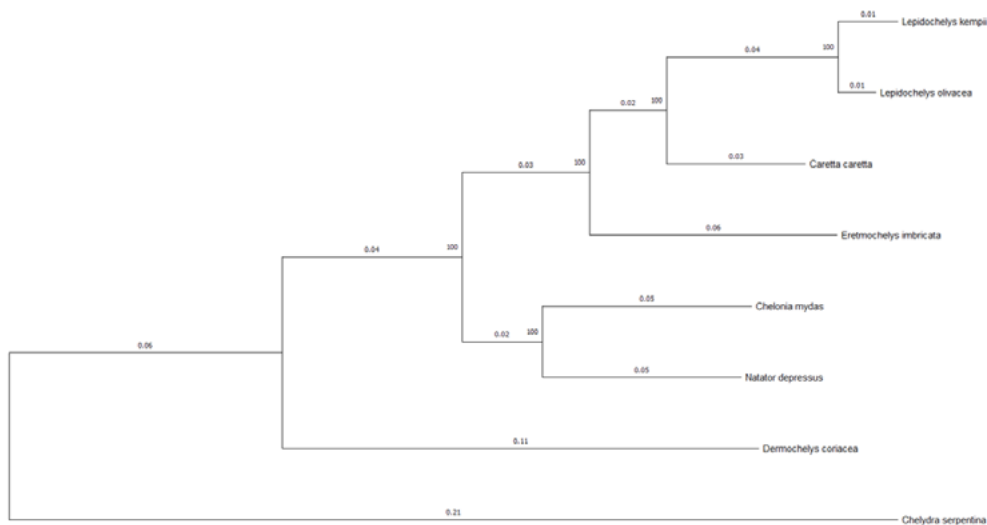
Figura 7: Filogenia das tartarugas-marinhas com base na região-controle da mitocôndria.



Fonte: imagem gerada pelos autores.

F) Agora com o conjunto de dados do genoma mitocondrial completo, refaça o passo a passo e compare os resultados, a topologia pode ser verificada na Figura 8.

Figura 8: Filogenia das tartarugas-marinhas com base em genomas mitocondriais completos.



Fonte: imagem gerada pelos autores.

3.6. Discussão e avaliação da aprendizagem

Tendo executado a atividade corretamente, os alunos terão obtido a monofilia de todos os gêneros de tartarugas-marinhas. Os indivíduos de oceanos distintos foram recuperados pertencendo ao mesmo grupo monofilético em ambos conjuntos de dados. Uma vez que ESUs são identificadas, a implementação de estratégias de conservação pode ser realizada para beneficiar todos aqueles que compartilham das mesmas adaptações e necessidades ecológicas.

Apesar de recuperar o mesmo relacionamento em ambos os marcadores, podemos verificar que os ramos com menos de 80% de

suporte nas replicatas de *bootstrap*²¹ foram rearranjados na filogenia do genoma mitocondrial com 100% de suporte. Esse resultado exemplifica a necessidade de explorar diferentes tipos de dados e métodos antes da tomada de decisões acerca do planejamento e da alocação de recursos para conservação. Ao integrar análises filogenéticas e dados filogeográficos, por exemplo, é possível entender o histórico de conectividade ou isolamento entre populações. No caso de populações historicamente isoladas, é importante evitar a introdução de indivíduos geneticamente diferentes, para não comprometer adaptações locais. Já para populações recentemente isoladas por causas antropogênicas, recomendam-se estratégias de reconexão de habitats, como corredores de dispersão.

A topologia filogenética pode revelar que uma espécie ocupa uma posição basal na filogenia. Espécies com alta divergência (ramos compridos) de suas espécies-irmãs poderiam ser priorizadas em razão da sua importância à diversidade do grupo (*D. coriacea* no nosso exemplo). O lobo-marinho-peruano é um exemplo do uso da análise filogenética na conservação. Durante décadas, a espécie foi considerada uma população pertencente ao lobo-marinho-sul-americano. Utilizando um conjunto de análises genômicas, chegou-se à conclusão de que a espécie teria uma origem híbrida, com características genéticas, comportamentais e morfológicas únicas, que necessitam de atenção especial na sua conservação.

A identificação de um grupo monofilético é o primeiro passo para a identificação correta de ESUs para destinar esforços de conservação. Os próximos passos envolvem estimativas de diversidade e demografia da ESU em questão. O MEGA 11 também é capaz de realizar essas análises populacionais (estimativa de diversidade nucleotídica, índice de diferenciação, teste de seleção, etc.), sendo resultados essenciais para elaboração de planos de conservação.

Para guiar a discussão e a avaliação de aprendizagem, os alunos podem ser instruídos a:

²¹ Método que faz uma subamostragem das sequências e replica a análise para verificar se o resultado se repete sem a necessidade de analisar os dados completos.

- investigar os relacionamentos obtidos com dados genéticos e combiná-los com as informações disponíveis na internet sobre diferenças morfológicas e a distribuição geográfica de cada uma das espécies (ex: características da carapaça, tamanho corporal, textura da pele, habitat (costeiras ou de oceano aberto) e termorregulação;
- comparar as filogenias obtidas com os dois conjuntos de dados e descrever se houveram diferenças na topologia ou suporte dos ramos, bem como explicar a importância de testar diferentes tipos de dados;
- falar sobre o papel da genômica na delimitação de espécies, discutindo sobre as diferenças na utilização de poucos marcadores e genoma completo e as consequências de hibridização e radiações de linhagens nas filogenias;
- discutir também como a delimitação correta ou incorreta pode influenciar na conservação dessas espécies;
- refletir sobre as filogenias e se elas sozinhas fornecem informações suficientes para a tomada de decisões de conservação, ou que tipo de informações também devem ser levadas em consideração;
- caso os alunos queiram um **desafio**, é possível sugerir que eles explorem adicionar outras espécies de tartarugas na filogenia, o que pode ser feito acessando o GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) e procurando por outras espécies de tartarugas na ferramenta de busca do site (Figura 9). Depois de encontrar a espécie de interesse, atentar-se ao marcador que a sequência representa (Figura 10). De preferência, escolher os mesmos marcadores das análises anteriores. Quando encontrar a espécie e o marcador, clicar em “FASTA” (Figura 11). O aluno será redirecionado para uma nova página, na qual ele deve selecionar a região do cabeçalho (“>”) e a sequência (Figura 12). A porção copiada deve ser colada no arquivo contendo as outras espécies e o mesmo marcador (Figura 13), como no passo 3.3E. Depois da realização desses passos,

continuar com os seguintes passos, como descrito no exercício anterior, com o alinhamento (tópico 3.4) e a análise de dados (tópico 3.5). Realizar a mesma comparação e discussão. Comparar as filogenias entre os alunos.

Figura 9: Busca por uma espécie, ou grupo, e um marcador no site do GenBank.



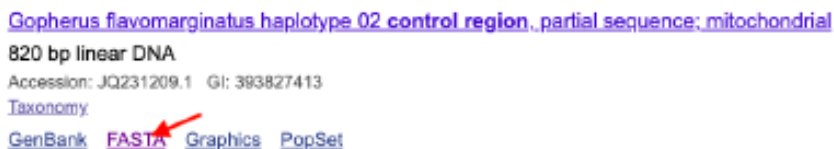
Fonte: imagem gerada pelos autores.

Figura 10: Checagem do marcador.



Fonte: imagem gerada pelos autores.

Figura 11: Abrir o arquivo FASTA da sequência.



Fonte: imagem gerada pelos autores.

Figura 12: Abrir o arquivo FASTA da sequência para copiar.

Região de controle do haplótipo 02 de *Gopherus flavomarginatus*, sequência parcial; mitocondrial

GenBank: JQ231209.1

[Gráficos](#) [GenBank](#) [PopSet](#)

```
>JQ231209.1 Região de controle do haplótipo 02 de Gopherus flavomarginatus,
sequência parcial; mitocondrial
CTAAACCCCATCCTCGGCCCCCAAGCCGAATTTTACTAACTACCTCTTGACCATACCTAACTTA
CTCTCCGCGCCCAAGAGATATATTTCCACTATACCTATGTAATTTACATTCATTTATTAC
ACTAAGATATATLACTAGTAAATATTTALLLTTAALTTTACTAAATACATATTAALTTAAATAGTTTATCTGA
AAGATTTGTTAAGCATGAATATTTTACTTAATTCAGAGTTAATGATCTAAGGACATAAAGGACATAG
CATCAATATTTATCAATACAACATGACTATTAAGCAAGTATCAGTTGATTTATTAATTTAACTGATCAC
GAGAAATANGCAACCTTGTTAATAAGATACAATATTACCAAGTTTCAAGTCCAATGATTGATGGGTACA
TAACGTATCTATTCTGGCCTCTGGTTGTTTTTCAGGCACATGATTAACAAGTTTCATTCATTTATCT
TTAAAGAGGCTCTGGTTAATGTGTTCTATACATTAATTTAATTAACCTGGCTATTCATGTTCTTAATGCA
TATAGTAGCTTTTTTCTCTTCTGTTCTCAGGCCCTCATAACTGATACCTGCCAATTAACCTGAACCT
GGACCCAAATCTAMGTGATTAATCTTGCAGGTATGATATGATATTTGAAATTAATGCTTGTGAAGACAT
AAATTTTACAAATATTCTCAACAGATAACTTCAAACTATTTATTCACACACATATTATTTAAGCTAA
CCCCCGACCCCATAGACTAACAATAATCCGAATAGCTGACTACTCT
```

Fonte: imagem gerada pelos autores.

Figura 13: Colar a sequência junto às sequências do mesmo marcador das demais espécies de tartarugas.

```
>Lepidochelys olivacea_ind
CATCTACTTATTACGGTTAGCATATGATCAGTAATATTGTCGATTAAATTTGGCTTTAAACATAAAAAAT
CATTAAATTTACATAAACTGTTTTTACCAATGAATATTATAAGATGATAATAGTGAATGATATAGGA
CATAAAATTAACCATTAATCTCGACCATGAATATCGTCACAGTAATTTGGTTATTTCTTAGTTCAACTCA
TCACGAGAAATAAGCAATCCTTGTTAACAGATACAACTTACCAGTTTCAGGCCCATTAACCTACGACG
TACATAACTGATCTATTCTGGCCTCTGGTTGTTTTTTCAGGCACATTAAGGTAATGAAGTTTCATTCGTT
CTCTTTAAAGGCTTGTGGTTGCACGATAAATGAGTTCTATACATTGAATTTATAACCTGGCATACT

>Nidopressus
AATTAACATATCCTTTGACGAGGAATAAAGCGTCCAAACACAAAACCTACCTATATCCTCTGCCGTGC
CCAAACAGAGTAATGCGCTGCAATACCCAGCTATGTATTATTGTACATCTACTTATTTACCACATGCATATG
ATCGGTAATGTAGTCGATTAAATTTAACCTAAACATAAAACCATTAAGTACCATCAATATTCTAATTAC
ATGAATATTATGCAAGTAATCAAGGTGAATGATATAGGACATAAAATTAACCATTAATCTCAACCATG
AATATTGTCACAGTAATTTGGTTATTTCTTACTCTAGCTAATCAAGAGAAATAAGCAACCTTTGTTAATAA
GATACGATATTACAGTTTCAAGTCCATTTAACTGTGGCGTACATAAAATGATCTATTCTGGCCTCTGGT
TGCTTTTTCAGGCACATATAAGTAACAACTTCATTCTGTTTCTCTTTAAAGGCTTTGGTTGAATGAGT
TCTATCCATCAAAATTTA

>20231209.1 Reação de controle do haplótipo 02 de Gopherus flavomarginatus. sequência parcial: mitocondrial
CTAAACCCCCCTCTCGGCCCCAAAGCGGAATTTTTTACTAACTACCTCTTGACCATACCTAACTTA
CTCTCCGCGGCCCAAGAGATAAATTTCCACTATACTATGCTATGTAATTTTACATTCAATTTATTTACC
ACTAGCATATCAGTAGTAATATTACCGCTTAACCTTACTAAATACATAGTAACTGAATAGTTGTACTGA
AAGATTTGTTAAAGCATGAATATTATTTACTTAATTCAAAGTTAATGATCTAAGGACATAAAGGACATAG
CATCAAAATATTACAAATACAACATGACTATTAAAGCAAGTATCAGTTGATTTATTAAATTTAACTGACAC
GAGAAATAAGCAACCTTGTTAATAGATACAATATTACCAAGTTTCAAGTCCAATGATTGATGGCGTACA
TAACTGATCTATTCTGGCCTCTGGTTGTTTTTTCAGGCACATGATATAACAAAGTTTCATTCTATTATCT
TTAAAGGCCCTCTGGTTAATGTGTTCTATACATTAATTTAACTTGGCTATTCATGTCTTAAATGCA
TATAGTAGCTTTTTTCTCTTTCTGTTCTCAGGCCCTCATAACTGATACCTGCCAATTAACGTAAACT
GGACCCAAATCTAAGTGATTAATCTTGCAAGTTATGATATGGTATTATTGAATTAATGCTTTGAAGACAT
AAATTTTACAAATATTCTCAACAGATAACTTCAAACATATTATTCAACACATATTATTTAAAGTAA
CCCCCCCCCATAGACTAAACACTAATCCGAATAGCTGACTACTCT
```

Fonte: imagem gerada pelos autores.

4. Considerações finais

É possível reconhecer diferentes espécies através da aplicação de dados genômicos com análises de reconstrução filogenética. A delimitação imprecisa de espécies leva à má gestão e ao desperdício de recursos para a conservação, ao proteger equivocadamente espécies não ameaçadas. Por isso, é importante identificarmos a real classificação das ESUs e, logo, o *status* de conservação mais aproximado da realidade da espécie.

A utilização de dados genômicos permite desvendar histórias evolutivas complexas que não são possíveis de compreender quando utilizamos apenas alguns *loci* ou aspectos morfológicos isoladamente.

Referências

- CEBALLOS, Gerardo; EHRLICH, Paul R. Mutilation of the tree of life via mass extinction of animal genera. *PNAS*, v. 120, n. 39, p. 6, set. 2023. DOI: 10.1073/pnas.2306987120.
- FRANKHAM, Richard; BALLOU, Jonathan. D.; BRISCOE, David A. *Fundamentos de Genética da Conservação*. Ribeirão Preto: Sociedade Brasileira de Genética, 2008.
- INTERNATIONAL UNION FOR CONSERVATION OF NATURE AND NATURAL RESOURCES (IUCN). *IUCN Red List Categories and Criteria*: Version 3.1. 2. ed. Gland: IUCN, 2012.
- HAMANN, Mark *et al.* Global research priorities for sea turtles: informing management and conservation in the 21st century. *Endangered Species Research*, v. 11, n. 3, p. 245-269, maio 2010.
- HOGG, Carolyn J. Translating genomic advances into biodiversity conservation. *Nature Reviews Genetics*, [S. l.], nov. 2023. DOI: 10.1038/s41576-023-00671-0.
- HOHENLOHE, Paul A.; FUNK, W. Chris; RAJORA, Om P. Population genomics for wildlife conservation and management. *Molecular Ecology*, [S. l.], v. 30, n. 1, p. 62-82, nov. 2021. DOI: 10.1111/mec.15720.
- JACKSON, Jeremy B. C. Historical Overfishing and the Recent Collapse of Coastal Ecosystems. *Science*, v. 293, n. 5530, p. 629-637, jul. 2001.
- LOPES, Fernando *et al.* Genomic evidence for homoploid hybrid speciation in a marine mammal apex predator. *Science advances*, v. 9, n. 18, maio 2023. DOI: 10.1126/sciadv.adf6601.
- MATIOLI, Sergio R.; FERNANDES, Flora M. C. *Biologia molecular e evolução*. 2. ed. Ribeirão Preto: Holos, Editora, 2012.
- MUSHER, Lukas J. *et al.* River network rearrangements promote speciation in lowland Amazonian birds. *Science Advances*, v. 8, n. 14, p. eabn1099, 2022.
- RIDLEY, M. *Evolução*. 3. ed. Porto Alegre: Artmed, 2013.
- SEGELBACHER, Gernot *et al.* New developments in the field of genomic technologies and their relevance to conservation management. *Conservation Genetics*, [S. l.], v. 23, n. 2, p. 217-242, nov. 2022. DOI: 10.1007/s10592-021-01415-5. Disponível em: <https://doi.org/10.1007/s10592-021-01415-5>.
- SUPPLE, Megan A.; SHAPIRO, Beth. Conservation of biodiversity in the genomics era. *Genome Biology*, [S. l.], v. 19, n. 1, p. 1-12, set. 2018. DOI: 10.1186/s13059-018-1520-3.

TRINDADE, Fernanda J., *et al.* Genome-wide SNPs clarify a complex radiation and support recognition of an additional cat species. *Molecular Biology and Evolution* v. 38, n. 11, nov. 2021. DOI: 10.1093/molbev/msab222.

WALLACE, Bryan P. *et al.* Global Conservation Priorities for Marine Turtles. *PLoS ONE*, v. 6, n. 9, p. e24510, set. 2011.

WASSWA, Fredrickson B.; KASSAZA, Kennedy; NIELSEN, Kirsten; BAZIRA, Joel. MinION Whole-Genome Sequencing in Resource-Limited Settings: Challenges and Opportunities. *Current Clinical Microbiology Reports*, [S. l.], v. 9, n. 4, p. 52-59, set. 2022. DOI: 10.1007/s40588-022-00183-1. Disponível em: <https://doi.org/10.1007/s40588-022-00183-1>.

Coleções biológicas: os bancos de dados da biodiversidade



Taciane Schröder¹, Fernando Fernandes¹ &
Julia Gabriele Dani¹

¹Programa de Pós-graduação em Botânica, Universidade
Federal do Rio Grande do Sul, Porto Alegre (Brasil)

1. Contextualização

Os bancos de dados da biodiversidade armazenam informações extremamente valiosas para diversas pesquisas e estudos, além de serem utilizados para tomada de decisão em órgãos públicos (Briggs, 2006). Dentre suas diversas aplicações, destacam-se a utilização para elaborar e monitorar ações de conservação da biodiversidade (Pougy *et al.*, 2014), a produção de listas de espécies ameaçadas de extinção (Martinelli *et al.*, 2014), a análise da distribuição das espécies (Barros *et al.*, 2012), os estudos evolutivos dos táxons (Meineke *et al.*, 2018), além de servirem como instrumento para o ensino e a divulgação científica (Kawakita, 2022).

Coleções biológicas têm como objetivo registrar a diversidade biológica e servir de referência para taxonomistas (Ariño, 2010). São, assim, umas das fontes mais importantes de conhecimento sobre os organismos vivos, disponibilizando diversas informações sobre eles, desde a composição dos ambientes até os dados sobre a distribuição e a variação da biodiversidade ao longo do tempo (Peixoto *et al.*, 2009).

A digitalização das coleções ampliou as colaborações internacionais e o acesso aos países localizados nos trópicos aos dados que antes estavam disponíveis apenas em museus e herbários da América do Norte e Europa (Meineke *et al.*, 2018). Devido à falta de recursos, durante muito tempo pesquisadores dos países neotropicais não conseguiam acessar esses materiais, o que impactava

diretamente suas pesquisas. Hoje em dia, muitos trabalhos são facilitados pela eliminação da necessidade de deslocamento geográfico até as coleções para obter as informações desejadas. Um exemplo de sucesso na questão da repatriação é o projeto Plantas do Brasil: Resgate Histórico e Herbário Virtual para o Conhecimento e Conservação da Flora Brasileira (REFLORA), cujo principal objetivo é digitalizar e repatriar espécimes de plantas brasileiras que se encontram depositadas em museus de outros países (BFG, 2015). A repatriação dos exemplares é um passo fundamental para garantir a preservação da maior flora nativa do mundo, que o Brasil abriga, ajudando a evitar a biopirataria e a promover o uso sustentável das espécies, contribuindo para que o Brasil mantenha seu patrimônio natural.

Com o avanço da tecnologia e da diversificação das formas de disponibilização de dados, atualmente existem inúmeras plataformas que oferecem uma vasta gama de informações sobre a biodiversidade. Um exemplo notável é a base de dados on-line *Global Biodiversity Information Facility* (GBIF), que permite o acesso a uma grande quantidade de dados sobre diversas espécies. Essa facilidade de acesso a registros aprimora e possibilita a realização de uma ampla variedade de estudos com maior precisão.

Além disso, o GBIF é uma rede internacional de dados que fornece acesso aberto a informações disponíveis sobre todos os tipos de vida no planeta (GBIF, [2024]). Dados globais como esses são essenciais para reduzir ao máximo possível vieses nas pesquisas. Ao realizar um trabalho utilizando apenas registros de determinado local, o pesquisador fica restrito a esse banco de dados, que muitas vezes pode ser limitado, e fica sujeito ao viés de coletor. O viés de coletor é a tendência de amostragem em se limitar a locais próximos a rotas de acesso (Oliveira *et al.*, 2016).

Entretanto, mesmo esses registros globais demandam um cuidado especial ao serem utilizados devido a erros de digitação, pois podem ser encontrados registros com coordenadas incorretas. Um exemplo são os pontos de coordenadas geográficas que podem

ser encontrados no meio do oceano para plantas terrestres (Zizka *et al.*, 2019).

Os registros de ocorrência são fundamentais para a documentação da biodiversidade. Neste capítulo, portanto, utilizaremos dados disponíveis na plataforma GBIF para criar um mapa de distribuição de uma determinada espécie, comparando os dados globais com os regionais. A partir deles, o professor poderá discutir com a turma a importância do acesso aos dados globais e seus impactos no resultado final das análises.

2. Objetivo

Capacitar os alunos a utilizar dados de ocorrência de espécies disponíveis em plataformas de biodiversidade para criar mapas de distribuição e realizar análises comparativas entre dados regionais e globais.

3. Descrição da atividade

A atividade está dividida em três etapas: (I) a obtenção dos dados de distribuição global e regional para a espécie de interesse; (II) a criação do mapa de distribuição utilizando um script em linguagem R; e (III) a análise e a discussão dos dados obtidos. A primeira etapa tem como objetivo familiarizar os alunos com as plataformas GBIF e *speciesLink* no processo de extração de dados. Assim, os alunos terão mais conhecimento das plataformas e poderão obter dados de qualquer espécie de interesse que disponha de registros on-line. A segunda etapa tem como foco ensinar aos alunos a utilizar a linguagem R para criar mapas de distribuição a partir dos dados obtidos. Nessa etapa será fornecido um *script* básico para a criação de mapas de distribuição e a explicação de cada parte do código. Além disso, será possível introduzir os conceitos básicos da linguagem R e a instalação dos pacotes necessários para a elaboração dos mapas. A terceira etapa será o momento de desenvolver habilidades analíticas nos alunos, permitindo-lhes interpretar e discutir os resultados obtidos. Nesse momento será

importante orientar os alunos na análise dos mapas de distribuição gerados, destacando os padrões de distribuição e as possíveis discrepâncias entre os dados globais e os regionais. Ademais, será possível promover uma discussão em grupo sobre os fatores que podem influenciar a distribuição das espécies, como o clima, o habitat e as atividades humanas.

3.1. Explorando a biodiversidade na prática

A primeira etapa consiste em obter as tabelas com dados de distribuição para a espécie-alvo. Vamos utilizar duas plataformas para esta etapa: GBIF e *speciesLink*. Para o nosso exemplo, utilizaremos a espécie *Calydorea campestris* (Klatt) Baker. Os dados obtidos a partir do GBIF serão referenciados como “Conjunto de Dados 1”. O passo a passo para baixar os dados de distribuição para *Calydorea campestris* na plataforma GBIF encontra-se na primeira etapa do *script* disponibilizado no Anexo I.

Sugerimos que, primeiramente, sejam baixados os dados do *speciesLink* (Conjunto de Dados 2). Em seguida, utilize o *script* disponível em <https://github.com/EbookBioinformaticaBiodiversidade/Cap5> para baixar os dados do GBIF. Após obter os dois conjuntos de dados, prossiga com a confecção dos mapas de distribuição, comparando as informações de ambos. A parte do *script* referente à criação do mapa está disponível na segunda etapa.

Passo a passo para baixar dados do *speciesLink*: (chamaremos os dados obtidos dessa plataforma de “Conjunto de Dados 2”):

- abra seu navegador e vá para o site www.specieslink.net;
- faça login, se você já tem uma conta no *speciesLink*, ou, se não, cadastre-se para poder realizar o download dos arquivos;
- clique no botão “busca”, que irá abrir a janela de filtros
- na área de identificação, clique na barra “nome científico” e insira o nome da espécie *Calydorea campestris*;

- na área de coleção, insira o acrônimo de um herbário, como, por exemplo, ICN (Herbário ICN – UFRGS), e selecione a caixa de marcação;
- clique no botão “REGISTROS” para observar as coletas registradas do espécime;
- para realizar o download, clique no botão “BAIXAR” – no formato do arquivo, clique em “Excel”, na opção de conjunto de campos, clique em “para Maxent”, nos campos *darwin core*, utilizaremos para a atividade prática os campos de “scientificname”, “latitude” e “longitude”;
- após, clique em “CONFIRMAR”;
- você receberá um e-mail com um link para baixar os dados quando o processamento estiver concluído.

3.2. Do download à visualização: criação de mapas de distribuição com R

Em plataformas de dados como GBIF e *speciesLink*, além de aprender a manipular os dados, na segunda etapa faremos uso das planilhas de distribuição que baixamos na etapa anterior. Para isso, será necessário que os alunos instalem o R e o RStudio (<https://www.r-project.org/>). O *script* deverá ser rodado no RStudio. Além disso, é necessário que os alunos tenham o conjunto de dados salvo no computador. O *script* que servirá como base para a análise encontra-se disponível em: <https://github.com/EbookBioinformatica/Biodiversidade/Cap5>. Os pacotes necessários estão listados nas primeiras linhas do código e podem ser instalados utilizando a função `install.packages()`, a qual baixa e instala pacotes do repositório oficial do R ou de outras fontes. Após a instalação, os pacotes não são carregados automaticamente, e para utilizá-los no seu *script* é necessário carregá-los com a função `library()`, que torna as funções e recursos do pacote disponíveis para uso no ambiente de trabalho.

Após isso, abriremos o mapa mundial com o comando `data(wrld_simpl)`. Logo em seguida, iremos plotar o mapa com o comando

`plot(wrld_simpl)`). Com o comando `read.csv` e `read_excel`, iremos carregar as planilhas de distribuição. Lembre-se de copiar o endereço do local onde estão armazenadas as planilhas. Na etapa “ajustando o mapa” será possível editar as cores do mapa e dos pontos de distribuição para o conjunto de dados. Após as edições necessárias, é esperado que o mapa de distribuição dos conjuntos de dados seja obtido.

3.3. Análise e discussão dos dados

Após a execução do *script* com os dois diferentes conjuntos de dados, o docente, juntamente com os alunos, poderá comparar os mapas de distribuição gerados. Com base no Conjunto de Dados 2, que contém dados de amostras de apenas um herbário, pode-se inicialmente concluir que a espécie *Calydorea campestris* é restrita à Região Sul do Brasil e que pode estar ameaçada (Figura 1). No entanto, ao visualizar o mapa gerado a partir do Conjunto de Dados 1 (Figura 2), é possível observar que a espécie possui muitos registros em várias regiões, incluindo São Paulo. Dessa forma, conclui-se que *Calydorea campestris* não é uma espécie endêmica da Região Sul do Brasil e que apresenta um número maior de registros, não estando, assim, ameaçada de extinção.

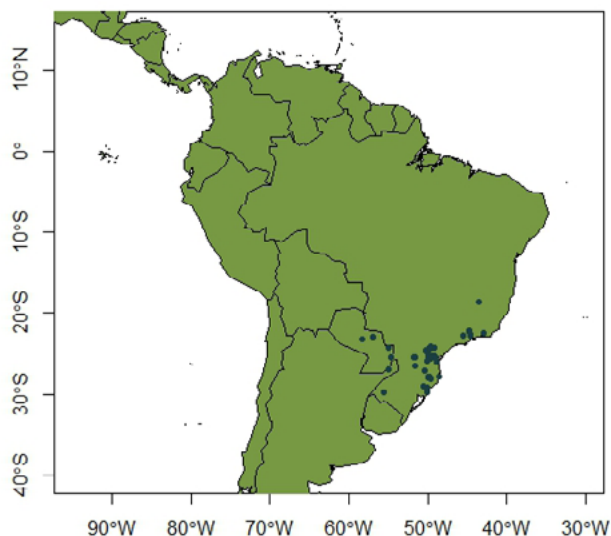
Além disso, com esse tipo de atividade é possível estudar os efeitos da fragmentação do habitat e discutir estratégias para mitigar seus impactos na biodiversidade. Essas atividades podem ser adaptadas para atender os objetivos específicos do curso e o nível de conhecimento dos alunos.

Figura 1: Mapa de distribuição de *Calydorea campestris* com base nos dados obtidos na plataforma *speciesLink*, referente aos registros disponíveis no Herbário ICN (Conjunto de Dados 2).



Fonte: imagem gerada pelos autores.

Figura 2: Mapa de distribuição de *Calydorea campestris* com base nos dados obtidos na plataforma GBIF (Conjunto de Dados 1).



Fonte: imagem gerada pelos autores.

4. Considerações finais

A atividade proposta visa capacitar os alunos no uso de dados de ocorrência de espécies para criar mapas de distribuição e realizar análises comparativas, além de desenvolver uma série de habilidades essenciais para a pesquisa científica em biodiversidade. Ao longo das três etapas, os alunos têm a oportunidade de se familiarizar com importantes comandos da linguagem R.

A familiarização com a extração de dados de diferentes fontes proporciona aos alunos uma compreensão prática sobre a disponibilidade e o uso de registros de ocorrência, ampliando sua capacidade de realizar estudos com base em dados reais. A criação de mapas de distribuição, por sua vez, não só ensina técnicas de programação e visualização de dados como também reforça a importância de ferramentas computacionais na biologia.

A etapa final de análise e discussão dos resultados é crucial para desenvolver o pensamento crítico e as habilidades analíticas dos alunos. Interpretar padrões de distribuição e identificar discrepâncias entre dados globais e regionais promove uma compreensão mais profunda sobre os fatores que influenciam a distribuição das espécies. Esse processo de análise e discussão também encoraja a colaboração e a troca de ideias entre os alunos, fortalecendo suas competências em trabalho em equipe e comunicação científica.

Em conclusão, a atividade, além de atingir o objetivo de capacitar os alunos na utilização de dados de ocorrência de espécies, também os prepara para enfrentar desafios reais na pesquisa em biodiversidade. Ao integrar a coleta, a análise e a interpretação dos dados, os alunos estarão mais preparados para contribuir significativamente na conservação da biodiversidade e nas futuras pesquisas científicas.

Referências

ARIÑO, A. H. Approaches to estimating the universe of natural history collections data. *Biodiversity Informatics*, v. 7, n. 2, 2010.

BARROS, F. S. M. *et al.* Modeling the potential geographic distribution of five species of Metzgeria Raddi in Brazil, aiming at their conservation. *The Bryologist*, v. 115, n. 2, p. 34-349, 2012.

THE BRAZIL FLORA GROUP (BFG). Growing knowledge: an overview of seed plant diversity in Brazil. *Rodriguésia*, v. 66, n. 4, p.1085-1113, 2015.

BRIGGS, S. V. Integrating policy and science in natural resources: why so difficult?. *Ecological Management & Restoration*, v. 7, n. 1, p. 37-39, 2006.

GLOBAL BIODIVERSITY INFORMATION FACILITY (GBIF). Página inicial. [2024]. Disponível em: <http://www.gbif.org>. Acesso em: 27 jun. 2024.

KAWAKITA, K. Coleção botânica e a popularização do conhecimento. *Arquivos do Mudi*, v. 27, n. 2, p.74-87, 2022.

MARTINELLI, G. MORAES, M. A. *Livro vermelho das plantas do cerrado*. Rio de Janeiro: Instituto de Pesquisas Jardim Botânico do Rio de Janeiro, 2013. *E-book*. Disponível em: <https://dspace.jbrj.gov.br/jspui/handle/doc/26>. Acesso em: 27 jun. 2024.

MEINEKE, E. K. *et al.* Biological collections for understanding biodiversity in the Anthropocene. *Philosophical Transactions of the Royal Society B*, v. 374, n. 1763, p. 20170386, 2019.

OLIVEIRA, U. *et al.* The strong influence of collection bias on biodiversity knowledge shortfalls of Brazilian terrestrial biodiversity. *Diversity and Distributions*, v. 22, n. 12, p. 1.232-1.244, 2016.

PEIXOTO, A. L. *et al.* Coleções Botânicas: objetos e dados para a ciência. In: GRANATO, M.; RANGEL, M. (orgs.). *Cultura Material e Patrimônio de C&T*. Rio de Janeiro: Museu de Astronomia e Ciências Afins, 2009.

POUGY, N. *et al.* Urban forests and the conservation of threatened plant species: the case of the Tijuca National Park, Brazil. *Natureza & Conservação*, v. 2, n. 12, p.170-173, 2014.

SPECIESLINK NETWORK. Página Inicial. [2024]. Disponível em: www.specieslink.net. Acesso em: 27 jun. 2024.

Zizka, A. *et al.* CoordinateCleaner: Standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, v. 10, n. 5, p. 744-751, 2019.

EVOLUÇÃO



Olhando para o presente e desvendando o passado com macroevolução



Patrícia Sperotto¹ & Sandra Reinales²

¹Programa de Pós-graduação em Botânica, Universidade Federal do Rio Grande do Sul, Porto Alegre (Brasil)

²Programa de Pós-graduação em Botânica, Universidade de São Paulo, São Paulo (Brasil)

1. Contextualização

Historicamente, a evolução pode ser abordada de duas perspectivas diferentes: estudando os processos que geram mudanças evolutivas ao nível intraespecífico, como entre populações, ou estudando a trajetória das espécies ou táxons de classificação superior ao longo do tempo geológico (Mayr, 1982; Gould, 1985). O termo *macroevolução* foi introduzido pelo botânico e entomólogo Yuri Filipctshenko (1927), para se referir justamente à evolução dos táxons acima do nível de espécie na hierarquia Lineana. No entanto, Dobzhansky (1937, p) adicionou a perspectiva de tempo ao conceito, sugerindo que as mudanças macroevolutivas requerem tempo numa escala geológica. A definição de macroevolução baseada em uma escala temporal permite a incorporação de todos os processos que afetam os padrões de evolução a longo prazo, desde interações bióticas até mudanças ambientais globais (Hautmann, 2019).

Embora a discussão sobre a separação entre essas duas perspectivas da evolução seja vasta na literatura – microevolução *vs.* macroevolução (Futuyma, 2015) –, um dos principais objetivos da macroevolução na atualidade é entender como mudanças em escala geológica como orogenia, erosão e flutuações climáticas, além da aquisição de novidades morfológicas, modelam a distribuição e diversidade das espécies (Vamosi *et al.*, 2018). Tradicionalmente, as hipóteses nessa escala eram descritivas, suportadas pela

ocorrência de padrões (e.g., diferentes grupos seguindo a mesma trajetória evolutiva) e defendidas de forma argumentativa pelos pesquisadores que se baseavam em observações sobre a distribuição geográfica e o conjunto de caracteres presentes nas espécies viventes e fósseis. Nas décadas de 1970 e 1980, métodos de correlação e regressão foram amplamente usados nos trabalhos em evolução, visando testar o efeito da variação de uma variável contínua sobre outra (e.g., tamanho corporal e densidade populacional) (Damuth, 1981). No entanto, dada a natureza histórica das hipóteses macroevolutivas e a dependência hierárquica entre as espécies, percebeu-se que entender e considerar as relações de parentesco entre os grupos de interesse é imprescindível, sendo este um passo-chave no estabelecimento da biologia evolutiva e da biogeografia moderna (Felsenstein, 1985).

O desenvolvimento da sistemática filogenética baseada no princípio de homologia da Cladística (Hennig *et al.*, 1966), junto à aplicação de métodos probabilísticos como a máxima verossimilhança (Edwards; Cavalli-Sforza, 1964; Felsenstein, 1981) ou a estatística bayesiana (Rannala; Yang, 1996) na inferência das árvores filogenéticas, além da disponibilidade de ferramentas computacionais cada vez mais poderosas (e.g., computação paralela), permitiu importantes avanços nos estudos macroevolutivos. São alguns exemplos o uso de grandes quantidades de dados morfológicos e moleculares, assim como a inclusão de diferentes fontes de evidência (e.g., fósseis) (Heath; Huelsenbeck; Stadler, 2014); a modelagem de processos evolutivos como especiação, extinção e diversificação, taxas de preservação no registro fóssil (Silvestro; Salamin; Schnitzler, 2014) e processos de evolução molecular como hibridização e duplicação gênica (Solís-Lemus; Bastide; Ané, 2017); a inferência de “*time trees*”, em que os comprimentos dos ramos representam tempo absoluto (Rannala; Yang, 2007); e o mapeamento da distribuição das espécies no tempo e no espaço (Hijmans; Van Etten, 2015). Essa crescente complexidade das análises macroevolutivas permite testar estatisticamente hipóteses clássicas sobre

a origem e diversificação da biota como o gradiente latitudinal na riqueza de espécies, “berçários” e “museus” da biodiversidade (Stebbins, 1974), ou o “abominável mistério” da origem das angiospermas (Darwin; Darwin; Seward, 1903; Vamosi *et al.*, 2018).

Também, a disponibilidade de repositórios de acesso livre contendo dados morfológicos (e.g., MelastomaTRAITS) (Reginato *et al.*, 2024), moleculares, taxonômicos, geográficos e climáticos revolucionou os estudos em ecologia e evolução. Porém, foi a incorporação de rotinas automatizadas de acesso e processamento desses dados, juntamente com o contínuo desenvolvimento de pacotes em diferentes linguagens de programação como R (e.g., CoordinateCleaner (Zizka *et al.*, 2019) e taxize (Chamberlain *et al.*, 2020) e técnicas de Machine Learning (e.g., Hussein *et al.*, 2022), que permitiu na última década a compilação de grandes “datasets” visando responder perguntas cada vez mais abrangentes e multidisciplinares. Ainda, a popularização das estratégias de ciência reprodutível mudou a forma como coletamos, processamos e analisamos dados, permitindo a geração de conhecimento científico mais robusto mediante a disponibilização dos dados e a documentação dos métodos usados em um estudo. Repositórios de código aberto como o GitHub são uma ferramenta importante para atingir esse objetivo. No início pode ser assustador se deparar com a quantidade de técnicas e ferramentas novas que precisamos aprender para desenvolver pesquisa moderna em biologia evolutiva, mas recursos on-line como Stack Overflow e as “wiki pages” dos diversos programas facilitam o aprendizado coletivo e autônomo. Toda essa conjuntura contribui para que a macroevolução venha se tornando cada vez mais presente e popular no âmbito da biologia evolutiva.

Uma abordagem central em estudos macroevolutivos é inferir as características de linhagens ancestrais com o intuito de compreender como essas linhagens mudaram através do tempo e originaram os descendentes que observamos no presente. Isso pode ser feito através de reconstruções de caractere ancestral. Essas análises consistem na inferência das características de linhagens an-

cestrais a partir das características observadas em linhagens atuais ou extintas, e são muito úteis para testarmos hipóteses relativas à evolução e adaptação dos organismos (Pagel, 1999), como, por exemplo, no caso das flores da tribo Miconieae (Melastomataceae) (Gavrutenko *et al.*, 2020) ou das mudanças na morfologia das glândulas de espécies carnívoras de Caryophyllales (Renner; Specht, 2011).

As primeiras reconstruções foram feitas baseadas no princípio da parcimônia, ou seja, em minimizar o número de mudanças entre estados de um caractere para produzir o padrão de apresentação do caractere que é observado nas linhagens atualmente (Maddison, 1990; Currie; Meade, 2014). Porém, os métodos baseados em parcimônia assumem que todas as mudanças entre estados do caractere têm a mesma chance de ocorrer, isto é, que as taxas de transição são as mesmas. Portanto, essa é uma abordagem considerada limitada, por isso há a preferência por outros métodos. Já métodos de máxima verossimilhança e estatística bayesiana funcionam baseando-se em modelos de evolução de caractere, o que permite a inclusão de diferentes taxas de transição (ou seja, permite mais e diferentes parâmetros) e o uso da informação dos comprimentos de ramo da filogenia nas reconstruções (Pagel 1999; Currie; Meade, 2014), o que torna esses métodos mais robustos. A máxima verossimilhança nos retorna o conjunto de valores dos parâmetros dentro do modelo evolutivo (e, portanto, apenas uma reconstrução) que maximiza a probabilidade de se obterem os dados observados. Por outro lado, o mapeamento estocástico baseado em estatística bayesiana (Huelsenbeck; Nielsen; Bollback., 2003) nos retorna a probabilidade posterior do conjunto de parâmetros a partir de várias reconstruções, incorporando, assim, a incerteza na estimativa dos parâmetros de interesse (Currie; Meade, 2014). Cada reconstrução pode ser apresentada separadamente ou sumarizada e apresentada como uma única reconstrução mostrando as probabilidades de cada estado do caractere nos ancestrais. Essa é a abordagem que usaremos na atividade descrita a seguir.

2. Objetivo

O objetivo deste capítulo é discutir como a bioinformática está inserida no desenvolvimento das pesquisas em macroevolução, além de exercitar uma das análises mais recorrentes dentro do campo, que é a reconstrução de caractere ancestral.

3. Descrição da atividade

A atividade prática consistirá em realizar uma reconstrução de caractere ancestral do número de peças do perianto (pétalas e/ou sépalas, ou tépalas) das angiospermas apenas com o intuito de explorar como teria sido o ancestral do grupo com relação a essa característica. Portanto, reconstruiremos um caractere discreto.

A árvore filogenética que usaremos contém uma amostra aleatória de 50 espécies da filogenia datada (*time tree*) de Smith e Brown (2018), uma das maiores filogenias de plantas com sementes já construídas, incluindo quase 80 mil espécies. Via de regra, usamos filogenias datadas porque elas nos permitem fazer inferências sobre as condições climáticas e geológicas à época das mudanças de estado de caractere e levantar hipóteses sobre o porquê de estas terem acontecido, mas é possível utilizar filogenias não datadas em alguns casos (Cusimano; Renner, 2014). Os dados sobre o número de peças do perianto foram obtidos da Flora e Funga do Brasil ([2024]), Plants of the World Online (POWO, [2024]), APG IV (Stevens, 2017) e da análise on-line de exsicatas do herbário US (<https://collections.nmnh.si.edu/search/botany>). Alguns estados de caractere foram codificados de forma a simplificar a nossa análise (como no caso de espécies cujo número de peças do perianto varia entre flores “masculinas” e “femininas”, em que foi escolhido apenas um dos números para representar a espécie como um todo). Por fim, o método de reconstrução que usaremos é o de mapeamento estocástico (Huelsenbeck; Nielsen; Bollback *et al.*, 2003) e a análise será implementada no R (R Core Team, 2023) (instruções para a instalação do programa em: <https://www.r-project.org/>).

Os dados usados podem ser encontrados em <https://github.com/EbookBioinformaticaBiodiversidade/Cap6>.

3.1. *Leitura dos pacotes necessários, dos dados, e visualização*

A primeira coisa a se fazer é instalar e ler os pacotes necessários para realizar as análises e ler os dados. O *script* com o código a seguir, a planilha de dados morfológicos e a árvore devem estar na mesma pasta. Depois, é sempre bom visualizar os dados para conferir se estão corretos (Figura 1).

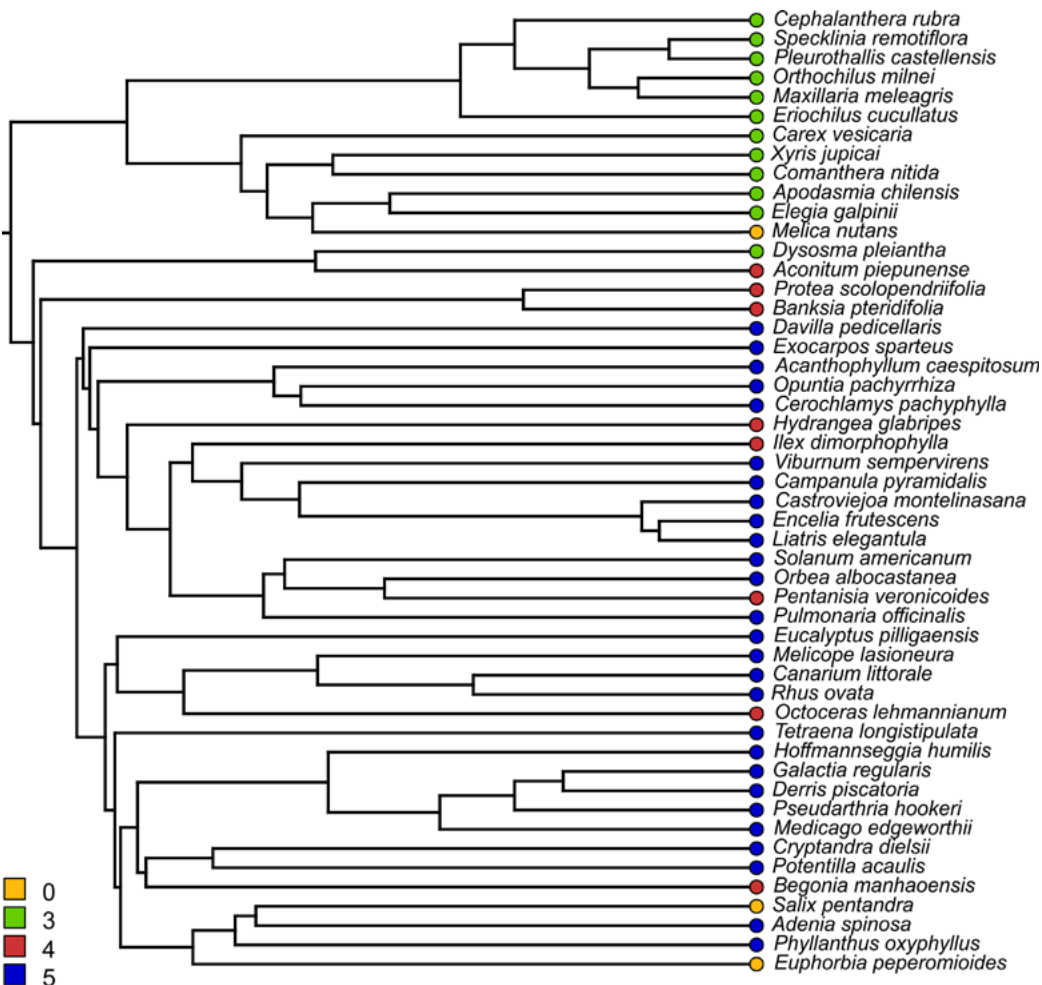
```
#### Instalar os pacotes necessários:
install.packages("ape")
install.packages("phytools")
install.packages("phangorn")
install.packages("geiger")

#### Ler os pacotes necessários:
library(phytools)
library(ape)
library(phangorn)
library(geiger)

#### Lendo a árvore e os dados de número de peças do perianto em cada espécie
arvore<-read.nexus(file="arvore.nex")
dados.reais<-read.csv("dados_reais.csv", header = T)

#### Visualizando dados
## Renomeando linhas de acordo com as espécies
rownames(dados.reais)<-dados.reais$especie
## Criando o vetor com nomes das espécies e estados de caráter
petalas<-setNames(as.factor(dados.reais[, "petalas"]), rownames(dados.reais))
## Delegando uma cor para cada estado de caráter
cols.r<-setNames(c("darkgoldenrod1", "chartreuse3", "brown3", "blue3"),
  levels(petalas))
## Plot da árvore com os dados
plotTree(arvore, fsize=0.9, ftype="i", lwd=2, offset=0.4)
tiplabels(pie=to.matrix(petalas[arvore$tip.label],
  levels(petalas)), piecol=cols.r, cex=0.4)
add.simmap.legend(colors=cols.r, prompt=FALSE, x=1, y=6, fsize=0.9)
```

Figura 1: Visualização dos dados na árvore filogenética. A seta preta indica o ancestral comum das Monocotiledôneas e Eudicotiledôneas.



Fonte: imagem gerada pelas autoras.

Perceba que a nossa filogenia é totalmente resolvida, isto é, cada ancestral sempre dá origem diretamente a duas linhagens. Isto é fundamental para uma boa análise de reconstrução. Cada cor corresponde a um número de peças do perianto; portanto, estamos trabalhando com um caractere que pode se apresentar em quatro

estados: 0 peças = aclamídea; 3 ou múltiplas de = trímera; 4 ou múltiplas de = tetrâmera; e 5 ou múltiplas de = pentâmera).

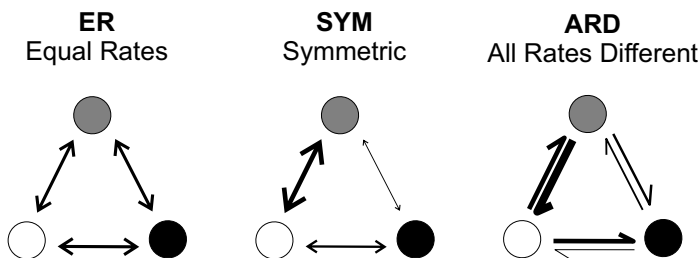
Olhando a distribuição dos estados de caractere na filogenia, poderíamos nos perguntar: 1) Considerando as Eudicotiledôneas, o que é mais provável: as flores com cinco peças terem evoluído a partir daquelas com quatro peças ou o contrário, isto é, que o seu ancestral tinha flores com cinco peças e as flores com quatro peças evoluíram independentemente? 2) O número de peças da corola das angiospermas evoluiu crescente e ordenadamente, isto é, linhagens mais cedo-divergentes apresentam zero peças e linhagens mais derivadas apresentam cinco? 3) Considerando as Monocotiledôneas, o perianto com zero peças é uma plesiomorfia ou uma apomorfia? Essas são apenas algumas questões que podemos explorar usando métodos filogenéticos comparativos como este.

3.2. *Testando modelos de evolução de caractere*

Para reconstruir um caractere numa filogenia, seja ele contínuo ou discreto, precisamos levar em consideração como este pode ter mudado ao longo do tempo. Para isso, testamos diferentes modelos de evolução. Para caracteres discretos, as mudanças entre estados são instantâneas, então falamos de transição entre estados de caractere. Portanto, os modelos de evolução devem levar em conta as taxas de transição entre estados, isto é, com que frequência o caractere muda de um estado para o outro. O número de taxas totais possíveis na análise depende do número de estados do nosso caractere: se é binário (tem apenas dois estados), teremos no máximo duas taxas de transição; se tem três estados, o número de taxas possíveis já aumenta para seis. Quanto mais estados, mais taxas possíveis e mais complexa fica a análise em termos de processamento, pois mais parâmetros vão ser estimados. O número de taxas depende também do modelo de evolução. Há três modelos básicos de evolução de caracteres discretos (Figura 2): *Equal Rates* (ER), em que é permitida apenas uma taxa de transição para transições entre todos os estados; *Symmetric* (SYM), em que as taxas

de ida e volta entre dois estados são as mesmas, mas variam a cada dupla de estados; e *All Rates Different* (ARD), em que é permitido que todas as taxas de transição entre quaisquer dois estados sejam diferentes. Como mostra a Figura 2, o ARD é o modelo mais complexo, pois permite a existência de seis taxas diferentes, enquanto o SYM permite três e o ER apenas uma, sendo o menos complexo.

Figura 2: Exemplo dos modelos básicos de evolução para um caractere com três estados (branco, cinza e preto). A grossura das linhas representa o tamanho das taxas.



Fonte: imagem gerada pelas autoras.

Vamos, agora, testar esses três modelos com os nossos dados usando a função `fitDiscrete()` do pacote *geiger* (Pennel *et al.*, 2014). O modelo melhor ajustado aos nossos dados será escolhido de acordo com o AICc, i.e., critério de Akaike (1987) corrigido para amostras pequenas, e usado na análise de reconstrução.

```
#### Testando 3 modelos básicos de evolução de caracter
```

```
fitER.r<-fitDiscrete(arvore,petalas, model = "ER")
```

```
fitSYM.r<-fitDiscrete(arvore,petalas, model = "SYM")
```

```
fitARD.r<-fitDiscrete(arvore,petalas, model = "ARD")
```

```
#### Avaliando qual o melhor modelo segundo o AIC
```

```
fitER.r$opt$aic
```

```
## [1] 82.33524
```

```
fitSYM.r$opt$aic
```

```
## [1] 87.94458
```

```
fitARD.r$opt$aic
```

```
## [1] 101.7732
```

De acordo com o resultado do teste, o modelo de evolução melhor ajustado aos nossos dados é o ER, pois ele apresentou o

menor valor de AICc, 82.33524, em comparação com 87.94458 do SYM e 101.7732 do ARD. Agora, estamos prontos para realizar a reconstrução.

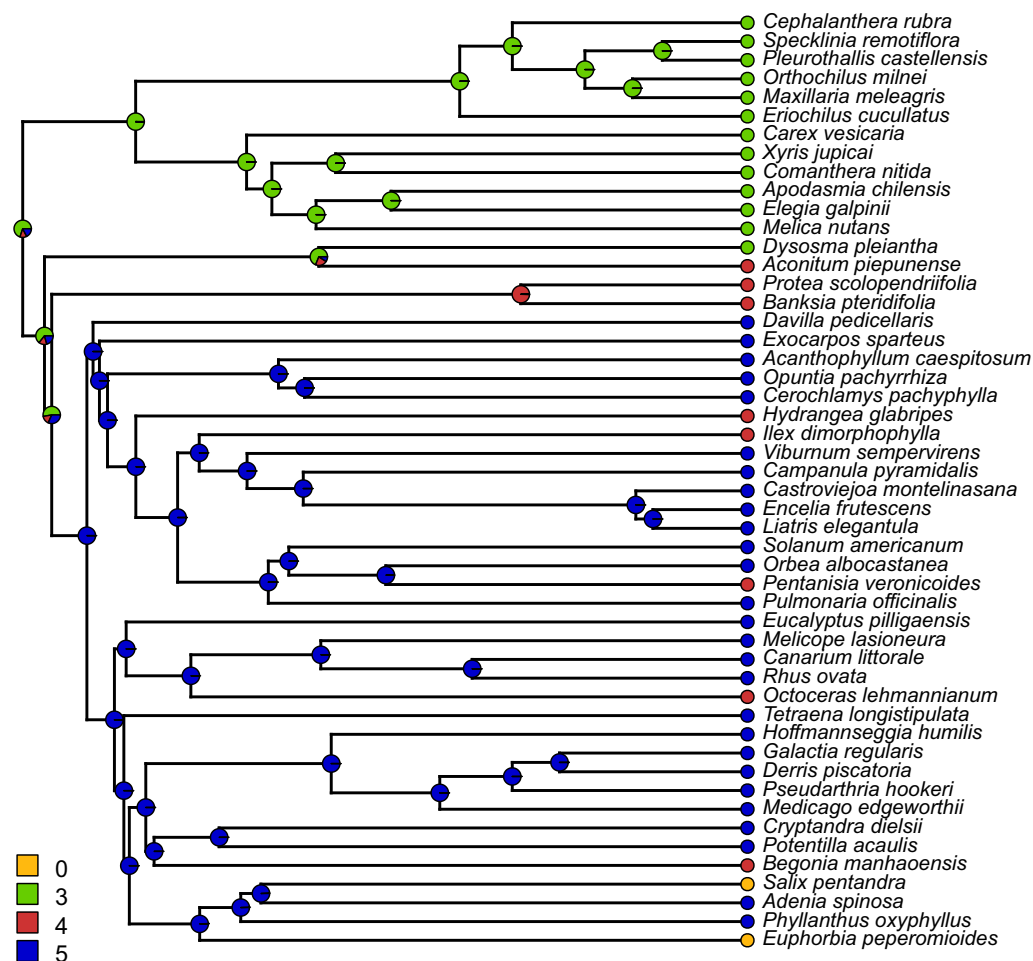
3.3. Reconstrução de caractere ancestral e discussão dos resultados

Faremos a reconstrução utilizando mapeamento estocástico implementado na função **make.simmap()** do pacote *phytools* (Revell, 2012). Como a nossa filogenia é pequena, podemos fazer apenas 100 simulações, mas conforme o número de espécies aumenta é necessário fazer mais simulações para que se tenha uma análise confiável. Visualizaremos o resultado das 100 reconstruções sumarizado em apenas uma reconstrução (Figura 3).

```
#### Reconstruindo o caractere ancestral com mapeamento estocástico
simmap.petalas<-make.simmap(arvore,petalas,model="ER",nsim=100)

#### Visualizando o resultado
sum.petalas<-summary(simmap.petalas)
plot(sum.petalas,fsize=0.9,ftype="i",colors=cols.r,
      ylim=c(-2,Ntip(arvore)))
add.simmap.legend(colors=cols.r,prompt=FALSE,x=0,
                  y=6,vertical=T)
```

Figura 3: Resultado da reconstrução de caracteres ancestral do número de peças do perianto das angiospermas. A seta preta indica o ancestral de todas as linhagens da filogenia.



Fonte: imagem gerada pelas autoras.

Os resultados da nossa análise indicam que o ancestral das angiospermas (seta preta, Figura 3) teria maior probabilidade de apresentar um perianto trímero (verde), seguido de um perianto pentâmero (azul) e, por último, uma pequena chance de um perianto tetrâmero (vermelho). Na realidade, a morfologia do ancestral

das angiospermas ainda é bastante debatida e incerta. Um fator importantíssimo que não consideramos na interpretação dos resultados da nossa análise, uma vez que este é apenas um exercício, é a inclusão de outras fontes de evidência, como do registro fóssil, por exemplo, dado que esse tipo de análise depende totalmente da amostragem. Diversos trabalhos empíricos mostraram que, embora as filogenias moleculares sejam úteis para testar hipóteses macroevolutivas, as inferências derivadas apenas de dados sobre as espécies viventes podem ser altamente imprecisas, devido, principalmente, às limitações dos métodos para modelar o processo de extinção (Leslie *et al.*, 2018; Lloyd; Slater, 2021; Coiro *et al.*, 2023). Também, se este fosse um estudo real, deveríamos atentar ao fato de que a nossa amostragem é extremamente pequena: estão amostradas apenas 50 espécies de angiospermas das cerca de 400.000 existentes (Stevens, 2017). Na esmagadora maioria das vezes, não será possível amostrar a totalidade das espécies de um grupo numa filogenia; porém, quanto mais espécies amostradas, mais perto chegaremos de desvendar a história evolutiva real daquele grupo. É importante lembrar que, independentemente do tipo de estudo, uma abordagem integrativa, com metodologias complementares e fontes diversas de evidência, torna nossos trabalhos mais robustos e embasados, contribuindo, assim, de forma mais sólida para o avanço do nosso conhecimento sobre o mundo que nos cerca. Para encerrar, podemos retomar as perguntas feitas a partir da Figura 1 e comparar as respostas dadas antes e depois da análise dos resultados.

4. Considerações finais

A reconstrução de caractere ancestral é apenas uma das diversas análises possíveis na área da macroevolução. Pode-se, por exemplo, estimar a diversificação de linhagens ao longo do tempo (e.g., BaMM; Rabosky, 2014), avaliar se essa diversificação está associada a algum caractere específico (métodos de *State Speciation and Extinction*) (Maddison; Midford; Otto, 2007; Beaulieu;

O'Meara, 2016) e até testar se caracteres contínuos e discretos evoluíram de maneira correlacionada (Boyko *et al.*, 2023). A macroevolução nos dá a oportunidade de fazer algumas das perguntas mais empolgantes dentro da biologia evolutiva, e, com o avanço no desenvolvimento de ferramentas bioinformáticas progressivamente mais robustas, é cada vez mais fácil respondê-las.

5. Para saber mais

5.1 Recursos on-line

- Tutorial do pacote *phytools* para análises filogenéticas comparativas: <http://blog.phytools.org>
- Aulas sobre métodos filogenéticos comparativos: <https://phylo-meth.info/syllabus.html>
- Mapa mental didático sobre métodos filogenéticos comparativos: https://coggle.it/diagram/W_hbkkxE2BAABOR0m/t/summary-of-phylogenetic-comparative-methods-diogo-b-provete
- Tutorial do pacote *ggtree* para manipulação de árvores filogenéticas: <https://bioconductor.statistik.tu-dortmund.de/packages/3.3/bioc/vignettes/ggtree/inst/doc/treeAnnotation.html>
- Curso on-line de R: <https://www.datacamp.com/learn/r>
- Curso online de R em português: <http://ecor.ib.usp.br/doku.php>

5.2. Literatura

- GARAMSZEGI, L. Z. (Ed.). *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*. New York: Springer, 2014.
- HARMON, L. J. *Phylogenetic comparative methods*. Traverse City, MI, USA: Independent. 2019.
- PARADIS, E. *Analysis of Phylogenetics and Evolution with R*. New York: Springer. 2012. v. 2.
- REVELL, L. J.; HARMON, L. J. *Phylogenetic comparative methods in R*. Princeton: University Press, 2022.

Referências

- AKAIKE, H. Factor analysis and AIC. *Psychometrika*, v. 52, p. 317-332, 1987.
- BEAULIEU, J.M.; O'MEARA, B.C. Detecting hidden diversification shifts in models of trait-dependent speciation and extinction. *Systematic Biology*, v. 65, n. 4, p. 583-601, 2016.
- BOYKO, J. D. *et al.* The evolutionary responses of life-history strategies to climatic variability in flowering plants. *New Phytologist*, v. 240, n. 4, p.1587-1600. 2023. DOI: <https://doi.org/10.1111/nph.18971>.
- CHAMBERLAIN, S. *et al.* taxize: Taxonomic information from around the web. *R package*, 2020. COIRO, M. *et al.* Reconciling fossils with phylogenies reveals the origin and macroevolutionary processes explaining the global cycad biodiversity. *New Phytologist*, v. 240, p.1616-1635, 2023.
- CURRIE, T. E.; MEADE, A. Keeping yourself updated: Bayesian approaches in phylogenetic comparative methods with a focus on Markov chain models of discrete character evolution. In: GARAMSZEGI, L. Z. (Ed.). *Modern phylogenetic comparative methods and their application in evolutionary biology: Concepts and practice*. New York: Springer, 2014, p. 263-286.
- CUSIMANO, N.; RENNER, S. S. Ultrametric trees or phylograms for ancestral state reconstruction: does it matter?. *Taxon*, v. 63, n. 4, p. 721-726, 2014.
- DAMUTH, J. Population density and body size in mammals. *Nature*, v. 290, p.699-700, 1981.
- DARWIN, C.; DARWIN, F.; SEWARD, A. C. *More Letters of Charles Darwin: A Record of his work in a series of hitherto unpublished Letters*. London: J. Murray, 1903.
- DOBZHANSKY, T. *Genetics and the origin of species*. New York: Columbia University Press, 1937.
- EDWARDS, A.W.F.; CAVALLI-SFORZA. Reconstruction of evolutionary trees. Phenetic and Phylogenetic Classification. *The Systematics Association Publication*, n. 6, p. 67-76, 1964.
- FELSENSTEIN, J. Evolutionary trees from DNA sequences: a maximum likelihood approach. *Journal of Molecular Evolution*, v. 17, n. 6, p.368-76, 1981.
- FELSENSTEIN, J. Phylogenies and the Comparative Method. *The American Naturalist*, v. 125, n. 1, p. 1-15, 1985.

- HUELSENBECK, J.P. Stochastic mapping of morphological characters. *Systematic biology*, v. 52, n. 2, p. 131-158, 2003.
- LESLIE, A; BEAULIEU, J.; HOLMAN, G., *et al.* An overview of extant conifer evolution from the perspective of the fossil record. *American Journal of Botany*, v. 105, n. 9, p. 1531-1544, 2018.
- FILIPTSCHENKO, Y. *Variabilität und Variation*. Berlin: Borntraeger, 1927.
- LLOYD, G.; SLATER, G. A Total-Group Phylogenetic Metatree for Cetacea and the Importance of Fossil Data in Diversification Analyses. *Systematic Biology*, v. 70, n. 5, p. 922-939, 2021.
- FLORA E FUNGA DO BRASIL. Página inicial. [2024]. Disponível em: <http://floradobrasil.jbrj.gov.br/>. Acesso em: 28 jun. 2024.
- FUTUYMA, D. J. Can Modern Evolutionary Theory Explain Macroevolution?. In: SERRELLI, E.; GONTIER, N. (Eds.). *Macroevolution: Interdisciplinary Evolution Research*. New York: Springer, Cham, 2015. v. 2.
- GAVRUTENKO, M. *et al.* Evolution of floral morphology and symmetry in the Miconieae (Melastomataceae): multiple generalization trends within a specialized family. *International Journal of Plant Sciences*, v. 181, n. 7, p. 732-747, 2020.
- GOULD, S. J. The Paradox of the First Tier: An Agenda for Paleobiology. *Paleobiology*, v. 11, n. 1, p. 2-12, 1985.
- HAUTMANN, M. What is macroevolution?. *Palaeontology*, v. 63, p. 1-11, 2020.
- HEATH, T. A.; HUELSENBECK, J. P.; STADLER, T. The fossilized Birth-Death Process: A coherent model of fossil calibration for divergence time estimation. *Proceedings of the National Academy of Sciences*, v. 111, n. 29, p. e2957-e2966, 2014.
- HENNIG, W. *et al.* *Phylogenetic systematics*. Champaign-Urbana: University of Illinois, 1966.
- HIJMANS, R. J.; VAN ETEN, J. raster: geographic data analysis and modeling. *R package*, 2015
- HUELSENBECK, J. P.; NIELSEN, R.; BOLLBACK, J. P. Stochastic mapping of morphological characters. *Systematic biology*, v. 52, n. 2, p.131-158, 2003.
- HUSSEIN, B. R. *et al.* Applications of computer vision and machine learning techniques for digitized herbarium specimens: A systematic literature review. *Ecological Informatics*, v. 69, p. 101641, 2022.

MADDISON, W. P. A method for testing the correlated evolution of two binary characters: are gains or losses concentrated on certain branches of a phylogenetic tree? *Evolution*, v. 44, n. 3, p. 539-557, 1990.

MADDISON, W. P.; MIDFORD, P. E.; OTTO, S. P. Estimating a binary character's effect on speciation and extinction. *Systematic Biology*, v. 56, n. 5, p. 701-710, 2007.

MAYR, E. Speciation and Macroevolution. *Evolution*, v. 36, n. 6, p. 1.119-1.132, 1982.

PAGEL, M. The maximum likelihood approach to reconstructing ancestral character states of discrete characters on phylogenies. *Systematic biology*, v. 48, n. 3, p. 612-622, 1999.

PENNELL, M. W. *et al.* geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, v. 30, p. 2.216-2.218, 2014.

PLANTS OF THE WORLD ONLINE (POWO). Facilitated by the Royal Botanic Gardens, Kew. Plants Of The World Online, [2024]. Disponível em: <https://powo.science.kew.org>. Acesso em: 28 jun. 2024.

R CORE TEAM. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria, [2023]. Disponível em: <https://www.R-project.org/>. Acesso em: 05 de maio de 2024

RABOSKY, D.L. Automatic detection of key innovations, rate shifts, and diversity-dependence on phylogenetic trees. *PloS one*, v. 9, n. 2, p. e89543, 2014.

RANNALA, B.; YANG, Z. Probability distribution of molecular evolutionary trees: a new method of phylogenetic inference. *Journal of Molecular Evolution*, v. 43, n. 3, p. 304-11. 1996.

RANNALA, B.; YANG, Z. Inferring speciation times under an episodic molecular clock. *Systematic Biology*, v. 56, n. 3, p. 453-466, 2007.

REGINATO, M. *et al.* MelastomaTRAITs 1.0: A database of functional traits in Melastomataceae, a large pantropical angiosperm family. *Ecology*, v. 105, n. 6, p. e4308, 2024.

RENNER, T.; SPECHT, C. D. A sticky situation: assessing adaptations for plant carnivory in the Caryophyllales by means of stochastic character mapping. *International Journal of Plant Sciences*, v. 172, n. 7, p. 889-901, 2011.

REVELL, L. J. phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution*, v. 3, p. 217-223, 2012.

SILVESTRO, D.; SALAMIN, N.; SCHNITZLER, J. PyRate: a new program to estimate speciation and extinction rates from incomplete fossil data. *Methods in Ecology and Evolution*, v. 5, p. 1.126-1.131, 2014.

SOLÍS-LEMUS, C.; BASTIDE P.; ANÉ C. Phylonetworks: a package for phylogenetic networks. *Molecular Biology and Evolution*, v. 34, p. 3.292-3.298, 2017.

STEBBINS, G. L. *Flowering Plants: Evolution Above the Species Level*. Cambridge: Belknap Press of Harvard University Press, 1974.

STEVENS, P. F. Angiosperm Phylogeny Website. Version 14, July 2017 [and continuously updated since]. Disponível em: <http://www.mobot.org/MOBOT/research/APweb/>. Acesso em: 07 maio 2024.

VAMOSI, J. C. *et al.* Macroevolutionary Patterns of Flowering Plant Speciation and Extinction. *Annual Review of Plant Biology*, v. 69, p. 1-22, 2018.

ZIZKA, A. *et al.* CoordinateCleaner: standardized cleaning of occurrence records from biological collection databases. *Methods in Ecology and Evolution*, v. 10, p. 744-751, 2019.

O código de barras da vida: reconhecendo espécies de fungos a partir do DNA



Nicolas do Carmo Regio¹ & Mauro Carpes Westphalen¹

¹Programa de Pós-graduação em Botânica, Universidade
Federal do Rio Grande do Sul, Porto Alegre (Brasil)

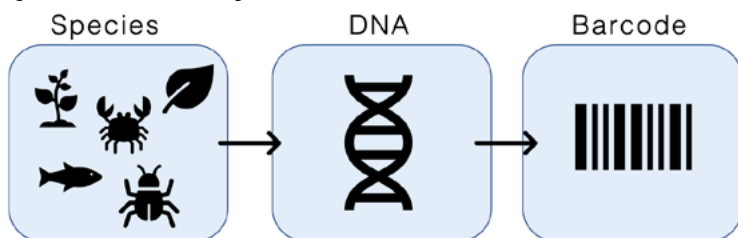
1. Contextualização

A categorização de organismos a nível de espécie é uma das bases dos estudos sobre a biodiversidade do planeta Terra. Desde Lineu, no século XVIII, inúmeros naturalistas e cientistas se dedicaram a descrever e classificar em categorias taxonômicas os mais diversos seres vivos de acordo com suas supostas semelhanças biológicas e morfológicas. No entanto, com o avanço do conhecimento científico e o desenvolvimento de novas técnicas de estudo, os limites de diferenciação entre espécies se tornaram cada vez mais dúbios, de tal forma que o próprio conceito de espécie não é uma unanimidade entre os cientistas.

Nessa perspectiva, análises envolvendo sequências de DNA, tido como o “código da vida” dos seres vivos, naturalmente surgem como grandes aliados na identificação de espécies por taxonomistas. Uma das ferramentas utilizadas é o “*DNA-barcoding*”, que busca identificar organismos distintos a partir de fragmentos de regiões específicas do material genético, funcionando quase que literalmente como um “código de barras da vida” (Figura 1). Essa abordagem, no entanto, não é tão simples, uma vez que não se aplica a todos os grupos de organismos nem substitui a descrição morfológica. Ainda assim é uma ferramenta muito útil, sobretudo ao se trabalhar com análises ambientais e metagenômicas que habitualmente englobam milhares de sequências de DNA dos mais

diversos grupos de seres vivos, em especial microrganismos como bactérias, vírus e fungos.

Figura 1: “DNA-barcoding”



Fonte: File:DNA_Barcoding.png (2019).

Os fungos, por sua vez, representam o segundo maior grupo de eucariotos da natureza, com estimativas de cerca de 2,5 milhões de espécies existentes (Niskanen *et al.*, 2023). No entanto, apenas cerca de 155 mil destas encontram-se formalmente descritas, fazendo com que aproximadamente 94% da funga mundial permaneça desconhecida para a ciência (Antonelli *et al.*, 2023). Parte dessa diversidade oculta vem sendo descrita principalmente em regiões historicamente subamostradas, como o Neotrópico, tratando-se de uma grande parcela a partir de estudos moleculares envolvendo complexos de espécies. Esses complexos, também chamados de espécies crípticas, se referem a grupos de organismos proximalmente relacionados e com limites morfológicos pouco claros. Isso muitas vezes leva a confusões taxonômicas e, consequentemente, subestimação da diversidade existente, já que diversas espécies morfológicamente semelhantes acabam sendo identificadas com um único nome.

Nesse sentido, a região ITS (*Internal Transcribed Spacer*), proveniente do DNA nuclear fúngico, é considerada o “barcode” ou “código de barras” dos fungos, haja vista seu comprimento relativamente curto e boa taxa de diversificação entre espécies, além de ser facilmente obtida em laboratório e até em campo (Schoch *et al.*, 2012). Mesmo assim, embora a região ITS seja amplamente aceita

e utilizada pela comunidade científica, vale ressaltar que a identificação e classificação de espécies utilizando um único marcador genético nem sempre é efetiva e tampouco representa a verdadeira história de vida das linhagens amostradas. Por isso, análises multigênicas, ou até genômicas, são essenciais para a realização de estudos filogenéticos mais completos. Nesse sentido, diversos marcadores moleculares secundários podem ser utilizados para auxiliar na identificação de fungos, como as regiões LSU, TEF1- α , RPB1, entre outras, sendo que a aplicação de cada região costuma variar de acordo com o grupo de fungos estudado (Blaalid *et al.*, 2013; Stielow *et al.*, 2015).

No presente capítulo, será demonstrado como a região ITS pode ser utilizada na identificação de espécies de fungos a partir de um exemplo prático baseado em espécies neotropicais do gênero *Favolus* Fr., recentemente estudadas por Palacio *et al.* (2021).

2. Objetivo

Promover um exercício de comparação simples e visual da região ITS de diferentes espécies neotropicais pertencentes ao gênero *Favolus* Fr, a fim de familiarizar o leitor com o processo de identificação e descrição de espécies da funga.

3. Descrição da atividade

Neste item, será proposta a construção de um alinhamento de sequências utilizando diferentes ferramentas de bioinformática aplicadas em parte do processo de identificação e descrição de novas espécies de fungos realizado por micólogos taxonomistas da atualidade. As sequências utilizadas seguirão o estudo feito por Palacio *et al.* (2021), que poderá ser utilizado na íntegra como material auxiliar.

3.1. Explorando a biodiversidade na prática: *Favolus* como exemplo

Os fungos do gênero *Favolus* Fr. são conhecidos por decompor madeira, formando corpos de reprodução carnosos, vulgarmente conhecidos por “orelhas-de-pau”, cujo himenóforo (região reprodutiva do basidioma) é formado por tubos alongados e similares a favos de mel, característica que dá nome ao táxon. Esses organismos possuem ampla distribuição e são comumente encontrados em ecossistemas florestais do Neotrópico e do Brasil, possuindo um histórico bem documentado de associações com diversos grupos de artrópodes. Além disso, muitas espécies são reconhecidamente comestíveis, sendo apreciadas e até mesmo cultivadas por diversos povos originários da América, como os Yanomami (Palacio *et al.*, 2021; Sanuma *et al.*, 2016).

Apesar de a espécie-tipo do gênero possuir o epíteto “*brasilensis*”, o grupo teve sua diversidade negligenciada por muito tempo no país. Foi só recentemente que Palacio *et al.* (2021) analisaram filogeneticamente diversos espécimes neotropicais, encontrando pelo menos sete espécies diferentes com ocorrência confirmada para o Brasil, das quais apenas cinco possuem dados moleculares disponíveis. As espécies apresentam diferenças morfológicas sutis para quem não é especialista, mas são suportadas por análises moleculares, tornando-as um exemplo da biodiversidade brasileira para ser explorado neste exercício prático.

3.2. De onde vêm as sequências?

O material genético dos fungos, assim como dos demais organismos, é composto por milhares ou até milhões de pares de bases de nucleotídeos, o que dificulta sua utilização em análises, devido a sua longa extensão e complexidade. Nesse sentido, um dos primeiros métodos desenvolvidos para obtenção desse dado, conhecido como método Sanger, envolve justamente a fragmentação e replicação de regiões específicas do DNA, usualmente obtidas por meio de reações de PCR (Reação em Cadeia da Polimerase). O produto

desse processo é constituído por milhões de cópias de fragmentos de DNA referentes à região desejada, que posteriormente são “lidos” por equipamentos denominados “sequenciadores”, dando origem às tão mencionadas “sequências de DNA” utilizadas nos estudos moleculares.

As “sequências” nada mais são do que arquivos digitais que possuem literalmente a sequência de nucleotídeos de uma região específica referente à espécie ou ao táxon do qual o material genético em questão foi extraído. Assim, cada sequência representa exclusivamente o espécime amostrado em uma determinada região, sendo identificado por um código específico chamado de *voucher*. Com isso, é de extrema importância que as sequências sejam disponibilizadas em repositórios públicos para que possam ser acessadas e utilizadas por outros pesquisadores, garantindo o princípio de repetição inerente a qualquer estudo científico.

Nessa perspectiva, o GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) é um dos maiores e mais acessíveis repositórios de sequências de material genético do mundo, sendo amplamente utilizado por pesquisadores para depositar as sequências geradas em seus estudos (Figura 2). A base de dados é mantida pelo *National Center for Biotechnology Information* (NCBI), que também dispõe de outras ferramentas interessantes, como o BLAST (*Basic Local Alignment Search Tool*) que possibilita a busca por similaridade de sequências depositadas no acervo do repositório. No GenBank é possível buscar e baixar sequências de regiões e genomas inteiros dos mais diversos organismos. A busca pode ser realizada pelo nome do táxon ou espécie, pelo *voucher* do espécime ou diretamente pelo “*Accession Number*”, que nada mais é do que o código de registro da sequência na plataforma do site.

Figura 2: Tela inicial do GenBank

GenBank Overview

What is GenBank?

GenBank[®] is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences ([Nucleic Acids Research](#), 2013 Jan 41(01):D36-42). GenBank is part of the [International Nucleotide Sequence Database Collaboration](#), which comprises the DNA DataBank of Japan (DDBJ), the European Nucleotide Archive (ENA), and GenBank at NCBI. These three organizations exchange data on a daily basis.

A GenBank release occurs every two months and is available from the [ftp site](#). The [release notes](#) for the current version of GenBank provide detailed information about the release and notifications of upcoming changes to GenBank. Release notes for [previous GenBank releases](#) are also available. GenBank growth [statistics](#) for both the traditional GenBank divisions and the WGS division are available from each release.

An [annotated sample GenBank record](#) for a *Saccharomyces cerevisiae* gene demonstrates many of the features of the GenBank flat file format.

Access to GenBank

There are several ways to search and retrieve data from GenBank.

- Search GenBank for sequence identifiers and annotations with [Entrez Nucleotide](#).
- Search and align GenBank sequences to a query sequence using [BLAST](#) (Basic Local Alignment Search Tool). See [BLAST info](#) for more information about the numerous BLAST databases.

GenBank Resources

- [GenBank Home](#)
- [Submission Types](#)
- [Submission Tools](#)
- [Search GenBank](#)
- [Update GenBank Records](#)

Fonte: imagem gerada pelos autores.

Esses dados devem ser disponibilizados nos estudos científicos que comumente trazem em seu texto, ou como material auxiliar, tabelas contendo as informações referentes às sequências utilizadas nas análises. A Tabela 1 apresenta as sequências que serão utilizadas neste exercício, representando as cinco espécies do gênero *Favolus* que ocorrem no Brasil e possuem dados moleculares disponíveis. O quadro também é um exemplo das informações que devem estar disponíveis nos artigos.

Tabela 1: Sequências de *Favolus* spp. do Brasil

Espécie	Voucher	Código GenBank
<i>F. brasiliensis</i> (Fr.) Fr.	MP 197	MN648683
<i>F. brasiliensis</i> (Fr.) Fr.	NMJ 348	OR353445
<i>F. ianthinus</i> (Gib. & Ryv.) Zmitr. & Kovalenko	DS 1677	MN648691
<i>F. ianthinus</i> (Gib. & Ryv.) Zmitr. & Kovalenko	DS 1700	MN648690
<i>F. pseudogrammocephalus</i> Palacio & Drechsler Santos	MP 218	MN648689
<i>F. rugulosus</i> Palacio & R.M. Silveira	MP 191	MN648684
<i>F. rugulosus</i> Palacio & R.M. Silveira	HUA 228632	ON897745
<i>F. yanomamii</i> Palacio & Menolli	MEE s.n.	MN648688
<i>F. yanomamii</i> Palacio & Menolli	ACM 1295	MN648686

Fonte: Palacio *et al.*, 2021.

Utilizando as informações fornecidas acima, encontre as sequências citadas na Tabela 1 através da ferramenta de busca do GenBank. É necessário estar conectado à internet para acessar a base de dados. Os arquivos das sequências indicadas serão utilizados para construção da matriz de caracteres, conforme apresentado no item a seguir.

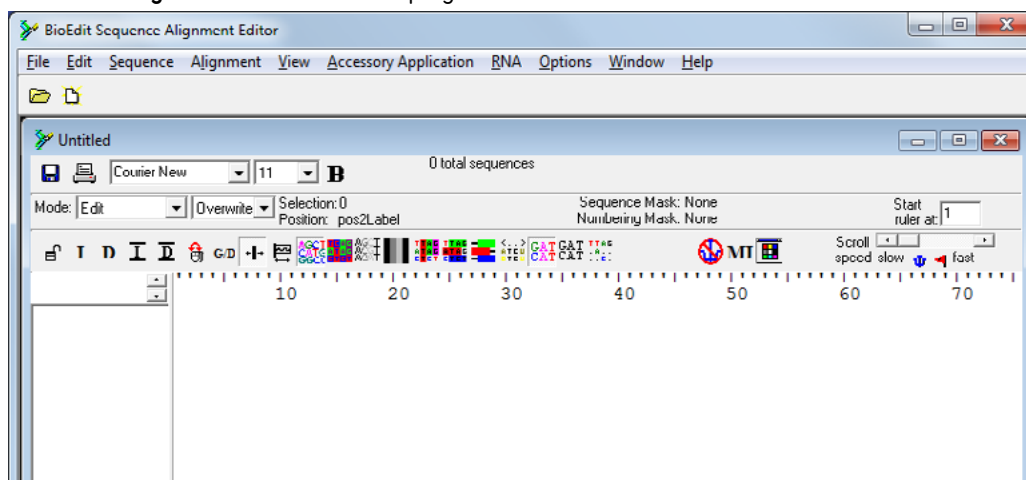
3.3. Construção do alinhamento

Uma vez encontradas as sequências de interesse, o próximo passo é a criação de uma matriz de caracteres. Essa matriz será futuramente utilizada para criar um alinhamento, ou seja, um pareamento das sequências por similaridade. Para construir esse arquivo, diversos *softwares* podem ser utilizados, como o BioEdit, o Geneious, o Aliview ou o ChromasPro. Esses programas costumam apresentar múltiplas funções além da simples construção e edição de alinhamentos, no entanto é importante atentar para a disponibilidade e a facilidade de uso de cada um, uma vez que alguns programas, como o Geneious, são pagos.

Sendo assim, neste exercício será utilizado o BioEdit (Hall, 1999), um dos programas de edição e análise mais antigos e de

livre acesso. O programa apresenta uma interface simples, mas bem intuitiva e relativamente fácil de usar (Figura 3), encontrado em repositórios on-line via *Google* ou qualquer outro site de pesquisa. No momento da escrita deste capítulo, sugere-se realizar o seu download a partir de algum dos links a seguir: <https://thalljscience.github.io/> ou <https://bioedit.software.informer.com/Download-gr%C3%Altis/>.

Figura 3: Interface inicial do programa Bioedit



Fonte: Hall (1999).

Uma vez instalado, basta abrir o programa e clicar no botão “*New alignment*”, localizado no canto esquerdo superior e representado pelo ícone de folha de papel. O programa abrirá um arquivo em branco que deverá ser utilizado para a construção do alinhamento. O próximo passo é copiar e colar as sequências disponibilizadas na Tabela 1 e encontradas no GenBank de acordo com o item anterior.

A Figura 4 ilustra a página encontrada a partir da busca de uma das sequências da Tabela 1 com seu código do Genbank. Nela é possível encontrar diversas informações, desde o nome dos autores e da publicação, bem como do marcador molecular referente à

sequência em questão, até o *voucher* do espécime e local de coleta, informações essenciais na realização de estudos taxonômicos.

Figura 4: Exemplo de página contendo o arquivo da sequência no GenBank

GenBank Send to: ▼ Change region shown ▼

Favolus brasiliensis voucher MP197 small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence

GenBank: MN648683.1

[FASTA](#) [Graphics](#)

[Go to: ☺](#)

LOCUS MN648683 626 bp DNA linear PLN 13-SEP-2021

DEFINITION Favolus brasiliensis voucher MP197 small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence.

ACCESSION MN648683

VERSION MN648683.1

KEYWORDS

SOURCE Favolus brasiliensis

ORGANISM Favolus brasiliensis

REFERENCE 1 (bases 1 to 626)

AUTHORS Palacio, M., Drechsler Dos Santos, E.R., Menolli, N. Jr. and da Silveira, R.M.O.

TITLE An overview of Favolus from the Neotropics, including four new species

JOURNAL Mycologia 113 (4), 759-775 (2021)

PUBMED 33945432

REFERENCE 2 (bases 1 to 626)

Analyze this sequence ▲

[Run BLAST](#)

[Pick Primers](#)

[Highlight Sequence Features](#)

[Find in this Sequence](#)

Related information ▲

[PubMed](#)

[Taxonomy](#)

LinkOut to external resources ▲

[SH0968154.10FU](#) [UNITE]

Recent activity ▲

[Turn Off](#) [Clear](#)

Favolus brasiliensis voucher MP197 small subunit ribosomal RNA gene, partial Full Record

Fonte: imagem gerada pelos autores.

Para copiar a sequência de interesse, basta clicar no link azul denominado “FASTA” (nome do formato de arquivo utilizado nos alinhamentos) e selecionar todo o texto, incluindo o cabeçalho com a sequência em questão, como ilustrado na Figura 5. Ademais, as sequências também podem ser baixadas individualmente através do link denominado “send to”, mas devem estar sempre em formato FASTA para que o alinhamento funcione.

Figura 5: Seleção e cópia de todo o texto, incluindo o cabeçalho da sequência

FASTA ▾

Send to: ▾

Favolus brasiliensis voucher MP197 small subunit ribosomal RNA gene, partial sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence

GenBank: MN648683.1

[GenBank](#) [Graphics](#)

```
>MN648683.1 Favolus brasiliensis voucher MP197 small subunit ribosomal RNA gene, partial
sequence; internal transcribed spacer 1, 5.8S ribosomal RNA gene, and internal transcribed
spacer 2, complete sequence; and large subunit ribosomal RNA gene, partial sequence
AGGACATTATCGAGTTTTGAAAGGGTGTAGCTGGCCCTTACGGAGGCATGTGCACGCCCTGCTCAAT
CCALCTACACCTGTGAACCTTACGTGGGTTTCGAGAGGGGGCCGTAAAAAGCCTCAGATTGAAGCCCA
CGCTAACCACAACACTGTTATTGTATTGTAATGTATCGCGATGTTAATAAACGCATCTATATACAAC
TTCAGCAACGGATCTCTTGGCTCTGCAATCGATGAAGAACGACGCAAAATGCGATAAGTAATGTGAATTG
CAGAATTCAGTGAATCATCGAATCTTTGAACGCACCTTGCGCTCTCTTGGTATTCGAGGAGCACACTGT
TTGAGTGTGCTGAACCTCAACCTACAAGTCCTTTGTTGGACTTTGTTAGGCTTGGATATTGGAGGCT
TTGCTGGCTCTTACTAGYAGCGAGCTCTCTTAAATGCATTAGCTTGGTCTCTGCGAATCGGCCTC
GGTGTGATAAATGTCTACGCCGCGACCGTGAAGCGTTATGGGCGAGCTTCAACAGTCTCTTTGTTGAG
ACAGTAAACACTTTGACATCTGACCTCAAATCAGGTGGGACTACCCGCTGAACCTTAAGCATATCAA
```

Fonte: imagem gerada pelos autores.

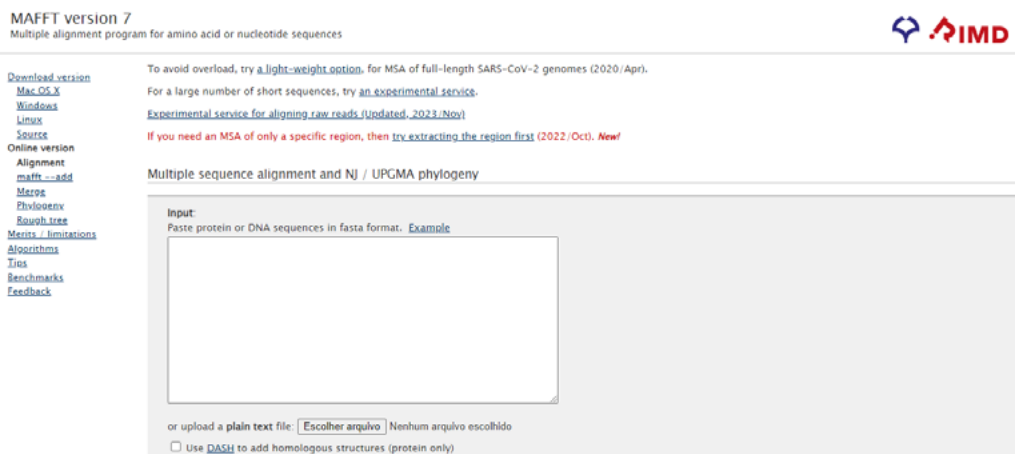
Após copiar a sequência, vá até o arquivo recentemente criado no BioEdit, selecione a opção “File”, no canto superior esquerdo, e clique na opção “Import from clipboard”. Com isso, a sequência completa deverá aparecer no documento. Repita o mesmo processo com as demais sequências presentes na Tabela 1 e a matriz de caracteres estará completa.

É recomendado editar o nome das sequências, haja vista que nomes longos, com espaços ou caracteres especiais, podem gerar erros em futuras análises filogenéticas. Por isso, é importante remover todos os espaços e substituir por traços “*underline*” (_). Sugere-se identificar as sequências apenas com o nome das espécies e o voucher do espécime. Ao final desta parte, as sequências não possuirão o mesmo tamanho e tampouco estarão alinhadas. Para isso, é recomendado utilizar outros programas que buscam alinhar as sequências devidamente, como apresentado a seguir.

3.4. Alinhando as sequências no MAFFT

Com a matriz construída, o próximo passo é alinhar as sequências selecionadas. Para isso, é necessário o uso de programas que possuem algoritmos capazes de encontrar e posicionar as sequências de acordo com a similaridade entre seus caracteres. Nesse sentido, existem diversos métodos que podem ser empregados, como o Muscle e o CustalW, que costumam estar embutidos nos programas de alinhamento. No entanto, neste exercício será utilizado o MAFFT, um método confiável e amplamente usado para alinhar sequências de DNA nuclear, como a região ITS (Katoh; Standley, 2013). Nesse caso, basta acessar o servidor on-line do programa pelo link: <https://mafft.cbrc.jp/alignment/server/index.html> (Figura 6).

Figura 6: Tela inicial do servidor on-line do MAFFT v.7



Fonte: imagem gerada pelos autores.

Para alinhar as sequências, será necessário salvar o arquivo criado no item anterior em formato FASTA com o nome da sua preferência. Feito isso, basta clicar em “Escolher arquivo” no servidor on-line do MAFFT e fazer *upload* do arquivo FASTA recém-criado. O MAFFT oferece a possibilidade de escolha de diferentes defini-

ções e modelos iterativos que podem ser selecionados de acordo com a especificidade de cada *dataset*. No entanto, outros testes em outros programas são necessários para determinar qual modelo seria o mais indicado. Por isso, não é necessário alterar nenhuma das predefinições indicadas no site, basta clicar em “*submit*”, ao final da página, que a análise rodará conforme a estratégia automática do programa.

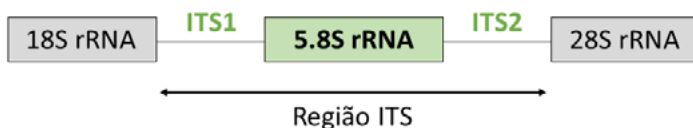
Terminado o serviço, basta baixar novamente o arquivo em formato FASTA e abri-lo no BioEdit. As sequências estarão com as posições de seus nucleotídeos alinhadas, mas ainda possuirão tamanhos diferentes. Nessa perspectiva, é interessante deixar todas as sequências com um tamanho similar, uma vez que a região ITS dos fungos costuma apresentar cerca de 600 pares de base. Por isso, aconselha-se que sejam cortados os excedentes das sequências primeiramente do final e posteriormente do início do alinhamento. Para fazer isso, basta selecionar no topo do alinhamento a última ou a primeira posição do alinhamento que se deseja manter, clicar em “*edit*”, no canto esquerdo superior, selecionar “*select to end*” ou “*select to beginning*” para selecionar a área a ser cortada e clicar no botão de deletar ou apagar do teclado. É importante notar que para fazer as edições o BioEdit deve estar no modo “*Edit*”, que pode ser selecionado na caixa à esquerda, conforme visto na Figura 4.

Ao alinhar as sequências, o MAFFT, ou qualquer outro programa, cria espaços, chamados de “*gaps*”, visando alinhar a posição exata ocupada por cada nucleotídeo em comparação com as demais sequências. Esses *gaps* tendem a variar de acordo com a distância filogenética entre os táxons analisados. Logo, no caso de sequências de um mesmo gênero, as distâncias tendem a ser pequenas, embora ainda existam. Com as sequências alinhadas, agora é possível analisar as diferentes regiões em que o gene ITS está organizado e identificar possíveis espécies distintas, conforme apresentado no item a seguir.

3.4. Comparando a região ITS de espécies diferentes

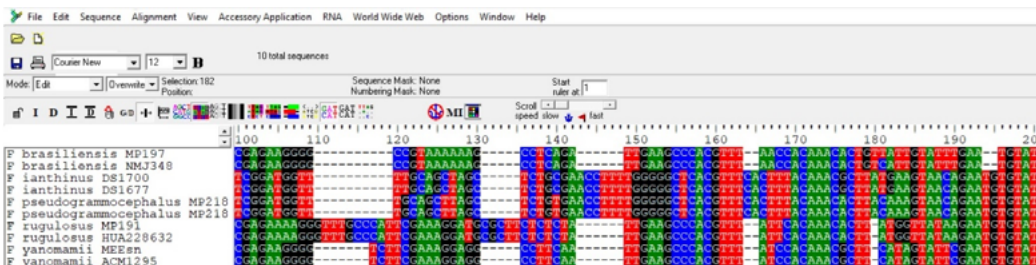
A região ITS é comumente dividida em três sub-regiões: ITS1, 5.8s e ITS2 (Figura 7). Enquanto as regiões ITS1 e ITS2 tendem a ser mais variáveis entre diferentes espécies, a 5.8S, que se localiza ao centro, é mais conservada. Dessa forma, em geral, espécies de um mesmo gênero terão a região 5.8S praticamente idêntica ou apenas com poucas mudanças de bases.

Figura 7: Detalhe da região ITS



Fonte: imagem gerada pelos autores.

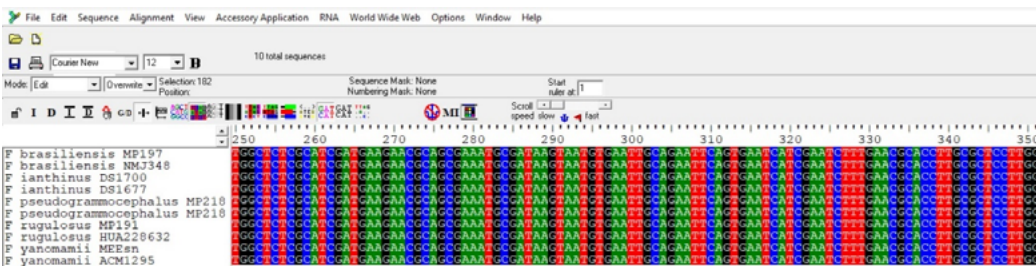
No caso das espécies de *Favolus* utilizadas neste exemplo, as regiões do ITS ficam claramente visíveis. Observe o alinhamento obtido e verifique: 1. as regiões variáveis no começo do alinhamento, referente ao ITS1 (Figura 8); 2. um bloco de sequências quase idênticas em todas as espécies, referente ao 5.8S (Figura 9); 3. uma segunda região variável após o 5.8S, referente ao ITS2 (Figura 10). Na Figura 8, pode-se observar duas regiões ao final do ITS1, dentre as posições 120 a 140 e 180 a 190, em que cada uma das espécies possui uma sequência única. O mesmo pode ser visto na Figura 10, próximo à posição 410 e entre as posições 440 e 465 da região ITS2. Essas regiões únicas auxiliam na identificação e no reconhecimento de táxons, além de se tratar de uma ferramenta importante na comparação de espécies morfológicamente semelhantes. Observados os alinhamentos obtidos, tente localizar essas e outras regiões únicas e verifique como algumas espécies apresentam sequências mais semelhantes entre si do que outras.

Figura 8: Seção do alinhamento da região ITS1

Fonte: imagem gerada pelos autores.

Figura 9: Seção do alinhamento da região 5.8S

Fonte: imagem gerada pelos autores.

Figura 10: Seção do alinhamento da região ITS2

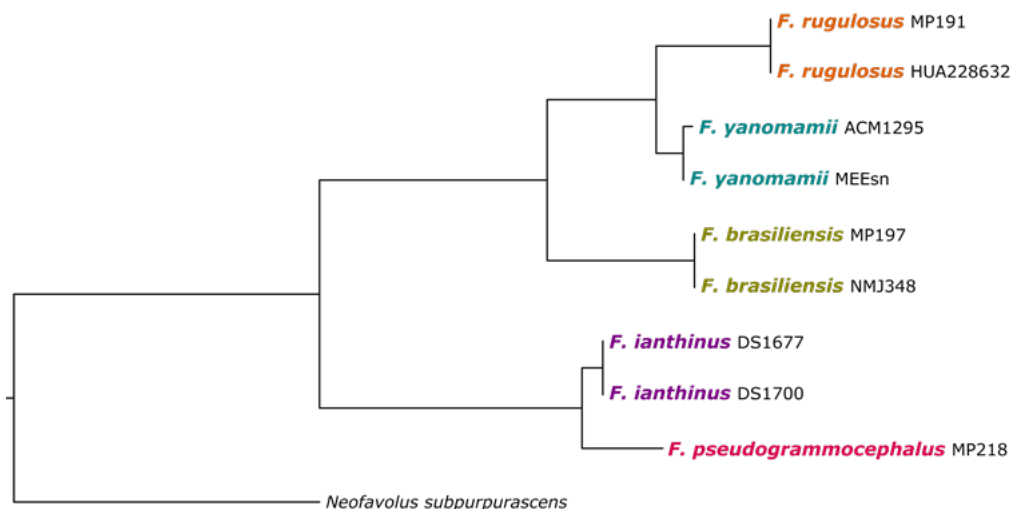
Fonte: imagem gerada pelos autores.

A partir do alinhamento, pode-se observar a similaridade das sequências entre as diferentes espécies. Fica claro, por exemplo, que *F. ianthinus* e *F. grammocephalus* são mais similares entre si do que com as demais espécies, da mesma forma que *F. rugulosus* e *F. yanomamii*. Também é possível que haja pequenas diferenças entre sequências de uma mesma espécie. Isso pode ser observado

na Figura 8, onde *F. brasiliensis* possui bases diferentes após a posição 180. Essas variações são comuns entre diferentes indivíduos de uma mesma espécie e podem representar sua variabilidade genética. No entanto, em alguns casos também podem ocorrer pequenos erros de sequenciamento, que levam à presença de bases diferentes em uma mesma posição de uma determinada espécie. Por esse motivo, é importante sempre observar um número maior de sequências para fazer essa comparação. Em geral, uma espécie pode apresentar até cerca de cinco bases diferentes na região ITS entre distintos indivíduos.

Com os alinhamentos prontos, a matriz de caracteres pode ser utilizada para realização de inferências filogenéticas. No caso dos fungos, os métodos mais utilizados são os de Máxima Verossimilhança, no programa RaxML, e de Inferência Bayesiana, no programa MrBayes. Esta parte, no entanto, exige um tempo considerável para ser realizada, além de conhecimento mais aprofundado acerca dos métodos de análises filogenéticas e dos programas citados, por isso não será tratada neste capítulo. Mesmo assim, a título de curiosidade, abaixo segue um exemplo gráfico de como ficariam as filogenias construídas a partir do alinhamento construído neste exercício (Figura 11). Na figura, é possível observar que as espécies que apresentaram sequências mais semelhantes entre si, como observado nas diferentes regiões do alinhamento, estão mais proximalmente relacionadas, formando clados separados e corroborando sua separação em táxons distintos.

Figura 11: Relações filogenéticas entre espécies brasileiras do gênero *Favolus*



Fonte: imagem gerada pelos autores.

4. Considerações finais

O exercício construído ao longo deste capítulo representa uma pequena parte dos métodos aplicados por micólogos taxonomistas na incessante busca pela elucidação da diversidade da funga neotropical. Mesmo com o avanço das técnicas e da aplicação cada vez maior de dados genômicos nos estudos científicos, o reconhecimento de espécies de fungos utilizando marcadores moleculares, como a região ITS, ainda se faz essencial tendo em vista as estimativas e o alto número de espécies que ainda permanecem desconhecidas para a ciência.

Nesse sentido, ratifica-se que análises moleculares e filogenéticas devem idealmente agregar mais de uma região do material genético, além do ITS, e estar sempre integradas a outras formas de análises, como estudos morfológicos e biológicos dos espécimes, a fim de elucidar e evidenciar na prática as relações evolutivas compartilhadas entre as espécies. Por fim, espera-se que o presente

capítulo tenha sido de bom proveito no aprendizado de diferentes ferramentas de bioinformática e, principalmente, na metodologia aplicada na longa área da taxonomia e da sistemática.

Referências

ANTONELLI, A. *et al.* State of the World's Plants and Fungi 2023. *Royal Botanical Garden, Kew*, 2023.

BLAALID, R. *et al.* ITS1 versus ITS2 as DNA metabarcodes for fungi. *Molecular Ecology Resources*, v. 13, n. 2, p. 218-224, 2013. Disponível em: <https://doi.org/10.1111/1755-0998.12065>. Acesso em: 28 jun. 2024.

FILE:DNA Barcoding.png. *WikiMedia*, 2019. Disponível em: https://commons.wikimedia.org/wiki/File:DNA_Barcoding.png. Acesso em: 28 jun. 2024.

HALL, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series*, v. 41, p. 95-98, 1999.

KATOH, K. & STANDLEY, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Molecular Biology and Evolution*, v. 30, n. 4, p. 772-780, 2013.


NISKANEN, T. *et al.* Pushing the Frontiers of Biodiversity Research: Unveiling the Global Diversity, Distribution, and Conservation of Fungi. *Annual Review of Environment and Resources*, v. 48, n. 22.1, p. 1-28, 2023. Disponível em: <https://doi.org/doi.org/10.1146/annurev-environ-112621-090937>. Acesso em: 28 jun. 2024.

PALACIO, M. *et al.* An overview of *Favolus* from the Neotropics, including four new species. *Mycologia*, v. 113, n. 4, p. 759-775, 2021. Disponível em: <https://doi.org/10.1080/00275514.2021.1878797>. Acesso em: 28 jun. 2024.

SANUMA, O. I. *et al.* *Sanöma samakönö sama tökö nii pewö oa wi i tökö waheta: Ana amopö* = Enciclopédia dos alimentos Yanomami (Sanöma): cogumelos. São Paulo: Instituto Socioambiental, 2016. (Série uli tä uli naha tökö kupai i tökö pewö taö wi: saberes da floresta; 4).

SCHOCH, C. L. *et al.* Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences*, v. 109, n. 16, p. 6.241-6.246, 2012. Disponível em: <https://doi.org/10.1073/pnas.1117018109>. Acesso em: 28 jun. 2024.

STIELOW, J.B. *et al.* One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. *Persoonia – Molecular Phylogeny and Evolution of Fungi*, v. 35, n. 1, p. 242-263, 2015. Disponível em: <https://doi.org/10.3767/003158515X689135>. Acesso em: 28 jun. 2024.



As organizadoras agradecem aos
revisores pela revisão e pelas valiosas
contribuições aos capítulos

REVISORES:

- Aline Zanetti
- Emanuel Scherdien da Rosa
- Jean Lucas Benvenuti
- Joelmir Mazzoni
- Johnatan Vilasboa
- Letícia Osório
- Mariane Dorneles

A Editora

A Editora da Universidade de Caxias do Sul, desde sua fundação em 1976, tem procurado valorizar o trabalho dos professores, as atividades de pesquisa e a produção literária dos autores da região. O nosso acervo tem por volta de 1.600 títulos publicados em formato de livros impressos e 600 títulos publicados em formato digital. Editamos aproximadamente 1.000 páginas por semana, consolidando nossa posição entre as maiores editoras acadêmicas do estado no que se refere ao volume de publicações.

Nossos principais canais de venda são a loja da Educs na Amazon e o nosso site para obras físicas e digitais. Para a difusão do nosso conteúdo, temos a publicação das obras em formato digital pelas plataformas Pearson e eLivro, bem como a distribuição por assinatura no formato streaming pela plataforma internacional Perlego. Além disso, publicamos as revistas científicas da Universidade no portal dos periódicos hospedado em nosso site, contribuindo para a popularização da ciência.

Nossos Selos



EDUCS/Ensino, relativo aos materiais didático-pedagógicos;



EDUCS/Origens, para obras com temáticas referentes às memórias das famílias e das instituições regionais;



EDUCS/Pockets, para obras de menor extensão que possam difundir conhecimentos pontuais, com rapidez e informação assertiva;



EDUCS/Pesquisa, referente às publicações oriundas de pesquisas de graduação e pós-graduação;



EDUCS/Literário, para qualificar a produção literária em suas diversas formas e valorizar os autores regionais;



EDUCS/Traduções, que atendem à publicação de obras diferenciadas cuja tradução e a oferta contribuem para a difusão do conhecimento específico;



EDUCS/Comunidade, cujo escopo são as publicações que possam reforçar os laços comunitários;



EDUCS/Internacional, para obras bilíngues ou publicadas em idiomas estrangeiros;



EDUCS/Infantojuvenil, para a disseminação do saber qualificado a esses públicos;



EDUCS/Teses & Dissertações, para publicação dos resultados das pesquisas em programas de pós-graduação.



Conheça as possibilidades de formação e aperfeiçoamento vinculadas às áreas de conhecimento desta publicação acessando o QR Code.

