



DEPARTAMENTO DE FITOPATOLOGIA - UFV
FIP-680 POPULATION BIOLOGY OF PLANT PATHOGENS



EXERCÍCIO 1

Prof. Eduardo S. G. Mizubuti
DFP-UFV

Atenção! Todos os exercícios deverão ser feitos individualmente. Caso seja detectada evidência de “coincidências de alto grau”, os trabalhos “coincidentes” serão penalizados.

1. Descritores gerais de variabilidade

Utilize os pacotes `vegan`, `poppr` e `iNEXT` para analisar os dados quanto à diversidade genética, em sentido amplo. Use as análises que julgar pertinente. Escreva um rápido relatório, apresente os dados utilizando figuras e tabelas e escreva um parágrafo para interpretar cada um dos conjuntos. Imagine cenários, pense em perguntas que você poderia fazer (e responder!). Descreva os resultados por meio de tabelas e/ou figuras e faça uma breve discussão dos resultados.

Para fazer as análises, você deverá escrever um *script* em R para que seja lido o arquivo que contém os dados que deseja analisar. Então, o primeiro passo é preparar o arquivo de dados:

O arquivo de dados deverá ser no formato texto e organizado da seguinte maneira:

cada população em uma linha.
nas colunas insira as frequências dos genótipos/fenótipos.

exemplo: para uma população de 31 indivíduos, 6 genótipos foram encontrados:

3 indivíduos têm o genótipo 1 (G1); 1 → G2 (apenas um indivíduo do genótipo 2); 18 → G3; 3 → G4; 5 → G5; 1 → G6

OU, visto de outra forma:

Genótipo	Frequência
1	3
2	1
3	18
4	3
5	5
6	1

Número de genótipos: 6

Tamanho da amostra: 31 indivíduos

O código de rarefação trabalha com frequências, não interessa o marcador utilizado. O que interessa são as frequências com que os genótipos foram observados. No arquivo haverá apenas uma linha com as frequências:

```
3 1 18 3 5 1
```

Após preparar o arquivo, salve como texto e use a extensão “.txt”

Agora, de volta ao R. Leia os dados do arquivo preparado. Veja a sequência de comandos no R para obter esse resultado (tecle "Enter" ao final de cada linha):

```
-- entrada de dados (vegan)
```

```
pop1 <- as.matrix(read.table("C:/Meus Documentos/nome_da_pasta/nome_do_arquivo.txt") )
```

Índice de Shannon-Wiener H'

```
pop1_H <- diversity(pop1, "shannon", MARGIN = 1, base = exp(1))
```

```
pop1_H
```

Índice de Stoddart & Taylor

```
pop1_G <- diversity(pop1, "invsimpson")
```

```
pop1_G
```

Usando esses dados, calcule o número esperado de genótipos para uma amostra de tamanho 16. Ou seja, estimar a riqueza [Expected number of genotypes - $E(g(n))$] para uma amostra de 16 indivíduos.

```
rarefy(pop1, 16, se=FALSE, MARGIN = 1)
```

Estimativa de intervalo de confiança para índice de diversidade

As funções do pacote *vegan* para obter os índices e as estimativas de erro padrão por *bootstrap* (1000 replicações) são:

```
> # ===== indice de Shannon-Wiener
>
> H(dados1, lev="alpha", q = 1, boot=TRUE, boot.arg=list(num.iter=1000))

> #===== indice N1 de Hill
>
> d(dados1, lev="alpha", q = 1, boot=TRUE, boot.arg=list(num.iter=1000))

> # ===== indice de Simpson
>
> H(dados1, lev="alpha", q = 2, boot=TRUE, boot.arg=list(num.iter=1000))

> # ===== indice de Stoddart & Taylor
>
> d(dados1, lev="alpha", q = 2, boot=TRUE, boot.arg=list(num.iter=1000))
```

Com o valor do Erro padrão (StdErr), calcule o intervalo de confiança usando a fórmula tradicional. Há várias formas de calcular intervalo de confiança baseado em *bootstrap*. Assumindo que a estimativa tenha distribuição normal (o que pode ser razoável na maioria dos casos), os limites inferiores (LI) e superiores (LS) do IC podem ser obtidos como:

LI = Estimativa do índice - 1,96.(StdERR)

LS = Estimativa do índice + 1,96.(StdERR)

2. Poppr

Explore os conjuntos de dados usando o pacote "poppr". Atente para o uso das funções `poppr`, `diversity_stats` e `diversity_boot`.

3. DnaSP

Clique em File → Open Data File -- selecione o arquivo de sequências já alinhadas.

Uma janela abrirá e você deverá conferir como os dados estão formatados. "Sequential" quer dizer que a sequência de nucleotídeos não tem "quebra de linhas". "Interleaved" é um formato no qual a sequência é representada em bloco, como se fosse num parágrafo, com várias linhas. Clique OK.

Clique no menu "Data" "Format"; formate os dados conforme solicitado; clique OK.

Clique em "Data" "Define sequence sets" para informar como os isolados estão agrupados. Ex.: quais isolados são de dicotiledôneas e quais são de monocotiledôneas. Selecione os isolados de uma hospedeira, clique no botão >> e depois clique em "Add new sequence set" e forneça o nome da subpopulação (dico ou mono). Repita o procedimento para a outra subpopulação. Para finalizar, clique em "Update all entries".

Clique no menu "Display" "Data info".

Clique no menu "Display" "Data view". Explore as possibilidades de visualização dos dados. Por exemplo, clique na caixa "Select sites/Codons..." (canto inferior direito) e selecione "Variable (polymorphic) sites" e role a barra horizontal para visualizar quais posições são variáveis.

Clique no menu "Overview", opção "Intraspecific Data" para iniciar suas análises. Neste caso as análises descritivas serão realizadas para todos os isolados.

Clique no menu "Analysis" e selecione opções referentes às análises que deseja realizar.

Atenção! Leia o Help para conhecer como as análises estão implementadas, os modelos, pressuposições e parametrização.

4. Rarefação

iNEXT

Baseado no roteiro discutido durante o laboratório, calcule os números efetivos de genótipos com base em riqueza genotípica, Shannon-Wiener e Simpson. Faça os cálculos para obter as curvas de interpolação e extrapolação.