

MESQUITE

A modular system for evolutionary analysis

W.P. Maddison University of British Columbia D.R. Maddison Oregon State University

Curso Mesquite Aula 1-2

1. Introdução ao funcionamento do programa

O programa Mesquite tem a grande vantagem de ser modular, isto é, tem um núcleo central de programas essenciais ao seu funcionamento aos quais se pode depois acrescentar módulos adicionais, especializados, por vezes escritos por <u>autores diferentes</u> dos autores do Mesquite. É conveniente consultar o sítio do Mesquite (<u>http://mesquiteproject.org</u>) para ver que módulos existem (surgem módulos novos com alguma regularidade) e que funcionalidades executam. Podem consultar os módulos disponíveis indo à pasta «Mesquite», dentro da pasta «Mesquite_Folder». Novo módulos deverão ser instalados na pasta «Mesquite».

- 1. Abrir programa e verificar versão (pode haver uma versão mais actualizada; conferir no sítio do <u>Mesquite</u>.
- 2. Abrem-se três janelas:
 - a. uma janela que se fecha sozinha, mas onde podem ver os módulos que são carregados no lançamento do programa. Consoante as necessidades do operador podem ser acrescentados mais módulos no lançamento inicial. Por defeito, porém, convém manter o número de módulos limitado ao necessário para não tornar o programa mais exigente em termos de memória.
 - a janela de fundo preto documenta instruções executadas; não devem precisar de recorrer a esta janela, mas NÃO a fechem, pois é necessária ao funcionamento do programa; podem minimizá-la para simplificar o monitor.
 - c. a janela inicial do Mesquite, onde é possível abrir um ficheiro de dados, gravar, e gerir os módulos ("Activate/Desactivate Packages"), etc. Além da barra de menus File, Edit, Window, Help a janela tem vários separadores Log, Projectos and Files, Search, Simplification uma janela de pesquisa, e uma zona de anotações. Notem que passando com o cursor por cima de diferentes áreas da janela surgem informações de ajuda. Esta janela será sobretudo utilizada para abrir o ficheiro inicial e para consultar a rede de projectos e ficheiros, no separador desse

nome. De resto, após aberto o ficheiro inicial, esta janela também poderá ser minimizada.

- 3. O programa é instalado com um tutorial interactivo que explora as suas múltiplas funcionalidades. Para o colocar a correr:
 - a. File > Open File > navegar até à pasta «Mesquite_folder» > «examples» > abrir ficheiro "Introduction.nex"
 - b. Abre-se uma nova janela. Esta será a janela central de funcionamento no projecto de trabalho, neste caso o tutorial. Notem de novo a estrutura de barra de menus e separadores. A barra de menus é bastante dinâmica, variando consoante o separador aberto, análises já feitas, etc., pelo que vão prestando atenção.
 - c. Notem que na janela central existem dois separadores:
 - i. o segundo, no qual a janela abre automaticamente, é onde poderão clicar para seguir o tutorial na direcção desejada
 - ii. o primeiro separador «Project of "Introduction.nex» será um separador importante mais tarde quando combinarmos diferentes tipos de ficheiros e dados. Carreguem na seta ao lado de "File incorporation" e verão vários icons para incorporar diferentes tipos de ficheiros e dados. Notem que existe a opção de incluir ("Include") ou ligar ("link") a um ficheiro.

O Mesquite tem 3 formas de ler um ficheiro (M32¹)

- 1. Abrir ("Open File ...") o ficheiro é aberto ficando independente de quaisquer outros ficheiros que estejam abertos, não partilhando com estes informação; pertencerá a outro Projecto.
- Incluir ("Include File ...") o conteúdo do ficheiro é lido e incorporado no principal ficheiro do projecto (em geral o ficheiro *.nex). Toda a informação do ficheiro passa a fazer parte do ficheiro principal e é gravada quando este é gravado.
- 3. Ligar ("Link File ... ") os conteúdos do ficheiro são lidos e a sua informação está disponível para uso no projecto a que está associado, mas mantém-se como um ficheiro separado, a sua informação não é integrada no ficheiro principal do projecto. Alterações feitas a esse ficheiro não são necessariamente gravadas quando o ficheiro principal é gravado. Ao ligar ficheiros (por exemplo, ficheiro com sequências e ficheiro com árvores), os ficheiros mantêm-se separados, mas a informação comporta-se como se fosse um único ficheiro (o projecto). Quando os ficheiros são ligados, essa informação (a ligação) é gravada no ficheiro inicial do projecto.Quando este for aberto de novo, as ligações são estabelecidas de novo.
 - d. Se regressarem à janela inicial, e escolherem o separador «Projects and Files» verão a estrutura de projectos e ficheiros.
 - e. Explorar início do tutorial
 - i. clicar no botão "Go" ao lado de "Basic Examples"
 - ii. clicar no botão "Go" ao lado de "Introducing the basic functioning of Mesquite"
 - iii. clicar no botão "Next", no canto superior direito, para o menu seguinte
 - iv. clicar no botão "Show page", no canto inferior esquerdo, para ver página do manual
 - v. clicar no botão "Next", no canto superior direito, para o menu seguinte
 - vi. Menu de taxa \ caracteres. Explorem com o cursor os diferentes botões e vejam as

¹ O formato M# refere-se ao número de página do pdf "Mesquite2Manual" versão 2.0 de 23.10.2007

anotações de ajuda na janela inferior. Notem que estão no Separador «Character Matrix» e que este tem por sua vez também separadores – Graphics, Text, Parameters, Modules, Citations.

- vii. Abram o separador «Project of ...» e vejam que o tutorial abriu um ficheiro com uma matriz e árvores. Cliquem nas setas junto a «Matrix ...» e «Trees ...»
- viii. Voltem ao separador «Character Matrix» e cliquem no botão "Next"
- ix. Esta é a janela árvores. Notem no canto superior direito o número da árvore e que, por debaixo a seta direita, está activa. Cliquem para explorar as árvores que foram carregadas.
- x. Sigam o tutorial, usando os botões "Next", até chegar à página "Web Links". Experimentem o botão «Go to file on website».
- xi. Fechem o tutorial: File > Close File

2. Taxa, Caracteres e Árvores

2.1 Caracteres Categóricos

O termo «Taxa» no Mesquite é usado da forma mais geral, como OTU. Pode significar espécie ou outro nível taxonómico, pode significar um indivíduo ou cópia de gene. Um taxa pertence a um bloco de taxa ("taxa block").

- Na janela inicial do Mesquite: File > New, dêm um nome ao novo ficheiro e gravem numa pasta própria. (É boa prática com o Mesquite, e não só, ter projectos nas suas pastas próprias, para evitar confusão entre ficheiros.)
- Vamos criar um exemplo com 5 taxa, chamados «Vertebrados», e vamos seleccionar «Make Character Matrix» para o Mesquite criar não só uma matriz para os taxa mas uma matriz associada de caracteres.
- Surge nova janela sobre a matriz de caracteres, ao qual podemos dar um nome, determinar o número de caracteres, e definir que tipo de caracteres são. Criemos uma matriz com 3 caracteres categóricos.
- Abre-se então a janela do nosso novo projecto, no separador da matriz de caracteres. Explorem os diferentes botões no painel da esquerda, sobrevoando o cursor sobre os botões e lendo a informação no painel de ajuda (M93).
- 5. Podemos agora dar nomes aos nossos taxa, fazendo duplo-clique sobre as janelas com "taxon 1", etc. Vamos introduzir 5 nomes: lampreia, sapo, ave, baleia e porco.
- Se quisermos adicionar um novo taxa, clicamos no botão "Add Taxa", o icone muda, e clicamos na matriz; uma janela pergunta-nos quantos taxa cremos adicionar. Adicionemos mais um, para incluir o *Homo*. Agora para mudarmos o nome, precisamos de mudar novamente o icone do cursor para a seta ou o icon de edição.
- 7. Do mesmo modo, podemos agora dar nomes aos três caracteres, e.g., "terrestre", "pêlo", "ovo amniótico". Podemos acrescentar um novo carácter de forma semelhante ao adicionar taxa, mas

escolhendo o botão apropriado. Podemos assim acrescentar a coluna "antebraço"

- 8. Agora podemos preencher a matriz, clicando nas células, e preenchendo com 1's e 0's. Podem usar as setas do teclado e a tecla TAB para moverem entre células.
- 9. Na ausência de informação, podemos deixar o "?" (missing data). No caso de poliformismo podemos introduzir mais do que um estado, e.g., "0&1" no caso do "sapo" no carácter "terrestre". Podem também introduzir um polimorfismo com "0.1". Notem que tal não será interpretado como um décimo visto tratar-se de uma matriz de caracteres categóricos.
- 10. Vamos agora para o separador do projecto. Aí seleccionem a matriz de taxa clicando em "List and Manage Taxa". Cliquem na pequena seta preta na célula "Group" no topo dessa coluna. Surge a opção "New Group". Podemos então dar-lhe um nome e escolher uma cor. Criado o grupo, podemos selecionar um ou mais taxa na matriz, clicar de novo na seta preta de "Grupo", escolher "Set Group" e atribuir um grupo aos taxa escolhidos. A cor dos taxa muda.
- 11. Nesta matriz podemos criar grupos que nos façam sentido, e.g. mamíferos. Se os nossos "taxa" fossem várias sequências, poderiamos indicar a espécie a que pertenciam que sequências/taxa.
- 12. Notem que alternando entre o separador da matriz de caracteres e matriz de taxa se alteram as opções na Barra de Menus. Há Menus e submenus específicos para cada tipo de matriz. Por exemplo, se fosse necessário adicionar texto a alguns nomes de taxa, na matriz Taxa seleccionariamos esses taxa, e iriamos ao menu List > Taxon Names > Add to Names e introduziriamos o texto.
- 13. Num projecto poderemos querer associar ficheiros nos quais os nomes dos taxa são ligeiramente diferentes. Para tal é útil a função «Alternative Taxon Names»: Columns > Alternate Names cria uma nova coluna na matriz de taxa. É possível criar múltiplas colunas de nomes alternativos. Aí podemos introduzir os nomes usados, por exemplo, como OTU's num ficheiro de uma árvore que tenhamos produzido noutro programa. Ou, porque há programas que apenas aceitam nomes com 8 caracteres podemos querer ter nomes alternativos com esse limite. Criem um nova coluna para nomes alternativos. Cliquem na pequena seta preta na célula "Alternate Names" e selecionem "Replace Alternatives by Taxon Names". Isto copia os nomes originais para a nova coluna. Agora para truncar os nomes extensos fariamos List > Taxon Names > Truncate. Desta forma mantemos ambas versões dos nomes. Se preferirmos os nomes extensos nas visulalizações no Mesquite, podemos voltar à seta preta e seleccionar "Trade Taxon Names with Alternatives"
- 14. Podemos também associar palavras aos diferentes estados numéricos. No separador de caracteres ir a Matrix > Edit State Names e introduzir descrições nas células relevantes. Notem nesta janela que existem 55 estados possíveis (0-9, A-H, K-N, P-Z, a-h. k-n, p-z). Assim, se tiverem um carácter codificado com números 0-13, para introduzir o estado '13' terão de usar a letra 'D'. A introdução do número '13' será interpretada como um polimorfismo '1&3'. E tomem atenção às maiúsculas/minúsculas.
- 15. Após fechar a janela de edição dos níveis de caracteres, vemos que na matriz de caracteres os algarismos foram substituídos pelos nomes. Ao clicar numa célula vemos de novo o algarismo correspondente. Os estados podem ser à mesma alterados usando algarismos, mas estes são agora associados aos nomes criados para os diferentes estados.
- 16. Um outro menu importante de edição na matriz de caracteres está em Matrix > Alter/Transform, por exemplo, recodificar, para resolver polimorfismos ou células vazias, ou preencher aleatoriamente estados de um carácter, i.e. simular um carácter. Por exemplo, criem um novo carácter. Seleccionem o novo carácter clicando no número da coluna. Escolham Matrix > Alter/Transform > Random Fill

Categorical e dê parâmetros para a simulação do carácter.

17. Uma funcionalidade útil sobretudo com caracteres morfológicos é o painel de anotações, que permite associar a um taxa, carácter ou célula comentários, referências, figuras e citações. Para abrir o painel de anotações Matriz > Show Annotation Panel. Depois é seleccionar uma célula – de um taxa, carácter ou célula na matriz; clicar no botão (+) e introduzir as informações. Estas depois ficam visíveis ao passar com o cursor "seta" sobre as células.

| 0 | 00 | | | | 11a-an | notations.nex | | |
|---|-----|------------------|---------|-----------|-----------------|-----------------|---|--|
| Project of "11a-annotations.nex" Character Matrix "Structure & Color" | | | | | | | | |
| Graphics Text Parameters Modules Citations O 📀 | | | | | | d Search Data | | |
| | ¥ | | 1 | 2 | 3 | 4 | | 🕱 🕀 📆 2 💷 💷 🖿 t.1/c.1 character *ki |
| | T | Taxon \ Characte | knobble | gizmoid | zibinator shape | zibinator setae | | Comment |
| 10 | Æ | 1 Fubonidia | | pointy | smooth | double sided | 4 | This is a distinct dimple unlike others |
| 1 | 8° | 2 Snorolopa | dimpled | pointy | edgy | double sided | | among the fubonidians. Quite pretty, actually. |
| | ő | 3 Quidnubia | flat | spatulate | smooth | single sided | | |
| 0 | • | 4 Zorabynyx | flat | spatulate | smooth | double sided | U | Reference |
| (E) | | | | | | | | Illustration from specimen AMNH |
| | * | | | | | | | 6788975 |
| * | 2. | | | | | | | Image 🔟 |
| 1 | IX. | | | | | | | distinct dimple |
| | | | | | | | | Author & Modification Dates |
| | | | | | | | | |
| | | | 1 | | | | 7 | Created Fri Apr 29 18 Fri Apr 29 18:11:14 F |
| This striking dimple was first noticed by Snerp (1965) | | | | | | | | |
| S [t.1 c.1 s.1] knobble: (0) flat; (1)* dimpled; [in taxon "Fubonidia"] | | | | | | | | |
| | | | | | | | | |

- 17. Tal como criamos grupos de taxa, podemos criar grupos de caracteres. Seleccionem o separador do Projecto e cliquem em "List and Manage Characters". Depois é proceder como para os Taxa. Vejam que há também uma coluna onde podem incluir/excluir um carácter das análises subsequentes. Seleccionem a fila do carácter, cliquem na pequena seta preta na coluna «Inclusion» e escolham a opção.
- 18. Com o Mesquite podemos querer modelar a evolução de caracteres. Um passo importante é portanto definir qual o modelo de evolução de cada carácter. Para ver os modelos disponíveis: Characters > List of Character Models
- 19. Na janela de caracteres, por defeito, cada carácter tem associado um modelo de probabilidade Markov de um parâmetro Mk1 (est). Seleccionando o carácter:
 - a. podem alterar o modelo de probabilidade (relevante em análises de verosimilhança ou *likelihood*) seleccionando a pequena seta preta junto a «Probability Model».
 - b. podem visualizar o modelo segundo o critério de parcimónia escolhendo Columns > Current Parsimony Models. Surge então uma nova coluna, onde podem também escolher diferentes modelos de parcimónia.
 - c. podem criar um novo modelo de probabilidade indo a Characters > New Character Model > escolhendo o modelo base e alterando os parâmetros. Depois de criado e guardado o novo modelo este aparece entre as opções quando seguirem o passo (a) ou (b)
 - d. podem alterar o peso de cada carácter escolhendo Columns > Current Weights e alterando os valores ou pesos.

2.2 Árvores

Vamos agora iniciar um novo projecto, partindo de ficheiros já existentes

- Abram a pasta «Parablennius» e abram, usando o «Bloco de Notas», o ficheiro nexus "s71.parabl.46.nex". Notem a estrutura: um bloco "data" com sequências e um bloco "paup" com instruções para análise das sequências pelo programa PAUP.
- Ainda usando o «Bloco de Notas» abram o ficheiro "S7.parabl.pars.tre". Notem o bloco "trees" contendo as 206 árvores mais parcimoniosas. Notem o formato usado para descrever uma árvore. Este formato, usando parêntesis, vírgulas e dois pontos designa-se Newick. Notem a notação "#taxa:comprimento de braço". Tentem localizar uma politomia. Qual seria uma notação alternativa para uma politomia?
- 3. Voltemos agora ao Mesquite para abrir o ficheiro nexus, usando File > Open file ...
- 4. O projecto abre-se com uma matriz de taxa e uma matriz de caracteres, neste caso de sequências de DNA.
- 5. Abram o separador de sequências e explorem os diferentes botões. [Notem que a função "Undo" só memoriza o último passo.]
- 6. Vamos agora associar o ficheiro com as árvores de parcimónia, <u>incluindo</u> o ficheiro "S7.parabl.pars.tre" (recordem-se que existem diferentes formas de juntar ficheiros num projecto: incluir e ligar). Um modo de o fazer é File > Include File. Surge agora no separador do projecto uma secção com as árvores.
- Cliquem em "View Trees". Mais uma vez notem as novidades na barra de menus e os diferentes botões. Explorem o uso dos botões para manipular as árvores (M108). Notem as setas no canto superior esquerdo para correr entre as 206 árvores mais parcimoniosas.
- 8. Voltem à pasta «Parablennius» e abram, usando o «Bloco de Notas», o ficheiro nexus "s71.parabl.46.nex.con". Este ficheiro contem duas árvores de consenso produzidas usando MrBayes. A segunda tem a topologia e os comprimentos de braço. A primeira tem também as probabilidades posteriores, os valores de suporte bayesianos associadas aos diferentes nodos. Reparem na notação NEWICK. Neste caso, os OTUs vêm indicados alfabéticamente (não com algarismos). Junto a cada OTU vem um número, correspondente ao comprimento de braço, separados por um ":"

tree con_all_compat = (COGA_EF52:0.064442,((...

Mas surgem também números a seguir aos parênteses direitos, correspondentes aos valores de suporte

... (HYFI:0.002841,HYFI_D45:0.014430)**0.85**:0.008487, ...

- Voltando ao Mesquite, vamos agora associar estas árvores ligando o seu ficheiro ao nosso projecto, fazendo File > Link File ...
- 10. Temos agora uma segunda secção de árvores no nosso projecto. Reparem que vêm apenas listados dois ficheiros sob "Project". O primeiro ficheiro com as árvores parcimoniosas foi integrado, incluido no ficheiro mestre inicial "s71.parabl.46.nex", enquanto o ficheiro de árvores bayesianas, que foi ligado, se manteve como ficheiro separado.
- 11. Abram as árvores bayesianas. A primeira surge com os valores de suporte. Se visularizarem a

segunda árvore desse ficheiro verão que estes já não estão presentes, como seria de esperar dada a estrutura NEWICK das duas árvores.

- 12. Vamos agora gravar o projecto. Vejam File > Save ou Save As. Vemos opções consonantes com a estrutura do projecto (ficheiros ligados e incluidos). Vamos gravar o projecto com um nome diferente. Façam File > Save As > "s71.parabl.46.nex" e alterem o nome para "s71.parabl.46.mesquite.nex". Fechem o projecto.
- 13. Usando o «Bloco de Notas» abram o ficheiro "s71.parabl.46.mesquite.nex". Tem um bloco de TAXA, um bloco CHARACTERS, com as sequências, depois um bloco TREES, com as árvores mais parcimoniosas que incluímos no projecto, <u>não tem</u> o bloco de árvores bayesianas. Seguem-se blocos criados pelo Mesquite: CODON, MESQUITECHARMODELS e MESQUITE. Notem que desapareceu o bloco PAUP. Por essa razão, se usarem ficheiros nexus com instruções PAUP que depois abram e gravem no MESQUITE é importante dar-lhe um nome novo, podendo assim manter o ficheiro nexus original, com as instruções do PAUP. File > Defaults > Automatic Nexus Backups ???
- 14. Ainda usando o «Bloco de Notas» abram o ficheiro "S7.parabl.nj.1000.boot.tre". Esta é uma árvore resultante de uma análise neighbor-joining usando bootstraps. Podem consultar o ficheiro nexus para ver as instruções no PAUP que lhe deu origem. Como foi dada instrução ao PAUP para guardar os valores de suporte como comprimentos de braço

savetrees file=S7.parabl.nj.1000.boot.tre from=1 to=1000 brlens;

estes surgem a seguir aos ":"

... (2:100.000000,3:100.000000):87.900002) ...

- 15. Vamos agora voltar ao MESQUITE e reabrir o projecto que gravámos "s71.parabl.46.mesquite.nex". Notem que ele abriu o projecto e associou também o ficheiro ligado com as árvores bayesianas.
- 16. Vamos agora incluir a árvore NJ clicando em "File incorporation" e "Include File"
- 17. Cliquem em «Tree block "S7.parabl.nj.1000.boot.tre"»
- 18. Vemos a árvore, mas não os valores de suporte. Recordem-se da estrutura NEWICK. Para o Mesquite interpretar os valores depois do ":" como valores de suporte e os desenhar na árvore, façam Drawing > Branch Length Labels > Show Labels
- 19. Explorem os menus para ver as opções disponíveis e explorar formas de desenhar árvore.

2.3 Dados Moleculares

- 1. Abram a janela de sequências do projecto "s71.parabl.46.mesquite.nex".
- Existe um módulo que podem adicionar que permite analisar chromatogramas
 <u>http://mesquiteproject.org/packages/chromaseq/</u> e outro que permite fazer alinhamentos
 <u>http://mesquiteproject.org/packages/opalescent/</u> O pacote básico do Mesquite permite também enviar
 uma matriz de sequências para o Clustal Ver Matrix > Align Multiple Sequences …
- 3. Recordem-se que o Mesquite não está, de momento, construído para fazer inferência filogenética a partir de sequências (ou outros caracteres). (Já existe porém um módulo para construir árvores de neighbor-joining <u>http://nimbletwist.com/software/ninja/mesquite.html</u>) O Mesquite permite porém fazer gestão de sequências que podem ser úteis, podendo depois os dados ser exportados para uso noutros programas File > Export ... Permite também fazer análises dos dados moleculares.
- 4. Existem importantes funcionalidades no submenu Matrix > Alter/Transform que se podem aplicar a

todas as sequências ou apenas às sequências seleccionadas.

- 5. Notem as funcionalidades para resolução de ambiguidades (M196).
 - a. «Convert Uncertainties to Polymorphisms» transforma códigos ambíguos (e.g., M, R, W, etc) num polimorfismo (e.g., A&C, A&G, A&T, etc)
 - b. «Convert Polimorphisms to Uncertainties» faz o inverso de (a)
 - c. «Arbitrarily Resolve State Ambiguities» transforma uma célula com uma ambiguidade (e.g. M) num dos estados possíveis (A ou C), com igual probabilidade.
- 6. Podemos adicionar, a esta matriz, sequências adicionais do mesmo marcador, usando File > Include Taxa from File ..., ou mesmo arrastando o icon de um ficheiro fasta sobre a matriz de sequências.
- Também podemos adicionar sequências interactuando com o GenBank (se o computador estiver ligado à rede), usando Matrix > Utilities > Fetch and Add GenBank e escrevendo na janela os números das sequências, separadas por vírgulas (e.g., GU983860.1, EF527590.1 ; sequência s7 de *Lipoprhys pholis* e *Coryphoblennius galerita*)
- 8. O Mesquite também interactua com GenBank permitindo fazer BLAST a partir de sequências na matriz. Seleccionem uma sequência ou parte de uma sequência e façam Matrix > Search
- 9. Podemos também concatenar outro marcador molecular. Temos diferentes cenário possíveis.
 - a. No mais simples, o ficheiro com as sequências do segundo marcador tem os mesmos taxa com *exactamente* os mesmo nomes. (Segundo consegui apurar, nem serve se tivermos os nomes alternativos definidos). Vamos criar uma tal matriz. Façam Characters > Make New Matrix From > Simulated Matrix on Current Tree > Evolve DNA Characters > Escolhendo o modelo de evolução (e.g., Jukes-Cantor) > Confirmando a árvore sobre a qual os caracteres vão ser simulados > estipulando o número de caracteres (nº de bases, neste caso) > dar um nome à nova matriz > escolher o ficheiro onde a nova matriz será guardada. Temos agora duas matrizes com sequências com os mesmos nomes (por agora, vamos ignorar o alinhamento das novas sequências).

Abram a janela das sequências de S7. Vamos primeiro criar uma cópia desta matriz, para ficarmos com uma matriz apenas com este marcador Characters > Make New Matrix from > Stored Matrices.

Agora escolham Matrix > Utilities > Concatenate Other Matrix Escolham a matrix de sequências simuladas. Vão até à direita da matriz de characteres e verão ai as sequências simuladas.

- b. Se a segunda matriz de sequências tiver um bloco de taxa sobreposto mas não idêntico (mas os nomes são *idênticos*), i.e., se tiver alguns taxa em comuns, mas não todos. Neste caso, façam primeiro uma cópia do ficheiro de sequências S7 e dêem-lhe um nome indicativo que virá a ser uma matriz concatenada. Depois liguem ao projecto o ficheiro contendo as novas sequências (Link File...). Isso implica que o Mesquite irá manter esse ficheiro ligando, mas mantendo a integridade do novo ficheiro, incluindo o seu bloco de taxa. Sigam então o mesmo passo de concatenação. Vamos simular esta situação
 - i. Primeiro gravem o projecto.
 - Vamos agora à matrix de sequências simuladas, e vamos apagar uns taxa.
 Seleccionem uns taxa não sequenciais usando CTRL e Matrix > Delete Selected Chars or Taxa
 - iii. Agora vamos gravar esta matriz como ficheiro separado, usando Export.
 - iv. Notem que os taxa foram apagados em todos os ficheiros integrados no projecto. Não

era bem isso que queriamos. Vamos então fechar o projecto, sem o gravar, e reabrí-lo.

- Agora vamos <u>ligar</u> este ficheiro ao projecto. Notem que no separador do projecto agora parecem dois grupos de characteres, e há dados associados a dois "Taxa" independentes.
- vi. Abram agora o ficheiro de sequências originais e façam concatenar. Como os nomes dos taxa são idênticos ele concatenou as sequências simuladas apenas nos taxa correspondentes.
- c. Se a segunda matrix tem nomes (ainda que ligeiramente) diferentes para as sequências, elas serão representadas por um bloco de taxa diferente, o bloco de taxa mestre. Neste caso:
 - i. (criar um novo bloco de taxa, com os nomes das espécies ou indivíduos Taxa&Trees > New Block of Taxa, escolher o ficheiro, dar um nome e escolher o número desejado a ter na matriz concatenada)
 - ii. (Dar os nomes aos taxa.)
 - iii. Ligar o ficheiro com o outro marcador.
 - iv. Na janela do bloco de taxa mestre, criar uma Associação entre taxa Taxa&Trees > New Association, escolher os blocos de taxa para a associação (os OTUs associados a cada um dos marcadores moleculares), dar um nome à associação
 - v. Ir criando a associação entre OTUs dos dois blocos de taxa
 - vi. Feita a Associação podemos fazer File > Export > Export Matrix Fused (NEXUS).
 - vii. Podem ligar a nova matrix ao projecto para ver como ficou.
- d. As <u>análises</u> gerais de caracteres categóricos podem ser aplicadas a sequências moleculares, e existem adicionalmente estatísticas descritivas para sequências de DNA e proteínas. Para explorar estas funcionalidades, abram o exemplo do Mesquite abrindo o ficheiro «introduction.nex» que está na pasta «Molecular».
- e. Para produzir o gráfico do exemplo Analysis > Characters > Stored Characters > Character Compositional Bias . (Ainda) Não se parece nada com o gráfico do exemplo. Notem os eixos. Temos de mudar a orientação do gráfico Chart > Orientation > Values (Y) by Items (X). Para alterar o tamanho das barras (dos bins) Chart > Grouping on X > Moving Window
- f. Na representação seguinte das sequências, primeiro, alterou-se Color Matrix Cells de "Character State" para "Cell Value", e depois especificou-se que o cell value é o o conteúdo GC into a Matrix > Values for Colors > GC (Notem que esta opção no menu Matrix não estava presente antes; os menus são <u>muito</u> dinâmicos.)
- g. O gráfico seguinte foi feito clicando no pequeno icone com cara de ave, na matriz de caracteres.
- h. Abram um nova janela com a árvore armazenada e tentem reproduzir o seguinte gráfico com árvore circular.
- Solução: Drawing > Tree Form > Circular tree Analysis > Tace Character History > ACGT Compositional Bias > Stored Matrices Drawing > Backgroud color > Light Blue

Por fim, vamos associar novos caracteres a estes taxa

1. Até agora estivemos a trabalhar com caracteres categóricos. Podemos criar uma nova matriz de

caracteres contínuos. Podemos criar uma matriz do zero, preenchendo valores indo a Characters > New Empty Matrix

 Ou simular dados na árvore, Characters > Make New Matrix from ... > Simulated Characters on Current Tree, escolher Caracteres Contínuos, optar pelo único modelo disponível para caracteres continuos (Brownian), escolher 2 characteres, dar nome à nova matriz, escolher em ficheiro vamos guardar os novos dados,

Charting information about characters http://mesquiteproject.org/Mesquite_Folder/docs/mesquite/characters.html#charting