

Aide à l'utilisation de la fonction gapw()

Le gap-weighting (Thiele, 1993) :

Le traitement des caractères continus en analyse cladistique est très discuté. Une des méthodes les plus utilisées est le gap-weighting (Thiele, 1993). Elle permet de conserver les différences entre taxons dans le codage. Thiele souhaitait que l'on puisse utiliser le maximum d'états possible pour coder les différences entre les taxons de manière à avoir la plus meilleure résolution possible. Wiens (2001) propose le step-matrix-gap-weighting, un dérivé de la méthode de Thiele augmentant beaucoup la résolution (1000 états).

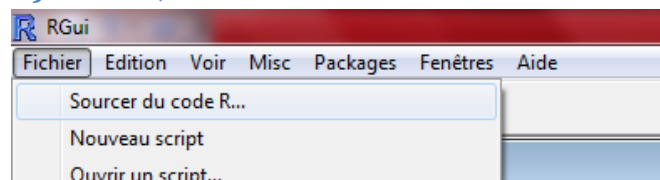
La fonction gapw() :

La fonction « gapw » permet de transformer des données selon la méthode de gap-weighting jusqu'à une résolution maximale de 65000 états. Les effets de la résolution sur les résultats des analyses phylogénétiques étant discutés, cette fonction permet de transformer automatiquement le jeu de données par plusieurs nombres d'états, c'est-à-dire plusieurs résolutions. Toutes les matrices sont générées dans des fichiers nexus ou tnt directement utilisables via les logiciels PAUP et TNT.

Deux types d'études sont facilités par cette fonction. D'une part, l'utilisateur peut traiter un jeu de données réels contenant des caractères continus et discrets. D'autre part, des études plus théoriques sur le gap-weighting peuvent être menées grâce à la genèse de jeu de données artificielles. Dans ce second cas, l'utilisateur crée une matrice en choisissant le nombre de caractères continus et discrets, le nombre de taxons, les valeurs minimales et maximales ainsi que le nombre de décimales pour le tirage des états de caractères continus, le nombre d'états différents ainsi que la probabilité de tirage de chacun d'entre eux pour les caractères discrets. Les caractères générés sont tous informatifs.

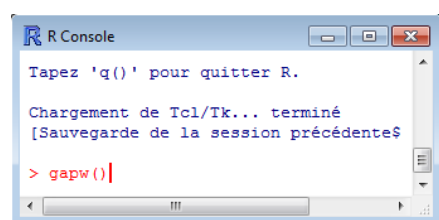
Utilisation de la fonction :

1) Sous R, vous devez sourcer la fonction :



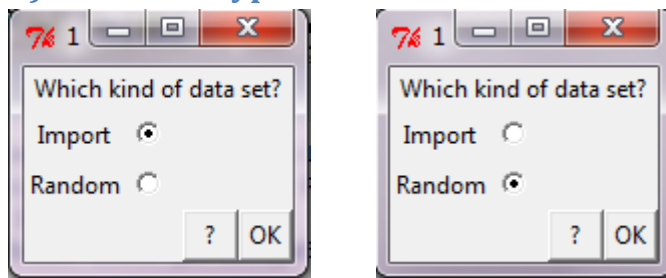
Commencer par sourcer la fonction gapw comme indiqué ci-contre :

Fichier->Sourcer du code R et sélectionner le fichier gapw.r que vous avez téléchargé.



Une fois sourcée, la fonction doit être appelée sur la console. Pour cela il suffit de taper : gapw() puis .

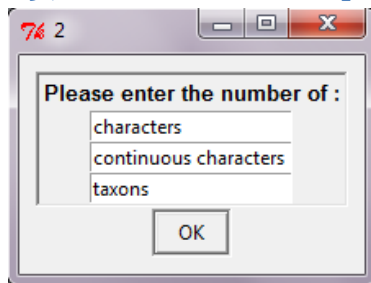
2) Différents types de données



Vous pouvez importer un jeu de données, pour cela cliquez sur « Import » et . Le jeu de données doit être sous la forme d'une matrice en format .csv. Les taxons sont placés en lignes et les caractères en colonnes. Les décimaux doivent être séparés des unités par une virgule. Les caractères continus sont placés avant les caractères discrets.

Vous pouvez générer un jeu de données aléatoires, pour cela cliquez sur « Random » et . Les paramètres du jeu de données générées sont choisis ci-après.

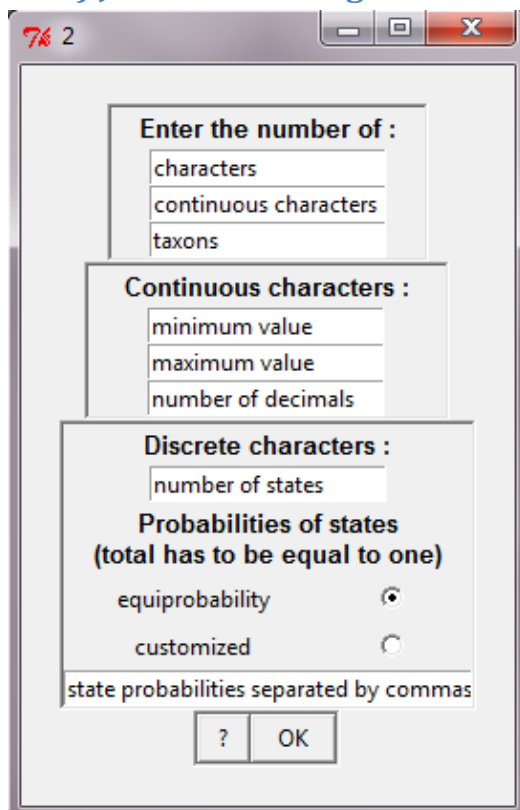
3) Jeu de données importé :



Compléter les trois champs suivants :

- Nombre de caractères de la matrice
- Nombre de caractères continus (Rappel : ils doivent être placés au début de la matrice)
- Nombre de taxons

3bis) Jeu de données généré :



Compléter les trois champs suivants pour déterminer la taille de la matrice :

- Nombre de caractères de la matrice
- Nombre de caractères continus (Rappel : ils doivent être placés au début de la matrice)
- Nombre de taxons

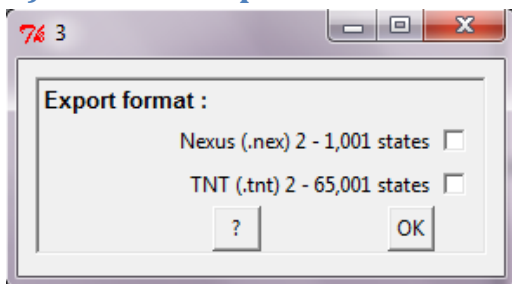
La deuxième sous-fenêtre comporte les paramètres pour le tirage des états des caractères continus de la matrice :

- Valeur minimal
- Valeur maximal
- Nombre de décimaux

La troisième sous-fenêtre comporte les paramètres pour les caractères discrets :

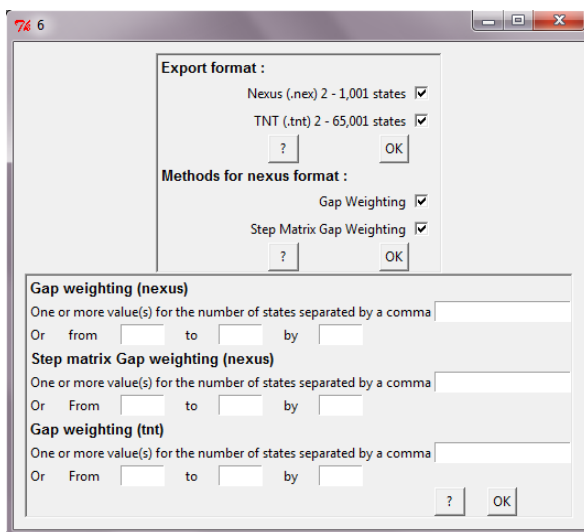
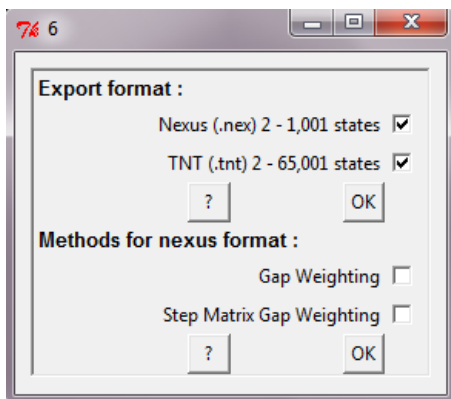
- Nombre d'états différents
- Probabilité de chaque état lors des tirages :
 - o équiprobabilité
 - o personnalisées : vous devez remplir le champ au-dessous avec la probabilité de tirage de chaque état séparée par une virgule. ATTENTION : le séparateur des décimaux est le point « . ».

4) Format d'export et discrétisation :



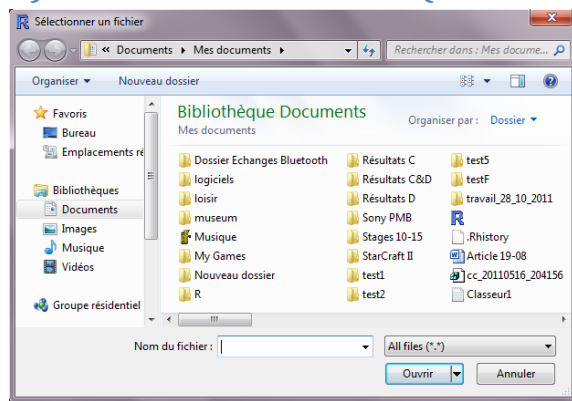
Choisir le format d'export :

- Nexus ou tnt (.nex ou .tnt)
- Les nombres d'états indiqués correspondent aux capacités maximales des programmes de reconstruction phylogénétiques lors de la sortie de la fonction.
- Pour plus d'informations sur ces méthodes, voir Thiele (1993) pour le gap-weighting et Wiens (2001) pour le step-matrix-gap-weighting.



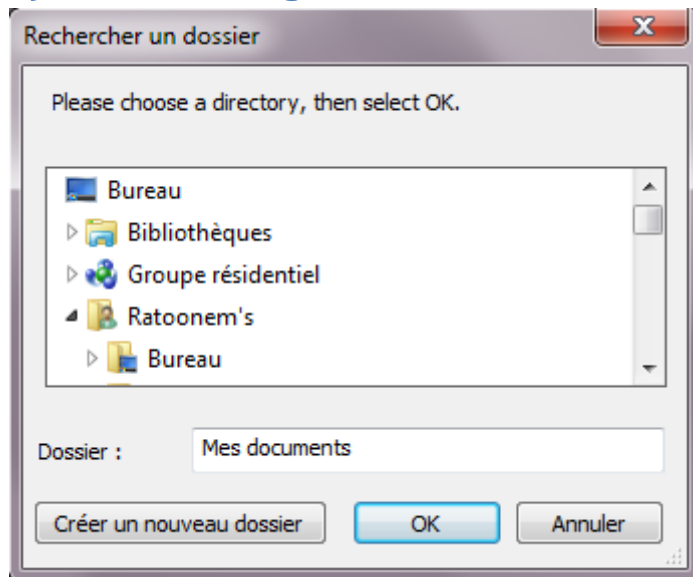
Pour chaque couple (méthode/ format) sélectionné, il existe deux manière d'entrer les nombres d'états souhaités pour la discrétisation. 1) Vous pouvez entrer une liste contenant une ou plusieurs valeurs. Deux valeurs sont séparées par une virgule. Exemple : « 2,5,10 ». 2) Vous pouvez également entrer une séquence. Exemple : toutes les valeurs entre 10 (From) et 50 (to) toutes les 20 unités (by) correspond à : 10, 30, 50.

5) Sélection de la matrice (seulement pour les jeux de données importés) :



Une fenêtre s'ouvre, vous devez sélectionner votre matrice. Elle doit être sous format .csv (export possible à partir de tous les tableurs). Les taxons et les caractères avec leurs noms doivent être respectivement en lignes et colonnes. Le caractère utilisé pour les décimaux est la virgule.

6) Dossier d'enregistrement :



Indiquez le dossier-cible pour l'enregistrement des fichiers exportés.

Références :

- Thiele, K. 1993. The holy grail of the perfect character: The cladistic treatment of morphometric data. *Cladistics* 9, 275-304.
- Wiens, J. J. 2001. Character analysis in morphological phylogenetics: Problems and solutions. *Syst. Biol.* 50, 689-699.