

Mesures de biodiversité

S. Pavoine & A.B. Dufour

A travers les différentes échelles (gène vs espèce vs règne ; espèce vs région) et les différents critères (e.g. phénotypiques, fonctionnels) de mesure de la biodiversité, une structure générale de données apparaît autour de trois objets : entité, catégorie, et collection. Une collection (e.g. communauté) est un ensemble d'entités (e.g. organismes) réparties entre catégories (e.g. espèces). L'entité est l'objet sur lequel nous pouvons faire des mesures, c'est l'objet que nous échantillons.

Table des matières

1	Introduction	2
2	Trois indices traditionnels de biodiversité	2
2.1	Présentation	2
2.2	La fonction de rareté	5
2.2.1	Etude de la fonction	5
2.2.2	Communautés d'oiseaux et habitat	6
3	L'entropie quadratique	8
3.1	Présentation	8
3.2	La mesure des dissimilarités	11
3.3	ADN mitochondrial et zones géographiques [2]	12
	Références	14

1 Introduction

Nous avons dans le tableau ci-dessous un aperçu très simplifié de différents niveaux de biodiversité. Ces niveaux s'emboîtent parfaitement. On peut ainsi souligner la complémentarité des études faites par des disciplines séparées par des traditions universitaires. La biodiversité existe à différentes échelles avec un fort continuum. Chaque niveau fait l'objet d'une spécialité représentée par une discipline à l'université. De nombreuses équipes de spécialistes ont développés des méthodes ad hoc pour leur domaine d'étude. Ce que l'on peut déplorer c'est que peu de généralisation ont encore été proposées. De telles généralisations permettraient d'analyser les différents niveaux avec des mêmes méthodes favorisant ainsi la comparaison, et la cohérence.

Collection	Entité	Catégorie	Objectif d'étude possible
Paysage	Relevé	Habitat	Etude de la diversité des habitats d'un paysage
Habitat	Relevé	Biotope	Etude de l'hétérogénéité d'un habitat en terme de biotopes
Biotope	Organisme	Espèce	Etude de la diversité des espèces présentes dans un biotope
Espèce	Organisme	Population	Etude de la diversité des populations d'une espèce
Population	Séquence d'ADN	Forme allélique	Etude de la diversité allélique d'une population
Forme allélique	Locus	Nucleotide (A, G, T ou C)	Etude de la diversité des nucléotides d'une forme allélique

2 Trois indices traditionnels de biodiversité

2.1 Présentation

Nous nous intéressons ici à trois indices très utilisés en écologie et en génétique.

La Richesse

L'indice le plus simple pour mesurer la biodiversité est le nombre S de catégories diminué de 1 afin qu'une collection comprenant une seule catégorie ait une biodiversité nulle. Cet indice s'écrit

$$H_r = S - 1,$$

et est appelé richesse.

Exemple.

Considérons trois espèces de poissons.



FIGURE 1 : Vues *in situ* du poisson lune (*Mola mola*), du thon rouge (*Thunnus Thynnus*) et du marlin bleu (*Makaira nigricans*).

Dans les deux situations définies ci-dessous, il y a trois catégories : $S = 3$ donc la richesse vaut $H_r = 2$.

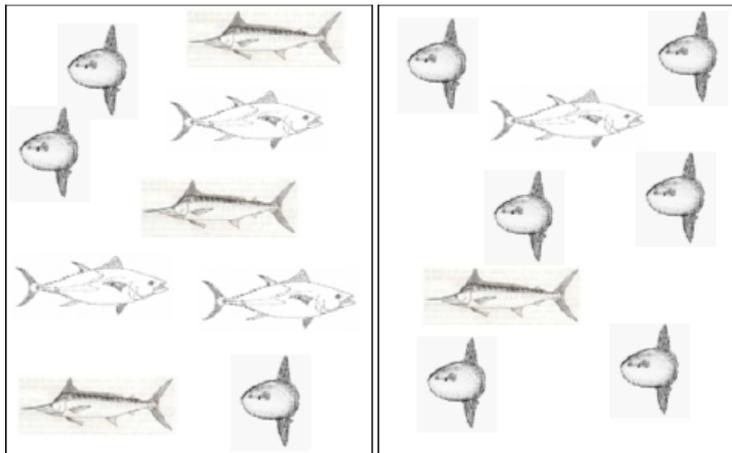


FIGURE 2 : La diversité dépend des fréquences des catégories. La diversité de l'assemblage théorique de gauche (**siteG**) est plus grande que celle de l'assemblage de droite (**siteD**), ce dernier possédant une catégorie dominante. Pour information, actuellement *Mola mola* (poisson lune) est une espèce commune ; les deux autres *Thunnus thynnus* (thon rouge) et *Makaira nigricans* (marlin bleu) sont classées vulnérables dans la liste de l'IUCN ("International Union for the Conservation of Nature"). Nous sommes donc plutôt dans la situation de droite.

Beaucoup ont argumenté contre ce type d'indices en affirmant que les fréquences des catégories doivent être considérées pour mesurer la biodiversité. Shannon [7], Gini et Simpson [3, 8] ont proposé des indices corrigeant la richesse par les fréquences relatives des catégories. Ces indices sont donc tous définis sur

l'ensemble

$$\mathcal{P} = \left\{ \mathbf{p} = (p_1, \dots, p_k, \dots, p_S), p_k \geq 0 \forall k = 1, \dots, S, \sum_{k=1}^S p_k = 1 \right\}$$

qui désigne l'ensemble des vecteurs de fréquences de longueur S .

L'indice de Shannon

Shannon [7] a développé son indice H_S dans le cadre de la théorie de l'information qui suppose que la diversité peut être mesurée de la même façon que l'information contenue dans un code ou un message. Les indices développés dans le cadre de cette théorie sont qualifiés de fonctions d'entropie. Soient p_k la fréquence de la catégorie k , et $\mathbf{p} = (p_1, \dots, p_k, \dots, p_S)$ la distribution de fréquences des catégories, l'indice de Shannon est

$$H_S(\mathbf{p}) = - \sum_{k=1}^S p_k \ln(p_k).$$

Exemple. On a calculé l'indice de Shannon dans le site de gauche.

```

siteG <- rep(c("poisson_lune", "thon_rouge", "marlin_bleu"), c(3,3,3))
siteG <- as.factor(siteG)
siteG
[1] poisson_lune poisson_lune poisson_lune thon_rouge thon_rouge thon_rouge
[7] marlin_bleu marlin_bleu marlin_bleu
Levels: marlin_bleu poisson_lune thon_rouge
(probG <- summary(siteG)/length(siteG))
marlin_bleu poisson_lune thon_rouge
 0.3333333 0.3333333 0.3333333
-sum(probG*log(probG))
[1] 1.098612

```

L'indice de Shannon est égale à 1.0986.

Exercice. Calculer l'indice de Shannon dans le site de droite.

Cette mesure peut être interprétée de plusieurs façons.

- L'indice de Shannon mesure la perte d'information due à la perte d'une entité.
- L'indice de Shannon est une mesure d'incertitude. Si nous tirons au hasard une seule entité de la collection, l'indice de Shannon mesure l'incertitude que nous avons sur le résultat, c'est-à-dire : quelle catégorie allons-nous tirer ?

L'indice de Gini-Simpson

L'indice H_{G-S} de Gini-Simpson [3, 8] est égal à la probabilité de tirer avec remise dans une collection deux entités appartenant à deux catégories différentes :

$$H_{G-S}(\mathbf{p}) = 1 - \sum_{k=1}^S p_k^2.$$

Gini [3] l'avait suggéré comme mesure de diversité écologique. Simpson [8] propose de nouveau cet indice à partir d'une mesure de concentration $\lambda = \sum_{k=1}^S p_k^2$ qui donne une valeur à la répartition des entités entre les catégories.

Exemple. On calcule l'indice de Gini-Simpson dans le site de gauche. et on trouve 0.6667.

Exercice. Calculer l'indice de Gini-Simpson dans le site de droite.

2.2 La fonction de rareté

2.2.1 Etude de la fonction

L'indice de Havrda et Charvat [5] généralise la richesse, l'indice de Shannon et celui de Gini-Simpson, dans une seule formule :

$$H_{H-C}(\mathbf{p}) = \frac{1 - \sum_{k=1}^S p_k^\alpha}{\alpha - 1}, \alpha \geq 0, \alpha \neq 1.$$

- ★ Lorsque $\alpha = 0$, $H_{H-C}(\mathbf{p}) = H_R$, la richesse ;
- ★ quand $\alpha \rightarrow 1$, $H_{H-C}(\mathbf{p}) \rightarrow H_S(\mathbf{p})$, l'indice de Shannon ;
- ★ et si $\alpha = 2$, alors $H_{H-C}(\mathbf{p}) = H_{G-S}(\mathbf{p})$, l'indice de Gini-Simpson.

A partir de H_{H-C} , Patil et Tallie [6] montrent que les trois indices, H_R , H_S et H_{G-S} , peuvent être réécrits comme des moyennes d'une fonction de rareté $R(p_k)$, la rareté d'une catégorie diminuant lorsque sa fréquence augmente :

$$H_{H-C}(\mathbf{p}) = \sum_{k=1}^S p_k \left(\frac{1 - p_k^{\alpha-1}}{\alpha - 1} \right).$$

Selon ces indices, la rareté d'une catégorie est donc égale à

$$R(p_k) = \begin{cases} (1 - p_k^{\alpha-1}) / (\alpha - 1) & \text{si } \alpha \neq 1 \\ -\ln(p_k) & \text{si } \alpha = 1 \end{cases}$$

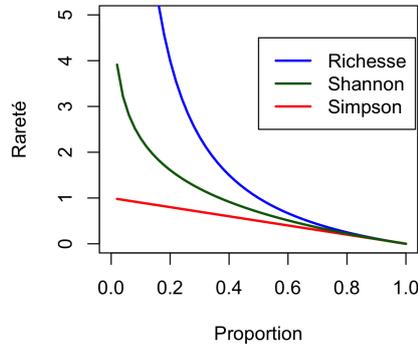
Les trois indices H_R , H_S et H_{G-S} ont les fonctions de rareté respectives :

$$R_R(p_k) = \frac{1}{p_k} - 1$$

$$R_S(p_k) = -\ln(p_k)$$

$$R_{G-S}(p_k) = 1 - p_k$$

Exercice. On trace les représentations graphiques des trois fonctions de rareté.



- 1) Comparer les fonctions.
- 2) Conclure sur l'impact des catégories rares sur la valeur de chaque indice.
- 3) Lequel est le moins sensible aux espèces rares ?

2.2.2 Communautés d'oiseaux et habitat

Blondel et ses collaborateurs [1] se sont intéressés aux cortèges avifaunistiques le long de successions écologiques. Ils ont démontré l'existence d'une convergence des communautés d'oiseaux dans l'ensemble des milieux fermés (forestiers) des 4 successions analysées.

Les données se trouvent dans le fichier de données `bf88` de la librairie `'ade4'`. Elles sont constituées d'une liste de 6 tableaux où chaque tableau contient en lignes une liste d'espèces d'oiseaux et en colonnes une liste de pays (Pologne, Bourgogne, Provence et Corse). Les 6 tableaux correspondent à 6 stades de végétation, de S1 habitat ouvert à S6 habitat fermé.

Exercice.

- 1) Récupérer les données et afficher un des tableaux, par exemple `bf88$S1`.
- 2) Calculer la diversité dans chacune des 24 communautés correspondant aux 6 habitats des 4 successions, en vous aidant des fonctions  ci-dessous.

★ Fonction permettant de calculer l'indice de Richesse sur les colonnes d'un data frame `df`

```
richness <- function(df) {
  res <- apply(df, 2, function(i) length(i[i!=0])-1)
  return(res)
}
```

★ Fonction permettant de calculer l'indice de Shannon sur les colonnes d'un data frame `df`

```
shannon <- function(df){
  interm <- function(y) {
    resi <- sapply(y, function(u) if(u != 0)
      (- (u / sum(y)) * log(u / sum(y))) else 0)
    return(sum(resi))
  }
  res <- apply(df, 2, interm)
  return(res)
}
```

★ Fonction permettant de calculer l'indice de Gini-Simpson sur les colonnes d'un data frame df

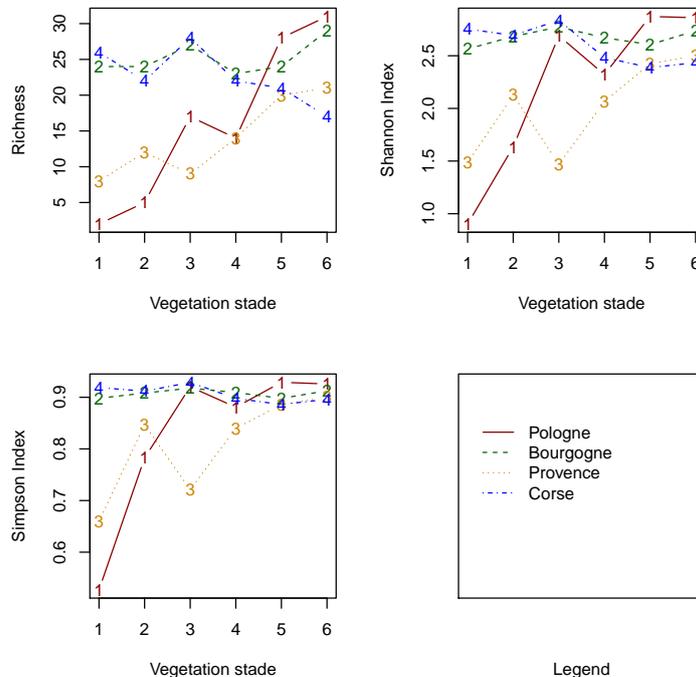
```
simpson <- function(df){
  interm <- function(y) {
    resi <- sapply(y, function(u) if(u != 0)
      ((u / sum(y))^2) else 0)
    return(1-sum(resi))
  }
  res <- apply(df, 2, interm)
  return(res)
}
```

★ Fonction permettant de calculer l'indice de biodiversité par l'une des trois méthodes (meth) à partir d'une liste de data frames (liste)

```
fun <- function(liste, meth) {
  res <- lapply(liste, meth)
  res <- as.data.frame(res)
  rownames(res) <- names(liste[[1]])
  return(res)
}
```

3) Comparer les valeurs données par les différents indices.

4) Analyser les résultats et ... comme un graphique est toujours plus parlant, voici une proposition de représentation (utiliser la fonction `matplot`).



5) Conclure.

Remarque. Regarder en particulier les valeurs de diversité attribuées aux communautés du troisième stade de végétation (S3) par les trois indices.

3 L'entropie quadratique

3.1 Présentation

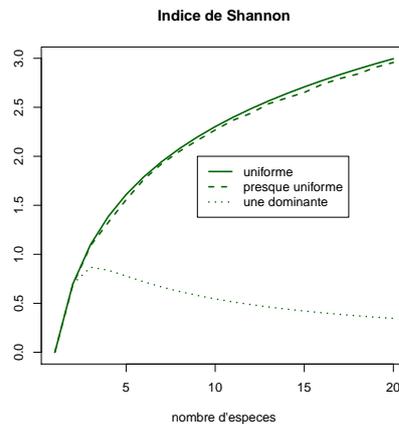
Deux critiques peuvent être faites à tous ces indices traditionnels, définis sur \mathcal{P} .

1. La première est que plus une collection possède de catégories, moins les indices de Shannon et de Gini-Simpson sont sensibles aux différences de fréquences entre catégories dans cette collection. La mesure de diversité qu'ils fournissent est alors très proche de la richesse.

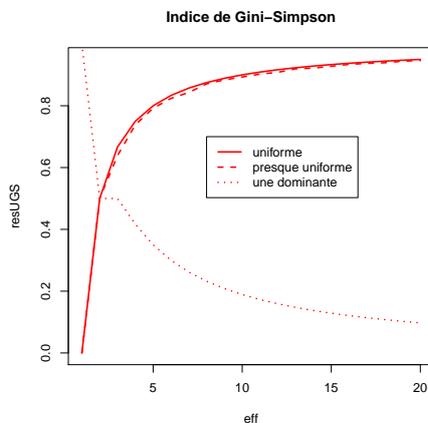
Exemple. Afin d'appréhender cette première critique, on construit les trois situations suivantes :

- (a) les catégories sont réparties uniformément,
- (b) les catégories varient autour de la loi uniforme,
- (c) une des catégories domine les autres qui sont elles réparties uniformément.

On trace l'indice de Shannon pour des collections variant de 5 à 20 catégories dans les trois cas proposés. Commenter.



On trace l'indice de Gini-Simpson pour des collections variant de 5 à 20 catégories dans les trois cas proposés. Commenter.



2. La deuxième critique concerne une des propriétés fondamentales de ces indices : ils sont invariables par permutation des catégories. D’après les définitions de ces indices, les catégories sont interchangeable : si dans une collection, une catégorie est remplacée par une autre alors qu’elle a des caractéristiques très différentes des autres, les indices précédents ne détecteront pas ce changement. La deuxième critique est donc que ces indices attribuerait la même diversité à une région dans laquelle seraient présents une autruche, un mulot et un lion, qu’à une région dans laquelle se trouveraient un campagnol, un mulot et un rat. Alors que d’un certain point de vue, nous avons l’intuition que la première région est plus diverse.

Pour avoir une mesure exhaustive de la diversité, il nous faudrait connaître les génomes de tous les êtres vivants, leurs phénotypes, leurs comportements donc leur rôle fonctionnel vis à vis des autres et du fonctionnement de la planète. Nous savons qu’une telle connaissance est actuellement impossible. Cette complexité de la diversité nous indique que chaque catégorie que nous considérons en écologie (espèce, allèle, etc) possède un rôle particulier et, en règle générale, n’est pas interchangeable. Pour tenir compte de ce fait, une suggestion est qu’une mesure générale de biodiversité doit non seulement inclure une mesure du nombre d’espèces, mais aussi une mesure du degré de différences entre ces espèces [9]. Ces différences peuvent être calculées par exemple à partir d’information génétique, phylogénétique, taxonomique, morphologique, comportementale, etc. (Fig. 3). L’interchangeabilité des espèces est une propriété traditionnellement prônée pour les mesures de biodiversité. Mais nous verrons pourtant que l’avantage des indices prenant en compte les différences entre les espèces est justement de ne plus vérifier cette propriété.

”A measure of the biodiversity of a site ought to say something about how different the inhabitants are from each other.” [4]

Considérons une collection d’entités regroupées en catégories. Soient $\mathbf{p} = (p_1, \dots, p_k, \dots, p_S)^\top$ la distribution de fréquences des catégories et $d_{k\ell}$ une mesure de la différence entre deux catégories k et ℓ , l’**entropie quadra-**

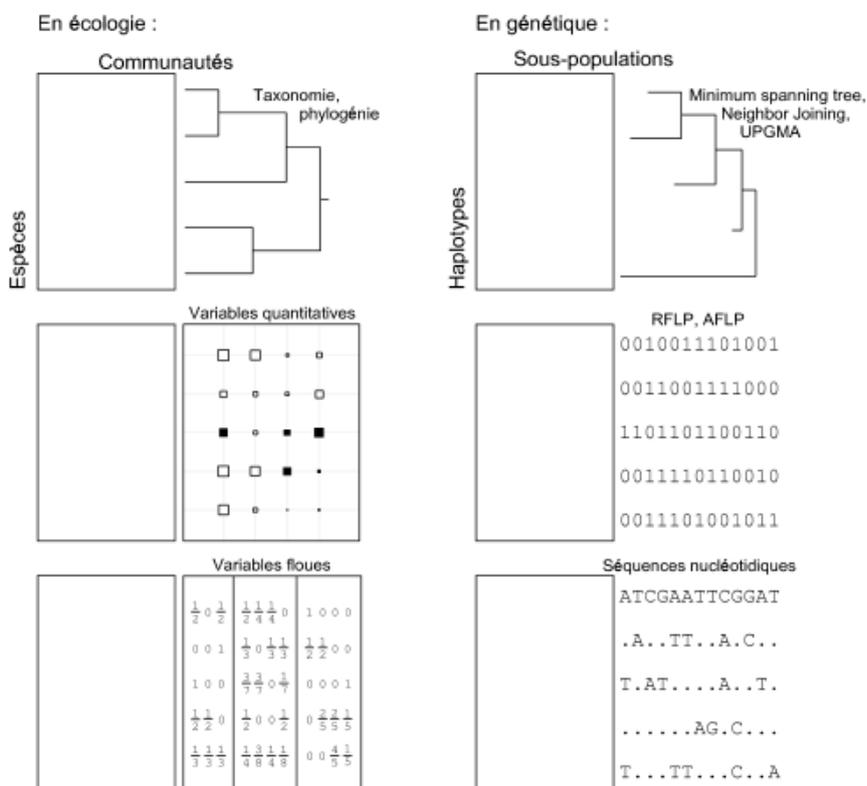


FIGURE 3 : Exemples de jeux de données utilisés pour mesurer la biodiversité. Les rectangles vides représentent des tableaux d'abondance ou de présences/absences.

tique est égale à :

$$H_{\mathbf{D}}(\mathbf{p}) = \sum_{k=1}^S \sum_{\ell=1}^S p_k p_{\ell} d_{k\ell},$$

où \mathbf{D} est la matrice contenant les dissimilarités entre catégories. Avec une écriture matricielle,

$$H_{\mathbf{D}}(\mathbf{p}) = \mathbf{p}^{\top} \mathbf{D} \mathbf{p}$$

Sous \mathbb{R} et dans la librairie `ade4`, l'entropie quadratique encore appelée coefficient de diversité de Rao se calcule à l'aide de la fonction `divc`.

3.2 La mesure des dissimilarités

Une similarité ou dissimilarité est une application numérique qui permet de mesurer le lien entre espèces. Pour une similarité, le lien est d'autant plus fort que sa valeur est grande.

Une dissimilarité d définie entre deux catégories k et ℓ vérifie les propriétés suivantes :

- ★ La dissimilarité d'un individu avec lui-même est nulle : $d_{kk} = 0$,
- ★ la dissimilarité entre deux individus différents est positive $d_{k\ell} \geq 0$,
- ★ la dissimilarité est symétrique : $d_{k\ell} = d_{\ell k}$.

Une distance vérifie en plus les deux propriétés suivantes :

- ★ la distance entre deux individus différents est **strictement** positive $d_{k\ell} = 0 \Rightarrow k = \ell$,
- ★ l'inégalité triangulaire : $d_{km} \leq d_{k\ell} + d_{\ell m}$. De nombreux indices de dissimilarité ne vérifient pas cette propriété.

Les distances entre les espèces prises deux à deux sont stockées dans une matrice dite matrice des dissimilarités. C'est une matrice carrée (S lignes et S colonnes). Dans \mathbb{R} , cette matrice est un objet de la classe `dist`.

Exemple. On considère la longitude et la latitude de cinq villes françaises.

```
longitude <- c(-4.4861,3.0573,4.8357,5.3698,7.7479)
latitude <- c(48.3904,50.62925,45.7640,43.2965,48.5831)
coordxy <- cbind(longitude,latitude)
villes <- c("Brest","Lille","Lyon","Marseille","Strasbourg")
rownames(coordxy) <- villes
coordxy
      longitude latitude
Brest      -4.4861 48.39040
Lille       3.0573 50.62925
Lyon        4.8357 45.76400
Marseille   5.3698 43.29650
Strasbourg  7.7479 48.58310
```

La distance à vol d'oiseaux c'est-à-dire la distance euclidienne entre Lyon (ligne 3) et Brest (ligne 1) est :

```
sqrt(sum((coordxy[3,]-coordxy[1,])^2))
[1] 9.684727
```

que l'on peut calculer automatiquement à l'aide de la fonction `dist` :

```
dist(coordxy)
      Brest      Lille      Lyon Marseille
Lille      7.868630
Lyon       9.684727  5.180093
Marseille 11.094439  7.688750  2.524642
Strasbourg 12.235518  5.117466  4.053176  5.796853

dist(coordxy, diag=T, upper=T)
      Brest      Lille      Lyon Marseille Strasbourg
Brest      0.000000  7.868630  9.684727 11.094439 12.235518
Lille      7.868630  0.000000  5.180093  7.688750  5.117466
Lyon       9.684727  5.180093  0.000000  2.524642  4.053176
Marseille 11.094439  7.688750  2.524642  0.000000  5.796853
Strasbourg 12.235518  5.117466  4.053176  5.796853  0.000000
```

3.3 ADN mitochondrial et zones géographiques [2]

Les données `humDNAM` se trouvent dans la librairie `ade4` et se présentent sous la forme de trois objets :

- des haplotypes d'ADN mitochondrial (`humDNAM$samples`) échantillonnés dans dix populations représentant cinq groupes régionaux de deux populations chacun,
- la structure des dix populations définies ci-dessus stockée dans `humDNAM$structures` : "Asia" (populations "Tharu" et "Oriental"), "West Africa" (population "Wolof" et "Peul"), "America" (populations "Pima" et "Maya"), "Europe" (populations "Finnish" et "Sicilian"), et "Middle-East" (populations "Israeli Jews" et "Israeli Arabs"),
- les dissimilarités entre haplotypes calculées en terme de nombre de sites de restrictions différents entre séquences. La matrice contenant les racines carrées de ces dissimilarités est euclidienne. En d'autres termes, \mathbf{D} est définie par la racine carré de l'objet (`humDNAM$distances`).

```
data(humDNAM)
names(humDNAM)
[1] "distances" "samples"   "structures"
is.euclid(humDNAM$distances)
[1] FALSE
is.euclid(sqrt(humDNAM$distances))
[1] TRUE
```

Les différences entre les haplotypes peuvent être visualisées par un réseau de longueur minimale (Fig. 4).

- 1) En utilisant les distances évolutives entre haplotypes, calculer la diversité génétique dans chaque population à l'aide de la fonction `divc` d'`ade4`.
- 2) Calculer aussi les diversités génétiques des populations à l'aide des fonctions `shannon` et `simpson`.
- 3) Commenter les résultats obtenus pour les trois indices.

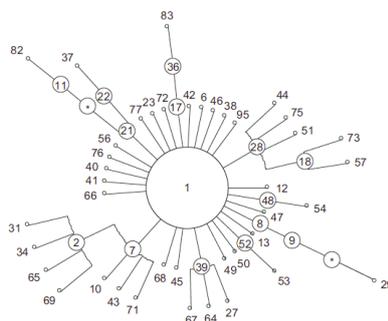


FIGURE 4 : Réseau de longueur minimale des 56 haplotypes trouvés dans les 10 populations étudiées. Chaque lien représente une mutation. Les astérisques indiquent deux haplotypes non trouvés dans les populations échantillonnées. L'haplotype 1 central a été trouvé dans toutes les populations. Les feuilles (haplotypes situés en extrémité du réseau) sont indiquées par des petits cercles, les nœuds intermédiaires par des cercles moyens et l'haplotype central par un grand cercle.

	Rao	Shannon	Simpson
oriental	0.43100189	1.2067368	0.5009452
tharu	0.48255042	1.5118191	0.6518536
wolof	0.65884298	1.7220666	0.7570248
peul	0.55952920	1.4346931	0.7134450
pima	0.06198035	0.3024839	0.1214412
maya	0.17092768	0.6142440	0.3243243
finnish	0.33280992	0.8989400	0.3659504
sicilian	0.61913580	1.6722281	0.6604938
israelij	0.67061144	1.5324659	0.7087442
israelia	0.64036818	1.6086954	0.6495726

L'entropie quadratique n'est pas bornée. Ses valeurs numériques dépendent de l'échelle des dissimilarités. Pour normaliser cet indice, une solution est de diviser sa valeur observée dans une collection par sa valeur théorique maximale :

$$H_{\Delta}^*(\mathbf{p}_{\text{obs}}) = \frac{H_{\Delta}(\mathbf{p}_{\text{obs}})}{\max_{\mathbf{p}} H_{\Delta}(\mathbf{p})}.$$

Dans cette formule, $H_{\Delta}(\mathbf{p}_{\text{obs}})$ est la valeur observée de l'entropie quadratique dans la collection, et $H_{\Delta}^*(\mathbf{p}_{\text{obs}})$ la valeur observée normalisée. La démarche utilisée pour maximiser l'entropie quadratique comprend trois étapes :

1. choisir des catégories dont le nombre est fini ;
2. choisir une matrice de dissimilarités entre ces catégories ;
3. rechercher une distribution de fréquences qui maximise la valeur de l'entropie quadratique pour ces catégories et cette matrice de dissimilarités.

En utilisant uniquement la matrice de distances entre haplotypes et la fonction `divcmx` (lire l'aide de la fonction pour bien comprendre les objets à extraire),

1. retrouver la valeur maximale théorique de l'entropie quadratique donnée ci-dessous

[1] 2.565789

2. retrouver le vecteur de fréquences qui conduit à cette valeur.

```

  1  2  6  7  8  9 10 11 12 13 17 18
  0  0  0  0  0  0  0  0  0  0  0  0

 21  22  23  27  28  29  31  34  36  37  38  39
0.0000 0.0000 0.0000 0.0000 0.0000 0.2293 0.0000 0.0000 0.0000 0.0790 0.0000 0.0000

40 41 42 43 44 45 46 47 48 49 50 51
  0  0  0  0  0  0  0  0  0  0  0  0

 52  53  54  56  57  64  65  66  67  68  69  71
0.0000 0.0000 0.0602 0.0000 0.1240 0.0000 0.0000 0.0000 0.0000 0.0000 0.1804 0.0000

 72  73  75  76  77  82  83  95
0.0000 0.0000 0.0000 0.0000 0.0000 0.2030 0.1241 0.0000

```

3. Utiliser la figure 4 pour commenter ce résultat.

Note

Les images du paragraphe 2 proviennent des différents sites suivants :

<http://www.gatech.edu/upload/pr/tbl58806.jpg>

http://www.coolwaterphoto.com/art_print/000889_print.jpg

<http://www.univ-ubs.fr/ecologie/Photos/visco2.jpg>

www.mazatlan.com.mx/.../fishing/blue-marlin.jpg

www.fishbase.org/.../gif/tn_Momol_u1.gif

www.gma.org/fogm/images/tuna.gif

Références

- [1] J. Blondel and H. Farr. The convergent trajectories of bird communities along ecological successions in european forests. *logia (Berlin)*, 75 :83–93, 1988.
- [2] L. Excoffier, Smouse P.E., and J.M. Quattro. Analysis of molecular variance inferred from metric distances among dna haplotypes : application to human mitochondrial dna restriction data. *Genetics*, 131 :479–491, 1992.
- [3] C. Gini. Variabilit e mutabilit. Studi economicoaguridici delle facoltta di giurizprudenza dell, Universite di Cagliari III, Parte II, 1912.
- [4] John L. Harper and David L. Hawksworth. Preface. In David L. Hawksworth, editor, *Biodiversity mesurement and estimation*, pages 5–12. Chapman & Hall, London, 1995.
- [5] M. Havrda and F Charvat. Quantification method of classification processes : concept of structural a-entropy. *Kybernetika*, 3 :30–35, 1967.
- [6] G.P. Patil and C. Taillie. Diversity as a concept and its measurement. *Journal of the American Statistical Association*, 77 :548–561, 1982.

- [7] C. E. Shannon. A mathematical theory of communication. *Bell System technical journal*, 27 :379–423, 623–656, 1948.
- [8] E.H. Simpson. Measurement of diversity. *Nature*, 163 :688, 1949.
- [9] P.H. Williams and C.J. Humphries. Biodiversity, taxonomic relatedness, and endemism in conservation. In P.L. Forey, C.J. Humphries, and R.I. Vane-Wright, editors, *Systematics and Conservation Evaluation*, volume 31, pages 269–287. Clarendon, Oxford, 1994.