

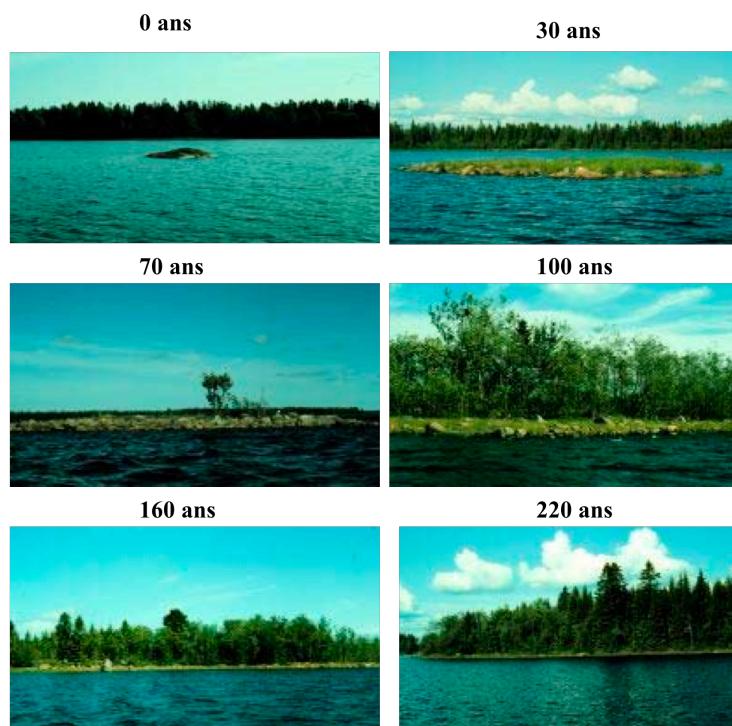
# Module « Génétique des populations moléculaire et coalescence »

## EXAMEN 2008-2009

Durée : 3h (documents autorisés)

### Sujet 1 :

Giles et Goudet (1997) [American Naturalist 149 : 507-526] ont étudié une espèce de campanule (*Silene dioica*, une petite herbe pérenne), dans des îles au large de la Suède. On connaît assez précisément l'âge de ces îles, qui émergent suite à l'élévation constante du plateau continental. Les photographies de la Figure 1, ci-dessous, représentent l'évolution typique de ces îles en fonction du temps :



**Figure 1 :** Photographies représentant des îles de différents âges, sur lesquelles ont été échantillonnées des spécimens de *Silene dioica*.

Giles et Goudet ont génotypé des échantillons de *Silene dioica*, grâce à six marqueurs enzymatiques. Ils ont ensuite mesuré des *F*-statistiques, et leurs résultats sont présentés dans le Tableau 1, ci-dessous.

**Tableau 1 :** *F*-statistiques en fonction de l'âge des îles.

Age des îles	Nbre d'îles	$F_{ST}$	$F_{IS}$
Jeunes (- 30 ans)	13	0.057	0.052
Intermédiaire (30 – 250 ans)	30	0.030	0.080
âgées (+ 250 ans)	9	0.066	0.098

1.- Du point de vue de la différentiation génétique entre les populations (mesurée par  $F_{ST}$ ) : les résultats du Tableau 1 correspondent-ils à vos attendus lorsque vous comparez les populations jeunes et les populations d'âge intermédiaire ? Justifiez votre réponse.

2.- De même, les résultats du Tableau 1 correspondent-ils à vos attendus lorsque vous comparez les valeurs de  $F_{ST}$  des populations d'âge intermédiaires et des populations les plus âgées ? Justifiez votre réponse.

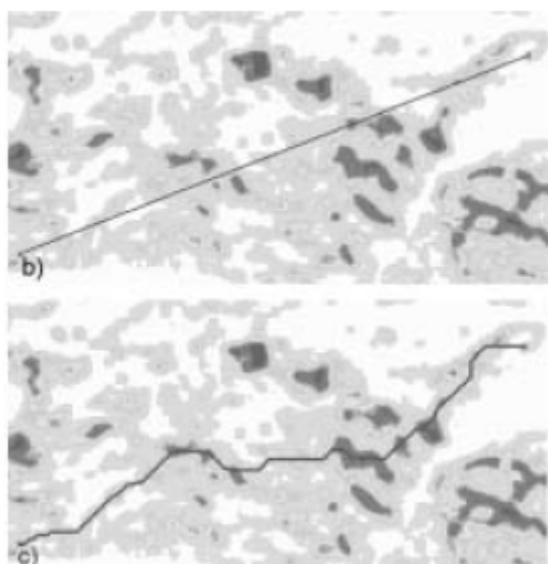
3.- Dans les populations âgées, le phénomène de succession écologique (un processus naturel d'évolution de l'écosystème, qui se traduit par un remplacement d'espèces au cours du temps) conduit à une survie réduite des plantules de *Silene* dans un stade avancé de la succession et à une mortalité accrue, suite au déplacement des *Silene* par des espèces arbustives. A partir de ces éléments, proposez des interprétations possibles au résultat identifié dans la question 2.

4.- Quelle tendance observez-vous pour  $F_{IS}$ , dans le Tableau 1 ? Que représente ce paramètre ?

5.- Il se trouve qu'aucune de ces valeurs de  $F_{IS}$  n'est significativement différente de zéro. Attendez-vous une relation particulière entre ce paramètre et l'âge des îles ? Pourquoi ? Justifiez précisément votre raisonnement.

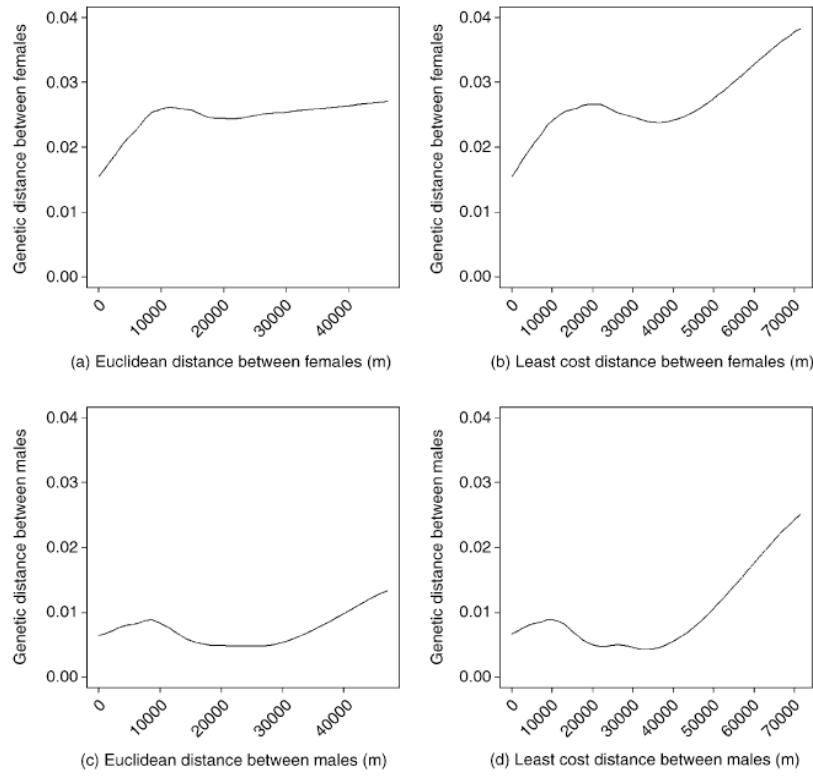
## **Sujet 2 :**

Dans une étude récente, Coulon *et al.* (2004) [*Molecular Ecology* **13** : 2841-2850] se sont intéressés à l'influence du paysage sur la dispersion chez le chevreuil. A partir d'un échantillon de 514 chevreuil, génotypés à 12 locus microsatellites, ils ont étudié la corrélation entre la différenciation génétique, mesurée par  $a_r$  (Rousset 2000), un équivalent de  $F_{ST}$  / (1 -  $F_{ST}$ ) calculé entre paires d'individus, et la distance géographique. Pour tester l'influence du paysage, ils ont considéré deux types de distance géographiques entre paires d'individus comme illustré sur la Figure 1 : (b) la distance Euclidienne entre individus, indépendante du paysage ; et (c) la distance de moindre coût, calculée en prenant en compte le fait que le chevreuil disperserait préférentiellement par les habitats lui étant favorable (ici la forêt).



**Figure 1:** Représentation des 2 types de distances géographiques estimées entre paires de chevreuils. Les zones de forte couverture forestière sont représentées en gris foncé. (b) distance Euclidienne entre 2 individus. Cette distance correspond à la distance minimale (ligne droite) entre ces points. (c) distance de moindre coût (courbe). Cette distance est estimée en minimisant le coût, calculé en fonction des valeurs numériques représentant le taux de couverture forestière de chaque pixel traversé par la courbe.

- 1.- D'après la Figure 1, les différences entre distances Euclidiennes et distances de moindre coût seront-elles équivalentes à petite et grande échelle géographique ?
- 2.- La représentation de la régression polynomiale entre les distances génétiques  $a_r$  et la distance géographique montrée en Figure 2 vous semble t-elle pertinente dans le cadre d'une analyse en isolement par la distance ? Notez que les auteurs n'ont pas représenté sur cette figure le nuage de points (distance génétique, distance géographique) représentant chaque paire d'individus analysés. Que pourrait apporter la représentation de ce nuage de points sur cette figure ?
- 3.- Que peut-on conclure des comparaisons entre les différentes distances sur la Figure 2 et dans le Tableau 1 ?
- 4.- Qu'apportent les résultats du Tableau 1 par rapport à ceux présentés sur la Figure 2 ?
- 5.- Que peut-on conclure des comparaisons entre mâles et femelles sur la Figure 2 et dans le Tableau 1 ? Quelles hypothèses peut-on faire à partir de ces résultats sur la biologie du chevreuil ?



**Figure 2 :** Représentation de la régression polynomiale entre les distances génétiques ( $a_r$ ) et les distances géographiques entre individus, pour les échantillons femelles (a, b) et mâles (c, d) en considérant les distances Euclidiennes (a, c) et les distances de moindre coût (b, d).

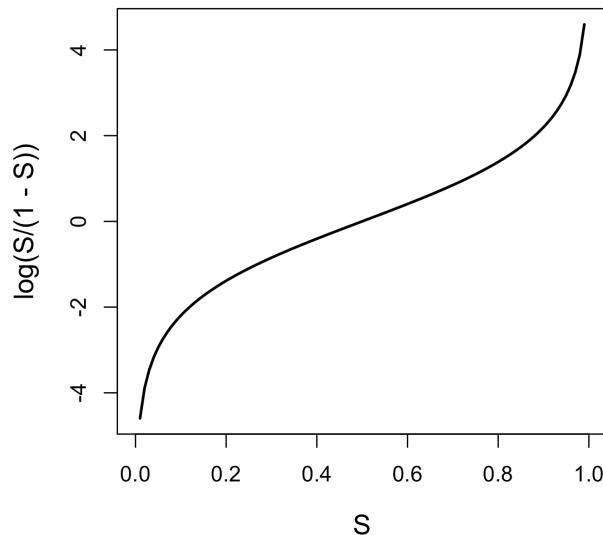
**Tableau 1 :** Probabilités associées aux tests de Mantel testant la corrélation entre les distances génétiques ( $a_r$ ) et le logarithme de la distance géographique dans les échantillons mâles et femelles, en fonction du type de distances considérées.

	Femelles	Mâles
Distances Euclidiennes	0.118	0.5
Distances du moindre coût	0.005**	0.401

10 000 permutations de la matrice des distances géographiques ont été réalisées pour le calcul de ces probabilités et les corrélations significatives au seuil de 0.05 sont indiquées par \*\*.

### **Sujet 3 :**

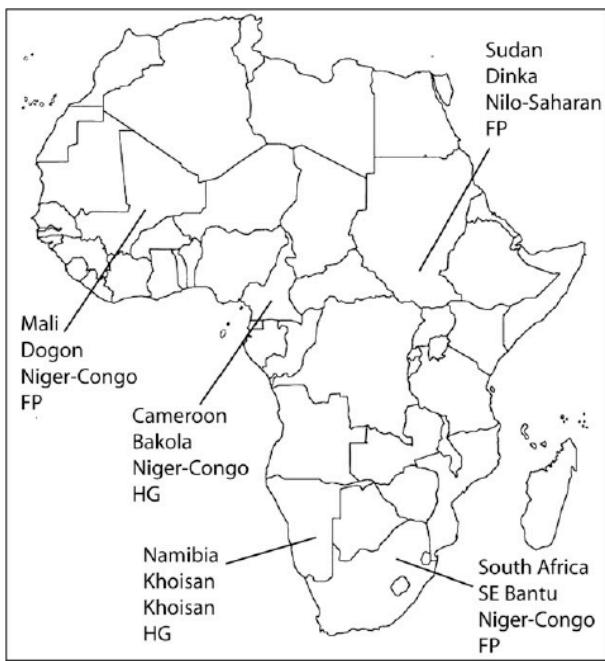
Fu (1997) [*Genetics* **147** : 915-925] a proposé un test de neutralité qui repose sur le calcul de la probabilité  $S$  d'observer dans un jeu de données neutres un nombre équivalent ou plus petit d'allèles que dans le jeu de données observé. La statistique  $F_s$  qu'il a définie est égale à  $F_s = \log(S / (1 - S))$



**Figure 1 :** relation  $\log(S / (1 - S))$  en fonction de la probabilité  $S$ .

1.- La Figure 1 ci-dessus vous montre la forme de la relation  $\log(S / (1 - S))$  en fonction de la probabilité  $S$ . Il se trouve que cette statistique  $F_s$  est également très sensible au régime démographique des populations. Notamment, des populations en expansion montrent des valeurs fortement négatives de  $F_s$ . Expliquez pourquoi. Quelle(s) autre(s) statistiques apporterai(en)t une information équivalente ?

2.- Pilkington *et al.* (2008) [*Molecular Biology and Evolution* **25** : 517-525] ont étudié des populations humaines en Afrique (Figure 2 ci-dessus), et ont réalisé des séquences sur l'ADN mitochondrial et le chromosome Y. Les résultats de leurs analyses statistiques sont donnés dans le Tableau 1 ci-dessous. Pour l'ADN mitochondrial (mtDNA) : qu'indiquent les valeurs du  $F_s$  de Fu ? Commentez en vous aidant de la carte et des indications qui y sont données.



**Figure 2 :** Carte géographique indiquant le pays d'origine et le nom des populations échantillonnées. Est également indiqué le mode de vie : FP (*Food-Producer*) = agriculteur ; HG (*Hunter-Gatherer*) = chasseur-cueilleur.

**Tableau 1 :** Analyse du polymorphisme sur l'ADN mitochondrial et le chromosome Y, dans les 5 populations africaines de la Figure 2.

	Population	N	Sites	Singletons	$F_s$ de Fu
mtDNA	Khoisan	25	10	2	-1.651
	Bakola	24	6	2	-1.956
	SE Bantu	50	16	6	-8.768**
	Dinka	23	15	12	-7.135**
	Dogon	49	8	6	-3.842**
Chrom. Y	Khoisan	25	10	1	0.183
	Bakola	25	14	3	-0.127
	SE Bantu	46	10	3	1.525
	Dinka	23	8	2	1.771
	Dogon	40	12	4	-0.071

N : nombre d'individus ; Sites : nombre de sites polymorphes ; Singletons : nombre de singulons.

3.- Comparez ces résultats avec ceux obtenus sur le chromosome Y (NRY dans le Tableau 1) : sachant que l'ADN mitochondrial et le chromosome Y sont des molécules différentes, que ni l'un ni l'autre ne recombine, qu'ils diffèrent par leur hérédité, etc. : proposez des hypothèses pour expliquer les différences que vous observez dans le Tableau 1 entre chromosome Y et ADN mitochondrial.

4.- Ray *et al.* (2003) [*Molecular Biology and Evolution* **20** : 76-86] ont étudié par des simulations l'effet d'une expansion dans une population subdivisée. Pour cela, ils ont simulé une population initiale (un dème) à partir de laquelle se produit une expansion spatiale (de nouvelles populations sont fondées, successivement, pour envahir l'espace adjacent ; les populations nouvellement fondées échangent ensuite des migrants à un taux  $m$  avec les populations adjacentes). A la fin d'une expansion, la population centrale est échantillonnée, et des mesures de polymorphisme peuvent être faites. Les résultats sont indiqués dans le Tableau 2, ci-dessous. Commentez ces résultats.

**Tableau 2 :** Valeurs du  $F_s$  de Fu, en fonction du produit  $Nm$  où  $N$  est la taille du dème échantillonné, et  $m$  le taux d'immigration dans ce dème.

$Nm$	$F_s$ de Fu
5	1.56
10	-0.42
20	-3.63
50	-11.38
100	-15.93
200	-19.61
250	-21.81
500	-23.13
1000	-23.98

5.- A la lumière des résultats de Ray *et al.* (2003), discutez des différences observées entre les données sur l'ADN mitochondrial et le chromosome Y. Proposez des scénarios évolutifs et/ou démographiques pour expliquer les différences entre marqueurs mitochondriaux et ceux liés au chromosome Y.