

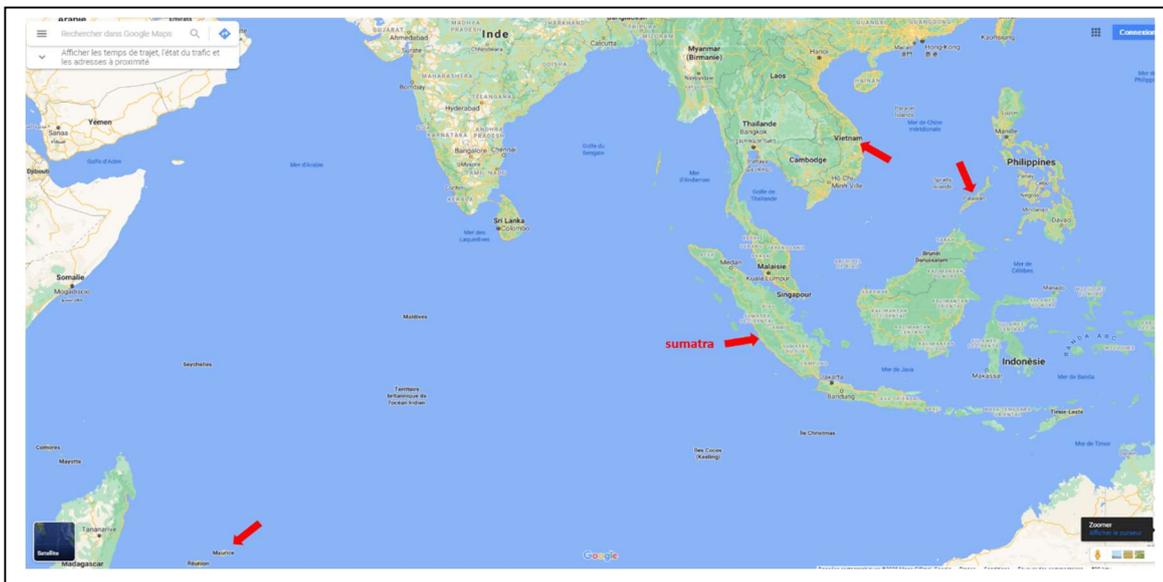
UE « Génétique évolutive et quantitative »

TP 1: génétique des populations

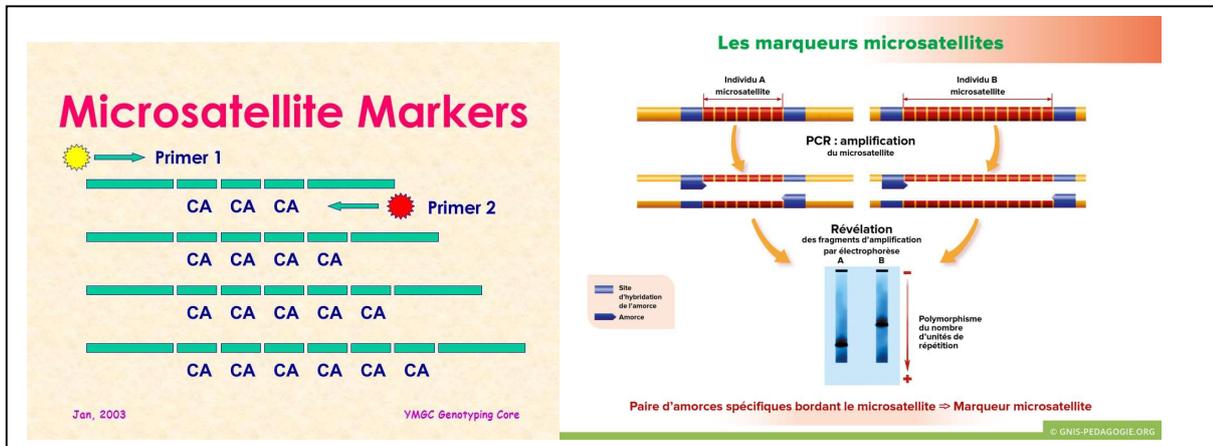
I - Description des systèmes de reproduction et de la différenciation génétique des populations d'une espèce végétale à l'aide de marqueurs moléculaires.



Dipterocarpus sp est une espèce diploïde d'arbre monoïque à fleurs bisexuées, communément répandue en **Asie du Sud-Est**. Des botanistes ont observé que la pollinisation croisée était essentiellement assurée par différentes espèces d'insectes. Pour étudier les systèmes de reproduction et la différenciation génétique à l'échelle des populations chez cette espèce, 4 populations originaires de l'île de Sumatra **(1)**, de l'île de Palawan dans l'archipel des Philippines **(2)**, du Viet Nam **(3)** ainsi que de l'île Maurice **(4)** où l'espèce a été implantée récemment, ont été génotypées avec **16 marqueurs microsatellites**.



Un marqueur microsatellite est un locus présentant un polymorphisme de longueur (allèles de différentes tailles), dû à un nombre de répétitions variables d'un motif nucléotidique (ex : **GCGCGCGC**, ou **GATAGATAGATAGATA**,...).



Vous allez réaliser l'analyse du polymorphisme génétique de ces populations avec le logiciel GENETIX. Lancer l'exécutable Genetix.exe. Puis, dans « Fichier », importer le fichier de données « plant_16_microsatellite_data.txt » (au format SEPARATEUR « tabulation », et ALLELES CODES PAR « 3 chiffres ») : votre fichier est prêt à être analysé. Les allèles aux marqueurs microsatellites analysés sont composés de répétitions de 4 nucléotides (**GATA_(n)**), et sont codés par la taille du fragment PCR amplifié. Ainsi un individu de génotype «203207» est hétérozygote, avec les allèles de 203pb et 207pb (= 1 répétition d'écart entre les 2 allèles). Un individu de génotype «271271» est homozygote pour l'allèle de 271pb.

1 – Analyser la diversité allélique et l'hétérozygotie.

- Pour cela cliquez sur l'onglet « Variabilité ».
- Observez les résultats, et notamment le tableau de bas de page montrant les valeurs moyennes sur l'ensemble des locus (hétérozygoties et nombre d'allèles par locus).
- Quelles sont les populations les plus (moins) polymorphes ? Qu'en déduisez-vous ?

2 – Test de panmixie sur le F_{IS} .

- Pour cela cliquez sur « Fstats », puis « Test sur Fis », puis sélectionnez « Fis par pop sur permutations » (nombre de permutations = 1000, ne pas cocher la case « Calcul locus par locus »).

On rappelle que $F_{IS} = \frac{(H_e - H_o)}{H_e}$.

- Sachant que « Réel » est le F_{IS} moyen (sur l'ensemble des locus) dans chaque population, et que « % val. > » est la proportion de valeurs de F_{IS} permutés qui sont > à la valeur réelle de F_{IS} (et donc que cette proportion correspond à quoi ?), que déduisez-vous concernant le système de reproduction dans ces populations ?

- Estimez le taux d'autofécondation dans chaque population. Pour cela, utilisez la formule $F_{is} = \frac{s}{2-s}$. Interprétez les résultats.

3 – Estimation de la différenciation génétique (F_{ST}).

- Calculez le F_{ST} moyen (sur l'ensemble des locus) par paires de populations. Pour cela cliquez sur « Fstats », puis « Traitement par paires », puis sélectionnez « Theta par paires / permutations ».
 - Qu'observez-vous ? Essayez d'expliquer quelles forces évolutives sont à l'origine des différences génétiques entre ces populations.

- Sous l'hypothèse d'un modèle en îles (où $F_{ST} = \frac{1}{1+4Nm}$), estimez le nombre de migrants par génération (Nm) pour les paires de populations 1-4 et 1-3. Pour cela, cliquez sur « Fstats », puis « Traitement par paires », puis sélectionnez « Theta, Nm par paires ». Que concluez-vous ?

II – Simulation de la dérive génétique.

Simulez avec R/Rstudio, l'évolution de la fréquence d'un allèle à un locus biallélique de type SNP sous dérive génétique uniquement, à l'aide du script « TP1_drift_simulations.R ». Sous R (ou RStudio), chargez le script, sélectionnez tout le code (CTRL + A) et lancez la commande (CTRL + R, ou bouton « Run »). Simulez 200 réalisations (= populations indépendantes) sur 1000 générations (nbgen=1000), avec une fréquence initiale de l'allèle de 0.5 ($p_0=0.5$), et déterminez l'effet de la taille de la population ($n=1000, 10000$), sur l'évolution ainsi que sur la valeur finale de l'hétérozygotie (H_e) et du coefficient de consanguinité (F).

Remplissez le tableau ci-dessous (H_e et F calculés sur les simulations, à la dernière génération):

	nbgen=1000
n=1000 et $p_0=0.5$ ($H_{e\text{initial}} = 0.5$)	$H_e=$ $F=$
n=10000 et $p_0=0.5$ ($H_{e\text{initial}} = 0.5$)	$H_e=$ $F=$