

# M1 MABS

## Examen terminal de Génomique et génétique statistique (durée 2h)

(EM7BBSBME2) - Janvier 2012

**Pour les exercices 3 à 8, inscrivez vos réponses directement sur le sujet.**

### Exercice 1

La figure ci-dessous représente les arbres de coalescence à un locus donné pour deux échantillons de 4 copies du locus. Chacun des échantillons provient d'une population différente. Les deux arbres sont à la même échelle (une unité de hauteur correspond au même nombre de générations pour l'arbre A et pour l'arbre B).

- 1) Que peut-on conclure de ces arbres concernant l'histoire démographique des deux populations?

Pour l'un de ces deux échantillons, on observe 5 sites polymorphes au locus. Les séquences observées pour ces 5 sites sont les suivantes :

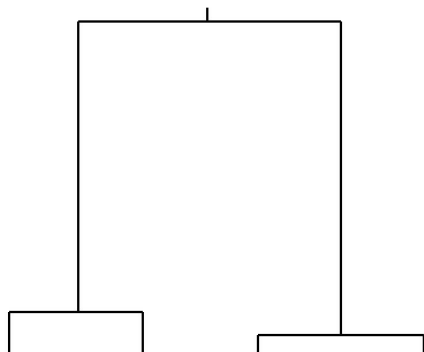
**copie 1** : 0 - 1 - 0 - 1 - 0

**copie 2** : 0 - 1 - 0 - 1 - 0

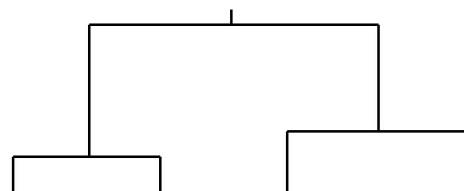
**copie 3** : 1 - 0 - 1 - 0 - 1

**copie 4** : 1 - 0 - 1 - 0 - 1

- 2) Au vu de ces données, pensez-vous qu'il s'agisse de l'échantillon A ou de l'échantillon B (justifier qualitativement votre réponse) ?



Arbre A



Arbre B

## Exercice 2

On a décidé de conduire une analyse de génétique d'association avec des lignées de maïs qui proviennent de deux régions du monde : le Mexique et la France. Le phénotype étudié est la hauteur de la plante et les lignées ont été génotypées pour 10 000 marqueurs polymorphes.

Les résultats de trois modèles différents vous sont proposés :

**Modèle (1) :** l'analyse a été menée avec un modèle linéaire simple d'association comportant la moyenne générale et l'effet d'un marqueur. Pour 100 marqueurs, l'hypothèse nulle est rejetée au risque choisi, ces 100 marqueurs sont déclarés associés au phénotype.

**Modèle (2) :** préalablement à l'analyse d'association, on a étudié la structuration des lignées avec le logiciel STRUCTURE. L'analyse d'association a été menée avec un modèle linéaire comportant la moyenne générale, la matrice de structuration en deux groupes donnée par le logiciel et l'effet du marqueur. 11 marqueurs sont déclarés associés au phénotype.

**Modèle (3) :** l'analyse a été menée avec un modèle linéaire comportant une moyenne générale, un effet fixe modélisant l'origine géographique des lignées et l'effet marqueur. 12 marqueurs sont déclarés associés au phénotype.

### 1) Quel(s) modèle(s) vous semble(nt) adapté(s) à l'analyse d'association de ces lignées ? Et pourquoi ?

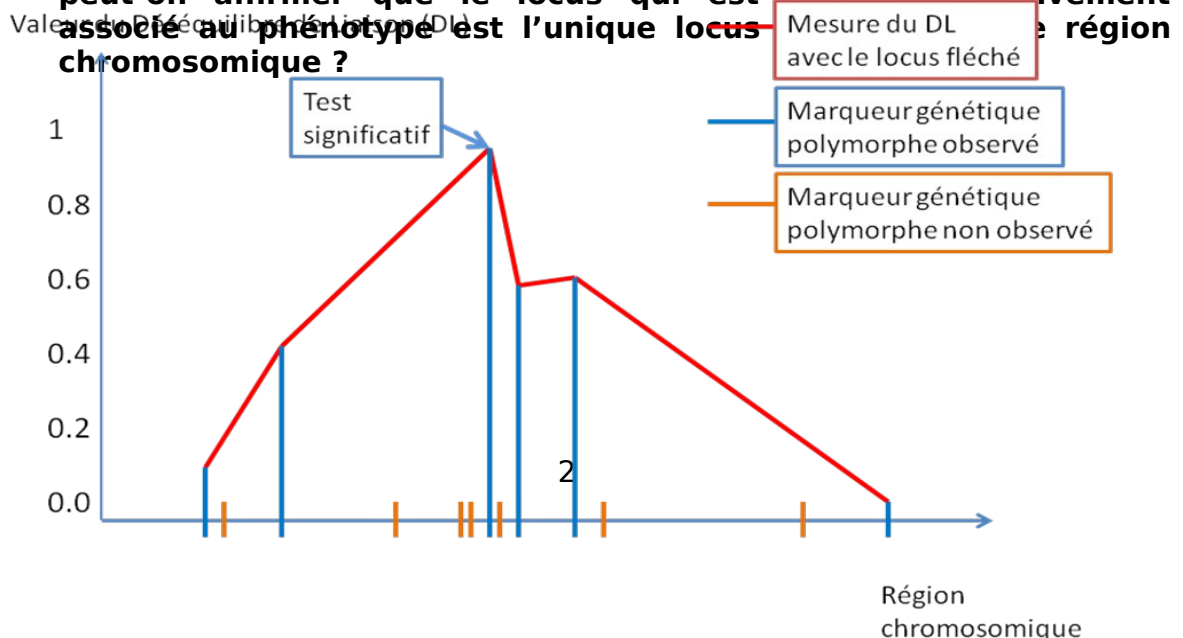
Une erreur a été commise dans la population des lignées françaises et on vous signale que 20 des lignées françaises sont des pleines-sœurs.

### 2) Que faites vous de cette information ?

On s'intéresse à une région chromosomique pour laquelle un locus a été jugé significativement associé au phénotype pour le modèle (2), c'est-à-dire que l'hypothèse nulle a été rejetée au risque choisi. La figure qui suit montre le déséquilibre de liaison de ce locus avec les marqueurs qui lui sont proches physiquement.

### 3) Quelle est la mesure du déséquilibre qui est liée à la puissance du modèle (2) ?

### 4) Sachant que c'est cette mesure qui est présentée dans la figure 2, peut-on affirmer que le locus qui est jugé significativement associé au phénotype est l'unique locus de la région chromosomique ?



**NOM :**

**Prénom :**

## **M1 MABS**

### **Examen terminal de Génomique et génétique statistique (durée 2h)**

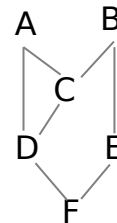
**(EM7BBSBME2) - Janvier 2012**

**Pour tous les exercices suivants, répondez directement  
sur le sujet**

#### **Exercice 3**

Soit la généalogie :

- en utilisant la méthode des chemins, calculez le coefficient de consanguinité des individus C et F. Détaillez les calculs que vous réaliserez.



#### **Exercice 4**

a)

Il ne peut y avoir de consanguinité chez un individu que si ses géniteurs sont apparentés.

☐ vrai

☐ faux

**b)**

La valeur génétique additive (A) représente la fraction de la valeur génétique (G) dont on peut prédire la transmission de parent à descendant : un parent transmet, en espérance, la moitié de sa valeur génétique additive.

☐ vrai

☐ faux

**c)**

L'héritabilité ( $h^2$ ) d'un caractère

- est spécifique d'une population donnée

☐ vrai

☐ faux

- représente la proportion de la variance des phénotypes pour ce caractère qui est de nature additive

☐ vrai

☐ faux

- est constante pour un caractère donné

☐ vrai

☐ faux

- permet de prédire si l'amélioration génétique par sélection sera efficace

☐ vrai

☐ faux

## **Exercice 5**

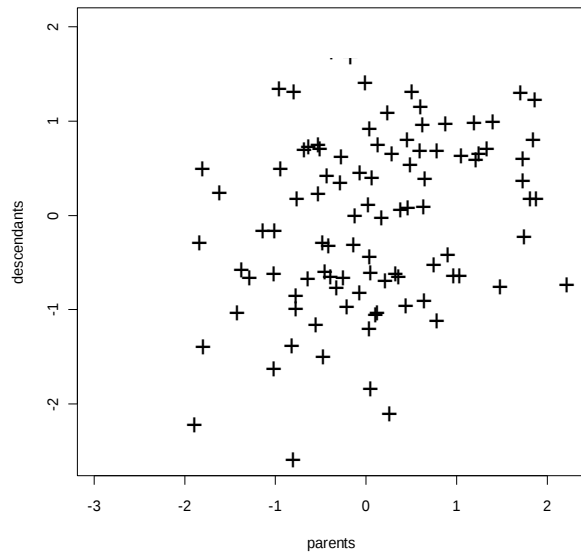
Le graphique ci-joint représente la relation entre un caractère (centré et réduit) mesuré chez une centaine de couples père (x, en ordonnées) -descendant (y, en abscisses). Pensez-vous que le caractère testé est

☐ la taille ( $h^2=0.75$ ) ?

☐ le poids à la naissance ( $h^2=0.30$ ) ?

☐ le niveau d'études atteint ( $h^2=0$ ) ?

Dans le graphique, la covariance entre les variables  $x$  et  $y$  est de 0.35 ; la variance de  $x$  et de  $y$  est de 1.

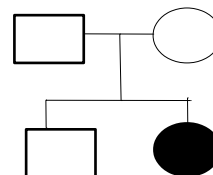


## Exercice 6

Soit une maladie due à un gène  $T$  à 2 allèles,  $D$  et  $d$ . On note  $d$  l'allèle de susceptibilité et  $q$  sa fréquence. Les probabilités d'être malade conditionnelles au génotype (pénétrances) sont :  $f_{dd}=1$ ,  $f_{Dd}=0$ ,  $f_{DD}=0$ .

**1. Le mode de transmission de  $T$  est récessif :** Vrai      Faux

Soit la famille F1 : le 1<sup>er</sup> enfant, non malade, est noté E1. Le 2<sup>ème</sup> enfant, malade, est noté E2. Le père et la mère ne sont pas malades.



**2. Dans le tableau suivant, identifiez par une croix le(s) génotype(s) possible(s) :**

Génotype	DD	Dd	dd
Père			
Mère			
E1			
E2			

### 3. Donnez les probabilités du (des) génotype(s) possible(s).

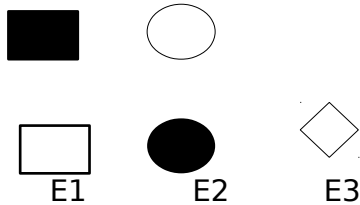
Génotype	DD	Dd	dd
Père			
Mère			
E1			
E2			

## Exercice 7

Soit une maladie dominante due à un gène T à 2 allèles, D et d.

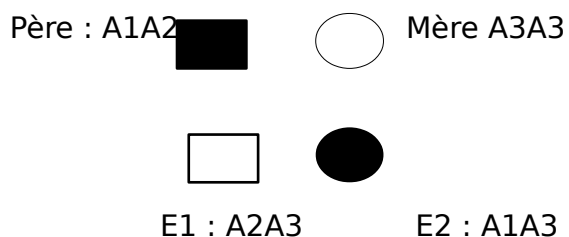
On note d l'allèle de susceptibilité et q sa fréquence. Les probabilités d'être malade conditionnelles au génotype (pénétrances) sont :  $f_{dd}=1$ ,  $f_{Dd}=1$ ,  $f_{DD}=0$ .

Soit la famille F2 : le père est malade, la mère non malade. L'enfant (E1) est non malade, le 2<sup>ème</sup> (E2) est malade. Le phénotype au 3<sup>ème</sup> enfant (E3) n'est pas connu



### 1. Donnez la probabilité que l'enfant E3 soit malade.

On analyse un marqueur M qui a 4 allèles (A1, A2, A3, A4) pour tester si Le locus T est génétiquement lié ou non à M. Dans la famille F2, les génotypes au marqueur sont :



- Soit les phases alléliques (arrangement des allèles sur le chromosome) suivantes :

Phase 1: A1D | A2d

Phase 2: A1d | A2D

Phase 3: A1d | A2d

**2. Listez celles qui sont possibles chez le père et donnez leur probabilité a priori :**

- Le tableau suivant liste l'ensemble des gamètes parentaux possibles de cette famille.

**3.a Reportez dans le tableau les probabilités de transmission des gamètes du père et de la mère en fonction du taux de recombinaison ( $\theta$ ) entre M et T. Vous distinguerez plusieurs cas selon les différentes phases possibles chez le père et/ou la mère.**

Gamète :	A1D	A1d	A2D	A2d	A3D	A3d
Père						
Mère						

## **Exercice 8**

Soit  $j$  le nombre d'allèles qui peuvent être identiques par descendance (IBD) chez 2 frères et sœurs.

**1. Donnez les valeurs possibles de  $j$  et leurs probabilités  $P(j= .)$**

**2. A priori, quelle est la proportion ( $\pi$ ) d'allèles IBD entre 2 frères?**

Soit l'échantillon de familles nucléaires : père, mère et 2 enfants atteints. Tous les membres de la famille sont génotypés pour un marqueur M ayant 4 allèles (A1, A2, A3, A4)

Le tableau suivant liste pour chaque famille (Fi) les génotypes au marqueur du père, de la mère et des 2 enfants atteints (E1, E2).

**3. Pour chaque famille Fi, reportez dans le tableau les probabilités  $P(j=0)$ ,  $P(j=1)$ ,  $P(j=2)$  et la proportion ( $\pi$ ) d'allèles IBD.**

	<b>Père</b>	<b>Mère</b>	<b>E1</b>	<b>E2</b>	<b><math>P(j=0)</math></b>	<b><math>P(j=1)</math></b>	<b><math>P(j=2)</math></b>	<b><math>\pi</math></b>
F1	A1A 2	A3A 4	A1A3	A2A3				
F2	A1A 2	A3A 4	A2A3	A2A3				
F3	A1A 2	A3A 4	A1A3	A2A4				
F4	A1A 3	A3A 4	A3A3	A3A3				
F5	A1A 3	A3A 4	A1A3	A1A4				
F6	A1A 2	A3A 3	A1A3	A2A3				
F7	A1A 2	A3A 3	A1A3	A1A3				

En supposant que l'échantillon contient 10 familles F1, 10 familles F2, 5 familles F3, 10 familles F4, 5 familles F5, 2 familles F6 et 2 familles F7.

**4.a Quel est le nombre total  $n_j$  de paires de germains pour  $j=0, 1$  ou  $2$  ?**

**4.b Que vaut  $\pi$  ?**