

FST

Dans la nature, les individus d'une même espèce se dispersent et se répartissent en plusieurs populations. Chacune des populations peut évoluer et acquérir une originalité génétique, caractérisé par des fréquences alléliques plus ou moins différentes de celles des populations voisines. On assiste à une diversification des populations, cependant, le plus souvent, elles continuent à échanger des individus par migration (pas d'isolement total). Ce flux migratoire tend à homogénéiser les populations (brassage génétique) et à limiter leur diversification .

I- LES MIGRATIONS

Soit une espèce fractionnée en plusieurs populations, pouvant échanger des individus migrants.

Soit m , le flux génétique d'une population j vers une population i . (taux d'individus se reproduisant, à chaque génération, dans la population i mais qui sont issus de la population j)

Si $m = 0$ alors les populations sont isolées

Si $m = 0,5$: les 2 populations n'en forment qu'une, panmixique

Si les populations sont très grandes, (peu de dérive génétique), toutes convergent vers la même fréquence allélique. (moyenne des fréquences initiales de chaque population)

- **Exemple de migration unidirectionnelle :**

Soit une île recevant un flux migratoire constant du continent.

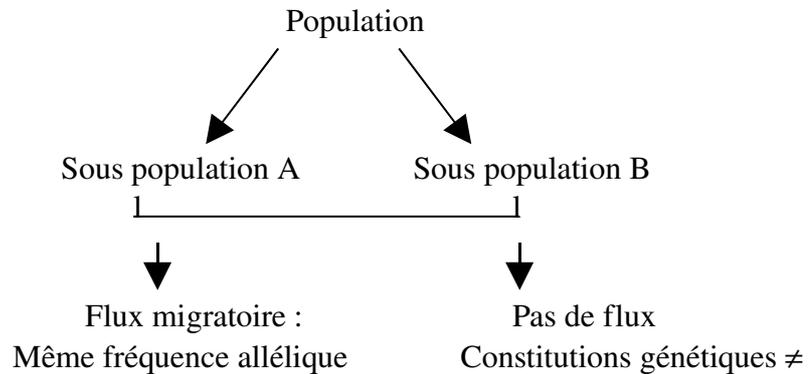
- Au bout d'un grand laps de temps, la fréquence allélique de l'île tend vers la fréquence allélique du continent
- Si les migrations ont lieu dans les 2 sens, l'équilibre est fonction des fréquences initiales des populations et des taux respectifs de migration.

Exemple d'application de ce modèle : population noire aux USA. Dans la population noire, la proportion de gènes originaires de la population blanche est de 11% si on se fonde sur le locus de Duffy. Réciproquement, des gènes d'origine africaine sont introduits dans la population blanche , mais plus faiblement (effectif réduit)

II- CONSTITUTION GENETIQUE

La colonisation d'une région aboutit à la formation de 2 sous populations, qui tendent à avoir les mêmes fréquences allélique si le flux migratoire persiste.

Si il n'y a plus de flux, les populations évoluent de manière indépendante, menant à des constitutions génétiques différentes.

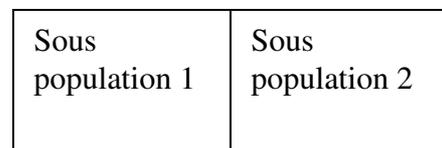
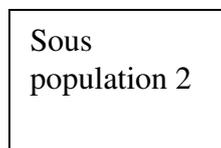
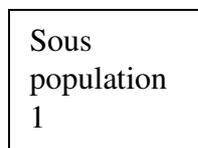


- **Effet Wahlund**

Soit les données suivantes : pour 2 allèles A et a, on a

	Population 1	Population 2
Fréquence allèle A	0,6	0,2
Fréquence allèle a	0,4	0,8
Fréquence AA	0,36	0,04
Fréquence Aa(Hétérozygotie)	0,48 (H1)	0,32 (H2)
Fréquence aa	0,16	0,64

Soit 2 sous populations A et B : on compare la **population globale** à **la population pancmitique qui résulteraient de la fusion des deux sous-populations :**



Population globale

AA : 0,2
 Aa : 0,4 et $\bar{A} = 0,4$ et $\bar{a} = 0,6$
 aa : 0,4

Population fusionnée

AA : 0,16
 Aa : 0,48 (HT)
 aa : 0,36

- Les valeurs de la population globale ne correspondent pas aux proportions de **Hardy-Weinberg** ($p^2, q^2, 2pq$)

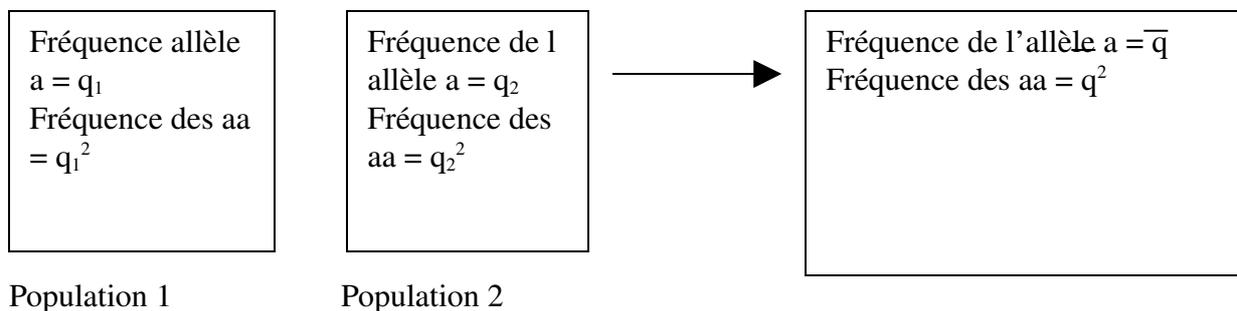
- la population globale présente plus d'homozygotes que la population fusionnée
- la population globale présente moins d'hétérozygotes que la population fusionnée

Le déficit en hétérozygotes est appelé **effet Wahlund**. Cet effet est observé dans le cas d'une divergence des fréquences alléliques.

La fusion des populations diminue donc la fréquence des génotypes homozygotes. Il en résulte une diminution de la naissance d'enfants atteints de déficiences génétiques dues à des gènes récessifs.

NB : utilisation de la variance :

Pour illustrer la rupture d'isolement (effet Wahlund) on peut utiliser la variance :



Fréquence moyenne de l'allèle a = $\bar{q} = \frac{1}{2}(q_1 + q_2)$

Fréquence moyenne du génotype aa = $q^2 = \frac{1}{2}(q_1^2 + q_2^2)$

Par conséquent, la fréquence des aa avant la fusion – la fréquence des aa après la fusion :

$$= q^2 - \bar{q}^2$$

$$= \sigma^2 \text{ la variance.}$$

III- DEFINITION FST

Le déficit en hétérozygotes est exprimé par Wright par :

$$F_{st} = \frac{(H_T - H_S)}{H_T}$$

Cet indice mesure la réduction relative des hétérozygotes, lié aux différences alléliques entre populations. **C'est un indice de diversification** des populations.

Si les populations sont identiques, $F_{st} = 0$

L'effet Wahlund doit être utilisé pour étudier des sous populations. Si on applique ce principe à une population unique, la baisse d'hétérozygotes peut être interprété comme consanguinité alors que cela est du aux différences de fréquences alléliques.

Fst est un indice de déviation par rapport aux proportions de Hardy-Weinberg.

Sans la connaissance exacte de la structure de la population (souvent fractionnée en sous populations), il est impossible d'attribuer ce phénomène à un effet de consanguinité ou d'isolement génétique.

- **Hétérozygotie**

HI : hétérozygotie observée en moyenne par individu sur l'ensemble des sous populations

Hi : hétérozygotie observée dans la ième sous-population

$HI = \sum(\text{de } i \text{ à } k) Hi/k$ (k populations)

HS : hétérozygotie attendue par individu pour une sous population supposée à l'équilibre de HW

$HS = 1 - \sum(\text{de } i) xi^2$, S (xi = fréquence du ième allèle) (et pour la Sième population)

Et HS est la moyenne des HS sur les k sous populations :

$\overline{HS} = \sum(\text{de } S \text{ à } k) HS/k$

HT : hétérozygotie attendue par individu, en supposant la population globale à l'équilibre de HW. Avec xi, fréquence moyenne de l'allèle Ai sur l'ensemble des k sous populations :

$HT = 1 - \sum(\text{de } i \text{ à } k) xi^2$

On a :

F_{IS} : $(\overline{HS} - HI) / \overline{HS}$: qui mesure la réduction éventuelle d'hétérozygotie des individus à l'intérieur de leur sous-populations. Si les populations sont à l'équilibre HW, alors $F_{IS} = 0$. Si F_{IS} est négatif les sous populations présentent un excès d'hétérozygotie

F_{ST} : $(HT - \overline{HS}) / HT$: donne l'effet de la subdivision entre les sous-populations et la population, et donne la réduction d'hétérozygotie dans les sous populations, lié aux différences de fréquences alléliques. Si toutes les populations sont à l'équilibre et possèdent les mêmes fréquences alléliques: $F_{ST} = 0$, sinon si les fréquences alléliques moyennes sont différentes, (dérivé génétique), il est positif. Cela reflète l'action de la dérive (diversification) et des migrations (homogénéisation). **F_{ST} est un indicateur de la cohésion de l'ensemble des sous populations considérées.**

F_{IT} : $(HT - HI) / HT$. Donne la réduction d'hétérozygotie globale entre l'individu et la population théorique.

Les trois indices sont liés par :

$(1 - F_{IT}) = (1 - F_{IS})(1 - F_{ST})$: si toutes les populations sont à l'équilibre de HW, $F_{IS} = 0$ et $F_{IT} = F_{ST}$. Si toutes les populations sont à l'équilibre de HW et ont les mêmes fréquences alléliques, les indices ont une valeur nulle.

Pour les locus multialléliques, on utilise une généralisation de F_{ST} qui est le G_{ST} de Nei : $G_{ST} = (HT - \overline{HS}) / HT$

Données : On étudie un locus diallélique (a et A) dans 12 populations, 9 d'entre elles sont monomorphes. (population de souris)

Fréquences observées	Fréquence de l'allèle A	Fréquence de l'allèle	Fréquence des
----------------------	-------------------------	-----------------------	---------------

		a	hétérozygotes
Population 1	0,04	0,96	0,08
Population 2	0,08	0,92	0,15
Population 3	0,07	0,93	0,14
Populations 4 à 12	-	-	-

- $H_I : \Sigma(\text{de } i \text{ à } k)H_i/k$ soit : $H_I = [(9 \times 0) + 0,08 + 0,15 + 0,14]/12 = 0,0275$
- $\bar{H}_S : 1 - \Sigma(\text{des } i) x_i^2$
- $H_S : \Sigma(\text{de } S \text{ à } k)H_S/k = [9 \times 0 + [(0,04)(0,96) \times 2 + 2(0,08)(0,92) + 2(0,07)(0,93)]]/12 = 0,0295$
- $H_T : 1 - \Sigma(\text{de } i \text{ à } k) x_i^2$ avec $x_1 = (9 \times 0 + 0,04 + 0,08 + 0,07)/12 = 2 \times 0,0158 \times 0,9842 = 0,0312$

On a donc :

$F_{IS} = 0,07$, exprime l'effet du aux croisements non panmixtiques dans les sous populations
 $F_{ST} = 0,05$ exprime les effets du à la subdivision, la variation moyenne dans tous les locus est de 0,465 : niveau de différenciation très grand.

$\overline{F_{IT}} = 0,12$

NB : pour les populations humaines, la valeur F_{ST} est plus élevée car elle inclut différentes données (groupes sanguins...) qui augmentent l'hétérozygotie. Chez les indiens Yanomana du Brésil, par exemple, vivant dans des petits villages, H_S est faible (du au faible effectif). Cependant la diversité totale n'est pas élevée. La population totale doit descendre d'un petit nombre d'individus et n'est pas revenue au niveau originel d'hétérozygotie.

IV- DIVERSIFICATION

Caractérisation du degré de diversification selon Wright :

F_{ST} : de 0 à 0,05 : faible, de 0,05 à 0,15 modérée ; de 0,15 à 0,25 importante et au delà, très importante. Dans les populations humaine, il existe une faible diversification.

La majeure partie des différenciations se trouve à l'intérieur des groupes majeurs

Exemple chez les indiens Yanomana :

7,7% de la variabilité du groupe est due à la différenciation entre villages et 92% de la variabilité totale se trouve dans un même village.

Dans l'espèce humaine, la majorité de la diversité génétique réside au niveau des individus de chaque population locale (Nei 1982) : environ 85% de la diversité totale apparaît liée aux différences entre individus appartenant à un même groupe régional (tribu ou nation).

Plusieurs mesures de distances génétiques ont été proposées, souvent liées au F_{ST} .

La plus utilisée est celle de Nei (1987) : D.

Avec $D = \ln(I)$ avec $I =$ identité normalisée.

Les distances calculées correspondent aux dates estimées par les paleo-anthropologues pour les divergences entre groupes. Les estimations sont également en accord avec les études menées sur l'ADN mitochondrial (Cann 1987), qui situent en Afrique, il y a 200 000 ans le groupe ancêtre de l'espèce humaine.

V- MUTATIONS ET MIGRATIONS

Les mutations et les migrations agissent sur la diversification génétique des populations qui tendent alors vers un état d'équilibre stationnaire.

Soit n populations, d'effectif N_e , avec un taux de migration m et un taux de mutation μ d'un allèle vers l'un des $k-1$ autres allèles.

$$G*ST = 1/(4N_e\mu + 1)$$

GST est l'expression de FST sous sa forme multiallélique et $a = [n/(n+1)]^2$

VI- POLYMORPHISME PAR INSERTION ALU

Utilisation de FST :

« Polymorphisme par insertion Alu dans trois populations du Nord de l'Italie »
Santovito A Al. ; 2004 (Colloque)

Le polymorphisme à six différents locus d'insertion de séquence Alu a été examiné dans 3 populations humaines du Piémont nord oriental : Postua (village de 570 habitants à isolement géographique et culturel partiel, fermé à l'immigration et avec un degré élevé de mariages consanguins et, comme contrôle, deux agglomérations urbaines limitrophes, Cavaglia (3690 habitants) et Bella ville de 47 000 habitants.

Les valeurs d'hétérozygotie moyenne pour chaque locus et pour chaque population et la valeur de F_{ST} par locus sont calculés. F_{ST} varie pour les locus de 0,004 à 0,0593. La valeur moyenne est de 0,038.

Comparé aux valeurs de grandes régions du monde, F_{ST} de 0,038, a une valeur plus faible que celle des autres régions du monde, à titre indicatif, celle du sud est de l'Asie est de 0,067 et celle de l'Afrique 0,086 (Nasidze et al, 2001)

Postua présente une valeur d'hétérozygotie moyenne inférieure (0,291) à celle de Cavaglia et de Biella (0,329 et 0,343). Les valeurs les plus basses d'hétérozygotie peuvent être dues à la consanguinité

A ces données correspond cependant une valeur moyenne de F_{ST} des populations étudiées, de 0,038. Cette donnée peut être expliquée par le caractère strictement limitrophe des populations examinées, la valeur globale de F_{ST} ne pouvant être donc très élevée.

Sources :

- **Génétique des populations, Daniel L. Hartl**
- **Cours migrations et division des populations, université de Tours**
- **Alu polymorphisms in three population from Northern Italy, Santovito.A (colloque)**