

APPORT DE LA GENETIQUE DES POPULATIONS A LA COMPREHENSION DE LA PREHISTOIRE

Charles SUSANNE*

Nous avons souvent tendance à décrire des individus, leur morphologie, leur pathologie. En fait, nous ne vivons pas en tant qu'individus mais en tant qu'êtres sociaux, notre unité est la population. Les restes fossiles n'étaient pas "fossiles", ils faisaient partie de populations vivantes.

L'anthropologiste est intéressé à décrire la variabilité des populations humaines et son origine. La définition d'une population est cependant difficile.

L'espèce humaine est un groupe d'individus, où tous les mariages sont potentiellement fertiles. Cependant, ces unions sont limitées par de nombreuses barrières d'origine géographique, socio-économique, ethnique et psychologique par exemple. Ces barrières, si elles fonctionnaient totalement, limiteraient des isolats. En réalité, les barrières ne sont jamais totalement imperméables et l'endogamie n'est jamais parfaite. Il en résulte que les populations doivent être définies par leurs barrières.

Dimensions des populations

La présence des différentes barrières résulte en une limitation du choix du partenaire et donc de la dimension d'une population.

Nous pouvons estimer que pendant toute notre évolution le nombre d'individus dans chaque population a été très limité (tableau 1), en fait même dans nos sociétés industrielles les cercles de mariages sont réduits.

Pour appréhender l'évolution, le nombre d'individus dans une population n'est pas une information suffisante, la répartition d'âge et de sexe est importante également. Si dans une grande population, on peut estimer que la distribution des sexes soit équilibrée au moins à l'âge de la procréation, dans une population limitée, la distribution aléatoire peut être totalement biaisée. En cas de guerre par exemple, une sérieuse perte de jeunes hommes peut former un biais important.

Il en résulte la nécessité de définir un nombre effectif d'individus. La probabilité que deux gènes de deux individus différents proviennent du même individu masculin est de $1/4 N_m$ (avec N_m le nombre d'individus masculins) et $1/4 N_f$ du même individu féminin (avec N_f le nombre d'individus féminins). Les deux gènes proviennent donc du même individu avec la probabilité de

$$\frac{1}{4N_m} + \frac{1}{4N_f} = \frac{1}{4N_e}, \text{ où } N_e \text{ est la dimension effective de la population.}$$

* Vrije Universiteit Brussel, Laboratorium voor Antropogenetica, Pleinlaan 2, 1050 Bruxelles, Belgique.

$$Ne = \frac{4NmNf}{Nm + Nf}$$

Cette valeur dépendra en fait essentiellement des valeurs les moins élevées de Nm ou de Nf, et peut diverger profondément du nombre total d'individus (tableau 2).

Loi de Hardy et Weinberg

La théorie de la génétique de populations est basée sur la loi classique de Hardy et Weinberg, qui définit une situation où la fréquence des gènes reste constante de génération en génération (tableau 3).

Cette non variabilité de fréquence constitue une situation d'absence d'évolution réalisée aux conditions qu'il n'y ait pas de migration, de sélection, de mutation, que la population soit grande et que les mariages se déroulent au hasard (panmixie).

Cette situation s'avère donc théorique et cette loi nous permet plutôt d'étudier l'influence que le non respect des conditions pourrait avoir en termes d'évolution.

L'étude de l'évolution ne peut se séparer de la compréhension de la sélection naturelle, des mutations, des migrations. Nous voudrions cependant insister sur l'aspect aléatoire lié à la présence de populations limitées.

Histoire d'une famille ou histoire de gènes

Une famille a toujours une histoire complexe où chaque individu a deux parents et où après n générations, chaque individu peut avoir 2ⁿ ancêtres. L'histoire d'un gène décrit cependant mieux l'évolution biologique et est plus simple en ce sens que chaque gène n'a qu'un ancêtre gène.

La fig. 1 compare un exemple de généalogie et l'histoire de ces gènes où l'on voit que certains gènes peuvent disparaître et que d'autres seront présents en de multiples copies.

Dérive génique.

Comme nous l'avons vu, une partie essentielle de l'histoire de l'Homme et de sa préhistoire se déroule dans des populations à effectif très limité, des groupes de 50 individus sont la norme (tableau 1). Lorsque la population devient plus grande, des migrants la quittent pour de nouveaux territoires et cette migration résulte en un choix au hasard d'un petit nombre de gènes. La population fille aura donc un pool génétique différent de celui de la population mère.

La formation d'une nouvelle génération résulte toujours d'un choix aléatoire de gamètes impliqués dans la fécondation et la formation des nouveaux zygotes.

Un tel choix aléatoire considéré dans une population infinie n'influence pas la fréquence des gènes. Mais lorsque la population est très peu nombreuse, le hasard peut résulter en de grandes fluctuations de fréquences de gènes. Ce processus est connu sous le nom de dérive génique, il correspond à la non transmission aléatoire de certains gènes et à la fixation d'autres.

A la limite, il importe plus de suivre la non transmissibilité que la transmission elle-même. La chance pour un gène à un locus déterminé de ne pas être transmis

à un enfant est de $1/2$

à n enfants est de $(1/2)^n$

Au niveau d'une population la probabilité de non transmissibilité est de

$$P = \sum_1^x \frac{\lambda x}{2^x}$$

où λx est la proportion de la population ayant x enfants. Cette probabilité dépend de x le nombre d'enfants mais aussi de la variation de ce nombre d'enfants (tableau 4).

Les facteurs aléatoires peuvent influencer aussi la fréquence d'un gène dans une population, leurs influences seront inversement proportionnelles à la dimension de la population. Dans une population de grandeur N où un gène a_i est présent en n_i copies, sa fréquence est de $p_i = n_i/N$.

Dans la nouvelle génération, l'espérance de la probabilité du gène est égale à la fréquence précédente mais sa variance est une fonction de cette fréquence mais aussi de la dimension de cette population.

$$E(p_i) = p_i$$

$$V(p_i) = p_i(1-p_i)/2N$$

Le tableau 5 illustre cette variation.

Si les fluctuations aléatoires peuvent être estimées d'une génération à l'autre suivant les règles de la distribution binomiale, il est plus difficile d'estimer l'évolution de la fréquence des gènes après de nombreuses générations. Des simulations informatiques ont été utilisées à cette fin, elles montrent que le résultat de la dérive génique serait la fixation d'un allèle et la perte de l'autre.

Effet fondateur.

Un cas particulier de la dérive génique est l'effet fondateur : notre histoire regorge d'exemples de migrations de nombre limité d'individus créant une nouvelle population, et participant donc ainsi à un effet de dérive génique.

La population Jicaque fournit un exemple d'effet fondateur, mais également d'histoire de gènes, où en fonction de variation de fécondité et de mortalité la structure d'une population se modifie.

Les Jicaques ont été fondés il y a environ un siècle par seulement 7 personnes, qui se sont isolées volontairement ainsi que leurs descendants, leur généalogie a pu être reconstruite. Le tableau 6 fournit le nombre de personnes par génération ainsi que le niveau de consanguinité. Celui-ci est naturellement fonction aussi de la faible dimension de la population, il diminue en 5e génération sous l'effet de l'introduction de quelques migrants. Le tableau 7 est plus riche d'informations lorsqu'il reconstruit la probabilité de l'origine des gènes issus des 7 fondateurs. Il montre que dès la 2e génération le poids des fondateurs peut être très différent (facteur 2,5 en 2e génération, facteur 5 en 5e génération).

Les Touaregs Kell Kummer (Chaventré, 1972) décrivent un exemple relativement similaire d'effet fondateur et de dérive génique. A la fin du 17e siècle, un groupe dominant de Touaregs (les Kel Tademaket) vivent au Sud Sahara. Un groupe fait sécession sous la conduite de Kari-Denna, ce groupe impose petit à petit sa domination et migre vers le Sud Mali. Sous l'occupation française de la fin du 19e siècle, les Kel Kummer résistent et beaucoup seront tués, en 1970 seulement 367 individus (171 hommes et 196 femmes) vivent encore. Toute la population ne connaît que 156 fondateurs.

Le tableau 8 donne un résumé de l'origine des gènes de ces différents fondateurs. L'asymétrie est flagrante, puisque le seul Kari Denna contribue pour 20 à 11% au pool génétique, seul 15 fondateurs contribuent à 70% du gene pool et 25 fondateurs à 80% de celui-ci.

En fait, bien d'autres exemples d'effet fondateur sont connus. Ainsi, les Huttérites, un groupe d'anabaptistes migrant d'Europe vers les Etats Unis en 1872, forment 3 colonies dont la variabilité génétique est importante dès l'origine. La secte religieuse du Old Order Amish migre en Pennsylvania entre 1720 et 1770, et crée de nombreuses colonies isolées où une pathologie récessive est fréquemment observée. Les Dunker sont une autre secte migrant au 18e siècle et dont 55 groupes de moins de 100 personnes se sont créés, où une dérive génique importante explique aussi les variations géniques observées entre groupes.

La fréquence élevée de retinitis pigmentosa dans la population de Tristan da Cunha, de porphyrie chez les Afrikaners, de Chorée d'Huntington chez les Tasmaniens ressort également d'un effet fondateur. La porphyrie par exemple, a été introduite chez les Afrikaners par un couple de migrants, Gerrit Jansz, un fermier hollandais arrivant à Kaapstad en 1688 et épousant une orpheline du nom de Ariaantje Jacobs. La chorée d'Huntington trouve son origine chez une femme Huguenotte quittant l'Angleterre en 1848.

La distribution de certains groupes sanguins trouve également son origine dans la dérive génique : il en est ainsi du groupe O extrêmement fréquent chez les Amerindiens, et l'absence du gène r' et r chez les mêmes Amerindiens.

Une étude du DNA non codant situé près du gène de la β globine montre quatre variants très fréquents dans l'ensemble des populations humaines : trois sont présents avec des fréquences semblables dans toutes les populations mondiales, le quatrième n'est présent que chez les Africains. Il est possible d'estimer qu'un effet fondateur a été engendré par la migration d'Afrique des

populations humaines d'*Homo erectus*, il est même possible d'estimer ce nombre de migrants. Ils auraient été au nombre de 50 pour 70 années de migration ou de 500 pour 200 années de migration : ces chiffres sont étonnamment bas mais ne sont pas improbables.

Migration

Si l'effet fondateur correspond à une histoire de fissions de populations, la migration correspond à une histoire de fusion. Les isolats ne sont jamais totalement fermés, des échanges entre populations ont toujours lieu. Ces migrations peuvent tamponner l'effet génétique de la dérive génique, elles ont un effet homogénéisateur.

Conclusion

Une population tend toujours vers un équilibre sous l'influence de la sélection, de la migration et de la dérive génique. Mais plusieurs populations se retrouvent dans des équilibres différents par des effets variés de sélection, de migration et de dérive : les facteurs aléatoires ne sont plus des facteurs d'uniformisation mais de diversification.

L'isolement d'une population n'est jamais totale, une contribution génétique de migrants sera toujours observée, même si ce niveau de migration est en continuel changement. Il en résulte que nulle région n'est suffisamment stable que pour résulter en une situation d'équilibre.

La (pré)histoire des groupes humains est une histoire de groupes relativement isolés, de groupes de nombre d'individus limité et donc de dérive génique.

Elle est aussi une histoire où les facteurs culturels et sociologiques résultent en des fissions et des fusions : des fissions seront facteurs d'ouverture d'endogamie et donc des facteurs de diversification, et des fusions en tant que facteurs d'homogénéisation.

Notre unité est et a toujours été la population. Les populations n'étaient pas fossiles, l'homme préhistorique ne savait d'ailleurs pas qu'il était préhistorique.

Addendum : Ce travail a été réalisé dans le cadre d'un programme Human Capital Mobility ERBCHRXCT 930193

Dimensions

Chimpanzé	20 - 40
Homo récent	20 - 50
Néolithique (village)	50 - 1000
Amazonie	20 - 80
Pygmées (campement)	10 - 100
Australian aborigènes	20 - 50
Maya (village)	500
New Guinea (village)	100 - 300
Belgique (groupe de mariages)	300 - 1000
<u>Densité par km²</u>	
Chimpanzé	0,07 - 0,09
Omo (Australopithecus)	0,006 - 0,016
Chasseur cueilleur	1
agriculture récente	10
agriculture traditionnelle	40
Belgique	1000
<u>Chasseur -cueilleur</u>	
Pygmées	0,2
Esquimau (Caribou)	0,04
(Groenland)	0,06
(Aleut)	0,6
Aborigènes Australiens	0,03
<u>Agriculture primitive</u>	
Nouvelle Guinée	4
Maya	20

Tableau 1 : Estimation des dimensions et de la densité de populations humaines

N	Nm	Nf	Ne
100	1	99	3,96
100	5	95	19
100	10	90	36
100	30	70	84
100	50	50	100
100	70	30	84
100	90	10	36
100	5	95	19
100	1	99	3,96

Tableau 2 : Grandeur effective (Ne) en fonction du nombre d'hommes (Nm) ou de femmes (Nf) pour un total N=100

1) supposons 2 allèles	A ₁ et A ₂	
de fréquence	p ₁ et p ₂	avec p ₁ +p ₂ =1
Les génotypes	A ₁ A ₁ A ₁ A ₂	A ₂ A ₂
auront comme fréquence	p ₁ ² 2p ₁ p ₂	p ₂ ²

Dans la nouvelle génération, la nouvelle fréquence de A₁ est

$$p'_1 = \frac{2p_1^2 + 2p_1p_2}{2} = p_1(p_1 + p_2) = p_1$$

et de A₂
$$p'_2 = \frac{2p_1^2 + 2p_1p_2}{2} = p_2(p_1 + p_2) = p_2$$

Les fréquences des gènes sont constantes.

2) Supposons des mariages panmictiques		Fréq. des enfants		
	Fréq.	A ₁ A ₁	A ₁ A ₂	A ₂ A ₂
A ₁ A ₁ × A ₁ A ₁	p ₁ ⁴	p ₁ ⁴		
A ₁ A ₁ × A ₁ A ₂	4p ₁ ³ p ₂	2p ₁ ³ p ₂	2p ₁ ³ p ₂	
A ₁ A ₁ × A ₂ A ₂	2p ₁ ² p ₂ ²		2p ₁ ² p ₂ ²	
A ₁ A ₂ × A ₁ A ₂	4p ₁ ² p ₂ ²	p ₁ ² p ₂ ²	2p ₁ ² p ₂ ²	p ₁ ² p ₂ ²
A ₁ A ₂ × A ₂ A ₂	4p ₁ p ₂ ³		2p ₁ p ₂ ³	2p ₁ p ₂ ³
A ₂ A ₂ × A ₂ A ₂	p ₂ ⁴			p ₂ ⁴
-		p ₁ ²	2p ₁ p ₂	p ₂ ²

Les fréquences des génotypes sont constantes.

Tableau 3 : Loi de Hardy et Weinberg

Au niveau individuel

D' un parent vers un enfant

$$1/2$$

D'un parent vers n enfants

$$(1/2)^n$$

Au niveau populationnel

La probabilité de non transmission est égale à $P = \sum x \frac{\lambda^x}{2^x}$

avec λ^x la proportion de la population ayant x enfants

Exemples de 2 populations ayant une moyenne de 2 enfants :

A avec $\lambda 2 = 1$

$$B \text{ avec } \lambda 0 = \frac{1}{2} \lambda 4 = \frac{1}{2}$$

$$P = \frac{2}{2^2} = 0,25$$

$$P = \frac{1}{2} \cdot \frac{1}{2^0} + \frac{1}{2} \cdot \frac{1}{2^4} = 0,53$$

Tableau 4 : Probabilité de non transmission

$$V(pi) = \frac{pi(1 - pi)}{2N}$$

Exemple 1 : $pi = 1/2$
avec 95%

Exemple 2 : $pi = 1/4$
avec 95%

N=5	0,3 à 0,7	N=5	0 à 0,52
N=10	0,32 à 0,68	N=10	0,06 à 0,44
N=50	0,42 à 0,58	N=20	0,12 à 0,38
N=100	0,44 à 0,56	N=50	0,17 à 0,33
N=500	0,48 à 0,52	N=100	0,19 à 0,31

Tableau 5 : Variance de la fréquence pi du gène après une génération de reproduction dans une population de dimension N

		Consan- guinité	N♂	N♀	Ne	N
	1	0	4	3	6,94	7
	2	0	19	15	33,5	34
	3	0,068	96	80	174,5	176
	4	0,092				247
	5	0,065+				101

+ Cette diminution est due à l'impact de migrants

Tableau 6 : Consanguinité et grandeur effective (Ne) de la population Jicaque.

Génération		Nb Fondateurs								Migrants	
		Léon	Franc.	Caci ana	Juan	Poli- naria	Pedro	Petro -na	Total	Indien s	Hybri des
1	7	143	143	143	143	143	143	143	1000		
2	34	88	74	74	191	176	199	199	1000		
3	176	66	78	78	167	151	208	208	957	43	
4	247	54	68	68	162	146	195	195	888	81	31
5	101	24	31	31	132	126	129	129	602	186	212

Tableau 7 : Probabilité des gènes trouvant leur origine chez les fondateurs de la population Jicaque.

Génération	1 à 5	6-7	8-9	10-11	12-13	14 à 16
Grandeur de la population	126	376	601	653	268	245
Fondateurs						
1093 & 1096	198)	142)	158)	146)	155)	186)
1 & 2	199) 571	110) 350	116) 370	105) 357	105) 361	116) 426
1331 & 1332	174)	98)	106)	106)	101)	124)
1919	30	66	76	64	71	107
1959	47	55	70	60	63	71
2060	16	60	48	33	37	33
2067	16	21	49	45	43	46
1968	29	35	39	34	40	53
2062	0	44	25	24	25	27
1628 & 1629	38	20	19	15	19	24
2063	0	23	23	15	22	20
Les 15 principaux	747	674	729	647	681	807
Les 10 suivants	103	112	102	111	101	78
Les 131 autres	150	214	169	242	218	115

1-2 = Kaari-Denna et son épouse

1331-32 = parents de l'épouse de Kaari-Denna

Tableau 8 : Touaregs Kel Kummer : Origine des gènes (en ‰)

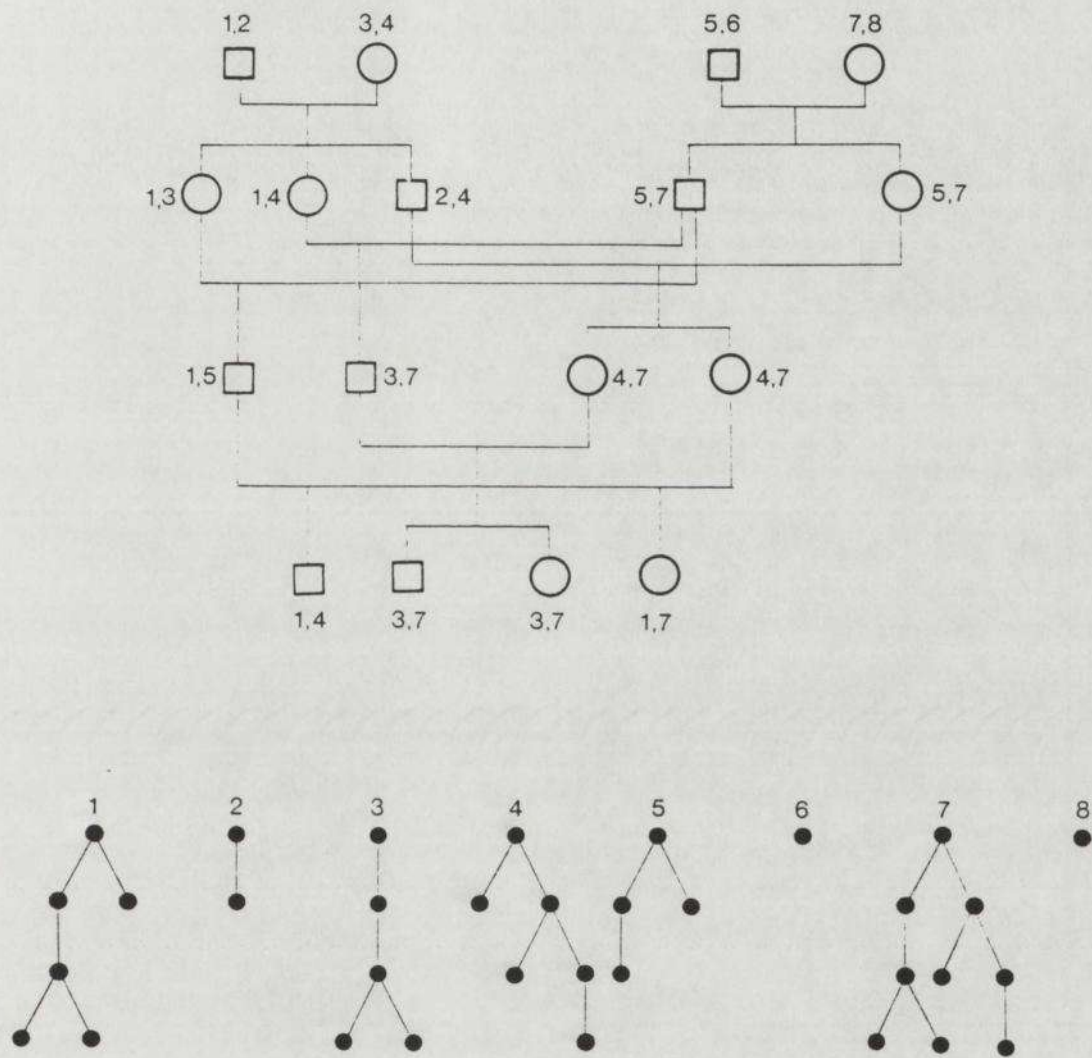


Fig. 1 : Au-dessus, histoire d'une famille, chaque individu pourrait avoir $2n$ ancêtres
 En-dessous, histoire des gènes (représentés par des nombres), un gène ne peut avoir qu'un "ancêtre" gène.