

# L'erreur du procureur-1

**Niveau :** première générale, terminale générale (spécialité. Maths complémentaires) ou technologique.

**Lien avec le programme :** probabilité, conditionnement, indépendance, loi binomiale.

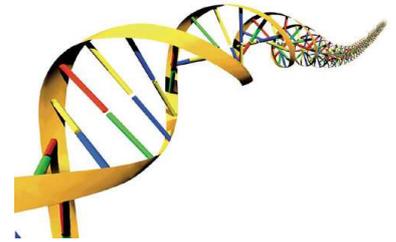
**Lien avec Les maths au quotidien :** Nature, Société.

Bien que deux êtres humains aient une large majorité de leur patrimoine génétique identique, un certain ensemble de séquences dans leur ADN reste spécifique à chaque individu : c'est le *polymorphisme*.

Ces séquences polymorphes d'ADN, aisément détectables grâce à un emplacement connu sur un chromosome, s'appellent des **marqueurs génétiques** (ou locus).

On appelle empreinte génétique (ou profil génétique), le résultat de l'analyse de plusieurs marqueurs génétiques. Selon les situations le nombre de marqueurs utilisés pour définir une empreinte est variable.

Grâce aux empreintes génétiques, il est possible d'identifier un individu et aussi de comparer des empreintes génétiques par exemple dans des cas de crimes.



Prenons l'exemple d'un profil ADN « partiel » obtenu avec l'analyse de seulement quatre marqueurs :

CSF1PO, TPOX, TH01 et VWA

Un crime a eu lieu dans une grande ville.

Une trace relevée sur la victime possède les allèles 10 et 11 sur le marqueur CSF1PO, 8 et 8 sur le marqueur TPOX, 6 et 7 sur le marqueur TH01, 5 et 8 sur le marqueur VWA (allèle donné par chaque parent, voir tableau en annexe et doc loi de Hardy Weinberg pour la transmission des allèles).

Les profils génétiques de plusieurs personnes susceptibles de se trouver sur la scène de crime sont établis...

Parmi eux, un homme possède ces mêmes allèles pour ces mêmes marqueurs et est aussitôt inculpé.

Une étude de population montre que, pour un individu tiré au hasard, la probabilité de :

- l'évènement *A* : « obtenir les allèles 10 et 11 pour le marqueur CSF1PO » est de 0,16 ;
- l'évènement *B* : « obtenir les allèles 8 et 8 pour le marqueur TPOX » est de 0,28 ;
- l'évènement *C* : « obtenir les allèles 6 et 7 pour le marqueur TH01 » est de 0,07 ;
- l'évènement *D* : « obtenir les allèles 5 et 8 pour le marqueur VWA » est de 0,05.

Les évènements *A*, *B*, *C* et *D* sont considérés comme indépendants.

## A. Le raisonnement du procureur.

Lors du procès, le procureur de l'accusation affirme :

« La fréquence d'apparition du profil dans la population est de 0,016 % donc la probabilité que quelqu'un d'autre ai laissé cette trace est de 0,016 %. Au vu de son profil génétique, on est donc sûr à 99,984 % que le suspect a laissé cette trace ».

Voyons cela d'un point de vue probabiliste :

On analyse le profil génétique d'un individu potentiellement présent sur la scène de crime.

On note *C* : « l'individu est l'auteur de la trace » (= « l'individu est coupable »)

*M* : « l'individu présente le même profil génétique que celui trouvé dans la trace »



1. Etablir que  $P_M(C) = \frac{P(C)}{P(M)}$ .

2. a. Comment, dans son réquisitoire, le procureur arrive-t-il (avec raison) au chiffre de 0,016 % ?
- b. De quelle probabilité parle selon lui le procureur quand il donne la valeur 99,984 % ? De quel évènement cette valeur est-elle en réalité la probabilité ?
- b. Conclure quant au raisonnement du procureur.

## B. Quelqu'un d'autre pouvait-il vraisemblablement laisser cette trace ?

On suppose que 1 500 personnes aient pu accéder à la scène du crime, c'est par exemple le nombre d'habitants du quartier.

On suppose que chaque personne de la population totale a la même probabilité a priori d'avoir le même profil génétique que la victime, et que la population totale est suffisamment importante pour que le choix de ces 1 500 personnes soit assimilé à 1 500 tirages avec remise.

Soit  $X$  la variable aléatoire égale au nombre de personnes n'ayant pas le même profil génétique que la victime parmi ces 1 500 personnes.

1. Calculer la probabilité qu'au moins une personne parmi les 1 500 ait le même profil génétique que la victime.
2. Quel argument pourrait être alors développé par l'avocat de la défense ?
3. Le tribunal fédéral de Karlsruhe (Allemagne) en 1992, a refusé de condamner un homme accusé de viol sur la seule « preuve » d'une analyse génétique établissant qu'il y avait 99,986 % de chance que le sperme trouvé sur la victime fut le sien. Rapporté au nombre d'habitants de Hanovre, ville où le viol avait été perpétré, ce pourcentage signifiait que 35 hommes pouvaient avoir les mêmes empreintes (*authentique*).  
À combien d'hommes le tribunal a-t-il estimé le nombre dans la population d'Hanovre ?

**Point info :** Le FNAEG (fichier national automatisé des empreintes génétiques) contenait en 2016 plus de trois millions d'empreintes génétiques portant sur une quinzaine de marqueurs... la probabilité d'avoir deux séquences identiques pour deux individus différents est dans ce cas inférieur à  $1/10^{15}$ ... ce qui est quasiment négligeable.

**Annexe :** (extrait traduit de la page <http://dna-view.com/profile.htm>)

### Probabilité de profil d'ADN

Locus	Allèles	effectif	Taille de la base de données	La fréquence	formule	nombre
CSF1PO	dix	109	432	$P = 0,25$	$2 pq$	0,16
	11	134		$Q = 0,31$		
TPOX	8	229	432	$P = 0,53$	$P^2$	0,28
	8					
THO1	6	102	428	$P = 0,24$	$2 pq$	0,07
	7	64		$Q = 0,15$		
VWA	16	91	428	$P = 0,21$	$P^2$	0,05
	16					

**Fréquence de profil = 0.00014**

L'allèle 10 au locus CSF1PO a été observé 109 fois dans un échantillon de population de 432 allèles (216 personnes). Par conséquent, il est raisonnable d'estimer qu'il existe une chance  $p = 0,25$  que tout allèle CSF1PO particulier, sélectionné au hasard, serait un 10. De même, la probabilité est d'environ  $q = 0,31$  pour un allèle aléatoire CSF1PO à 11. Avant d'inculper le suspect, si nous supposons qu'il n'est pas l'auteur de la trace, alors on peut penser qu'il a reçu un allèle CSF1PO au hasard de chacun de ses parents. La probabilité de recevoir 10 de sa mère et 11 de son père est donc  $pq$ , et recevoir 11 de la mère et 10 du père est aussi  $pq$ , donc la probabilité d'obtenir les allèles 10 et 11 par hasard est de  $2 pq$ . Par conséquent, environ 16 % des personnes ont le génotype 10,11 au locus CSF1PO.

Au locus TPOX, puisque les deux allèles sont les mêmes, il n'y a qu'un seul terme :  $pp$  ou  $p^2$ , ce qui représente la probabilité combinée d'hériter de l'allèle 8 de chaque parent. Par conséquent, environ 28 % des personnes ont le même génotype TPOX que la trace. On s'attend à ce que la proportion de personnes TPOX 8,8 soit encore de 28 %, même si l'attention est limitée uniquement aux personnes ayant un génotype CSF1PO particulier tel que 10,11. Par conséquent, la probabilité pour une personne d'avoir le génotype combiné dans les deux loci est de 28 % de 16 % - environ 4%.

Les calculs pour les lois THO1 et VWA sont similaires, et en prenant en compte la possibilité globale pour une personne aléatoire d'avoir le génotype combiné de 4 % à environ 1/7000.