

# Evaluation de la preuve génétique par l'utilisation des probabilités

## 1 La fréquence d'un profil génétique

### 1.1 Utilisations erronées

Partons d'un exemple concret :

On a découvert sur la scène du crime des traces de sang humain. Un test ADN a montré que ces traces de sang s'accordaient avec le sang du suspect. L'expert a estimé la probabilité d'une coïncidence génétique due au hasard c'est à dire la probabilité que deux personnes non apparentées aient le même profil génétique à 1/10 millions.

L'accusation pourrait alors argumenter :

- La fréquence d'apparition du profil est de 0,01%.
- La probabilité que quelqu'un d'autre ait laissé cette trace est de 0,01%.
- On est sûr à 99,99% que le suspect a laissé cette trace.

La Défense lui répondrait que :

- Le profil génétique de la trace est présent chez environ 1 individu sur 10 millions personnes.
- Si on considère comme population d'intérêt 20 millions personnes qui peuvent avoir laissé cette trace, il y a un deuxième individu présentant ce profil génétique.
- Dès lors la probabilité de se tromper est de 50%

Le raisonnement de l'accusation est faux car l'expert a estimé la probabilité que les deux empreintes génétiques concordent sachant que le suspect n'est pas l'auteur de la trace. Alors que l'accusation la considère égale à la probabilité que le suspect n'est pas l'auteur de la trace sachant que les empreintes coïncident.

Le raisonnement de la défense oublie que les autres indices ont sans doute limité la population d'intérêt à une taille inférieure et que si une autre personne a le même profil, on ne peut pas forcément supposer que leur implication dans le crime soit équiprobable.

### 1.2 Probabilité d'une autre correspondance

A partir des données fournies par l'expert on peut calculer la probabilité que quelqu'un d'autre ait le même profil génétique. On reste sous les hypothèses de l'exemple

précédent c'est à dire que la fréquence du profil génétique de la trace est de 1/10 millions. Supposons que la population d'intérêt (celle qui a eu accès à la scène du crime) comporte  $n$  personnes, la probabilité  $p$  que quelqu'un d'autre que le suspect ait pu laisser cette trace est égale à :

$$p = 1 - \left(1 - \frac{1}{10000000}\right)^n.$$

Si par exemple on considère que la population d'intérêt correspond aux habitants de l'agglomération dont on estime la taille à 1 millions de personnes, il y a 10% de chances pour que dans l'agglomération, une autre personne ait la même empreinte génétique.

## 2 L'approche bayésienne

Nous allons présenter une autre approche probabiliste de l'information fournie par l'expert. L'expert ne peut estimer que les probabilités suivantes :  $\mathbb{P}(E|A)$  et  $\mathbb{P}(E|D)$  où

- $E$  est l'événement : "l'empreinte du suspect et de la trace coïncident" ;
- $A$  est l'événement : "le suspect est l'auteur de la trace" c'est l'hypothèse de l'accusation ;
- $D$  est l'événement : "le suspect n'est pas l'auteur de la trace" c'est l'hypothèse de la défense ;
- $\mathbb{P}(B|C)$  est la probabilité de  $B$  sachant que l'événement  $C$  est réalisé.

On définit le rapport de vraisemblance  $R$  par :

$$R = \frac{\mathbb{P}(E|A)}{\mathbb{P}(E|D)}.$$

Celui-ci est estimable par l'expert.

La question que se pose le jury est quel est le rapport de probabilité entre l'hypothèse de l'accusation et l'hypothèse de la défense. Plus ce rapport est grand (très supérieur à 1), plus l'hypothèse de l'accusation est probable par rapport à celle de la défense, plus ce rapport est faible (inférieur à 1) plus l'hypothèse de la défense est plus probable.

Nous allons maintenant isoler le rôle joué par les empreintes génétiques. Notons  $I$  l'ensemble des indices en possession du tribunal à part le fait que les empreintes coïncident. On obtient alors une décomposition du rapport des probabilités des hypothèses de l'accusation et la défense en deux parties :

$$\begin{aligned} \frac{\mathbb{P}(A|E, I)}{\mathbb{P}(D|E, I)} &= \frac{\mathbb{P}(A|I)}{\mathbb{P}(D|I)} \times \frac{\mathbb{P}(E|A, I)}{\mathbb{P}(E|D, I)} \\ &= \frac{\mathbb{P}(A|I)}{\mathbb{P}(D|I)} \times R. \end{aligned}$$

Le premier terme du produit correspond à l'information apportée par les autres indices, le deuxième est ce qu'apporte la coïncidence des empreintes de plus ce rapport de vraisemblance est estimable par l'expert. Si  $R > 1$  le fait que les empreintes coïncident renforce l'hypothèse de l'accusation et si  $R < 1$  cela renforce l'hypothèse de la défense. Cette approche permet de mieux comprendre quel est le rôle joué par l'analyse génétique des empreintes dans la preuve d'une culpabilité.

## Références

- [1] [http://www.assemblee-nationale.fr/rap-ocst/empreintes\\_genetiques/r3121-1.asp](http://www.assemblee-nationale.fr/rap-ocst/empreintes_genetiques/r3121-1.asp)