

# L'évolution de la biodiversité génétique : le principe de Hardy-Weinberg

Karim ZAYANA<sup>1,2</sup> et Victor RABIET<sup>1,3</sup>

<sup>1</sup> Ministère de l'Éducation nationale, Paris

<sup>2</sup> LTCl, Télécom Paris, Institut Polytechnique de Paris

<sup>3</sup> DMA, École normale supérieure, Paris



Ce texte revient sur le principe d'équilibre de Hardy<sup>1</sup>-Weinberg<sup>2</sup> [1] tel qu'il est énoncé dans le programme d'enseignement scientifique de terminale [9] : en l'absence de forces évolutives et pour peu que le hasard joue correctement

1. Godfrey Harold HARDY ; mathématicien britannique, 1877 – 1947. Il s'est illustré en théorie des nombres et en analyse, entretenant une collaboration fructueuse avec ses collègues John Edensor LITTLEWOOD et Edward Maitland WRIGHT ainsi qu'avec le mathématicien indien Srinivasa RAMANUJAN. C'est de manière tout à fait anecdotique, au cours d'un repas au Trinity College de Cambridge avec son ami et partenaire de cricket le généticien Reginald Crundal PUNNETT, que le problème de la structure génotypique d'une population lui fut soumis. Il résolut en quelques minutes la question de sa stabilisation théorique et, se pensant insuffisamment légitime dans un domaine aux frontières du sien, hésita même à publier ce résultat ! [10, 3]

2. Wilhelm WEINBERG ; médecin allemand, 1862 – 1937. C'est parallèlement aux travaux de Hardy qu'il postula lui aussi le principe d'équilibre auquel il donne également son nom.

son rôle, la structure génotypique d'une grande population de même espèce se stabilise dès la première descendance, c'est-à-dire à partir de la génération  $F_1$ <sup>3</sup>.

Ce principe tout théorique est rarement respecté dans la réalité, précisément parce que son cadre d'application est trop étroit. Mais avant, nous en réaliserons la démonstration formelle en l'accordant avec le langage des probabilités vu au lycée [6]. Nous raisonnerons avec un nombre  $p$  d'allèles quelconque plutôt que de nous limiter au seul cas où  $p = 2$ . Plus réaliste, ce choix nous permettra, en outre, de suivre plus aisément le fil de l'exposé. Interviennent en particulier la notion de probabilité conditionnelle<sup>4</sup>, le concept d'indépendance et la formule des probabilités totales. Nous verrons qu'une fois les conventions d'écriture, certes un peu fastidieuses, bien posées, la preuve coule de source. Les hypothèses mathématiques nécessaires à sa mise en œuvre s'imposeront d'elles-mêmes. Elles fixeront de facto les contours du modèle. Assez nombreuses, nous les porterons *en italiques* à mesure que nous les rencontrerons.

## 1 Prérequis

Un gène est une unité d'information constituée par une séquence d'ADN [7, 8]. Il est responsable d'un caractère héréditaire se traduisant par une spécificité anatomique ou physiologique chez une espèce donnée<sup>5</sup>. Un même gène admet en général plusieurs variants, mettons  $p$ , comme autant de déclinaisons possibles d'une matrice commune  $G$ . Appelons-les  $G_1, G_2, \dots, G_p$  et signalons dès maintenant qu'on pourrait aussi les noter en lettres minuscules, la raison en sera clarifiée tout à l'heure.

Par exemple, le gène  $G$  déterminant le groupe sanguin chez l'Homme est porté par le chromosome 9 et peut apparaître sous trois formes :  $G_1 = A$ ,  $G_2 = B$  ou  $G_3 = O$  (ce dernier figurant provisoirement en majuscule).

- Le variant  $G_1 = A$  induit la synthèse d'une certaine molécule  $A$  sur la membrane des hématies (ou globules rouges), les cellules qui renferment l'hémoglobine ;

---

3.  $F_1$  pour "first Filial generation", première descendance de la population d'origine,  $F_0$ . Sur le même modèle,  $F_2$  est la "second Filial generation",  $F_3$  la "third Filial generation", etc.

4. Dans tout le texte, la probabilité d'un événement  $A$  conditionnellement à un événement  $B$  sera notée  $\mathbb{P}(A | B)$  comme c'est l'usage dans la littérature, et non  $\mathbb{P}_B(A)$ , comme c'est l'habitude dans les manuels scolaires, afin de ne pas surcharger l'espace alloué aux indices qui sera déjà occupé par divers compteurs.

5. L'être humain possède environ 20 000 gènes.

- Le variant  $G_2 = B$  induit la synthèse d'une certaine molécule B sur la membrane des hématies ;
- Le variant  $G_3 = O$  n'induit rien de son propre fait, mais n'inhibe rien non plus.

Dans son patrimoine génétique, un être humain détient deux exemplaires<sup>6</sup> de chaque chromosome, dits homologues car portant les mêmes gènes. En particulier, il possède deux exemplaires du chromosome 9 : l'un provient de son père, l'autre de sa mère. De son père il a reçu une première version, ou premier allèle, du gène  $G$  : potentiellement  $A$ ,  $B$  ou  $O$ . De sa mère, il a reçu une seconde version, ou second allèle, du gène  $G$  : potentiellement encore  $A$ ,  $B$  ou  $O$ . Soit une combinatoire de  $\binom{3}{2} = 6$  paires d'allèles, ou génotypes, possibles :

$$\{A, A\}, \{B, B\}, \{O, O\}, \{A, B\}, \{A, O\}, \{B, O\}$$

Nous avons noté les paires entre accolades comme c'est de circonstance en mathématiques. Toutefois, en biologie, elles apparaîtraient différemment, par exemple  $A \parallel B$ , voire  $AB$  tout court en lieu de  $\{A, B\}$ . Nous opterons pour  $A \parallel B$ , car les mots «parallélisme» et «allèle» ont une racine grecque commune et car le parallélisme est une relation symétrique tandis que l'autre écriture,  $AB$ , laisse penser que l'ordre d'exposition compte.

Les individus  $A \parallel A$ ,  $B \parallel B$  et  $O \parallel O$  sont de génotype «pur» : on les dit homozygotes pour le gène  $G$ , à la différence des individus  $A \parallel B$ ,  $A \parallel O$  et  $B \parallel O$  aux profils hybrides, dits hétérozygotes. Un individu homozygote, par exemple de type  $A \parallel A$ , ne produira que des cellules reproductrices, les gamètes (spermatozoïdes ou ovules), possédant l'allèle  $A$ . Un individu hétérozygote, par exemple de type  $A \parallel B$ , produira à *parité* des gamètes portant l'allèle  $A$  ou  $B$  : *on stipule à cet endroit une hypothèse d'équiprobabilité*, dite loi de ségrégation (ou de disjonction) de Mendel, loi que nous évoquerons encore un peu plus loin.

Un individu du type  $A \parallel A$  (respectivement  $B \parallel B$ ) fabrique des molécules  $A$  (respectivement  $B$ ) à la surface de ses globules rouges en réponse à l'ordre qu'il reçoit (deux fois) de ses gènes. Un individu du type  $O \parallel O$  ne fabrique pas de molécule particulièrement remarquable. Un individu du type hybride  $A \parallel O$  (resp.  $B \parallel O$ ) fabriquera des molécules  $A$  (resp. des molécules  $B$ ) ; un individu

---

6. On dit que les cellules de l'être humain (exception faite de ses cellules sexuelles) sont diploïdes. Beaucoup d'autres êtres vivants fonctionnent ainsi. Mais pas tous, comme les abeilles mâles (les faux bourdons) et la majorité des bactéries par exemple.

hybride  $A \parallel B$  fabriquera des deux molécules A et B. De ce fait les variants A et B, qui s'expriment toujours quand ils apparaissent, sont dits codominants et le variant O est dit récessif. L'usage veut qu'on puisse alors écrire ce dernier avec une minuscule :  $o$ , et non O. Nous venons incidemment de dégager 4 traits de caractère visibles extérieurement pour les 6 génotypes dénombrés : le trait **A** issu de  $A \parallel A$  ou de  $A \parallel o$ , le trait **B** issu de  $B \parallel B$  ou de  $B \parallel o$ , le trait **o** issu de  $o \parallel o$  uniquement et le trait **AB** issu de  $A \parallel B$  uniquement. Ces traits sont qualifiés de phénotypes. En biologie, on les note plus volontiers entre crochets, soit [A], [B], [o] et [AB], et on les qualifie présentement de groupes. Ils sont d'une grande importance du fait des réactions immunitaires qu'ils peuvent induire chez le receveur d'une transfusion sanguine et soulèvent alors des questions de compatibilité.

La nature dominante ou non d'un allèle ne remet pas en question la règle d'équiprobabilité posée plus haut. Cela tient au mécanisme de la méiose, le processus de division cellulaire qui donne naissance aux gamètes et sur lequel la loi de Mendel se fonde. Concrètement, un individu  $A \parallel o$  a autant de chances de produire un gamète avec l'allèle dominant A qu'un gamète avec l'allèle récessif  $o$ . Évidemment, un individu  $A \parallel A$  ne peut produire qu'un gamète avec l'allèle A, un individu  $o \parallel o$  qu'un gamète avec l'allèle  $o$ . Au regard des probabilités, écrire en majuscule un certain variant  $G_i$  ou l'écrire en minuscule,  $g_i$ , n'a donc aucune incidence. Aussi, nous l'indiquerons toujours en majuscule dans ce qui suit. Et nous *considérerons que l'allèle que produit un individu qui en aurait deux distincts est une issue de Bernoulli<sup>7</sup> de paramètre  $\frac{1}{2}$ .*

Le modèle construit par Hardy et Weinberg pour expliquer la stabilité de la structure génétique d'une population au cours des générations reposera sur des hypothèses supplémentaires. La première d'entre elle est la *panmixie*. Elle consiste à considérer que les *individus s'accouplent sans tenir aucun compte de leur génotype ni même de leur phénotype (qui en est la partie visible)*. Cela va de soi pour le groupe sanguin, mais interroge davantage sur des caractères plus voyants. Ainsi, nous dirons par exemple qu'un individu aux yeux bleus n'ira pas préférentiellement vers un individus aux yeux également bleus, dans la mesure où il aurait le choix entre des individus au yeux bleus, marrons, verts, etc. : « aveugle », il est juste insensible à ce caractère. Ajoutons que s'il fréquente un individu aux yeux marrons, les gamètes que chaque partenaire a produits se rencontrent, donc

---

7. Jakob BERNOULLI, mathématicien suisse et premier d'une longue lignée de savants, 1654 – 1705.

se transmettent, à leur tour au hasard : c'est l'*hypothèse de pangamie*. Panmixie et pangamie maximisent les croisements possibles et permettent idéalement, par le métissage, à la diversité de se renouveler ou de se maintenir.

Bref, de la panmixie, de la loi de discrimination de Mendel, et de la pangamie, retenons que tout se passe comme si les deux futurs parents étaient juste tirés au sort dans la population mâle et femelle sans qu'aucun des deux tirages n'influence l'autre (ils ne «cherchent» pas à se ressembler ni à être complémentaires par exemple), et comme si les allèles qu'ils transmettent ensuite étaient eux aussi tirés au sort (parmi deux possibles pour chaque partenaire) sans qu'aucun des deux tirages n'influe non plus sur l'autre. **Du point de vue des probabilités, nous poserons qu'au moment d'une fécondation, les événements « le père transmet l'allèle  $G_i$  » et « la mère transmet l'allèle  $G_j$  » sont indépendants**<sup>8</sup>. Du point de vue de la biologie, c'est cette «grande loterie» du vivant, répétée sur tous les gènes de tous les chromosomes, qui assure que chaque individu est un être inédit.

## 2 Le principe de Hardy-Weinberg

Soit une population « zéro »,  $F_0$ , composée d'individus d'une même espèce. Une expérience aléatoire peut consister à les tirer au sort. Aussi, nous introduisons les probabilités suivantes, que nous *supposerons indifférentes au sexe, mâle ou femelle*. Pour tous  $i, j$  dans  $\llbracket 1, p \rrbracket$ , :

- $P_{\{i,j\},0} = \mathbb{P}(\text{«un individu a le génotype } G_i \parallel G_j\text{»})$ . Lorsque  $i \neq j$ , notons qu'un individu  $G_i \parallel G_j$  peut être d'allèle  $G_i$  par son père et  $G_j$  par sa mère, ou vice versa et que, vu des probabilités, père et mère sont interchangeable.
- $P_{i,0} = \mathbb{P}(\text{«un individu transmet dans son gamète l'allèle } G_i\text{»})$

À l'origine, la population étudiée pourrait réunir différentes souches très homogènes, essentiellement homozygotes pour le gène  $G$ . Il est donc possible qu'au sein du grand groupe ainsi reconstitué, les densités soient inégalement distribuées, avec des probabilités homozygotes  $P_{\{i,i\},0}$  significatives et des probabilités hétérozygotes  $P_{\{i,j\},0}$ ,  $i \neq j$ , insignifiantes.

---

8. Attention toutefois. Pour  $i \neq j$ , les événements « le père transmet l'allèle  $G_i$  » et « la mère transmet l'allèle  $G_j$  » ne sont en revanche pas indépendants. Certains gènes «voyagent» ensemble, portés qu'ils sont par un même chromosome. Et tout cela peut être complexifié par le "crossover" (enjambement), quand des morceaux de chromosomes homologues s'échangent.

Nous admettons que la population se régénère à un rythme régulier *par vagues successives*  $F_1, F_2, \dots, F_n, \dots$  *étanches les unes aux autres* : une génération donnée se reproduit au même moment quand les générations passées ne sont, elles, plus fécondes. On dit que les *générations successives sont discrètes et non chevauchantes*, hypothèse bien sûr discutable. Dès lors, nous définissons à l'avenant pour la génération  $F_n$  les probabilités  $P_{\{i,j\},n}$  et  $P_{i,n}$  quand  $i, j \in \llbracket 1, p \rrbracket$ .

### Principe de Hardy-Weinberg

Dans un cadre idéal, les probabilités  $P_i$  sont stables dès le départ, tandis que les probabilités  $P_{\{i,j\}}$  se stabilisent à partir de la première génération filiale,  $F_1$ . Ainsi la répartition génotypique de la population se fige-t-elle à l'image du premier brassage qui se sera produit. En particulier, un variant dominant du gène étudié ne prend pas le dessus, en termes statistiques, sur les autres. Et un variant qui serait récessif ne disparaîtrait pas non plus.

On démontre ce principe en passant d'un degré  $n$  de filiation au degré  $n + 1$  pour les  $P_i$ , puis au degré  $n + 2$  pour les  $P_{\{i,j\}}$ ,  $n \geq 0$ . Trois étapes nous seront nécessaires.

**Étape 1 : calcul de  $P_{\{i,j\},n+1}$ .** Tout d'abord, compte-tenu des définitions ci-avant, de la similarité des rôles du père et de la mère, des propriétés d'indépendance déjà soulevées, quand  $i, j \in \llbracket 1, p \rrbracket$  avec  $i \neq j$ ,

$$\begin{aligned} P_{\{i,j\},n+1} &= \mathbb{P}(\text{«un individu de } F_{n+1} \text{ est du génotype } G_i \parallel G_j \text{»}) \\ &= \mathbb{P}(\text{«le père transmet l'allèle } G_i \text{»} \cap \text{«la mère transmet l'allèle } G_j \text{»}) \\ &\quad + \mathbb{P}(\text{«le père transmet l'allèle } G_j \text{»} \cap \text{«la mère transmet l'allèle } G_i \text{»}) \\ &= 2\mathbb{P}(\text{«le père transmet l'allèle } G_i \text{»} \cap \text{«la mère transmet l'allèle } G_j \text{»}) \\ &= 2\mathbb{P}(\text{«le père transmet l'allèle } G_i \text{»}) \times \mathbb{P}(\text{«la mère transmet l'allèle } G_j \text{»}) \\ &= 2P_{i,n}P_{j,n}. \end{aligned}$$

De même, si  $i \in \llbracket 1, p \rrbracket$ ,

$$P_{\{i,i\},n+1} = P_{i,n}^2.$$

**Étape 2 : calcul de  $P_{i,n+1}$  et première stabilité annoncée.** Prenons  $i \in \llbracket 1, p \rrbracket$ . D'après la formule des probabilités totales,

$$\begin{aligned} P_{i,n+1} &= \mathbb{P}(\text{«un individu de } F_{n+1} \text{ transmet l'allèle } G_i\text{»}) \\ &= \sum_{k=1}^p \mathbb{P}(\text{«l'allèle transmis est } G_i\text{»} \mid \text{«l'individu est de génotype } G_i \parallel G_k\text{»}) \\ &\quad \times \mathbb{P}(\text{«l'individu est de génotype } G_i \parallel G_k\text{»}) \end{aligned}$$

Quand  $k \neq i$ , la première probabilité vaut  $\frac{1}{2}$ , et la seconde  $P_{\{i,k\},n+1} = 2P_{i,n}P_{k,n}$  d'après l'étape 1. Quand  $k = i$ , la première probabilité vaut 1 et la seconde  $P_{\{k,k\},n+1} = P_{k,n}^2$  toujours d'après l'étape 1. La somme se simplifie donc en

$$P_{i,n+1} = \sum_{k=1}^p P_{i,n}P_{k,n} = P_{i,n} \sum_{k=1}^p P_{k,n} = P_{i,n}. \quad (1)$$

$\rightsquigarrow$  Ceci prouve la première stabilité annoncée.

**Étape 3 : calcul de  $P_{\{i,j\},n+2}$  et deuxième stabilité annoncée.** Comme à l'étape 1, pour tous  $i, j \in \llbracket 1, p \rrbracket$ ,  $i \neq j$ ,

$$P_{\{i,j\},n+2} = 2P_{i,n+1}P_{j,n+1}.$$

Mais alors, d'après l'étape 2, suivie de l'étape 1,

$$P_{\{i,j\},n+2} = 2P_{i,n}P_{j,n} = P_{\{i,j\},n+1}. \quad (2)$$

Le raisonnement est analogue quand  $i = j$ .

$\rightsquigarrow$  Ceci prouve la deuxième stabilité annoncée.

Une autre façon de mener la preuve consistait à suivre un variant quelconque  $V$ , symbolisant un certain  $G_i$  où  $i \in \llbracket 1, p \rrbracket$ , à désigner sous l'appellation générique  $\bar{V}$  tous les autres dont on aurait fait masse, puis à ne manipuler que des probabilités du type  $P_{V \parallel V}$ ,  $P_{\bar{V} \parallel \bar{V}}$ ,  $P_{V \parallel \bar{V}}$ ,  $P_V$  et  $P_{\bar{V}}$ . Ceci nous affranchit des sommations et de leur indexation, mais conduit à des informations un peu moins fines en termes de distributions.

Notre démonstration contient par ailleurs plusieurs implicites :

- Que de sa naissance à son âge de procréer, une génération *n'ait pas été perturbée par des facteurs extérieurs*. Migrations, décès prématurés ou mutations peuvent en effet en modifier la structure. Tel génotype supportera moins tel climat, tandis que des mutations (à faible probabilité sur une courte échelle de temps) peuvent altérer ou modifier les gènes et créer un nouveau réservoir de variants, etc. ;
- Que *tous les individus d'une génération soient également viables et fertiles*. Savoir que tel génotype ne serait pas fécond changerait certainement la donne.

Enfin, si les probabilités «zéro» correspondent bien à des fréquences empiriques, c'est-à-dire à des proportions calculées sur une population effectivement existante, les probabilités concernant les générations suivantes ne sont que prospectives. C'est par la grâce de la loi des grands nombres, dans *l'hypothèse d'une très grande population* – échantillon statistique dont la taille  $N$  tend vers l'infini, qu'elles coïncideraient, asymptotiquement, avec les fréquences génotypiques et alléliques observées lorsque se réalise l'expérience.

## Conclusion

En pratique, nous étudions un échantillon de population. Nous y mesurons des effectifs que nous rangeons dans une table, par exemple selon que le variant  $V$  est présent ou non comme dans le tableau 1.

Génotypes	$V \parallel V$	$V \parallel \bar{V}$	$\bar{V} \parallel \bar{V}$	Total
Effectifs	$n_{V \parallel V}$	$n_{V \parallel \bar{V}}$	$n_{\bar{V} \parallel \bar{V}}$	$N$

TABLE 1 – Tableau d'effectifs d'un échantillon de  $N$  individus d'une même espèce classés selon leur génotype. L'attention est ici portée sur un variant  $V$  d'un certain gène  $G$ .

Les fréquences  $\hat{p}$  et  $\hat{q}$  des allèles  $V$  et  $\bar{V}$  dans cet échantillon valent par conséquent :

$$\hat{p} = \frac{2n_{V \parallel V} + n_{V \parallel \bar{V}}}{2N} \text{ et } \hat{q} = \frac{2n_{\bar{V} \parallel \bar{V}} + n_{V \parallel \bar{V}}}{2N}.$$

avec  $N = n_{V \parallel V} + n_{V \parallel \bar{V}} + n_{\bar{V} \parallel \bar{V}}$ .

En considérant que  $\hat{p}$  et  $\hat{q} = \hat{p}$  sont des estimateurs des probabilités  $p = P_V$  et  $q = P_{\bar{V}}$  qu'un individu choisi au hasard transmette l'allèle  $V$  ou l'allèle  $\bar{V}$



à l'équilibre de Hardy-Weinberg, nous attendons les probabilités génotypiques  $P_{V||V} = \hat{p}^2$ ,  $P_{V||\bar{V}} = 2\hat{p}\hat{q}$  et  $P_{\bar{V}||\bar{V}} = \hat{q}^2$ , et donc les effectifs génotypiques renseignés dans le tableau 2 ci-après.

Génotypes	$V  V$	$V  \bar{V}$	$\bar{V}  \bar{V}$	Total
Effectifs attendus	$\hat{p}^2 N$	$2\hat{p}\hat{q} N$	$\hat{q}^2 N$	$N$

TABLE 2 – Tableau des effectifs génotypiques attendus à l'équilibre de Hardy-Weinberg.

Sur des données expérimentales, on peut comparer visuellement les deux tableaux. Mais on peut aussi, plus rigoureusement, utiliser un indicateur de conformité. Typiquement

$$\begin{aligned} \chi_c^2 &= \sum_{\text{génotypes}} \frac{(\text{effectif attendu} - \text{effectif réel})^2}{\text{effectif attendu}} \\ &= \frac{(\hat{p}^2 N - n_{V||V})^2}{\hat{p}^2 N} + \frac{(2\hat{p}\hat{q} N - n_{V||\bar{V}})^2}{2\hat{p}\hat{q} N} + \frac{(\hat{q}^2 N - n_{\bar{V}||\bar{V}})^2}{\hat{q}^2 N}, \end{aligned}$$

censé suivre une loi du  $\chi^2$  à  $3 - 1 = 2$  degrés de liberté, ce qui permet d'élaborer un test dont l'hypothèse nulle  $H_0$  [2] affirmerait par exemple que la population étudiée répond à la loi de Hardy-Weinberg.

Concrètement, on observera plutôt des écarts au modèle de Hardy-Weinberg : à de grandes échelles de temps, les populations, de tailles aussi grandes soient-elles, n'ont pas de structure allélique stable. Ceci démontre que des forces évolutives sont à l'œuvre [4, 5] : sélection, migrations, mutations, dérive génétique. Il reste ensuite à les déterminer, mais aussi à reconsidérer les hypothèses formulées quant à la panmixie et la pangamie par exemple.

On pourra faire l'analogie avec la physique et son principe d'inertie : quand aucune force résultante ne s'applique à un mobile, ce dernier suit un mouvement rectiligne uniforme. Tout écart à ce modèle prouve, a contrario, la présence d'interventions extérieures.

Karim ZAYANA est inspecteur général de l'éducation, du sport et de la recherche (groupe des mathématiques) et professeur invité à l'institut polytechnique de Paris (Palaiseau) et Victor RABIET est responsable éditorial de CultureMath (école normale supérieure, Paris).

## Remerciements

Les auteurs remercient tout particulièrement Jean-François BEAUX, professeur de SVT en classe de spéciale BCPST au lycée Henri IV, Bruno ANSELME, professeur de SVT en classe de spéciale BCPST au lycée Fénelon, Pascal COMBEMOREL, responsable éditorial du site ens-dgesco Planet-Vie, Vincent PANTALONI, inspecteur d'académie – inspecteur pédagogique régional de mathématiques dans l'académie d'Orléans-Tours, Astrid ZAYANA, étudiante ingénieur à l'Agro Nancy (ENSAIA) et Gaëlle GUILLOTIN, inspectrice d'académie – inspectrice pédagogique régionale dans l'académie de Versailles pour leur relecture minutieuse (et bienveillante!).

## Références

- [1] Bruno ANSELME. *Biomathématiques*. Dunod, 2015.
- [2] Françoise COUTY-FREDON, Jean DEBORD et Daniel FREDON. *Probabilités et statistiques*. Dunod, 2018.
- [3] James F. CROW. « Hardy, Weinberg and language impediments ». In : *Genetics* (1999).
- [4] Patrice DAVID et Sarah SAMADI. *La théorie de l'évolution*. Champs sciences. Flammarion, 2021.
- [5] Yuval Noah HARARI. *Sapiens*. Albin Michel, 2015.
- [6] « Programme d'enseignement de spécialité de mathématiques de la classe de Première de la voie générale ». In : *Bulletin officiel spécial n°1 du 22 janvier 2019* (2019). URL : <https://www.education.gouv.fr/bo/19/Special1/MENE1901632A.htm>.
- [7] « Programme de l'enseignement de spécialité des Sciences de la Vie et de la Terre de Première des voies générale et technologique ». In : *Bulletin officiel spécial n°1 du 22 janvier 2019* (2019). URL : [https://cache.media.eduscol.education.fr/file/SP1-MEN-22-1-2019/54/2/spe648\\_annexe\\_1063542.pdf](https://cache.media.eduscol.education.fr/file/SP1-MEN-22-1-2019/54/2/spe648_annexe_1063542.pdf).

- [8] « Programme de l'enseignement de spécialité des Sciences de la Vie et de la Terre de Terminale des voies générale et technologique ». In : *Bulletin officiel spécial n°8 du 25 juillet 2019* (2019). URL : [https://cache.media.eduscol.education.fr/file/SPE8\\_MENJ\\_25\\_7\\_2019/11/4/spe252\\_annexe\\_1159114.pdf](https://cache.media.eduscol.education.fr/file/SPE8_MENJ_25_7_2019/11/4/spe252_annexe_1159114.pdf).
- [9] « Programme de l'enseignement scientifique de la classe de Terminale de la voie générale ». In : *Bulletin officiel spécial n°8 du 25 juillet 2019* (2019). URL : [https://www.education.gouv.fr/pid285/bulletin\\_officiel.html?pid\\_bo=39051](https://www.education.gouv.fr/pid285/bulletin_officiel.html?pid_bo=39051).
- [10] Angelo VULPIANI. *Hasard, probabilité et complexité*. Nouvelles visions des sciences. Spartacus-ldh, 2021, p. 172. ISBN : 978-2-36693-084-9. URL : <https://spartacus-idh.com/liseuse/084/>.