



Algorithmes et Optimisation

Que vous soyez...

... affairiste notoire, politicien corrompu, sportif dopé, escroc rutilant...
ou tout simplement prudent,

Cryptez vos données sensibles!

Belzebuth et la Société de Calcul Mathématique, SA, présentent

Sieve II le diabolique

La référence ultime en matière de cryptographie logicielle

• **simple d'installation** : Windows 95, 98, NT,

• **simple d'emploi** : mot de passe de 4 caractères,

• **inviolable** : Sieve II le diabolique utilise un procédé nouveau, le cryptage dynamique récursif, conçu par la SCM. Un prix de 10 000F est offert depuis deux ans à la première personne qui parviendra à décrypter le texte (écrit en français banal) qui accompagne le logiciel.

• **inconditionnellement garanti à vie**. Sieve II le diabolique crypte absolument tout : données (tout format), sons, images, exécutables, etc.

Livraison en recommandé, sous pli discret.

Prix de vente unitaire HT : 1 500F,
TTC : 1 809F.

(Pour quantités importantes, nous consulter)



BON DE COMMANDE

Je commande exemplaires du logiciel de cryptage Sieve II le diabolique, au prix de 1 809 F l'unité (1 500F HT).

NOM :

PRÉNOM :

ADRESSE DE LIVRAISON :

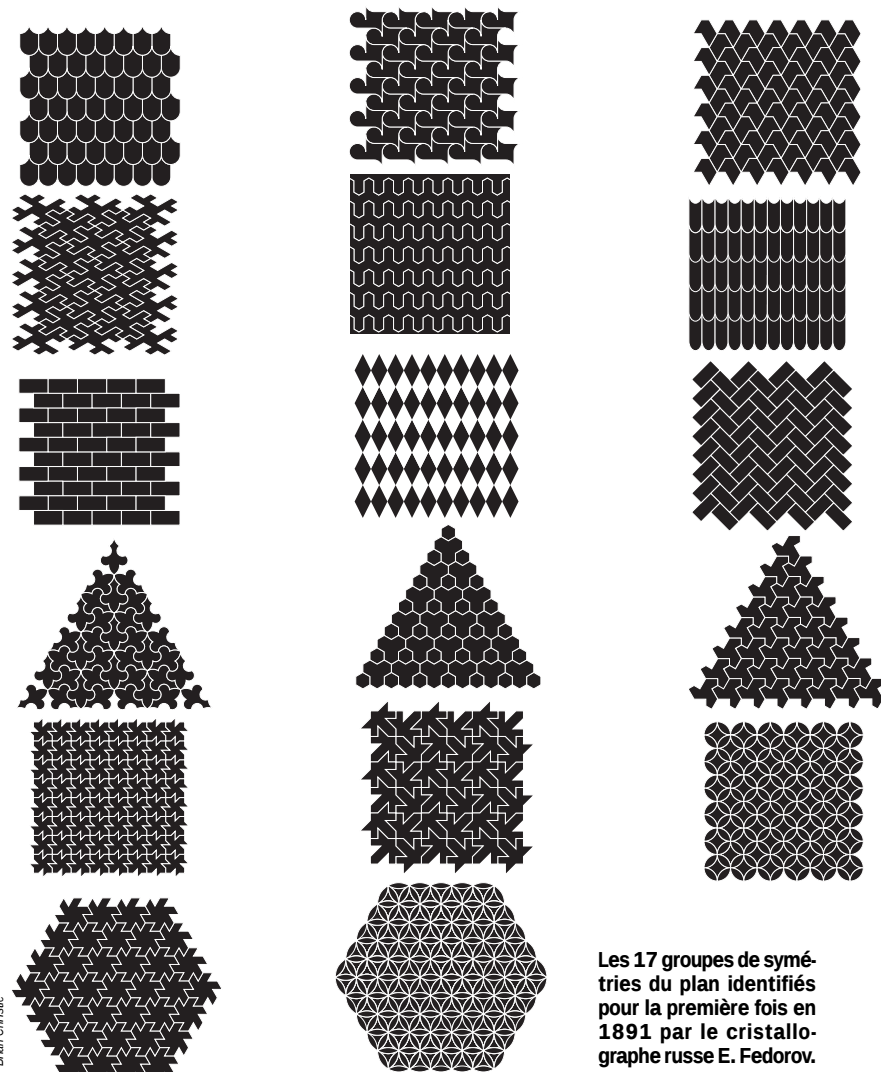
Bon à retourner, accompagné du règlement (chèque à l'ordre de SCM) à : Société de Calcul Mathématique, SA, 111, Faubourg Saint Honoré, 75 008 Paris.

de R. Grazebrook et des hexagones réguliers, on peut obtenir les 17 classes de symétries sauf une (Laquelle? Et comment obtient-on les autres? Je vous laisse le chercher).

R. Grazebrook doit l'idée initiale de ses pavages à Martin Gardner, qui fit les articles de la rubrique «Jeux mathématiques» de *Pour la Science* pendant des années. Elle étudiait l'art de l'Alhambra quand un article de Martin Gardner lui révéla les relations entre les motifs islamiques et les pavages du graveur Maurits Escher. Elle introduisit alors deux procédés distincts pour colorier ses pavés pentagonaux : soit en divisant les pavés en trois triangles (la série dite «Pentland»); soit en les divisant en quatre, (deux carés, un quadrilatère en forme de cerf-volant et un pentagone plus petit, de la série «Penthouse»). On peut naturellement colorier les pavés pentagonaux de R. Grazebrook d'autres manières, mais ces deux séries seulement engendrent une surprenante variété de motifs.

Réactions

L'article sur les lancers de pièces et les dés (*Abrogeons la loi des moyennes*, juin 1998, p. 122) a intéressé Tom Guldbrandsen, de Lyngby, au Danemark. Supposons que nous lancions un grand nombre de fois un dé, et que nous notions le nombre de 1, de 2, ..., de 6 obtenus. Quelle est la probabilité que les six lancers soient en nombreux égaux, quand vient le n -ième lancer? T. Guldbrandsen a d'abord observé que l'égalité ne peut se produire qu'après 6, 12, 18... lancers : le nombre de lancers doit être un multiple de 6. Il a trouvé une formule qui donne, en fonction de n , la probabilité que les six totaux soient égaux. Comme cette égalité peut être vérifiée plusieurs fois, il a obtenu comme probabilité : 0,021903735824.



Les 17 groupes de symétries du plan identifiés pour la première fois en 1891 par le cristallographe russe E. Fedorov.

Brigitte



La fin de la lignée humaine?

JACQUES NINIO

Les mutations délétères semblent si nombreuses que la lignée humaine devrait être éteinte depuis longtemps. Où est l'erreur?

La science-fiction prête aux mutants des capacités extraordinaires. Pourtant, nous sommes tous des mutants, mais dans quelle proportion? Lorsqu'un enfant naît, il hérite d'un mélange de gènes de son père et de sa mère. En raison des erreurs de reproduction, quelques gènes de l'enfant diffèrent légèrement de ceux de ses parents. Ces erreurs, les mutations, sont des modifications localisées des séquences d'ADN, la molécule en double hélice qui constitue les gènes. Certaines mutations, nocives, finissent par être éliminées. D'autres sont bénéfiques et permettent une meilleure adaptation de l'espèce (voir la figure 1). Le plus souvent, elles sont sans effet.

Comment évaluer la fréquence des mutations? Une première méthode consiste à recenser les individus porteurs d'une anomalie génétique absente chez les parents. Les statistiques portent sur les anomalies les plus fréquentes : l'achondroplasie, l'hémophilie A ou la dystrophie musculaire de Duchenne. Pour ces anomalies, le taux de mutations est compris entre un et 5×10^{-5} par gène et par génération. Toutefois, on ne peut pas généraliser ces fréquences à l'ensemble des mutations du génome, car les anomalies génétiques demeurent anecdotiques : elles concernent peu d'individus et des portions très courtes de certains gènes (une dizaine de nucléotides sur le millier que comporte le gène), de sorte que l'incertitude statistique est élevée.

Une évaluation plus sûre du taux de mutation est fondée sur le séquençage systématique d'un même gène chez un grand nombre d'individus et chez leurs parents. On recense chez les enfants les mutations dans un gène, puis on extrapole à l'ensemble du génome (voir la figure 2). Selon ces études, les biologistes estiment le taux M de mutations produites d'une génération sur l'autre à 10^{-8} par nucléotide et par génération, soit soixante mutations au total par

individu puisque le matériel génétique d'une cellule humaine comporte 23 paires de chromosomes, soit six milliards de nucléotides.

MUTATIONS TRANSITOIRES ET MUTATIONS FIXÉES

Les biologistes des populations évaluent différemment le taux de mutation. D'une part, la datation du plus jeune fossile connu commun à deux lignées indique quand ces lignées, deux espèces de singes par exemple, se sont séparées. D'autre part, on compare les séquences d'ADN actuelles de ces deux espèces. On déduit de ces mesures le taux moyen de mutations effectivement fixées au cours de l'évolution, qui se sont répandues dans une espèce. Ainsi, chez l'homme, en moyenne, deux mutations sont fixées par milliard de nucléotides et par an, ce qui, si l'on suppose

qu'une génération dure 20 ans, correspond à un taux extrapolé E_x de 4×10^{-8} mutations fixées par nucléotide et par génération.

De ce taux de mutations fixées dans l'espèce, comment déduit-on M , le taux qui nous intéresse, c'est-à-dire le taux de mutations produites «à la source» chez les individus? Le raisonnement nécessite l'examen du taux F de mutations fixées par un mécanisme purement aléatoire, en négligeant les effets de la sélection naturelle.

D'abord, on admet que, sur le long terme, le taux E de mutations effectivement fixées ne peut être supérieur au taux F de mutations fixées sans sélection naturelle : l'évolution trie et élimine, mais n'accroît pas la variabilité.

Chez les bactéries, la variabilité est faible, et la durée entre générations est courte. La sélection agit de manière féroce en éliminant toutes les mutations qui produiraient un désavantage, fût-il minime (E est bien inférieur à F). Chez l'homme, la variabilité est relativement grande, la durée entre les générations est longue, et quand au bout de quelques siècles ou quelques millénaires, la sélection a retenu une mutation avantageuse dans un gène, dix autres se sont fixées dans le même gène, mutations qui sont neutres ou légèrement délétères : la sélection naturelle n'a pas beaucoup joué et le taux de mutations effectivement fixées au cours de l'évolution est proche du taux sans sélection naturelle (E est proche de F).

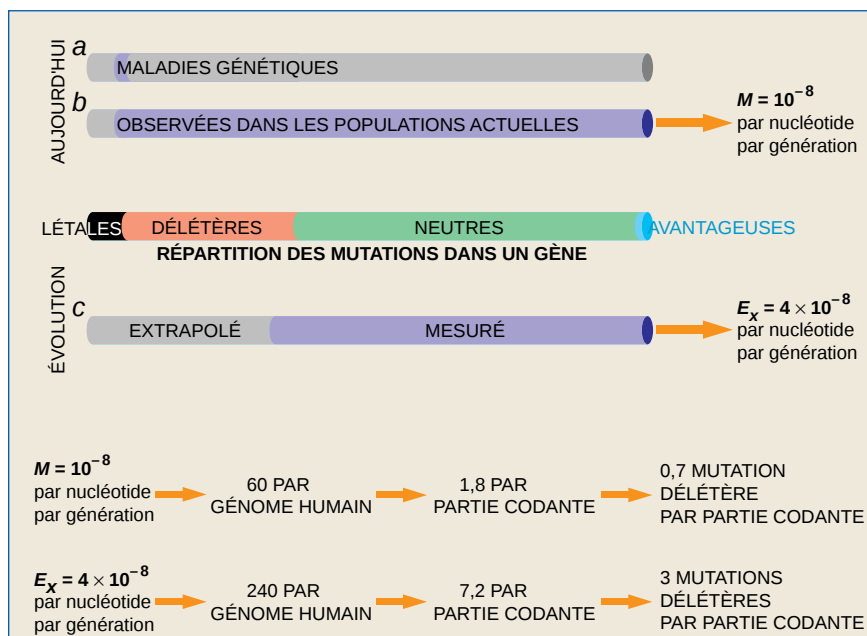
Par ailleurs, les biologistes des populations supposent que le taux F des mutations qui se seraient fixées sans sélection est égal au taux M de mutations produites. C'est un résultat intouchable de la théorie de l'évolution, au cœur de la génétique des populations actuelle. Cette égalité théorique relie les données mesurées sur les espèces (comparaison des séquences) aux données sur la production des



Stéphan Meyer

K. Stepien

1. Les mutations, dues à des modifications de l'ADN, sont la source de diversité entre les espèces, par exemple entre le premier mouton domestique d'il y a 8 000 ans, dont le mouflon de Corse (en haut) est le descendant direct, et le mouton à laine mérinos (en bas), race apparue au Moyen Âge.



2. Pour estimer le taux de mutation, on analyse les maladies génétiques, mais ces maladies concernent une toute petite partie du génome (a) et il est difficile de généraliser à l'ensemble du génome. On répertorie aussi les mutations entre parents et enfants (b). On en déduit M , le taux de mutation observé dans la population actuelle. Par ailleurs, les biologistes des populations estiment E_x le taux des mutations qui se sont fixées au cours de l'évolution (a). Étrangement, on trouve E_x est supérieur à M , ce qui signifierait que les mutations s'installent plus vite qu'elles ne se créent.

mutations à la source. Intuitivement, lorsque vous comparez l'un de vos gènes à celui d'un ancêtre d'il y a un million d'années, les différences sont d'une part dues aux mutations bénéfiques sélectionnées et conservées au cours du temps (une faible fraction du total), d'autre part, et pour l'essen-

tiel, aux mutations produites dans la descendance et qui n'ont pas eu d'effet sélectif. Cette égalité suppose, entre autres, que les mutations se produisent indépendamment.

Finalement, ce raisonnement aboutit à un taux de mutations fixées au cours de l'évolution égal au taux de mutations produites (E est égal à M). Or, le taux des mutations fixées fourni par l'évolution moléculaire ($E_x = 4 \times 10^{-8}$ mutation par nucléotide et par génération) est de deux à quatre fois supérieur au taux des mutations produites fourni par la génétique (M). Certes, ces deux valeurs sont incertaines, mais, même en tenant compte des incertitudes, un écart d'un facteur au moins deux demeure : les mutations s'installent (critères d'évolution moléculaire) plus vite qu'elles ne sont produites (critère génétique). Où est l'erreur ?

HYPOTHÈSES ERRONÉES ?

Le taux de mutations produites mesuré par la génétique (M) me semble correct, mais le raisonnement qui permet de relier E à M est suspect. Les hypothèses de travail de la biologie des populations pourraient être remis en cause, notamment celle selon laquelle les mutations surviennent indépendamment. Souvent, dans un intervalle de deux ou trois générations, deux mutations différentes

apparaissent sur un même gène, chez deux individus d'une petite population humaine. Or, si les mutations surviennent de manière indépendante, de telles occurrences seraient rarissimes.

Les généticiens des populations ne tiennent pas compte des mécanismes génétiques qui violent l'indépendance des mutations et les concentrent en certaines régions du génome. Comme ils estiment que leur évaluation du taux de mutations fixées (E_x) est fiable, ils concluent que l'estimation du taux de mutations engendrées (M) est erroné (ce qui n'est nullement exclu). En suivant ce raisonnement, on aboutit à un taux de mutation chez l'homme qui serait supérieur ou égal à 4×10^{-8} par nucléotide et par génération. D'où 240 mutations par génération, dont sept à huit, à chaque naissance, dans les régions codantes.

Selon une récente étude tirée des comparaisons de séquences dans un gène, 40 pour cent des mutations seraient délétères. C'est la mesure «en négatif» des mutations manquantes que l'on aurait dû trouver, et qui ont donc été éliminées. Ainsi, à chaque naissance, trois à quatre mutations sont désavantageuses. Le généticien des populations James Crow, de l'Université du Wisconsin, estime que chaque mutation délétère doit être éliminée par une mort prématurée ou par des chances de reproduction diminuée. Ce qui implique trois morts «génétiques» par personne ! Pourquoi l'espèce humaine n'est-elle pas éteinte ?

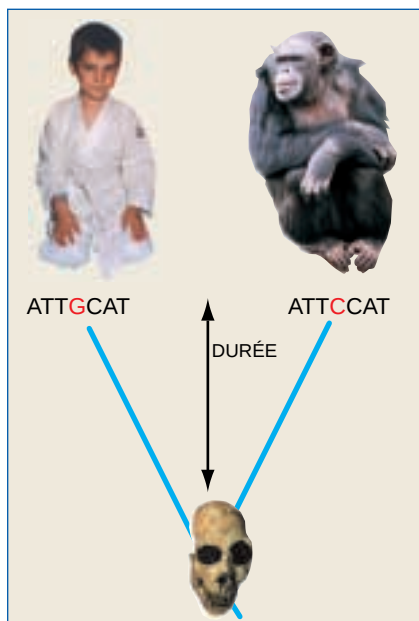
J. Crow invoque la reproduction sexuée, qui dilue les mauvais gènes, et inversement, des interactions entre mauvais gènes qui les élimineraient par paquets. Si ces effets jouent dans le sens voulu, ils ne me semblent pas suffisamment efficaces pour éliminer la difficulté. Pour l'avenir de l'espèce humaine, espérons que les généticiens des populations se trompent, lorsqu'ils passent des taux de mutations fixées aux taux de mutations produites.

Jacques NINIO est physicien au Laboratoire de physique statistique de l'École normale supérieure de Paris.

J.F. CROW, *The odds of losing at genetic roulette*, in *Nature*, vol. 397, pp. 293-294, 1999.

J. FEINGOLD, *Mutations multiples d'un même gène dans les petites populations*, in *Compte rendus de l'Académie des sciences, Sciences de la vie*, vol. 321, pp. 553-555, 1998.

J. NINIO, *The evolutionary design of error-rates and the fast fixation enigma*, in *Origins of Life*, vol. 27, pp. 609-621, 1997.



3. Les biologistes des populations comparent l'ADN de deux espèces dont on peut dater un ancêtre commun grâce à l'étude de fossiles. De la sorte, on estime le taux des mutations qui se sont installées au cours de l'évolution.

