

des banques de données. Ces comparaisons spécifiques de l'échantillon ont permis de découvrir un gène sur-exprimé dans les ostéoclastes : celui d'une classe de molécules déjà identifiées et nommées cathepsines. Sans la bio-informatique, qui a ainsi fourni une cible thérapeutique prometteuse en quelques semaines, des années d'expériences de laboratoire classiques et un peu de chance auraient été nécessaires.

Aujourd'hui, les biologistes cherchent un médicament pour bloquer la cible de la cathepsine K. Les recherches de principes actifs, ainsi que les évaluations de l'activité, de la toxicité et de l'absorption, se font encore dans un laboratoire de biochimie classique, c'est-à-dire sur une paillasse avec de la verrerie. Ce travail long et fastidieux sera progressivement confié aux ordinateurs, grâce aux nouveaux outils bio-informatiques et à la quantité croissante de données sur les structures des protéines et sur les voies métaboliques de l'organisme (voir l'encadré de la page 48) : après la biologie *in vivo* est venue la biologie *in vitro* ; bientôt, elle sera *in silico*.

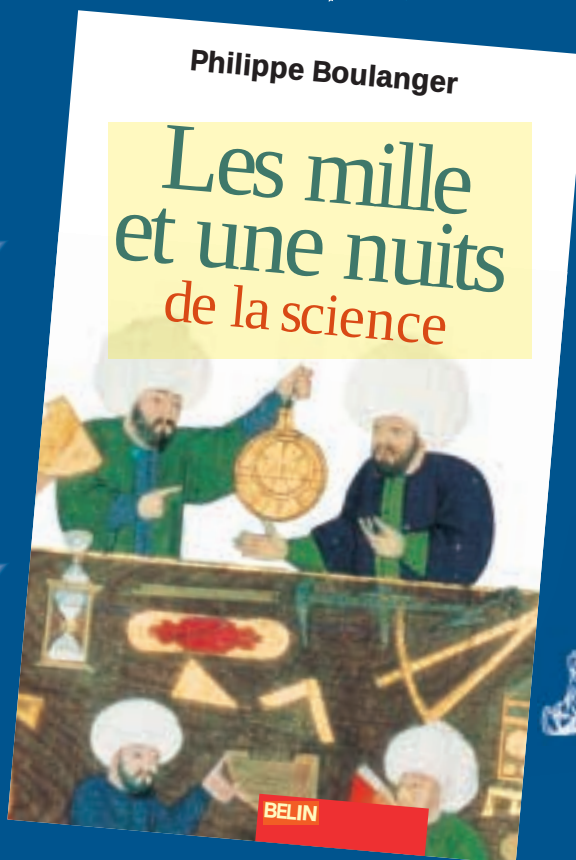
Beaucoup pensent que les promesses réelles de la génomique ne seront tenues que si la bio-informatique se développe : ce n'est pas la génomique qui est révolutionnaire, mais l'utilisation qui peut en être faite.

La bio-informatique à domicile

Cette révolution bio-informatique se fonde sur des groupes de toutes tailles. Certaines entreprises de bio-informatique pourvoient de gros clients, en ciblant leurs produits : elles créent des programmes personnalisés et des services de consultation adaptés à l'ensemble de l'entreprise. En revanche, d'autres visent des clientèles plus petites ou des organismes universitaires : elles proposent, par le réseau *Internet*, des services d'achat à la carte et de consultation de banque de données.

Enfin, de grands groupes pharmaceutiques valorisent leurs travaux en génomique en créant des unités internes de bio-informatique. Parmi eux, certains ont organisé de vastes services afin de mettre en œuvre les logiciels d'accès aux bases de données : ces groupes veulent ainsi rapprocher des services classiquement séparés, tels ceux qui sont chargés de la recherche de nouveaux

Entrez dans l'ambiance orientale
des mille et une nuits et laissez-vous
charmer par l'imagination
de Shéhérazade contant la science
et les tribulations des savants

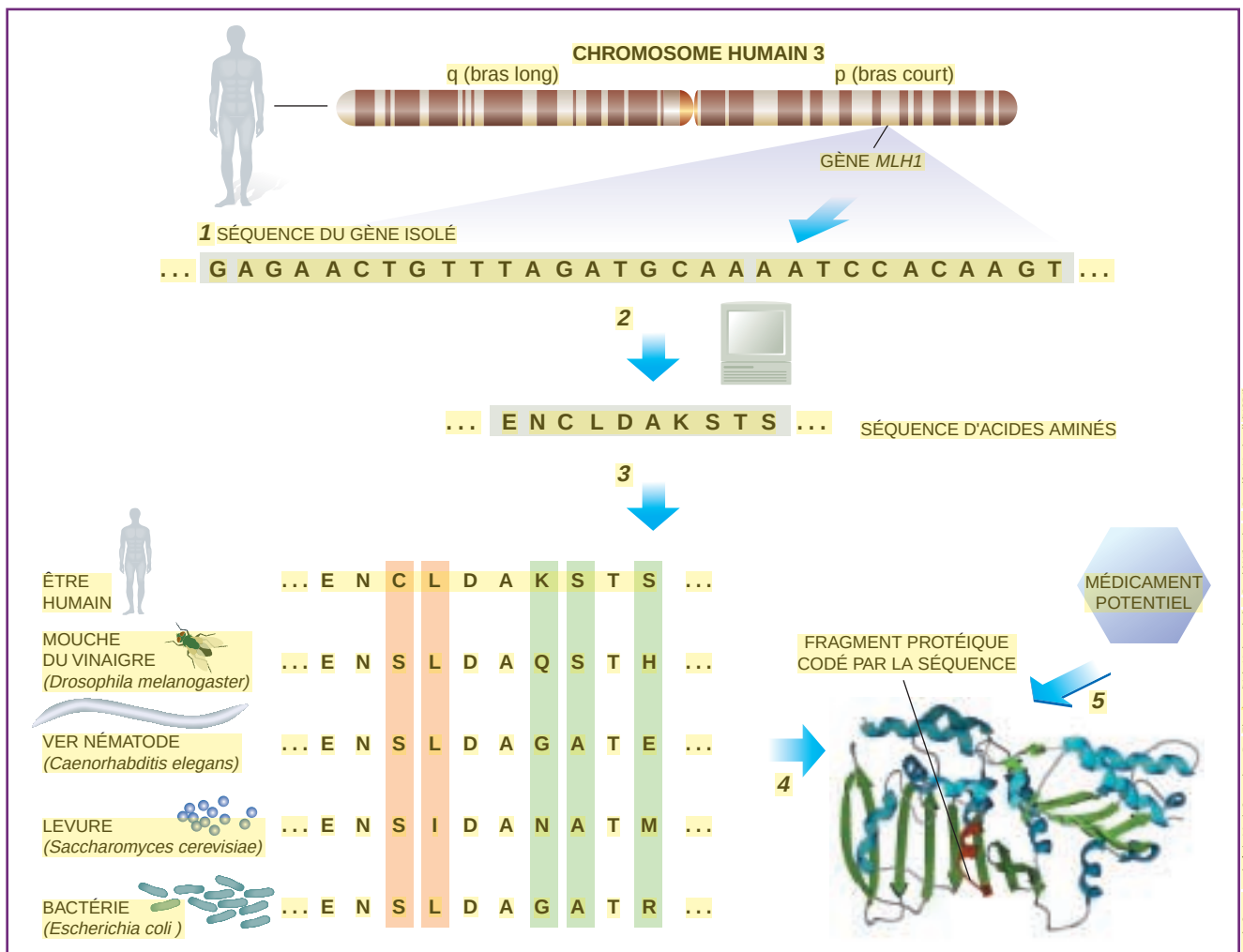


«Dans ce livre, Philippe Boulanger, Directeur de la revue *Pour la Science*, réussit le tour de force de vulgariser des concepts difficiles en charmant son lecteur par une subtile danse... des neurones».

Hervé Ponchelet
Le Point

Code 2445 • 240 pages • Prix 100 F
Voir bon de commande p. 98

BELIN • POUR LA SCIENCE



Laurie Grace, assistée de Mark Gerstein et Pat Fleming (Yale University) et David Wheeler et Jennifer Viskochil (NCBI)

3. LA COMPARAISON DE GÈNES HUMAINS avec ceux d'animaux modèles offre parfois des pistes pour la recherche de nouveaux médicaments. Par exemple, quand on connaît la séquence du gène humain *MLH1* (1), associé au cancer du côlon, un programme informatique peut traduire cette séquence d'ADN en une séquence d'acides aminés, qui compose la protéine codée (2). On cherche ensuite des séquences similaires dans des banques de données

de protéines synthétisées par des organismes modèles (3). Les bandes vertes indiquent des zones de grande variabilité (les acides aminés sont peu conservés d'un organisme à un autre), et les bandes orange, des zones de faible variabilité. On modélise la protéine humaine à partir des structures des protéines similaires des organismes modèles (4) et, enfin, on crée une molécule qui se fixe à la protéine modélisée (5).

produits, de la formulation, de la toxicologie et des essais cliniques.

Naguère, ces services étaient excessivement cloisonnés, et les données d'un service restaient inaccessibles aux autres équipes. La bio-informatique donnera la même information à tous, et chacun fera un usage individuel des données.

En outre, les services internes communs de bio-informatique réduisent les investissements informatiques. Par exemple, les moyens logiciels individuels, utilisés par les différents chercheurs d'une même entreprise pour l'accès aux bases de données et le traitement de ces données, sont remplacés par une plate-forme logicielle unique : les économies se comptent en millions de francs.

Afin d'intégrer pleinement la bio-informatique à leurs sociétés, des

groupes pharmaceutiques concluent parfois des alliances stratégiques ou même acquièrent des sociétés de génie génétique plus petites ; ainsi munis de partenaires et de réseaux commerciaux, ils augmentent leurs compétences bio-informatiques et s'adaptent aux nouvelles techniques sans réorganiser leurs propres structures.

Toutefois, le regroupement des banques de données est notoirement périlleux. On cherche à éviter les mauvaises communications à l'aide de systèmes de nomenclature des données comprise aussi bien des diverses banques de données que des systèmes d'appellation. Imaginons qu'une banque A communique avec une banque B, celle-ci avec une banque C, C avec D, et ainsi de suite. Si A change ses données, les références de la banque D

risquent d'être périmées, surtout quand le renouvellement des données est permanent. Ce problème s'aggrave, car les connaissances en biologie et les analyses par ordinateur sont de plus en plus complexes.

Des améliorations systématiques seront d'une grande aide, mais le vrai progrès semble reposer sur l'ingéniosité des utilisateurs : le succès de la bio-informatique n'est pas une affaire de matériel, ni de logiciel, mais bien d'intelligence.

Ken HOWARD est journaliste scientifique.

D. B. SEARLS, *Using Bioinformatics in Gene and Drug Discovery*, in *Drug Discovery Today*, vol. 5, n° 4, pp. 135-143, avril 2000.