



The DNA revolution

La révolution de l'ADN

This century has been dubbed the century of biology — and for very good reason. Since the discovery of the structure of DNA by Watson and Crick in 1953, scientists have continued their unrelenting assault on the information hidden in the DNA of humans, other animals, and plants. Numerous nucleic acid-based techniques have been developed to allow us to identify macro- and micro-organisms, diagnose infectious disease, trace sources of microbial agents of disease, identify susceptibility to disease, and improve our understanding of evolution. Much of this work is based on identification of a single gene or a small set of genes. That approach continues to be important, but there is now considerable activity in taking a genomic approach, which addresses all the genes or a large subset of the genes in an organism, whether it be a virus, a bacterium, an animal, or a person.

The genomic approach is feasible because of technological developments that make it possible to determine the nucleotide sequence of millions of base pairs of DNA at increasing speed and decreasing cost. The human genome project was a major driver in these developments. This project took 13 years to complete in 2003 and cost 2.7 billion dollars. By 2007, DNA discoverer James Watson had his genome sequenced in 2 months at a cost of \$2 million. There is now an international plan involving sequencing centres in the United States, the United Kingdom and China, to sequence 1000 genomes in 3 years. This will facilitate identification of moderately common and rare genetic variants and association of specific genetic variations with disease. This project will add 60 times the sequence data that have accumulated in public data bases over the past 25 years. The Wellcome Trust is funding a collaboration of some 60 institutes around the world to sequence 100 000 genomes in studies to associate genetic changes with 25 diseases. This will require the handling of over 100 billion pieces of genetic data. It appears that the \$1000 genome is not far away.

There is enormous interest in developing applications from genome sequence data. Although the genomes of any two individuals are about 99.9% identical, the 0.1% variation allows for changes in a massive number of nucleotides, given that the human genome is approximately 30 billion base pairs. Medicine and agriculture are prime areas for these applications, which include pharmacogenomics (custom drugs and new approaches to drug design), diagnostics, gene therapy, and improved selection of animals. One area that is garnering a great deal of

Ce siècle a été surnommé le siècle de la biologie — et pour de très bonnes raisons. Depuis la découverte de la structure de l'ADN par Watson et Crick en 1953, les scientifiques ont poursuivi leur implacable assaut sur l'information dissimulée dans l'ADN des humains, des autres animaux et des plantes. De nombreuses techniques fondées sur l'acide nucléique ont été mises au point pour nous permettre d'identifier les macro et les microorganismes, de diagnostiquer des maladies infectieuses, de retracer les sources d'agents microbiens des maladies, d'identifier la susceptibilité à la maladie et d'améliorer notre compréhension de l'évolution. Une bonne part de ce travail se base sur l'identification d'un seul gène ou d'un petit groupe de gènes. Cette approche continue d'être importante, mais il y a maintenant de l'activité considérable en vue de l'adoption d'une approche génomique, qui étudie tous les gènes ou un grand sous-groupe de gènes dans un organisme, qu'il s'agisse d'un virus, d'une bactérie, d'un animal ou d'une personne.

L'approche génomique est réalisable en raison des progrès technologiques qui permettent la détermination de la séquence des nucléotides de millions de paires de bases d'ADN de plus en plus rapidement et à un coût qui continue de baisser. Le projet du génome humain a été une force motrice majeure de ces progrès. Ce projet, terminé en 2003, a exigé treize ans et a coûté 2,7 milliards \$. En 2007, James Watson, le responsable de la découverte de l'ADN, a établi la séquence de son génome en deux mois à un coût de 2 millions \$. Il existe maintenant un projet international auquel participent des centres de séquençage aux États-Unis, au Royaume-Uni et en Chine pour réaliser le séquençage de 1000 génomes en trois ans. Cela facilitera l'identification de variantes génétiques modérément communes et rares et l'association de variations génétiques spécifiques avec la maladie. Ce projet ajoutera 60 fois les données de séquençage qui se sont accumulées dans les bases de données publiques au cours des 25 dernières années. Le Wellcome Trust finance une collaboration de quelque 60 instituts de par le monde pour séquencer 100 000 génomes dans des études visant à associer les changements génétiques avec 25 maladies. Cela exigera la manipulation de plus de 100 milliards d'éléments de données génétiques. Le génome à 1000 \$ ne semble plus une réalité lointaine.

Il existe un intérêt énorme pour le développement d'applications à partir des données de séquençage du génome. Même si les génomes de deux personnes sont identiques à 99,9 %, la

attention is the identification of genes that are associated in complex ways with disease. A haplotype map (a HapMap) that is being developed will reduce the search from some 10 million nucleic acid changes to about 500 000 haplotypes. The haplotype is a combination of alleles along the chromosome that is identified on the basis that a variation at one site is associated with certain alleles at nearby variant sites (because these sets of genes are inherited as a block). The map will facilitate recognition of genes associated with diseases that have a significant heritable component and will make it easier to investigate the influence of environment and of interventions on disease.

Sequencing of the genomes of animals is of immense importance to agriculture, veterinary medicine, and human life. Sequence data will allow better selection of genetic stock and will enhance investigation of environmental influences, as well as drug and vaccine effects. Following his return from a recent

“Biology of the Genomes” meeting in New York, Ewan Birney, senior scientist at the European Molecular Biology Laboratory noted in his blog that the most exciting session was on “non-traditional models — dogs, horses and cows, where the ability to do cost effective genotyping has completely revolutionized this field.” Genome sequences are now available for several animal species including the dog, cat, cow, chicken, pig and horse. In pet animals “personalized genetics” will permit more precise treatments; in food animals and horses we can expect more rapid and effective selection for desired traits. These developments will have significant implications for the way we educate veterinarians, for what clients will expect of veterinarians, and for the services that our profession will be able to offer. ■

Carlton Gyles

variation de 0,1 % permet le changement d'un nombre massif de nucléotides, vu que le génome humain se compose d'environ 30 milliards de paires de bases. La médecine et l'agriculture sont des domaines privilégiés pour ces applications, qui incluent la pharmacogénomique (médicaments sur mesure et nouvelles approches pour la conception de médicaments), les techniques diagnostiques, la thérapie génique et la sélection améliorée d'animaux. Un domaine qui suscite beaucoup d'attention est l'identification des gènes qui sont associés de manières complexes à la maladie. Une carte haplotype en cours d'élaboration réduira la recherche de quelque 10 millions de changements d'acide nucléique à environ 500 000 haplotypes. L'haplotype est une combinaison d'allèles le long du chromosome qui est identifiée selon qu'une variation à un site est associée à certains allèles dans les sites variants environnants (parce que ces groupes de gènes sont hérités en bloc). La carte facilitera la reconnaissance des gènes associés aux maladies qui ont une composante héréditaire importante et permettra d'étudier plus facilement l'influence de l'environnement et des interventions sur la maladie.

Le séquençage des génomes des animaux est d'une importance immense pour l'agriculture, la médecine vétérinaire et la vie humaine. Les données de séquençage permettront une meilleure

sélection des souches génétiques et amélioreront l'étude des influences environnementales ainsi que des effets des médicaments et des vaccins. Après son retour d'une réunion récente sur la biologie des génomes à New York, Ewan Birney, scientifique principal au Laboratoire européen de biologie moléculaire a fait remarquer dans son blogue que la séance la plus excitante portait sur «les modèles non traditionnels — les chiens, les chevaux et les vaches, où la capacité de faire du génotypage rentable a complètement révolutionné ce domaine». Des séquences de génomes sont maintenant disponibles pour plusieurs espèces d'animaux, incluant le chien, le chat, la vache, le poulet, le cochon et le cheval. Chez les animaux de compagnie, la «génétique personnalisée» permettra des traitements plus précis; chez les animaux de consommation et les chevaux, nous pouvons nous attendre à une sélection plus rapide et efficace des caractéristiques souhaitées. Ces progrès auront des répercussions importantes sur la façon dont nous formons les vétérinaires, les attentes futures des clients à l'égard des vétérinaires et les services que notre profession sera en mesure d'offrir. ■

Carlton Gyles