

La bio-informatique au Québec :

un levier essentiel du développement
des bio-industries

Avis



Québec 
Conseil de la science
et de la technologie

20 14 234
54
3201 56
67898
45 29 1 23
65
444 534
8795
5622 12873
24
78621 55
6744
34 9785
27897
2756 24
24975
7856 354
2867

Conseil de la science
et de la technologie
1200, route de l'Église
3^e étage – Local 3.45
Sainte-Foy (Québec)
G1V 4Z2

Téléphone : (418) 644-1165
Télécopieur : (418) 646-0920

Ce document est disponible sur le site Web
du Conseil de la science et de la technologie
<http://www.cst.gouv.qc.ca>

**Conception graphique,
typographie et mise en pages**
Bruno Balatti Design

© Gouvernement du Québec 2001
Premier tirage, janvier 2001
Dépôt légal : 1^{er} trimestre 2001

Bibliothèque nationale du Québec
Bibliothèque nationale du Canada

ISBN : 2-550-36957-2

Janvier 2001

Monsieur Jean Rochon
Ministre de la Recherche, de la
Science et de la Technologie
Gouvernement du Québec
Québec

Monsieur le Ministre,

Conformément aux dispositions de l'article 16 de la Loi du ministère de la Recherche, de la Science et de la Technologie, j'ai l'honneur de vous transmettre l'avis du Conseil de la science et de la technologie intitulé *La bio-informatique au Québec : un levier essentiel du développement des bio-industries*.

Je vous prie d'agréer, Monsieur le Ministre, l'expression de ma haute considération.

La présidente,
Hélène P. Tremblay

Remerciements

Ce nouvel avis du Conseil est le fruit de collaborations multiples, tout au long de sa conception et de son édition, qui doivent être soulignées.

Je tiens donc en premier lieu à remercier les membres du Conseil pour leur participation au devis initial, de même qu'aux orientations et au contenu spécifique du document.

Je remercie également les membres du comité de pilotage, le président, M. Martin Godbout, président de Hodran inc., M. Jean-Marc Proulx, vice-président R-D du Groupe Conseil DMR inc., tous deux membres du Conseil, de même que MM. Paul Beaulieu, directeur général de la Chaire en gestion des bio-industries de l'Université du Québec à Montréal, Thomas Hudson, directeur du Centre génomique de Montréal, Jacques Hurtubise, directeur du Centre de recherches mathématiques de l'Université de Montréal, et Jean Morrissette, chercheur senior au Centre de recherche en endocrinologie moléculaire et oncologie de la Faculté de médecine de l'Université Laval.

Leur expertise reconnue et leur disponibilité tout au long des sessions de travail auront permis de proposer au Conseil une lecture véritablement prospective de ce secteur en émergence.

Je tiens ensuite à remercier M. Alain Bergeron, agent de recherche au Conseil, pour la coordination des travaux et la rédaction du document. Sa maîtrise du sujet et sa capacité à extraire l'essentiel d'une information encore disparate et mouvante ont rendu sa contribution extrêmement précieuse.

La mise en forme matérielle du document a été effectuée par Hélène Lafrance.

La présidente

La bio-informatique au Québec :
un levier essentiel du développement des bio-industries

Introduction	10
Chapitre 1	
Perspectives de la bio-informatique dans le monde	13
1.1 Qu'est-ce que la bio-informatique?	13
1.2 Bio-informatique et industrie	17
1.3 Efforts nationaux et internationaux	21
1.4 Une pénurie universelle de ressources humaines	23
Chapitre 2	
Perspectives de la bio-informatique au Canada et au Québec	25
2.1 Situation au Canada	25
2.2 Situation au Québec	28
2.3 Enjeux et atouts du Québec en bio-informatique	32
Chapitre 3	
Une stratégie de démarrage pour la bio-informatique au Québec	35
Volet formation	36
Volet recherche et réseautage	40
Volet transfert de compétences vers l'industrie	42
Suivre de près l'évolution du domaine	44
Annexe 1 Membres du Conseil de la science et de la technologie	46
Annexe 2 Membres du Comité sur la bio-informatique	48

Avant-propos

Au cours des dernières années, le Québec a réussi à se doter d'une industrie forte et réputée en biotechnologies, notamment en biopharmaceutique, qui le place au premier rang des provinces canadiennes. Il s'est doté de plus d'une nouvelle organisation, Génome Québec, pour appuyer son développement en génomique et protéomique. Il s'agit là d'atouts majeurs, mais encore fragiles, vu la rapidité avec laquelle les grands acteurs mondiaux se font concurrence actuellement pour détenir le leadership scientifique et économique dans ces créneaux.

Or, force est de constater que l'essor des bio-industries repose en grande partie sur le déploiement d'un secteur en amont encore très embryonnaire au Québec, celui de la bio-informatique.

La bio-informatique désigne globalement l'ensemble des applications de l'informatique aux sciences biologiques et apparaît comme champ du savoir au croisement des sciences de la vie, de l'informatique et des mathématiques. Sans former, pour l'instant du moins, un secteur industriel spécifique – la nomenclature de ses composantes demeure encore arbitraire –, la bio-informatique se révèle déjà une plate-forme obligée pour asseoir la recherche et le développement dans les différentes branches des sciences de la vie, que ce soit en santé humaine, en agriculture ou en matière de protection de l'environnement.

La bio-informatique devient ainsi une clé stratégique pour le développement scientifique et technologique d'un des nouveaux fleurons québécois. Toutefois, le temps constitue une dimension essentielle : si le Québec réussit à acquérir une masse critique de ressources spécialisées, il pourra alors prétendre, non pas devenir un acteur de premier plan, certes, sa taille ne le lui permettant

pas, mais bien supporter la croissance des industries en aval déjà présentes, se doter d'une expertise propre et déterminer en toute connaissance de cause les choix sociaux les plus appropriés.

Développer la bio-informatique signifie en tout premier lieu former rapidement des chercheurs universitaires et industriels en nombre suffisant pour satisfaire à la fois les besoins de R-D et ceux de formation d'une main-d'œuvre hautement qualifiée. Cela signifie également constituer des lieux et réseaux de recherche importants, puisque les découvertes en bio-informatique et dans les domaines apparentés dérivent actuellement des résultats de la recherche fondamentale. Finalement, il sera utile, croyons-nous, d'appuyer durant quelques années les entreprises québécoises utilisatrices dans leurs efforts pour embaucher des bio-informaticiens : la concurrence internationale est déjà féroce, et les armes dans certains cas sont inégales.

Voilà en résumé les principales conclusions auxquelles le Conseil de la science et de la technologie arrive dans cet avis intitulé *La bio-informatique au Québec : un levier essentiel du développement des bio-industries*.

Le lecteur constatera qu'il s'agit d'un avis différent des précédents. En effet, il s'agit d'un sujet très spécialisé, sur un objet encore embryonnaire. Mais les enjeux qui seront énumérés dans l'avis sont tels et les actions si pressantes, que le Conseil, pleinement conscient de ses responsabilités en matière de vision prospective, a cru bon de partager ses préoccupations et d'énoncer les pistes de solutions qu'il croit les plus pertinentes à court terme.

La présidente

Résumé

Objet de l'avis

Le présent avis porte sur la bio-informatique, une plate-forme technologique multidisciplinaire qui joue un rôle stratégique central dans le développement de la génomique et de la protéomique. Son rôle est de pouvoir organiser, gérer et analyser efficacement des masses d'informations considérables, générées par ces domaines de recherche de pointe. La bio-informatique doit être considérée comme un des fondements actuels et futurs de la biotechnologie, de la recherche pharmaceutique et du développement général de toutes les bio-industries. La rareté des ressources humaines qualifiées dans ce domaine encore en émergence demeure cependant le principal obstacle au développement de la bio-informatique à travers le monde.

Structure de l'avis

Le rapport comprend trois chapitres. Le premier effectue un survol de la conjoncture scientifique et industrielle internationale dans laquelle se développe actuellement la bio-informatique. Le deuxième chapitre présente la situation canadienne et québécoise. Il s'efforce de mettre en relief les atouts et les opportunités que présente le développement de la bio-informatique au Québec, ainsi que les principaux obstacles à lever. Le troisième chapitre propose une stratégie d'ensemble pour organiser et structurer rapidement le démarrage de la bio-informatique au Québec.

Les grands enjeux pour le Québec

Les grands enjeux du développement de la bio-informatique au Québec sont les suivants.

- 1) Assurer la croissance de l'industrie québécoise des biotechnologies.** Comme tout indique que l'expansion des biotechnologies reposera de plus en plus sur la maîtrise des outils et des connaissances de la bio-informatique, les perspectives de développement de l'industrie québécoise des biotechnologies dépendront pour une bonne part de la disponibilité au Québec des ressources dont elle aura besoin dans ce domaine.
- 2) Maintenir la vigueur d'une R-D pharmaceutique de pointe au Québec.** Les entreprises pharmaceutiques mondiales ont surtout implanté des établissements de R-D au Québec. On sait qu'elles doivent réorienter leurs stratégies de recherche pour exploiter le potentiel de la génomique/protéomique. Le Québec a donc tout intérêt à développer une expertise en bio-informatique pour répondre à cette demande.
- 3) Offrir aux entreprises québécoises des TIC des possibilités d'expansion intéressantes.** Une tendance récente est l'entrée massive des grands joueurs des TIC dans la bio-informatique (IBM, Compaq, Hitachi...). La demande mondiale est appelée à croître rapidement. Si on parvient à développer celle du marché québécois, l'industrie des TIC québécoise devrait y trouver une perspective d'expansion très intéressante.
- 4) Accélérer l'intégration de la génomique/protéomique dans d'autres bio-industries québécoises.** Même si c'est surtout autour du biopharmaceutique que se développe mondialement la bio-informatique, les effets de débordement se font déjà sentir dans l'agroalimentaire, l'industrie de l'environnement, la sylviculture, etc. D'ici quelques années,

de plus en plus de bio-industries québécoises devront intégrer et exploiter à leur tour les connaissances en génomique et en protéomique, et donc devenir assez rapidement des utilisatrices de bio-informatique.

Des atouts non négligeables pour le Québec

Bien qu'il soit aux prises lui aussi avec le problème de pénurie des ressources humaines en bio-informatique, le Québec possède des atouts sur lesquels il peut miser pour construire une compétence en bio-informatique.

- 1) **Une qualité de recherche scientifique reconnue mondialement** dans le secteur biomédical et plus particulièrement dans des spécialités porteuses de la bio-informatique, comme le pharmaceutique.
- 2) **Une main-d'œuvre scientifique et technique de grande compétence**, dans les sciences de la vie et les bio-industries.
- 3) **La présence au Québec d'industries biopharmaceutiques et biotechnologiques** dynamiques qui devront compter de plus en plus sur la fonction bio-informatique pour assurer leur développement.
- 4) **Une localisation géographique** qui donne un accès rapide aux principaux centres de recherche universitaires et industriels, ainsi qu'aux marchés d'Amérique du Nord.
- 5) **Une panoplie de mesures de soutien au développement de la haute technologie** dont la bio-informatique pourrait bénéficier, notamment des mesures fiscales très favorables et la présence de sociétés de capital de risque.
- 6) **La présence d'une industrie des TIC très importante**, ainsi que de centres de calcul de haute puissance.
- 7) **Les programmes de Genome Canada et Génome Québec** qui vont soutenir dans les années qui viennent des projets suscitant une forte demande en bio-informatique.

Présentation de la stratégie

La stratégie retenue par le Conseil constitue une proposition d'investissement rapide et ciblée sur les quelques éléments clés qui pourront stimuler efficacement la structuration du domaine de la bio-informatique au Québec. Cette stratégie permet d'agir sur trois fronts simultanément : celui de la formation des ressources humaines, celui de l'organisation de l'activité bio-informatique comme volet constitutif du réseau de recherche québécois en génomique et protéomique ainsi que comme champ de recherche spécifique, et celui du soutien au développement d'une compétence industrielle en bio-informatique.

Rôle clé de Génome Québec

Pour réussir, une telle stratégie exige un maître d'œuvre qui prendra la responsabilité de l'organisation ou de la coordination d'ensemble. Le Conseil de la science et de la technologie propose de confier cette mission à Génome Québec. En effet, cet organisme est l'instrument dont le gouvernement du Québec s'est doté pour assurer le développement de la génomique et des domaines connexes au Québec et aucune autre organisation existante ne détient un mandat aussi bien adapté aux exigences formulées ici.

Le Conseil émet donc les recommandations suivantes.

Volet formation

Recommandation 1

En toute priorité, mettre en place, d'ici septembre 2001, un programme intensif de formation-perfectionnement en bio-informatique au Québec.

Volet recherche
et réseautage

Recommandation 2

Permettre, si possible dès l'hiver 2002, à une dizaine de boursiers d'acquérir une formation supérieure en bio-informatique à l'extérieur du Québec.

Recommandation 3

Mettre en place le plus rapidement possible des activités et des programmes universitaires de formation de deuxième et troisième cycles en bio-informatique au Québec.

Volet transfert de
compétences vers
l'industrie

Recommandation 4

Développer l'activité de bio-informatique au Québec à l'intérieur d'un réseau de projets de recherche structurants, en génomique, en protéomique, de même qu'en bio-informatique proprement dite.

Recommandation 5

Fournir aux entreprises des bio-industries du Québec, à partir de 2002, des incitatifs fiscaux spécifiques reliés à l'embauche de bio-informaticiens.

Suivre l'évolution du
domaine

Recommandation 6

Suivre de très près l'évolution de la génomique, de la protéomique et de la bio-informatique au Québec, tant du point de vue scientifique que du point de vue industriel et social, au cours des trois ou quatre prochaines années.

Introduction

Objet de l'avis

Avec la création de Génome Québec, le gouvernement du Québec vient de manifester son intérêt pour la génomique et la protéomique, qui sont des domaines scientifiques en pleine effervescence et dont les retombées économiques seront extrêmement importantes à moyen et à long termes.

Une telle initiative vise à favoriser la structuration d'un réseau de compétences ainsi qu'une participation accrue du Québec au développement de ce vaste domaine de recherche et d'applications.

Le présent avis du Conseil s'inscrit directement dans cette perspective. Il porte plus spécifiquement sur le développement de la bio-informatique, qui est une compétence absolument essentielle pour assurer l'essor de la génomique et de la protéomique au Québec. Cette compétence doit pouvoir se retrouver dans les centres de recherche comme dans l'industrie.

En fait — et il faut l'affirmer très clairement — il ne saurait être question d'investir adéquatement en génomique ou en protéomique au Québec sans s'appuyer sur une capacité d'analyse solide en bio-informatique. La mise en place de cette infrastructure nécessite des mesures précises et ciblées à très court terme. Il faut organiser rapidement et efficacement une capacité de formation dans le domaine, arrimer l'activité de bio-informatique au développement d'un réseau de recherche en génomique et protéomique, et faciliter l'embauche de bio-informaticiens dans les bio-industries¹. Le temps est un facteur prédominant ici. Tout va se jouer dans les deux ou trois prochaines années.

Le Conseil estime que le Québec a les atouts nécessaires pour devenir un joueur important en bio-informatique à l'échelle mondiale et favoriser l'essor de ses bio-industries, à condition d'agir rapidement et efficacement. La stratégie qu'il propose dans le chapitre 3 du présent rapport est à cet égard incontournable. Tous les acteurs interpellés doivent en prendre conscience et collaborer de façon étroite pour atteindre l'objectif.

Importance de la bio-informatique

La bio-informatique est une plate-forme technologique multidisciplinaire qui réunit des informaticiens, des développeurs de logiciels et d'équipements, des mathématiciens, des statisticiens et des biologistes. Elle s'occupe essentiellement de la gestion et de l'analyse des données qui sont générées par les sciences de la vie. Il s'agit donc d'un ensemble de connaissances et de techniques intégrées qui font appel à des spécialistes formés au confluent de toutes ces disciplines, les bio-informaticiens.

Des retombées économiques et sociales

Les retombées économiques de la bio-informatique sont surtout indirectes. La création d'une véritable industrie de la bio-informatique proprement dite paraît encore imprécise. Par contre, tout le vaste champ des bio-industries aura de plus

¹ Le terme de bio-industries regroupe habituellement l'ensemble des entreprises du vaste secteur des biotechnologies, dans ses très nombreux champs d'application. Par exemple, les entreprises membres de BIOQuébec, anciennement Association québécoise des bio-industries, se retrouvent dans des secteurs comme la santé humaine (y compris le bio-pharmaceutique), la santé animale, l'agriculture, le bioalimentaire, la foresterie et l'environnement.

en plus besoin de compter sur les ressources de la bio-informatique pour la conception et le développement de ses produits. L'industrie pharmaceutique, par exemple, connaît actuellement une mutation importante dans l'organisation des différentes étapes de sa recherche-développement. L'identification des cibles thérapeutiques, de même que la mise au point de médicaments, deviennent de plus en plus dépendantes de la ressource informationnelle, c'est-à-dire de l'exploitation massive de gigantesques bases de données informatisées.

À l'échelle mondiale, l'essor de la génomique et de la protéomique est porteur de grandes promesses pour le bien-être des populations. Dans le domaine de la santé, par exemple, on s'attend à ce que le traitement et la prévention des maladies génétiques, du cancer et des maladies dégénératives du système nerveux connaissent des percées spectaculaires dans les prochaines années. Ceci est particulièrement important pour le Québec qui se trouve à la fine pointe de plusieurs secteurs de la recherche biomédicale. Ici encore, le développement de la bio-informatique va s'avérer outil essentiel.

Nous ne pouvons ignorer que la manipulation du code génétique de certains organismes vivants soulève des controverses. Toutes les percées qui ont lieu dans les sciences et technologies du vivant ne sont pas nécessairement désirables ou acceptables par la société. Même si ces questions d'ordre éthique ne concernent pas directement la bio-informatique comme tel, le Conseil tient à souligner qu'il est conscient de la question. Il croit également que l'existence de risques ou d'inquiétudes en ce domaine constitue précisément une très bonne raison pour le Québec d'être présent dans le développement de la génomique et de la protéomique. C'est en développant sa propre expertise dans ces secteurs de pointe que la société québécoise pourra le mieux s'informer de ce qui se passe à l'échelle internationale et déterminer ses choix de la façon la plus éclairée possible. Génome Québec a une responsabilité importante à ce chapitre, puisqu'il doit communiquer avec le citoyen et se préoccuper des impacts sociaux et des dimensions éthiques de la recherche en génomique et protéomique.

À ce propos, le ministre de la Recherche, de la Science et de la Technologie a soumis à une vaste consultation un projet de politique scientifique et technologique au cours de l'été 2000, dans lequel il annonçait son intention de créer une commission de l'éthique de la science et de la technologie qu'il propose de rattacher au Conseil de la science et de la technologie. Même si les suites concrètes qu'il compte donner à ce projet, dans la version finale de la politique, ne sont pas encore connues au moment d'écrire ces lignes, il est clair que le Conseil s'intéressera de près, au cours des prochaines années, aux dimensions éthiques de ses objets d'avis.

Plan du rapport

Le présent rapport comprend trois chapitres. Le premier effectue un survol de la conjoncture scientifique et industrielle internationale dans laquelle se développe actuellement la bio-informatique. Le deuxième chapitre présente la situation canadienne et québécoise. Il s'efforce de mettre en relief les atouts et les opportunités que présente le développement de la bio-informatique au Québec, ainsi que les principaux obstacles à lever. Le troisième chapitre propose une stratégie d'ensemble pour organiser et structurer rapidement le démarrage de la bio-informatique au Québec.

20 14 234
54
3201 56
67898
45 29 1 23
65
444 534
8795
5622 12873

Chapitre 1

Perspectives de la bio-informatique dans le monde

1.1 Qu'est-ce que la bio-informatique?

Une science
interdisciplinaire

Le terme de bio-informatique (ou biologie computationnelle) remonte au milieu des années 1980 et désigne globalement l'ensemble des applications de l'informatique aux sciences biologiques. La bio-informatique est une science interdisciplinaire qui se situe au point de jonction des sciences de la vie et des technologies de l'information. Elle concerne l'organisation, l'entreposage, le traitement, l'analyse et la diffusion de données biologiques par ordinateur et s'applique à des domaines aussi diversifiés que la génétique et la biologie moléculaire, la neurophysiologie, l'écologie ou la simulation de processus vitaux sur ordinateur (« vie artificielle »).

La bio-informatique est pratiquée partout dans le monde, par des groupes de recherche dans les universités, les entreprises, les consortiums nationaux et internationaux. Elle est considérée comme un des fondements actuels et futurs de toute la biotechnologie.

Bio-informatique
et génomique

Depuis quelques années, c'est par sa contribution à l'essor de la génomique que l'importance de la bio-informatique a été particulièrement mise en évidence. Au cours de la dernière décennie, les progrès de la biologie moléculaire se sont faits sur deux grands fronts, celui d'une automatisation accélérée des procédés de séquençage et celui du développement des technologies de l'information nécessaires pour en compiler et analyser les résultats. La quantité phénoménale et la complexité des informations recueillies et analysées dans le cadre d'une opération d'envergure comme le Programme de séquençage du génome humain a créé une forte demande pour l'utilisation d'une assise informatique de très haute capacité. En même temps, le génome de plusieurs autres organismes (dont la drosophile et la souris) faisait également l'objet d'un séquençage systématique dans les laboratoires américains, européens et japonais, donnant lieu à la mise en place de bases de données comparatives de plus en plus importantes.

De la génomique
à la protéomique

Bien que fortement médiatisée, la grande percée scientifique annoncée par le président des États-Unis, le 26 juin 2000, constitue moins un véritable point d'arrivée qu'un point de départ. Le séquençage du génome humain, actuellement en voie d'achèvement, a constitué une vaste opération de collaboration (et de compétition) scientifique internationale dont le principal résultat sera de livrer très bientôt une description très précise et complète de notre identité génétique. Toutefois, pour rendre cette information pleinement utilisable, il reste encore à déterminer la fonction de chaque gène (et dans quelles circonstances un gène s'exprime), ainsi que les relations possibles de ses mutations avec des désordres physiologiques connus. Une meilleure compréhension des fonctions du gène viendra à son tour appuyer une autre phase de recherche, encore plus imposante et complexe, qui consistera à analyser la structure et les fonctions des protéines qui sont exprimées par les gènes.

Équipement haute puissance et réseaux

L'ampleur des données générées par de telles opérations est une des raisons qui font de la bio-informatique une des plates-formes technologiques essentielles de soutien à la recherche en génomique et en protéomique. Aussi, l'un des éléments clés de toute activité de bio-informatique est l'accès à des capacités de traitement de données de forte puissance. Ainsi, en 1998, on a pu constater que ce sont les applications biomoléculaires qui ont accaparé la plus forte proportion de temps-machine des ordinateurs de la National Science Foundation aux États-Unis. On sait également que la réussite de la célèbre compagnie Celera Genomics dans le séquençage du génome humain repose en très grande partie sur un puissant réseau d'ordinateurs interfacés avec des appareils de séquençage. Toutefois, ce genre d'équipements étant extrêmement dispendieux, la plupart des laboratoires doivent souvent rechercher des solutions alternatives, comme des réseaux de micro-ordinateurs ou de stations de travail. On fait aussi de plus en plus appel à Internet, comme lieu d'accès et d'échange de données ou de logiciels, ainsi que pour distribuer de forts volumes d'opérations de calcul et de traitement d'informations.

De grandes bases de données publiques, accessibles sur Internet

La génomique et la bio-informatique comptent parmi les domaines scientifiques qui ont le plus bénéficié de l'existence d'Internet, non seulement pour les collaborations et les échanges d'informations entre laboratoires et centres de recherche, mais également parce que plusieurs des grandes bases de données sont accessibles sur le Web: la GDB (Genome Data Base) de l'Université Johns Hopkins, la Genbank du National Center for Biotechnology Information, l'EMBL du Laboratoire européen de biologie moléculaire, SWISS-PROT de l'Université de Genève et de l'Institut européen de bio-informatique, etc.

Genbank, par exemple, constituée il y a une vingtaine d'années, a vu ses entrées s'accroître de façon exponentielle depuis la mise en marche du projet Génome humain, au début des années 1990, et comprendrait aujourd'hui plus de six millions de séquences d'ADN et sept milliards de nucléotides. On dit que sa taille double tous les sept mois. On s'attend également à ce que la Protein Data Bank (PDB) administrée par le consortium RCSB (Research Collaboratory for Structural Bioinformatics), triple de taille au cours des cinq prochaines années².

Un tsunami d'information

Le rôle fondamental de la bio-informatique est de pouvoir organiser, gérer et analyser efficacement des masses d'informations considérables. Celera Genomics, qui a annoncé avoir complété un premier « brouillon » de la séquence du génome humain en avril dernier, disait disposer de quelque 50 tera-bytes d'information, soit l'équivalent de 80 000 disques compacts. Gene Myers, vice-président de la compagnie, parle d'un « tsunami d'information »³.

Le *data mining*

Toutefois, il ne suffit pas de stocker correctement l'information dans des bases de données. Il faut pouvoir explorer les données brutes, obtenues par séquençage, et les transformer en connaissances utilisables. Pour cela, des programmes doivent être conçus et utilisés pour fouiller intelligemment ces données, les comparer et les analyser. Le *data mining*, version moderne de l'analyse de données, qui fait appel à des techniques statistiques, à l'intelligence artificielle et aux réseaux neuronaux, permet d'extraire des

² Helen Berman et al. : « The Protein Data Bank and the Challenge of Structural Genomics », *Nature : Structural Biology*, volume 7, Supplément, novembre 2000, p. 957.

³ Cité dans Ken Howard : « The Bioinformatics Gold Rush », *Scientific American*, juillet 2000 p. 58.

masses d'informations stockées dans des « entrepôts de données » (*data warehouses*), les schémas relationnels les plus pertinents, qui ne seraient pas détectables autrement. Également utilisé dans d'autres domaines qui dépendent d'une analyse rapide et optimale de l'information (finance, marketing, transports...), le *data mining* est en pleine expansion et son développement est directement lié à la puissance de calcul des ordinateurs.

Intégrer les données

Un autre défi important de la bio-informatique est l'intégration de données de sources différentes. Dans les bases de données génomiques et protéomiques, on doit « annoter » les différentes séquences répertoriées, c'est-à-dire adjoindre à chacune l'état des connaissances qui existent sur elle. À chaque séquence sont attachées diverses informations, telles que l'organisme dont elle provient, une description plus ou moins succincte de ses fonctions, ainsi que les références bibliographiques. La structure et la qualité de ces annotations sont cependant variables. Systématiser le processus d'annotation nécessite qu'on recoupe des données et des connaissances de provenances très diverses, structurées de manière différente. Une des opérations de base de la bio-informatique sera donc la recherche de similarités ou homologues entre de nouvelles séquences d'ADN et des séquences existantes. Ce type d'intégration est ce qui permet, par exemple, de déterminer quel type de protéines la nouvelle séquence est susceptible de coder et donc d'ouvrir des pistes additionnelles à la recherche.

Standardiser

Un autre grand défi de la bio-informatique consiste à assurer l'interopérabilité des différentes bases de données, c'est-à-dire la capacité de les exploiter conjointement pour répondre aux besoins des utilisateurs. Le défi est de taille et des efforts sont actuellement tentés pour standardiser les formats des bases de données génomiques et protéomiques, notamment par les grandes organisations internationales. Le projet BioStandards de l'Institut européen de bio-informatique (EBI), par exemple, financé conjointement par la Commission européenne et les nombreuses compagnies pharmaceutiques participantes, vise à normaliser la pratique de la bio-informatique dans l'industrie, en agissant à la fois sur la structure des bases de données, les outils logiciels et la formation des praticiens.

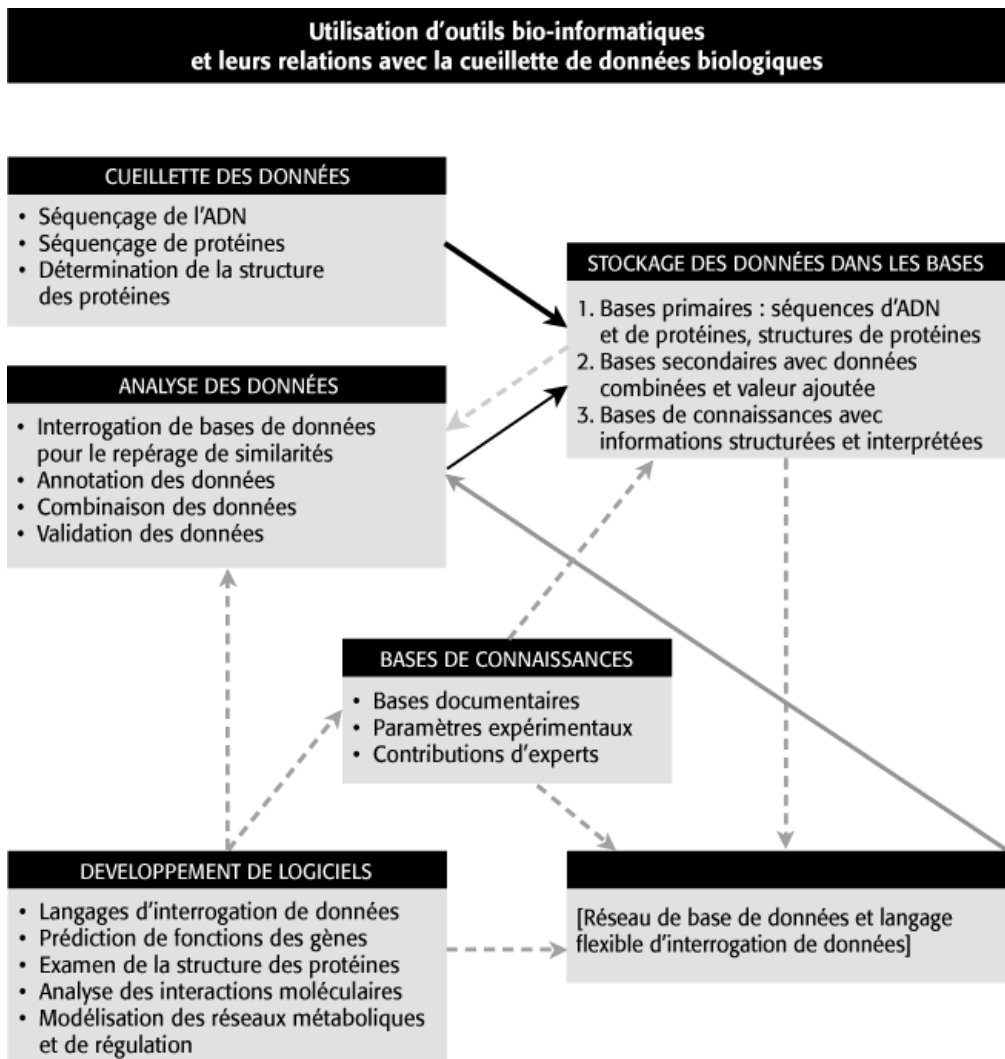
Quoiqu'en laissent entendre parfois les médias, la génomique n'est pas une discipline vouée à disparaître parce que le séquençage du génome humain arrive à terme. Le séquençage du génome d'autres espèces demeure en effet hautement pertinent, notamment pour des applications dans le domaine de l'agriculture, de la sylviculture et de l'élevage, par exemple. Mais de plus, une utilisation efficace des données du génome humain exige qu'on puisse les comparer à des génomes d'autres espèces.

Une importance appelée à s'accroître

On ne peut nier cependant que le centre d'attraction se soit déplacé depuis quelques mois de la génomique proprement dite à ce qu'on appelle la post-génomique ou la protéomique. Cette « seconde phase », axée sur l'investigation de l'expression des gènes en protéines et enzymes, ainsi que sur l'analyse fonctionnelle et structurale des protéines elles-mêmes, est appelée à accroître encore l'importance du rôle de la bio-informatique et de la demande en bio-informaticiens. Dans l'industrie, comme dans les centres de recherche, il faudra monter et exploiter des bases de données beaucoup plus complexes, mettre au point de nouveaux logiciels pour toutes les étapes de l'analyse, ainsi que mobiliser des puissances de calcul encore supérieures à celles qu'on a connues au cours de la période précédente.

L'analyse linéaire de la composition de séquences d'ADN, de même que l'identification et la localisation des gènes sur les chromosomes, constituaient déjà un grand défi à relever (et qui n'est pas terminé). Identifier comment ces gènes, isolément ou par groupes, s'expriment sous la forme de protéines et dans quelles circonstances ils le font, nécessite un travail encore plus complexe. Analyser la composition des protéines (mise en ordre des acides aminés) et surtout leur architecture (leur forme et la façon dont elles se déploient dans l'espace) représente un niveau encore supérieur de complexité.

Pourtant, c'est à cette étape lointaine que les vrais dividendes de la protéomique et de la bio-informatique se feront vraiment sentir, lorsqu'on parviendra non seulement à décrire correctement la structure des protéines, mais à en déduire les fonctionnalités.



D'après : Petra Dux et François Moille, *Bioinformatics : An Inventory and Analysis of Recent Developments in Bioinformatics and Related Areas of Research and Development*, Institute for Prospective Technological Studies, Séville, Septembre 1999, p. 12

In silico

La bio-informatique est en constante évolution. Certains n'hésitent pas à affirmer qu'elle sera au cœur de tous les grands développements en biologie au 21^e siècle et qu'elle y jouera un rôle comparable à celui des mathématiques pour les sciences physiques. D'autres parlent d'un changement de paradigme. On emploie de plus en plus l'expression *in silico* (par opposition à *in vivo* et *in vitro*) pour désigner cette toute nouvelle façon d'analyser les informations recueillies sur les processus biologiques par un recours intensif à l'ordinateur.

La bio-informatique est appelée aussi à devenir un lieu de convergence des technologies de pointe. La photonique, par exemple, est impliquée de près dans la lecture des biopuces (*biochips, microarrays*), ces matrices de sondes biochimiques qui permettent d'effectuer des analyses génétiques rapides. À long terme, on peut envisager une jonction avec les nanotechnologies et, à plus long terme encore, la construction d'ordinateurs biomoléculaires ou d'ordinateurs à base d'ADN.

1.2 Bio-informatique et industrie

Un domaine en pleine émergence

Il est extrêmement difficile d'évaluer la croissance de la bio-informatique en termes économiques. On ne doit pas s'étonner qu'il n'existe pas encore de données statistiques fiables sur le sujet. Il s'agit d'un domaine en pleine émergence, dans un secteur d'activités complètement nouveau et dont les standards industriels sont encore à établir. Par exemple, une bonne partie de la technologie en usage (un grand nombre de logiciels) se développe et circule gratuitement, via Internet en particulier, et échappe aux analyses de marché habituelles. Malgré cela, à la fin de 1999, on pouvait constater que les compagnies de bio-informatique (ou, plus justement, celles qui ont un fort volume d'activités en bio-informatique) commençaient à proliférer, mais que la plupart n'étaient pas encore inscrites en bourse. Pour l'année 1999, la performance de ces firmes dites de bio-informatique avait fortement surpassé celle de l'ensemble du secteur des biotechnologies⁴.

Comme toutes les technologies qui dérivent d'approches multiples et qui ont des applications génériques potentielles considérables, l'avenir exact de la bio-informatique est impossible à prédire. Si tous les observateurs s'entendent pour affirmer qu'elle est appelée à jouer un rôle crucial dans la période nouvelle de la postgénomique, les projections varient du simple au double. Au début de 1999, la firme Strategic Management Consulting Inc. prévoyait que le marché global de la bio-informatique allait quadrupler en cinq ans, passant de 290 millions de dollars, en 1998, à 1,2 milliard en 2005⁵. Pour sa part, SRI Consulting croit que ce marché devrait approcher du milliard en 2004⁶. Par contre, d'autres données, comme celles de la firme Oscar Gruss & Son, prévoient jusqu'à deux milliards d'ici 2005⁷.

⁴ Lulu Pickering : « BioInformatics Public Firm Stocks », *Genetic Engineering News*, vol.19 n° 20, 15 novembre 1999 p. 17.

⁵ Reproduit dans le *Genetic Engineering News*, vol.19 no 4, 15 février 1999, p. 1.

⁶ Andrew Broderick : *Bioinformatics : An Emerging Role in Life-Sciences Research*, SRI Consulting, 2000, p. 1.

⁷ Cité dans Ken Howard, *op.cit.*, p. 58.

Une étude effectuée par la Chaire de gestion des bio-industries de l'UQAM propose de découper l'industrie de la bio-informatique en cinq segments :

- 1) **Les pourvoyeurs de technologie** : systèmes d'exploitation bio-informatique et autres plates-formes technologiques (biopuces, séquençage).
- 2) **Les fournisseurs d'instruments d'analyse** : conception de logiciels, *data mining*, visualisation...
- 3) **Les fournisseurs de bases de données**, sous forme de séquences annotées.
- 4) **Les pourvoyeurs de connaissances** (*knowledge providers*), qui transforment les données brutes en connaissances utilisables par leurs clientèles.
- 5) **Les entreprises de génomique intégrée**, qui font à la fois du séquençage et de l'identification de gènes, de la génomique fonctionnelle, de la protéomique et du développement d'agents thérapeutiques.

Le quatrième segment est considéré comme particulièrement prometteur. Les entreprises fournissent de l'information structurée sur mesure pour leurs différentes clientèles, par un système d'abonnement, une entente de collaboration ou encore via un portail Internet⁸.

Une industrie de la bio-informatique?

Il est sans doute prématuré de parler des possibilités de formation d'une industrie spécifique de la bio-informatique. Un certain nombre d'entreprises ont développé et mis en marché des logiciels d'analyse de données, mais il s'agit de produits très spécialisés pour lesquels il n'existe pas nécessairement de grand marché. Les solutions bio-informatiques sont encore développées en bonne partie sur mesure, pour des besoins spécifiques, avec peu d'espoir de standardisation à court terme. Les analystes ont remarqué que les compagnies qui se créent pour offrir des services en bio-informatique ont encore quelque difficulté à se maintenir pour le long terme⁹. Les raisons peuvent être multiples. L'une d'entre elles est qu'une grande partie des logiciels qui ont été développés pour la bio-informatique ont été rendus accessibles gratuitement sur Internet. L'Institut européen de bio-informatique dresse une liste de plus de 500 de ces logiciels¹⁰.

Même si ses perspectives industrielles directes demeurent assez incertaines pour l'instant, il est incontestable que la bio-informatique est une composante essentielle du développement de plusieurs secteurs industriels qui reposent sur le développement de la génomique et de la protéomique. « La capacité de rassembler, entreposer, classer, analyser et distribuer l'information biologique générée par les projets de séquençage et d'analyse fonctionnelle est devenue si indispensable pour le développement de la biotechnologie, qu'on voit les compagnies et les organismes de financement conjuguer leurs efforts comme ils ne l'ont jamais fait pour s'aider les uns les autres et aider la communauté scientifique à mettre en place cette capacité¹¹. »

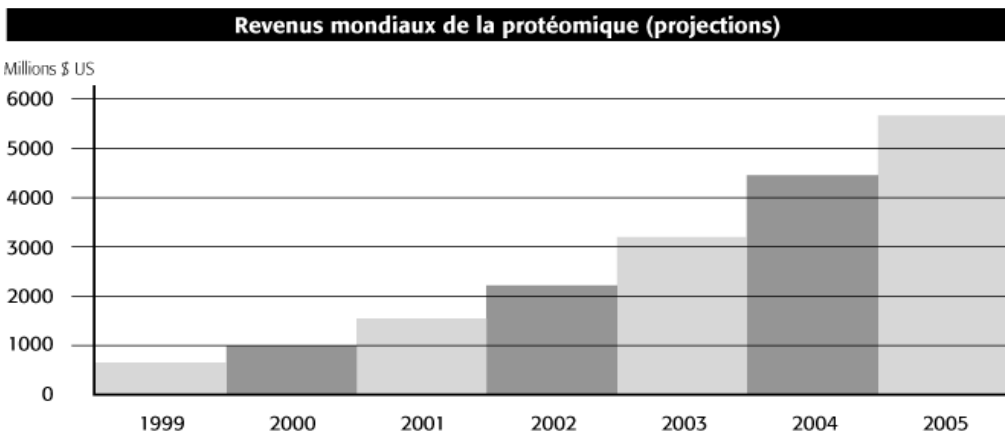
⁸ *Développement de la bio-informatique au Québec : Opportunités et priorités*, Chaire en gestion des bio-industries, École des sciences de la gestion, Université du Québec à Montréal, à paraître.

⁹ Petra Dux et François Moille : *op.cit.*, p. 29-31.

¹⁰ Cité dans « Bioinformatics », *Nature Biotechnology*, vol.18, Supplement 2000, p. IT31.

¹¹ *Ibid.*

C'est dans ce rôle de soutien, absolument fondamental, qu'il faut voir les retombées économiques de la bio-informatique. Dans les années qui viennent, l'essor de cette plate-forme paraît étroitement lié à celui de ses industries clientes, qui voudront profiter de cette ressource stratégique. La protéomique en particulier, plus proche que la génomique d'une commercialisation possible des applications, devrait connaître une croissance phénoménale. Le domaine est si vaste, les créneaux si nombreux que de très nombreuses petites entreprises se créent autour des technologies de la protéomique à travers le monde. Les grandes entreprises également : après s'être illustrée dans le séquençage du génome humain, la firme Celera Genomics vient de créer une filiale, PE Corp., dont on attend qu'elle arrive à séquencer 30 000 échantillons de protéines à l'heure.



Source : Frost & Sullivan, *World Proteomics Markets*, Report # 5822-55, octobre 2000; reproduit dans Stefan H. Unger et Justin Saeks, « Proteomics Research and Development Takes Off », *Genetic Engineering News*, vol. 20 n° 16, 15 septembre 2000, p. 17.

L'industrie pharmaceutique : une utilisatrice de toute première importance

L'industrie pharmaceutique est une de celles qui devraient bénéficier directement le plus de l'apport des découvertes en génomique et de la bio-informatique. Cette industrie a éprouvé des difficultés de croissance ces dernières années, en raison notamment des coûts de plus en plus importants du développement des médicaments. Les vingt plus puissantes sociétés pharmaceutiques ont pratiquement doublé leurs dépenses de R-D au cours des sept dernières années. Une étude de Price Waterhouse Coopers démontre que, pour maintenir un rythme de croissance raisonnable de leurs ventes (environ 7 % par année pour l'ensemble de l'industrie), les grandes compagnies pharmaceutiques devront repenser complètement leurs façons de faire en recherche et diminuer fortement les coûts¹².

Un changement de paradigme

L'avènement de la génomique et de la protéomique permettra de mieux rationaliser le processus de mise au point de nouveaux médicaments autour de cibles thérapeutiques extrêmement précises. Plusieurs observateurs n'hésitent pas à parler d'un « changement de paradigme ». La recherche et le développement peuvent être conçus de façon plus « intelligente », en s'appuyant sur une meilleure compréhension des interactions moléculaires. L'approche dite « par essais et

¹² *Pharma 2005, An Industrial Revolution in R&D*, Price Waterhouse Coopers, 1999.

erreurs » peut être remplacée par une approche prédictive, qui est en même temps plus efficace, plus sécuritaire et plus économique¹³. Grâce à ce qu'on appelle maintenant la « pharmacogénomique », soit l'étude des variations génétiques interindividuelles et de leur influence sur les réponses aux médicaments, les compagnies pharmaceutiques seront à même de développer des médicaments nouveaux, spécifiques et adaptés au profil de chaque patient, avec moins d'effets secondaires.

Le SNP Consortium

Pour l'industrie pharmaceutique, les enjeux sont considérables. Récemment, au lieu de dupliquer séparément leurs efforts, une dizaine parmi les plus grandes entreprises mondiales de ce secteur se sont alliées avec cinq grands laboratoires publics de génomique pour former un consortium international (SNP Consortium) dont l'objectif est de répertorier d'ici deux ans 300 000 SNP. Les SNP (*single nucleotide polymorphism*; se prononce « snip ») sont les sites du génome où se localisent les variations génétiques entre individus. Leur connaissance est d'une très grande importance stratégique pour accélérer la localisation des gènes associés aux pathologies et permettre le développement de médicaments adaptés au profil des patients. La base de donnée du Consortium, également financée par le Wellcome Trust, sera publique.

La bio-informatique se développe *intra-muros*

Pour se doter de compétences en bio-informatique, les entreprises du secteur pharmaceutique s'associent avec d'autres firmes, mais elles ont surtout tendance à développer cette fonction à l'intérieur de leurs propres laboratoires, plutôt que par la sous-traitance comme elles le font pour d'autres services de technologies de l'information. Il en va de même des grandes entreprises biotechnologiques, surtout lorsqu'elles ont des activités de génomique¹⁴. Cette tendance à constituer une compétence interne en bio-informatique a été observée notamment par l'Institut de prospective technologique de Séville qui l'associe en bonne partie à la difficulté d'établir un dialogue fructueux entre spécialistes des sciences de la vie et informaticiens. Ce dialogue aurait de meilleures chances de donner des résultats lorsque les parties appartiennent à la même entreprise. Mais même lorsque la bio-informatique se développe comme une compétence interne, la compréhension des besoins des utilisateurs et leur traduction en outils appropriés de traitement de l'information ne se fait pas de façon automatique¹⁵.

Autres champs d'application

À moyen terme, les applications de la bio-informatique dans l'industrie déborderont largement le secteur pharmaceutique. On peut penser au domaine de la protection de l'environnement, par exemple. En agriculture, le rôle de plus en plus important que jouent les biotechnologies dans le développement de nouveaux produits ou la lutte contre les agents pathogènes ou les parasites, par exemple, s'appuiera de façon croissante sur l'apport des connaissances en génomique et en protéomique. Comme dans le secteur biopharmaceutique, le séquençage du génome des espèces végétales appelle la constitution de banques de données d'envergure considérable, que les entreprises sont appelées à utiliser systématiquement. Ici, les frontières entre secteurs industriels sont loin d'être étanches. La mise en commun des connaissances accumulées sur les gènes et protéines des espèces animales et végétales peut servir aussi bien les entreprises

¹³ Voir Broderick, *op.cit.*, p. 9.

¹⁴ « Bioinformatics », p. IT33.

¹⁵ Dux et Moille, *op.cit.*, p. 27-28.

pharmaceutiques que celles de l'agroalimentaire. On voit des compagnies comme Incyte Pharmaceuticals développer des activités en génomique végétale, ou encore les géants Monsanto et Millennium Pharmaceuticals s'associer pour former une nouvelle entreprise de biotechnologie agricole, Cereon Genomics¹⁶.

Des investissements majeurs du côté des firmes de TIC

Signalons enfin que le développement de la bio-informatique commence à attirer les entreprises du domaine du logiciel et de l'électronique. Au Japon, Hitachi vient d'annoncer des investissements de 56 millions en R-D pour la bio-informatique. Partenaire de Celera Genomics à qui elle a fourni une bonne partie de son parc informatique pour l'opération de séquençage du génome humain, la compagnie Compaq Computers a créé un centre d'expertise en bio-informatique à Marlboro (Massachusetts). En décembre 1999, IBM annonçait la construction de Blue Gene, un ordinateur mille fois plus rapide que son célèbre Deep Blue, et dont la première tâche sera l'analyse du repliement tridimensionnel d'une protéine extrêmement complexe, formée de 324 acides aminés. IBM estime que le marché international des applications des technologies de l'information aux sciences de la vie va grimper de 3,5 à 9 milliards de dollars en 2003¹⁷.

1.3 Efforts nationaux et internationaux

Des efforts internationaux nombreux

Le développement de la bio-informatique fait l'objet d'efforts de recherche, de formation et de diffusion dans un grand nombre de pays. Il existe quelques organismes internationaux comme The International Society for Computational Biology, APBioNet — pour les pays du Pacifique, dont le Canada — et l'Institut européen de bio-informatique. Une recherche sur Internet révèle la présence d'organisations en bio-informatique dans plusieurs pays d'Europe, particulièrement l'Allemagne et la France, au Japon, en Australie et surtout aux États-Unis. On trouve des centres de bio-informatique dans des pays comme l'Inde, la Corée, la Malaisie, Singapour, Israël, etc.

Aux États-Unis

Aux États-Unis, les principaux intervenants sont des entreprises privées, mais les grandes universités américaines ont leurs propres centres de bio-informatique, liés à des groupes de recherche travaillant sur le séquençage de génomes. Certains de ces groupes ne font qu'utiliser des outils existants, d'autres développent des logiciels, des méthodes ou des services.

Les agences gouvernementales américaines sont présentes également. Le département de l'Énergie de même que le département de la Défense financent des projets en bio-informatique. Cependant, l'acteur le plus important dans le domaine est le réseau des National Institutes of Health (NIH) qui a patronné le projet Génome humain. Depuis 1988, une des constituantes de ce réseau, le National Centre for Biotechnology Information, qui est affiliée à la National Library of Medicine, constitue un des centres de références les plus marquants dans le domaine. Le NCBI développe et gère des bases de données publiques, met au point des outils d'interrogation et d'analyse, effectue de la recherche et diffuse l'information dans les milieux biomédicaux.

¹⁶ *Ibid.*, p. 13.

¹⁷ Anita Flanagan, « Racing to conquer Biodata With Computers », *Genetic Engineering News*, vol. 20 n° 17, 1^{er} octobre 2000, p. 8.

Les NIH ont lancé récemment la Protein Structure Initiative qui prévoit investir 125 millions de dollars sur cinq ans. Une des constituantes des NIH, le National Institute of General Medical Sciences, s'est donné pour objectif de déterminer la structure et le déploiement de 10 000 protéines en 10 ans et vient de consacrer 30 millions de dollars à une première ronde de recherche¹⁸. Une partie importante du financement en bio-informatique provient également de fondations privées, comme le Howard Hughes Medical Institute.

En Europe

En Europe, la principale référence en bio-informatique est l'Institut de bio-informatique européen, une constituante de l'European Molecular Biology Laboratory (EMBL) qui regroupe des organisations de quinze pays. L'Institut lui-même, logé sur le Wellcome Trust Genome Campus, à Hinxton (Angleterre), est un des centres de ressources les plus importants dans le domaine de la génomique, de la protéomique et des structures macromoléculaires. Sa mission est essentiellement de rendre accessibles les connaissances disponibles en génomique et en protéomique aux groupes de recherche et aux industries. Il effectue aussi lui-même de la recherche en bio-informatique. Son financement provient à 40 % de la Communauté européenne, à 30 % de l'EMBL et à 30 % de l'industrie.

Plusieurs centres de bio-informatique sont reliés par le réseau European Molecular Biology Network (EMBnet). Son noyau initial exclusivement européen s'est élargi avec le temps, et relie aujourd'hui des pays comme l'Argentine, le Canada, la Russie, la Chine, l'Inde et Israël. Le financement est partagé entre la Communauté européenne et les différents pays membres¹⁸.

INFOBIOGEN

Le centre de ressources INFOBIOGEN a pour principales missions :

- 1 de développer et d'exploiter des services bio-informatiques d'intérêt général, entre autres : de permettre à l'ensemble de la communauté scientifique française un accès aux bases de données, en particulier en développant la connexion des utilisateurs aux réseaux nationaux et internationaux et en facilitant l'accès à des données biologiques actualisées (accès aux séquences quotidiennes de l'EMBL et des autres grandes bases de données internationales), de mettre à disposition les environnements d'analyses et d'interrogations nécessaires à l'exploitation des banques de données disponibles;
- 2 d'exercer des actions de formation et d'assistance aux utilisateurs des ressources du Centre;
- 3 de coordonner le réseau informatique national de recherche et de service pour la Génomique; de participer à la concertation et à la coordination nationale et internationale en matière de bio-informatique;
- 4 de favoriser la valorisation des logiciels et des banques de données produites en France;
- 5 de développer et d'exploiter au niveau national les codes de calcul et les bases de données utilisés par la recherche Génomique;
- 6 de mener une activité de recherche et de développement dans le champ de l'informatique appliquée à la Génomique; notamment dans le domaine des logiciels d'analyse intensive des génomes, des modalités d'interconnexion et de diffusion des grandes bases de données, et des environnements favorisant la recherche coopérative à travers les réseaux de communication.

Source : <http://www.infobiogen.fr/presentation/infobiogen.html>

¹⁸ John C. Norvell et Alisa Zapp Malachuk, « Structural Genomics Programs at the US National Institute of General Medical Sciences », *Nature Structural Biology*, volume 7, Supplément, novembre 2000, p. 931.

En France

Certains pays ont lancé des actions de grande envergure en génomique et post-génomique. En France, par exemple, le ministère de l'Éducation nationale, de la Recherche et de la Technologie créait en 1999 un centre national de ressources en bio-informatique appelé INFOBIOGEN, qui remplace un consortium du même nom qui existait depuis 1995. Le centre a pour vocation d'accélérer la recherche coopérative en génomique en offrant des outils informatiques et en développant de nouveaux services accessibles à l'ensemble de la communauté scientifique.

1.4 Une pénurie universelle de ressources humaines

Une demande de ressources humaines très supérieure à l'offre

Le domaine de la bio-informatique est encore relativement nouveau et la demande mondiale de spécialistes devient rapidement considérable. Au fur et à mesure que se développent les bases de données en génomique et en protéomique, et que les possibilités d'exploitation industrielle des informations contenues dans ces bases s'accroissent, cette demande est appelée à s'amplifier. Déjà, la situation actuelle en est une de pénurie. En Europe, malgré la présence de plusieurs centres renommés dans le domaine et de nombreuses industries clientes de la bio-informatique, le manque de ressources humaines constitue le problème principal qu'ont identifié les auteurs de l'étude publiée par l'Institut de prospective technologique de Séville. Ce problème, affirme-t-on, est assez grave pour compromettre la compétitivité européenne en bio-informatique, par rapport aux États-Unis²⁰.

Mais les États-Unis ont eux aussi leurs problèmes de pénurie de personnel en bio-informatique. C'est le cas particulièrement en milieu universitaire où la demande industrielle en diplômés est si forte qu'il reste extrêmement difficile d'assurer une relève pour la recherche fondamentale et la formation. La situation est jugée suffisamment critique pour appeler une intervention spécifique des NIH et de fondations privées comme le Howard Hughes Medical Institute. Les NIH s'approprient à financer vingt centres nationaux de formation en bio-informatique. Une autre organisation, l'Alfred P. Sloan Foundation, a récemment lancé un appel d'offres pour créer des programmes de maîtrise de deux ans en bio-informatique qui auraient pour effet de répondre plus adéquatement aux besoins de l'industrie et de garder plus de Ph.D. à l'université.

La question de la formation

La bio-informatique fait appel à une combinaison relativement rare de compétences pour lesquelles il n'existait pas *a priori* de programmes de formation identifiés dans les grandes disciplines que sont la biologie moléculaire et l'informatique. Même aux États-Unis, malgré une croissance exponentielle de la demande, il semble que les universités américaines aient été lentes à réagir et à offrir des programmes de formation ou de spécialisation adéquates en bio-informatique²¹. C'est surtout depuis 1997 qu'on a commencé à donner des cours en bio-informatique dans plusieurs universités, tant aux États-Unis qu'ailleurs dans le monde. Ceci signifie aussi qu'une très forte partie des développements de la bio-informatique s'est faite par des gens qui se sont un peu formés « sur le tas »,

¹⁹ Dux et Moille, *op.cit.*, p. 21. Voir aussi Helen Gavaghan, « Europe Seeks Solution to Bioinformatics Shortfall », *Nature*, 404, 6 avril 2000, p. 688.

²⁰ *Ibid.* p. 6

²¹ Voir Paula E. Stephan et Grant Black, « Bioinformatics : Does the US System Lead to Missed Opportunities in Emerging Fields? A Case Study », *Science and Public Policy*, décembre 1999, p. 382-393.

soit des chercheurs en sciences de la vie qui se sont donné une formation additionnelle en informatique et dans les technologies de l'information, soit des informaticiens qui ont appris à travailler dans le contexte particulier des laboratoires de biotechnologie, de génomique et de biologie moléculaire.

À l'heure actuelle, des pays comme les États-Unis et l'Angleterre ont intégré la bio-informatique dans la formation universitaire, offrant des programmes mixtes de biologie et d'informatique qui permettent aux étudiants d'appivoiser les deux mondes. Selon l'étude de l'Institut de prospective technologique de Séville sur l'état de la bio-informatique en Europe, ce serait loin d'être le cas ailleurs dans les pays européens, où ce sont encore les compartimentations traditionnelles qui dominant dans la formation des spécialistes en sciences de la vie et en technologies de l'information. Les auteurs de l'étude attribuent à ce retard dans les programmes universitaires la pénurie importante de ressources professionnelles en bio-informatique qu'ils ont observée dans l'industrie européenne²².

La tendance universelle est d'apporter des solutions à ce problème. De plus en plus d'universités à travers le monde offrent des cours et des diplômes en bio-informatique. L'International Society for Computational Biology en tient la liste sur son site Web. On peut remarquer aussi qu'il se donne de plus en plus de formations (ou de perfectionnements) en ligne sur Internet.

La situation de la bio-informatique peut se résumer ainsi : une demande de compétences très forte et qui est appelée à s'accroître encore de façon phénoménale partout dans le monde, mais une offre de formation encore insuffisante qu'il faut de toute urgence organiser et financer.

²² Dúx et Moille, *op.cit.*, p.22.

20 14 234
54
3201 56
67898
45 29 1 23
65
444 534
8795
5622 12873

Chapitre 2

Perspectives de la bio-informatique au Canada et au Québec

1) Situation au Canada

Le Canada n'est pas un joueur de premier rang dans le domaine

Les listes existantes d'organismes, centres et groupes de recherche internationaux sur le Web ne mentionnent que très peu de références canadiennes en bio-informatique ou même en génomique. Toutefois, sans avoir été parmi les premiers joueurs dans la révolution génétique actuelle, le Canada et le Québec possèdent des atouts qui, moyennant les ressources et les stratégies appropriées, peuvent lui permettre d'en tirer profit sur le plan économique. D'une part, la demande internationale massive en bio-informatique va se faire sentir également de plus en plus dans l'industrie biopharmaceutique et biotechnologique canadienne et québécoise. D'autre part, les perspectives d'applications de la génomique-protéomique sont si vastes et si diversifiées que des joueurs secondaires ou entrés tardivement en course comme le Canada et le Québec ont encore la chance de se positionner avantageusement.

Au Canada comme au Québec, il est difficile de retracer l'activité bio-informatique en dehors du domaine de la génomique, puisque les deux sont extrêmement liés. Ainsi, une étude bibliométrique effectuée pour Genome Canada par l'Observatoire des sciences et des technologies montre que, de 1995 à 1997, le Québec (40 %) et l'Ontario (39 %) sont en tête de la recherche canadienne en génomique ou associée à la génomique. Cependant, seulement 0,7 % des articles canadiens dans ce domaine appartiendraient clairement au secteur de la bio-informatique.

Réseau du CNRC, l'IRB

Le réseau du Conseil national de recherches du Canada (CNRC) possède plusieurs centres d'expertise en bio-informatique. Il existe un groupe spécialisé dans le traitement informatique des données biologiques à l'Institut de biodiagnostic de Winnipeg. L'Institut des biosciences marines à Halifax a la responsabilité d'exploiter et d'administrer le Réseau de bio-informatique canadien (RBC), une installation nationale donnant accès à des bases de données biotechnologiques et à des logiciels de bio-informatique. Le RBC fournit également des services aux entreprises de biotechnologie et aux équipes de recherche universitaires.

Au Québec, l'Institut de recherche en biotechnologie (IRB) du CNRC a des activités en génomique et en bio-informatique dans son Département de génétique. On s'y intéresse notamment au développement de biopuces. Avec ses quelque 500 chercheurs et son budget de fonctionnement annuel d'environ 20 millions de dollars, l'IRB est considéré comme un des plus importants centres de recherche en biotechnologies en Amérique du Nord. Une forte partie de ses activités s'effectue de façon conjointe avec l'industrie ou pour le compte de l'industrie.

Réseaux de centres d'excellence

Le Réseau de centres d'excellence le plus important au Canada en bio-informatique est le Réseau canadien des maladies génétiques qui fait de la recherche de pointe dans le secteur de la santé humaine. Avec le Conseil des ressources humaines en biotechnologie, un organisme privé à but non lucratif qui fait la promotion de la qualité de la main-d'œuvre dans l'industrie des bio-

technologies, le Réseau canadien des maladies génétiques organise des ateliers de formation en bio-informatique dans toutes les régions du pays. L'objectif, défini dès 1998, est de former en quelques années au moins 300 experts et 1500 utilisateurs avancés.

Un autre Réseau de centres d'excellence porte sur le génie protéique (PENGE). Sa tête de réseau est à l'Université d'Alberta, mais il compte douze chercheurs universitaires au Québec, douze en Ontario et treize en Colombie-Britannique. Notons également qu'au début de 2000, le gouvernement fédéral a lancé un appel de propositions pour la création d'un nouveau Réseau de centres d'excellence portant sur « les technologies génomiques et la société ».

Autres centres d'expertise canadiens en bio-informatique

Parmi les principaux centres canadiens d'expertise en bio-informatique, il faut également mentionner :

- 1) Le Bioinformatics Supercomputing Centre (BiSC) de l'Hospital for Sick Children de Toronto, qui assure la gestion et une partie du développement du Genome Data Base, la base de données où sont rassemblées les données du projet du Génome humain et qui est située à l'Université John Hopkins de Baltimore.
- 2) Le Samuel Lunenfeld Research Institute (SCRI) du Mount Sinai Hospital de Toronto, fondé en 1985, un prestigieux centre de recherche qui a fait œuvre de pionnier au Canada en génomique, protéomique et bio-informatique.
- 3) Le Centre for Molecular Medicine and Therapeutics (CMMT) de Vancouver, affilié à l'Université de Colombie-Britannique, qui a développé des activités de bio-informatique en lien avec la recherche sur les maladies génétiques.
- 4) Le Centre des génomiques intégrées (CGI) de Vancouver, fondé en 1999, qui constitue un réseau de centres créé par l'Université de Colombie-Britannique et le Cancer Agency of BC; ses activités de bio-informatique sont liées à l'application de la génomique à la recherche en oncologie.

Génome Canada

Jusqu'à la formation récente de l'organisme privé à but non lucratif Génome Canada, le Canada était un des rares pays développés à ne pas posséder de programme national en génomique. Doté d'une subvention fédérale de départ de 160 millions de dollars sur cinq ans, Génome Canada entend assurer une expansion importante au domaine, tant dans le secteur universitaire que dans les entreprises, en réunissant des investissements de l'ordre de 800 millions de dollars. L'organisme créera un réseau de cinq centres canadiens de génomique dans les différentes régions canadiennes. Ces centres devraient à leur tour permettre le regroupement et la coordination des ressources en place²³.

Les Instituts de recherche en santé du Canada

Une autre occasion de développer la bio-informatique au Canada vient de la restructuration des programmes fédéraux de financement de la recherche en santé. Les nouveaux Instituts de recherche en santé du Canada, plus ou moins

²³ Génome Canada, *Plan d'affaires*, 29 novembre 1999.

inspirés des National Institutes of Health des États-Unis, se proposent de mettre sur pied treize instituts spécialisés à travers le Canada, dont un en génétique et génomique appliquées à la santé. Le budget total des Instituts pour 2000-2001 (incluant les subventions à la recherche) est de 340 millions de dollars.

Peu d'entreprises,
souvent de petite taille

Une recension des entreprises de bio-informatique canadiennes effectuée par la Chaire en gestion des bio-industries de l'UQAM montre un secteur d'activités encore très émergent. On ne compterait en effet que quatorze entreprises de bio-informatique proprement dite au Canada, la plupart de petite taille et d'âge très récent. Dix d'entre elles sont localisées en Ontario. Bon nombre de ces entreprises seraient nées par essaimage de centres ou de réseaux de centres universitaires. Leurs activités se concentrent surtout dans la catégorie des « fournisseurs d'instruments d'analyse » (voir nomenclature dans section 1.2). Si le Québec a encore peu d'entreprises de bio-informatique proprement dite, il concentre 40 % des entreprises de génomique fonctionnelle canadiennes, qui sont des grandes clientes de la bio-informatique. Toutefois, l'ensemble de ce marché intérieur canadien est extrêmement restreint et il constitue en bonne partie une extension tardive et encore modeste de l'expansion que connaît la « grappe » génomique-protéomique-bio-informatique en Amérique du Nord. Il est clair que, pour la très grande majorité des entreprises canadiennes, la compétition se trouve à l'extérieur du pays²⁴.

Avantages compétitifs

Des avantages compétitifs se dessinent néanmoins pour les entreprises canadiennes. D'abord, le potentiel de développement du domaine général des applications de la génomique-protéomique est lui-même si vaste et si diversifié qu'il y aura encore pendant plusieurs années amplement de place pour des entreprises nouvelles et spécialisées qui sauront se positionner par rapport aux besoins émergents. Ensuite, l'origine universitaire de ces entreprises et leur proximité des centres d'expertise de haut niveau leur confèrent un atout important, puisqu'elles sont dans un domaine industriel où c'est la recherche scientifique de pointe qui sert de principal vecteur de développement. Enfin, une des tendances fortes du marché mondial de la bio-informatique est l'essor important que prend de plus en plus le segment des « pourvoyeurs de connaissances » (*knowledge providers*), notamment celui des services spécialisés d'analyse et de mise en forme de l'information brute pour les besoins spécifiques de clientèles (identification de cibles thérapeutiques et autres). Il s'agit d'un segment où l'expertise canadienne peut se faire valoir.

Des enjeux importants
pour les bio-industries
canadiennes

Toutefois, au Canada comme ailleurs, un des enjeux majeurs du développement du domaine demeure la question des ressources humaines. BIOTECanada, l'organisation privée de promotion des biotechnologies au Canada a récemment soumis à Industrie Canada un mémoire sur le développement de la bio-informatique²⁵. Dans son survol de la situation canadienne, l'organisme identifie comme principal point fort l'excellence reconnue mondialement des quelques chercheurs et équipes scientifiques qui travaillent déjà dans le domaine de la bio-informatique, que ce soit dans les universités et les hôpitaux, les entreprises ou les centres de recherche

²⁴ Développement de la bio-informatique au Québec : Opportunités et priorités.

²⁵ BIOTECanada, *From Information to Knowledge : Canada's Capacity in Bioinformatics*, 31 mars 2000.

gouvernementaux. Le principal point faible, cependant, reste le manque flagrant de programmes de formation en bio-informatique, qui crée à son tour une pénurie de ressources humaines qualifiées pour l'industrie.

Ce diagnostic ressemble à celui qu'ont fait bien d'autres pays, mais l'enjeu pour le Canada est peut-être encore plus sérieux qu'ailleurs. Car si une offre de formation n'est pas rapidement organisée pour répondre à la demande, les bio-industries canadiennes — dont une forte partie est concentrée au Québec — devront aller à l'étranger pour trouver réponse à leurs besoins. La bio-informatique constitue une ressource stratégique et les compagnies actuelles, aussi bien que celles qui démarreront, pourraient être tentées de s'implanter là où elles seront en mesure de se la procurer. Un exode des cerveaux est aussi à craindre. En l'absence de structures au Canada et au Québec, les meilleurs spécialistes iront s'installer là où existent et se développent des centres de bio-informatique. À moyen terme, on risque ainsi de compromettre l'essor des entreprises qui reposent sur l'exploitation de la génomique et de la protéomique, et d'affaiblir l'ensemble des bio-industries canadiennes et québécoises.

2.2 Situation au Québec

Au Québec, une industrie biotech dynamique, mais encore fragile

Depuis vingt ans, le Québec a vu se développer un secteur de biotechnologies très dynamique. Selon une étude faite pour le compte de BIOTECCanada, en 1998, 42 % des firmes de biotechnologie au Canada se trouvaient au Québec, contre 24 % en Ontario et 18 % en Colombie-Britannique. L'Ontario devancerait cependant le Québec par le volume de R-D et par la proportion d'entreprises inscrites en bourse²⁶. Ces données sont un peu différentes de celles que publie Statistique Canada, sur la base d'une enquête de 1997.

Répartition des entreprises de biotechnologie au Canada, par province (1997)

	Nombre d'entreprises	%
Québec	86	30
Ontario	71	25
Colombie-Britannique	56	20
Alberta	19	7
Saskatchewan	21	7
Nouvelle-Écosse	9	3
Manitoba	8	3
Î.-P.-É., T.-N., N.-B.	11	4
Total	282	

Source : Statistique Canada, *Enquête sur les entreprises de biotechnologie - 1997*.

Le secteur des biotechnologies demeure encore loin de la maturité, dans la mesure où très peu d'entreprises existantes sont parvenues à l'étape de la commercialisation de leurs produits. Une étude récente de Jorge Niosi souligne la grande difficulté qu'ont ces firmes d'accéder au capital de risque (les deux tiers n'auraient pas de brevets sur leurs inventions)²⁷.

²⁶ BIOTECCanada, *op. cit.*, p. 8.

²⁷ Jorge Niosi, *Comment expliquer la croissance rapide parmi les entreprises canadiennes de biotechnologie?*, Document de recherche n° 8, Statistique Canada, février 2000.

Quelques entreprises québécoises ayant des activités de bio-informatiques²⁸

Chemical Computing Group (CCG)	Montréal (1994).	Développe et commercialise des logiciels scientifiques de pointe pour le criblage et l'analyse de données à haut débit dans les sciences de la vie et les matériaux.
DNA Landmarks Inc.	Saint-Jean-sur-Richelieu (1995).	Compagnie affiliée à Svalöf Weibull (Suède). Se spécialise dans le développement d'outils de cartographie génétique détaillée à base de marqueurs d'ADN pour l'agronomie.
Écopia Biosciences	Montréal (1998).	Identification de gènes microbiens ayant un potentiel pour l'industrie pharmaceutique. Applique des techniques de haut débit pour le séquençage et l'analyse bio-informatique.
Galileo Genomics Inc.	Montréal (1999).	Identification de cibles thérapeutiques pour le pharmaceutique. Utilise le concept de population fondatrice comme modèle génétique (échantillonnage dans la population québécoise).
Geneka Biotechnology Inc.	Montréal (1995).	Compagnie spécialisée dans les facteurs de transcription et leur rôle dans la régulation génétique. Gère des bases de données publiques et commerciales dans le domaine et a mis au point une plate-forme technologique de criblage à haut débit de molécules qui agissent sur la régulation des gènes et ont un potentiel thérapeutique.
GenomicsOne Corporation	Laval (1995).	Développe des plates-formes technologiques pour la manipulation et l'identification des gènes ou pour le développement de produits de type thérapeutique ou de diagnostic. Offre aussi des services de construction de bases de données génomiques.
Procréa Biosciences Inc.	Montréal (1990).	À la fois centre de recherche, de services cliniques et d'analyse diagnostique en génétique. Utilise notamment une technologie développée à l'Université de Montréal et offrant une toute nouvelle approche pour l'identification rapide de cibles géniques à fin thérapeutique.
SignalGene Inc.	Montréal (1991).	Compagnie de génomique appliquée et de pharmacogénomique. Développe des outils de dépistage génétique comportant des méthodes de génotypage à débit ultrarapide ainsi que des analyses statistiques et bio-informatiques.

²⁸ Liste non exhaustive. Ne comprend pas notamment les grandes compagnies pharmaceutiques.

L'Association québécoise des bio-industries compte actuellement environ 150 membres. Pourtant, il est difficile de repérer dans le monde de la recherche ou dans celui de l'industrie des activités spécifiquement en bio-informatique. Chose certaine, l'avenir des biotechnologies, au Québec comme ailleurs, passe en très grande partie par l'exploitation des filières génomique et protéomique et repose, par conséquent, sur des assises bio-informatiques fortes. Récemment, le directeur général du Centre québécois d'innovation en biotechnologie (CQIB), le premier incubateur d'affaires québécois spécialisé dans les biotechnologies, faisait valoir la nécessité de développer des entreprises spécialisées dans ce secteur²⁹.

L'industrie pharmaceutique aura de plus en plus besoin de la bio-informatique

Pour sa part, l'industrie biopharmaceutique constitue un des secteurs industriels de pointe les plus importants au Québec. Bénéficiant d'une assise de recherche biomédicale de haute qualité, de la présence de grands joueurs comme BioChem Pharma et l'Institut de recherche en biotechnologies du CNRC, ainsi que d'un régime fiscal éminemment favorable à la R-D, cette industrie — principalement concentrée dans la région de Montréal — représente presque la moitié de toute l'industrie pharmaceutique canadienne, avec 163 entreprises, 2,3 milliards de dollars d'expédition et 610 millions de dollars de R-D (chiffres de 1997). Plusieurs des plus grandes multinationales pharmaceutiques ont des établissements de R-D au Québec.

Les perspectives de développement de l'industrie pharmaceutique mondiale se retrouvent également au Québec. La nécessité d'un changement de paradigme dans la conception de la R-D est un impératif qui concerne directement l'industrie québécoise. Au fur et à mesure que s'effectuera le passage vers une approche de développement de produits fondée sur l'exploitation des bases de données génomiques et protéomiques, les entreprises pharmaceutiques québécoises et les filiales d'entreprises étrangères devront de plus en plus compter sur l'existence de ressources bio-informatiques.

Une industrie des TIC en croissance

Il faut se rappeler en outre qu'une des industries en croissance du Québec est celle des technologies de l'information. Un profil de ce secteur, publié par le MICST en juin 1998, faisait état de 3500 entreprises, dont plus de 80 % œuvraient dans le logiciel et les services informatiques, 10 % dans les services de télécommunication et 10 % dans les équipements et composants électroniques. En 1997, on estimait à 20 milliards les recettes et à 80 000 le nombre d'emplois dans l'ensemble de l'industrie. Les dépenses de R-D étaient de 600 millions en 1995.

Ces entreprises ne semblent pas avoir jusqu'ici porté beaucoup d'attention aux besoins informatiques des bio-industries, mais elles pourraient être de plus en plus sollicitées au fur et à mesure que s'exprimera la demande. Déjà, le Centre de recherche informatique de Montréal (CRIM), qui n'avait pas développé de projets dans ce secteur, vient de participer au financement d'une étude sur les opportunités et priorités de la bio-informatique au Québec. L'étude elle-même a été conduite par la Chaire en gestion des bio-industries de l'UQAM.

Des centres universitaires

Avec 38 % des dépenses (contre 25 % pour les sciences pures et 17 % pour les sciences appliquées), la recherche biomédicale est de longue date le secteur le

²⁹ *Les Affaires*, 11 novembre 2000, p. A9.

plus florissant de la recherche universitaire québécoise. C'est surtout dans ce milieu que s'est développée une expertise québécoise en génomique, en protéomique et en bio-informatique. Plusieurs centres de recherche universitaires travaillent par contre dans ces domaines, comme le Centre de recherche sur la fonction, la structure et l'ingénierie des protéines (Laval) ou le Groupe de recherche en modélisation biomédicale (Montréal).

Le Centre de génomique de Montréal, basé à l'Institut de recherche du Centre universitaire de santé McGill se compose de plusieurs équipes de recherche, notamment en génomique fonctionnelle. Il s'intéresse particulièrement à l'utilisation de l'informatique pour l'exploration des processus de contrôle génétique. Il exerce un leadership important dans les systèmes robotisés et les biopuces (*microarrays*).

Le Québec dispose depuis trente ans d'une forte expertise en génétique des populations ou démogénétique, qui s'est développée d'abord dans le cadre du Réseau de médecine génétique et qui regroupe maintenant un grand nombre de chercheurs autour de l'Institut interuniversitaire de recherches sur les populations (IREP). Cette spécialité joue un rôle stratégique extrêmement important dans la recherche génomique et protéomique, puisque les bases de données sur les pools génétiques des populations peuvent être exploitées dans la recherche de profils génétiques à risque et de cibles thérapeutiques pour le développement de nouveaux médicaments. Certains pays comme l'Islande et l'Estonie ont d'ailleurs capitalisé sur la relative homogénéité génétique de leurs populations pour se doter d'une industrie pharmacogénomique.

Calcul de
haute puissance

Le Québec a également investi dans le domaine de la modélisation et de l'utilisation d'ordinateurs à haute capacité, avec le Centre de recherche en calcul appliqué (CERCA), un centre de liaison et de transfert qui fournit notamment ses services à l'industrie pharmaceutique. Le CERCA constitue la tête de pont du Réseau québécois de calcul de haute performance (RQCHP) et s'est aussi associé à d'autres centres de recherche pour former le Réseau de calcul et de modélisation mathématiques (RCM²) qui fait partie du RQCHP. Soulignons enfin qu'une quarantaine de chercheurs québécois provenant de huit universités et de seize entreprises sont associés à un des réseaux de centres d'excellence canadiens qui œuvrent dans le domaine de la modélisation, « Mathématiques de la technologie de l'information et des systèmes complexes » (MaTISC).

Pour sa part, le Centre de calcul scientifique et de bio-informatique de l'Université Laval soutient une centaine de projets subventionnés dans différentes disciplines en leur fournissant la puissance de calcul et de traitement de données ainsi que le soutien technique qui leur sont nécessaires. Son champ d'expertise déborde cependant largement celui de la bio-informatique, puisqu'il s'affirme également dans des domaines comme la chimie des matériaux et la géomatique.

Génome Québec

Annoncé dans le budget 2000-2001 du gouvernement du Québec, le projet de centre québécois de génomique a fait l'objet d'un financement initial de 10 millions de dollars. Génome Québec est une société à but non lucratif chargée de promouvoir et de soutenir des activités structurantes et mobilisatrices dans des programmes de recherche de pointe en génomique et dans les disciplines

connexes, et d'améliorer ainsi la compétitivité et la croissance économique du Québec. Génome Québec est de création toute récente, mais il a déjà lancé un premier appel de propositions pour des projets dans deux grands volets :

- 1) **Grands projets et plates-formes technologiques** : un nombre restreint de projets de recherche de grande envergure et d'infrastructures de pointe en génomique et en protéomique. Ces projets doivent pouvoir augmenter l'impact du Québec dans ces disciplines, aux niveaux canadien et international. Ils doivent aussi avoir la capacité d'appuyer le développement de l'industrie biotechnologique québécoise.
- 2) **Centres d'expertise** : soutien à un nombre limité de centres d'expertise fournissant des services de grand débit (noyaux) ou d'une envergure plus ciblée (foyers). Les centres d'expertise doivent obligatoirement chapeauter une ou plusieurs demandes d'infrastructure et/ou de grands projets définis au volet 1.

Le FRSQ a reçu le mandat d'évaluer, au nom des trois fonds subventionnaires du Québec (FRSQ, FCAR et CQRS) et en collaboration avec eux, les demandes de projets structurants en génomique et/ou protéomique soumises à Génome Québec. Ce n'est pas uniquement le secteur biomédical qui est couvert, car des projets peuvent également être soumis dans d'autres secteurs d'application comme l'agroalimentaire et l'environnement. De plus, Génome Québec s'intéresse aux aspects éthiques, légaux et sociaux de la génomique.

2.3 Enjeux et atouts du Québec en bio-informatique

Enjeux et opportunités

Les grands enjeux du développement de la bio-informatique au Québec sont les suivants.

- 1) **Assurer la croissance de l'industrie québécoise des biotechnologies.** Comme tout indique que l'expansion des biotechnologies reposera de plus en plus sur la maîtrise des outils et des connaissances de la bio-informatique, les perspectives de développement de l'industrie québécoise des biotechnologies dépendront pour une bonne part de la disponibilité au Québec des ressources dont elle aura besoin dans ce domaine.
- 2) **Maintenir la vigueur d'une R-D pharmaceutique de pointe au Québec.** Les entreprises pharmaceutiques mondiales ont surtout implanté des établissements de R-D au Québec. On sait qu'elles doivent réorienter leurs stratégies de recherche pour exploiter le potentiel de la génomique/protéomique. Le Québec a donc tout intérêt à développer une expertise en bio-informatique pour répondre à cette demande.
- 3) **Offrir aux entreprises québécoises des TIC des possibilités d'expansion intéressantes.** Une tendance récente est l'entrée massive des grands joueurs des TIC dans la bio-informatique (IBM, Compaq, Hitachi...). La demande mondiale est appelée à croître rapidement. Si on parvient à développer celle du marché québécois, l'industrie des TIC québécoise devrait y trouver une perspective d'expansion très intéressante.

- 4) **Accélérer l'intégration de la génomique/protéomique dans d'autres bio-industries québécoises.** Même si c'est surtout autour du bio-pharmaceutique que se développe mondialement la bio-informatique, les effets de débordement se font déjà sentir dans l'agroalimentaire, l'industrie de l'environnement, la sylviculture, etc. D'ici quelques années, de plus en plus de bio-industries québécoises devront intégrer et exploiter à leur tour les connaissances en génomique et en protéomique, et donc devenir assez rapidement des utilisatrices de bio-informatique.

Le Québec peut miser sur ses atouts

Le Québec s'est affirmé au cours des années comme un lieu très compétitif en matière de R-D, notamment dans les technologies de pointe. L'argument cité le plus souvent (37 %) par les répondants de l'enquête effectuée par la Chaire en gestion des bio-industries de l'UQAM pour définir les avantages concurrentiels du Québec est le soutien financier des gouvernements, particulièrement en matière de mesures fiscales.

En plus d'un régime fiscal avantageux pour les compagnies de haute technologie, le Québec jouit d'une grande réputation pour la qualité de sa matière grise. Si on fait exception de la question particulière de la pénurie de bio-informaticiens, la disponibilité d'un personnel scientifique de haut niveau est donnée comme raison principale de leur implantation au Québec (à Montréal) par de nombreux responsables d'entreprises des bio-industries. L'expertise du milieu académique et son excellence scientifique viennent en deuxième place (30 %) comme avantage concurrentiel du Québec, selon l'enquête de la Chaire en gestion des bio-industries³⁰.

En fait, si on fait le bilan des atouts, on s'aperçoit que le Québec offre de nombreuses conditions favorables pour construire une compétence en bio-informatique.

- 1) **Une qualité de recherche scientifique reconnue mondialement** dans le secteur biomédical et plus particulièrement dans des spécialités porteuses de la bio-informatique, comme le pharmaceutique.
- 2) **Une main-d'œuvre scientifique et technique de grande compétence** dans les sciences de la vie et les bio-industries.
- 3) **La présence au Québec d'industries biopharmaceutiques et biotechnologiques** dynamiques qui devront compter de plus en plus sur la fonction bio-informatique pour assurer leur développement.
- 4) **Une localisation géographique** qui donne un accès rapide aux principaux centres de recherche universitaires et industriels, ainsi qu'aux marchés d'Amérique du Nord.
- 5) **Une panoplie de mesures de soutien au développement de la haute technologie** dont la bio-informatique pourrait bénéficier, notamment des mesures fiscales très favorables et la présence de sociétés de capital de risque.
- 6) **La présence d'une industrie des TIC très importante**, ainsi que de centres de calcul de haute puissance.

³⁰ *Développement de la bio-informatique au Québec : Opportunités et priorités.*

7) Les programmes Genome Canada et Génome Québec qui vont soutenir dans les années qui viennent des projets suscitant une forte demande en bio-informatique.

Équipement de haute puissance et réseaux

Le principal obstacle au développement de la bio-informatique au Québec est le même qu'au Canada et ailleurs : l'insuffisance de l'offre en personnel spécialisé dans le domaine. Comme on le dit plus haut, le Québec dispose d'une main-d'œuvre scientifique et technique de haute compétence dans les sciences de la vie. Mais les diplômés de ces disciplines qui possèdent en plus une expertise dans les technologies de l'information, spécialement dans les bases de données, sont extrêmement rares. Si on tient compte en plus de la pénurie généralisée des informaticiens dans tous les domaines, on comprend que la première action à entreprendre pour le développement de la bio-informatique au Québec sera du côté de la formation des ressources humaines. La demande est là, c'est l'offre qu'il faut stimuler de toute urgence. À ce sujet, tous les témoignages concordent : les besoins en bio-informaticiens dans l'industrie des biotechnologies et du pharmaceutique, sans parler de ceux des centres de recherche universitaires, sont considérables et vont encore augmenter au cours des prochaines années.

En 1998, le Réseau canadien des maladies génétiques avait estimé la demande en diplômés de deuxième et troisième cycles en bio-informatique à 300 postes à combler en cinq ans au Canada. Ce chiffre peut facilement être doublé, voire triplé aujourd'hui. On sait aussi qu'une forte proportion (jusqu'à deux tiers) des postes annoncés en bio-informatique ne sont pas comblés. Les carences de main-d'œuvre spécialisée en certains domaines, la difficulté de recruter des bio-informaticiens et l'insuffisance du capital de risque (du moins dans ce domaine précis) sont les trois facteurs désavantageux mis en relief par l'enquête de la Chaire en gestion des bio-industries³⁰.

³⁰ *Ibid.*

20 14 234
54
3201 56
67898
45 29 1 23
65
444 534
8795
5622 12873

Chapitre 3

Une stratégie de démarrage pour la bio-informatique au Québec

Agir rapidement
et efficacement

La stratégie de développement présentée dans ce chapitre constitue une proposition d'investissement rapide et efficace dans les quelques éléments clés qui pourront accélérer l'implantation du domaine de la bio-informatique au Québec. La stratégie proposée est essentiellement proactive. Elle mise sur les principaux atouts du Québec pour contrer le retard qu'il a pris et le fait que la bio-informatique ne soit à peu près pas prise en compte par les institutions actuelles de recherche et de soutien à la recherche. La stratégie proposée permettra d'agir sur trois fronts simultanément :

- 1) La formation des ressources humaines en bio-informatique;
- 2) L'organisation de l'activité bio-informatique comme volet constitutif de la génomique et de la protéomique au Québec, ainsi que comme champ de recherche spécifique;
- 3) Le soutien au développement d'une compétence industrielle en bio-informatique.

Ce sont là des mesures de démarrage, puisqu'on parle ici d'un domaine encore très embryonnaire. Il serait prématuré de proposer une stratégie de développement industriel complète. L'approche adoptée ici consiste à définir des moyens rapides et efficaces pour développer une expertise scientifique et technologique dans un champ qui offre un très fort potentiel de valeur ajoutée pour le Québec. Le Conseil estime que, pour l'instant, c'est du côté de la formation et de l'organisation de la recherche que doivent être mises les priorités. Cependant, il faut également s'assurer de maximiser les retombées industrielles de cette discipline. D'où, notamment, l'insistance qu'on voudra mettre à associer l'industrie aux deux premiers volets de la stratégie proposée.

Les découvertes en génomique et l'ouverture du champ de la protéomique dans les années qui viennent (et même les prochaines décennies) sont si vastes que les pays qui y mettent un effort sérieux profiteront nécessairement de ses retombées économiques. En pratique, ceux qui y consacreront les ressources les plus abondantes sont aussi ceux qui auront les meilleures chances de développer des créneaux intéressants. Mais il faut aussi agir avec rapidité et efficacité. Tous les observateurs le reconnaissent, le facteur temps est primordial dans ce domaine. En avril dernier, par exemple, le Ontario Research and Development Challenge Fund (ORDCF) annonçait un investissement de 74 millions de dollars en génomique. Parmi les projets retenus, se trouve un centre de bio-informatique (Bioinformatics Supercomputing Centre) situé au Toronto's Hospital for Sick Children.

Le Québec possède des atouts pour assurer ce développement, mais à la condition d'adopter une approche d'organisation et de coordination d'ensemble.

Confier la maîtrise d'œuvre à Génome Québec

Pour réussir, une telle stratégie exige un maître d'œuvre qui prendra la responsabilité de l'organisation et de la coordination du développement de la génomique au Québec. Le Conseil de la science et de la technologie propose de confier cette mission à Génome Québec. En effet, cet organisme est l'instrument dont le gouvernement du Québec s'est doté pour assurer le développement de la génomique et des domaines connexes au Québec, suite à la création de Génome Canada. Son rayon d'action est vaste, car il touche aussi bien le domaine pharmaceutique que l'agroalimentaire, les pêches, la foresterie et l'environnement. De plus, les objectifs des deux volets de son programme de financement, de même que ses orientations générales, vont tout à fait dans le sens d'une organisation du milieu de la recherche et du développement d'applications industrielles. Aucune autre organisation existante ne détient un mandat aussi bien adapté aux exigences formulées ici. Son statut corporatif lui permet d'agir rapidement et avec flexibilité, ce qui est capital ici.

De création très récente, Génome Québec est cependant encore en phase de démarrage, puisqu'il vient à peine de commencer ses activités. Pour s'acquitter d'un mandat comme celui de mettre en place les assises de la bio-informatique au Québec, il lui faudra s'associer à un certain nombre de partenaires appelés à jouer un rôle clé : les administrations universitaires, les centres de recherche comme l'IRB, les organismes subventionnaires (FRSQ, FCAR, CQRS), les principales entreprises et les associations industrielles, les sociétés de capital de risque, etc.

VOLET FORMATION

1- Un programme intensif de formation-perfectionnement en bio-informatique

Une urgence : former des bio-informaticiens

Le manque de ressources humaines en bio-informatique est un problème universel qui nécessite des actions correctives rapides et exige une ferme volonté d'agir. À moyen terme, on doit constituer des programmes de deuxième et troisième cycles dans les universités. À court terme cependant, il faut songer à des formules plus rapides (*fast track*).

Étant donné la pénurie dans le secteur informatique en général, il est clair que la formation en bio-informatique doit miser sur des candidats autres que ceux de la filière de formation générale en informatique, car les diplômés sont déjà fortement attirés vers les entreprises de communication, multimédia, etc. On peut croire qu'il y a plus de chances de former des bio-informaticiens en procurant une formation supplémentaire à des gens ayant déjà une spécialité en sciences de la vie qu'à tenter de convaincre une partie importante des étudiants en informatique de se spécialiser en bio-informatique. À plus long terme cependant, la croissance de la demande en bio-informaticiens dans les bio-industries, ainsi que les offres salariales qui leur seront associées, devraient attirer de plus en plus d'informaticiens proprement dits.

Impliquer les entreprises

Il faut dès à présent impliquer l'industrie dans l'opération, surtout les entreprises qui ont déjà des activités de bio-informatique et qui pourraient servir de sites de stages pratiques pour les étudiants en formation.

Recommandation 1

En toute priorité, mettre en place, d'ici septembre 2001, un programme intensif de formation-perfectionnement en bio-informatique au Québec.

Objectif

Former rapidement (en six à dix mois) entre dix et vingt personnes capables de maîtriser suffisamment les principaux éléments de la bio-informatique pour répondre aux besoins les plus urgents des équipes de recherche et de l'industrie au Québec.

Clientèle

Diplômés et /ou professionnels en exercice, universitaires et industriels possédant une formation en génétique, biochimie, biologie moléculaire, pharmacologie ou toute autre discipline pertinente.

Moyens

Recruter des spécialistes canadiens et étrangers, afin de concevoir et d'offrir un programme de formation dès l'automne 2001.

Négocier des ententes avec les universités pour trouver la structure d'accueil nécessaire à ce programme de formation; faire de même avec des entreprises consommatrices de bio-informatique, pour qu'elles puissent accueillir des stagiaires dans le cadre du programme de formation.

Fournir une bonne publicité à ce programme intensif de formation-perfectionnement; effectuer un recrutement actif des candidats.
Étudier la possibilité d'offrir des bourses d'appoint.

Coûts

De l'ordre de 600 000 \$.

Responsables

Génome Québec (maîtrise d'oeuvre), avec la collaboration de partenaires universitaires et industriels.

2. Des bourses pour acquérir une formation supérieure à l'extérieur du Québec

Formation à l'extérieur du Québec

En parallèle à la mise en place d'un programme de formation-perfectionnement, et pour des raisons similaires, un deuxième volet de la stratégie de formation devrait consister à financer la formation de bio-informaticiens dans des centres de recherche à l'étranger. Il est clair qu'une telle proposition présente un certain risque, celui de voir les meilleurs boursiers retenus par des entreprises ou des universités à l'extérieur du Québec. Mais tant et aussi longtemps qu'on n'offrira pas au Québec un solide programme universitaire de troisième cycle en bio-informatique, des bourses de doctorat et de postdoctorat demeurent encore un excellent moyen de former les scientifiques dans cette discipline. La meilleure

façon de faire revenir et de garder ensuite cette expertise au Québec est de travailler à l'organisation d'une demande forte et attrayante dans les milieux universitaires et industriels québécois. Ajoutons que, même lorsque des programmes de deuxième et troisième cycles en bio-informatique existeront dans les universités du Québec, on aura encore intérêt à envoyer des étudiants se former et se perfectionner à l'étranger.

Recommandation 2

Permettre, si possible dès l'hiver 2002, à une dizaine de boursiers d'acquérir une formation supérieure en bio-informatique à l'extérieur du Québec

Objectif

Former le noyau d'expertise avancée dont on aura besoin dans les années qui viennent pour développer la recherche et la formation en bio-informatique au Québec.

Clientèle

Candidats de niveau doctorat ou postdoctorat en sciences de la vie ou en informatique.

Moyens

Organiser un programme de bourses spécifiques qui pourrait être administré conjointement par le Fonds FCAR (pour les sciences de la vie et l'informatique) et le FRSQ (pour les sciences de la santé).

Publiciser le programme et effectuer un recrutement actif des candidats.
Tenir un premier concours au printemps 2001.

Coûts

Autour de 500 à 600 000 \$ annuellement (entre 10 et 20 candidats) pendant trois ans.

Responsables

Génome Québec (maîtrise d'oeuvre), avec la collaboration du Fonds FCAR et du FRSQ.

3- Des programmes de formation universitaire supérieure en bio-informatique

Programmes de 2^e
et 3^e cycles

Tout en prenant les mesures pour instituer une formation rapide en bio-informatique, il faut s'intéresser à l'organisation de programmes universitaires de formation de niveaux 2^e et 3^e cycles dans cette spécialité.

Il est clair que pour engager une initiative importante de formation en bio-informatique, il faut être capable de recruter rapidement de très bons professeurs. Étant donné le problème de la rareté dans cette spécialité, les mécanismes de financement de l'enseignement universitaire actuels sont inadéquats. On doit entrer en concurrence, non seulement avec les universités américaines, mais avec l'industrie. La formule des chaires industrielles pourrait peut-être apporter une solution. On peut penser également à un concours du type du programme stratégique de professeurs-chercheurs du Fonds FCAR. Il faut ici s'assurer d'un appui institutionnel fort de la part des universités.

L'essentiel cependant est d'organiser une offre de formation qui puisse répondre adéquatement à la demande. Pour cela, l'implication de l'industrie dans l'évaluation de cette demande, de même que dans les contenus des programmes et l'organisation de stages en entreprise, est absolument nécessaire.

Encore une fois ici, la contrainte de temps joue de façon primordiale. On ne fera pas beaucoup de progrès s'il faut attendre trois ans ou plus pour que se mette en place un programme universitaire en bio-informatique qui aurait suivi la filière d'approbation courante. On devra au contraire être très audacieux et ne pas hésiter à essayer des formules novatrices ou expérimentales. Entre le format du cours intensif proposé dans la première recommandation et les programmes de deuxième ou troisième cycle de type standard, il y a sûrement place pour tout un éventail de formules intermédiaires qui répondraient plus rapidement à la demande de personnel hautement qualifié en bio-informatique. On peut penser à un Diplôme d'études supérieures spécialisées (DESS) en bio-informatique. On peut aussi faire de la bio-informatique une composante importante à l'intérieur de programmes existants dans les sciences de la vie. Des ententes interuniversitaires pour monter rapidement un ou des programmes conjoints seraient également une voie à explorer.

En fin de compte, la collaboration étroite de tous les partenaires (les établissements universitaires eux-mêmes, le gouvernement du Québec, l'industrie) est ici essentielle pour que des programmes de formation universitaires aux cycles supérieurs puissent voir le jour au Québec d'ici deux ans ou moins. Un démarchage important auprès de toutes ces instances sera nécessaire de la part de Génome Québec.

Recommandation 3

Mettre en place le plus rapidement possible des activités et des programmes universitaires de formation de deuxième et troisième cycles en bio-informatique au Québec

Objectif

Répondre avec souplesse aux besoins de formation de pointe en bio-informatique.

Clientèle

Étudiants possédant un premier cycle en sciences de la vie ou en informatique.

Moyens

Rencontrer et convaincre les autorités universitaires de l'importance et de l'urgence de monter des programmes de formation avancée en bio-informatique.

Obtenir l'appui du ministre de l'Éducation dans cette opération.

Entreprendre le plus rapidement possible les démarches pour recruter des professeurs-chercheurs canadiens et étrangers. À ce chapitre, réserver une enveloppe pour des postes pouvant intéresser les boursiers envoyés à l'étranger (recommandation 2).

Négocier, s'il y a lieu, des ententes interuniversitaires pour l'organisation des programmes.

Associer l'industrie à la démarche, dès le début et tout au long du processus. Intégrer dans la formation des stages obligatoires en entreprise.

Coûts

Des montants substantiels devront être versés pour la conception et l'implantation la plus rapide possible des programmes, avec toutes les ressources requises.

Responsables

Génome Québec pour l'initiative et la coordination; les universités québécoises disposées à collaborer pour construire les programmes; les entreprises utilisatrices de bio-informatique en partenariat tout au long du processus; le ministre de la Recherche, de la Science et de la Technologie, pour appuyer les démarches de Génome Québec, notamment auprès du ministre de l'Éducation.

Perspectives

À plus long terme, on peut prévoir que la perspective *in silico* prendra de plus en plus d'importance dans les programmes de formation en sciences de la vie, car c'est là une tendance de fond qui se dessine dans l'évolution future de ces disciplines. Mais une stratégie de développement rapide de la bio-informatique au Québec ne saurait se contenter d'attendre que ce processus naturel se fasse. Sans en faire l'objet d'une recommandation, le Conseil estime qu'il faut déjà préparer l'avenir et s'assurer que les programmes de formation en sciences de la vie, dès le niveau collégial, intègrent de façon plus marquée la maîtrise des outils informatiques dans la formation des futurs diplômés.

VOLET RECHERCHE ET RÉSEAUTAGE

Le second volet de la stratégie proposée concerne le développement en réseau de l'activité de bio-informatique. La mise en place de ce réseau est absolument essentielle. On sait que la bio-informatique se développe principalement en symbiose avec les projets de génomique ou protéomique qui s'appuient sur elle; l'action structurante de Génome Québec en matière de génomique et protéomique devrait donc avoir des effets salutaires sur la consolidation de l'expertise québécoise de bio-informatique qui sera associée à tous ces projets. Génome Québec a aussi le mandat de soutenir des projets de R-D en bio-informatique proprement dite et il est impératif qu'il accorde à cette responsabilité une forte priorité.

Réseauter ou concentrer?

Pour structurer rapidement le milieu de la recherche en bio-informatique au Québec, deux types de solutions ont été envisagées :

- 1) Développer les ressources de bio-informatique à l'intérieur d'un réseau de grands projets comme celui de Génome Québec;
- 2) Concentrer ces ressources dans un lieu privilégié (centre, institut) et s'efforcer de lui fournir la masse critique nécessaire le plus rapidement possible.

Bien qu'elles se réfèrent à des stratégies d'organisation complètement différentes, ces deux options ne sont pas entièrement incompatibles. Le Conseil croit que, compte tenu de l'état de la bio-informatique au Québec et de l'urgence d'agir, la première option est celle qui peut donner de meilleurs résultats à court terme. Par contre, à plus long terme, la seconde option ne doit pas être écartée. Un réseau de recherche peut en effet se développer de plus d'une façon. Dans certains cas, on peut être amené à consolider progressivement une tête de réseau et donc à s'orienter vers une option plus concentrée. Parmi les formules de structure à envisager, on pourra alors considérer celle d'un centre de liaison et de transfert (comme le CRIM ou le CERCA), celle d'un centre péri-industriel comme l'INO ou d'un centre de recherche et de formation pluriuniversitaire.

Bénéfices recherchés

Dans l'immédiat, la structuration d'un réseau de recherche génomique-protéomique où la bio-informatique se développera de façon optimale constitue une action incontournable. Les effets d'entraînement recherchés sont en effet multiples et importants :

- Offrir des structures d'accueil et de rétention pour les chercheurs et praticiens de la bio-informatique que nous voulons former;
- Soutenir la formation aux 2^e et 3^e cycles en bio-informatique par l'activité de recherche de pointe;
- Créer une assise de collaboration universités-entreprises vigoureuse en bio-informatique;
- Offrir une pépinière de compétences pour l'industrie;
- Servir d'incubateur d'entreprises par essaimage;
- Mettre le Québec sur la carte internationale des compétences en bio-informatique;
- Servir d'attracteur pour quelques grosses entreprises qui recherchent précisément ces compétences;
- Mettre en place le noyau de base d'une future « grappe » industrielle en génomique-protéomique-bio-informatique.

Recommandation 4

Développer l'activité de bio-informatique au Québec à l'intérieur d'un réseau de projets de recherche structurants en génomique, en protéomique, de même qu'en bio-informatique proprement dite.

Objectif

Organiser une capacité québécoise de haut calibre en bio-informatique et en maximiser l'impact par un fonctionnement en réseau.

Clientèle

Les équipes de recherche universitaires et industrielles travaillant en génomique, protéomique ou en bio-informatique et possédant un fort potentiel en termes d'excellence et de retombées économiques.

Moyens

S'assurer que les projets de génomique et protéomique soutenus par Génome Québec ont un volet bio-informatique important et que ce volet bénéficie des ressources financières voulues.

Par ailleurs, accorder priorité au soutien d'un ou deux projets de recherche structurants et prometteurs en bio-informatique proprement dite.

Faire en sorte que des partenaires industriels soient liés à chacun des projets.

Concevoir éventuellement un volet spécifique au programme de Génome Québec pour soutenir un certain nombre de projets de recherche en bio-informatique, lorsque le nombre et l'excellence des équipes le justifieront.

Coûts

Il faudra attendre l'annonce des résultats du premier concours de Génome Québec pour être en mesure de chiffrer les sommes additionnelles qui seront nécessaires pour encourager particulièrement le développement de projets spécifiques en bio-informatique.

Responsables

Génome Québec, avec la collaboration des fonds FCAR et FRSQ.

VOLET TRANSFERT DE COMPÉTENCES VERS L'INDUSTRIE

Comme il a été dit au début de ce chapitre, il serait prématuré de proposer une stratégie de développement industriel pour la bio-informatique, puisqu'on parle ici d'un domaine encore très embryonnaire. Les mesures proposées précédemment, visant à mettre en place les conditions nécessaires pour structurer une compétence québécoise en bio-informatique, devraient avoir des répercussions importantes sur l'industrie, soit par l'intégration de cette compétence dans les entreprises qui en auront besoin pour se développer, soit par la création de nouvelles firmes misant spécifiquement sur cette compétence.

Dans quels créneaux des bio-industries la bio-informatique se développera-t-elle au Québec? Il est encore trop tôt pour le dire. Les trois prochaines années vont être cruciales à cet égard. Si un effort intensif est fait du côté de la formation et de la recherche, comme le propose le Conseil, on peut s'attendre à ce qu'un effet d'entraînement ait lieu du côté industriel. C'est d'ailleurs dans cette optique que l'on a insisté pour associer l'industrie à l'orientation et à l'élaboration des volets formation et recherche de la stratégie (recommandations 1 à 4).

La question du capital de risque

Certains problèmes surgiront possiblement dans les années qui viennent, mais ils ne peuvent être encore analysés ni, à plus forte raison, solutionnés. Par exemple, plusieurs observateurs au Canada et dans le monde soulignent la difficulté de trouver du capital de risque pour les entreprises en démarrage dans ce domaine. Il est encore trop tôt pour dire si ce problème va se manifester également au Québec, ou si les firmes de capital de risque existantes vont être prêtes à investir lorsqu'il le faudra. Cependant, si Génome Québec et ses partenaires parviennent à construire rapidement un réseau de recherche et de ressources de haut calibre en génomique, protéomique et bio-informatique au Québec, ce capital de compétences

devrait pouvoir exercer un attrait, à point nommé, chez les investisseurs. Sans minimiser l'importance de cette question, le Conseil n'a pas jugé bon d'émettre de recommandation particulière sur le capital de risque, mais il suggère fortement (voir recommandation 6) que le gouvernement suive de très près la situation au cours des trois prochaines années.

La cinquième recommandation vise spécifiquement à faciliter l'intégration de la bio-informatique dans le tissu industriel québécois. L'instrument fiscal est sans doute le plus approprié ici, comme il l'a été dans d'autres domaines (multimédia, par exemple). Les mesures fiscales proposées sont destinées à soutenir l'embauche de bio-informaticiens par les entreprises. Il s'agit d'une mesure temporaire de cinq ans dont il faudrait évaluer l'impact après quatre ans.

Pourquoi agir aussi sur la demande?

Pourquoi recommander de jouer sur la demande industrielle en bio-informatique, alors qu'il a été constaté que cette demande est importante et qu'elle est appelée à connaître une forte croissance dans les années qui viennent? Il faut le dire très clairement : la ressource bio-informatique est rare et chère. Partout dans le monde, les personnes qui vont se chercher une formation dans le domaine reçoivent tellement d'offres d'emploi qu'elles ont souvent du mal à terminer leur cours; bien entendu, les salaires sont à l'avenant. Il faut donc tenir compte du fait que, pour se doter de compétences en bio-informatique, les entreprises québécoises doivent concurrencer des compétiteurs extérieurs qui recherchent avidement le même type de personnel et ont souvent des capacités de payer bien supérieures. Pour cette raison, le Conseil estime qu'un incitatif fiscal spécifique pour la bio-informatique peut donner aux bio-industries québécoises le coup de pouce dont elles ont besoin.

Recommandation 5

Fournir aux entreprises des bio-industries du Québec, à partir de 2002, des incitatifs fiscaux spécifiques reliés à l'embauche de bio-informaticiens.

Objectif

Accélérer l'intégration de compétences en bio-informatique dans les bio-industries du Québec.

Clientèle

Les entreprises des bio-industries québécoises, y compris les filiales de compagnies étrangères, qui ont à se doter rapidement d'une capacité en bio-informatique pour maintenir ou accroître leur compétitivité.

Moyens

En complémentarité avec les mesures visant à accroître l'offre de ressources humaines, fournir un appui à la demande, en diminuant notamment les coûts de l'embauche de personnel en bio-informatique dans les entreprises. La mesure devrait être évaluée après quatre ou cinq ans.

Coûts

À préciser.

Responsables

Le MRST et ses partenaires, à savoir le ministère des Finances et le ministère de l'Industrie et du Commerce.

SUIVRE DE PRÈS L'ÉVOLUTION DU DOMAINE

- Une vigie essentielle La mise en place des mesures de démarrage proposées dans cet avis est nécessaire, mais elle n'est pas suffisante pour s'assurer que le développement de la bio-informatique au Québec va s'effectuer de façon optimale. L'importance des enjeux que nous avons dégagés à la fin du chapitre 2 nous incite à suggérer en outre un suivi excessivement attentif de l'évolution du domaine, tant du point de vue scientifique qu'industriel et social, au cours des trois ou quatre prochaines années.
- Un rôle partagé Le Conseil lui-même entend se préoccuper des suites que l'on apportera à ses recommandations, mais il ne lui revient pas d'effectuer un suivi régulier de l'évolution de la situation de la bio-informatique au Québec. Ce sont d'autres instances que le Conseil qui devraient assurer cette responsabilité : Génome Québec, bien sûr, parce qu'il en est le principal artisan, mais aussi et peut-être surtout le MRST, parce qu'il s'agit d'un domaine scientifique et technologique de pointe extrêmement prometteur pour le développement socio-économique du Québec. D'autres ministères sont concernés et devraient être impliqués, chacun dans son secteur : le ministère de l'Industrie et du Commerce, le ministère de l'Éducation, le ministère de la Santé et des Services sociaux, le ministère de l'Agriculture, des Pêcheries et de l'Alimentation, le ministère de l'Environnement...
- De nombreux aspects à couvrir Le *monitoring* et la vigie ne doivent pas porter exclusivement sur la bio-informatique : c'est tout le champ de la génomique et de la protéomique et de leurs différents secteurs d'application qu'il faudra suivre de très près, afin de comprendre comment se développe la bio-informatique au Québec, d'évaluer ses progrès et d'identifier ses difficultés. Parmi les dimensions qui requièrent une attention particulière, mentionnons l'organisation rapide d'une offre de formation avancée en bio-informatique, le démarrage d'entreprises et la réaction du capital de risque, les besoins en infrastructures physiques et en équipements, l'intervention d'autres joueurs comme les fonds subventionnaires, les organismes fédéraux, Valorisation-Recherche Québec et ses quatre sociétés de valorisation de la recherche universitaire, etc. Il faudra également se soucier d'assurer l'interface avec le grand public, afin de l'informer et de lui permettre de s'exprimer sur les questions sociales et éthiques liées à l'évolution du domaine. La nouvelle commission de l'éthique de la science et de la technologie, si elle voit effectivement le jour, devrait apporter une précieuse contribution à ce dernier aspect.

Recommandation 6

Suivre de très près l'évolution de la génomique, de la protéomique et de la bio-informatique au Québec, tant du point de vue scientifique que du point de vue industriel et social, au cours des trois ou quatre prochaines années.

Objectifs

Vérifier les résultats atteints, corriger le tir au besoin, en cours de route, et préparer la prochaine génération de mesures de soutien.

Clientèle

Tous les acteurs engagés dans la phase de démarrage de la bio-informatique au Québec.

Moyens

Faire le point au moins annuellement avec tous les intervenants sur l'état d'évolution du domaine.

Mettre en place des indicateurs de suivi sur chacune des dimensions les plus importantes.

Rester à l'écoute des difficultés qui pourraient survenir par un dialogue régulier avec les équipes de recherche scientifique et les industries. Organiser des forums de discussion avec les médias et le public.

Responsables

MRST et ses partenaires : Génome Québec et les autres ministères concernés.

Annexe 1

Membres du Conseil de la science et de la technologie

Le président par intérim

M. Camil Guy (jusqu'au 30 juin 2000)

La présidente

M^{me} Hélène P. Tremblay (depuis le 21 août 2000)

Les membres

M. Maurice Avery

Président

Soft Innove inc.

M^{me} Claude Benoît

Présidente et chef de direction

Société du Vieux-Port de Montréal

Directrice, Centre interactif des sciences de Montréal

M^{me} Louise Dandurand

Vice-rectrice à la recherche, à la création et à la planification

Université du Québec à Montréal

M. Jean-Guy Frenette

Vice-président à la concertation sectorielle et adjoint au premier vice-président

Fonds de solidarité des travailleurs du Québec

M. Martin Godbout

Président

Hodran inc.

M. Pierre-André Julien

Professeur et titulaire de la Chaire Bombardier

Université du Québec à Trois-Rivières

Mme Nicole Lafleur

Directrice générale

Cégep de Lévis-Lauzon

M. Germain Lamonde (jusqu'au 6 septembre 2000)

Président

Exfo ingénierie électro-optique inc.

M^{me} Maryse Lassonde (jusqu'au 10 novembre 2000)

Professeur-chercheur, Département de psychologie

Université de Montréal

M. Réginald Lavertu

Directeur général

Collège de Rosemont

M^{me} Louise A. Perras (jusqu'au 16 juin 2000)

Présidente-directrice générale
Consortium Multimédia CESAM

M. Denis Poussart

Professeur - Département de génie électrique et informatique
Université Laval

M. Jean-Marc Proulx

Vice-président R-D
Groupe Conseil DMR inc.

M^{me} Louise Proulx (jusqu'au 15 août 2000)

Vice-présidente, Développement de produits thérapeutiques
BioChem Pharma inc.

Les membres observateurs

M^{me} Pauline Champoux-Lesage

Sous-ministre
Ministère de l'Éducation

M^{me} Marie-France Germain

Sous-ministre adjointe à la planification
Ministère de la Recherche, de la science et de la Technologie

M. Michel Desrochers

Directeur général
Institut de recherche en biotechnologie

Le Secrétaire général

M. Camil Guy (jusqu'au 30 juin 2000)

Annexe 2

Membres du Comité sur la bio-informatique

Le président

M. Martin Godbout
Président
Hodran inc.

Les membres

M. Paul Beaulieu
Directeur général, Chaire en gestion des bio-industries
École des sciences de la gestion
Université du Québec à Montréal

M. Thomas Hudson
Directeur, Centre génomique de Montréal
Professeur adjoint, Université McGill
Médecin, Centre universitaire McGill
Directeur adjoint Whitehead Institute / MIT Genome Center

M. Jacques Hurtubise
Directeur, Centre de recherches mathématiques
Université de Montréal

M. Jean Morrissette
Chercheur senior
Centre de recherche en endocrinologie moléculaire et oncologie
Faculté de médecine - Université Laval

M. Jean-Marc Proulx
Vice-président R-D
Groupe Conseil DMR inc.

Le Secrétaire du Comité

M. Alain Bergeron
Agent de recherche
Conseil de la science et de la technologie