

Politique sur la diffusion des données et le partage des ressources

PRINCIPE STRATÉGIQUE

Génome Canada souscrit au principe de la diffusion rapide des données et de la mise en commun des ressources uniques avec la communauté scientifique; les projets financés par Génome Canada doivent donc, en temps opportun, partager les données et les ressources, tout en imposant que peu ou pas de restrictions. En assurant à la communauté scientifique un accès opportun aux résultats des projets qu'elle finance, Génome Canada vise, par la présente, à accélérer l'application de la recherche, au bénéfice de l'humanité.

La présente a initialement été établie par le Comité consultatif science et industrie de Génome Canada et approuvée par le conseil d'administration le 17 juin 2005. La politique révisée a été approuvée par le conseil d'administration de Génome Canada le 18 septembre 2008, et elle s'applique à tous les projets financés par Génome Canada après le 1^{er} septembre 2008.

MÉCANISMES DE PARTAGE DES DONNÉES ET DES RESSOURCES

Génome Canada s'attend à ce que les chercheurs partagent le plus rapidement possible les données et les ressources. Lorsque le projet a pour but de produire des données ou des ressources pour l'ensemble de la communauté scientifique, il doit se conformer aux principes de la diffusion des données et de la mise en commun des ressources d'un « projet de ressources communautaires » qui s'entend d'un « projet de recherche spécialement conçu et mis en œuvre pour créer un ensemble de données, des réactifs ou d'autre matériel qui seront d'abord et avant tout des ressources pour la communauté scientifique en général ». Cette définition et les principes connexes de diffusion des données et de mise en commun des ressources ont été élaborés à une rencontre qui a eu lieu en janvier 2003 à Fort Lauderdale¹. Génome Canada encourage l'application des principes de la diffusion rapide des données, avant leur publication, à d'autres types de projets et collabore avec d'autres bailleurs de fonds de recherche à la promotion de cette pratique.

Génome Canada reconnaît la publication comme mode de diffusion des données, et s'attend, à tout le moins, à ce que les données soient diffusées et partagées au plus tard à la date de publication originale des conclusions principales de tout ensemble de données produit par le projet. En ce qui concerne les grands ensembles de données recueillies au cours de plusieurs périodes ou phases distinctes, il est raisonnable de s'attendre à la publication de données par phases, lorsque les données sont connues ou lorsque les principales conclusions d'une phase de la recherche sont publiées. Toutes les données doivent toutefois être publiées sans restriction à la conclusion d'un projet.

Génome Canada reconnaît la nécessité de protéger les données brevetables et autres données exclusives; les données doivent cependant être diffusées au plus tard à la date du dépôt de la demande de brevet (y compris les brevets provisoires) ou au moment de la publication, selon la première éventualité à se produire. Une prolongation maximale de 90 jours est acceptable, si des circonstances atténuantes la justifient. Pour diffuser les données au-delà de ce délai, il faut obtenir l'autorisation écrite de Génome Canada, qui fera au besoin appel à des conseillers

¹ La définition de « projet de ressources communautaires » a été élaborée à une réunion qui a eu lieu les 14 et 15 janvier 2003 à Fort Lauderdale, en Floride. Le rapport (en anglais seulement) sur les conclusions de la réunion, intitulé *Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility*, est publié à http://www.wellcome.ac.uk/stellent/groups/corporatesite/@policy_communications/documents/web_document/wtd003207.pdf.

indépendants pour évaluer la demande. De plus, les ententes de confidentialité entre des commanditaires et des chercheurs ne doivent pas être indûment restrictives et doivent autoriser les publications dans un délai défini et raisonnable.

Les projets financés par Génome Canada doivent également tenir compte de la mise en commun des ressources produites par les projets, par exemple des spécimens biologiques uniques et des programmes informatiques d'analyse d'ensembles de données. Les réactifs biologiques, comme les souches uniques, doivent être déposés dans des dépôts tels qu'ATCC et les programmes informatiques conçus pour analyser les grands ensembles de données doivent être mis à la disposition des autres au moyen d'ententes de licence qui respectent les principes de la « source ouverte » (se reporter à <http://www.opensource.org/> pour en voir un exemple).

PLANS DE PARTAGE DES DONNÉES ET DES RESSOURCES

Les candidats à des concours de Génome Canada doivent joindre un plan de partage des données et des ressources à leur candidature. Ce plan doit aborder les aspects suivants :

- **Types de données et de ressources** - Quelles données et quelles ressources le projet produira-t-il?
- **Choix du moment et mécanisme de partage** – Pour chaque type de données et de ressources, quand, comment et où les données seront-elles mises à la disposition de la communauté scientifique? Les chercheurs doivent utiliser les bases et les dépôts de données publics reconnus, lorsqu'il en existe, sinon indiquer les plans prévus pour diffuser la ressource en question à l'ensemble de la communauté.
- **Qualité** – Quels seront les mécanismes de contrôle et d'assurance de la qualité mis en place?
- **Normes** – La communauté applique-t-elle des normes pour les données et/ou les ressources qui seront produites et comment le projet s'y conformera-t-il? Génome Canada s'attend à ce que les données et les ressources produites soient conformes aux normes reconnues à l'échelle internationale, et les chercheurs doivent préciser les normes appliquées, le cas échéant.
- **Éthique, protection des renseignements personnels et confidentialité** – Comment les données à caractère confidentiel seront-elles protégées? Lorsque la recherche fait intervenir des sujets humains, comment les intérêts des participants à la recherche seront-ils protégés? Comment le plan de partage des données et des ressources se conforme-t-il aux modalités du consentement?
- **Propriété intellectuelle** – Y aura-t-il des restrictions ou des délais concernant le partage des données et/ou des ressources pour garantir la protection de toute propriété intellectuelle ou données et/ou ressources exclusives².
- **Modalités et conditions** – Quelles sont les modalités et les conditions, le cas échéant, de l'accès et de l'utilisation des données et/ou des ressources? Il est à noter que les chercheurs, lorsqu'ils diffusent des données, ne peuvent pas imposer de limites aux questions posées ou aux méthodes utilisées, ni exiger, comme condition d'obtention des données ou des ressources, d'être cités comme coauteurs.

Génome Canada et ses comités d'évaluation examineront le plan de partage des données et des ressources proposé par les candidats et s'assureront qu'il est conforme à la politique; les fonds ne seront pas versés aux projets tant qu'un plan acceptable n'aura pas été approuvé et intégré aux modalités de l'octroi des fonds. Le personnel de Génome Canada et du centre de génomique veillera au respect de la présente par diverses méthodes, dont l'évaluation provisoire.

² Se reporter à la Politique sur la propriété intellectuelle de Génome Canada : <http://www.genomecanada.ca/medias/pdf/fr/IntellectualProperty.pdf>

UTILISATION DES DONNÉES ET DES RESSOURCES

Génome Canada croit que la communauté scientifique tire profit de la pratique du partage rapide des données et des ressources. Pour s'assurer de la perpétuation de cette pratique et respecter les intérêts des auteurs de données et/ou de ressources, Génome Canada s'attend à ce que les utilisateurs reconnaissent la source et se conforment à toutes les modalités et conditions qui se rattachent à l'utilisation.

Génome Canada réévaluera la présente et collaborera avec d'autres bailleurs de fonds de recherche à la promotion de pratiques exemplaires dans ce domaine.

EXEMPLES DE DÉPÔTS INTERNATIONAUX DE DONNÉES ET DE RESSOURCES

Le tableau suivant donne des exemples des dépôts où différents types de données ou des ressources uniques, produits par les projets financés par Génome Canada, peuvent être versés. Les exemples sont des bases de données ou des dépôts reconnus pour certains types de données et de ressources; cette liste n'est pas exhaustive.

Type de données/ressources	Base de données	Adresse URL en vue de la soumission de données
ADN, ARN, séquences de protéines	DDBJ/EMBL/GenBank ¹	http://www.ncbi.nlm.nih.gov/Genbank/submit.html
Séquences EST	DDBJ/EMBL/GenBank	http://www.ncbi.nlm.nih.gov/dbEST/how_to_submit.html
Séquences STS	DDBJ/EMBL/GenBank	http://www.ncbi.nlm.nih.gov/dbSTS/how_to_submit.html
Données sur puces à ADN (p. ex. expression des gènes et VCN)	GEO	http://www.ncbi.nlm.nih.gov/geo/
Polymorphismes nucléotidiques simples (SNP)	dbSNP	http://www.ncbi.nlm.nih.gov/SNP/
Génotype et phénotype humains (GWAS)	dbGaP	http://www.ncbi.nlm.nih.gov/sites/entrez?db=gap
Archives de traces et lectures abrégées des séquenceurs de la nouvelle génération	Trace Archive and Short Read Archive	http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?cmd=show&f=rfc&m=main&s=rfc And http://www.ncbi.nlm.nih.gov/Traces/sra/sra.cgi?cmd=show&f=main&m=main&s=main
Structures de protéines	PDB	http://www.wwpdb.org/
Interactions	IntAct	http://www.ebi.ac.uk/intact
Nomenclature du génome humain	HGNC	http://www.gene.ucl.ac.uk/nomenclature
Nomenclature du génome murin, données sur le phénotype de la souris et l'expression des gènes murins	MGI	http://www.informatics.jax.org/mgihome/nomen/index.shtml http://www.informatics.jax.org/submit.shtml
Souches biologiques	ATCC	http://www.atcc.org/
Codes de logiciels	SourceForge.net/ Canadian Bioinformatics Help Desk Software Repository	http://sourceforge.net/ http://www.gchelpdesk.ualberta.ca/repository/contents.php

¹ DDBJ, EMBL et GenBank sont des bases de données équivalentes qui font toutes partie de l'International Nucleotide Sequence Databases Collaboration, et soumettre des données à l'une ou l'autre équivaut à en soumettre à GenBank. Il est très commode pour les chercheurs canadiens de soumettre leurs données à GenBank, mais ils sont libres de les soumettre à l'une ou à l'autre de ces trois bases de données.