



GenAp

L'informatique au service des spécialistes en génomique et des cliniciens

Jamais le code génétique de l'être humain n'a-t-il été autant à portée de la main, essentiellement parce qu'on dispose depuis peu de technologies très pointues, mais aussi peu coûteuses pour séquencer l'ADN. La génomique est une science d'une utilité formidable pour étudier les origines de l'Homme, les maladies héréditaires et les systèmes biologiques. Cet outil puissant facilite aussi le travail clinique et ouvre la porte aux diagnostics, aux traitements et à l'évaluation des risques d'après le génome de l'individu. Malheureusement, la génomique n'a eu qu'un impact modeste sur la médecine clinique de tous les jours, car ses outils actuels s'adressent exclusivement à la recherche et leur usage s'avère complexe. De plus, exploiter les jeux de données colossaux en génomique n'est pas une mince affaire.

Voici GenAP (pour Genetics and Genomics Analysis Platform ou Plateforme d'analyse en génétique et en génomique), une plateforme qui, non seulement, repousse les limites de l'analyse des données en génomique, mais met ces dernières à la portée des biologistes et des médecins.

Des données volumineuses personnalisées

Il faudrait une pile de 50 DVD pour stocker le code génétique complet d'une personne; une étude sur un millier de patients engendrera des centaines de téraoctets de données. Transférer une telle masse d'informations avec les moyens usuels serait malcommode; il serait difficile de la stocker et, sans les outils adéquats, l'exploiter se révélerait tout simplement impossible.

GenAP réunit des centaines d'éléments de logiciels en génomique et des installations de calcul informatique de pointe de façon transparente pour en faire un outil de recherche unique qui possède la capacité de stockage et la puissance de calcul voulues. Ce système réparti garde les données médicales sur les malades à l'abri derrière un coupe-feu adhérent à des normes très sévères de confidentialité et d'accessibilité, tout en autorisant une vaste consultation des jeux de données publics sur le génome. Ainsi, les chercheurs et les cliniciens qui utilisent GenAP peuvent parcourir et partager les données

plus facilement, ce qui accélère les percées tant scientifiques que thérapeutiques.

Un jour, les découvertes en génomique permettront à l'omnipraticien de dépister de rares mutations lors de son auscultation et de prescrire un médicament spécialement conçu pour les soigner. Les familles touchées par ces maladies peu communes seront les premières à en profiter mais, au bout du compte, chacun jouira des retombées de tels outils. En effet, les affections qui affligent bon nombre d'entre nous – l'asthme, certaines formes de cancer, les cardiopathies et le diabète, par exemple – pourraient elles aussi résulter d'une récente ou ancienne variation du code génétique.

Travailler la main dans la main

GenAP est le produit d'une collaboration entre les équipes de génomique et de calcul informatique de pointe (CIP) de l'Université McGill University et de l'Université de Sherbrooke.

Il est possible d'accéder à la plateforme GenAP par le registre de logiciels de CANARIE. La plateforme intègre des services facilitant la gestion d'énormes fichiers répartis pour le traitement et l'analyse du génome, ainsi qu'une centaine d'outils en génomique.

GenAP réunit des centaines d'éléments de logiciels en génomique et des installations de calcul informatique de pointe de façon transparente pour en faire un outil de recherche unique.

Les chercheurs et les cliniciens qui utilisent GenAP peuvent parcourir et partager les données plus facilement, ce qui accélère les percées tant scientifiques que thérapeutiques.

Aspects techniques

Plateforme : GenAP (Genetics and Genomics Analysis Platform)

Description	Plateforme de calcul pour les chercheurs en sciences de la vie. Elle exploite le réseau ultrarapide de CANARIE et les installations de calcul informatique de pointe (CIP) de Calcul Canada en permettant aux chercheurs d'accéder à des services Web modernes et spécialisés ainsi qu'à des outils d'analyse étroitement intégrés aux ressources CIP.
Créateur(s)	Université McGill, Université de Sherbrooke
Collaborateur(s)	Génome Québec, Calcul Canada
Domaine de recherche	Bioinformatique
Version gérée ⁱ	Oui – disponible aux utilisateurs de Calcul Canada
Version autonome ⁱⁱ	Non
Soutien en nuage	Intégré à l'infrastructure CIP de Calcul Canada
Système d'exploitation	Linux
Licence	Plusieurs de source ouverte
Précisions	canarie.ca/fr/logiciels/genap-fr

Services contribués

	CVMFS stratum 0 et stratum 1	Antémémoire Squid pour machine virtuelle	Machine virtuelle avec image de Galaxy (plateforme en sciences de la vie)
Description	Répartissent les logiciels sur le Web grâce aux systèmes de fichiers des machines virtuelles (CVMFS). Aucune administration à distance supplémentaire des logiciels requise.	Place les ressources Squid en antémémoire pour une répartition rapide et robuste des logiciels avec le service CVMFS.	Donne accès à la plateforme Galaxy sur le Web.
Catégorie	Gestion des ressources/du nuage	Gestion de ressources/de nuage	Gestion de ressources/de nuage
Domaine de recherche	Génomique	Génomique	Génomique
Version gérée ⁱ	Oui – exploitée avec la plateforme GenAP pour configurer les logiciels sur des serveurs éloignés	Oui – exploitée comme un service par la plateforme GenAP pour mettre le contenu des CVMFS en antémémoire	Oui – exploitée avec la plateforme GenAP
Version autonome ⁱⁱ	Oui – logiciel disponible auprès de Calcul Canada	Oui – logiciel disponible auprès de Calcul Canada	Oui – disponible sous forme d'image de machine virtuelle auprès de Calcul Canada
Soutien en nuage	Intégré à l'infrastructure CIP de Calcul Canada	Intégré à l'infrastructure CIP de Calcul Canada	Intégré à l'infrastructure CIP de Calcul Canada
Système d'exploitation	Linux	Linux	Linux
Licence	Licence publique générale GNU	Licence publique générale GNU	Licence publique générale GNU
Details	canarie.ca/fr/logiciels/cvfms	canarie.ca/fr/logiciels/squid	canarie.ca/fr/logiciels/galaxy

ⁱVersion gérée : le créateur garde une instance du logiciel active sur son infrastructure afin que d'autres puissent s'en servir.

ⁱⁱVersion autonome : l'utilisateur garde une instance personnelle du logiciel sur sa propre infrastructure.