



Un autre projet mis de l'avant par **GenomeCanada**

Progiciels pour simplifier la prédiction de la fonction d'un gène

État	En cours
Concours	Développement de nouvelles technologies
Secteur	Développement de nouvelles technologies
Centre de génomique	Institut de génomique de l'Ontario
Chefs de projet	Morris Quaid et Gary Bader

Description du projet

À la fin des années 1990, Google a mis à la disposition de l'utilisateur d'Internet moyen un puissant moteur de recherche en offrant, pour la première fois, une interface intuitive et simple étayée par un logiciel d'analyse puissant et un dépôt de données imposant. En appliquant les prototypes de logiciels existants pour bâtir le « Google de la biologie », un progiciel qui permet aux utilisateurs d'utiliser toutes les données existantes en génomique et en protéomique pour répondre à des questions biologiques précises sur la fonction des gènes, nous proposons de mettre la puissance de la révolution génomique à la portée du biologiste moyen.

Jusqu'à maintenant, cette révolution a été à la fois une bénédiction et une malédiction. Même s'il est maintenant possible de rassembler des données sur tous les gènes du génome, les dépôts publics qui stockent ces données ont augmenté de façon exponentielle en volume et en complexité, ce qui fait que les biologistes ont de la difficulté à consulter les données pour les utiliser dans leurs recherches s'ils ne possèdent pas d'expertise statistique et computationnelle perfectionnée. C'est là une occasion perdue, car les analyses computationnelles nous ont montré que ces données peuvent receler d'importants aperçus du fonctionnement des cellules et des défaillances des systèmes biologiques dans les maladies. Comme peu de biologistes possèdent l'expertise d'analyse nécessaire, il faut mettre à leur disposition des outils efficaces pour consulter ces ressources et éviter ainsi que des chercheurs ne répètent des efforts déjà déployés par d'autres, ce qui ajoute beaucoup aux investissements internationaux déjà considérables dans les données en génomique et en protéomique.

Nous proposons de bâtir un progiciel qui aidera à aplanir deux obstacles qui nuisent aux progrès de la recherche en génomique et en protéomique : i) le biologiste moyen ne peut pas participer à la technologie de pointe en génomique et en tire rarement profit, et ii) la collecte et l'analyse d'ensemble de données pangénomiques sont coûteuses et énergivores. Étonnamment, des solutions semblables pourraient résoudre ces deux problèmes. Pour le biologiste moyen, nous voulons créer un site Web convivial où il pourra faire des prédictions sur la fonction des gènes en utilisant toutes les données disponibles en génomique et en protéomique. Pour le chercheur en génomique fonctionnelle, nous créerons un système d'appui aux décisions qui permettra d'évaluer différentes stratégies de collecte de données en comparant la qualité des données et le contenu nouveau en information des données

préliminaires avec des données publiées sur la génomique, la protéomique et la fonction des gènes. Ces deux outils sont basés sur la même architecture logicielle sous-jacente et nous la mettrons gratuitement à la disposition des chercheurs dans Internet. Pour nous assurer que nos progiciels sont les plus utiles à une communauté diversifiée d'utilisateurs en biologie, nous élaborerons et diffuserons de nombreuses ressources de formation des utilisateurs.

Notre site Web et le système de soutien aux décisions seront étayés par un dépôt de données en génomique et en protéomique qui seront automatiquement colligées et tenues à jour par une série d'agents logiciels explorateurs Web que nous mettrons au point. Les deux outils bénéficieront également de l'interface perfectionnée de visualisation en réseau qui aidera les utilisateurs à interroger les données et comprendre les résultats de leurs recherches.

Pour bâtir cette technologie habilitante, nous demandons du financement pour les programmeurs qui développeront ces outils, un spécialiste en formation qui rédigera des ressources pédagogiques et des serveurs qui prendront en charge notre service Web. Grâce au « Google de la biologie », nous augmentons la valeur des données en génomique et en protéomique en les rendant accessibles à toute la communauté des chercheurs en biologie. Il faudra deux ans pour bâtir et mettre à l'essai notre progiciel, mais une fois bâti, notre site Web se mettra à jour automatiquement et notre système de soutien aux décisions nécessitera peu d'entretien, de sorte que les avantages de nos travaux perdureront pendant de nombreuses années.