



## Infolettre

Édition 4

Juillet 2015

### Table des matières:

- L'application de la sélection génomique pour l'amélioration génétique des épinettes - des outils prometteurs
- Mission et objectifs du projet
- Données publiques
- Nouvelles, événements et publications

#### **L'application de la sélection génomique pour l'amélioration génétique des épinettes - des outils prometteurs**

par Patrick Lenz, Jean Beaulieu et Jean Bousquet

Les épinettes ont une grande importance écologique et économique à travers le Canada. Depuis le début des années 1950, des programmes ont été mis en place pour d'amélioration génétique de la croissance et, plus récemment, de certaines caractéristiques de bois des plants d'épinette destinés au reboisement (Mullin et al. 2011). Chaque année, plus de 300 millions de semis d'épinettes génétiquement améliorés sont plantés au Canada. Cependant, les efforts d'amélioration sont souvent ralentis par les coûts élevés et les délais très longs liés à l'évaluation de caractères matures, comme la qualité du bois, en raison de la croissance lente des épinettes en forêt boréale. Un des objectifs du projet SMarT Forests était de résoudre les problèmes inhérents à la longueur des cycles d'amélioration des épinettes et aux coûts qui y sont associés. La mise au point de nouveaux outils en génomique forestière au cours des dix dernières années et la possibilité de réaliser le génotypage d'individus pour un grand nombre de marqueurs à un coût raisonnable offrent un ensemble de nouvelles approches de sélection pour les améliorateurs.

**Les études d'association génétique :** Dans le cadre de notre précédent projet (Arborea II), nous avons mené des études d'association (corrélation entre des marqueurs génétiques et des caractéristiques du bois et de la croissance), et c'est dans SMarT Forests que nous avons étendu l'analyse à l'ensemble du génome afin de réaliser des études d'association pangénomiques (GWAS) (Beaulieu et al. 2011). Ces travaux antérieurs ont permis de découvrir quelques associations significatives entre les SNP de gènes (polymorphismes nucléotidiques simples) et les caractères de la qualité du bois. Les difficultés d'ordre statistique, à savoir les corrections nécessaires pour les tests multiples et la non-linéarité des effets des marqueurs, ont conduit à un faible nombre d'associations significatives, même si les caractères quantitatifs, tels que la croissance et la qualité du bois, sont contrôlés par un grand nombre de gènes (Beaulieu et al. 2011, Pelgas et al. 2011). Conséquemment, les marqueurs identifiés expliquent seulement une petite proportion de la variation des caractères. C'est un inconvénient majeur pour l'utilisation des résultats des études d'association dans l'amélioration des arbres, car les améliorateurs cherchent à prédire les caractères sur lesquels ils réalisent les sélections. Néanmoins, les études d'association sont utiles dans la découverte des gènes clés et le développement d'une compréhension du contrôle moléculaire des caractères complexes.

**La sélection génomique – les marqueurs utilisés pour l'amélioration des arbres :** Le petit nombre de marqueurs associés aux caractères d'intérêt entrave l'identification exacte d'individus supérieurs et nous a conduit à expérimenter la sélection génomique (Meuwissen et al. 2001, Grattapaglia et Resende 2011) pour l'amélioration génétique chez les arbres. Afin de prédire la valeur génétique des arbres dès le stade de la plantule, nous avons utilisé les profils génomiques des arbres candidats obtenus à partir de plusieurs milliers de marqueurs et un modèle multi-locus corrélant l'information phénotypique et génotypique. Deux études pilotes de la sélection génomique ont été menées chez l'épinette blanche et une autre chez l'épinette noire (voir encadré 1) dans le présent projet.



## Mission et objectifs du projet

Le projet SMarTForests s'appuie sur une décennie de découvertes scientifiques en génomique des épinettes découlant de deux projets antérieurs : [Arborea](#) (Université Laval) et [Treenomix](#) (University of British Columbia). L'équipe de recherche possède l'expérience et le savoir-faire nécessaires pour réaliser ce projet d'envergure. Notre mission est d'innover dans le séquençage du génome de l'épinette, de bien positionner le Canada au sein d'initiatives internationales touchant le génome des conifères et de réaliser un transfert efficace des résultats et des innovations aux utilisateurs finaux du secteur forestier canadien.

Le projet SMarTForests a trois grands objectifs :

1. Développer des systèmes de marqueurs moléculaires;
2. Séquençer le génome de l'épinette blanche;
3. Analyser les impacts de la génomique forestière sur l'économie et la société.

## Encadré 1

### Développement de ressources de génotypage propres à l'épinette noire (*Picea mariana*)

Au cours des 10 dernières années, les chercheurs ont développé des ressources génomiques d'envergure pour l'épinette blanche (*Picea glauca*). Cependant, un manque d'information génomique était évident pour l'épinette noire (*Picea mariana*), une espèce sympatrique à l'épinette blanche, avec un poids écologique et économique élevé dans la forêt boréale canadienne.

Avant de penser tester la sélection génomique chez l'épinette noire, il a fallu développer un catalogue de SNP (polymorphismes nucléotidiques simples) propre à cette espèce. Nous avons bâti un registre de SNP représentant l'épinette noire. Ce registre fut élaboré en compilant des SNP d'études antérieures sur l'épinette blanche (Pavy et al. 2008, Prunier et al. 2012, Prunier et al. 2013, Pavy et al. 2013a) et en découvrant de nouveaux SNP à l'aide d'un outil génomique appelé *exome capture*. Deux approches de capture (capture de la phase solide avec séquençage 454 GS-FLX-titane et capture en phase liquide avec le séquençage Illumina HiSeq) ont été utilisées avec des sondes conçues pour *Picea glauca* (Rigault et al. 2011). Ainsi, plus de 680 millions de séquences ont été générées.

Nos analyses bio-informatiques ont permis d'identifier plus de 97 000 SNP de grande qualité provenant de 21 000 gènes, avec un taux de réussite de 96% au niveau du génotypage, et dépassant de surcroît le taux de 92% obtenu pour l'épinette blanche (Pavy et al. 2013b). Près de 5 000 SNP ont été assemblés sur une puce *Illumina Infinium iSelect* qui est maintenant utilisée pour le génotypage des descendants lors des expériences de sélection génomique sur l'épinette

Une première étude menée sur l'épinette blanche (Beaulieu et al. 2014a) a été basée sur 1 694 arbres de 215 familles de demi-frères provenant de différentes sources de semences échantillonnées à travers le Québec. Deux autres études sur l'épinette blanche (Beaulieu et al. 2014b) et l'épinette noire (Lenz et al., en prép.) furent réalisées sur des populations avancées d'élevage en utilisant 1 748 et 734 arbres provenant de 59 et 34 croisements dirigés établis sur différents sites environnementaux. Nous avons génotypé des arbres dans les trois études avec respectivement 6 385, 6 932 et 4 993 marqueurs SNP à la plateforme de génotypage *Infinium iSelect* au Centre d'innovation Génome Québec. Les caractères retenus dans les trois études (croissance et autres caractéristiques citées ci-dessous) furent sélectionnés pour répondre aux besoins des améliorateurs des conifères. Les caractères tels que la densité du bois et l'angle des microfibrilles de cellulose ont été analysés à partir de carottes de bois. Ces caractères sont de bons indicateurs de la résistance mécanique et de la rigidité du bois, deux caractéristiques très recherchées par l'industrie forestière.

### La précision des modèles de prédiction génomique :

La précision des modèles de prédiction est généralement déterminée en utilisant des méthodes de validation croisée avec des groupes d'arbres indépendants et elle est estimée en comparant la valeur génétique prédite à la valeur génétique connue d'un individu dans les populations étudiées. La précision des modèles de sélection génomique était élevée pour les deux études basées sur des croisements dirigés chez l'épinette blanche et l'épinette noire et les corrélations variaient de 0,70 à 0,80. La précision des modèles basés sur les marqueurs était ainsi comparable (> 90%) à une précision obtenue par sélection classique. Cela signifie que les caractéristiques du bois et de la croissance pourraient être efficacement améliorées par la sélection génomique qui peut déjà être appliquée au stade de la plantule.

La précision des modèles de sélection génomique était « modérée » pour l'étude basée sur des familles de demi-frères (épinette blanche) et les corrélations entre la valeur

prédite et la valeur génétique connue se situaient entre 0,33 et 0,44. Les modèles basés sur les marqueurs étaient un peu moins précis que les modèles considérant l'apparentement connu ou supposé (pedigree) et ont atteint en moyenne 90% de l'amélioration prévue lors d'une sélection conventionnelle fondée sur le pedigree.

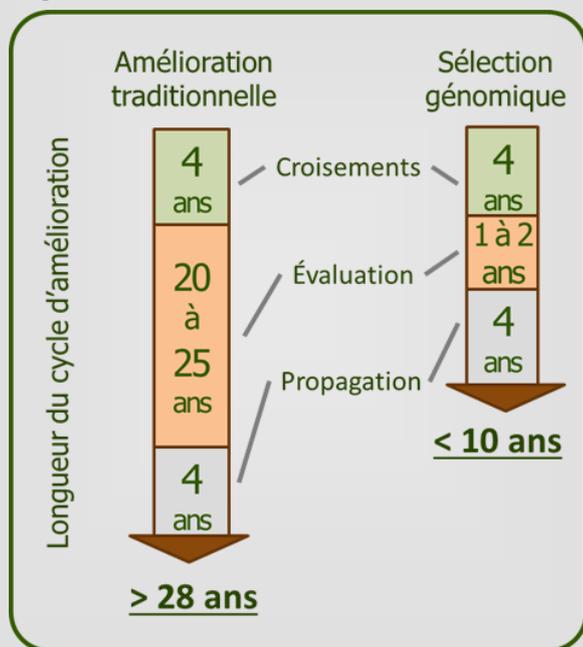
Nos résultats font ressortir à quel point la sélection génomique peut être plus rentable lorsque des croisements dirigés de génération avancée sont utilisés. À ce stade, la valeur génétique de la descendance peut être prédite avec une plus grande précision, comme indiqué par les différences dans la précision entre les résultats obtenus à partir de familles de demi-frères par rapport à ceux obtenus à partir des familles de pleins-frères. Autre avantage, lors du développement des modèles pour différents sites, nous avons constaté que la précision globale était maintenue, en particulier pour les caractères liés à la qualité du bois. Cela indique la présence de peu d'interaction génotype-environnement pour ces espèces et permet par conséquent l'application de modèles de sélection génomique dans différents environnements.

**Le rôle de l'apparentement dans la prédiction génomique :** La structure génétique des populations utilisées pour la création du modèle et celles de la validation est cruciale, tant du point de vue expérimental que d'opérationnalisation. Dans le scénario standard, les individus utilisés pour la construction du modèle et les individus utilisés pour la validation provenaient des mêmes croisements ou des mêmes familles. Il est à noter que la précision de la prédiction va baisser de façon marquée si le modèle est utilisé pour prédire la valeur génétique d'individus ne partageant aucune parenté avec les individus initialement utilisés pour la construction du modèle. Ce n'est pas surprenant, étant donné qu'une grande partie de la

puissance prédictive de la sélection génomique est basée, avec la densité de marqueurs actuelle, sur le déséquilibre de liaison sur de longues distances chromosomiques et l'apparentement. Toutefois, la construction de modèles de sélection génomique impliquant différents groupes d'élevage pourrait être possible avec seulement une faible perte de précision.

**Travaux en cours :** Nous tentons actuellement de déterminer le nombre optimal de marqueurs et d'individus nécessaires pour la construction de modèles de sélection génomique dans les populations avancées. Ceci nous aidera à trouver l'équilibre entre le besoin de précision dans nos modèles et les coûts engendrés pour les populations comprenant des milliers d'individus. Jusqu'à présent, nos résultats indiquent que le nombre de marqueurs, par conséquent le coût de génotypage, peut être réduit sans perte significative de précision, indiquant encore une fois que la majeure partie de la précision est due autant au déséquilibre de liaison qu'au lien de parenté existant entre les arbres étudiés et les populations test.

**Figure 1**



**Implications pour l'amélioration des conifères :** Nos résultats mettent en évidence la puissance de la sélection génomique pour l'amélioration génétique des arbres, en particulier pour les petites populations dans un programme d'amélioration avancé. Le principal avantage est évidemment la possibilité de sélectionner, parmi des milliers d'arbres candidats très tôt dans un cycle d'amélioration, les données de marqueurs pouvant être obtenues au stade de plantule. Il n'est plus nécessaire d'attendre d'avoir des arbres matures pour évaluer certains caractères; il est possible de le faire de façon précoce, une fois les modèles de sélection génomique calibrés. Ainsi, l'évaluation onéreuse des caractères comme la densité du bois de tous les individus deviendra, dans le futur, obsolète. Les cycles d'amélioration peuvent être réduits de deux tiers (voir la figure 1), passant de  $\pm 28$  ans à  $\pm 10$  ans, et le gain génétique obtenu par année grâce à une sélection de marqueurs est multiplié par un facteur similaire, toujours par rapport à l'amélioration conventionnelle faite sur des arbres matures. De plus, les modèles de sélection pourraient facilement être calibrés pour d'autres caractères d'intérêt dans l'avenir, les génotypes étant déjà obtenus, ce qui engendre un avantage opérationnel de long terme aux programmes de sélection. Toute diminution en temps du cycle d'amélioration se traduit ultimement par des économies de coûts pour les améliorateurs.

#### References:

- Beaulieu J, Doerksen T, Boyle B, Clement S, Deslauriers M, Beauseigle S, Blais S, Poulin PL, Lenz P, Caron S, Rigault P, Bicho P, Bousquet J, MacKay J (2011). Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. *Genetics*, 188: 197-214.
- Beaulieu J, Doerksen T, Clément S, MacKay J, Bousquet J (2014a). Accuracy of genomic selection models in a large population of open-pollinated families in white spruce. *Heredity*, 113: 342-352.
- Beaulieu J, Doerksen TK, MacKay J, Rainville A, Bousquet J (2014b). Genomic selection accuracies within and between environments and small breeding groups in white spruce. *BMC Genomics*, 15: 1048.
- Grattapaglia D, Resende MDV (2011). Genomic selection in forest tree breeding. *Tree Genetics & Genomes*, 7:241-255.
- Meuwissen THE, Hayes BJ, Goddard ME (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 157:1819-1829.
- Mullin TJ, Andersson B, Bastien J-C, Beaulieu J, Burdon RD, Dvorak WS, King JN, Kradowski J, Lee SJ, McKeand SE, Pâques L, Raffin A, Russel JH, Skroppa T, Stoehr M, Yanchuk A. (2011). Economic importance, breeding objectives and achievements, Chapter 2. In: Plomion C, Bousquet J, Kole C (eds), *Genetics, Genomics and Breeding of Conifers*. CRC Press and Science Publishers Inc., New York, pp 40-127.
- Pavy N, Pelgas B, Beauseigle S, Blais S, Gagnon F, Gosselin I, Lamothe M, Isabel N, Bousquet J (2008). Enhancing genetic mapping of complex genomes through the design of highly-multiplexed SNP arrays: application to the large and unsequenced genomes of white spruce and black spruce. *BMC Genomics*, 9: 21-38.
- Pavy N, Gagnon F, Rigault P, Blais S, Deschênes A, Boyle B, Beaulieu J, Isabel N, Mackay J, Bousquet J (2013a). Development of high-density SNP genotyping arrays for white spruce (*Picea glauca*) and transferability to subtropical and nordic congeners. *Molecular Ecology Resources*, 13: 324-336.
- Pavy N, Deschênes A, Blais S, Lavigne P, Beaulieu J, Isabel N, Mackay J, Bousquet J (2013b). The landscape of nucleotide polymorphisms among 13,500 genes of the conifer *Picea glauca*, relationships with functions, and comparison with *Medicago truncatula*. *Genome Biology and Evolution*, 5: 1910-1925.
- Pelgas B, Bousquet J, Meirmans PG, Ritland K, Isabel N (2011). QTL mapping in white spruce: gene maps and genomic regions underlying adaptive traits across pedigrees, years and environments. *BMC Genomics*, 12: 145.
- Prunier J, Gérardi S, Laroche J, Beaulieu J, Bousquet J (2012). Parallel and lineage-specific molecular adaptation to climate in boreal black spruce. *Molecular Ecology*, 21: 4270-4286.
- Prunier J, Pelgas B, Gagnon F, Despons M, Isabel N, Beaulieu J, Bousquet J (2013). Genomic architecture and association genetics of adaptive characters using a candidate SNP approach in boreal black spruce. *BMC Genomics*, 14: 368.
- Rigault P, Boyle B, Lepage P, Cooke JEK, Bousquet J, MacKay JJ (2011). A white spruce gene catalogue for conifer genome analyses. *Plant Physiology*, 157:14-28.

## Données publiques publiées par SMarTForests:

| Type de données  | Description   | Citation  |
|--|---|---|
| White Spruce ( <i>Picea glauca</i> ) PG29 4th draft assembly                         | Fourth draft assembly of <i>Picea glauca</i> (PG29) is available ( <a href="http://www.ncbi.nlm.nih.gov/bioproject/83435">http://www.ncbi.nlm.nih.gov/bioproject/83435</a> )<br><br>The data was generated using Illumina Hiseq technology.       | Warren, R., Keeling, C.I., Yuen, M.M.S., Raymond, A., Taylor G.A., Vandervalk, B.P., Mohammadi, H., Paulion D., Chiu R., Jackman S.D., Roberson, G., Yang, C., Hoffmann, M., Weigel, D., Nelson, D.R., Ritland, C., Isabel, N., Jaquish, G., Yanchuk, A., Bousquet, J., Jones, S.J., MacKay, J., Birol, I. and Bohlmann J. July 2015.<br>Improved white spruce ( <i>Picea glauca</i> ) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. The Plant Journal Vol 82 Issue 6 pp.xxx (currently online: <a href="http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291365-313X/earlyview">http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291365-313X/earlyview</a> ) |
| White Spruce ( <i>Picea glauca</i> ) WS77111 1 <sup>st</sup> draft assembly          | Revised first draft assembly of <i>Picea glauca</i> (WS77111) is available ( <a href="http://www.ncbi.nlm.nih.gov/bioproject/242552">http://www.ncbi.nlm.nih.gov/bioproject/242552</a> ). The data was generated using Illumina Hiseq technology. | Warren, R. Keeling, C.I., Yuen, M.M.S., Raymond, A., Taylor G.A., Vandervalk, B.P., Mohammadi, H., Paulion D., Chiu R., Jackman S.D., Roberson, G., Yang, C., Hoffmann, M., Weigel, D., Nelson, D.R., Ritland, C., Isabel, N., Jaquish, G., Yanchuk, A., Bousquet, J., Jones, S.J., MacKay, J., Birol, I. and Bohlmann J. July 2015.<br>Improved white spruce ( <i>Picea glauca</i> ) genome assemblies and annotation of large gene families of conifer terpenoid and phenolic defense metabolism. The Plant Journal Vol 82 Issue 6 pp.xxx (currently online: <a href="http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291365-313X/earlyview">http://onlinelibrary.wiley.com/journal/10.1111/%28ISSN%291365-313X/earlyview</a> )  |
| Analyse transcriptomique du bois juvénile et mature dans quatre espèces de conifères | Analyse transcriptomique, no d'accèsion GEO GSE51884, soumis pour publication. Les données deviendront publiques au moment de la publication.   | Beaulieu, J., Doerksen, T., Boyle, B., Clément, S., Deslauriers, M., Beauseigle, S., Blais, S., Poulin, P-L., Lena, P., Caron, S., Rigault, P., Bicho, P., Bousquet, J. and MacKay J. May 2011. Association genetics of wood physical traits in the conifer white spruce and relationships with gene expression. Genetics Vol 188 Issue 1 pp. 197-214.  |
| Analyse comparative du transcriptome de sept tissus de <i>Picea glauca</i>           | Analyse transcriptomique, no d'accèsion GEO GSE60277, soumis pour publication. Les données deviendront publiques au moment de la publication.   | Raherison, E., Rigault, P., Caron, S., Poulin, P-L., Boyle, B., Verta, J-P., Giguère, Is., Bomal, C., Bohlmann, J. and MacKay, J. 2012. Transcriptome profiling in conifers and the PiceaGenExpress database show patterns of diversification within gene families and interspecific conservation in vascular gene expression. BMC Genomics Vol 13 pp. 434-450/   |
| Pour toutes les données publiques  | SMarTForests ( <a href="http://www.smartforests.ca/en-ca/publications.aspx">http://www.smartforests.ca/en-ca/publications.aspx</a> )  |   |

VISITEZ-NOUS À:  
[www.smartforests.ca](http://www.smartforests.ca)

POUR NOUS  
CONTACTER:

GESTIONNAIRES  
DE PROJET

**Sophie Laviolette**  
Centre d'étude de la forêt  
Université Laval  
1030 av. de la Médecine  
Bureau 2153  
Québec (Qc), G1V 0A6  
Canada  
Téléphone: 418.656.2408  
Télécopieur: 418.656.7176  
Courriel:  
[Sophie.laviolette@sbf.ulaval.ca](mailto:Sophie.laviolette@sbf.ulaval.ca)

**Carol Ritland**  
#3611-2424 Main Mall  
Department of Forest  
Sciences  
University of British Columbia  
Vancouver, B.C. Canada  
V6T 1Z4  
Téléphone: 604.822.3908  
Télécopieur: 604.822.2114  
Courriel:  
[critland@mail.ubc.ca](mailto:critland@mail.ubc.ca)

## Nouvelles et événements récents

### Fin de SMarTForests

Le projet SMarTForests est terminé officiellement depuis le 30 juin 2015, après quatre années d'activités scientifiques. Nous tenons à remercier les membres de notre équipe, les organismes subventionnaires, les institutions d'enseignement impliquées ainsi que tous nos collaborateurs. Nous sommes très fiers du travail accompli et des avancées scientifiques réalisées au cours des dernières années. Notre réussite est une affaire d'équipe! Suivez-nous au cours des prochains mois, de nouvelles initiatives découlant du projet SMarTForests seront annoncées!

### Prix remarquable

Le SIB (Swiss Institute of Bioinformatics) a décerné son prix *SIB International Bioinformatics Resource Award 2015* au Dr Inanç Birol et son équipe pour **ABYSS** (Assembly By Short Sequences).

### Nouveaux articles (pour la liste complète, visitez [www.smartforests.ca](http://www.smartforests.ca))

- Warren, R.L., Kelling, C.I., Yuen, M.M.S., Raymond, A., Taylor, Greg A., Vandervalk, B.P., Mohamadi, H., Paulino, D., Chiu, R., Jackman, S.D., Robertson, G., Yang, C., Boyle, B., Hoffmann, M., Weigel, D., Nelson, D.R., Ritland, C., Isabel, N., Jaquish, B., Yanchuk, A., Bousquet, J., Jones, S.J.M., MacKay, J., Birol, I., Bohlmann, J., 2015. Improved white spruce (*Picea glauca*) genome assemblies and annotation of large gene families of conifer defense metabolism. *The Plant Journal* (<http://doi.org/10.1111/tpj.12886>)
- Raheison, E.S.M., Giguère, I., Caron, S., Lamara, M., MacKay, J., 2015. Modular organization of the white spruce (*Picea glauca*) transcriptome reveals functional organization and evolutionary signatures. *New Phytologist* (<http://doi.org/10.1111/nph.13343>).
- Porth, I., Bull, G., Ahmed, S., El-Kassaby, A.Y., Boyland, M., 2015. Forest Genomics Research and Development in Canada: Priorities for Developing an Economic Framework. *The Forestry Chronicle* (<http://doi.org/10.5558/tfc2015-011>)
- Mageroy, M.H., Parent, G., Germanos, G., Giguère, I., Deltas, N., Maaroufi, H., Bauce, E., Bohlmann, J., MacKay, J., 2015. Expression of the beta-glucosidase gene Pgbglu-1 underpins natural resistance of white spruce against spruce budworm. *The Plant Journal* (<http://doi.org/10.1111/tpj.12699>).

Continuez à suivre notre site Web, plusieurs articles seront publiés au cours des mois à venir!

## Événements à venir

### 3e Sommet sur les génomes des conifères

(28 au 30 septembre 2015 à Gysinge, en Suède)

Pour inscription à l'événement:

<http://login3.axaco.se/C12571F400166B19/registrationForm?openagent&unid=4F49A0412A5778A4C1257E1400560261>



## Un grand merci à nos organismes subventionnaires



GenomeCanada



GenomeQuébec



GenomeBritishColumbia



GenomeAlberta