

Nouvel algorithme pour la sélection de marqueurs géniques robustes du cancer

RÉSUMÉ

Une tendance croissante se manifeste à recourir à des biomarqueurs pour aider les cliniciens à poser des diagnostics et à traiter diverses maladies. Cependant, la détermination, la sélection et la validation de marqueurs valables et précis comportent encore leur lot de difficultés. Un algorithme nouveau et unique a été mis au point afin de produire des marqueurs géniques du cancer à partir des données de biopuces. Cet algorithme, a permis de générer plusieurs ensembles de biomarqueurs robustes qui présentent une corrélation avec la survie dans des cas de cancer du sein.

APPLICATIONS

- Mise au point de tests diagnostiques pour les patientes atteintes de cancer du sein.
- Détermination et validation de groupes de biomarqueurs propres à différents types ou sous-types de cancers.
- Mise au point de trousseaux d'analyse pour le pronostic d'un ou de plusieurs types de cancer.

CONCEPT

À l'heure actuelle, il est très difficile de déterminer les modalités de traitement les plus efficaces et les plus sûres pour chaque patient atteint de cancer. Les biomarqueurs apparaissent de plus en plus comme une solution face à ce problème. Toutefois, la détermination de biomarqueurs du cancer est limitée par plusieurs facteurs. En raison de l'hétérogénéité des tumeurs, il est difficile de trouver des biomarqueurs robustes. En plus de la variabilité individuelle des profils d'expression génique, les signatures géniques « réelles » sont enfouies dans les profils géniques globaux des tumeurs. De surcroît, le nombre d'ensembles d'échantillons de tumeurs pour la découverte de marqueurs est très faible. Devant ces difficultés, une nouvelle stratégie pour la recherche de biomarqueurs a été développée. En premier lieu, la démarche consiste à trouver des ensembles de gènes dont l'expression est corrélée avec une issue sur le plan de la survie, en effectuant à l'aide de biopuces un profilage du génome d'un ensemble d'échantillons de tissus cancéreux (« ensemble d'apprentissage initial ») provenant de patients chez qui l'issue clinique était connue. Dans un deuxième temps, on génère 1 million d'ensembles aléatoires de gènes à partir des ensembles de gènes choisis susmentionnés, et 34 ensembles aléatoires d'échantillons à partir de l'ensemble d'apprentissage initial ou de sous-ensembles. En dernier lieu, on détermine les ensembles de biomarqueurs optimaux en examinant, pour chaque ensemble aléatoire de gènes, la corrélation de l'issue sur le plan de la survie des patients cancéreux dans chaque ensemble aléatoire d'échantillons, et en retenant et en classant les ensembles aléatoires de gènes ayant donné les meilleurs résultats. Le

cancer du sein a servi de point de départ pour évaluer et valider la stratégie. Après avoir appliqué cet algorithme à deux ensembles d'apprentissage, neuf ensembles de marqueurs renfermant chacun 30 gènes ont été définis. NRC1 à NRC6 sont associés à des tumeurs exprimant le récepteur œstrogénique (ER+) tandis que NRC7 à NRC9 sont associés à des tumeurs ne l'exprimant pas (ER-). La robustesse des neuf ensembles de marqueurs de mauvais pronostics ont été validés au moyen de trois ensembles de test indépendants.

CARACTÉRISTIQUES ET AVANTAGES

Marqueurs validés pour le cancer du sein affichant une bonne performance pronostique

Les biomarqueurs du cancer du sein utilisés actuellement ont généralement une faible valeur prédictive lorsqu'ils sont mis à l'épreuve sur des échantillons autres que ceux qui ont servi à les définir. La robustesse des groupes de marqueurs sélectionnés a été validée à l'aide de trois ensembles d'échantillons (> 650) indépendants de cancers du sein. Les ensembles NRC1 à NRC6 ont permis de stratifier les sujets ER+ en groupes à risque faible, moyen ou élevé, tandis que les ensembles NRC7 à NRC9 ont permis de stratifier les sujets ER- en groupes à faible risque ou à risque élevé. Cette stratification était hautement significative dans les trois ensembles de données. En outre, des taux de prédiction élevés ont été obtenus; chez les groupes à faible risque l'exactitude était de 90 % pour les sujets ER+, et de 92 à 100 % pour les sujets ER- (voir le tableau).

Marqueurs particuliers pour le cancer du sein présentant un large spectre clinique

Contrairement à d'autres marqueurs servant au pronostic dans les cas de cancer du sein qui ne s'appliquent qu'à des groupes choisis de patientes, ces ensembles de marqueurs peuvent être utilisés chez toutes les patientes atteintes du cancer du sein (c.-à-d. NRC1-6 chez les sujets ER+ et NRC7-9 chez les sujets ER-). De plus, les listes de gènes des ensembles de marqueurs n'ont pas été signalées par d'autres.

Algorithme applicable à une grande échelle

Les biomarqueurs offrent des perspectives intéressantes comme outils de diagnostic et de pronostic pour l'avenir. Si l'on considère que la plupart des algorithmes mis au point jusqu'à présent n'ont pas été conçus pour trouver des marqueurs de sous-types de cancer, ce nouvel algorithme est une occasion privilégiée d'obtenir les biomarqueurs de pronostic nécessaires pour mieux classer les patients à l'intérieur du vaste éventail de types de cancer.

Meilleure prise en charge des patients atteints de cancer

Le fait de pouvoir prédire plus précisément l'issue du cancer aura certainement des répercussions importantes sur les soins de santé. Les biomarqueurs utilisés aux fins de pronostic permettront aux médecins de choisir le traitement le plus efficace pour un patient donné en déterminant, par exemple, s'il pourrait tirer avantage ou non d'une chimiothérapie additionnelle. Par ailleurs, cette capacité pourrait améliorer la durée de survie et la qualité de vie chez les patients et se traduire par une réduction des coûts médicaux.

PROPRIÉTÉ INTELLECTUELLE

Ensembles de marqueurs pour le pronostic dans les cas de cancer du sein et méthode pour les établir (CNRC n° 12020).

Capacité de prédire avec exactitude la non-récurrence des tumeurs du sein chez les groupes à faible risque établie au moyen des ensembles de marqueurs NRC1 à 9

Ensemble de marqueurs	Ensemble d'apprentissage	Ensemble de test		
		1	2	3
NRC1-6 (ER+)	89 %	94 %	91 %	87 %
Nb d'échantillons ER +	209	173	74	201
NRC7-9 (ER-)	93 %	100 %	92 %	100 %
Nb d'échantillons ER -	77	46	43	31

CONTACTS

Yves Quenneville

Tél. : (514) 496-8507

Agent de développement des affaires

Courriel : yves.quenneville@cnrc-nrc.gc.ca

Daniel Desmarteaux

Tél. : (514) 496-5300

Agent de développement des affaires

Courriel : daniel.desmarteaux@cnrc-nrc.gc.ca

Edwin Wang

Tél. : (514) 496-0914

Groupe Chimie et biologie computationnelles

Courriel : edwin.wang@cnrc-nrc.gc.ca