

Discovery of metabolic gene mutations causing intellectual delay

Wyeth W. Wasserman

David Aubert

Ahmed Rafik

University of Montpellier

Bio-informatique

05 jan 2015

Le Dr. Wasserman est à la tête d'un groupe travaillant sur la partie calcul et analyse du génome.

Aujourd'hui il existe quelques domaines qui occupent une place importante du monde de la recherche, à cause de leurs importances et de leurs complexités et qui font régulièrement appel aux centres de calculs haute performances. L'étude du génome en fait partie.

Aujourd'hui beaucoup des maladies difficilement curables sont dûes à des problèmes de mutations et de génétique (Trisomie, cancer etc...). Ces problèmes trouvent leurs sources dans l'ADN et/ou son expression.

A l'heure actuelle, le séquençage du génome pose des problèmes à la recherche et ralenti son avancée.

1 Medical Subject Heading Over-representation Profiles

1.1 Définition

MeSHOP est une application très complète. Celle-ci permet de centraliser les recherches et d'indexer les publications scientifiques. Celle-ci englobe des thématiques allant de la recherche génétique à la chimie en passant par les maladies, ce qui intéresse l'auteur.

1.2 fonctionnement

Le moteur de cette application repose sur une base de données MEDLINE et le moteur de recherche PubMed.

1.2.1 Annotations

Chaque article devra être annoté avec un vocabulaire particulier. En effet, c'est cette annotation qui permettra une recherche efficace. Le système MeSH a une liste de vocabulaire (*PMIDs*) qui servira d'index pour le fichier.

1.3 Editer un article

Chaque mot sera classé en fonction de son nombre d'ocurrence dans les *PMIDs*

Un point important est la thématique recherchée. En effet, comme vu précédemment, les sujets d'articles englobent un grand nombre de thématique (génétique, maladies, chimie etc...) et donc le contexte peut changer pour des mêmes mots ou combinaisons de mots.

2.1 Définition

2.2 allèles préférées

Mais comme les allèles fortes s'expriment plus que les allèles récessives cela peut être dû à plusieurs raisons mais la plus commune est que, par sélection naturelle, un gène subira

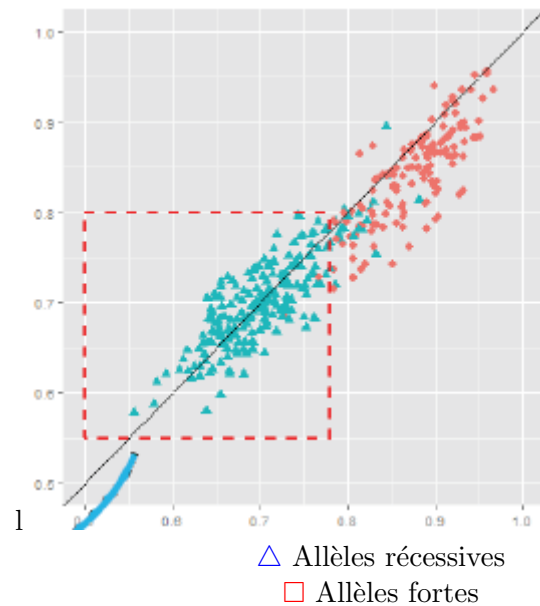


Figure 2.1: *Allèle score*

une mutation, et aura un nombre supérieur de site de transcriptions que le précédent et ainsi pourra se transmettre plus facilement de génération en génération. De ce fait l'étude à de meilleures chances de résultats si il est effectué sur des gènes forts.

Pour trouver les places de liaisons de transcription, une unité a été créée: **PWM**: La matrice de position de poids.

Il s'agit d'une matrice où sont marquées les probabilités de chaque nucléotide d'apparaître dans le site de liaison pour la transcription, ainsi que sa position.

Ceci a une importance particulière car grâce à ces matrices, il a été possible de remarquer que les liaisons se faisaient plus sur les allèles récessives que les autres.

Mais ces travaux sont ralentis par le fait qu'il peut exister de multiples facteurs qui modifient le taux de transcription d'un site comme les mutations, l'hérédité et l'épigénétique.

2.3 Recherche de site de liaison pour la transcription pour les lymphomes

C'est ici que la présentation recoupe avec les travaux du Dr. Wasserman.

En effet, grâce aux outils présentés précédemment, beaucoup de données ont été recueillies, notamment des échantillons d'ADN, d'ARN de malades atteints de lymphomes.

Ensuite certaines zones du génomes ont été ciblés, et une des premières remarques est que les sites de transcriptions ont eu des taux de mutations plus élevé comparés aux séquences saines.

La mort des cellules, à un rôle dans ce fonctionnement. En effet, la mort des cel-

lules est régulé d'une manière naturel et saine mais si cette régulation est perturbée cela peut causer des problèmes, dans le cas d'une diminution et dans le cas d'une augmentation un lymphome, par exemple. Une augmentation des cellules (ou plutôt une non décroissance normale du nombre de cellules) augmentent le nombre de site de liaison augmente.