# Discovery of metabolic gene mutations causing intellectual delay

Wyeth W. Wasserman

David Aubert
Ahmed Rafik
University of Montpellier
Bio-informatique

05 jan 2015

## 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. Celli ci englobe des thématiques allant de la recherche génétique à la chimie en passant par les maladies, ce qui interesse 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

Après l'écriture d'un article, il est nécessaire de le publier. Et c'est à ce moment là que va servir le système d'annotation. En effet une lecture par un robot de l'article sera fait et permettra, en fonction du vocabulaire utilisé, de classer l'article dans des catégories et de l'indexer pour des recherches.

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



Figure 1.1: exemple d'un nuage de mot

#### 1.4 Recherche

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 Site de liaison d'allèle spécifique

#### 2.1 Définition

Pour commencer un transcription, une protéïne est nécessaire. Celle-ci doit se fixer sur l'hélice. Une liaison d'allèle spécifique est ce même phénomène mais avec une protéine qui se liera d'avantages sur les allèles récessives.

#### 2.2 allèles préférées

Une des premières observations faites est que sur les individus hétérozygotes (ayant deux allèles différentes) on remarque que certaines allèlles créent plus de sites de connections que une autre, alors que il était possible de s'attendre à une distribution equi-probable.

Mais comme les allèles fortes s'expriment plus que les allèles récessives celà peut être dû à plusieurs raisons mais la plus commune est que, par selection naturelle, un gène subira 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 meilleurs chances de résultats si il est éffectué sur des gènes forts.

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

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

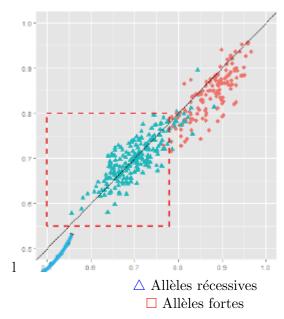


Figure 2.1: Allèle score

Ceci a une importance particulière car grâce à ces matrices, il a été possible de remarquer que les liaisons ce 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'epigenetique.

### 2.3 Recheche 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 recueillis, 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 cellules est régulé d'une manière naturel et saine mais si cette régulation est perturbée celà 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.