

Analyses de séquences ARN dans l'étude de l'hypertrophie cardiaque chez la souris

Objectif

Le but de notre projet est de voir s'il existe des différences dans l'expression des gènes au niveau du cœur lorsque différents traitements sont appliqués. Pour les traitements, nous avons un contrôle, un β -agoniste (l'isoproterenol) qui augmente la contraction du cœur et un β -bloqueur (l'atenolol) qui au contraire va la diminuer.

Design de l'expérience

Nous avons 160 transcriptomes de tissus de cœur provenant de 18 familles de souris, divisées en trois catégories, les contrôles (CTL), les isoproterenol (ISO) et les atenolol (ATE).

Analyses

Une fois en possession de notre matrice de données, nous avons dû la normaliser en utilisant la formule suivante : $\text{Log}(x + 1)$ avec $x = \text{niveau d'expression}$.

Pour avoir une vue d'ensemble des données, nous avons fait dans un premier temps des heatmaps, que nous avons ensuite réorganisés en groupant les transcriptomes les plus semblables. Nous avons ensuite fait une corrélation de l'expression des gènes pour visualiser si nos données étaient cohérentes en vérifiant que les trois réplicats soient bien corrélés avec eux-mêmes.

Dans un second temps, nous avons effectué une analyse en composante principale (PCA), qui nous permet de résumer l'information contenue dans les variables avec un minimum de pertes. Sur cette PCA, nous avons pu observer une différence entre les deux traitements, le traitement ISO ayant un effet beaucoup plus prononcé que le traitement ATE par rapport au contrôle. Bien sûr, il ne s'agit encore que d'une visualisation des données, et que cela n'a aucune valeur statistique.

Concernant l'analyse statistique, nous avons d'abord procédé sans tenir compte des souches. Nous avons comparé les deux traitements au contrôle, et comme observé sur la PCA, nous avons plus de gènes exprimés différemment pour le traitement ISO (1162 pour ISO et 17 pour ATE). Le seuil de significativité utilisé a été de 0.01 pour la p-value avec la correction de Bonferroni. Nous voyons que dans les deux groupes, nous avons un gène en commun qui est exprimé différemment dans les deux traitements. Il s'agit du gène ENSMUSG00000046709 qui code pour Mapk10 (Mitogen-activated protein kinase 10) et qui est entre autre exprimé dans le cœur.

Nous avons ensuite fait une analyse bioinformatique sur la « gene ontology ». Pour le traitement ATE, rien de concluant n'en est ressorti ; tandis que pour le traitement ISO, nous avons des groupes de gènes qui ont un rôle dans la liaison au calcium, dans le développement des muscles striés, dans la myofibrilles,... et donc qui jouent un rôle dans la fonction musculaire, ce qui nous donne une piste pour de plus amples études.

Pour l'analyse en tenant compte des souches, pour le traitement ISO, 13 souches sont exprimées différemment contre 9 pour le traitement ATE. Dans ces souches, 3 ont les mêmes gènes réprimés lors des deux traitements :

Souche	Gène	Infos
BALB_cByJ	ENSMUSG00000044709	Protéine exprimée dans le noyau Rôle dans le splicing des pre-ARN
LP_J	ENSMUSG00000025195	Dynamin binding protein
SWR_J	ENSMUSG00000030275 ENSMUSG00000059474	Ethanolamine kinase 1

Nous trouvons aussi, lorsque nous augmentons le seuil de significativité à 0.05, deux souches qui sont plus sensibles aux traitements :

Souche	DE
BALB_cJ	14 gènes exprimés différemment avec ATE et 12 avec ISO
NOD_ShiLtJ	13 gènes exprimés différemment avec ISO

Discussion des résultats

Si nous comparons les résultats des deux différentes analyses, nous voyons que nous obtenons plus de résultats significatifs lorsque nous ne tenons pas compte des souches, venant du fait que nous comparons deux groupes beaucoup plus grands (54 individus comparés avec 54 individus). Logiquement, plus l'effectif est grand, plus les chances d'avoir des résultats significatifs sont grandes. Par conséquent, les p-values en ressortiront plus basses. Lorsque l'on tient compte des souches, les comparaisons sont faites entre des groupes de 3 individus, ce qui fait que l'on a moins de résultats significatifs.

Perspectives

Il faudrait encore plus approfondir les recherches bioinformatiques, plusieurs gènes n'étant pas connus dans les bases de données utilisées, ou encore certains gènes ayant des résultats qui ne sont pas attendus ou cohérents dans le contexte de notre projet.

Des études sur les gènes qui se recoupent entre les comparaisons seraient aussi très intéressantes à envisager, pour voir si spécifiquement ces gènes jouent un rôle primordial dans la réponse à un traitement précis. Ceci pourrait aussi être étendu aux deux souches trouvées qui sont plus sensibles aux traitements.

Il serait également pertinent d'étudier l'effet des deux traitements utilisés ici sur le phénotype de la souris, c'est-à-dire la réponse du cœur par rapport à la substance utilisée.