

UNIVERSITÉ GRENOBLE-ALPES

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE-ALPES

Spécialité : **Modèles, méthodes et algorithmes en biologie, santé et environnement**

Arrêté ministériel : ?

Présentée par

Thomas Karaouzene

Thèse dirigée par **Pierre Ray**

Thèse co-dirigée par **Nicolas Thierry-Mieg**

préparée au sein du laboratoire
et de l'école doctorale "**Ingénierie de la Santé, de la Cognition et Environnement**" (EDISCE)

Écrire le titre de la thèse ici

Thèse soutenue publiquement le 31 octobre 2017,
devant le jury composé de :

Préface

This is an example of a thesis setup to use the reed thesis document class (for LaTeX) and the R bookdown package, in general.

Table des matières

Chapitre 1 : Delete line 6 if you only have one advisor	1
Remerciements	3
Résumé	5
Chapitre 2 : Introduction	7
Chapitre 3 : Investigation génétique et physiologique de la globo- zoospermie	9
Chapitre 4 : Mise en place d’une stratégie pour l’analyse des données exomiques – application en recherche clinique	11
4.1 Intro	11
4.2 Résultats	12
4.2.1 Description de la pipeline	12
4.2.2 utilisations de la pipeline pour l’identifications de variants pathogènes et l’identification de nouveaux gènes impliqués dans l’infertilité	12
Etude familiale MMAF -> DNAH1	13
Etude familiale echec de fécondation -> PLCzeta	14
Etude familiale azoospermie : SPINK2	15
Etude d’une large cohorte de patients MMAF	19
Chapitre 5 : MutaScript	21
Conclusion	23
Chapitre 6 : The First Appendix	25
References	27

Liste des tableaux

4.1	Comptage des variants communs à B1 et B2	16
4.2	Liste des variants ayant passé l'ensemble des filtres	17

Table des figures

4.1	Arbre généalogique des deux frères azoospermes B1 et B2	15
4.2	Comptage des variants retrouvés sur les frères B1 et B2 avec leur génotypes associés	16
4.3	Expression du gène *SPINK2* dans plusieurs tissus	17
4.4	Représentation du gène *SPINK2*	18

Chapitre 1

Delete line 6 if you only have one advisor

Remerciements

Résumé

Chapitre 2

Introduction

Chapitre 3

Investigation génétique et physiologique de la globozoospermie

Chapitre 4

Mise en place d'une stratégie pour l'analyse des données exomiques – application en recherche clinique

4.1 Intro

Dans ce chapitre, je détaillerai les résultats de 4 articles dont je suis coauteur :

1. **Whole-exome sequencing of familial cases of multiple morphological abnormalities of the sperm flagella (MMAF) reveals new DNAH1 mutations** : [todo]
2. **Homozygous mutation of PLCZ1 leads to defective human oocyte activation and infertility that is not rescued by the WW-binding protein PAWP** : Dans cet article j'ai, comme précédemment, effectué l'intégralité des analyses bioinformatiques des données d'exomes effectués sur deux frères infertiles présentant des échecs de fécondation.
3. **SPINK2 deficiency causes infertility by inducing sperm defects in heterozygotes and azoospermia in homozygotes** : Dans cet article j'ai effectuer non seulement l'intégralité des analyses bioinformatiques des données d'exomes de deux frères infertiles présentant un phénotype d'azoospermie mais aussi séquencer en Sanger les séquences codantes du gène *SPINK2* pour une parie des 611 individus analyser ainsi que contribué à l'extraction de l'ARN testiculaire des souris pour l'analyse fonctionnelle du gène *Spink2* sur le modèle murin.
4. **** : [todo]

4.2 Résultats

4.2.1 Description de la pipeline

4.2.2 utilisations de la pipeline pour l'identification de variants pathogènes et l'identification de nouveaux gènes impliqués dans l'infertilité

Etude familiale MMAF -> DNAH1

Etude familiale echec de fécondation -> PLCzeta

Etude familiale azoospermie : SPINK2

Description de la cohorte familiale Comme nous avons pu le voir, l'azoospermie est un phénotype d'infertilité masculine caractérisé par l'absence de spermatozoïde dans l'éjaculat. En 2017, notre équipe a publié des résultats liant une mutation du gène *SPINK2* à ce phénotype. Pour cette étude, nous avons effectué un séquençage exomique de deux frères azoospermes (B1 et B2) issus d'un union consanguin puisque leurs parents sont cousin au deuxième degré (**Figure : 4.1**). Étant donné l'historique consanguin de cette famille nous avons émis l'hypothèse d'un phénotype à transmission autosomique récessif.

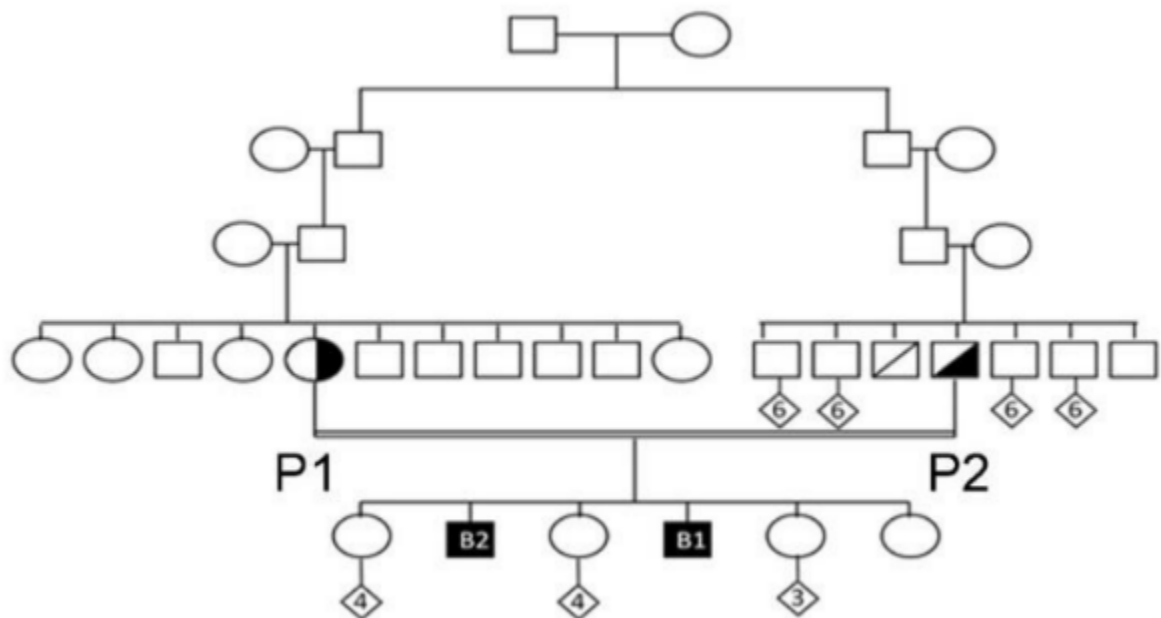


Figure 4.1 – Arbre généalogique des deux frères azoospermes B1 et B2 : Sur cet arbre nous pouvons observer la parenté des parents (P1 et P2), cousins au deuxième degré

Séquençage WES et analyse bioinformatique Le séquençage exomique des deux frères B1 et B2 fut réalisé au Mount Sinai Institutes en [TODO : mettre l'année du séquençage]. Suite à cela nous avons, comme pour les études précédentes nous avons appliqué notre pipeline d'analyse nous permettant de recenser un total de 88546 variants différents pour B1 et 66521 pour B2 (**Figure : 4.2**). Ayant émis l'hypothèse d'une cause génétique commune expliquant le phénotype des deux frères, nous nous sommes concentrés uniquement sur les variants homozygotes communs aux deux frères réduisant ainsi notre liste de variant à un total de 27604 variants dont 26718 SNVs et 886 indels (**Table : 4.1**). À partir de cette liste, nous avons soustrait l'ensemble des variants prédit par VEP comme n'impactant pas un transcrypt codant pour une protéine. De même, nous avons filtré l'ensemble des variants présents

Table 4.1 – Comptage des variants communs à B1 et B2

Variant type	Genotype	Count
SNV	Heterozygous	27056
SNV	Homozygous	26718
Indel	Heterozygous	874
Indel	Homozygous	886

dans les séquences UTRs ou encore ceux causant une substitution synonymes ou une substitution faux-sens prédite “benign” par PolyPhen2 et “tolerated” par SIFT. Suite à cela, nous avons filtrer les variants fréquents dans la population générale en filtrant l’ensemble des variant ayant une $MAF \geq 0.01$, de même, nous avons confrontés les variants restant à une base de donnée interne ressentant les variant de 83 individus non atteint d’azoospermie afin de filtrer les variants homozygotes retrouvés dans cette basse de données de contrôle.

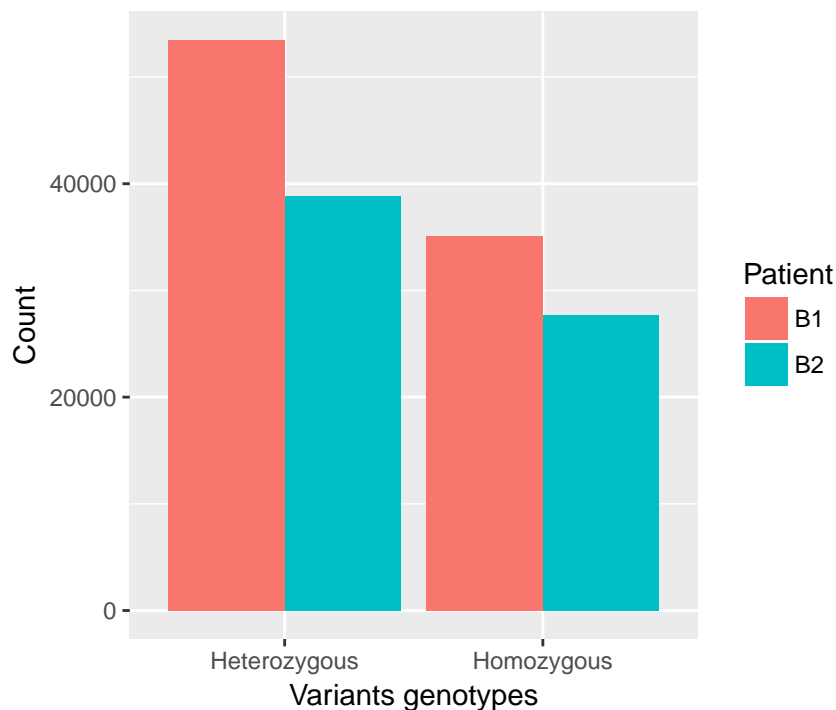


Figure 4.2 – Comptage des variants retrouvés sur les frères B1 et B2 avec leur génotypes associés

Après avoir appliqué l’ensemble de ces filtres, nous sommes arrivés à une liste de 2 variants impactant 2 gènes différents : *SPINK2* et *GUF1* (**Table** : 4.2). Parmi ces deux gènes, seul *SPINK2* était décrit comme fortement exprimé dans le testicule [INS2RER FIGURE ACEVIEW]. Nous avons d’ailleurs pu confirmer cette forte expression par RT-PCR quantitative en temps réel forte à la fois chez l’Homme (**Figure** : 4.3 - **A**) et chez la souris (**Figure** : 4.3 - **B**). Ces données ont donc fait de *SPINK2* le seul

Table 4.2 – Liste des variants ayant passé l'ensemble des filtres

Chromosome	Position	Reference allele	Altered allele	Gene
4	57686748	G	C	SPINK2
4	44683156	G	T	GUF1

candidate évident pouvant expliquer ce phénotype. Le variant partagé par les deux frères : Chr4 :57686748G>C n'a été recensé dans aucune des bases de données que sont ExAC, 1000Genomes et ESP6500. Le gène *SPINK2* est localisé sur le chromosome 4 et contient 4 exons (**Figure : 4.3**). Sa localisation intronique à 3 pb du 2^{ème} exon indique que ce variant pourrait avoir un effet sur l'épissage de l'ARNm

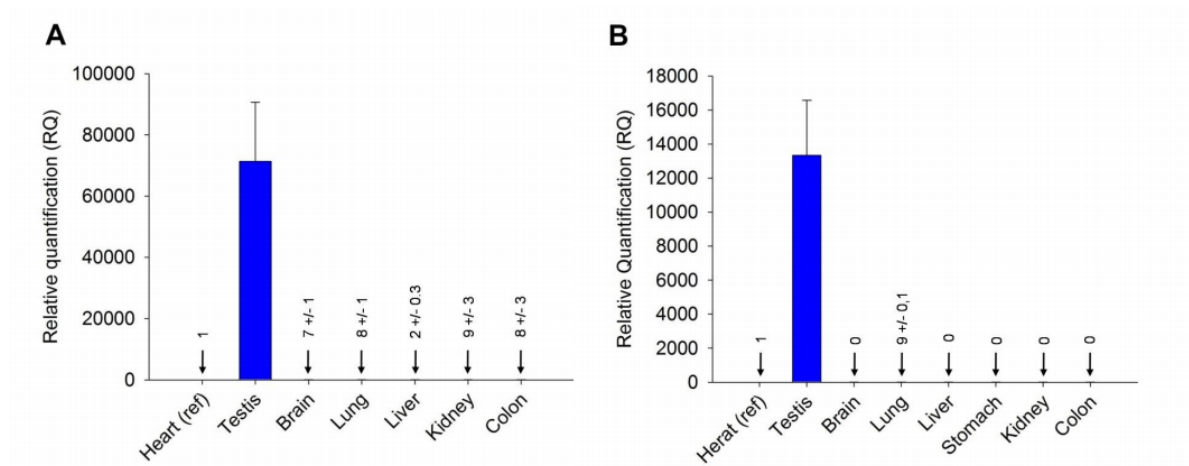


Figure 4.3 – Expression du gène **SPINK2** dans plusieurs tissus : On peut constater que chez l'humain (**A**) comme chez la souris (**B**), le gène **SPINK2** a non seulement une forte expression exclusive au testicule

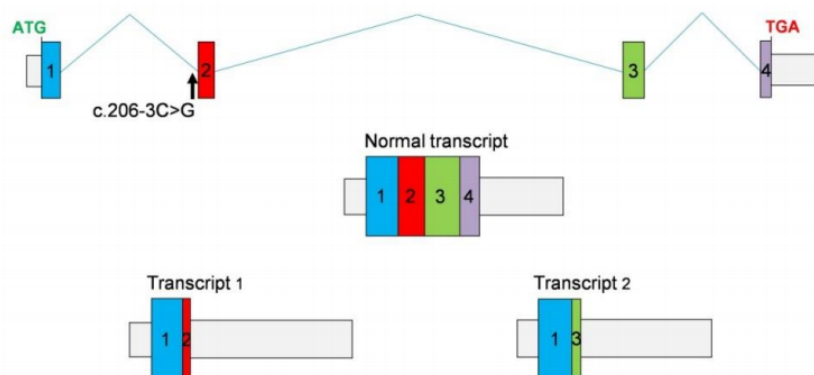


Figure 4.4 – Représentation du gène **SPINK2** : L'épissage du gène **SPINK2** crée un transcrit de 4 exons (Transcrit normal). Cependant, le variant c.206-3C>G observé chez les frères B1 et B2 crée un nouveau site accepteur d'épissage ajoutant 2 nucléotides à l'exon 2 induisant un décalage du cadre de lecture menant à un codon stop 3 nucléotides plus loin (Transcrit 1) et / ou causant le saut de l'exon 2 menant à un codon stop prématuré (Transcrit 2)

Estimation de l'incidence du gène *SPINK2* chez des patients azoospermes

Afin de déterminer l'incidence des variants du gène *SPINK2* dans la population azoospermiques nous avons séquencé l'intégralité des séquences codantes de ce gène chez 611 individus comprenant 210 patients azoospermiques, 393 oligozoospermiques et présentant un phénotype non spécifié. Parmi cet ensemble de patient, seul le patient 105 (P105) présentant un phénotype d'oligozoospermie s'est révélé porteur d'un variant non répertorié dans les bases de données et présentant un impact prédit comme délétère. Ce variant, c.1A>T, présent à l'état hétérozygote chez P105 affecte le codon start décalant ainsi le démarrage de la traduction produisant ainsi une protéine tronquée.

Autres résultats Dans cette même étude nous avons également pu étudié les caractéristiques reproductives de souris KO *Spink2*^{-/-} et de souris hétérozygotes *Spink2*^{+/-}. Ainsi, nous avons pu confirmer que les souris mâles KO étaient infertiles et présentaient un phénotype d'azoospermie ainsi qu'une diminution de la taille de leurs testicules tandis que les hétérozygotes étaient parfaitement fertiles bien leur nombre de spermatozoïdes par mL de sperme soit plus faibles que les souris sauvages. Les souris femelles, elles, présentaient des caractéristiques reproductives tout à fait normal. [TODO : finir cette partie].

Etude d'une large cohorte de patients MMAF

Chapitre 5

MutaScript

Conclusion

Chapitre 6

The First Appendix

References