

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 :



**Université
Grenoble
Alpes**

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
3.1 Introduction sur la globozoospermie	9
3.2 Résultats	12
3.2.1 Les mécanismes mutationnels	12
Confirmation de l'excès de délétion	12
Détermination du taux de recombinaison <i>de novo</i> de <i>DPY19L2</i>	14
Autres résultats	15
3.2.2 La transcriptomique	15
3.3 Conclusion	17
Chapitre 4 : Mise en place d'une stratégie pour l'analyse des données exomiques – application en recherche clinique	19
Chapitre 5 : MutaScript	21
Conclusion	23
Chapitre 6 : The First Appendix	25
References	27

Liste des tableaux

Table des figures

3.1	Observation au microscope à balayage d'un spermatozoïde normal (**A**) et d'un spermatozoïde globozoocéphale (**B**) (changer les photos avec celles sur lesquelles on voit l'acrosome colorés)	10
3.2	Comparaison entre les spermatozoïdes des souris <i>Dpy19l2</i> ^{-/-} (à gauche) et les souris sauvages <i>Dpy19l2</i> ^{+/+} (à droite)	11
3.3	Représentation schématique du mécanisme de NAHR causé par les séquences LCR flanquant le gène <i>DPY19L2</i>	13
3.4	Ratio des délétions / duplications <i>de novo</i> observées au locus <i>DPY19L2</i> déterminé par PCR digital à partir d'ADN spermatique de trois donneurs	15

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

3.1 Introduction sur la globozoospermie

Comme expliqué précédemment, La globozoospermie est phénotype rare ($< 0.1\%$ des patients infertiles) mais néanmoins sévère (Sen, Holstein, & Schirren, 1971) de teratozoospermie menant à l'infertilité masculine. Cette anomalie est caractérisée par la présence de spermatozoïdes présentant une tête ronde dépourvue d'acrosome et d'une pièce intermédiaire désorganisée dans l'éjaculat (Singh, n.d., Pedersen & Rebbe (1974)) (**Figure : 3.1**). En plus des anomalies morphologiques, les spermatozoïdes globozoocéphales présentent également des désorganisations au niveau moléculaire. Par exemple, le facteur spermatique PLC ζ requis pour l'activation ovocytaire, est absent ou en quantité infime dans les spermatozoïdes globozoocéphales (Heytens et al., 2009, Taylor et al. (2010), S.-Y. Yoon et al. (2008)) compromettant ainsi l'activation ovocytaire et expliquant le faible taux de fécondation observés en IVF (*in vitro* fertilization) et en ICSI (intra cytoplasmic sperm injection) (A. Dam et al., 2006). On distingue la globozoospermie totale avec 100% des spermatozoïdes présentant le phénotype ou partielle en fonction du taux de spermatozoïdes atteints. Bien que l'infertilité masculine soit souvent la résultante de plusieurs facteurs, les premières études présentant des patients atteints par un phénotype complet (Sen et al., 1971) suggéraient que la globozoospermie était une exception. De plus les caractéristiques morphologiques très typiques des spermatozoïdes laissaient penser à une cause monogénique. En 2007, une étude portant sur une famille juive ashkénaze comprenant six frères dont trois atteints a pu lier ce phénotype à une mutation homozygotes sur le gène *SPATA16* présente chez les trois frères atteints (A. H. D. M. Dam et al., 2007). Cependant, dans la même étude, 29 autres patients présentant le même phénotype ont été analysés, et pour ceux-ci, aucun variant du gène *SPATA16* n'a

pu être lié au phénotype (A. H. D. M. Dam et al., 2007) indiquant clairement que les mutations de ce gène n'étaient pas les seules responsables. En 2011, une autre étude portant sur une cohorte de 20 patients Tunisiens a pu mettre en évidence une délétion homozygote de 200kb emportant la totalité du gène *DPY19L2* chez 15 des 20 patients analysés (Harbuz et al., 2011). Les études effectuées ultérieurement sur ce phénotype ont ensuite pu montrer que les altérations du gène *DPY19L2*, et notamment cette délétion, étaient responsables de la majorité des cas de globozoospermie (P. F. Ray & Arnoult, 2011, Ellnati et al. (2012)).

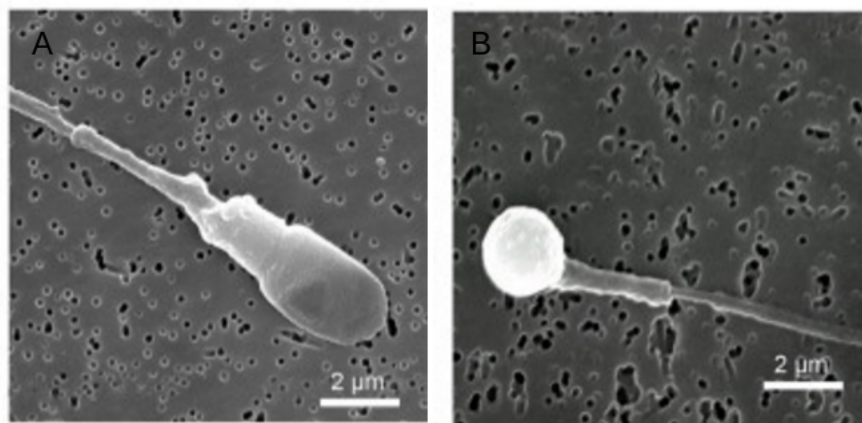


Figure 3.1 – Observation au microscope à balayage d'un spermatozoïde normal (**A**) et d'un spermatozoïde globozoocéphale (**B**) (changer les photos avec celles sur lesquelles on voit l'acrosome colorés) adapté d'après [Harbuz2011]

En 2012, le développement d'un modèle murin KO *Dpy19l2*^{-/-} a permis de mieux comprendre les mécanismes moléculaires impliqués dans la globozoospermie causée par la délétion du gène *DPY19L2* chez l'humain (V. Pierre et al., 2012). Tout d'abord car ce modèle de souris KO présentait les mêmes caractéristiques que les patients humains. Tout d'abord, ces souris étaient infertiles et présentaient des spermatozoïdes globozoocéphales (**Figure : 3.2**) mais aussi et surtout, l'ensemble des autres dysfonctionnements étaient retrouvés, c'est à dire : l'absence de l'acrosome, les défauts morphologiques du noyau, de l'enveloppe nucléaire et de l'acroplaxome ainsi que le mauvais positionnement de la manchette (V. Pierre et al., 2012). Ainsi il a pu être démontré que la protéine Dpy19l2 était principalement exprimée dans les spermatozoïdes et plus spécifiquement dans la membrane nucléaire interne faisant face à la vésicule acrosomale et que l'absence de cette protéine entraînait la déstabilisation à la fois de la lamine nucléaire, de la jonction entre l'acroplaxome et l'enveloppe nucléaire (V. Pierre et al., 2012).

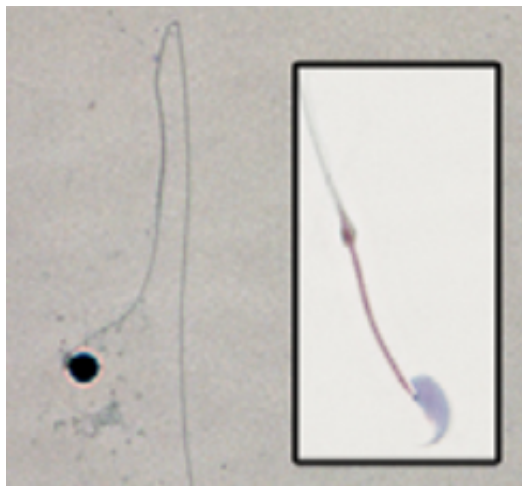


Figure 3.2 – Comparaison entre les spermatozoïdes des souris $*Dpy19l2^{* - / -}$ (à gauche) et les souris sauvages $*Dpy19l2^{* + / +}$ (à droite) d'après [Pierre2012]

Dans ce chapitre je détaillerai les résultats de trois articles dont je suis co-auteur :

1. **Fine Characterisation of a Recombination Hotspot at the DPY19L2 Locus and Resolution of the Paradoxical Excess of Duplications over Deletions in the General Population** : au cours de cette étude j'ai participé à divers manipulation de biologie moléculaire tel que l'extraction d'ADN spermatique, quantification des délétions / duplications *de novo*. De même, j'ai pu contribuer au divers analyses statistiques.
2. **Comparative testicular transcriptome of wild type and globozoospermic Dpy19l2 knock out mice** : Dans cette étude j'ai pu effectuer l'intégralité des manipulation de biomoléculaire (extraction de l'ARN testiculaire de souris et analyse sur puce) de même que l'intégralité de l'analyse bioinformatique des résultats.

3.2 Résultats

3.2.1 Les mécanismes mutationnels

Confirmation de l'excès de délétion

Chez les mammifères il existe trois paralogues de *DPY19L2* de fonction encore inconnue et un pseudogène présentant une très forte homologie de séquence ($> 95\%$) (Carson, Cheung, & Scherer, 2006). Chez l'Homme, ce gène est flanqué de deux séquences présentant une forte homologie ($>95\%$) d'une taille de 28kb (**Figure : 3.3 - 1**). Ces séquences appelées LCRs (*low copy repeats*) représentent une large portion du génome humain (Cheung et al., 2003, Bailey et al. (2002)) et vont, de par leur homologie favoriser les duplications de gènes jouant ainsi un rôle important dans l'évolution des génomes des vertébrés (Walsh, 2003, Ohno (1970)). Dans le cas de *DPY19L2*, ces LCRs vont, au cours de la méiose entraîner la venue de recombinaison homologues non-allélique (NAHR) donnant lieu soit à une délétion du gène *DPY19L2* et la formation d'un ADN circulaire comprenant le gène (**Figure : 3.3 - 2**) soit à un allèle possédant deux copies du gène tandis que l'autre n'en possède aucune (**Figure : 3.3 - 3**).

Ainsi, le mécanisme de NAHR devrait, en théorie, engendrer la formation de plus d'allèle délété que d'allèle dupliqué puisque les cas présentés en figures 3.3 - 2 et 3.3 - 3 induisent la formation d'un allèle délété tandis que seul le cas 3.3 - 3 forme un allèle dupliqué. Cependant, les données mises à disposition par la base de donnée *Database of Genomic Variants* (MacDonald, Ziman, Yuen, Feuk, & Scherer, 2014) indique un excès de duplication puisque sur un total de ... individus analysés, un total de 83 duplications et de 26 délétions hétérozygotes ont été observées pour le locus de *DPY19L2*. Afin de confirmer ces résultats et ainsi écarter l'hypothèse d'un biais causé par la présence du pseudogène *DPY19L2P1* très homologue avec *DPY19L2* (Carson et al., 2006) notre équipe a procédé à l'analyse du locus de *DPY19L2* de 1699 individus par CGHarray. Un total de 15 duplication et de 3 délétions hétérozygotes furent observés

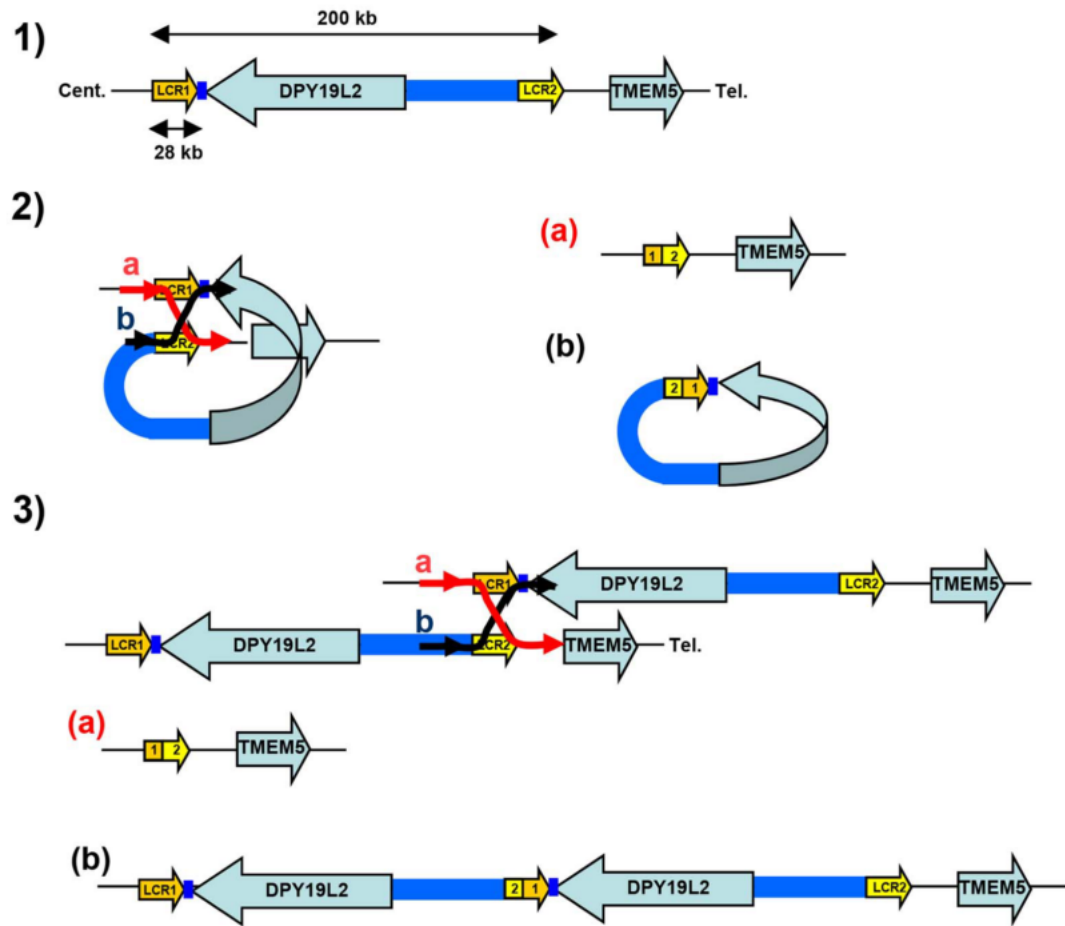


Figure 3.3 – Représentation schématique du mécanisme de NAHR causé par les séquences LCR flanquant le gène *DPY19L2* : ****1**** : Le gène *DPY19L2* est entouré par les séquences LCR1 et LCR2 qui correspondent respectivement au LCR centromérique et télomérique. Ces LCRs sont séparés par environ 200kb et chacun d’eux mesurent approximativement 28kb. ****2**** : NAHR résultant du mauvais alignement des LCRs 1 et 2 du même chromatide entraînant la formation d’un allèle délété (a) et d’un ADN circulaire comprenant le gène *DPY19L2*. Cet ADN circulaire sera éliminé au cours du cycle cellulaire. ****3**** : NAHR résultant du mauvais alignement des LCRs 1 et 2 provenant de deux chromatides différents conduisant à la formation d’un allèle délété (a) et d’un allèle dupliqué (b)

Détermination du taux de recombinaison *de novo* de *DPY19L2*

Alors que les résultats précédents confirment un excès de l'allèle dupliqué de *DPY19L2* dans la population générale, nous avons par la suite cherché à déterminer les fréquences de duplications / délétions *de novo* de ce même locus. Ceci ayant pour but de déterminer si cet excès est dû à une sélection de l'allèle dupliqué ou au fait que celui-ci était effectivement produit plus fréquemment que l'allèle délété. Pour ce faire nous avons quantifié le taux d'apparition de ces événements génétiques à partir d'ADN spermatique. Les spermatozoïdes étant le produits direct de la méiose, ils sont donc les reflets d'haplotypes produits *de novo*. Pour cela, nous avons analysé par PCR digitale l'ADN spermatique de trois donneurs. Leur ADN a tout d'abord été dilué en série de sorte à ce qu'environ 25% des 96 puits de la PCR contiennent un événement (délétion ou duplication). Ainsi, en acceptant qu'un génome haploïde humain représente 3pg, 50ng d'ADN spermatique furent déposés dans chaque puit pour la PCR spécifique à la délétion, et 100ng dans chaque puit spécifiques à la duplication. Chaque puit contient donc une partie de cette charge d'ADN initiale. La distribution de cette charge d'ADN au sein des 96 puits peut donc s'apparier à un tirage sans remise, la probabilité qu'un puit soit positif pour un événement chromosomique (duplication ou délétion) peut donc être modélisé par une loi hypergéométrique (**Équation** : (3.1)). Nous permettant ainsi d'estimer la fréquence duplication / délétion λ pour chaque donneur (**Équation** : (3.2)) (**Figure** : 3.4).

$$\frac{\frac{(N-R)!}{W!(N-R-W)!}}{\frac{N!}{W!(N-W)!}} = \frac{(N-R)!(N-W)!}{N!(N-W-R)!} = \prod_{i=0}^{R-1} \frac{N-W-i}{N-i} \quad (3.1)$$

$$\lambda = \frac{R}{N} \quad (3.2)$$

Où :

. N : représente le nombre de copie de chromosome 12 dans la charge d'ADN initiale (1.6×10^6 pour la PCR spécifique à la délétion, 3.2×10^6 pour la PCR spécifique à la duplication)

. $W = \frac{N}{96}$ correspond au nombre de copie de chromosome 12 par puit

. R représente le nombre total de recombinaison observées

L'intervalle de confiance (IC) à 95% est ensuite calculé grâce à une loi binomiale de sorte à modéliser la dilution initiale pour obtenir l'ADN d'entrée. Le puit contenant le *pool* des trois ADN spermatique est donc celui ayant les résultats les plus robustes, l'IC étant le plus resserré (**Figure** : 3.4), et permet donc d'établir le taux de délétion *de novo* à 1.8×10^{-5} (IC 95% : 1.4×10^{-6} ; 2.2×10^{-6}) tandis que le taux de duplication *de novo* est estimé à 7.7×10^{-6} (IC 95% : 6.1×10^{-6} ; 9.7×10^{-6}) montrant un enrichissement environ deux fois supérieur des délétions par rapport aux duplications sur le site de *DPY19L2* (**Figure** : 3.4).

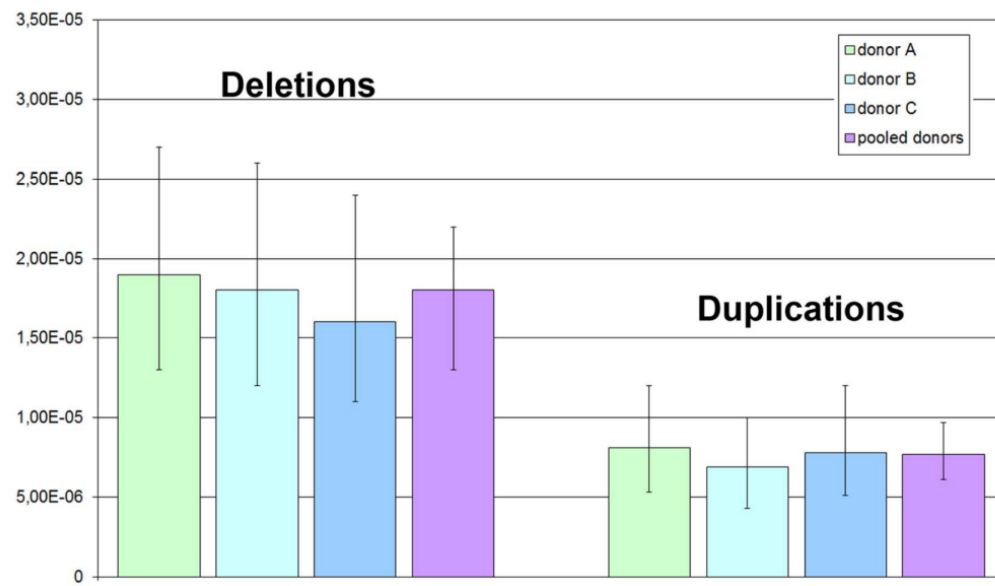


Figure 3.4 – Ratio des délétions / duplications *de novo* observées au locus *DPY19L2* déterminé par PCR digital à partir d'ADN spermatique de trois donneurs : [TODO]

Autres résultats

3.2.2 La transcriptomique

Karaouzen et al. *Basic and Clinical Andrology* 2013, **23**:7
<http://www.bacandrology.com/content/23/1/7>



BASIC AND CLINICAL
ANDROLOGY

RESEARCH ARTICLE

Open Access

Comparative testicular transcriptome of wild type and globozoospermic *Dpy19l2* knock out mice

Thomas Karaouzen^{1,2}, Michèle El Atifi^{3,4,5}, Jean-Paul Issartel^{3,4,5}, Marianne Grepillat^{1,2,6}, Charles Coutton^{1,2,7}, Delphine Martinez^{1,6}, Christophe Arnoult^{1,2} and Pierre F Ray^{1,2,6*}

Abstract

Background: Globozoospermia is a male infertility phenotype characterized by the presence in the ejaculate of near 100% acrosomeless round-headed spermatozoa with normal chromosomal content. Following intracytoplasmic sperm injection (ICSI) these spermatozoa give a poor fertilization rate and embryonic development. We showed previously that most patients have a 200 kb homozygous deletion, which includes *DPY19L2* whole coding sequence. Furthermore we showed that the DPY19L2 protein is located in the inner nuclear membrane of spermatids during spermiogenesis and that it is necessary to anchor the acrosome to the nucleus thus performing a function similar to that realized by Sun proteins within the *LINC-complex* (Linker of Nucleoskeleton and Cytoskeleton). SUN1 was described to be necessary for gametogenesis and was shown to interact with the telomeres. It is therefore possible that *Dpy19l2* could also interact, directly or indirectly, with the DNA and modulate gene expression during spermatogenesis. In this study, we compared the transcriptome of testes from *Dpy19l2* knock out and wild type mice in order to identify a potential deregulation of transcripts that could explain the poor fertilization potential of *Dpy19l2* mutated spermatozoa.

Methods: RNA was extracted from testes from *DPY19L2* knock out and wild type mice. The transcriptome was carried out using GeneChip® Mouse Exon 1.0 ST Arrays. The biological processes and molecular functions of the differentially regulated genes were analyzed with the PANTHER software.

Results: A total of 76 genes were deregulated, 70 were up-regulated and 6 (including *Dpy19l2*) were down-regulated. These genes were found to be involved in DNA/RNA binding, structural organization, transport and catalytic activity.

Conclusions: We describe that an important number of genes are differentially expressed in *Dpy19l2* mice. This work could help improving our understanding of *Dpy19l2* functions and lead to a better comprehension of the molecular mechanism involved in spermatogenesis.

Keywords: Male infertility, Globozoospermia, Spermatogenesis, *Dpy19l2*, Transcriptome

* Correspondence: pray@chu-grenoble.fr

¹Université Joseph Fourier, Grenoble F-38000, France

²Laboratoire AGIM, CNRS FRE3405, Equipe "Génétique, Infertilité et Thérapeutiques", La Tronche F-38700, France

Full list of author information is available at the end of the article

3.3 Conclusion

Chapitre 4

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

Chapitre 5

MutaScript

Conclusion

Chapitre 6

The First Appendix

References

- Bailey, J. A., Gu, Z., Clark, R. A., Reinert, K., Samonte, R. V., Schwartz, S., ... Eichler, E. E. (2002). Recent Segmental Duplications in the Human Genome. *Science*, 297(5583), 1003–1007. <http://doi.org/10.1126/science.1072047>
- Carson, A. R., Cheung, J., & Scherer, S. W. (2006). Duplication and relocation of the functional DPY19L2 gene within low copy repeats. *BMC Genomics*, 7, 45. <http://doi.org/10.1186/1471-2164-7-45>
- Cheung, J., Estivill, X., Khaja, R., MacDonald, J. R., Lau, K., Tsui, L.-C., & Scherer, S. W. (2003). Genome-wide detection of segmental duplications and potential assembly errors in the human genome sequence. *Genome Biology*, 4(4), R25. <http://doi.org/10.1186/gb-2003-4-4-r25>
- Dam, A. H. D. M., Kosciński, I., Kremer, J. A. M., Moutou, C., Jaeger, A.-S., Oudakker, A. R., ... Viville, S. (2007). Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *American Journal of Human Genetics*, 81(4), 813–20. <http://doi.org/10.1086/521314>
- Dam, A., Feenstra, I., Westphal, J., Ramos, L., Golde, R. van, & Kremer, J. (2006). Globozoospermia revisited. *Human Reproduction Update*, 13(1), 63–75. <http://doi.org/10.1093/humupd/dml047>
- Ellnati, E., Kuentz, P., Redin, C., Jaber, S., Vanden Meerschaut, F., Makarian, J., ... Viville, S. (2012). Globozoospermia is mainly due to DPY19L2 deletion via non-allelic homologous recombination involving two recombination hotspots. *Human Molecular Genetics*, 21(16), 3695–3702. <http://doi.org/10.1093/hmg/dds200>
- Harbuz, R., Zouari, R., Pierre, V., Ben Khelifa, M., Kharouf, M., Coutton, C., ... Ray, P. F. (2011). A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. *American Journal of Human Genetics*, 88(3), 351–61. <http://doi.org/10.1016/j.ajhg.2011.02.007>
- Heytens, E., Parrington, J., Coward, K., Young, C., Lambrecht, S., Yoon, S.-Y., ... De Sutter, P. (2009). Reduced amounts and abnormal forms of phospholipase C zeta (PLCzeta) in spermatozoa from infertile men. *Human Reproduction (Oxford, England)*, 24(10), 2417–28. <http://doi.org/10.1093/humrep/dep207>
- MacDonald, J. R., Ziman, R., Yuen, R. K. C., Feuk, L., & Scherer, S. W. (2014).

The Database of Genomic Variants : a curated collection of structural variation in the human genome. *Nucleic Acids Research*, 42(Database issue), D986–92. <http://doi.org/10.1093/nar/gkt958>

Ohno, S. (1970). *Evolution by Gene Duplication*. Berlin, Heidelberg : Springer Berlin Heidelberg. <http://doi.org/10.1007/978-3-642-86659-3>

Pedersen, H., & Rebbe, H. (1974). Fine structure of round-headed human spermatozoa. *Journal of Reproduction and Fertility*, 37(1), 51–4. <http://doi.org/10.1530/JRF.0.0370051>

Pierre, V., Martinez, G., Coutton, C., Delaroche, J., Yassine, S., Novella, C., ... Arnoult, C. (2012). Absence of Dpy19l2, a new inner nuclear membrane protein, causes globozoospermia in mice by preventing the anchoring of the acrosome to the nucleus. *Development*, 139(16), 2955–2965. <http://doi.org/10.1242/dev.077982>

Ray, P. F., & Arnoult, C. (2011). La délétion homozygote du gène *DPY19L2* est responsable de la majorité des cas de globozoospermie. *Médecine/Sciences*, 27(8-9), 692–693. <http://doi.org/10.1051/medsci/2011278004>

Sen, C. G. S., Holstein, A. F., & Schirren, C. (1971). über die Morphogenese rundköpfiger Spermatozoen des Menschen. *Andrologia*, 3(3), 117–125. <http://doi.org/10.1111/j.1439-0272.1971.tb02106.x>

Singh, G. (n.d.). Ultrastructural features of round-headed human spermatozoa. *International Journal of Fertility*, 37(2), 99–102. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/1349598>

Taylor, S., Yoon, S., Morshedi, M., Lacey, D., Jellerette, T., Fissore, R., & Oehninger, S. (2010). Complete globozoospermia associated with PLC ζ deficiency treated with calcium ionophore and ICSI results in pregnancy. *Reproductive BioMedicine Online*, 20(4), 559–564. <http://doi.org/10.1016/j.rbmo.2009.12.024>

Walsh, B. (2003). Population-genetic models of the fates of duplicate genes. *Genetica*, 118(2-3), 279–94. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12868616>

Yoon, S.-Y., Jellerette, T., Salicioni, A. M., Lee, H. C., Yoo, M.-S., Coward, K., ... Fissore, R. A. (2008). Human sperm devoid of PLC, zeta 1 fail to induce Ca(2+) release and are unable to initiate the first step of embryo development. *The Journal of Clinical Investigation*, 118(11), 3671–81. <http://doi.org/10.1172/JCI36942>