THÈSE

Thomas Karaouzene

31 octobre 2017

Delete line 6 if you only have one advisor

Remerciements

Résumé

Introduction

Investigation génétique et physiologique de la globozoospermie

Introduction sur la globozoospermie

Comme expliqué précédemment, La globozoospermie est phénotype rare (< 0.1% des patients infertiles) mais néanmoins sévère (Sen, Holstein, & Schirren, [1971](#ref-Sen2009)) 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](#ref-Pedersen1974))) (**Figure :** 1). En plus des anomalies morphologiques, les spermatozoïdes globozoocéphales présentent également des désorganisations au niveau moléculaire. Par exemple, le facteur spermatique PLC requit pour l'activation ovocitaire, est absent ou en quantité infime dans les spermatozoïdes globozoocéphales (Heytens et al., [2009](#ref-Heytens2009), Taylor et al. ([2010](#ref-Taylor2010)), S.-Y. Yoon et al. ([2008](#ref-Yoon2008))) 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](#ref-Dam2006)). 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](#ref-Sen2009)) 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 atteins a pu lier ce phénotype à une mutation homozygotes sur le gène *SPATA16* présente chez les trois frères atteint (A. H. D. M. Dam et al., [2007](#ref-Dam2007)). Cependant, dans la même étude, 29 autres patients présentant le même phénotype ont été analysé, et pour ceux-ci, aucun variant du gène *SPATA16* n'a pu être lié au phénotype (A. H. D. M. Dam et al., [2007](#ref-Dam2007)) 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](#ref-Harbuz2011)). 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](#ref-Ray2011), ElInati et al. ([2012](#ref-ElInati2012))).

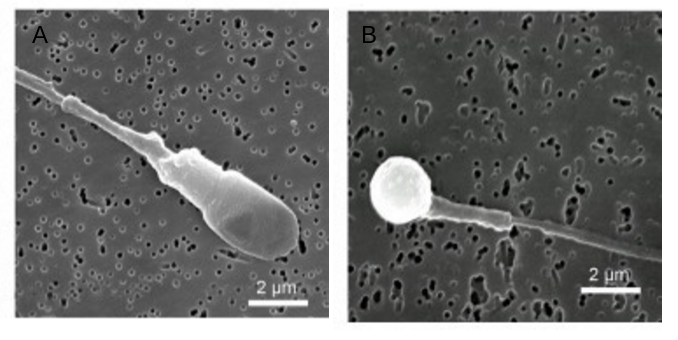


Figure 1 Observation au microscope à balayage d'un spermatozoïde normal (**A**) et d'un spermatozoïde globozoocéphale (**B**) (changer les photos avec celles sur lesquelles on voit l'acrosome colorés) adapté d'après (Harbuz et al., [2011](#ref-Harbuz2011))

En 2012, le développement d'un modèle murin KO *Dpy19l2* a permis de mieux comprendre les mécanismes moléculaires impliqué dans la globozoospermie causée par la délétion du gène *DPY19L2* chez l'humain (V. Pierre et al., [2012](#ref-Pierre2012)). 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 :** 2) mais aussi et surtout, l'ensembles 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](#ref-Pierre2012)). Ainsi il a pu être démontré que la protéine Dpy19l2 étaient principalement exprimé dans le spermatides et plus spécifiquement dans la membrane nucléaire interne faisant face à la vésicule acrosomale et que l'absence de cette protéine entrainait la déstabilisation à la fois de la lamine nucléaire, de la jonction entre l'acroplaxome et l'enveloppe nucléaire (V. Pierre et al., [2012](#ref-Pierre2012)).

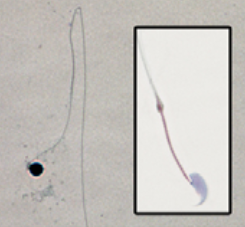


Figure 2 Comparaison entre les spermatozoïdes des souris *Dpy19l2* (à gauche) et les souris sauvages *Dpy19l2* (à droite) d'après (V. Pierre et al., [2012](#ref-Pierre2012))

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**](#mecamut) : 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**](#transcriptome) : 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.

Résultats

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](#ref-Carson2006)). Chez l'Homme, ce gène est flanqué de deux séquences présentant une forte homologie (>95%) d'une taille de 28kb (**Figure :** 3 - **1**). Ces séquences appelées LCRs (*low copy repeats*) représentent une large portion du génome humain (Cheung et al., [2003](#ref-Cheung2003), Bailey et al. ([2002](#ref-Bailey2002))) 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](#ref-Walsh2003), Ohno ([1970](#ref-Ohno1970))). Dans le cas de *DPY19L2*, ces LCRs vont, au cours de la méiose entrainer 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 - **2**) soit à un allèle possédant deux copies du gène tandis que l'autre n'en possède aucune (**Figure :** 3 - **3**).

Ainsi, le mécanisme de NAHR devrait, en théorie, engendrer la formation de plus d'allèles délétés que d'allèles dupliqués puisque les cas présentés en figures 3 - **2** et 3 - **3** induisent la formation d'un allèle délété tandis que seul le cas 3 - **3** forme un allèle dupliqué. Cependant, les données mises à disposition par la base de donnée [*Database of Genomic Variants*](http://dgv.tcag.ca/dgv/app/home) (DGV) (MacDonald, Ziman, Yuen, Feuk, & Scherer, [2014](#ref-MacDonald2014)) indique un excès de duplication puisque sur un total de 6575 individus analysés, 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](#ref-Carson2006)) notre équipe a procedé à la ré-analyse du locus de *DPY19L2* de 1699 individus par CGHarray. Un total de 15 duplication et de 3 délétions hétérozygotes furent observés coroborant les données fournies par DGV et confirmant ainsi l'excès de l'allèle dupliqué dans la population générale.

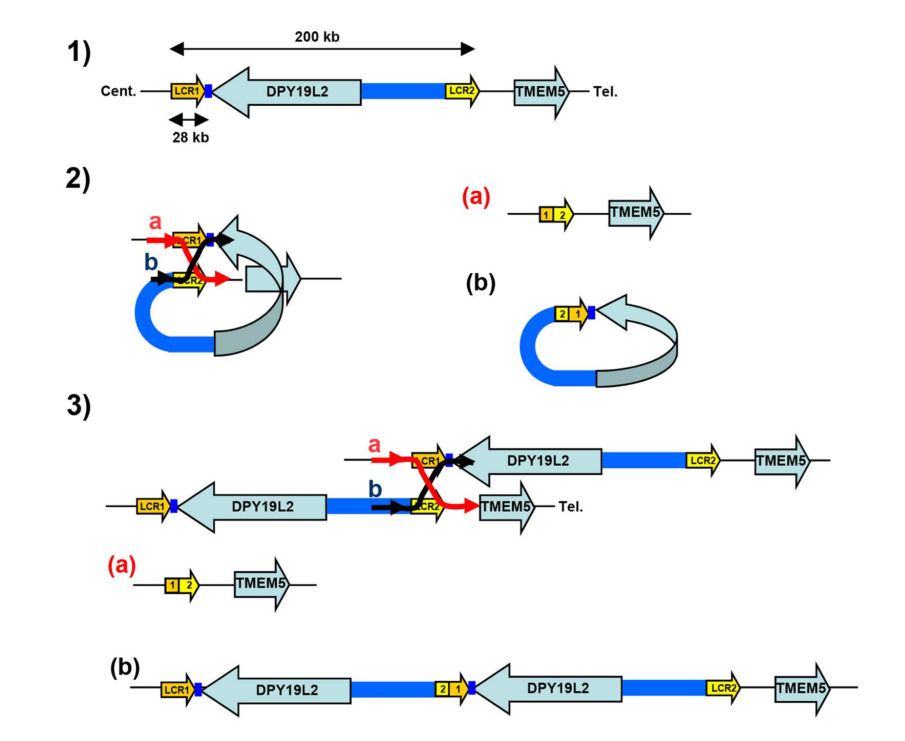


Figure 3 Représentation schématique du mécanisme de NAHR causé par les séquences LCR flanquant le gène *DPY19L2* : **1** : Le gène *DPY19L2* est entouré par les séquences LCR1 et LCR2 qui correspondent respectivement au LCR centromérique et télomérique. Ces LCRs sont séparés par environs 200kb et chacun d'eux mesurent approximativement 28kb. **2** : NAHR résultant du mauvais alignement des LCRs 1 et 2 du même chromatide entrainant 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 exces est dû à une séléction de l'allèle dupliqué ou au fait que celui-ci était effectivement produit plus fréquement 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 ainsi que l'ADN spermatique constitué d'un mix provenant de ces trois donneurs. Leur ADN a tout d'abor 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 a 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** : (1)). Nous permettant ainsi d'estimer la fréquence duplication / délétion pour chaque donneur (**Équation** : (2)) (**Figure :** 4).

$$

$$

Où :  
. : représente le nombre de copie de chromosome 12 dans la charge d'ADN initiale (1.6x10 pour la PCR spécifique à la délétion, 3.2x10 pour la PCR spécifique à la duplication)  
. correspond au nombre de copiede chromosome 12 par puit  
. représente le nombre total de recombinaison observées

L'interval 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 resséré (**Figure :** 4), et permet donc d'établir le taux de délétion *de novo* à 1.8 x 10 (IC 95% : 1.4x 10; 2.2x 10) tandis que le taux de duplication *de novo* est estimé à 7.7 x 10 (IC 95% : 6.1 x 10; 9.7 x 10) montrant un enrichisement environ deux fois supperieur des délétion par rapport aux duplications sur le site de *DPY19L2* (**Figure :** 4).

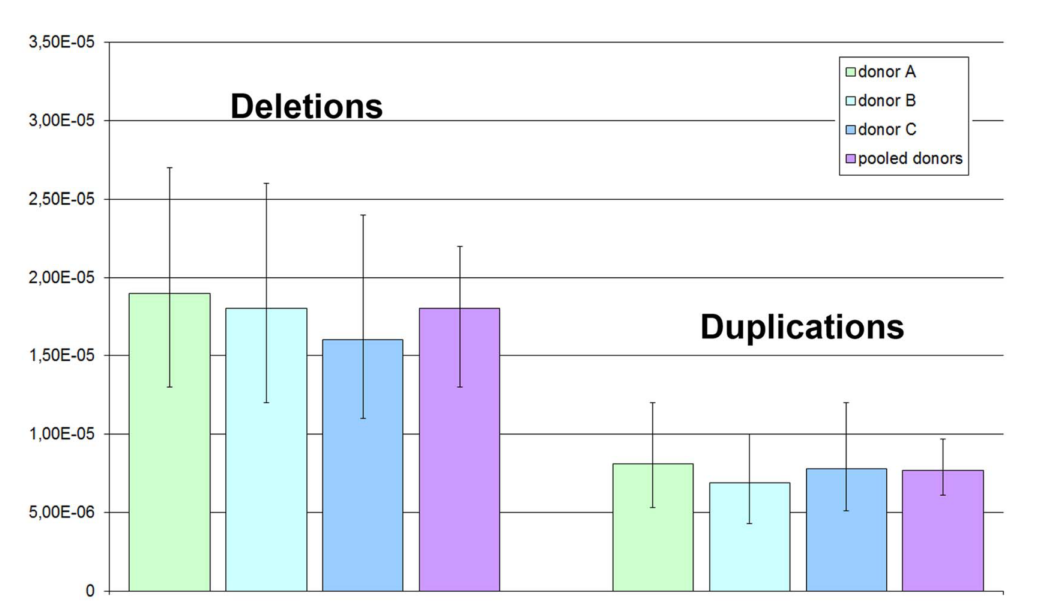


Figure 4 Ratio des délétions / duplications *de novo* observées au locus *DPY19L2* déterminé par PCR digital à partir d'ADN spermatique de trois donneurs : [TODO]

[TODO: expliquer brievement notre hpothèse expliquant ce phénomène]

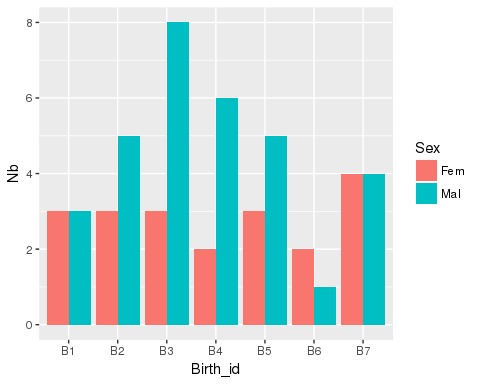
Autres résultats

Cette étude a également été pour notre équipe l'occasion d'effectuer une étude plus approfondie des LCRs flanquant le locus de *DPY19L2*. Ainsi, nous avons pu mettre en évidence que les LCRs 1 et 2 contenaient 5 répétitions du site de reconnaissance consensus de PRDM9 (CCNCCNTNNCCNC), une protéine connue pour [TODO : décrire le rôle de PRDM9]. De même, nous avons pu mettre en évidence que les recombinaisons s'effectuaient le long de 5 points de cassures distincts répartis le long des LCRs 1 et 2 et que comme attendu, l'incidence des points de cassures dans les recombinaisons du locus *DPY19L2* était proportionelle à la taille des régions [TODO : expliquer mieux].

La transcriptomique

Dans des études précédentes, notre équipe à réussit à démontrer que la protéine *DPY19L2* était localisée dans la membrane interne des noyaux des spermatides pendant la spermatogénèse et qu’elle est nécessaire pour fixer l’acrosome au noyau [TODO: insert ref]. De même, nous avons pu mettre en évidence que dans des cellules HEK cette protéine colocalisait avec la protéine SUN5 et que *Dpy19l2* pourrait être un partenaire de SUN5(V. Pierre et al., [2012](#ref-Pierre2012)). nous avons cherché à observer si, comme la protéine SUN5. Chez la souris, la protéine Sun1 est elle aussi nécéssaire à la gamétogénèse et est connue pour permettre l'intéraction entre le noyau et les télomères (Ding et al., [2007](#ref-Ding2007)). Dans cette étude nous avons donc chercher à savoir si l'absence de la protéine Dpy19l2 pouvait entrainer des dérèglement transcriptionelle qui pourrait, entre autre, expliquer l'absence de la protéine PLC dans les spermatozoïdes globozoocéphales. De plus, au cours de l'elevage des souris *Dpy19l2* KO au sei de notre laboratoir nous avons pu observé un excès de naissance de souris mâle lorsque l'on croisait deux sours *Dpy19l* [TODO: insérer tableau et figure montrant le biais de sexe].  
Afin de tenter d'expliquer ces observations, nous avons effectué une analyse comparative du transcriptome testiculaire de deux souris *Dpy19l2* (S1+ et S2+) et deux souris *Dpy19l2* (S1- et S2-) ayant pour but de mettre en évidence d'eventuels dereglement transcriptionels chez la souris KO. Pour effectuer ces analyses, nous avons donc extrait l'ARN testiculaire des 4 souris que nous avons ensuite hybridé sur des puces à ADN Affymetrix GeneChip® Mouse Exon 1.0 contenant des sondes pour 35.557 gènes murins. Cette étape nous a alors permi d'obtenir pour chacune des 4 souris les valeurs d'expression testiculaire de l'ensemble de leurs gènes. Pour chacun de ces gènes, nous avons donc chercher à soivoir s'ils étaient différentiellement exprimés chez les sours S1- et S2- lorsqu'on comparait leur expression avec celle des souris S1+ et S2+. Pour cela, nous avons calculé quatre ratios (R1, R2, R3 et R4) (**Équation** : (3)). Les gènes pour lesquels au moins 3 de leurs ratio étaient 1,7 furent considérés comme sur-exprimés tandis que ceux pour lesquels 3 de leurs ratio étaient 0,58 () furent considérés comme sous-exprimés.

$$



Ainsi cette étude a pu mettre en évidence la sous-expression de 6 gènes (incluant *Dpy19l2*) et la sur-expression de 70 gènes chez les souris *Dpy19l2*. Parmi ces gènes, nous ne figurait pas *Plc* indiquant que l'absence de cette protéine chez les spermatozoïdes globozoocéphales n'étaient pas directement dûe à un dysfonctionnement transcriptionel. Afin de prédire les fonction moléculaire dans lesquels étaient impliqués ces gènes, nous nous sommes servi du logiciel PANTHER (Mi et al., [2017](#ref-Mi2017)). Ainsi, nous avons pu constater que 23 gènes codant pour des protéines de liaison étaient dérégulés (**Figure :** 5), dont 12 sont des protéines de liaison aux acides nucléiques (**Figure :** 6) suggérant que *Dpy19l2* pourrait effectivement intéragir avec l'ADN. D'autre fonctions moléculaires telles que l'activité catalytique, la régulation de la transcription et des protéines ayant des fonctions structurelles étaient égallement dérégulées chez les souris KO. Ces derniers sont particulièrement intéréssant lorsque l'on sait que les spermatozoïdes globozoocéphales sont caractérisés par plusieurs défauts structurels.

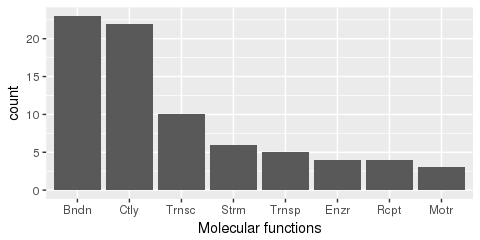


Figure 5 Fonctions moléculaires affectées par au moins un gène dérégulé : Bndn = Binding, Ctly = Catalytic, Trnsc = Transcription, Strm = Structural molecule, Enzr = Enzyme regulator, Rcpt = Receptor, Motr = Motor

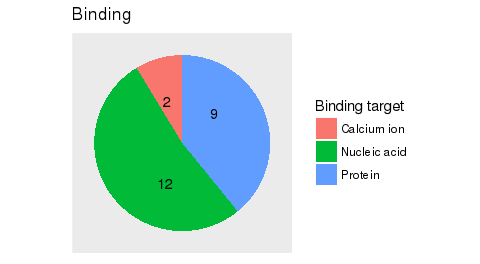


Figure 6 Détails des sites ciblés par les gènes dérégulés ayant la fonction moléculaire"Binding" : Les nombres présent au milieu des parts indiquent la quantité de gène impliqués dans cette fonction moléculaire

Au cours de cette étude nous n'avons pas pu mettre en évidence la cause expliquant le biais de sexe observé dans les naissance des souris. De même aucune dérégulation dans l'expression du gène *Plc* n'a pu être observé. Cependant, et que ceux-ci soient modestes, nous avons montrés que les souris *Dpy19l2* présentaient des déréglements transcriptionels affectant plusieurs fonctions moléculaires pouvant ainsi expliquer, du moins en partie, les nombreux défauts morphologiques caractérisant les spermatozoïdes globozoocéphales. De même, nous avons pu observer un déréglement de nombreux gènes impliqués dans la liaison d'acide nucléique et de protéine pouvant ainsi expliquer les défauts d'ancrage de l'acrosome au noyau chez les spermatozoïdes globozoocéphales.

Conclusion

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

MutaScript

Conclusion

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., Koscinski, 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>

Ding, X., Xu, R., Yu, J., Xu, T., Zhuang, Y., & Han, M. (2007). SUN1 Is Required for Telomere Attachment to Nuclear Envelope and Gametogenesis in Mice. *Developmental Cell*, *12*(6), 863–872. <http://doi.org/10.1016/j.devcel.2007.03.018>

ElInati, 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>

Mi, H., Huang, X., Muruganujan, A., Tang, H., Mills, C., Kang, D., & Thomas, P. D. (2017). PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Research*, *45*(D1), D183–D189. <http://doi.org/10.1093/nar/gkw1138>

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 <i>DPY19L2</i> 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 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>