

THÈSE

Pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ DE GRENOBLE

Spécialité : Biotechnologie, instrumentation, signal et imagerie
pour la biologie, la médecine et l'environnement

Arrêté ministériel : 7 août 2006

Présentée par

Charles COUTTON

Thèse dirigée par **Pierre RAY**

préparée au sein de l'**Equipe "Génétique, Epigénétique et Thérapie de l'Infertilité"**, Institut Albert Bonniot, INSERM U823, La Tronche, F-38706
dans l'**Ecole Doctorale Ingénierie pour la santé la Cognition et l'Environnement**.

**Caractérisation génétique et moléculaire
de l'infertilité masculine : applications à
plusieurs formes sévères de
térazoospermie.**

Thèse soutenue publiquement le **22 Juin 2015** devant le jury composé de :

Pr Pierre HAINAUT

Professeur des Universités–Praticien Hospitalier, Université Grenoble Alpes, Président

Dr Ken McELREAVY

Directeur de Recherche, Université Pierre et Marie Curie, Paris, Rapporteur

Pr François VIALARD

Professeur des Universités–Praticien Hopsitalier, Université Versailles St Quentin en Yvelines, Rapporteur

Pr Serge AMSELEM

Professeur des Universités–Praticien Hospitalier, Université Pierre et Marie Curie, Paris, Examinateur

Pr Pierre-Simon JOUK

Professeur des Universités – Praticien Hospitalier, Université Grenoble Alpes, Examinateur

Dr Pierre RAY

Maître de conférence – Praticien Hospitalier, Université Grenoble Alpes, Directeur de thèse



REMERCIEMENTS

Je remercie Messieurs le Professeur François VIALARD et le Docteur Kenneth McELREAVEY d'avoir accepté la charge d'être rapporteurs de ma thèse.

Je remercie Messieurs les Professeurs Pierre HAINAUT et Serge AMSELEM d'avoir accepté de participer à mon jury et juger ce travail.

Je remercie Monsieur le Professeur Pierre-Simon JOUK d'avoir accepté, encore une fois, d'être dans mon jury de thèse.

A Pierre, un immense merci de m'avoir encadré (et supporté) pendant ces 5 années de thèse. J'ai énormément appris à tes côtés, tu m'as transmis ton goût de la recherche, fait profiter de tes qualités de fin stratège et initié à l'art délicat de la rédaction d'articles. Il ne manque plus qu'à me faire partager ton optimisme à toutes épreuves même face aux remarques acerbes des reviewers !! Je suis ravi de poursuivre l'aventure scientifique à tes côtés.

A Florence, Françoise et Véronique, merci de votre patience, de votre affection quasi-maternelle et de votre enthousiasme. Sans votre aide et votre grande compréhension, rien n'aurait été possible. C'est presque fini, reste encore l'HDR (hein Véro ?), le post-doc, le CNU... bref, encore pleins de choses palpitantes à partager à vos côtés durant les prochaines années à venir.

A l'ensemble des personnes du Département de Génétique et Procréation et en particulier du laboratoire de Génétique Chromosomique, qui m'ont aidé et soutenu au cours de ces années. Je suis ravi de poursuivre ma route avec vous.

A toute l'équipe GETI et spécialement à Christophe pour nos longues journées derrière l'ordinateur à tenter de décortiquer la structure de l'axonème des flagelles de *Chlamydomonas* et autres protistes aux noms imprononçables ! To be continued...

A l'ensemble des personnes du laboratoire de Biochimie et Génétique Moléculaire qui ont su faire preuve de beaucoup de patience et tolérer l'exilé de l'HCE...

A mes amis, toujours présents, toujours importants.

A mes cousins adorés, à nos souvenirs partagés, à ceux qui restent à inventer.

A Thomato, pour m'avoir fait découvrir le KFC, et la chicha à la pomme ! Comme dirait notre mentor Ali G « impose ton style » !!

A Sophie, pour ton amitié retrouvée. Va falloir se supporter encore un moment, le DPI nous attend...

A John, je ne sais pas où on va ni comment on y va mais c'est avec la tête pleine d'idées folles et le corps plein de caféine que se dessinent de grands projets !

A Annick et Michel, si j'en suis là c'est un peu aussi grâce (à cause ?) à vous !

A mes parents, soyez enfin rassurés, ma vie d'éternel étudiant s'arrête ici. Non, ce n'est plus un rêve mais la réalité, cette thèse sera la dernière je vous le jure ! Je vous sais toujours présents à mes cotés quelque soit mes choix et d'un amour indéfectible. Merci d'avoir été des parents aussi exemplaires et désormais des grands-parents au dévouement sans limite pour vos deux petits trésors.

A Marie, toi qui a connu le fervent critique du monde HU le voilà qui soutient sa thèse d'Université aujourd'hui... Ma petite basque adorée tu es venue pimenter ma vie, tu m'as offert mes plus beaux cadeaux, ton amour, Camille, Antoine, une famille... comment te remercier. En cette période délicate, ton courage force l'admiration et n'a d'égale que l'amour que tu portes à tes enfants chéris. La vie continuera à nous réservier des surprises, des bonnes et des mauvaises, mais sache que tu pourras toujours t'appuyer sur l'amour inconditionnel de ta petite famille, à jamais éternel.

A mes deux petits amours, vous avez bouleversé ma vie, vous êtes sources de toutes mes pensées, de mes joies et mes angoisses aussi... Que votre vie puisse être heureuse, sans limite et accomplie.

Je dédie ce travail à la mémoire de mes grands-parents adorés.

RESUME

L'infertilité masculine concerne plus de 20 millions d'hommes à travers le monde et représente un véritable enjeu de santé public. Bien que multifactorielle, l'infertilité masculine a une composante génétique importante qui jusqu'à présent n'a été que peu étudiée. L'objectif de mon travail a été d'initier et de développer les investigations génétiques sur trois phénotypes responsables de tétratozoospermie: les spermatozoïdes macrocéphales, la globozoospermie et les anomalies morphologiques multiples des flagelles (AMMF).

Pour le premier phénotype, nous avons étudié 87 patients, dont 83 cas-index, présentant un phénotype de macrozoospermie. Une mutation pathogène dans le gène *AURKC* a été identifiée chez 82% des patients (68/83) confirmant qu'il s'agit de la cause génétique prépondérante pour ce phénotype. Une nouvelle mutation récurrente, p.Y248*, entraînant la dégradation totale du transcrit anormal par nonsense-mediated mRNA decay, a été retrouvée chez 10 patients non-apparentés. L'identification de ces deux mutations ancestrales dans leur maintien au cours de l'évolution malgré leur effet délétère sur la reproduction masculine à l'état homozygote, pose la question d'un potentiel avantage sélectif procuré par l'haplo-insuffisance d'*AURKC*.

Pour le second phénotype, une cohorte de 34 patients globozoospermiques a été analysée par séquençage et MLPA (multiplex ligation-dependent probe amplification). La délétion homozygote de *DPY19L2* a été retrouvée chez 22 des 30 patients non apparentés (73.3%) et 3 nouvelles mutations ponctuelles ont été identifiées dans ce gène. Ces résultats indiquent que l'analyse moléculaire de *DPY19L2* ne devrait pas être limitée à la recherche de la délétion homozygote de *DPY19L2* chez les patients globozoospermiques. Dans un second temps, nous avons démontré que la délétion récurrente de *DPY19L2* était médiée par le mécanisme de recombinaison homologue non allélique (NAHR) entre deux séquences répétées homologues (LCR) de 28kb situées de chaque côté du gène. La très grande majorité des points de cassure survient dans une région de 1,2 kb située dans la partie centrale des LCRs. Cette région minimale de recombinaison est elle-même centrée sur une séquence consensus de 13 nucléotides reconnue par PRDM9, une protéine à doigts de zinc qui favorise la survenue des cassures doubles brins à l'origine des processus de recombinaisons. Les modèles théoriques prédisent que, lors de la méiose, le mécanisme NAHR génère *de novo* plus d'allèles recombinés délétés que dupliqués. Étonnamment, dans la population générale les allèles *DPY19L2* dupliqués sont trois fois plus fréquents que les allèles délétés. Nous avons développé une PCR digitale sur le sperme afin de mesurer le taux de délétions et de duplications *de novo* à ce locus chez des témoins. Tel qu'il était prévu par le modèle de NAHR, nous avons identifié un taux de délétions supérieur à celui des duplications. Ce paradoxe peut s'expliquer par la sélection négative qui s'opère à l'encontre des hommes infertiles porteurs de la délétion homozygote et potentiellement des hommes porteurs d'une délétion hétérozygote.

Enfin pour le troisième phénotype, nous avons réalisé l'analyse par cartographie par homozygotie de 20 patients dont 18 cas-index, avec des spermatozoïdes présentant des anomalies morphologiques du flagelle. Cinq mutations homozygotes ont été identifiées dans le gène *DNAH1* parmi les 18 patients non-apparentés (28%). Ce gène code pour une chaîne lourde des bras internes de dynéine exprimée dans le testicule. Des analyses d'immunofluorescence et de RT-PCR ont confirmé le caractère pathogène d'une de ces mutations situées sur un site donneur d'épissage. Les analyses par microscopie électronique ont révélé une désorganisation générale de l'axonème incluant une disparition des doublets centraux et des bras internes de dynéine suggérant que *DNAH1* est une protéine clé dans la biogénèse du flagelle du spermatozoïde.

ABSTRACT

Male infertility affects more than 20 million men worldwide and represents a major health concern. Although multifactorial, male infertility has a strong genetic basis which has so far not been extensively studied. The objectives of my thesis were to initiate and conduct some genetic investigations on three specific phenotypes of teratozoospermia: macrozoospermia, globozoospermia and multiple morphological abnormalities of the flagella (MMAF).

For the first phenotype, we studied 87 patients with macrozoospermia, including 83 index cases. We identified a pathogenic mutation in the *AURKC* gene in 82% of patients (68/83) confirming that molecular alteration of this gene is the main cause of macrozoospermia. A new recurrent mutation, p.Y248*, leading to the degradation of the mutant transcripts by non-sense mediated mRNA decay was identified in 10 unrelated patients. Patients with no identified *AURKC* mutation have a decreased rate of spermatozoa with a large head and multiple flagella. Identification of two ancestral mutations in *AURKC* maintained during evolution despite their negative effect on reproduction in homozygous men, raises the question of a potential selective advantage provided by the *AURKC* haploinsufficiency.

For the second phenotype, we first analyzed 34 patients presenting with globozoospermia using MLPA (multiplex ligation-dependent probe amplification) and Sanger sequencing. In total, 22 of the 30 unrelated patients were homozygous for the *DPY19L2* deletion (73.3%) and 3 novel point mutations were identified. These results suggest that the molecular investigation of the *DPY19L2* gene in globozoospermic patients should not be limited to the detection of the *DPY19L2* genomic deletion and open interesting perspectives for the identification of *DPY19L2* partners during acrosome biogenesis. Subsequently, we demonstrated that the genomic deletion was mediated by Non-Allelic Homologous Recombination (NAHR) between two homologous 28-Kb Low Copy Repeats (LCRs) located on each side of the gene. The vast majority of genomic breakpoints fell within a 1.2-Kb region central to the 28-Kb LCR. A 13-mer consensus sequence is located in the centre of that 1.2-Kb region recognized by PRDM9, a multi-unit zinc finger binding protein that promotes the formation of double-strand breaks (DSBs) initiating the homologous recombination process. The accepted theoretical NAHR model predicts that during meiosis, NAHR produces more deleted than duplicated alleles. Surprisingly, array-CGH data show that, in the general population, *DPY19L2* duplicated alleles are approximately three times as frequent as deleted alleles. In order to shed light on this paradox, we developed a sperm-based digital PCR to measure the *de novo* rates of deletions and duplications at this locus. As predicted by the NAHR model, we identified an excess of *de novo* deletions over duplications. These seemingly discordant results may be explained by the purifying selection against sterile, homozygous deleted men. Heterozygous deleted men might also suffer a small fitness penalty.

Lastly, for the third phenotype, homozygosity mapping was carried out on a cohort of 20 North African individuals, presenting with primary infertility resulting from impaired sperm motility caused by a mosaic of multiple morphological abnormalities of the flagella (MMAF). Five unrelated subjects out of 18 (28%) carried a homozygous variant in *DNAH1*, which encodes an inner dynein heavy chain and is expressed in testis. RT-PCR and immunostaining studies confirmed the pathogenic effect of one of these mutations located on a donor splice site. Electronic microscopy revealed a general axonemal disorganization including mislocalization of the microtubule doublets and loss of the inner dynein arms suggesting that *DNAH1* plays a critical role in sperm flagellum biogenesis and assembly.

TABLE DES MATIERES

<u>Remerciements</u>	2
<u>Résumé</u>	6
<u>Abstract</u>	8
<u>Tables des matières</u>	10
<u>Table des figures</u>	14
<u>Tableaux</u>	17
<u>Liste des abréviations</u>	19
<u>Préambule</u>	21
<u>Introduction</u>	23
A. Rappels généraux sur la spermatogénèse	24
1. Fonctions et structure du testicule humain : rappels	24
2. Des cellules germinales aux spermatozoïdes	29
2.1. La phase de multiplication	29
2.2. Méiose ou phase de maturation	30
2.3. La spermiogénèse ou phase de différenciation	33
2.4. La spermiation	35
2.5. Cinétique de la spermatogénèse	35
2.6. Régulation hormonale de la spermatogénèse	37
3. Anatomie du spermatozoïde	38
3.1. La tête	38
3.1.1. L'acrosome	38
3.1.2. Morphologie de la tête : rôle du complexe acrosome-acroplaxome-manchette	40
3.2. Le noyau	44
3.3. La thèque périnucléaire	46
3.4. Le flagelle	46
3.5. L'axonème	47
3.5.1. Les dynéines axonémiales	49
3.5.1.1. Bras de dynéines	50
3.5.2. Liens de nexine et complexes de régulation des dynéines	53

3.5.3. Ponts radiaires	53
3.5.4. Le complexe central	54
3.5.5. Formation de l'axonème	55
3.6. Le manchon de mitochondrie	56
3.7. Les fibres denses	57
3.8. La gaine fibreuse	58
B. Infertilité masculine	60
1. Moyen d'exploration de l'infertilité masculine	60
1.1. Le spermogramme	60
1.2. Le spermocytogramme	62
2. Causes génétiques de l'infertilité masculine	64
2.1. Les causes génétiques fréquentes	65
2.1.1. Les microdélétions du chromosome Y	65
2.1.2 Anomalies chromosomiques	67
2.1.2.1. Syndrome de Klinefelter	67
2.1.2.2. Anomalies de structures	69
2.1.2.3. Autres anomalies chromosomiques	69
2.1.3. Mutations CFTR	70
2.2. Identification de nouveaux gènes : approches et résultats	70
2.2.1. Approche « gènes candidats »	70
2.2.2. Cartographie par homozygotie	75
2.2.3. Séquençage nouvelle génération	77
2.3. Causes syndromiques de l'infertilité	82
2.4. Infertilité et polymorphismes	82
2.5. Les défauts épigénétiques	83
2.6. MicroARN: vers de nouvelles pistes	84
Objectifs de la thèse	86
Phénotype 1: Les spermatozoïdes macrocéphales	89
<i>Article 1 “Identification of a new recurrent aurora kinase C mutation in both European and African men with macrozoospermia.”</i>	90

<u>Phénotype 2: Globozoospermie</u>	104
<u>Article 2:</u> “MLPA and sequence analysis of DPY19L2 reveals point mutations causing globozoospermia.	105
<u>Article 3:</u> “Fine characterisation of a recombination hotspot at the DPY19L2 locus and resolution of the paradoxical excess of duplications over deletions in the general population.”	119
<u>Phénotype 3: Anomalies morphologiques multiples du flagelle du spermatozoïde</u>	144
<u>Article 4:</u> « Mutations in DNAH1, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella.”	145
<u>Discussion générale et perspectives</u>	172
<u>Article (5) de revue:</u> “Teratozoospermia: spotlight on the main genetic actors in the human. “	174
<u>Références</u>	211
<u>Annexes</u>	229
Liste et résumés des publications sur la tératozoospermie non incluses dans la thèse.	230

TABLE DES FIGURES

Figure 1. Schéma anatomique et organisation du testicule humain.	25
Figure 2. Tubes séminifères.	27
Figure 3: Schéma de la barrière hémato-testiculaire.	28
Figure 4. Les différentes phases de la spermatogénèse.	30
Figure 5. Les différentes étapes de la méiose masculine.	31
Figure 6. Assemblage et dissociation du complexe synaptonémal durant la prophase I.	32
Figure 7. Principales étapes et modifications structurales lors de la spermio-génèse.	34
Figure 8. Résumé de la spermatogénèse humaine et sa cinétique.	36
Figure 9. Régulation hormonale de la spermatogénèse.	37
Figure 10. Anatomie du spermatozoïde humain.	38
Figure 11. Schéma de l'acrosome.	40
Figure 12. Schéma de l'acropaxome.	41
Figure 13. Schéma des forces de constriction endogènes et exogènes ainsi que le transport intra-manchette et intra-flagellaire de protéines au cours de l'elongation de la spermatide.	42
Figure 14 : Schéma de la compaction de l'ADN dans les cellules somatiques et dans les spermatozoïdes.	45
Figure 15. Structure du flagelle d'un spermatozoïde.	47
Figure 16. Schéma de l'axonème du flagelle de spermatozoïde.	48
Figure 17. Organisation de l'unité fondamentale de 96 nm de l'axonème.	49
Figure 18. Schéma de l'organisation des domaines d'une chaîne lourde de dynéine.	51
Figure 19. Représentation des bras de dynéine externe chez <i>Chlamydomonas</i> .	52
Figure 20. Représentation schématique des bras internes chez <i>Chlamydomonas</i> .	53
Figure 21. Structure et représentation 3D des ponts radiaires chez <i>Chlamydomonas</i> .	54
Figure 22. La structure de la paire centrale de microtubule chez <i>Chlamydomonas</i> .	55
Figure 23. Représentation schématique du transport intraflagellaire (IFT).	56
Figure 24. Observation en microscopie électronique du manchon mitochondrial.	57
Figure 25. Coupe transversal du flagelle au niveau de la pièce intermédiaire.	58

Figure 26. Représentation schématique de la gaine fibreuse.	59
Figure 27. Représentation schématique du chromosome Y et de la région AZF.	67
Figure 28. Stratégie de localisation d'un gène pathogène par cartographie d'homozygotie	76
Figure 29. Principe général du séquençage nouvelle génération ou NGS.	78
Figure 30. Principe général du système CRISPR/Cas9.	207
Figure 31. Stratégie générale d'identification et de caractérisation des gènes de l'infertilité masculine appliquée dans notre laboratoire.	208
Figure 32. Importance du diagnostic génétique dans la prise en charge du patient infertile	209
Figure 33. Illustration de la thérapie génique par utilisation de vecteurs viraux.	210

TABLEAUX

Tableau 1. Durée de vie moyenne des cellules germinales humaines.	36
Tableau 2. Limites basses de référence d'après l'OMS.	62
Tableau 3. Différentes anomalies morphologiques du spermatozoïde selon la classification de David modifiée.	63
Tableau 4. Tableau récapitulatif des causes géniques dans l'infertilité masculine.	80

LISTE DES ABREVIATIONS

A: Adénine
ADN: Acide désoxyribonucléique
ADNc: ADN complémentaire
ARNm: ARN messager
ART: Assisted reproductive technologies
AURKA: Aurora Kinase A
AURKB: Aurora Kinase B
AURKC: Aurora Kinase C
BP: Breakpoints
C: Cytosine
CGH-array: Array comparative genomic hybridization
CI: Confidence intervals
CNV: Copy Number Variation
CP: Central pair microtubules
CPC: chromosomal passenger complex
CSC: Calmodulin- and spokeassociatedcomplex
DFS: Dysplasia of the fibrous sheath
DGV: Database of Genomic Variants
DNAH1: Dynein heavy chain 1
DP: Distal protrusion
DRC: Dynein regulatory complex
DSBs: Double strand breaks
EVS: Exome Variant Server
FISH: Fluorescent in situ hybridization
FS: Fibrous sheath
G: Guanine
GSM: General Selection Model
HC: heavy chains
HRM: High-resolution melting
IC/LC: Intermediate chain/light chain
ICSI: Intracytoplasmic sperm injection
IDA: Inner dynein arms
IF : Immunofluorescence
IVF: In vitro fertilization
Kb: Kilobases
KO: Knock-out
LCR: Low-copy repeats
Mb: Mégabases
MIA: modifier of inner arms complex
MLPA: Multiplex Ligation-dependent probe amplification
MMAF: Multiple morphological abnormalities of the flagella
NAHR: Non-allelic homologous recombination
NHEJ: Non Homologous End Joining
NMD: Nonsense-mediated mRNA decay
OAT: Oligoasthenoteratozoospermia
ODA: Outer dynein arms
PCD: Primary ciliary dyskinesia
PCR: Polymerase chain reaction
PLC ζ : Sperm-specific phospholipase C ζ
RS: Radial spokes
RT-PCR: Reverse transcriptase PCR
SAC: Spindle assembly checkpoint
SNP: Single nucleotide peptide
SPZ: spermatozoa
T: Thymine
TEM: Transmission electron microscopy
Tm: Température de fusion
UTR: UnTranslated Region
WHO: World Health Organization
WT: wild type

PREAMBULE

Ce travail de thèse a été réalisé parallèlement à mon activité hospitalière dans le Laboratoire de Génétique Chromosomique au sein du Département de Génétique et Procréation (DGP) du CHU de Grenoble. L'infertilité masculine et/ou féminine constitue l'un des principaux motifs de consultation dans notre service et reste l'une des principales indications de réalisation d'un caryotype constitutionnel. Notre Département hospitalier a des liens forts avec le laboratoire de Biochimie et Génétique Moléculaire et en partie avec Pierre Ray qui est biologiste au sein de ce laboratoire. Ce dernier est également co-directeur de l'équipe de recherche « Génétique, Epigénétique et Thérapeutique de l'Infertilité » (GETI). C'est donc naturellement que ma thématique de recherche s'est centrée sur les aspects génétiques de l'infertilité masculine.

L'infertilité masculine regroupe de nombreux phénotypes dont les causes sont extrêmement variées et dans certains cas multifactorielles. Historiquement, notre équipe a commencé à travailler avec succès sur les défauts qualitatifs des spermatozoïdes ou tétratozoospermie. En effet, les gènes *AURKC* et *DPY19L2* ont pu être identifiés respectivement chez des patients présentant des spermatozoïdes macrocéphales ou une globozoospermie. Le travail de recherche que j'ai réalisé durant ma thèse fut de développer les investigations génétiques et moléculaires sur ces deux phénotypes. Enfin, j'ai pu initier l'identification de nouveaux gènes impliqués dans un troisième phénotype de tétratozoospermie, les anomalies morphologiques multiples du flagelle.

Cette thèse a été rédigée entièrement sous format « article » et s'articule autour de 5 articles (1 article de revue et 4 articles originaux). La discussion générale de la thèse a été rédigée sous la forme d'un article de revue publié dans un journal international. Ainsi, le plan général de la thèse s'articule de la façon suivante :

- Introduction
- Phénotype 1. Spermatozoïdes macrocéphales : Article 1
- Phénotype 2. Globozoospermie : Articles 2 et 3
- Phénotype 3. Anomalies morphologiques multiples du flagelle : Article 4
- Discussion générale et Perspectives : Article (5) de revue

INTRODUCTION

A. Rappels généraux sur la spermatogénèse

Chez l'homme, la spermatogenèse est un processus physiologique de différenciation cellulaire mis en place lors de la puberté qui permet la production continue et la transformation des cellules germinales en cellules hautement différencierées appelées spermatozoïdes. La spermatogenèse se déroule dans la paroi des tubes séminifères du testicule.

1. Fonctions et structure du testicule humain : rappels

Les testicules sont des organes qui possèdent principalement deux fonctions: une fonction exocrine, ou gamétogenèse, caractérisée par la production des spermatozoïdes et une fonction endocrine, avec synthèse des hormones stéroïdes sexuelles mâles, les androgènes. Chez l'homme adulte, chaque testicule est une masse ovoïde qui mesure environ 5 cm de long, 3 cm de large et 2,5 cm d'épaisseur. Comme chez de nombreuses espèces animales, ils sont situés dans une poche de peau appelée scrotum. Cette externalisation des testicules permet leur maintien à une température plus basse que celle du reste du corps nécessaire à la spermatogénèse. Les testicules sont encapsulés dans une enveloppe conjonctive dense et épaisse, connue sous le nom de tunique albuginée (figure 1). Elle est constituée de fibres de collagène et de cellules musculaires lisses responsables de contractions rythmiques spontanées qui participent à la propulsion des spermatozoïdes et du liquide testiculaire hors du testicule.

L'intérieur du testicule contient des tubes séminifères enroulés ainsi que du tissu entre les tubules appelé espace interstitiel. Les tubes séminifères sont de longs tubes compactés sous forme de boucles et dont les deux extrémités débouchent sur le *rete testis* (figure 1). Ils sont constitués d'un épithélium composé de deux types de cellules: les cellules germinales et les cellules de Sertoli qui leur servent de soutien.

Les cellules germinales sont composées des différentes cellules correspondant aux différents stades de leur différenciation en spermatozoïdes. Les cellules souches (spermatogonies de type A) sont situées à la base du tube puis viennent les spermatogonies de type B, les spermatocytes primaires et secondaires, les spermatides rondes et allongées et enfin les spermatozoïdes. Ceux-ci se trouvent sur la partie la plus interne de l'épithélium d'où ils sont libérés dans la lumière du tube (figure 2).

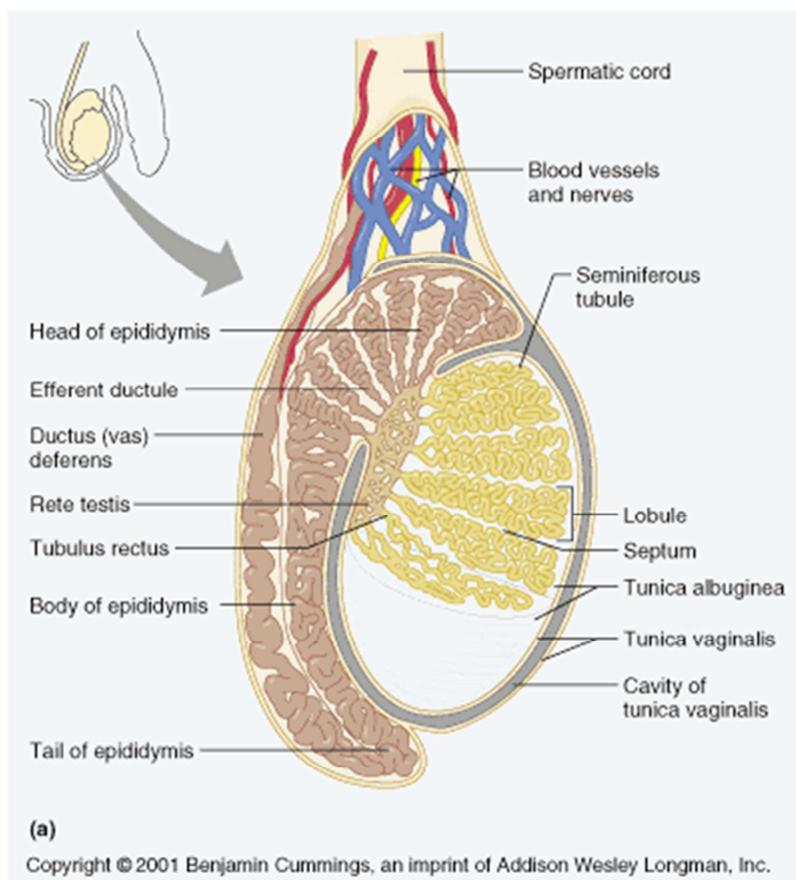


Figure 1. Schéma anatomique et organisation du testicule humain.

Les cellules de Sertoli ont leur corps cellulaire situé contre la membrane basale qui entoure les tubules séminifères (figure 2). Ces cellules possèdent des prolongements cytoplasmiques s'étirant jusqu'à la lumière du tube. La cellule de Sertoli joue donc un rôle fondamental dans le contrôle de la spermatogénèse. Leur fonction est de servir de soutien à la lignée germinale, de les nourrir, de contrôler leur développement et de participer à leur différenciation morphologique. Une cellule de Sertoli peut ainsi soutenir jusque à cinquante cellules germinales à différents stades de maturation (Mruk et Cheng 2004). Elles sont aussi responsables de la phagocytose des cellules germinales dégénérantes ainsi que des corps résiduels résultant de la compaction du cytoplasme des cellules germinales éliminés au cours de la spermio-génèse. Enfin, elles sont impliquées dans la production de protéines régulant et/ou répondant à la libération d'hormones hypophysaires (Holdcraft et Braun 2004). Les fonctions des cellules de Sertoli sont ainsi contrôlées par l'hormone hypophysaire FSH (folliculo-stimulating hormone) et la testostérone produite par les cellules de Leydig. Les cellules de Sertoli synthétisent de très nombreuses protéines différentes associées à la fonction

de la reproduction dont l'inhibine, une hormone qui exerce un rétrocontrôle négatif de la production de FSH par l'hypophyse, l'androgen-binding protein (ABP) impliquée dans les transports des androgènes et l'hormone anti-mullérienne (AMH) (Sharp *et al.*, 2003).

Au niveau de la membrane basale constituée de tissus conjonctifs se trouvent des cellules contractiles myoïdes (figure 2). Leur rôle serait d'aider à la circulation du fluide des tubes séminifères et donc au drainage des spermatozoïdes vers le *rete testis* (Russell *et al.*, 1989). Les cellules de Sertoli participent avec la membrane basale des tubes séminifères, à la barrière hémato-testiculaire.

Enfin, entre les tubes séminifères se trouve l'espace interstitiel du testicule qui est un tissu conjonctif lâche contenant de nombreux vaisseaux sanguins et lymphatiques ainsi que les nerfs du parenchyme testiculaire (figure 2). Au niveau cellulaire, l'espace interstitiel comprend des macrophages et des mastocytes et les cellules de Leydig. Les cellules de Leydig sont les principales cellules productrices des hormones androgènes comme la testostérone sous le contrôle directe de l'hormone lutéinisante (LH). La testostérone ainsi produite va diffuser dans l'épithélium des tubes séminifères où elle va se lier aux récepteurs aux androgènes situés dans les cellules de Sertoli. Elle va ainsi réguler la spermatogénèse. Les cellules de Leydig synthétisent aussi de l'ocytocine et des cytokines impliquées dans la régulation paracrine des fonctions testiculaires.

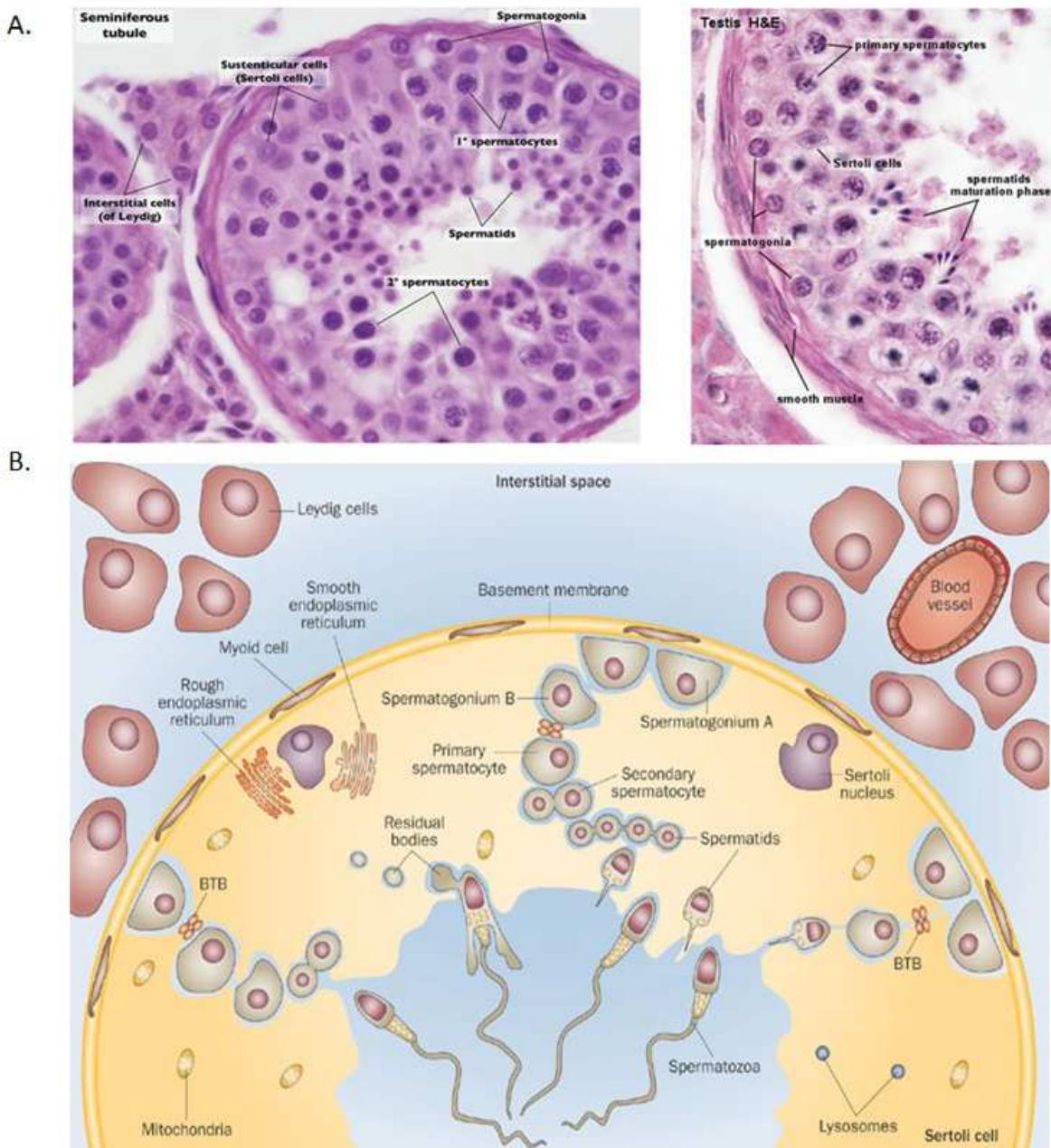


Figure 2. Tubes séminifères. A. Coupe histologique de testicules et de tubes séminifères (UWA Blue Histology); B. Schéma simplifié d'une portion transversale de tube séminifère montrant les différentes étapes de la spermatogenèse, d'après Rato *et al.*, 2012.

Chaque cellule de Sertoli est en contact avec près d'une douzaine d'autres cellules de Sertoli. Entre les cellules de Sertoli adjacentes, des jonctions serrées constituent une barrière entre deux compartiments (la barrière hémato-testiculaire, figure 3). Cette barrière sépare les spermatogonies et les spermatocytes primaires préleptotènes, situés dans le compartiment basal, des autres stades des cellules spermatogéniques, situés dans le compartiment adluminal. Un relâchement des jonctions serrées permet la migration des spermatocytes primaires

préleptotènes dans le compartiment adluminal (Cheng al., 2011). Cet arrangement crée une barrière immunologique qui sépare les cellules germinales en cours de différenciation. Elle apparaît seulement à la puberté, afin d'empêcher le système immunitaire d'être stimulé par les protéines spermatiques très immunogènes et de générer une réponse auto-immune (Kaur *et al.*, 2014). Cette barrière est maintenue dans l'épididyme.

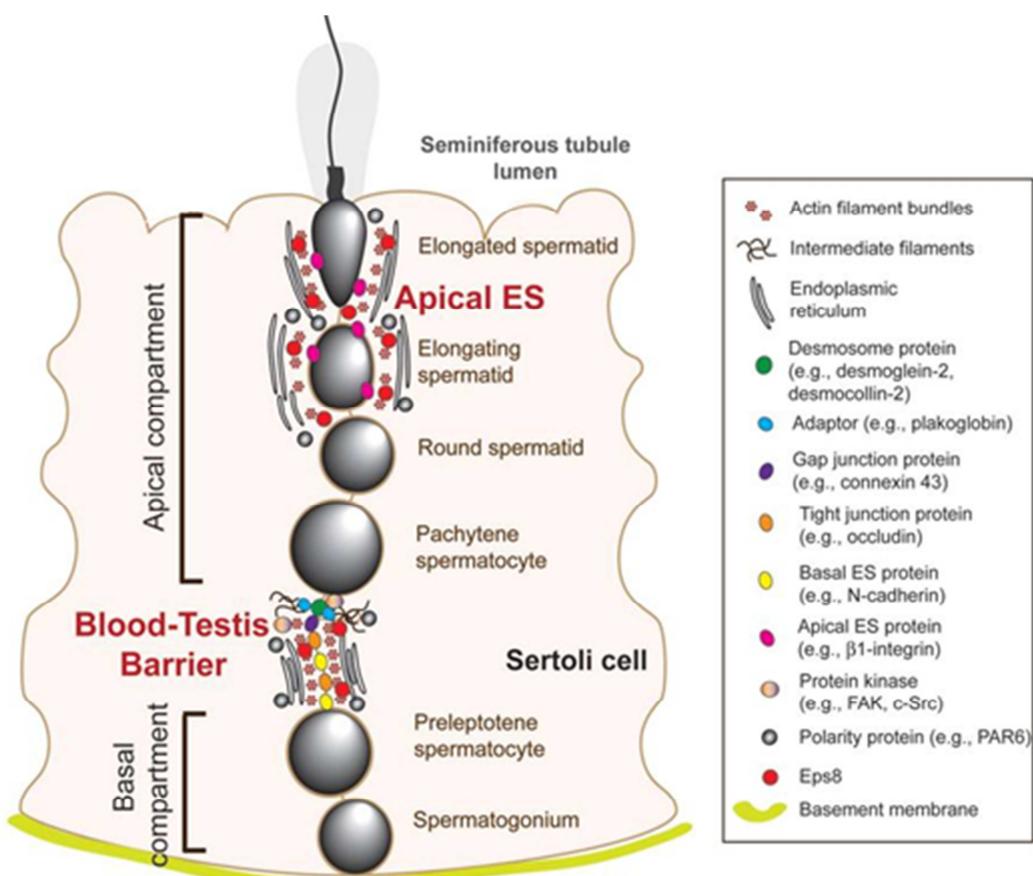


Figure 3: Schéma de la barrière hémato-testiculaire de l'épithélium des tubes séminifères dans des testicules de mammifère adulte. De manière anatomique, la barrière testiculaire divise le tube séminifère en deux compartiments, le compartiment basal et le compartiment apical (ou adluminal). Les tubes séminifères sont composés de cellules de Sertoli et de cellules germinales. La barrière testiculaire est constituée de jonctions serrées (tight junction), de desmosomes et de jonctions communicantes et de spécialisations ectoplasmiques basales. La barrière hématotesticulaire est souvent associée à des filaments d'actine parallèles à la membrane plasmique de deux cellules de Sertoli adjacentes et prise en sandwich en deux citernes de réticulum endoplasmique ces structures sont appelées les spécialisations ectoplasmiques basales (Cheng al., 2011).

2. Des cellules germinales aux spermatozoïdes

La production des spermatozoïdes à partir des cellules germinales se déroule en 3 phases : la multiplication, la méiose (appelée aussi phase de maturation nucléaire) et la spermiogénèse (appelée aussi phase de différenciation). Trois types de cellules germinales sont impliqués dans la spermatogenèse : les spermatogonies, les spermatocytes et les spermatides.

2.1. La phase de multiplication

Au cours de la phase de multiplication, les cellules germinales souches ou spermatogonies se divisent par mitoses pour aboutir au stade de spermatocytes primaires. Les spermatogonies sont de cellules diploïdes disposées en périphérie des tubes séminifères. Ce sont des cellules arrondies, d'un diamètre de 10 à 15 µm, avec un cytoplasme clair et un noyau ovoïde. Les analyses histologiques ont permis de distinguer trois types de spermatogonies (Clermont 1963) :

- les spermatogonies de type Ad (d=dark) d'aspect sombre et qui possèdent un noyau arrondi avec une chromatine dense, finement granuleuse, occupée par une vacuole centrale.
- les spermatogonies de type Ap (p=pâle) qui ont un noyau ovalaire, clair, avec une chromatine fine et dispersée, renfermant un ou plusieurs nucléoles.
- les spermatogonies de type B qui sont caractérisées par un noyau arrondi, foncé avec une chromatine en mottée et irrégulière.

Les spermatogonies Ad se divisent en une spermatogonie Ap et en une autre spermatogonie Ad (figure 4). Cette propriété permet à la fois de se différencier en spermatocytes tout en constituant un compartiment de réserve de spermatogonies Ad permettant d'assurer la production permanente de spermatozoïdes. Les spermatogonies Ap ainsi produites se diviseront successivement en deux spermatogonies B qui elles-mêmes se diviseront en deux spermatocytes primaires diploïdes (figure 4). Les spermatogonies issues d'une même génération sont reliées entre elles par des ponts cytoplasmiques qui persisteront entre ces cellules tout au long de la spermatogénèse.

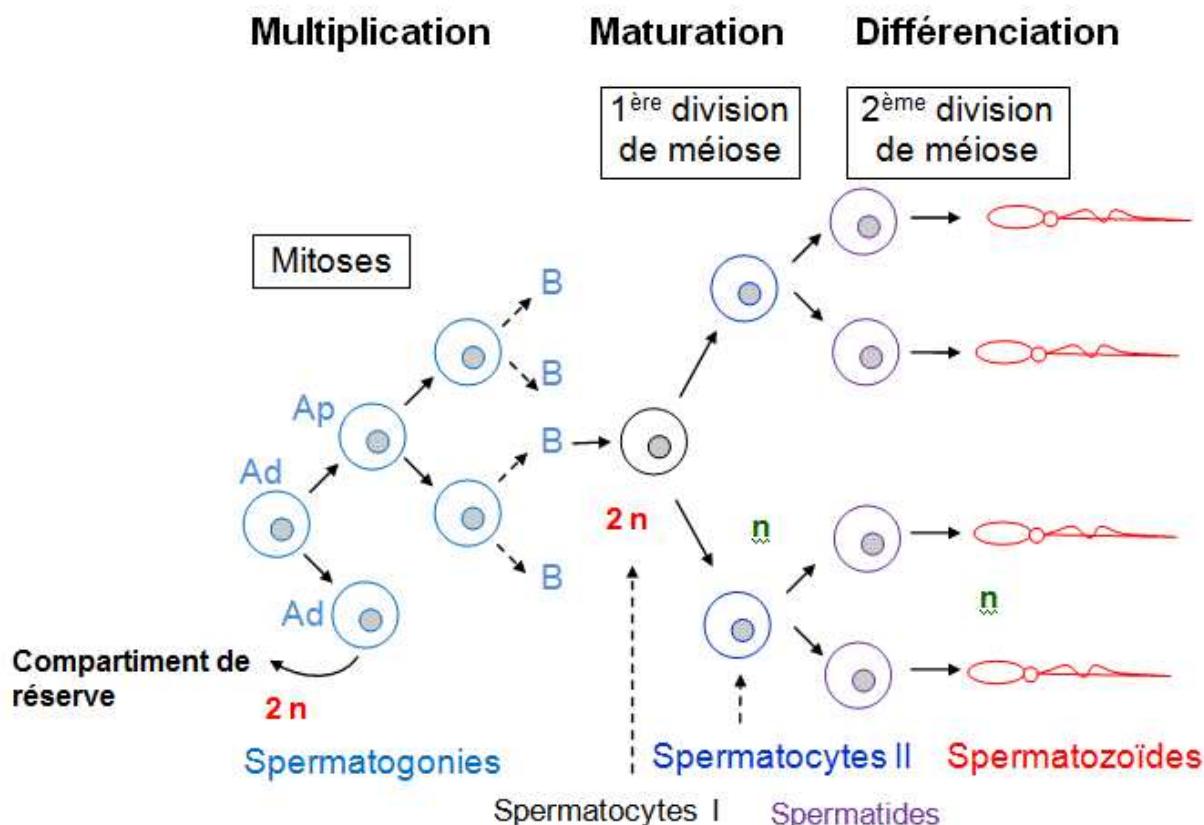


Figure 4. Les différentes phases de la spermatogénèse. Ad : spermatogonies Ad avec noyau sombre ; Ap : spermatogonies Ap avec noyau pâle. B : spermatogonies de type B ; n : lot haploïde de chromosomes.

2.2. Méiose ou phase de maturation

La méiose ou phase de maturation concerne les deux générations de spermatocytes : les spermatocytes primaires (spermatocytes I) et les spermatocytes secondaires (spermatocytes II).

La méiose consiste en deux divisions successives (figure 5) : la méiose I (MI) ou méiose réductionnelle aboutissant à la séparation des chromosomes homologues parentaux appariés et à la formation des spermatocytes secondaires (spermatocytes II) et la méiose II (MII) ou méiose équationnelle permettant la séparation de chacune des chromatides sœurs et à la formation des spermatides. Le rôle de la méiose est d'assurer la réduction chromatique et le brassage et la transmission de l'information génétique.

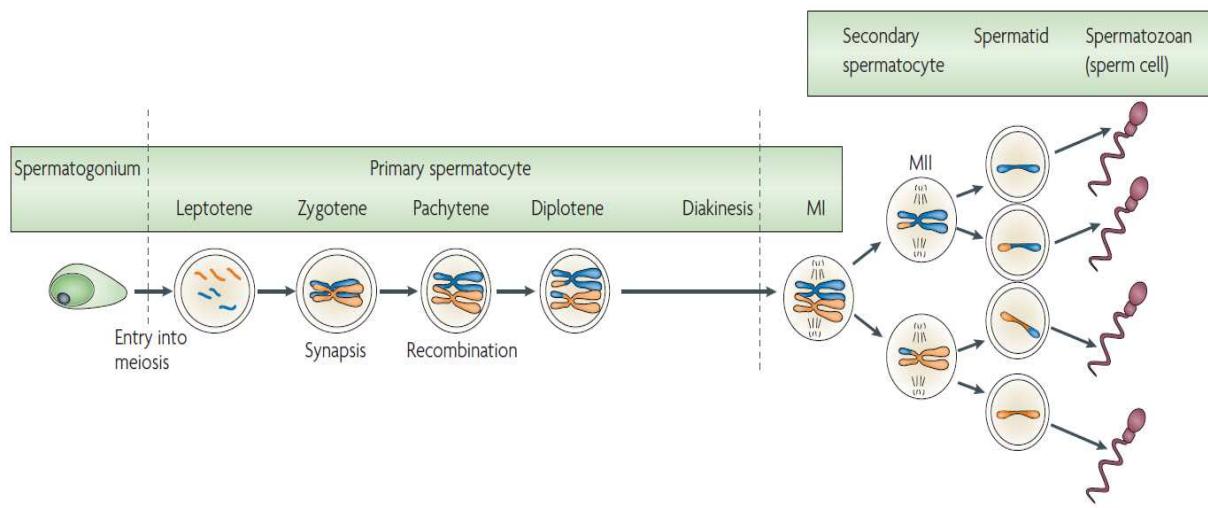


Figure 5. Les différentes étapes de la méiose gamétique masculine. D'après Sasaki et Matsui, 2008.

La méiose va être initiée par les spermatocytes primaires issus de la dernière division des spermatogonies B. Les spermatocytes primaires sont de grandes cellules au noyau arrondi contenant plusieurs nucléoles. Les spermatocytes primaires, qui se situent dans le compartiment basal du tube séminifère, ont une courte interphase (stade pré-leptotène qui dure entre 2 et 4 jours), pendant laquelle s'effectue une réplication d'ADN. Les chromosomes sont alors constitués de deux chromatides liées par leur centromère. À leur entrée en phase de méiose, les spermatocytes I se situent dans le compartiment adluminal du tube séminifère. Par la suite, chacune des deux divisions méiotiques est décomposée, comme lors la mitose, en 4 étapes distinctes : prophase, métaphase, anaphase, télophase.

La prophase de la méiose I ou prophase I est une étape longue qui dure chez l'homme 23 jours et se divise en 5 stades successifs : leptotène, zygotène, pachytène, diplotène et diacinese. Très schématiquement, les chromosomes homologues après avoir commencé à se condenser durant le stade leptotène, vont s'appareiller (synapsis) durant le stade zygotène par l'intermédiaire d'une structure multi-protéique : le complexe synaptonémal (figure 6). Les paires de chromosomes homologues appareillés sont appelées des bivalents. C'est au stade de pachytène que se produit l'échange de matériel génétique (crossing-over) entre les chromatides non-sœurs appelés noudes de recombinaison. Le stade de pachytène est le stade le plus long et dure 16 jours chez l'homme. Les chromosomes vont commencer à se séparer au stade diplotène en raison de la dissociation du complexe synaptonémal. Les chromosomes homologues conservent cependant encore

certains sites d'appariement étroits nommés chiasmas. On notera d'ailleurs que la présence de ces chiasmas assure une séparation plus progressive des chromosomes. En l'absence de ces chiasmas, la séparation rapide et prématûre des chromosomes favorisent la survenue d'aneuploïdies (nombre anormal de chromosomes) (Handysite, 2012). Enfin, la diacénèse marque la fin de la prophase I et la transition avec la métaphase I. Elle est caractérisée par une condensation maximale des chromosomes et la disparition de la membrane nucléaire et du nucléole. Le fuseau méiotique commence à s'assembler, les centromères des chromosomes homologues s'éloignent et les chiasmas glissent progressivement vers les télosomes (terminalisation des chiasmas).

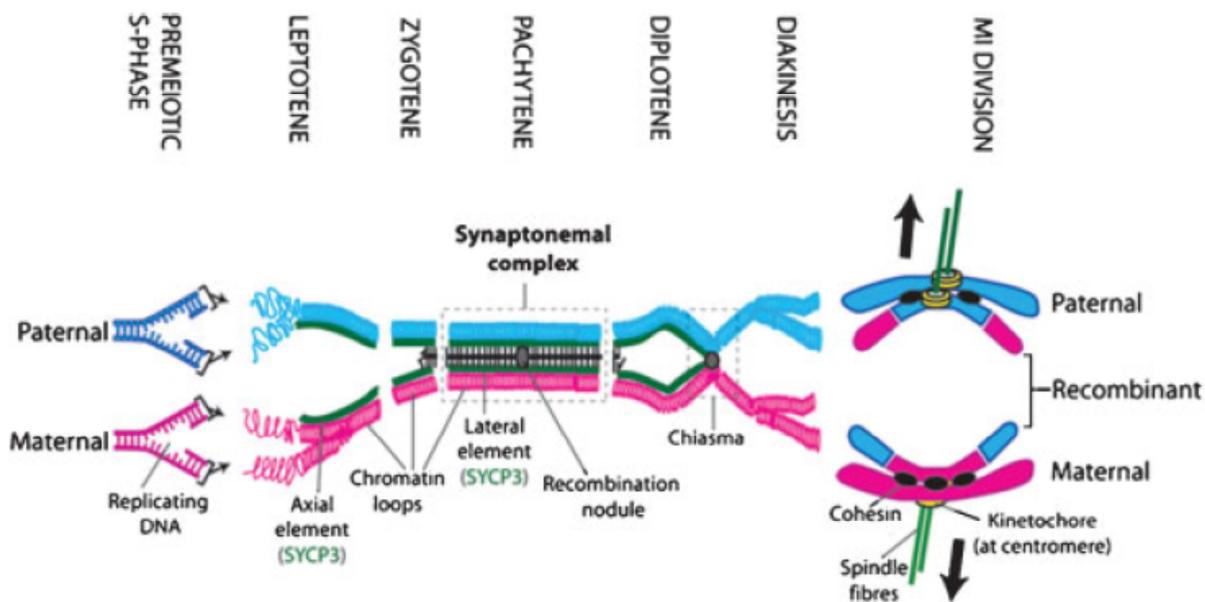


Figure 6. Assemblage et dissociation du complexe synaptonémal durant la prophase I.
D'après Burgoyne *et al.*, 2007

Lors de la métaphase I, les chromosomes sont positionnés de part et d'autre de la plaque équatoriale. Seuls les chiasmas sont encore alignés sur la plaque équatoriale.

En anaphase I, les chromosomes homologues se séparent et s'orientent vers les pôles du fuseau de façon aléatoire.

En télophase I, les deux contingents de chromosomes sont situés à des pôles opposés et la cellule se divise en deux grâce à un anneau contractile composé notamment d'actine (cytodiérèse). Cette cytodiérèse est cependant incomplète avec persistance d'un pont cytoplasmique entre les deux cellules filles, appelées spermatocytes secondaires.

La première division méiotique about ainsi à la formation de spermatocytes

secondaires dans lesquels chaque chromosome est formé de deux chromatides sœurs. Les spermatocytes II ont une existence brève (1 jour) ce qui les rend très rarement observables en histologie sur les coupes de biopsie testiculaire. Après une très courte interphase sans réplication d'ADN, les spermatocytes secondaires vont entrer en deuxième division méiotique. Cette deuxième division est très semblable à une division mitotique. La prophase II, à la différence de la prophase I, est très courte. Lors de cette étape, les chromosomes constitués de chromatides sœurs se dirigent vers la plaque équatoriale. En métaphase II, les chromosomes s'alignent au niveau de leurs centromères. En anaphase II, le centromère de chaque chromosome se rompt séparant les chromatides sœurs l'une de l'autre et permettant leur déplacement vers les pôles opposés des spermatocytes II. Lors de la télophase II, on observe la formation de cellules filles haploïdes appelées spermatides, contenant chacune n chromosomes.

2.3. La spermiogénèse ou phase de différenciation

La spermiogénèse est la phase de différenciation post-méiotique des spermatides en spermatozoïdes. Chez l'homme, elle dure environ 23 jours et peut être subdivisée en plusieurs étapes (figure 7). Lors de la spermiogénèse, les spermatides vont subir une série de modifications nucléaires, cytoplasmiques et membranaires. La spermiogénèse est caractérisée par trois évènements majeurs : la compaction de l'ADN nucléaire, la formation de l'acrosome et la formation du flagelle.

Les spermatides sont situées dans le compartiment adluminal, à proximité de la lumière du tube séminifère. Les spermatides sont incluses dans les dépressions (cryptes) cytoplasmiques des cellules de Sertoli auxquelles elles sont reliées par des jonctions communicantes. Les spermatides sont des cellules de petites tailles, de 8 à 10 µm que l'on peut schématiquement diviser en trois classes :

- les spermatides rondes (figure 7, étapes 1-2), possèdent un noyau rond avec une chromatine pâle et homogène. C'est à partir de ces étapes que démarre la biogénèse de l'acrosome avec la production par l'appareil de Golgi des vésicules pro-acrosomales (phase de Golgi). Les deux centrioles contenus dans le cytoplasme vont se déplacer au futur pôle caudal. Le centriole proximal est inactif alors que le centriole distal donne naissance à un ensemble de microtubules à l'origine de l'axonème du futur flagelle;

- les spermatides en élongation (figure 7, étapes 3-4), le noyau s'allonge et la chromatine devient plus sombre. L'acrosome va s'étendre le long du noyau (phase de la

coiffe). Des microtubules se forment autour du noyau et s'organisent en faisceaux formant la manchette. Cette structure transitoire se développe du pôle antérieur (face à l'acrosome) vers le pôle opposé et permet un transport actif de l'appareil de Golgi, localisé au pôle antérieur vers le flagelle en formation. Elle va aussi participer à l'allongement de la tête du spermatozoïde et permettre la migration des mitochondries vers la pièce intermédiaire du flagelle où elles se regrouperont pour former le manchon de mitochondries (Moreno *et al.*, 2006). Finalement, l'appareil de Golgi arrête la synthèse de protéines et migre aussi via la manchette du côté postérieur avec l'axonème.

- les spermatides en condensation (figure 7, étapes 5-7), le noyau est très allongé, avec une partie caudale globulaire et une partie antérieure saillante. La chromatine est sombre et condensée. L'axonème va continuer à s'allonger pour former le flagelle mature. Les différentes organelles inutiles pour la physiologique spermatique et l'excès de cytoplasme vont former la gouttelette cytoplasmique qui va se détacher et donner le corps résiduel qui va ensuite être phagocyté par les cellules de Sertoli (Hermo *et al.*, 2010).

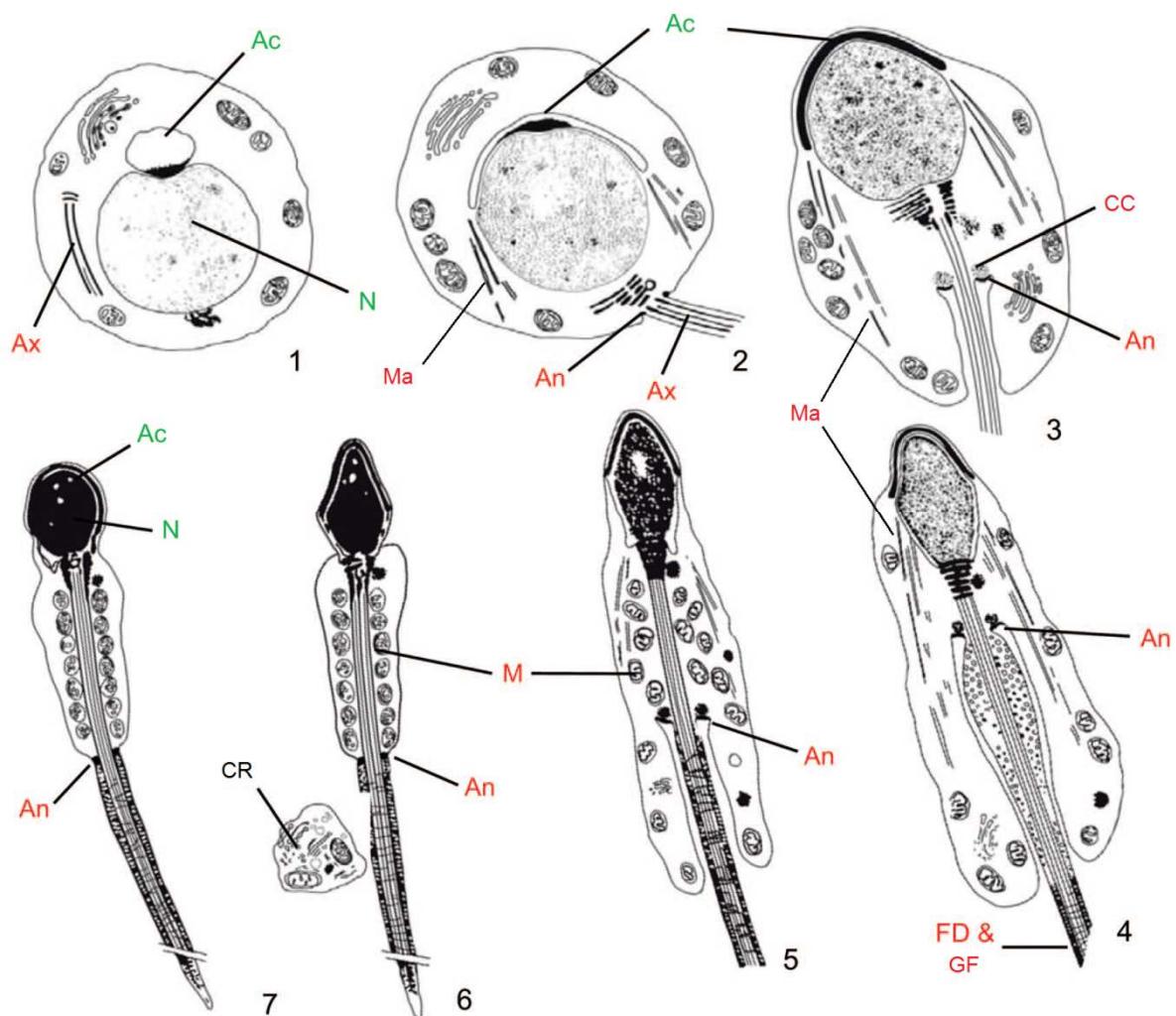


Figure 7. Principales étapes et modifications structurales lors de la spermiogénèse. 1. La spermatide immature avec un gros noyau arrondi. La vésicule acrosomale est attachée au noyau, l'ébauche du flagelle n'atteint pas le noyau. 2. La vésicule acrosomale a augmenté de taille et apparaît aplatie au niveau du noyau. Le flagelle entre en contact avec le noyau. 3-7. Formation de l'acrosome, condensation du noyau et développement des structures flagellaires. Ac, acrosome; Ax, axonème; CC, corps chromatoïdes; CR, corps résiduel; FD, fibres denses; GF, gaine fibreuse; M, mitochondrie; Ma, manchette. D'après Touré *et al.*, 2011.

2.4. La spermiation

Une fois les étapes de la spermiogénèse terminée, le spermatozoïde mature va être libéré dans la lumière des tubes séminifères. Cette phase est appelée spermiation. Elle est régulée par des hormones comme la FSH et la testostérone (O'Donnell *et al.*, 2011). Les spermatozoïdes libérés dans la lumière du tube séminifère vont être drainés jusqu'à l'épididyme via le canal efférent. Lors de leur passage dans l'épididyme les spermatozoïdes vont subir les premières modifications biochimiques (Kopera *et al.*, 2010). Les observations les plus notables sont une acquisition progressive de la motilité et une diminution de la taille du noyau. La composition de la membrane plasmique va aussi être modifiée (Jones *et al.*, 2007). Après l'éjaculation, le spermatozoïde subira encore de dernières modifications lors de la traversée des voies génitales femelles comme la capacitation. La capacitation correspond à l'ensemble des modifications physiologiques post-éjaculation qui vont permettre au spermatozoïde d'acquérir son pouvoir fécondant. A la suite de la capacitation il va être capable de réaliser la réaction acrosomique nécessaire à la pénétration de la zone pellucide. Il devient aussi capable d'entrer dans un état d'hyperactivation flagellaire qui lui permettra de progresser dans le cumulus et de traverser la zone pellucide (De Jonge, 2005). L'ensemble de ces processus sont mal connus, mais certains mécanismes ont été mis à jour et semblent impliquer des modifications de la composition lipidique de la membrane plasmique et un fort rôle du calcium (Bailey 2010).

2.5. Cinétique de la spermatogénèse

La spermatogenèse humaine dure 74 jours. Le cycle spermatogénétique est défini comme la succession chronologique des différents stades de différenciation d'une génération de cellules germinales (depuis la spermatogonie jusqu'au spermatozoïde). Chacune des étapes

du cycle spermatogénétique a une durée fixe et constante selon les espèces (tableau 1).

Tableau 1. Durée de vie moyenne des cellules germinales humaines

Cellules germinales	Durée de vie moyenne (jours)
Spermatogonies Ap	16-18
Spermatogonie B	7.5-9
Spermatocytes primaires	23 (dont 16 jours en pachytène)
Spermatocyte secondaires	1
Spermatides	23

En réalité, les cellules souches entrent en spermatogénèse périodiquement et à des intervalles réguliers. Chez l'homme, le départ d'une nouvelle lignée est déclenché tous les 16 jours : ceci constitue le cycle de l'épithélium séminal et explique l'association préférentielle de certaines cellules entre elles à un stade donné. Ainsi, on distingue 6 associations cellulaires possibles (ou stades) au cours de la spermatogenèse humaine (figure 8). Globalement, 25 000 spermatozoïdes sont produits chez l'homme chaque minute (Amann, 2008).

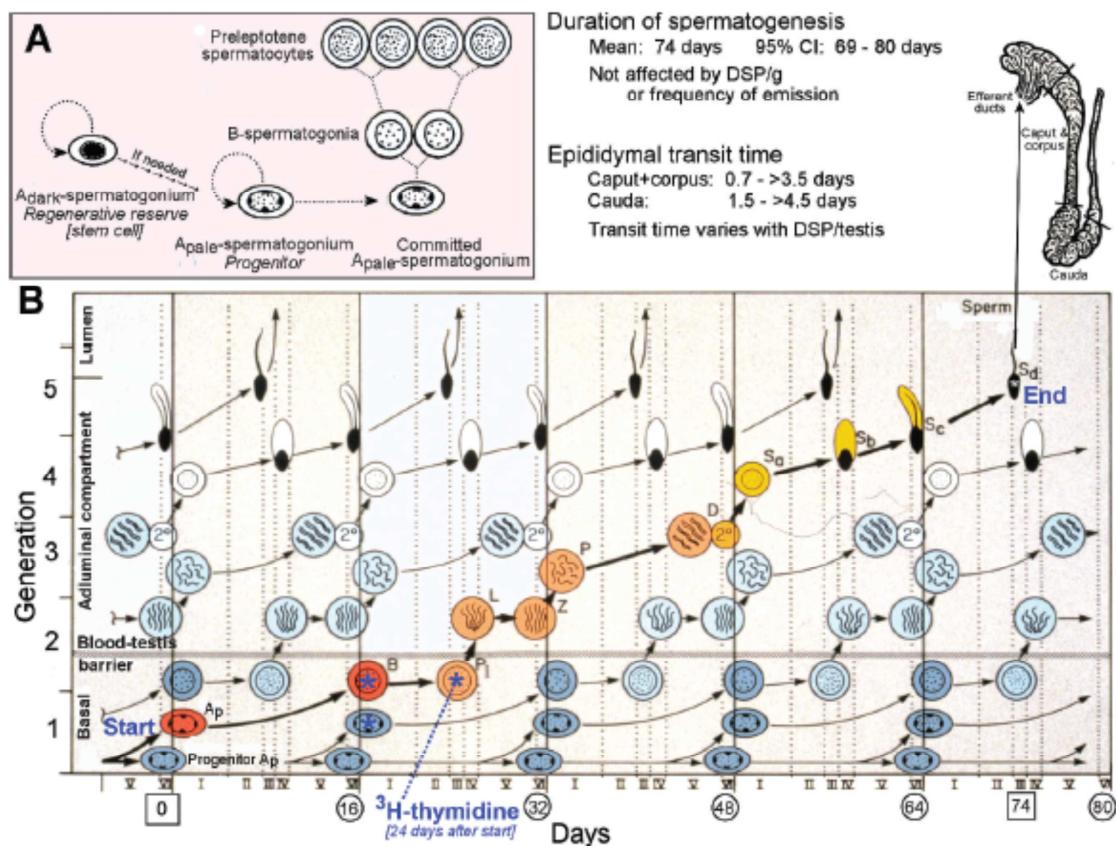


Figure 8. Résumé de la spermatogénèse humaine et sa cinétique. D'après Amann, 2008.

2.6. Régulation hormonale de la spermatogénèse

L'axe hypothalamo-hypophysio-gonadique contrôle la spermatogenèse. La régulation de la spermatogenèse implique la sécrétion de gonadolibérine (GnRH) par l'hypothalamus. La GnRH stimule ensuite la sécrétion à la fois de l'hormone folliculo-stimulante (FSH) et de l'hormone lutéinisante (LH) par l'hypophyse antérieure (de Kretser *et al.*, 1998). Cette dernière agit directement sur les cellules de Leydig et va stimuler la production de testostérone qui va agir sur les cellules germinales (figure 9). Les cellules germinales ne possédant pas de récepteurs pour la testostérone, les cellules de Sertoli et les cellules péri-tubulaires vont produire l'ABP (androgène-binding protein), une protéine de liaison aux androgènes qui va permettre l'internalisation des hormones par les spermatogones (Sofikitis, Giotitsas *et al.*, 2008). Les cellules de Sertoli produisent aussi l'activine qui est capable de stimuler directement la production de FSH. La spermatogenèse est également régulée par un retrocontrôle négatif à plusieurs niveaux. En effet, la testostérone inhibe directement la production des hormones par l'hypothalamus et l'hypophyse. Enfin, les cellules de Sertoli peuvent également stimuler la production d'inhibine inhibant directement la production de FSH par l'hypophyse (Figure 9).

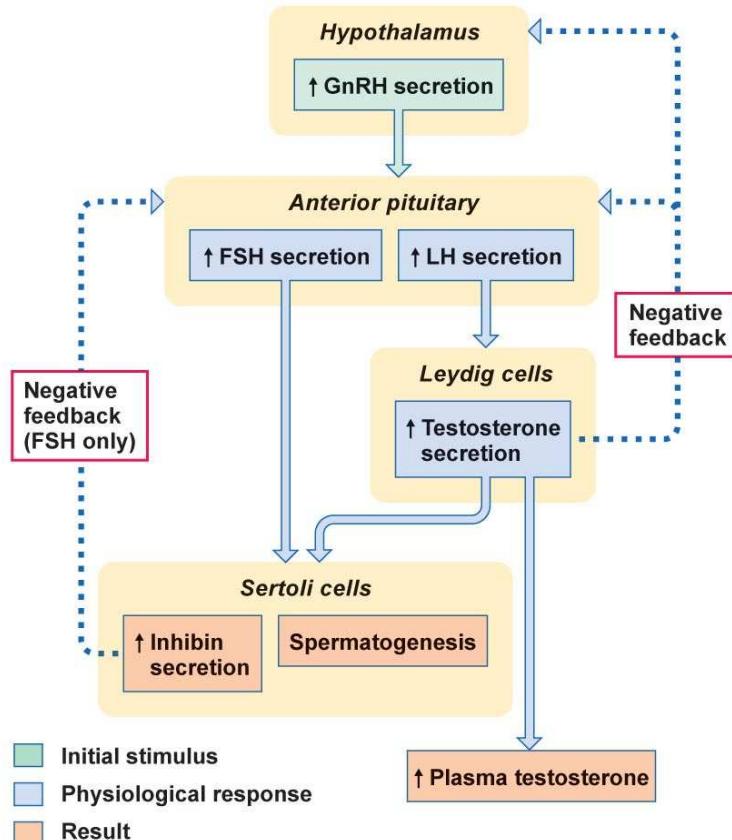


Figure 9. Régulation hormonale de la spermatogénèse (Pearson Education, 2011).

3. Anatomie du spermatozoïde

Le spermatozoïde, est une cellule hautement différenciée et orientée, avec une taille, une forme et des axes de symétrie déterminés. La morphologie générale du spermatozoïde éjaculé est similaire à celle du spermatozoïde testiculaire. Le spermatozoïde humain normal mature mesure environ 60 µm de long et est essentiellement constitué de trois parties : la tête, le cou et le flagelle (figure 10). Le flagelle est lui-même partitionné en trois parties : la pièce intermédiaire, la pièce principale et la pièce terminale.

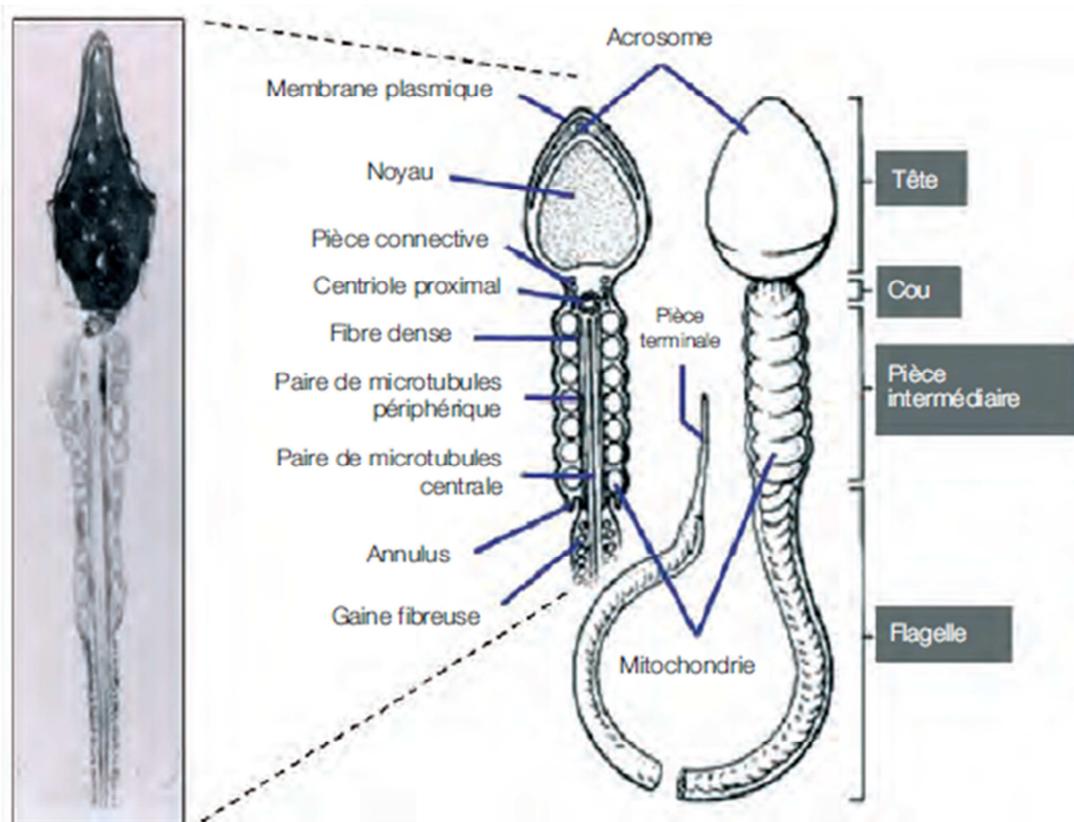


Figure 10. Anatomie du spermatozoïde humain (bioforma.org).

3.1. La tête

3.1.1. L'acrosome

L'acrosome est une vésicule géante de sécrétion qui recouvre entre 40 et 70 % du noyau délimitée par une double membrane interne et externe autour de la matrice

acrosomiale. Son rôle est fondamental dans le processus de la fécondation en permettant la progression du spermatozoïde à travers la zone pellucide de l'ovocyte. La synthèse des protéines impliquées dans la biogénèse de l'acrosome débute au stade pachytène et se poursuit jusqu'au stade tardif de la spermiogénèse (Anakwe et Gerton, 2009). De nombreux modèles de souris KO ont permis de mettre en lumière de nombreux acteurs de la biogénèse de l'acrosome. Cette partie est plus longuement détaillée dans l'article de revue « Teratozoospermia: spotlight on the main genetic actors in the human » dans la partie discussion de la thèse.

L'acrosome est apparenté à la famille des lysosomes. Cette vésicule a un pH acide et contient de nombreuses enzymes hydrolytiques (hyaluronidase, neuraminidase, phosphatase acide) qui sont libérées lors de son exocytose (Berruti et Paiardi, 2011). L'acrosome contient aussi de nombreuses enzymes spécifiques comme l'acrosine, l'acrin1, l'acrogranine (Moreno et Alvarado, 2006).

La membrane qui entoure l'acrosome est divisée en deux parties : la membrane de l'acrosome interne (MAI) et la membrane de l'acrosome externe (MAE) (figure 11). La MAI est la partie la plus interne de l'acrosome, située juste au-dessus du noyau. Elle est ancrée à la surface du noyau grâce à une structure de cytosquelette, la thèque périnucléaire (Yoshinaga et Toshimori, 2003). La thèque périnucléaire contient la phospholipase zeta qui est responsable de la cascade d'activation de l'ovocyte dont l'élément essentiel est l'initiation d'oscillations du flux de Ca^{2+} (Escoffier *et al.*, 2015). La membrane de l'acrosome externe (MAE) correspond à la partie de la membrane qui sépare l'acrosome du cytoplasme. Au cours de la réaction acrosomique (i.e. l'étape d'exocytose de l'acrosome) la MAE fusionne avec la membrane plasmique grâce au complexe de protéines spécialisées dans la fusion membranaire, les protéines SNARE. Il faut noter que la fusion membranaire de l'acrosome fait intervenir de très nombreuses protéines spécifiques au spermatozoïde comme Munc18 (Tomes *et al.*, 2015). Après la réaction acrosomique, il ne reste plus que la MAI qui demeure intacte jusqu'à la fusion de la spermatide avec l'ovocyte (figure 11) (Yoshinaga et Toshimori, 2003). Il faut noter qu'afin d'assurer l'intégrité de l'homéostasie cellulaire, la MAI fusionne avec la membrane plasmique dans la partie postérieure de l'acrosome. La MAI joue alors le rôle de membrane plasmique jusqu'à la fusion du spermatozoïde avec l'ovocyte.

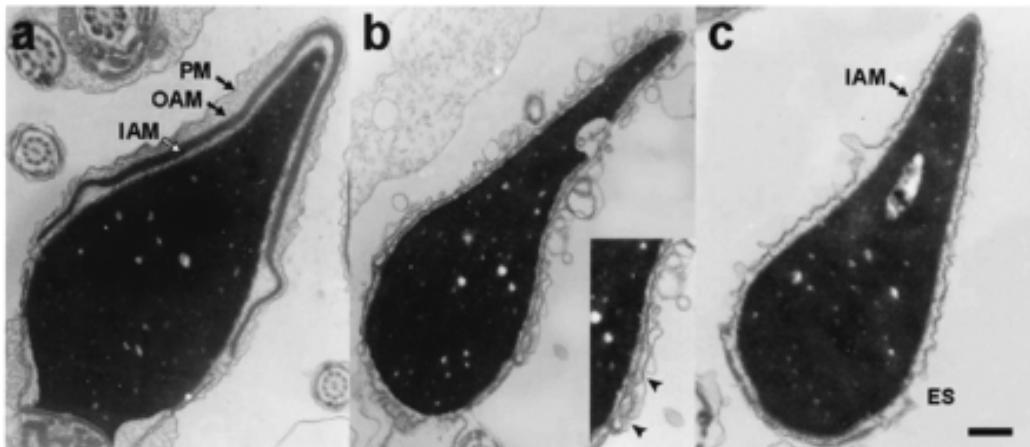


Figure 11. Schéma de l'acrosome et de ces différentes membranes avant et après réaction acrosomique. (A) acrosome intact, (B) au cours de la réaction acrosomique et (C) après la réaction acrosomique. La vésicule de l'acrosome est constituée de la membrane externe de l'acrosome (OAM) située sous la membrane plasmique (PM) et la membrane interne de l'acrosome (IAM) collée au noyau. D'après Michaut *et al.*, 2000.

3.1.2. Morphologie de la tête : rôle du complexe acrosome-acroplaxome-manchette

L'acquisition de la morphologie de la tête spermatique est contrôlée par le complexe acrosome – acroplaxome-manchette (Kierszenbaum et Tres, 2004). L'acroplaxome est une structure de cytosquelette constituée de microfilaments d'actine (F- actine) et de kératine 5 positionnés en face de l'appareil de golgi et contre le noyau. L'apparition de cette structure précède l'attachement et la fusion de vésicules pro-acrosomales auxquelles elle servirait de point d'attachment. Elle se développe et s'étend sur le noyau à partir du stade précoce de spermatide ronde. Elle est constituée de deux parties : l'anneau marginal qui se situe à l'interface entre l'acrosome et le noyau et la plaque acroplaxomique qui occupe tout l'espace sub-acrosomique (figure 12) et servirait de point d'ancre et de guide aux vésicules pro-acrosomales (Kierszenbaum et Tres, 2004). C'est une structure transitoire qui disparaît remplacée par la thèque périnucléaire dans le spermatozoïde mature. La manchette est constituée de deux structures de tubuline associées. Un anneau de tubuline alpha, l'anneau périnucléaire, encerclant le noyau juste en dessous de l'anneau marginal de l'acroplaxome et un réseau de tubulines delta attachés à l'anneau périnucléaire descendant dans le cytoplasme de la spermatide en élargissement et parallèlement au flagelle en formation (figure 12). Les

spermatozoïdes de différents modèles murins inactivées pour des gènes codant pour des protéines de la manchette (Hook1, Rimbp3, E-MAP-115, Clip-170) présentent toutes des anomalies morphologiques de la tête et du noyau renforçant l'importance de cette structure dans la morphogénèse de la tête et/ou du noyau spermatique (Komada *et al.*, 2000 ; Mendoza-Lujambio *et al.*, 2002, Akhmanova *et al.*, 2005 ; Zhou *et al.*, 2009).

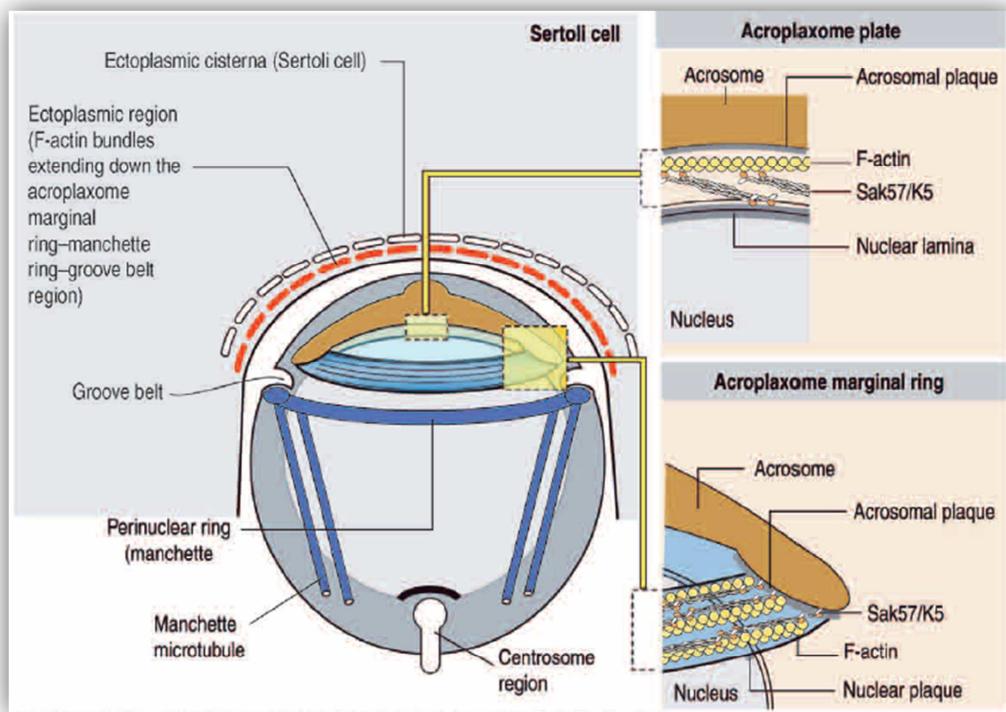


Figure 12. Schéma de l'acoplaxome. L'acoplaxome est une plaque située entre l'acrosome et le noyau et constituée de cytosquelette d'actine (F-actine) et de kératine 5/Sak57. D'après (Kierszenbaum et Tres, 2004).

Le noyau va s'allonger sous l'action de forces externes exercées par la cellule de Sertoli qui l'entoure et par des forces internes à la spermatide (figure 13). Les forces internes de remodelage de la spermatide sont les forces d'elongation engendrées par l'extension de la manchette ainsi que des forces de constriction engendrées par l'anneau marginal de l'acoplaxome et périnucléaire de la manchette. Egalelement, à ce stade de maturation des spermatides, les cellules de Sertoli dans lesquelles sont insérés les spermatozoïdes vont développer des citernes de réticulum endoplasmique auxquelles sont associées des boucles de F-actine. Ces boucles enserrent la tête du spermatozoïde et génèrent des forces qui sont transmises au spermatozoïde grâce à des jonctions adhérentes (figure 13). La dysfonction de molécules responsables des jonctions adhérentes comme certaines nectines par exemple, ont

été décrites pour entraîner des malformations de la tête du spermatozoïde (Ozaki-Kuroda *et al.*, 2002).

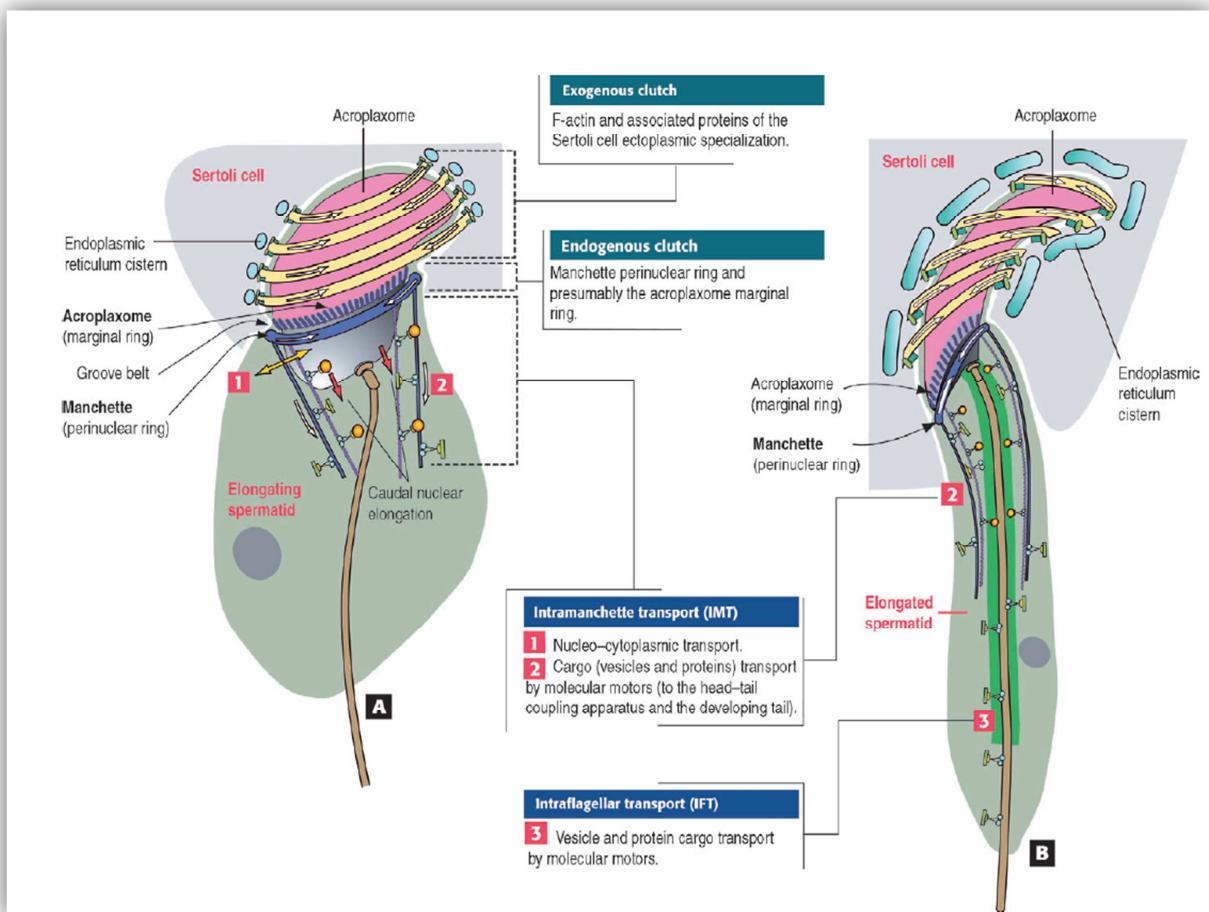


Figure 13. Schéma des forces de constrictions endogènes et exogènes ainsi que le transport intra-manchette et intra-flagellaire de protéines au cours de l'élongation de la spermatide.
A. Des boucles d'actines situées dans les cellules de Sertoli enveloppent la tête de la spermatide en élongation (forces de constrictions exogènes). L'anneau marginal de l'acroplaxome ainsi que l'anneau périnucléaire de la manchette voient leurs diamètres diminuer au cours de l'élongation de la spermatide (forces de constrictions endogènes). B. La manchette s'étend progressivement dans le cytoplasme et sert de rail pour le transport de vésicules et de protéines nécessaires au développement du centrosome et du flagelle. D'après Kierszenbaum et Tres, 2004.

3.2. Le noyau

Le noyau du spermatozoïde est caractérisé par une compaction extrêmement importante de l'ADN. Dans les cellules somatiques l'ADN est enroulé par unité de 146 paires de bases autour d'un octamère d'histones dit de cœur (H2A, H2B, H3 et H4) afin d'organiser les 3 milliards de paires de bases du génome humain dans un noyau de quelques microns.. Cette association forme le nucléosome qui constitue le premier niveau de compaction de la chromatine. La répétition des nucléosomes constituera ensuite une fibre de 30 nm de diamètre, la fibre chromatinienne qui représente le niveau supérieure de compaction et qui pourra se replier selon différents modèles dit en solénoïde ou en zig-zag (figure 14) (Wolffe 1995).

L'ADN des spermatides va subir une réorganisation chromatinienne plus importante au cours de la spermatogénèse afin d'augmenter sa compaction. Cette compaction extrême du génome paternel à plusieurs rôles. Il permet de réduire la taille du noyau, mais aussi de protéger l'ADN d'agents de dégradation comme l'oxydation des bases. Cette réorganisation s'articule en trois temps (figure 14). Dans un premier temps, des modifications épigénétiques comme l'acétylation, la méthylation ou l'ubiquitination de certains résidus des histones et notamment des résidus lysines vont participer fortement à la modification de la structure de la chromatine en la faisant passer d'un état condensé à un état décondensé. Cette décondensation permet l'échange des histones canoniques en histones testiculaires (Gaucher *et al.*, 2010). Plusieurs variants ont été décrits pour les histones de cœur H2A, H2B, et H3, et pour l'histone de liaison H1. Certains sont somatiques d'autres sont spécifiques du testicule comme H2AL1 and H2AL2 (Govin *et al.*, 2004; Rousseaux *et al.*, 2005). L'incorporation de ces variants va déstabiliser les nucléosomes. (Govin *et al.*, 2004). Ces étapes vont permettre dans un deuxième temps le remplacement de plus de 90% des variants d'histones par des protéines basiques, les protéines de transition (TP1 et TP2) puis dans un troisième temps par des protamines (PRM1, PRM2), protéines encore plus basiques, riches en arginines et cystéines (Carrell 2012). Cela a pour conséquence l'hyper-compaction de l'ADN des spermatides. Seul 1% du génome chez la souris reste compacté sous forme de nucléosome (van der Heijden *et al.*, 2005), chez l'homme cette proportion est estimée entre 10 et 15% (Brykczynska *et al.*, 2010). Le remplacement des histones par les protéines de transition est contrôlé par une hyperacétylation globale de la chromatine, elle-même contrôlée par la protéine Brdt (Bromodomaine factor1) (Pivot-Pajot *et al.*, 2003 ; Gaucher *et al.*, 2012). La protéine Brdt est non seulement importante pour l'organisation de l'ADN spermatique mais elle permet

également le contrôle de l'expression de nombreux gènes au cours de la phase méiotique. Cette protéine a donc une dualité de fonction qui la rend essentielle au cours de la spermatogénèse (Berkovits et Wolgemuth, 2012). Les séquences d'ADN qui restent attachées aux nucléosomes ne sont pas dues au hasard, mais semblent au contraire correspondre à des régions dans lesquelles de nombreux gènes ont été identifiés comme important dans le développement embryonnaire précoce (Berkovits et Wolgemuth, 2012 ; Gaucher *et al.*, 2012).

Parallèlement à cette condensation chromatinienne se produit un arrêt des processus de transcription cellulaire (Kierszenbaum et Tres 1978). Le noyau du spermatozoïde est donc un noyau au repos, transcriptionnellement inactif (Ward 1994). Cette mise au repos du génome est réversible suite à la fécondation de l'ovocyte. On sait cependant que le noyau renferme un certain nombre de petits ARNm et d'ARN non codants comme les miARN et les piARN (Piwi interactings RNA) (Sendler *et al.*, 2013). Plusieurs études ont confirmé que le spermatozoïde n'avait pas d'activité transcriptionnelle ni traductionnelle suggérant que ces ARNs ont été synthétisés en amont, durant la spermatogénèse (Grunewlad *et al.*, 2005 ; Cappallo-Obermann *et al.*, 2011). Les spermatozoïdes permettraient d'apporter un pool d'ARNs libérés lors de la fécondation et probablement importants pour le développement précoce du zygote (Barroso *et al.*, 2009). Des données récentes ont démontré par des analyses protéomiques la présence de protéines ribosomales suggérant toutefois une possible activité traductionnelle résiduelle (De Mateo *et al.*, 2011).

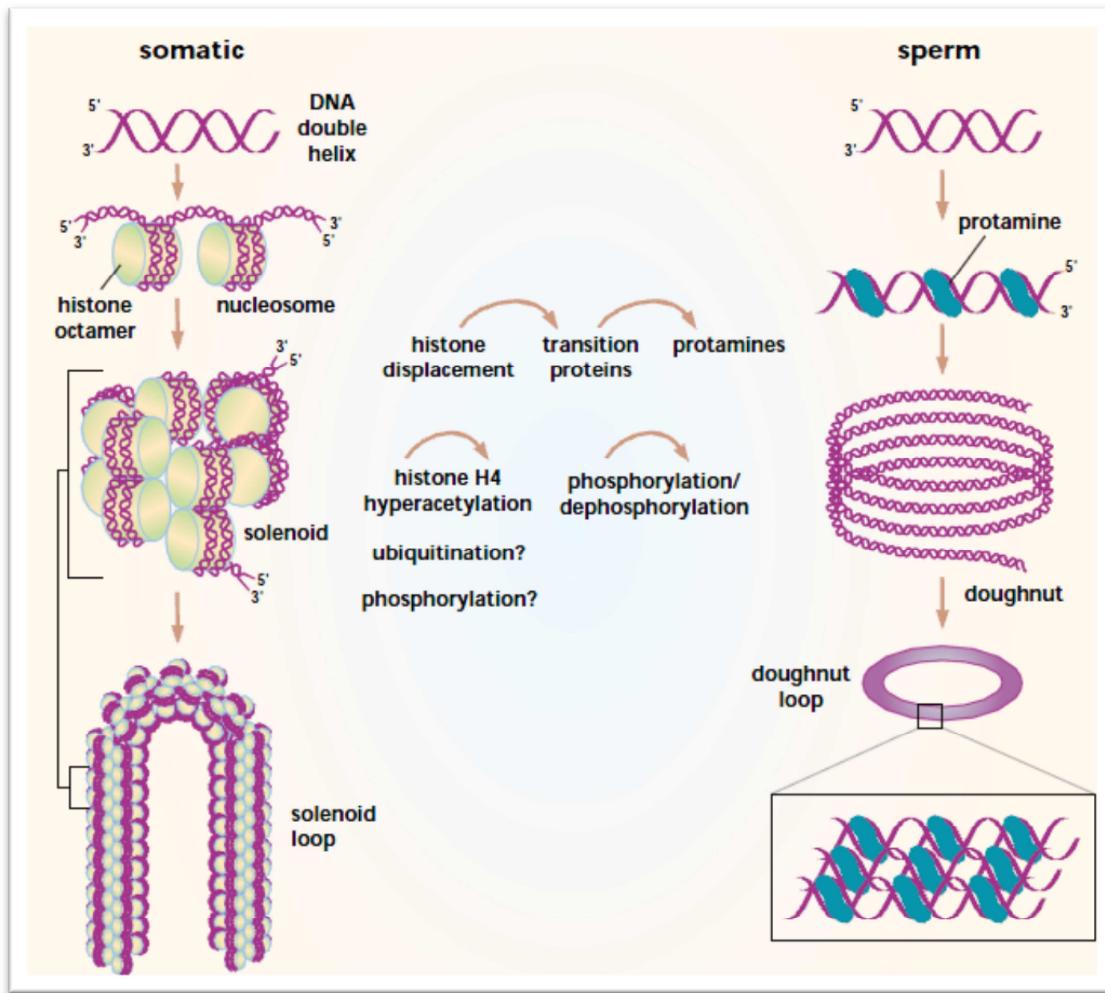


Figure 14 : Schéma de la compaction de l'ADN dans les cellules somatiques et dans les spermatozoïdes. Dans les cellules somatiques, l'ADN est enroulé sous forme de nucléosome. Les nucléosomes vont s'agencer entre eux pour former un solénoïde qui sera attaché à la matrice nucléaire par sa base. Dans le noyau spermatique les nucléosomes sont remplacés par des protamines qui vont compacter l'ADN sous forme de « doughnut ». Le remplacement des histones est facilité par des acétylations, des ubiquitinisations et des phosphorylations (Braun 2001).

Il est important de noter que la compaction de l'ADN dépend à la fois de mécanismes spécifiques (BRDT, Acétylase, histones spécifiques,...) mais aussi d'un déroulement harmonieux de la spermiogénèse. En effet, lorsque les acteurs impliqués dans la biogénèse de l'acrosome sont défectueux, un phénotype de mauvaise compaction de l'ADN spermatique est systématiquement associé.

3.3. La thèque périnucléaire

La thèque périnucléaire est une structure qui se trouve à l'interface noyau acosome. Elle se forme au stade tardif de la spermiogénèse. Elle est majoritairement constituée de cytosquelette (Oko et Sutovsky, 2009). Sa structure se compose de deux parties, une partie sub-acrosomique située sous l'acosome (appelée aussi *perforatum*) et une partie post-acrosomique située en périphérie de l'acosome (appelée *calix*). La thèque périnucléaire pourrait jouer un rôle dans l'ancrage de l'acosome, dans le remodelage du noyau et enfin dans l'activation ovocytaire (Oko, 1995).

Elle est constituée de très nombreuses protéines aux fonctions comme des protéines interactions avec l'actine (ex : calcine, cyclin II), des protéines impliquées dans la formation de l'acosome, certains histones ou variants d'histones ou encore la phospholipase zeta (PLC ζ) nécessaire à l'activation ovocytaire (Oko et Sutovsky, 2009 ; Escoffier *et al.*, 2015).

3.4. Le flagelle

Le flagelle est responsable de la mobilité du spermatozoïde. Il présente trois parties, la pièce intermédiaire (PI), la pièce principale (PP) et la pièce terminale (PT) (figure 15). La PI se trouve juste derrière la tête, c'est une partie plus épaisse que la PP. La PP se trouve dans le prolongement de la pièce intermédiaire et représente 5/6 de la longueur totale du spermatozoïde (figure 15).

La pièce intermédiaire est composée de trois types de structures : la gaine de mitochondries qui, notons le, ne va produire qu'une partie négligeable de l'énergie nécessaire au battement flagellaire (grâce à la phosphorylation oxydative qui produit de l'ATP), l'axonème qui se prolonge dans la pièce principale et un ensemble de neuf faisceaux de fibres denses.

Au niveau de la PP, le manchon de mitochondries a disparu. Une structure supplémentaire est présente, la gaine fibreuse, formée par l'association d'anneaux semi-circulaires et qui entoure l'ensemble axonème-fibres denses. Elle présente aussi deux épaississements adhérents aux fibres denses 3 et 8 constituant ainsi des colonnes longitudinales. Ainsi, la PP ne compte plus que 7 fibres denses au lieu de 9 pour la pièce intermédiaire. C'est au niveau de la gaine fibreuse qu'est produit la majorité de l'énergie

nécessaire au glissement des microtubules par l'intermédiaire d'enzymes appartenant à la voie de la glycolyse (Eddy, 2007).

La pièce terminale est située au niveau de l'extrémité distale du flagelle et ne contient que l'axonème (Inaba, 2003).

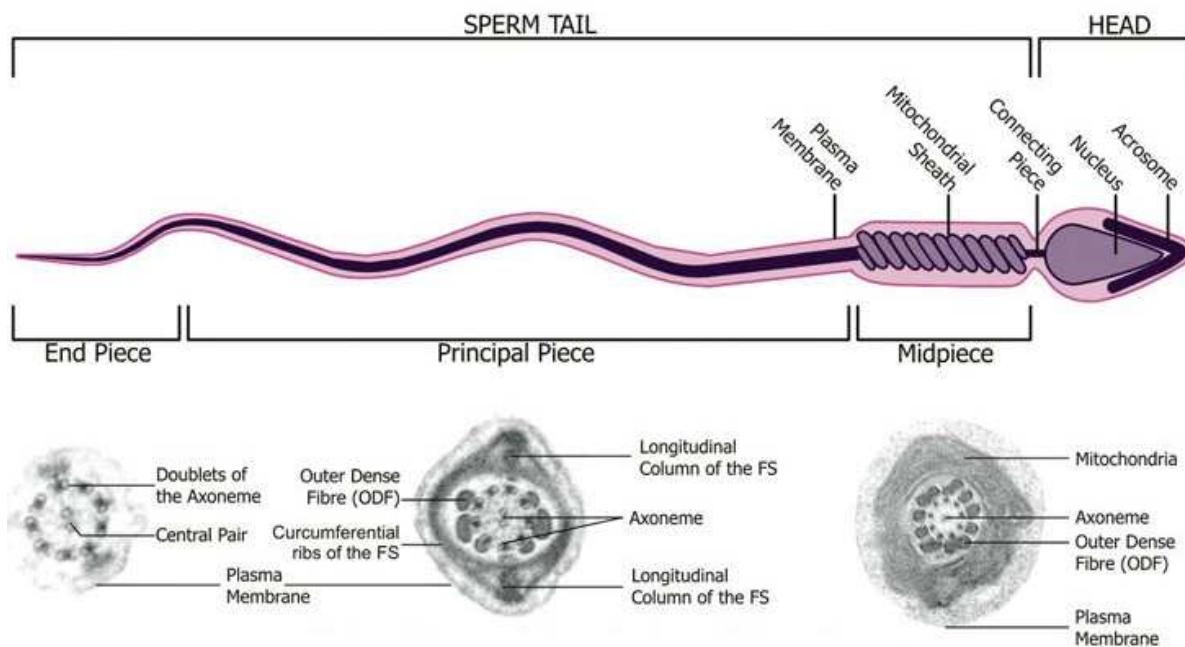


Figure 15. Structure du flagelle d'un spermatozoïde. Coupes transversales en microscopie électronique. Le flagelle se compose de trois parties: la pièce intermédiaire, contenant les mitochondries, la pièce principale et la pièce terminale. L'axonème, en position centrale, parcourt tout le flagelle. Des structures périaxonémiales sont observables : les fibres denses dans la pièce intermédiaire et principale, et la gaine fibreuse dans la pièce principale seulement. D'après Borg *et al.*, 2010.

3.5. L'axonème

L'axonème est une structure hautement spécialisée qui se développe dès les phases précoces de la spermiogénèse à partir de l'un des deux centrioles formant le corps basal du flagelle. Le corps basal est le centre organisateur des microtubules permettant l'initiation de la polymérisation des microtubules de l'axonème. Ces centrioles sont associés à de très nombreuses protéines péri-centriolaires indispensables à la formation du flagelle (Schatten et Sun, 2009). L'axonème est la seule structure présente sur l'intégralité du flagelle du spermatozoïde. Parmi les structures ciliées, il existe deux formes d'axonème, de conformation dite « 9+0 » ou « 9+2 », et qui se différencient par la présence d'un doublet

central de microtubules (Fliegauf *et al.*, 2008). Au niveau structural, l'axonème du spermatozoïde est constitué de 9 doublets de microtubules (A et B) organisés en cercle et de deux microtubules centraux et est donc de conformation dite « 9+2 ». Chaque doublet périphérique comprend des protéines motrices, les dynéines, qui vont conférer au flagelle son mouvement oscillatoire. On distingue les bras externes de dynéine (ODA pour Outer Dynein Arm) situés du côté de la membrane plasmique et les bras internes de dynéines (IDA pour Inner Dynein Arm) situés du côté du doublet central (figure 16). Les doublets périphériques sont reliés entre eux par les liens de nexine et au doublet central par les ponts radiaux (figure 16) (Inaba, 2003; Satir and Christensen, 2008).

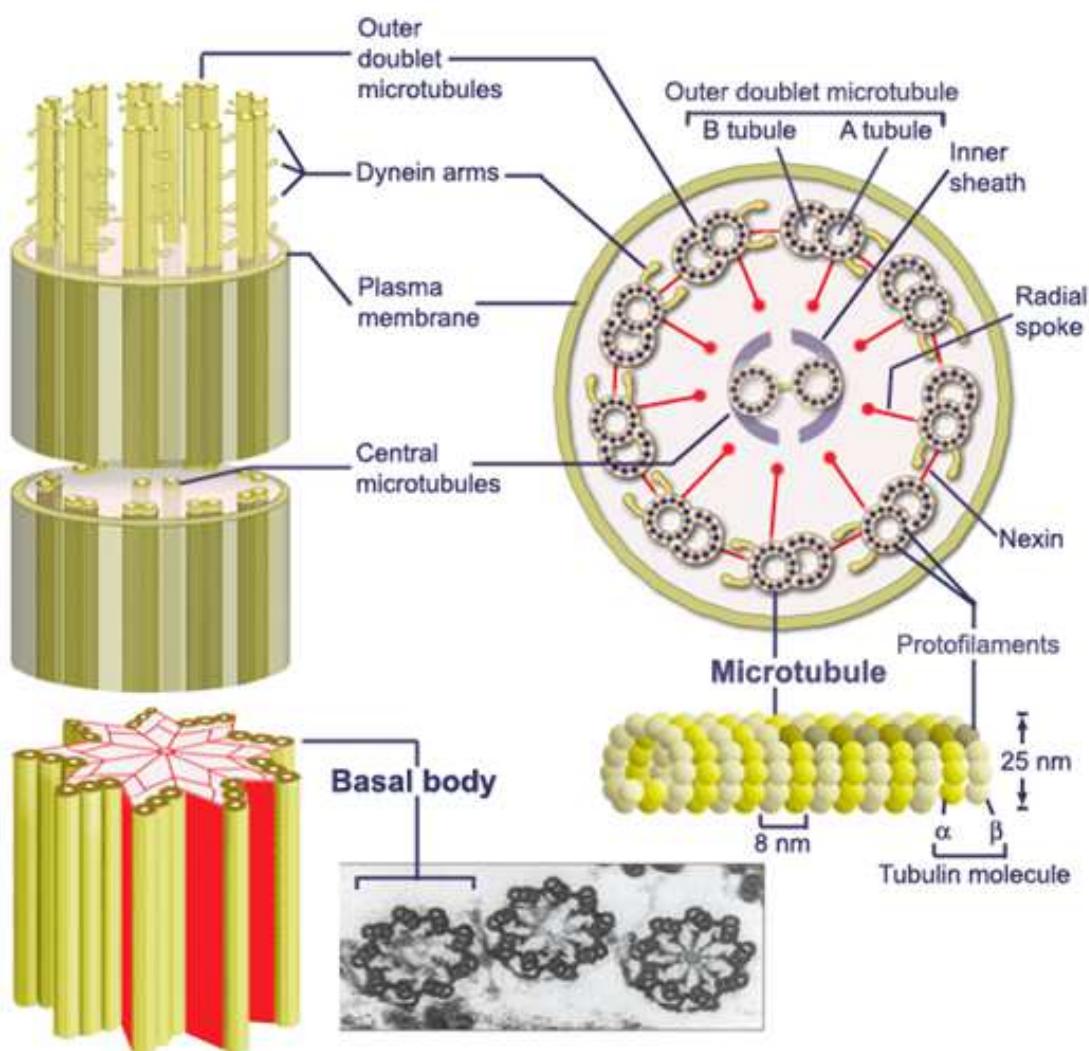


Figure 16. Schéma de l'axonème du flagelle de spermatozoïde. 9 doublets de microtubules reliés entre eux par de la nexine entourent une paire centrale de microtubules. Les microtubules sont constitués de tubuline α et β , disposés alternativement et font 25 nm de diamètre. Le doublet de microtubule est constitué d'un tubule A et B. Le tubule A est directement relié à la paire centrale par les ponts radiaux et supporte les bras de dynéine. D'après Inaba, 2007.

Les bras de dynéines sont des complexes multiprotéiques ancrés à intervalles réguliers sur l'axonème. L'unité fondamentale de l'axonème est répétée tous les 96 nm et contient 4 bras externes de dynéines, sept isoformes de bras internes de dynéines, les ponts radiaux (2 à 3 selon les espèces), le lien de nexine – complexe de régulation des dynéines, et le complexe associé à la calmoduline et au pont radiaire (figure 17).

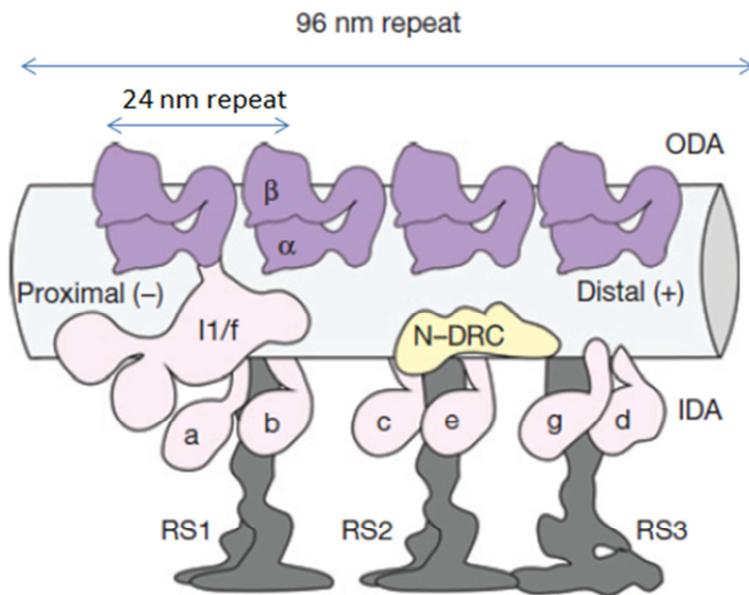


Figure 17. Organisation de l'unité fondamentale de 96 nm de l'axonème du flagelle du sperme d'oursin (*Strongylocentrotus purpuratus*). Les bras de dynéine externes (en violet, ODA) composés de 2 chaînes lourdes (α et β) sont ancrés sur le microtubulue A tous les 24 nm. Les 7 isoformes des bras internes de dynéine (IDA, rose) sont représentées ainsi que le complexe de régulation des dynéines (N-DRC, jaune) et 3 ponts radiaux (RS1-RS3, gris). Le complexe associé à la calmoduline et au pont radiaire n'est pas visible. Adapté de Lin *et al.*, 2014.

3.5.1. Les dynéines axonémiales

Les connaissances de la composition structurale de l'axonème et plus particulièrement de l'organisation des bras de dynéines le long de l'axonème sont issues principalement des études faites sur certains protozoaires comme *Chlamydomonas reinhardtii*, *Tetrahymena sp.*, *Trypanosoma sp.* ou encore sur le sperme d'oursin (*Echinus sp.* ou *Strongylocentrotus purpuratus*) (Vincensini *et al.*, 2011), grâce à une remarquable conservation de cette structure au cours de l'évolution. De très nombreuses protéines du

flagelle chez ces espèces présentent des orthologues dans le flagelle du spermatozoïde humain avec des très fortes homologies (supérieure à 75%) (Inaba, 2007, 2011).

Les dynéines font partie des protéines motrices de l'axonème qui vont conférer au flagelle son mouvement oscillatoire. Bien que les deux bras fonctionnent ensemble, les bras internes conféreraient l'amplitude au mouvement alors que les bras externes participeraient à la fréquence et à la vitesse du battement (Satir and Christensen, 2008).

3.5.1.1. Bras de dynéines

Les bras de dynéines axonémiales sont composés de plus d'une vingtaine de sous-unités protéiques distinguées selon leur masse moléculaire parmi lesquelles on retrouve les chaînes lourdes (HC pour heavy chain, ~500 kDa), les chaînes intermédiaires (IC, intermediate chains, 120-156 kDa) et enfin les chaînes légères (LC, light chains, 30-80 kDa) (Bisgrove and Yost, 2006).

Les chaînes lourdes des dynéines sont bien conservées dans l'évolution et sont responsables de l'activité motrice des dynéines (Mazumdar *et al.*, 1996). Les chaînes lourdes contiennent plusieurs domaines distincts (figure 18) :

- La partie C-terminale de chaque chaîne lourde est constituée d'une tête globulaire contenant les domaines moteurs ATPasiques, de deux hélices formées par des domaines coiled-coil (CC1, CC2) permettant la flexion et la mobilité de la tête et enfin d'une petite unité globulaire de liaison aux microtubules B (MTBD). On dénombre 6 domaines moteurs ATPasiques nommés AAA1-6 (ATPase Associated with cellular Activities) et contenant chacun un motif P-loop capable de se lier à l'ATP. Cependant, il semblerait que seul le premier domaine (AAA1) puisse être capable d'hydrolyser l'ATP (Kon *et al.*, 2004).

- La partie N-terminale forme la queue de la molécule et constituerait le principal site d'interaction avec les autres sous-unités de la dynéine (Habura *et al.*, 1999 ; Tynan *et al.*, 2000).

Les chaînes intermédiaires sont situées à la base des bras de dynéines et possèdent un domaine d'interaction avec la partie N-terminale des chaînes lourdes. Elles participent à l'assemblage et à l'ancre des bras sur le microtubule A.

Les chaînes légères interagissent directement avec les chaînes intermédiaires (Makokha *et al.*, 2002). Le rôle principal serait de réguler l'activité des dynéines, notamment via des interactions avec des protéines régulatrices dépendantes du calcium comme la calmoduline.

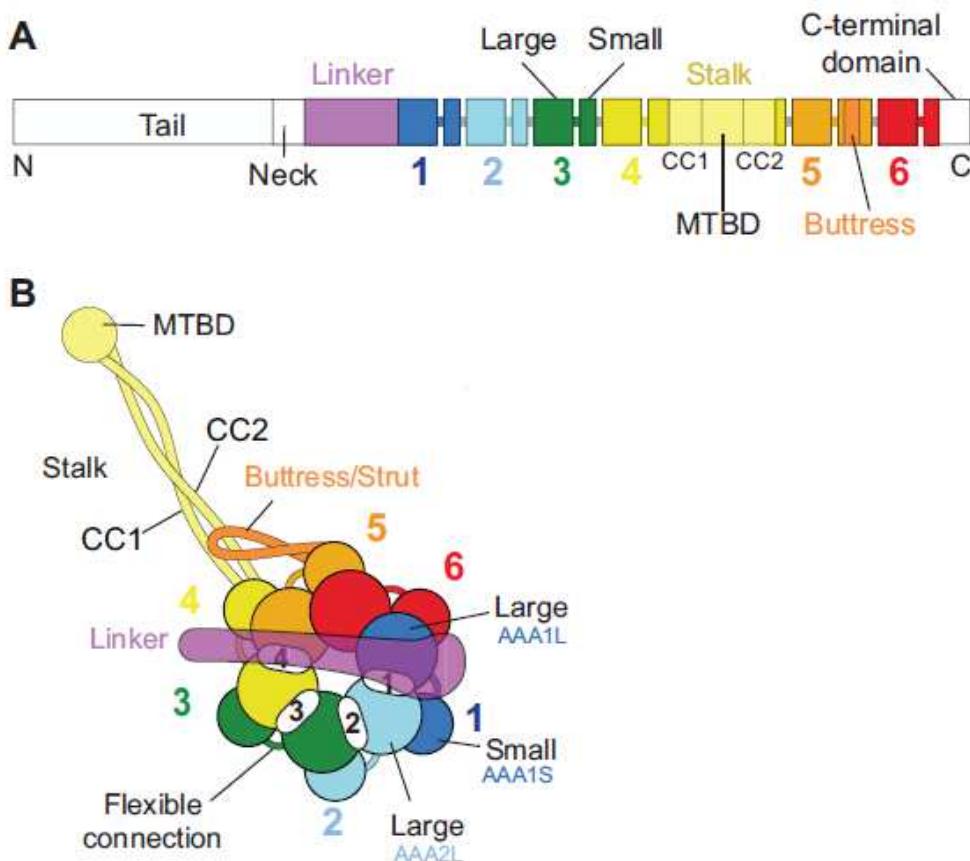


Figure 18. Schéma de l’organisation des domaines d’une chaîne lourde de dynéine (A) et sa structure tertiaire (B). CC : coiled-coil domain ; MTBD : microtubule binding domain. D’après Carter, 2013.

- **Bras externes de dynéines**

Il existe un seul type de bras externes composés de trois chaînes lourdes chez *Chlamydomonas* (2 chez l’homme), au moins deux chaînes intermédiaires et au moins onze chaînes légères (figure 19) (King, 2012). Les bras externes sont attachés aux microtubules A à des sites précis et selon une périodicité de 24 nm. Le complexe d’ancrage (ODA-DC, ODA docking complex) composé de trois sous-unités protéiques (DC1, DC2 et DC3) coïncide avec le site d’attache des bras de dynéines externes (Takada *et al.*, 2002). Il existe également des liens entre les bras de dynéines externes (appelées OOD, outer-outer dynein). Leur composition et leur fonction ne sont pas complètement élucidées mais ils pourraient participer à l’assemblage ou à la coordination de l’activité des bras externes (Nicastro *et al.*, 2006).

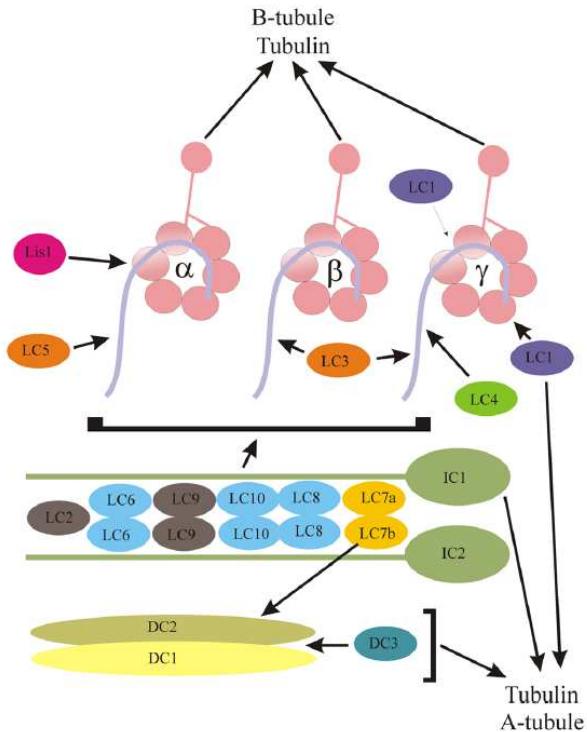


Figure 19. Représentation des bras de dynéine externe chez *Chlamydomonas*. En rose, sont représentées les trois chaînes lourdes (α , β et γ). LC (1-10) : chaînes légères ; IC (1-2) : chaînes intermédiaires ; DC (1-3) : outer arm docking complex. D'après King 2012.

- **Bras internes de dynéines**

L'organisation des bras internes de dynéine est plus complexe que celle des bras externes. En effet, il existe 7 isoformes de bras internes de dynéines différentes réparties régulièrement sur le microtubule A en répétition de 96 nm. Chaque isoforme est composée de chaînes lourdes différentes, de chaînes intermédiaires et légères. Chez *Chlamydomonas*, on dénombre 8 chaînes lourdes différentes au total pour les bras internes, la première isoforme (IDA1) en possédant deux (Bui *et al.*, 2008). Le bras interne I1 est composé de deux DHC (isoforme à deux têtes), trois chaînes intermédiaires et au moins 5 chaînes légères DLC (figure 20). Les six autres isoformes correspondent aux bras I2 et I3 (isoformes à tête unique) et sont composées d'une seule chaîne lourde associée à une chaîne intermédiaire et à une chaîne légère qui peut être soit la centrine soit p28 (Bui *et al.*, 2008).

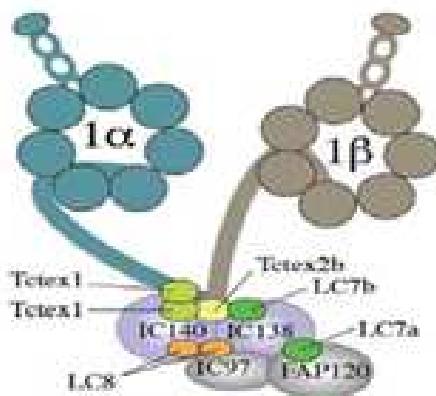


Figure 20. Représentation schématique des bras internes chez *Chlamydomonas reinhardtii*. Il existe au moins 7 sous-espèces de BDI. Les bras de dynéines internes sont composés par deux chaines lourdes (HC α et HC β), trois chaines intermédiaires (IC140, IC138 et IC97) et par cinq chaines légères (Tctex1, Tctex2b, LC7b, LC7a et LC8). D'après King *et al.*, 2000.

3.5.2. Liens de nexine et complexes de régulation des dynéines

Les liens de nexine relient entre eux les microtubules A avec les microtubules B de deux doublets de microtubules périphériques. Ces liens contribuent à la résistance élastique permettant de convertir les glissements des doublets de microtubules en flexion axonemale. Ces liens de nexine font partie du complexe de régulation des dynéines (N-DRC, nexin-dynein regulatory complex) composé d'une dizaine de protéine et dont la base est ancrée sur les microtubules au niveau du site d'ancrage du pont radiaire 2 (RS2) avec lequel il interagit (Bower et al., 2013).

3.5.3. Ponts radiaux

Les ponts radiaux sont des complexes multiprotéiques structurés en « T » reliant les doublets périphériques au doublet central. Les ponts radiaux agissent comme des transducteurs mécaniques entre les doublets périphériques et le doublet central. Ils contribuent à la régulation de l'activité motrice des dynéines et ainsi à la mobilité du flagelle. Chez *Chlamydomonas* on dénombre 2 types de ponts radiaux (RS1 et 2 – RS : radial spoke) alors qu'on en retrouve 3 (RS1-3) chez d'autres espèces de protistes comme *Tetrahymena sp.* ou chez les Mammifères. Le RS3 est présent chez *Chlamydomonas* sous forme « atrophiée » (RS3 stump) et ne relie pas la paire centrale (Pigino *et al.*, 2011). Les ponts radiaux se composent de 4 domaines : une base ancrée sur les microtubules, la tige, le cou et la tête orthogonale (figure 21). Chaque domaine est composé de protéines différentes pouvant

interagir directement entre elles. Chez *Chlamydomonas* au moins 23 protéines ont déjà été identifiées (Yang *et al.*, 2006). Les ponts radiaux semblent directement connectés entre eux par leur tête suggérant des interactions fortes et étroites en RS adjacents (Pigino *et al.*, 2011).

Les ponts radiaux sont aussi connectés aux bras de différentes dynéines internes par leur base (Pigino et Ishikawa, 2012) (figure 21). A ces ponts radiaux sont également associés le complexe associé à la calmoduline (CSC : calmodulin and spoke-associated complex) qui s'étend de la base de RS2 à la base du RS3. Ce complexe formé de trois protéines (FAP61, 91 et 251) chez *Chlamydomonas* semble essentiel à la motilité du flagelle par régulation calcique, à l'assemblage des ponts radiaux et à leur ancrage aux microtubules (Dymek *et al.*, 2011).

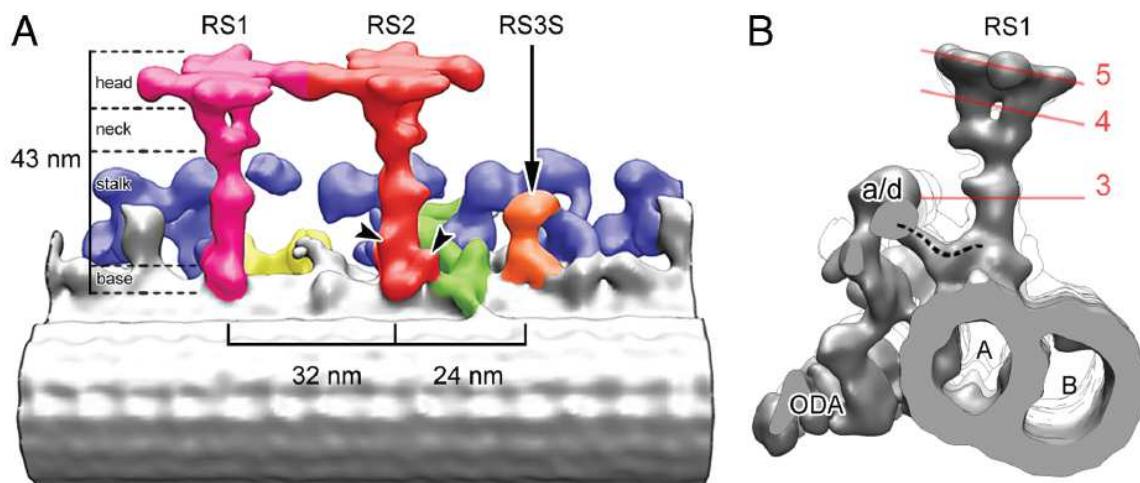


Figure 21. Structure et représentation 3D des ponts radiaux chez *Chlamydomonas reinhardtii*. A. Vue longitudinale présentant les RS1 (rose), RS2 (rouge) et RS3 stump (orange), les bras de dynéine interne (bleu), les chaînes légères et intermédiaires (jaunes) et le complexe N-DRC (vert). B. Vue en coupe transversale du pont radial RS1. On observe la connexion entre RS1 et la queue terminale du bras de dynéine interne IDA a/d (pointillés noirs). D'après Pigino et Ishikawa, 2012.

3.5.4. Le complexe central

La paire centrale de microtubules est composée de microtubules isolés appelés C1 et C2 entourés d'une gaine centrale associée à des projections asymétriques orientées vers les doublets périphériques. Les microtubules C1 et C2 sont reliés entre eux par des « ponts ». Les études chez *Chlamydomonas* ont montré que le microtubule C1 supporte six projections (C1a-C1f) et 5 pour le microtubule C (C2a-C2e) (figure 22). Ces projections sont reliées entre elles transversalement et longitudinalement et interagissent avec la tête des ponts radiaux (Wargo

et Smith, 2003). La paire centrale comprend au moins 23 protéines très conservées lors de l'évolution et distribuées différemment entre les deux microtubules. Ces données démontrent que les microtubules C1 et C2 sont structurellement distinctes et possèdent probablement une fonction spécifique. La paire centrale joue un rôle essentiel dans la motilité flagellaire (Carbajal-Gonzalez *et al.*, 2013), mais aussi dans la cohésion des différentes microtubules périphériques lors de leur mouvement.

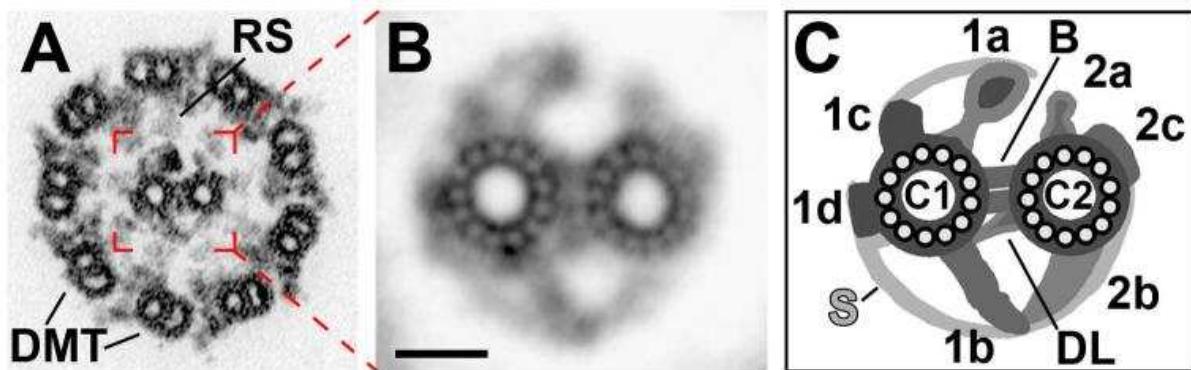


Figure 22. La structure de la paire centrale de microtubule chez *Chlamydomonas reinhardtii*. A-B : images en microscopie électronique. C : représentation schématique des projections de la paire centrale. D'après Carbajal-Gonzalez *et al.*, 2013.

3.5.5. Formation de l'axonème

Des analyses de protéomiques récentes ont rapporté la présence de plus de 700 protéines spécifiquement localisées dans le flagelle du spermatozoïde humain dont une grande partie est constituée de protéines axonémiales (Baker *et al.*, 2013). L'absence de machinerie de synthèse protéique au niveau des cils et des flagelles signifie que la formation de l'axonème nécessite le transport et l'assemblage de plusieurs centaines de protéines synthétisées dans le cytoplasme.

Ce transport des molécules dans le flagelle est appelé transport intraflagellaire (IFT : intraflagellar transport) et constitue un système de transport interne bidirectionnel. La biogenèse du cil requiert des complexes protéiques spécifiques propulsés par des moteurs moléculaires : le kinésines pour le transport antérograde (vers le sommet) et les dynéines cytoplasmiques pour le transport rétrograde (vers la base). Le complexe IFT est divisé en deux sous complexes : le complexe IFT-A associé au transport rétrograde et le complexe IFT-B au transport antérograde. Les molécules transportées appelées « cargos » sont prises charge par le complexe IFT-A de la base jusqu'au sommet du flagelle. Les molécules seront finalement libérées et un mécanisme de régulation permettra le recyclage des protéines

motrices et la mise en place du transport rétrograde par le complexe IFT-B (figure 23) (Ishikawa et Marshall, 2011).

Ce transport intraflagellaire permet d'acheminer différents composants essentiels à la formation et la fonction l'axonème. C'est le cas notamment des bras externes et internes de dynéine et des ponts radiaux pré-assemblés en amont dans le cytoplasme (Ishikawa et Marshall, 2011).

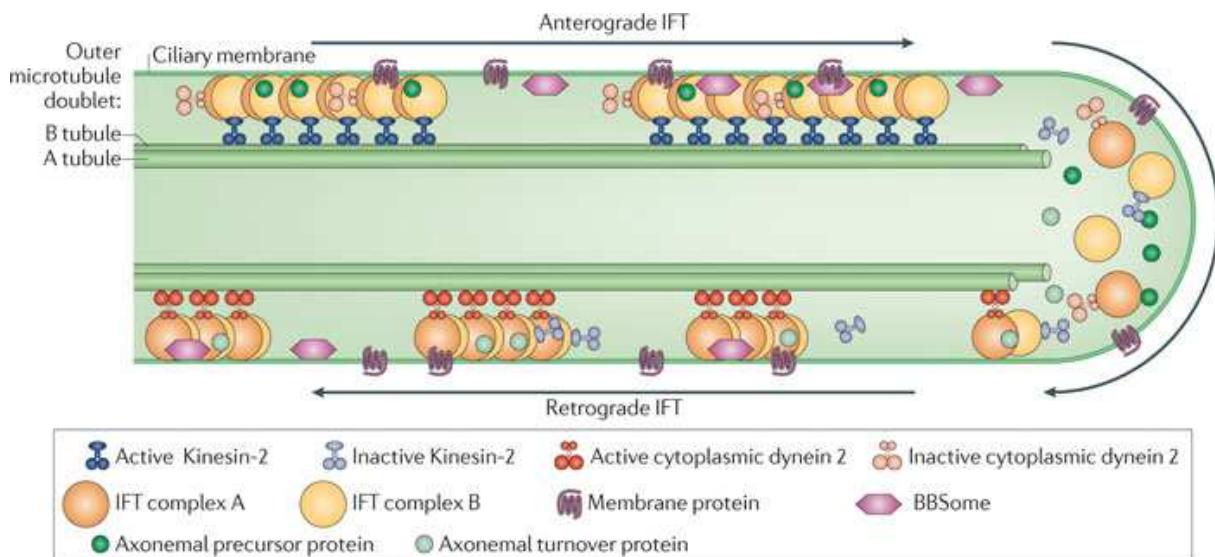


Figure 23. Représentation schématique du transport intraflagellaire (IFT). D'après Ishikawa et Marshall, 2011.

3.6. Le manchon de mitochondrie

Les mitochondries sont disposées de manière hélicoïdale autour de l'axonème dans la pièce intermédiaire (figure 24). La formation de la gaine de mitochondrie commence à la fin de la spermiogenèse. Les mitochondries vont permettre la production d'ATP indispensable à la mobilité spermatique. La gaine de mitochondrie va produire cette énergie grâce au processus de phosphorylation oxydative ou via le cycle de l'acide citrique (Piomboni *et al.*, 2012). Les mitochondries constituerait également une réserve calcique, le calcium étant un élément central de la capacitation, de l'hyperactivation du mouvement du flagelle, et de la réaction acrosomique (Costello *et al.*, 2009). Enfin, les mitochondries permettraient de réguler la production des espèces réactives de l'oxygène pouvant impacter la fluidité membranaire, la mobilité spermatique et générer des dommages de l'ADN (De Lamirande et O'Flaherty 2008).

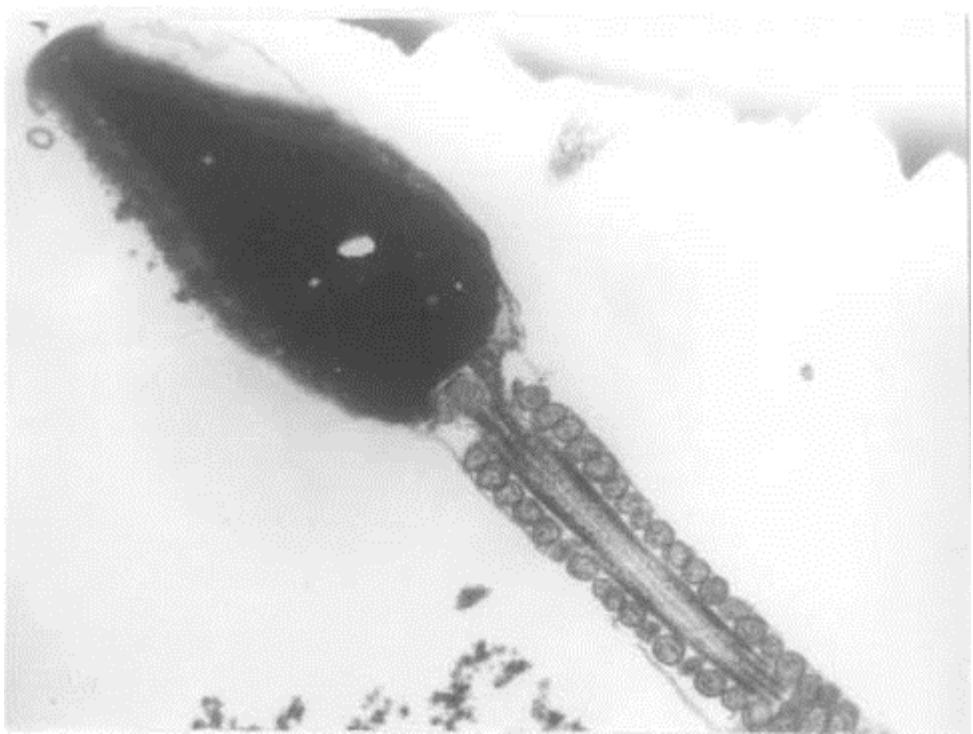


Figure 24. Observation en microscopie électronique du manchon mitochondrial. D'après Lohiya *et al.*, 2000.

3.7. Les fibres denses

Les fibres denses recouvrent l'axonème sur toute sa longueur à l'exception de la partie terminale du flagelle. On dénombre 9 fibres denses dans la pièce intermédiaire (figure 25) et 7 dans la pièce principale. Les fibres denses externes (ODF) sont des éléments du cytosquelette importants qui contribuent aux caractéristiques morphologiques distinctifs du flagelle du spermatozoïde. Des études ont indiqué que les fibres denses participeraient à l'amélioration des contraintes de flexion du flagelle (Lindemann, 1996) et /ou à la protection des contraintes mécaniques lors du transit épидydymaire du spermatozoïde (Baltz *et al.*, 1990). La protéine ODF2 a été identifiée comme le composant majeur des fibres denses dans le flagelle de mammifères (Hoyer-Fender *et al.*, 1998). Il a été démontré que les fibres denses étaient hautement phosphorylées mais la signification de ces modifications post-traductionnelles restent encore mal connue. La protéine CDK5 a été identifiée comme l'une des kinases capable de phosphoryler la protéine ODF1 et favoriserait sa dégradation et potentiellement le détachement du flagelle après la fécondation (Rosales *et al.*, 2008).

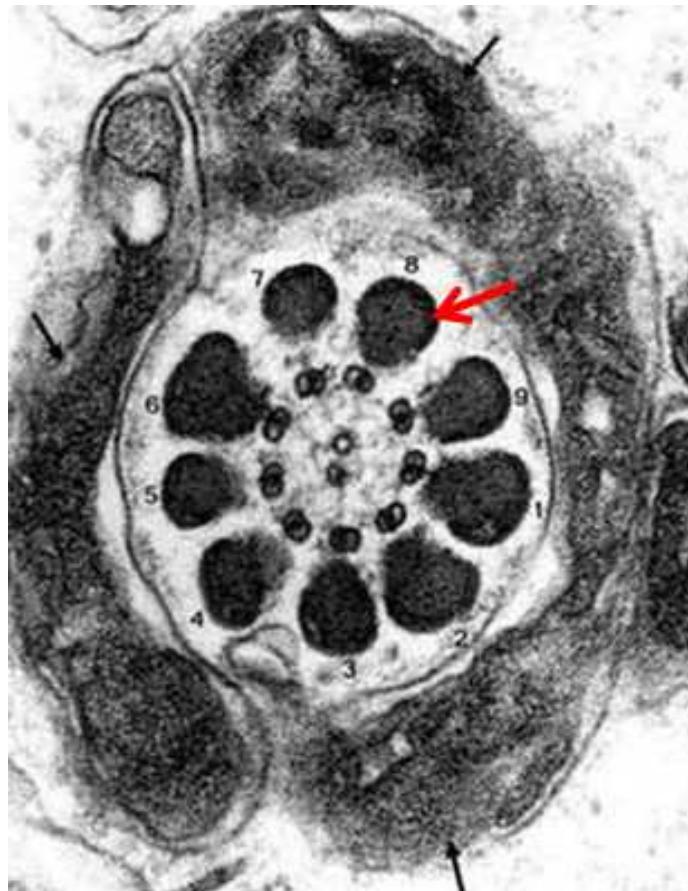


Figure 25. Coupe transversal du flagelle au niveau de la pièce intermédiaire où l'on visualise les 9 fibres denses (flèches rouges). D'après "Visual atlas of human sperm structure and function for assisted reproductive technology" Ed A.H. Sathanathan 1996.

3.8. La gaine fibreuse

La gaine fibreuse est une structure qui enveloppe l'axonème et les fibres denses uniquement au niveau de la pièce principale du flagelle. La gaine fibreuse est constituée d'anneaux circonférentiels qui s'étendent à partir des colonnes longitudinales équidistantes (figure 26). Le développement de la gaine fibreuse démarre au stade des spermatides rondes. La plupart des gènes codant pour la gaine fibreuse sont exprimés durant la phase de différenciation du spermatozoïde et seulement par les cellules spermatogéniques. La formation de la gaine fibreuse présente la particularité de commencer après la formation de l'axonème et progresse de la partie distale vers la partie proximale du flagelle (Eddy, 2007).

Il est admis que la gaine fibreuse a un rôle purement mécanique qui procure une rigidité élastique au flagelle et module l'amplitude de ses mouvements. Plus récemment, il a été établi que la gaine fibreuse constitue le support de très nombreuses protéines et notamment des enzymes indispensables à la mobilité spermatique. Aujourd'hui plus d'une

trentaine de protéines sont associées à la gaine fibreuse, notamment plusieurs enzymes glycolytiques (aldolase A, lactate déshydrogénase, glyceraldehyde 3-phosphate déshydrogénase ...) responsables de la production d'ATP ou des protéines de la voie de signalisation de la protéine kinase A (Eddy, 2007). Parmi ces dernières on va trouver les protéines AKAP3 et AKAP4 (Kinase Anchoring Protein) qui constituent les protéines les plus abondantes de la gaine fibreuse. La protéine AKAP3 participe à l'organisation de la structure de base de la gaine fibreuse alors qu'AKAP4 semble plutôt être impliquée dans le processus d'assemblage de la gaine (Brown *et al.*, 2003). De manière intéressante, il a été démontré que les souris inactivé pour le gène AKAP4 présentent des anomalies morphologiques et ultrastructurales sévères du flagelle du spermatozoïde (raccourci de moitié, pièce principale plus fine, manchon mitochondrial anormal) confirmant le rôle structural essentiel de ces protéines (Miki *et al.*, 2002).

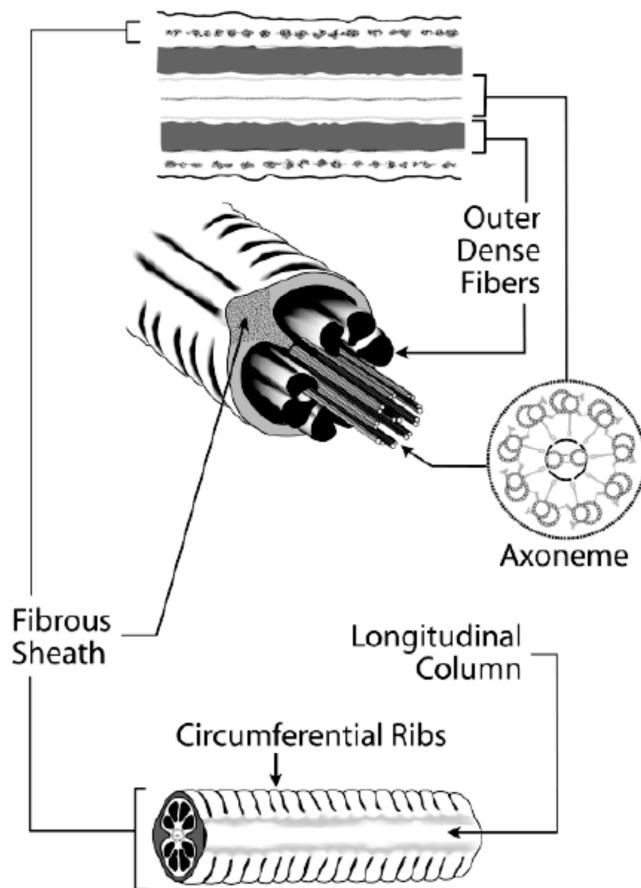


Figure 26. Représentation schématique de la gaine fibreuse. La gaine fibreuse est formée par l'association d'anneaux semi-circulaires et deux colonnes longitudinales épaissements adhérents aux fibres denses 3 et 8. D'après Eddy, 2007.

B. Infertilité masculine

L'infertilité, définie comme l'incapacité pour un couple à concevoir après 1 an de rapports sexuels non protégés, est un problème de santé publique majeur et représente un des enjeux médicaux et scientifiques importants de ces dernières années. On estime qu'au moins 9 % des couples seraient concernés par des problèmes de fertilité soit plus de 70 millions de personnes dans le monde (Boivin, 2007).

Chez l'homme, l'infertilité est en relation avec une altération quantitative et/ou qualitative des spermatozoïdes, d'origine congénitale ou acquise. Parmi ces altérations, on distingue une réduction (oligozoospermie) ou une absence totale de spermatozoïdes (azoospermie) (18% des cas), des spermatozoïdes immobiles (asthénozoospermie, 28%), une ou plusieurs anomalies morphologiques (térapozoospermie, 16%) ou un phénotype multiple (oligo-térapoasthénozoospermie, OAT, 34%) (Thonneau *et al.*, 1991). Ces altérations peuvent être détectées et quantifiées dans les laboratoires spécialisés par la réalisation d'un spermogramme et /ou d'un spermocytogramme.

1. Moyen d'exploration de l'infertilité masculine

1.1. Le spermogramme

Le spermogramme est l'étude du sperme frais éjaculé après un délai d'abstinence sexuelle de 3 jours minimum. Le recueil se fait au laboratoire dans un flacon stérile. Immédiatement après éjaculation, le sperme est déposé dans une étuve à 37° C pour assurer sa liquéfaction (environ 30 min). Au terme de celle-ci, l'examen est réalisé. Un aspect anormal, tel qu'une hémospermie, ou une forte viscosité doit être notée. Les paramètres étudiés sont :

- **pH** : Le pH du sperme est normalement compris entre 7,4 et 8,0. Des valeurs trop faibles peuvent être le reflet d'un défaut de sécrétion des vésicules séminales (normalement alcalines) ou des sécrétions prostatiques (légèrement acides).

- **Volume** : Il traduit essentiellement les capacités sécrétoires des glandes annexes. Une hyperspermie (volume > 6 ml) ne doit pas être considérée comme pathologique. Par contre, l'hypospermie (volume < 2ml) peut s'expliquer soit par un trouble de l'éjaculation, soit par une insuffisance sécrétoire de l'une des glandes annexes pouvant être liée à une infection (prostatite, vésiculite).

- **Nombre de spermatozoïdes** : Il est exprimé en concentration (millions/ml). Si aucun spermatozoïde n'est observé par la technique classique, il est nécessaire de rechercher les spermatozoïdes dans le culot de centrifugation du sperme, avant de conclure ou non à une azoospermie.

- **Mobilité** : La mobilité appréciée au microscope optique est exprimée en pourcentage de spermatozoïdes mobiles. L'examen est réalisé dans l'heure qui suit la liquéfaction avec un suivi de 4 heures. Une évaluation qualitative est réalisée en différenciant les spermatozoïdes se déplaçant activement suivant une trajectoire linéaire ou aléatoire (mobilité progressive, PR), ne progressant que très faiblement (non progressif, NP) et les spermatozoïdes totalement immobiles. Des systèmes d'analyse vidéomicrographique assistée par ordinateur (système CASA) permettent une mesure automatique objective.

- **Vitalité** : La vitalité renseigne sur le nombre de spermatozoïdes vivants et morts par une coloration vitale (éosine-nigrosine).

- **Cellules rondes** : Les cellules épithéliales de l'urètre, les cellules germinales immatures et les leucocytes sont regroupés sous ce terme de «cellules rondes». Dans les cas où ce nombre est élevé, les polynucléaires, témoins d'un foyer infectieux doivent être précisément recherchés

L'atteinte du profil spermatique peut être secondaire à l'altération d'un seul ou de plusieurs paramètres du spermogramme. Pour chaque paramètre est établi une valeur seuil en dehors de laquelle on peut définir une ou plusieurs anomalies (tableau 2).

Tableau 2. Limites basses de référence d'après l'OMS (WHO, 2010). PR : mobilité progressive ; NP : mobilité non progressive.

Paramètres spermatiques	Limites basses de références (WHO, 2010)	Définitions de l'anomalie
Volume	1.5 ml	<1.5 ml : hypospermie >6 ml : hyperspermie
Numération des spermatozoïdes	15 millions/ml 39 millions / ejaculat	Néant : azoospermie < oligozoospermie
Numération des cellules rondes	1 million / ml	
Numération des polynucléaires	1 million / ml	> leucospermie
Mobilité	PR \geq 32% ; (PR+NP) \geq 40%	< Asthénozoospermie
Vitalité	58%	< nécrozoospermie
Morphologie	>4% de forme normale	< tératozoospermie

Les résultats du spermogramme peuvent être influencés par plusieurs facteurs tels que le délai d'abstinence, la saison, l'âge, l'exposition professionnelle ou épisodique à des facteurs toxiques etc... L'examen du sperme sera le reflet du tractus génital masculin dans un ensemble à moment donné. Il est donc recommandé de faire pratiquer au moins deux spermogrammes à intervalle de 2 à 3 mois) pour mieux évaluer le profil et la qualité spermatique (Lornage, 2002; WHO, 2010).

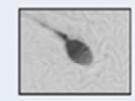
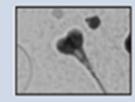
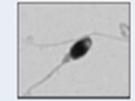
Il est important enfin de souligner que comme pour tous les examens reposant sur certains facteurs humains comme l'observation microscopique, il existe une certaine variabilité inter-laboratoire. De plus, les données du spermogramme ne reposent que sur petite fraction des spermatozoïdes pouvant entraîner un biais dans l'évaluation globale de la qualité du sperme.

1.2. Le spermocytogramme

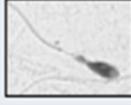
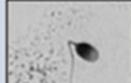
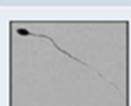
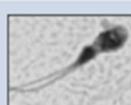
Le spermocytogramme consiste à rechercher des anomalies morphologiques des spermatozoïdes en microscopie optique (tableau 3). La morphologie spermatique peut être évaluée après coloration des spermatozoïdes sur un frottis. Plusieurs types de coloration existent, les plus utilisées étant les colorations de Shorr et de Papanicolaou modifiée.

Plusieurs classifications ont été établies. En France, c'est la classification de David modifiée qui est la plus largement répandue (Auger *et al.*, 2000). L'observation est généralement réalisée sur 100 spermatozoïdes. L'ensemble des anomalies pour chaque spermatozoïde est répertorié permettant de définir un index d'anomalies multiple (nombre total d'anomalies/nombre de spermatozoïdes anormaux) révélant le nombre moyen d'anomalies par spermatozoïde.

Tableau 3. Différentes anomalies morphologiques du spermatozoïde selon la classification de David modifiée. Adapté d'Auger, 2010.

	Anomalies	Illustrations
Anomalies de la tête	Amincie	
	Allongée	
	Microcéphale	
	Macrocéphale	
	Multiples	
	Acrosome anormal ou absent	
	Base anormale ou absente	

(suite page suivante)

Anomalies de la pièce intermédiaire	Grêle (gaine de mitochondries incomplètes)	
	Angulée	
	Restes cytoplasmiques (corps résiduel persistant)	
	Absent	
Anomalies de la pièce principale	Court	
	Enroulé	
	Calibre irrégulier	
	Multiples	

2. Causes génétiques de l'infertilité masculine

Une importante détérioration de la qualité spermatique a été observée ces dernières décennies, particulièrement dans les pays développés, tant chez l'homme (Rolland *et al.*, 2013) que dans les élevages animaux (Royal *et al.*, 2000). Cela souligne l'impact des facteurs environnementaux sur la qualité spermatique. Parallèlement, la composante génétique occupe aussi une part très importante dans les causes d'infertilité masculine puisque on estime que l'infertilité masculine d'origine génétique pourrait concerter près de 1'homme sur 40 (Tüttelmann *et al.*, 2011). Ces données sont appuyées par l'augmentation des cas d'infertilité au sein d'une même fratrie et l'étude des modèles animaux notamment murins (Matzuk et Lamb, 2008).

2.1. Les causes génétiques fréquentes

2.1.1. Les microdélétions du chromosome Y

Le chromosome Y est un petit chromosome d'une taille d'environ 53 Mb. Il représente 1,7% du génome humain et contient 78 gènes principalement impliqués dans la différenciation sexuelle masculine (SRY) et la spermatogénèse (Skaletsky *et al.*, 2003).

Des comparaisons menées sur les gonosomes suggèrent que le chromosome X et Y dérivent de chromosomes reptiliens (les protochromosomes X et Y) qui sont apparus il y a plus de 180 millions d'années (Cortez *et al.*, 2014). Au cours de l'évolution, les deux chromosomes ont divergés et le chromosome Y a perdu 97 % de ses gènes ancestraux. Ceux qui subsistent interviennent principalement dans la formation des testicules et la spermatogénèse. Depuis 25 millions d'années, le chromosome Y aurait cessé de dégénérer.

A l'exception des régions pseudoautosomiques le chromosome Y étant isolé ne recombine pas et à la différence du chromosome X, ne peut bénéficier de réparation au cours de la méiose. Ainsi les mutations affectant le chromosome Y sont souvent délétères ne pouvant être réparées ou compensées. Pour compenser l'absence de recombinaisons de nombreux gènes du chromosome Y ont été dupliqués et peuvent recombiner entre eux permettant les mécanismes de conversion génique.

Depuis longtemps, les anomalies de structures du chromosome Y (anneaux, isodicentrique, grande délétion) ont été associées à des oligozoospermies sévères ou à des azoospermies (Yatsenko *et al.*, 2010). Généralement, ces anomalies respectent le facteur de déterminisme testiculaire SRY, situé à l'extrémité du bras court. La caractéristique commune à ces remaniements de structure de l'Y chez des patients infertiles est la perte d'une partie de la région euchromatique du bras long suggérant la présence d'un facteur de contrôle de la spermatogénèse dans cette région critique de l'Y appelé facteur AZF (Azoospermia Factor) (Tiepolo et Zuffardi, 1976). L'évolution des techniques a permis de mettre en évidence chez un certain nombre de patients infertiles la présence de délétions de cette région qui n'étaient pas visibles au caryotype (on parle de microdélétions). La région AZF est subdivisée en trois sous-régions différentes : AZFa, AZFb et AZFc qui contiennent des gènes impliqués dans la spermatogénèse et le développement des gonades (figure 27). Depuis plusieurs années de nombreuses séries de patients azoospermiques ou oligozoospermiques ont été publiées et montrent que globalement les microdélétions du chromosome Y sont retrouvées chez 10% des

hommes avec une azoospermie non-obstructives et chez 5% des patients avec une oligozoospermie sévères (<5 millions de spermatoïdes/ml) (Hotaling et Carrell, 2015).

Les délétions de la région AZFa qui contient principalement les gènes *USPY9* et *DBY* (figure 27) sont le plus souvent associées à la perte des cellules germinales et à un syndrome des cellules de Sertoli seules (SCO) (Nuti et Krausz, 2008). Les délétions de la région AZFb entraînent un blocage en méiose dû à l'absence du gène *RBMY* et *PRY* impliqués dans la régulation de l'apoptose (figure 27). Enfin les délétions d'AZFc sont associées à des défauts de la spermatogenèse qui peuvent être relativement modérés. Près de 70% des patients avec une délétion de la région AZFc auront des spermatozoïdes retrouvés après extraction chirurgicale testiculaire (TESE) (Nuti et Krausz, 2008). Le principal gène impliqué dans la région AZFc est le gène *DAZ* (figure 27) exprimé à toutes les étapes de la spermatogenèse (Kee *et al.*, 2009). D'autres gènes ont probablement un rôle important dans la spermatogénèse : le gène *CDY*, un facteur de transcription qui intervient dans l'acétylation des histones et le gène *TSPY* (figure 27) qui déclenche l'entrée des spermatogonies en méiose (Vogt, 2005).

Le séquençage du chromosome Y en 2003 a révélé une structure globale du chromosome constituée d'une mosaïque de séquences uniques et d'autres répétées. La particularité notable du bras long de l'Y est la présence de 8 grands blocs palindromiques dans la région AZFc (figure 27) allant de plusieurs milliers à près d'un million et demi de paires de bases et constitués de séquences quasiment identiques dupliquées en miroir. (Skaletsky *et al.*, 2003). On sait que ces palindromes peuvent être le siège de recombinaisons intrachromatidiennes par mécanisme de recombinaison homologue non allélique (NAHR) ce qui explique le caractère récurrent des délétions d'AZFc. Des recombinaisons entre palindromes expliquent aussi les délétions AZFb ou AZFb+c. Les délétions AZFa, quant à elles, sont plus rares et mettent aussi en jeu des phénomènes de recombinaison qui ont lieu entre les séquences homologues de nature rétrovirale.

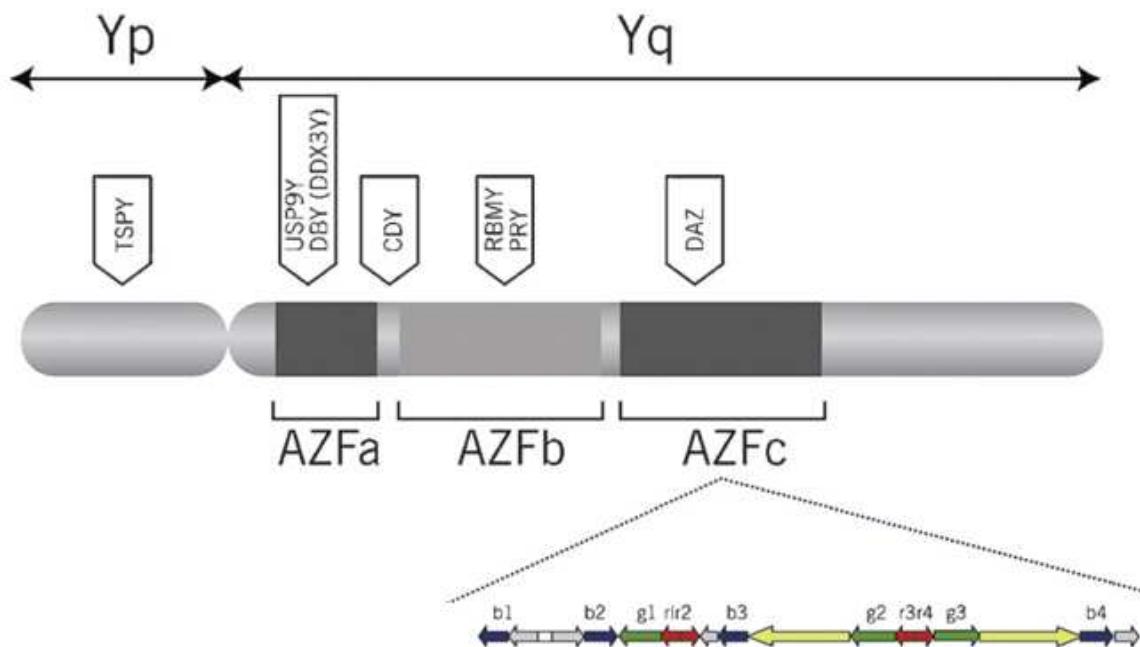


Figure 27. Représentation schématique du chromosome Y et de la région AZF ainsi que des trois sous-régions AZF a, b,c et des principaux gènes compris dans chacune des sous-régions. Un agrandissement montre les nombreuses séquences palindromiques situées dans la sous-région AZFc. Adpaté de O'Flynn *et al.*, 2010.

2.1.2 Anomalies chromosomiques

Des anomalies chromosomiques de nombre ou de structure des autosomes mais surtout des gonosomes peuvent être impliquées. Le pourcentage d'anomalies chromosomiques observé sur caryotype sanguin chez les patients infertiles s'échelonne entre 2 et 8 % et peut atteindre jusqu'à 15 % chez les patients azoospermiques, soit 10 à 20 fois la fréquence retrouvée dans la population générale (Ravel *et al.*, 2006).

2.1.2.1. Syndrome de Klinefelter

Le syndrome de Klinefelter est la cause la plus fréquente d'hypogonadisme et d'infertilité chez l'homme. Sa prévalence est 50 fois plus élevée chez les patients infertiles azoospermiques (14 %) que dans la population générale (~0,2 %) (Gekas *et al.*, 2001). Le syndrome de Klinefelter est dû à la présence d'un chromosome X supplémentaire (47,XXY) et a une prévalence estimée à 1/600 (Bojesen et Gravholt, 2007). On retrouve une formule chromosomique homogène dans 80 à 90% des cas, Les 10 à 20% restants sont des

présentations en mosaïques avec un contingent de cellules 46,XY (Lanfraco *et al.*, 2004). Ces formes en mosaïque sont généralement cliniquement moins sévères que les formes homogènes.

Les manifestations sont variables d'un individu à l'autre et n'apparaissent pas chez tous les porteurs du syndrome de Klinefelter. Les manifestations physiques sont souvent imperceptibles durant l'enfance et elles apparaissent à la puberté. La découverte se fait cependant dans 90 % des cas lors d'un bilan d'infertilité. Au niveau clinique, on va retrouver une grande taille à l'âge adulte, une gynécomastie (c'est-à-dire une augmentation du volume des glandes mammaires uni ou bilatérales), des testicules petits (hypogonadisme) avec un pénis de taille normale, une pilosité peu développée et au niveau des dents, la pulpe dentaire, peut être anormalement plus développée (taurodontisme). Les enfants porteurs du syndrome de Klinefelter n'ont pas de déficit intellectuel, avec des variations comme dans la population générale. Cependant, les enfants porteurs de ce syndrome peuvent présenter des retards dans les premières acquisitions : apprentissage du langage, de la lecture, développement de la motricité. Les hommes porteurs de ce syndrome sont infertiles avec en grande majorité une absence totale de spermatozoïdes (azoospermie) dans l'éjaculat. Des spermatozoïdes pourront être retrouvés dans l'éjaculat chez moins de 10% des patients Klinefelter avec une formule homogène (Hotaling et Carrell, 2015). Les biopsies testiculaires chez ces patients montrent un arrêt de la spermatogénèse au stage spermatocyte primaire voir une hyalinisation partielle ou totale des tubes séminifères (Huynh *et al.*, 2002).

Il est actuellement possible de proposer une prise en charge par AMP à certains couples, dont le conjoint est porteur d'un syndrome de Klinefelter, dès lors que quelques spermatozoïdes mobiles peuvent être obtenus, soit dans l'éjaculat, soit au niveau testiculaire. Cependant, le recours à ces techniques dans le cas d'une aneuploïdie des cellules germinales pose le problème de la transmission de cette aneuploïdie à la descendance et du risque de stérilité de la descendance masculine. Ainsi, dans le cas du syndrome de Klinefelter, le risque de transmission du chromosome X surnuméraire est directement lié au pourcentage de spermatozoïdes aneuploïdes 46,XY ou 46,XX. Il faut noter que cette proportion augmente fortement avec l'âge paternel (Arnedo, 2005). De plus, il a été observé chez ces patients une augmentation des spermatozoïdes diploïdes pour certains autosomes comme le chromosome 13, 18 et 21 (Hennebicq *et al.*, 2001; Morel *et al.*, 2003). Bien qu'aucun cas de naissances d'enfants avec une anomalie chromosomique après ICSI n'ait été recensé, un conseil génétique est indispensable chez ces patients ayant un projet parental et des solutions comme le diagnostic pré-implantatoire (DPI) pourra être alors envisagé (Hotaling et Carrell, 2015).

2.1.2.2. Anomalies de structures

Parmi les anomalies de structure principalement retrouvées chez les patients infertiles on distingue les translocations et les inversions.

Les translocations correspondent à un échange de matériel chromosomique entre deux chromosomes non homologues. Chez les patients infertiles, ces translocations sont généralement équilibrées c'est-à-dire qu'il n'y a pas de perte ou gain de matériel chromosomique. On va distinguer les translocations réciproques et les translocations robertsonniennes impliquant les chromosomes acrocentriques (13, 14, 15, 21, 22).

Les translocations réciproques sont retrouvées 4 à 10 fois plus fréquemment chez les patients infertiles que dans la population générale (Elliott et Cooke, 1997). Chez ces patients, le taux de spermatozoïdes avec un contenu chromosomique déséquilibré est important mais le taux est variable selon les chromosomes impliqués dans la translocation (Benet *et al.*, 2005). Ces déséquilibres exposent à un fort risque de fausse couche et il est recommandé d'évaluer le taux de spermatozoïdes par FISH avant toute prise en charge thérapeutique par des techniques d'AMP (Carrell, 2008). En fonction, des prises en charge alternatives comme le DPI ou le don de sperme seront à envisager.

Les translocations robertsonniennes peuvent entraîner aussi un défaut de la spermatogenèse et sont retrouvées chez 1.6% des patients oligozoospermiques et 0.09% des patients azoospermiques (O'Flynn O'Brien *et al.*, 2010). Comme pour les translocations réciproques, les porteurs de translocations robertsonniennes ont un risque de produire des gamètes déséquilibrés dont le taux dépendra du type de ségrégation lors de la méiose.

Les inversions chromosomiques sont également retrouvées plus fréquemment chez les patients infertiles que dans la population générale (Krausz et Forti, 2000). Les inversions correspondent à la cassure d'un fragment de chromosome, suivie de rotation à 180°C de ce puis de sa réintégration à la même position chromosomique. Ces inversions vont gêner l'appariement des chromosomes homologues (formation d'une boucle d'inversion) pendant la méiose.

2.1.2.3. Autres anomalies chromosomiques

On peut par exemple citer les hommes de formule 46,XX présentant un phénotype masculin normal. Le principal mécanisme en cause est la translocation d'une partie du chromosome Y contenant le gène SRY sur un autre chromosome. Ces patients sont

généralement totalement infertiles et présentent une azoospermie par absence des sous-régions AZFa,b et c (Vorona *et al.*, 2007).

2.1.3. Mutations CFTR

L'identification du gène *CFTR* (Cystic Fibrosis Transmembrane conductance Regulator) chez les patients atteints de mucoviscidose qui présentent habituellement une agénésie bilatérale des canaux déférents (ABCD) a permis d'associer ce gène aux agénésies des déférents isolées. Cette malformation est la cause d'environ 2 % des cas d'infertilité masculine et de près de 25 % des azoospermies obstructives (Yu *et al.*, 2012). Dans ce contexte, les hommes ABCD sont souvent porteurs d'au moins une mutation du gène *CFTR* majeure (ex : p.F508del). Le plus souvent cette première mutation est associée à une seconde mutation dite « mineure » (ex : R117H, allèle 5T) ce qui explique que ces patients présentent souvent uniquement un problème de fertilité et non l'ensemble des signes cliniques généralement associés à la mucoviscidose (Yu *et al.*, 2012).

La prévalence de ces anomalies génétiques varie selon le phénotype concerné mais de manière globale et tous phénotypes confondus, ces défauts sont retrouvés dans moins de 5 % des cas (Nieschlag *et al.*, 2010). Ces chiffres démontrent clairement l'implication certaine d'autres gènes encore inconnus dans les différents phénotypes d'infertilité masculine.

2.2. Identification de nouveaux gènes : approches et résultats

2.2.1. Approche « gènes candidats »

L'approche gène candidat consiste à rechercher des mutations chez un patient dans un gène cible. Le gène cible étudiée pourra se définir selon plusieurs critères. Le premier d'entre eux est l'étude de gènes reliés à des phénotypes d'infertilité similaires dans différentes modèles animaux et notamment murins. Dans ce cas, les mutations seront recherchées sur le gène orthologue humain (de Boer *et al.*, 2014). Une autre possibilité est de rechercher des mutations dans des gènes paralogues à un gène précédemment identifié supposant que leur structure proche implique une fonction similaire. Cette approche est bien souvent infructueuse dû en grande partie à la grande hétérogénéité génétique des phénotypes étudiés et du nombre

limité de patients testés (Elinati *et al.*, 2012). Cependant quelques succès retentissants utilisant l'approche gènes candidats peuvent être rapportés (tableau 4).

- **Résultats dans la téратozoospermie**

Dans la globozoospermie, un phénotype rare de tératozoospermie défini par la présence dans l'éjaculat de spermatozoïdes microcéphales dépourvus d'acrosome, différentes équipes ont étudié chez l'homme certains des gènes rapportés pour être impliqués dans la biogénèse de l'acrosome chez la souris. Ainsi, une mutation faux-sens homozygote a été retrouvée dans le gène *PICK1* dans une famille chinoise (Liu *et al.*, 2010).

Une seconde équipe a identifié des mutation faux-sens et d'épissage de le gène *ZPBP1* chez des patients avec des anomalies de la tête spermatique (Yatsenko *et al.*, 2012). Ces données n'ont cependant pas été validées fonctionnellement et jusqu'à présent aucune nouvelle mutation dans ces gènes n'a été rapportée.

Enfin, le gène *AKAP4* (A-kinase anchor protein 4) a été étudié chez des patients présents des anomalies morphologies et ultrastructurales des flagelles. Ces études s'appuient sur l'observation de spermatozoïdes présentant des flagelles courts et malformés chez les souris KO pour le gène *Akap4* (Miki *et al.*, 2002). Une délétion homozygote de la totalité du gène a pu être identifiée chez un patient (Baccetti *et al.*, 2005). Aucune autre mutation n'a été identifiée dans une cohorte supplémentaire de 9 patients avec des anomalies flagellaires (Turner *et al.*, 2001).

- **Résultats dans l'asthénozoospermie**

En 2012, Kuo et collaborateurs ont identifié des mutations faux-sens hétérozygotes dans le gène *SEPT12* codant pour une GTPase chez deux patients avec une oligoasthénozoospermie (Kuo *et al.*, 2012). Les auteurs avaient précédemment mise en évidence que l'absence de la protéine Septin12 chez la souris entraînait des défauts de mobilité du spermatozoïde, des flagelles coudés et des anomalies morphologiques de la tête (Lin *et al.*, 2009) . Les auteurs ont pu démontrer que ces deux mutations faux-sens impactaient fortement l'activité catalytique et les interactions au GTP de la protéine (Kuo *et al.*, 2012).

L'année suivante une seconde équipe a identifié 3 mutations faux-sens hétérozygotes dans le gène *SLC26A8* (connu aussi sous le nom de TAT1) chez 146 patients (Dirami *et al.*, 2013). La

protéine TAT1 est un transporteur d'anions exclusivement exprimé dans le testicule adulte humain au niveau de la membrane plasmique des cellules germinales. En 2007, cette même équipe avait observé chez la souris *Tat1*-/- des défauts majeurs de la structure du flagelle des spermatozoïdes associés à une absence de mobilité (asthénozoospermie) (Touré *et al.*, 2007). Plus récemment, une dernière équipe a identifié à partir des données produites chez la souris mutante hétérozygote, une délétion hétérozygote d'un nucléotide dans l'exon 6 du gène *GALNTL5* (N-acetylgalactosaminyltransferase-like protein 5) chez un patient parmi 200 patients asthénozoospermiques étudiés (Takasaki *et al.*, 2014). Les auteurs supposent que cette mutation a dû être transmise par la mère du patient bien que cela n'ait pu être vérifié.

- **Résultats dans l'azoospermie**

Dans l'azoospermie, plusieurs études basées sur une stratégie « gène candidat » identique ont été décrites. En 2003, Miyamoto et collaborateurs ont étudié le gène *SYCP3* chez 19 patients azoospermiques avec arrêt de maturation. Les études chez le modèle murin mutant avaient précédemment démontré que l'absence de la protéine Sycp3 entraînait une azoospermie par blocage méiotique consistant avec le rôle de Sycp3 dans le complexe synaptonémal (Yuan *et al.*, 2000). Une délétion hétérozygote de 643delA a été détectée chez un des 19 patients étudiés. La protéine mutante a été produite et a permis de démontrer un effet dominant négatif de la mutation inhibant la fonction de la protéine (Miyamoto *et al.*, 2003). Des études ultérieures sur 58 patients belges avec un blocage méiotique n'a pas permis de retrouver de mutations dans le gène *SYCP3* (Stouffs *et al.*, 2005).

En 2006, Choi et collaborateurs ont séquencé le gène *SOHLH1* chez 96 patients azoospermiques (non obstructive) basé uniquement sur le constat d'un phénotype similaire observé chez la souris knockout (Ballow *et al.*, 2006). Le gène *SOHLH1* est un facteur à domaine basique hélice-boucle-hélice (bHLH) spécifique des cellules germinales régulant l'expression de plusieurs gènes (Ballow *et al.*, 2006). Trois mutations hétérozygotes dans le gène *SOHLH1* ont pu être identifiées dont une mutation d'épissage et deux mutations faux-sens. Les études fonctionnelles ont démontré l'effet délétère de la mutation d'épissage produisant une protéine avec un domaine bHLH tronqué incapable d'interagir avec ces protéines partenaires. La présence des deux autres mutations faux-sens n'a pas montré d'effet significatif (Choi *et al.*, 2006).

Egalement en 2006, Yatsenko *et al.*, ont mis en évidence des mutations hétérozygotes dans le gène *KLHL10* chez 3 patients sur une cohorte de 270 patients avec une oligozoospermie sévère suggérant qu'elles étaient responsables du phénotype puisque l'haplo-insuffisance de *Klhl10* entraîne une infertilité chez la souris (Yan *et al.*, 2004). La protéine KLHL10 est fortement exprimée durant les étapes de la phase de différenciation des spermatozoïdes et interagit avec le protéine CUL3 (cullin3) pour former un complexe E3 ubiquitine ligase capable d'initier l'ubiquitination de certaines protéines durant la spermiogénèse (Wang *et al.*, 2006). Il a été montré que la présence des deux mutations faux-sens entraînait un défaut d'homodimerisation de la protéine confirmant leur probable pathogénicité (Yatsenko *et al.*, 2006). Aucune autre mutation n'a été retrouvée par ailleurs chez 325 patients infertiles (Qiu *et al.*, 2009).

En 2010, le gène *NR5A1* a été séquencé chez 315 patients avec une azoospermie idiopathique (Bashamboo *et al.*, 2010). Le gène *NR5A1* code pour le facteur stéroïdogénique 1 (SF-1) et agit comme un régulateur de la transcription de nombreux gènes impliqués dans l'axe hypothalamo-hypophysio-gonadique. Il a une expression forte au niveau testiculaire et notamment au niveau des cellules de Sertoli et de Leydig et régule l'expression de *SOX9* et de *l'AMH*. Les mutations dans ce gène avaient été préalablement identifiées dans les insuffisances ovariennes précoces ainsi que dans le cadre de certaines anomalies de la différenciation (Lin *et al.*, 2008 ; Lourenço *et al.*, 2009). Dans son étude, Bashamboo et collaborateurs (2010) ont identifié 7 patients avec des mutations faux-sens hétérozygotes dans le gène *NR5A1* toutes situées dans la région charnière (hinge) de la protéine. Les études fonctionnelles confirmeront un effondrement de l'activation de la transcription de certains gènes régulés par *NR5A1* comme *l'AMH*. Ces données démontrent une nouvelle fois que pour un même gène il peut exister un large spectre de manifestations cliniques en fonction du type et de la position des mutations incriminées. De nouvelles mutations dans le gène *NR5A1* ont été par la suite identifiées chez plusieurs cohortes de patients azoospermiques confirmant l'implication et la forte prévalence des altérations de ce gène dans l'azoospermie (Ropke *et al.*, 2013 ; Zare-Abdollahi *et al.*, 2014 ; Ferlin *et al.*, 2015). De manière intéressante et illustrant parfaitement la stratégie des gènes candidats, la même équipe publiera simultanément des mutations hétérozygotes dans le gène *GATA4* codant pour un facteur de transcription interagissant avec *NR5A1* chez des patients présentant des anomalies de la différenciation sexuelle (Lourenço *et al.*, 2009).

Une approche gène candidat reposant sur le rôle des gènes de la famille Nanos dans la spermatogénèse a été utilisée pour l'identification de mutations chez une cohorte de 195

patients avec une azoospermie ou une oligozoospermie (Kusz-Zamelczyk *et al.*, 2013). Sur les trois gènes Nanos chez la souris, l'altération des gènes *Nanos2* et *Nanos3* entraînent une infertilité alors que le gène *Nanos1* ne semble pas avoir d'effet sur la reproduction (Haraguchi *et al.*, 2003). Etrangement, les 3 mutations identifiées (2 délétions en phase à l'état hétérozygote et une double substitution) chez 3 patients non apparentés dans cette cohorte se situent sur le gène *NANOS1*. Dans cette étude, le niveau de preuve de la pathogénicité de ces trois mutations demeure insuffisante et nécessitent d'être confirmés sur de plus larges cohortes.

Enfin dernièrement, le gène *TEX11* a été aussi relié à l'azoospermie à partir d'une approche gène candidat (Yatsenko *et al.*, 2015). Chez la souris mutante inactivée pour *Tex11*, il a été décrit une infertilité due à un arrêt complet de la méiose de cellules germinales au stade pachytène (Adelman *et al.*, 2008; Yang *et al.*, 2008). Le gène a été étudié par CGH-array et séquençage sur une cohorte de 289 hommes azoospermiques. Deux délétions homozygotes intragéniques ainsi que 5 autres mutations ponctuelles homozygotes (épissage et faux sens) ont pu être identifiées. Dans environ 1 cas sur 3 ces mutations ont été retrouvés chez des patients avec un blocage méiotique (Yatsenko *et al.*, 2015).

On peut enfin citer les mutations identifiées dans les gènes *PRM1* et *UTP14c* chez des cohortes de 20 et 234 patients azoospermiques respectivement (Iguchi *et al.*, 2006 ; Rohozinski *et al.*, 2006). Cependant, les mutations trouvées dans ces publications ont été secondairement retrouvées dans une large cohorte de patients fertiles remettant fortement en question leur implications dans ces phénotypes (Kichine *et al.*, 2008)

- **Autre phénotype**

Un dernier exemple de l'utilisation avec succès de l'approche gène candidat est l'identification de mutations du gène *PLCZ1* chez des patients pour lesquels une absence d'activation ovocytaire a été diagnostiquée. La protéine PLC ζ (zeta) se situe au niveau de la thèque périnucléaire et a été identifié chez la souris et l'homme comme l'un des facteurs essentiels dans la cascade d'activation ovocytaire (Jessica *et al.*, 2015). L'étude du gène *PLCZ1* chez 9 patients avec absence d'activation ovocytaire a permis d'identifier 1 patient hétérozygote composite pour 2 mutations faux-sens (H398P et H233L) (Heytens *et al.*, 2009 ; Kashir *et al.*, 2012).

2.2.2. Cartographie par homozygotie

De nombreuses publications ont rapporté des cas d'infertilité au sein d'une même fratrie et notamment issus d'union consanguine (Matzuk et Lamb, 2008). Ces observations évoquent un mode de transmission autosomique récessive du trait pathologique responsable de l'infertilité. Dans les pathologies autosomiques récessives, l'utilisation de l'approche par cartographie d'homozygotie permet d'identifier des régions d'homozygotes communes à plusieurs patients et de localiser des gènes d'intérêt pouvant contenir la mutation responsable (figure 28). Cette approche repose sur l'analyse globale des polymorphismes du génome avec des puces à ADN de type SNP (single nucléotide polymorphisms) (figure 28). Les régions d'homozygotie ainsi définies sont généralement d'une taille supérieure à 10Mb. Cette région doit être ensuite décortiquée afin de mettre lumière les meilleurs gènes candidats susceptibles d'expliquer le phénotype. Cette étape peut être extrêmement laborieuse étant donné que certaines régions d'homozygoties peuvent contenir plusieurs centaines de gènes. Des éléments d'orientation comme l'expression du gène, sa fonction voir la publication d'un modèle animal permettra de cibler les recherches et de limiter le nombre de gènes à séquencer.

Cette stratégie a été appliquée avec succès dans plusieurs études (Tableau 4). Les principaux résultats ont été obtenus chez des patients présentant différentes formes homogènes de téратozoospermie. Ainsi, le gène *AURKC* a été identifié dans la macrozoospermie (Dieterich *et al.*, 2007), les gènes *SPATA16*, *DPY19L2* dans la globozoospermie (Dam *et al.*, 2007 ; Harbuz *et al.*, 2011 ; Koscinski *et al.*, 2011) et plus récemment le gène *DNAH1* car les anomalies flagellaires multiples (Ben Khelifa *et al.*, 2014). Cette partie sur la tératozoospermie sera traitée et discutée en détails dans l'article de revue (5) de la partie discussion (cf plus bas).

Dans les autres phénotypes comme l'asthénozoospermie, on citera l'identification de mutations dans le gène *CATSPER1* à partir de deux familles consanguine iraniennes. Deux insertions nucléotidiques homozygotes ont été mises en évidence chez 3 patients asthénozoospermiques dont 2 frères (Avenarius *et al.*, 2009). Les deux insertions (c.539-540insT et c.948-949insATGGC) entraînent l'apparition d'un codon stop prématûre et peuvent aboutir à la formation d'une protéine tronquée dépourvues de ses 6 domaines transmembranaires. Ce défaut abrogerait totalement la fonction de canal calcium voltage dépendant. En effet, la protéine *CATSPER1* permet l'entrée du calcium dans le flagelle nécessaire à la capacitation et à l'hyperactivation du spermatozoïde, des processus

fondamentaux pour la fécondation de l'ovocyte (Ren *et al.*, 2001; Carlson *et al.*, 2003).

Enfin dans l'azoospermie, 2 gènes ont été identifiés par l'étude de 2 familles consanguines d'origine turque (Ayhan *et al.*, 2014). Il s'agit de *TAF4B*, un facteur de transcription testiculaire et *ZMYND15*, un répresseur transcriptionnel dépendant des histone-déacétylases. Ces gènes se sont révélés être d'excellents candidats puisque les auteurs ont identifiés des mutations tronquantes dans les deux gènes et que les modèles KO pour ces deux gènes présentent une azoospermie isolée (Falender *et al.*, 2005; Yan *et al.*, 2010). Cependant, aucune mutation supplémentaire n'a été identifiée dans une cohorte complémentaire de 45 patients azoospermiques et 15 oligozoospermiques sévères.

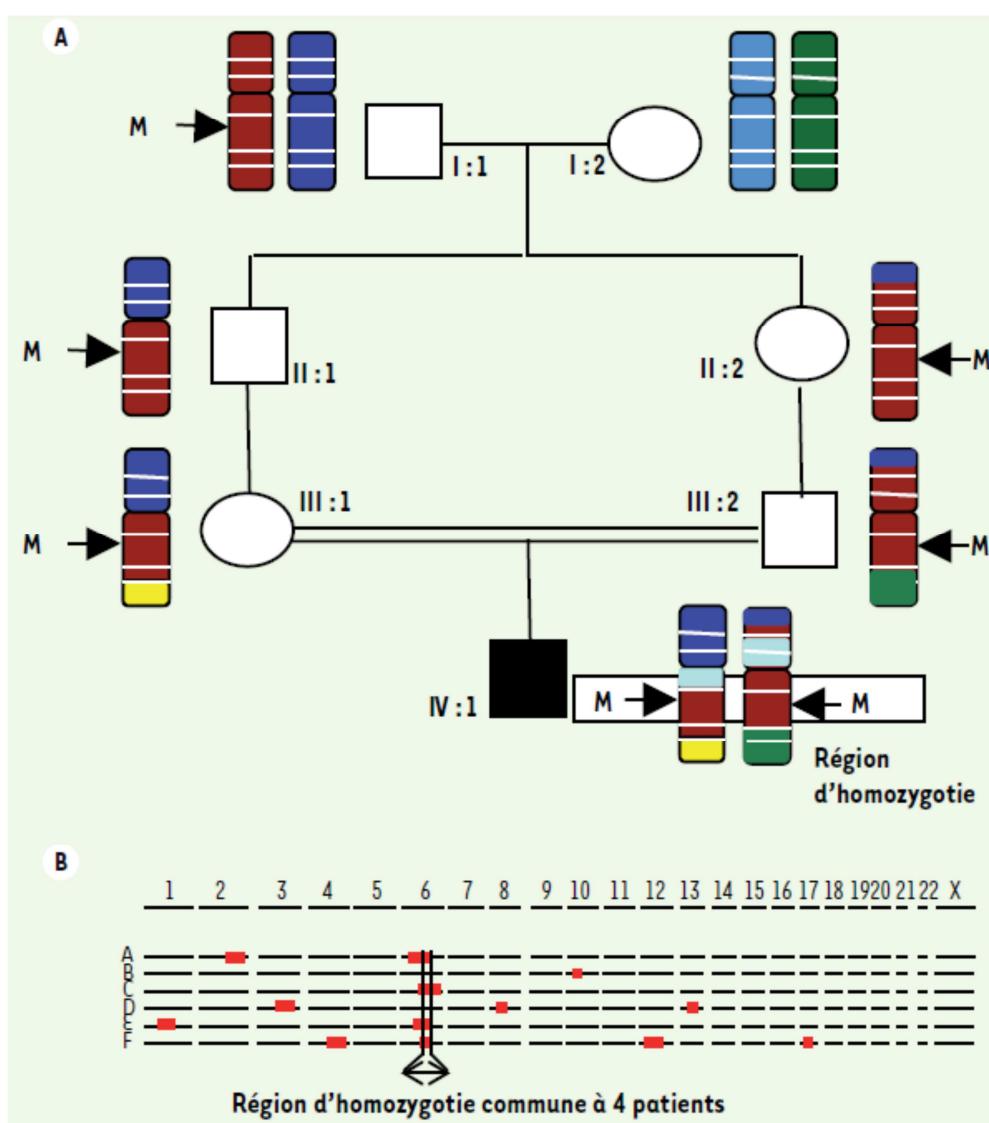


Figure 28. Stratégie de localisation d'un gène pathogène par cartographie d'homoygوتie.
A. Pour les patients atteints d'une maladie génétique rare dont les parents sont apparentés (cousins au 1er ou au 2ème degré), on postule que la pathologie ségrège sur un mode autosomique récessif et que les patients ont hérité de deux copies d'une même mutation

familiale. Pour simplifier on ne représente ici qu'un seul des 23 chromosomes, porteur de la mutation d'intérêt. L'arrière grand-père (I:1) du patient (IV:1) est porteur d'une seule copie de la mutation M. La mutation est transmise à l'état hétérozygote à ses enfants et à ses petits enfants avant d'être héritée à l'état homozygote par IV:I. Des recombinaisons entre chromosomes homologues se produisent à chaque méiose, contribuant à la diversité génétique. Deux copies d'une même région chromosomique, plus ou moins grande et entourant la mutation, sont héritées par le patient IV:1. Cette région, dépourvue des variations génétiques normalement présentes entre chromosomes homologues est appelée « région d'homozygotie ». B. Une analyse sur « puces à ADN » de type SNP (single nucleotide polymorphism) permet d'analyser des variants génétiques répartis sur tout le génome et de repérer les régions d'homozygotie. Cette analyse est réalisée sur plusieurs patients (A-F) présentant le même phénotype et issus d'une union consanguine. La présence d'une région d'homozygotie commune à plusieurs patients peut indiquer que la mutation causale se trouve dans cette région. Une étude *in silico* des gènes se trouvant dans la plus petite région d'homozygotie commune est réalisée. Si un gène, de par sa fonction et son profil d'expression paraît être un bon candidat, il sera séquencé chez tous les patients. Si une mutation ayant un effet délétère sur la protéine est identifiée chez plusieurs patients, on conclut que ce gène est impliqué dans la genèse des symptômes. D'après Coutton *et al.*, 2012.

2.2.3. Séquençage nouvelle génération

Récemment une nouvelle méthode à révolutionné les investigations génétiques et permis d'accélérer de manière exponentielle l'identification des gènes impliqués dans différentes pathologies humaines : il s'agit du séquençage nouvelle génération ou NGS (pour next generation sequencing). Cette technologie permet de réaliser l'étude simultanée de mutations à l'échelle d'un panel de gènes, de tous les exons (exomes) ou du génome entier.

Très schématiquement cette technique se décompose en 4 étapes (figure 29). Tout d'abord l'ADN génomique va être digéré afin d'obtenir des fragments de petite taille uniformes. Ces fragments d'ADN seront ensuite préparés (fixation d'adaptateurs, amplification) afin de constituer une banque d'ADN (on parle de librairie). Il est possible à cette étape d'effectuer un enrichissement de la librairie pour certaines régions d'intérêt par différentes approches (PCR multiplex ou hybridation en milieu liquide). Ainsi seules les séquences cibles seront sélectionnées et ensuite séquencées. L'enrichissement pourra concerner par exemple un panel de gènes ou l'ensemble des exons du génome (exomes). Une fois la librairie constituée, une étape d'amplification clonale va permettre d'amplifier de manière spécifique chacun des fragments à séquencer. Différentes méthodes existent pour réaliser cette amplification clonale comme la PCR en émulsion ou la bridge-PCR. Enfin, la totalité des fragments amplifiés seront séquencés générant des signaux protoniques ou photoniques qui seront ensuite traduits en séquences nucléotidiques. Les séquences obtenues seront ensuite alignées sur le génome de référence et l'application de différents algorithmes

permettra secondairement de détecter les possibles mutations causales parmi les milliers de variants détectés.

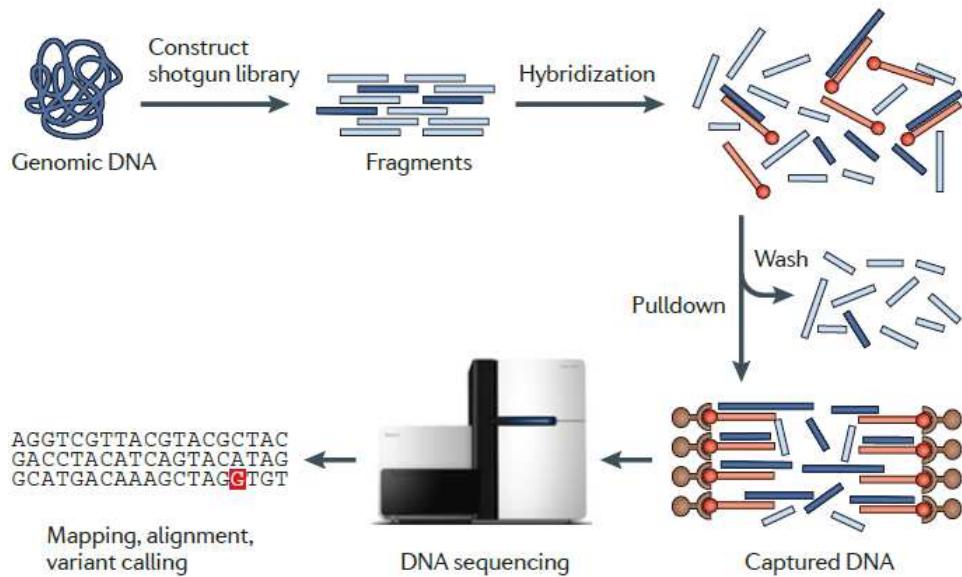


Figure 29. Principe général du séquençage nouvelle génération ou NGS. D'après Bamshad *et al.*, 2011.

Plusieurs stratégies existent pour l'identification des gènes responsables par NGS. Le séquençage simultané des tous les exons ou whole-exome sequencing (WES) est aujourd'hui la stratégie de première intention la plus employée pour l'identification de nouveaux gènes en génétique humaine (Robinson *et al.*, 2011). Le WES permettrait d'identifier le gène en cause dans plus d'un cas sur deux (60%) (Gilissen *et al.*, 2012). Cette approche globale bien qu'efficace requiert la capacité à trier les milliers de variants obtenus, étape qui peut s'avérer laborieuse. En effet, environ 25 000 variants par patient sont identifiés en moyenne pour chaque WES dont la très grande majorité est des polymorphismes bénins sans lien avec la pathologie (Singleton, 2011). Différents filtres techniques et biologiques peuvent être appliqués afin de trier ces variants et sélectionner que les mutations les plus significatives (Elinati *et al.*, 2012). Une autre stratégie pourrait être de se limiter à un large panel de gènes « candidats ». Cette stratégie présente l'avantage de sélectionner par leur localisation ou leur fonction présumée des gènes qui pourraient *a priori* être impliqués dans ces pathologies. Elle a été appliquée dans certains groupes de pathologies comme par exemple dans les ciliopathies avec l'étude des gènes du « ciliome » (Bredrup *et al.*, 2011; Perrault *et al.*, 2012; Schmidts *et al.*, 2013). Cependant, cette sélection entraîne un biais important et exclut de nombreux gènes

de fonction peu ou non connue, pourtant potentiellement pathogènes. Ces stratégies « ciblées » se justifient plus dans des études à visée diagnostique que dans les recherches de nouveaux gènes pathologiques (Berg *et al.*, 2011).

Grâce aux approches de séquençage haut-débit l'identification des gènes en génétique médicale humaine est devenue « simple » et efficace comme en témoigne le nombre croissant de gènes identifiés ces dernières années dans de multiples pathologies (Bamshad *et al.*, 2011). La principale difficulté est actuellement de valider les mutations identifiées dans les gènes candidats potentiels afin de confirmer le lien entre l'anomalie et le phénotype. Cette exigence incontournable passe entre autres par l'utilisation de différents modèles animaux ou cellulaires.

Les références relatives à l'utilisation et l'identification de nouveaux gènes impliqués dans l'infertilité par NGS sont pour le moment peu nombreuses. Cependant quelques résultats prometteurs ont été obtenus sur des cohortes de patients azoospermiques (Tableau 4). La première publication a rapporté l'indentification de gènes dans l'azoospermie par NGS est une large étude chinoise qui a procédé au séquençage de 600 gènes « liés à l'infertilité » (stratégie des panels) chez 766 patients azoospermiques et 521 contrôles (Mou *et al.*, 2013). L'étude présente uniquement l'identification des mutations dans un unique gène, le gène *HSF2* (Heat shock factor 2). Les auteurs identifient 5 mutations hétérozygotes et les études fonctionnelles indiquent qu'au moins une de ces mutations a un effet dominant négatif sur la forme sauvage de *HSF2*. L'interprétation des résultats obtenus restent cependant délicate puisque la perte de fonction de la protéine *Hsf2* chez la souris n'entraîne qu'un effet très modeste sur la fertilité (Wang *et al.*, 2004). Enfin très récemment, des mutations homozygotes dans le gène *NPAS2* ont été identifiées par WES à partir de l'étude d'une famille turque (Ramasamy *et al.*, 2015). Ainsi, une mutation faux-sens homozygote a été retrouvée chez 3 frères azoospermiques. La protéine *NPAS2* intervient dans le rythme circadien et interagit avec la protéine *BMAL1* dont le modèle murin KO est infertile. Hormis une validation *in silico* du potentiel effet délétère de la mutation, aucune preuve concrète de son implication dans le phénotype n'a été apportée.

Gènes	Phénotype	Méthodes d'identification	Nombre de patients mutés sur total	Niveau de preuve	Références
PICK1	Tératozoospermie (globozoospermie)	Gène candidat	1 (cas familial)	FAIBLE	Liu <i>et al.</i> , 2010
ZPBP1	Tératozoospermie (globozoospermie)	Gène candidat	15/381	FAIBLE	Yatsenko <i>et al.</i> , 2012
SPATA16	Tératozoospermie (globozoospermie)	Cartographie par homozygotie	3 (cas familial)	FORT	Dam <i>et al.</i> , 2007
DPYI9L2	Tératozoospermie (globozoospermie)	Cartographie par homozygotie	15/20 dont 2 frères	FORT	Harbuz <i>et al.</i> , 2011
AKAP4	Tératozoospermie (flagelles)	Gène candidat	1	MOYEN	Baccetti <i>et al.</i> , 2005
DNAH1	Tératozoospermie (flagelles)	Cartographie par homozygotie	7/20 dont 3 frères	FORT	BenKhelifa <i>et al.</i> , 2014
AURKC	Tératozoospermie (macrozoospermie)	Cartographie par homozygotie	14/14 dont 2 frères	FORT	Dieterich <i>et al.</i> , 2007
SLC26A8	Asthénozoospermie	Cartographie par homozygotie	3/146	FORT	Rode <i>et al.</i> , 2012
SEPT12	Asthénozoospermie	Gène candidat	2/200	FORT	Kuo <i>et al.</i> , 2012
GALNTL5	Asthénozoospermie	Gène candidat	1/200	MOYEN	Takasaki <i>et al.</i> , 2014
CATSPER1	Asthénozoospermie	Cartographie par homozygotie	3 (cas familial)	FORT	Avenarius <i>et al.</i> , 2009
SYCP3	Azoospermie	Gène candidat	1/9	FORT	Miyamoto <i>et al.</i> , 2003
SOHLH1	Azoospermie	Gène candidat	3/96	FORT	Bacetti <i>et al.</i> , 2005
KLHL10	Oligozoospermie sévère	Gène candidat	3/270	FORT	Yatsenko <i>et al.</i> , 2006
NR5A1	Azoospermie	Gène candidat	7/325	FORT	Bashamboo <i>et al.</i> , 2010

<i>NANOS1</i>	Azoospermie	Gène candidat	3/195	FAIBLE	Kusz-Zamelczyk <i>et al.</i> , 2013
<i>TEX11</i>	Azoospermie	Gène candidat	7/289	FORT	Yatsenko <i>et al.</i> , 2015
<i>TAF4B</i>	Azoospermie	Cartographie par homozygotie	4 (cas familial)	FORT	Ayhan <i>et al.</i> , 2014
<i>ZMYDM5</i>	Azoospermie	Cartographie par homozygotie	3 (cas familial)	FORT	Ayhan <i>et al.</i> , 2014
<i>HSF2</i>	Azoospermie	NGS	5/766	MOYEN	Mou <i>et al.</i> , 2013
<i>NPAS2</i>	Azoospermie	NGS	3 (cas familial)	FAIBLE	Ramasamy <i>et al.</i> , 2015
<i>PLCZ1</i>	Absence d'activation ovocytaire	Gène candidat	1/9	MOYEN	Kashir <i>et al.</i> , 2012

Tableau 4. Tableau récapitulatif des causes géniques identifiées dans l'infertilité masculine.

2.3. Causes syndromiques de l'infertilité

L'infertilité peut être associée à des tableaux cliniques plus complexes et rapportée dans de très nombreux syndromes (pour revue Matzuk et Lamb, 2008). Le mécanisme aboutissant à l'infertilité dans ces multiples syndromes peut être de nature très différente.

On peut par exemple citer le syndrome de dysplasie campomélique dû à des mutations dans le gène *SOX9*. Ce syndrome associe différentes anomalies squelettiques à des multiples malformations et à une réversion sexuelle chez des individus 46, XY. Le gène *SOX9* est responsable de la différenciation des chondrocytes mais aussi des cellules de Sertoli. Il est exprimé dans les crêtes génitales en amont du gène *SRY* et joue un rôle clé dans l'activation du gène de l'AMH (Hormone Anti Müllerienne) et la différenciation du testicule (Parker *et al.*, 1999).

Autre exemple, la maladie de Steinert (ou dystrophie myotonique de type 1) qui est une dystrophie musculaire myotonique due à l'amplification anormale d'une répétition du trinucléotide CTG situé dans la région 3' non codante du gène codant la DMPK (myotonic dystrophy protein kinase). L'atteinte gonadique est fréquente dans cette pathologie et est responsable généralement d'une azoospermie (Meola et Cardani, 2015). Cette atteinte gonadique est primaire et correspond à un hypogonadisme hypergonadotrope (augmentation LH, FSH, testostérone diminuée).

Enfin parmi les autres syndromes associés à l'infertilité on retrouve la large famille des syndromes entraînant un hypogonadisme hypogonadotrope (taux bas de testostérone, LH et FSH). Un exemple bien connu est le syndrome de Kallmann associant à un hypogonadisme à une anosmie. Les principales manifestations cliniques sont l'association d'un micropénis et d'une cryptorchidie, un retard pubertaire et un déficit de perception des odeurs chez les deux sexes. Le syndrome de Kallmann est dû à un défaut du développement du système olfactif et de la migration embryonnaire des neurones synthétisant la GnRH. A ce jour, six gènes responsables de la maladie ont été identifiés : *KAL1*, responsable de la forme liée au chromosome X, et *FGFR1*, *FGF8*, *PROKR2*, *PROK2*, *CDH7* impliqués dans les formes de transmission autosomique (Buck *et al.*, 2013).

2.4. Infertilité et polymorphismes

Ces dernières années ont vu l'explosion de très nombreuses publications présentant des études d'association entre différents variants ou polymorphismes et certains phénotypes d'infertilité masculine (Nuti et Krausz, 2008 ; Carell et Aston, 2011). Ces études concernent

des variants de différentes natures comprenant à la fois les variations nucléotidiques (SNP, single nucleotide polymorphism) et génomiques (CNV, copy number variation) (Aston, 2014). La plupart de ces études reposent sur une stratégie de gènes candidats basées sur les données de modèles animaux. Plusieurs dizaines de gènes ont été ainsi étudiés (Massart *et al.*, 2012).

Ces études présentent souvent de nombreuses faiblesses : un échantillonnage de patients et de contrôles trop faible, une méthodologie de l'étude parfois approximative (ethnies, âges, phénotypes différents entre les groupes) et l'absence de validation par des tests fonctionnels des variants identifiés. De plus, la plupart des variants retrouvés sont généralement à l'état hétérozygote et aucune étude familiale n'est ensuite faite afin d'identifier la présence ou l'absence du variant chez les autres membres de la famille. Certaines méta-analyses de plusieurs de ces études d'association révèlent déjà que seule une infime fraction des variants identifiés peuvent être considérés comme associés significativement à l'infertilité (Sato *et al.*, 2013 ; Lu *et al.*, 2014). Ainsi, des études sur de plus larges cohortes de patients correctement phénotypés et avec une méthodologie robuste sont nécessaires pour confirmer ou infirmer ces résultats (Carrell et Aston, 2011). De telles études plus rigoureuses ont été réalisées pour seulement quelques gènes comme par exemple *MTHFR* (Wei *et al.*, 2012), *GSTM1* (Song *et al.*, 2013) et *FSHB* (Tuttlemann *et al.*, 2012) révélant la présence de variants significativement associées à l'infertilité mais avec un impact toutefois très modéré.

D'autres études d'association plus globale type GWAS (Genome-wide association studies) ont également été réalisées chez des patients oligozoospermiques et azoospermiques (Aston *et al.*, 2010 ; Hu *et al.*, 2011). Ces études ont identifié finalement assez peu de nouveaux variants significatifs dont l'implication dans le phénotype reste à confirmer (Aston et Conrad, 2013). Les résultats peu concluants de ces études suggèrent cependant que certains phénotypes comme l'azoospermie sont des pathologies probablement « multigéniques » et qu'un seul variant ne peut expliquer leur origine.

2.5. Les défauts épigénétiques

Les défauts épigénétiques constituent une part croissante des causes retrouvées dans l'infertilité masculine (Carell 2012 ; Dada *et al.*, 2012). L'épigénétique correspond à l'étude des changements dans l'activité transcriptionnelle des gènes n'impliquant pas de modification de la séquence d'ADN. Les modifications épigénétiques peuvent résulter par exemple de la

méthylation de certains îlots CpG de régions promotrices des gènes ou encore des modifications chimiques post-traductionnelles des histones (acétylation, phosphorylation, méthylation, ubiquitination). Contrairement aux mutations qui affectent la séquence d'ADN, les modifications épigénétiques peuvent être réversibles.

Des modifications anormales des histones, des défaut de méthylation de gènes soumis à empreinte ou de séquence répétées comme les séquences Alu ont été identifiées dans plusieurs études chez des patients présentant une oligozoospermie et/ou des défauts de condensation de l'ADN (Kobayashi *et al.*, 2007 ; Marques *et al.*, 2008; Poplinski *et al.*, 2010 ; Hammoud *et al.*, 2011; El Hajj N. *et al.*, 2011 ; Nanassy et Carrell, 2011 ; Aston *et al.*, 2012 ; Montjean *et al.*, 2013). Ces données ont été confirmées sur des modèles murins infertiles dont les gènes codant pour des méthyltransférases ont été inactivés (Kato *et al.*, 2007 ; Hussain *et al.*, 203).

En plus d'affecter directement la spermatogénèse, les défauts épigénétiques du génome gamétique male sont aussi directement corrélés à un mauvais développement des embryons après ICSI (Kobayashi *et al.*, 2009; Hammoud *et al.*, 2011 ; Aston *et al.*, 2012 ; Jenkins et Carrell, 2012). De plus, des études d'ICSI chez la souris avec des spermes à l'ADN fragmenté ont prouvé des effets à long terme sur la descendance avec l'apparition de différentes pathologies chroniques (cancers) et des troubles du comportement (Fernandez-Gonzalez *et al.*, 2008). Ainsi, au-delà du risque de l'impact sur la fertilité, ces données doivent être aussi prises en compte pour établir le risque pour la descendance de développer certaines pathologies en lien avec les défauts épigénétiques de l'ADN spermatique (Balasch et Gratacos, 2011).

Enfin, il existe un lien étroit entre l'environnement et l'épigénome. En effet, certaines modifications épigénétiques sont directement induites par l'environnement au sens large (âge, alimentation, tabagisme, stress...) (Cortesis *et al.*, 2012). Ainsi, ce lien évident entre l'environnement et les processus épigénétiques permettrait d'expliquer comment certains facteurs environnementaux peuvent influer négativement sur la qualité spermatique.

2.6. MicroARN: vers de nouvelles pistes.

Les analyses de transcriptomiques sur les cellules germinales masculines ont mis en évidence la richesse extraordinaire de leur contenu en différents types ARN et notamment en ARN non codants comme les microARN ou piARN. Les micro-ARNs sont d'importants régulateurs de l'expression des gènes et agissent principalement en modulant la stabilité ou la

traduction de différents ARNm cibles. Plusieurs études chez des souris mutantes ont démontré l'importance des miARN dans la spermatogénèse (Kotaja, 2014). Un exemple récent démontre que la délétion des locus miR34b/c et miR-449 chez la souris est responsable de défauts de la méiose et de la spermogénèse entraînant une oligoasthenozoospermie (Comazzetto *et al.*, 2014).

De nombreux ARN sont également retrouvés dans le spermatozoïde mature. Toutefois, le rôle de ces ARNs dans le spermatozoïde reste mal compris, le spermatozoïde étant transcriptionnellement inactif. Il semble que ces ARNs conservent une importance fondamentale pour le développement du zygote et la transmission de certaines marques épigénétiques. L'étude de ces ARNs (ARNm et miARN) dans les spermatozoïdes de plusieurs patients infertiles avec différents phénotypes (azoospermie, asthénozoospermie, tétratozoospermie) a démontré des modifications importantes du profil transcriptionnel. A défaut d'être la cause formelle de l'infertilité chez ces patients, le profil d'expression de ces ARNs peut constituer un « biomarqueurs de la fertilité » intéressant (Jodar *et al.*, 2013).

OBJECTIFS DE LA THESE

Malgré l'identification régulière de nouveaux gènes depuis plusieurs années, la très grande majorité des cas d'infertilité masculine reste encore inexplicable et sans cause génétique connue. Ce faible taux de succès peut s'expliquer en partie par une grande hétérogénéité génétique pour de nombreux phénotypes. Cette hétérogénéité est consistante avec la complexité et le grand nombre de protéines impliquées dans la spermatogénèse (Wistuba *et al.*, 2007). Il est estimé qu'entre 1500 et 2000 gènes sont impliqués dans le contrôle de la spermatogénèse parmi lesquels 300 à 600 sont spécifiquement exprimés dans les cellules germinales masculines, on s'attend logiquement à ce que des anomalies génétiques portant sur ces gènes perturbent la fertilité masculine (Matzuk et Lamb, 2008). Des processus épigénétiques sont également en cause et représentent une part croissante des causes d'infertilité (Dada *et al.*, 2012). De plus, l'infertilité masculine revêt souvent un caractère multifactoriel avec l'implication de plusieurs facteurs associés, notamment environnementaux (Oliva *et al.*, 2001). Cette réalité semble néanmoins varier en fonction du phénotype concerné. Certains phénotypes comme l'oligo-astheno-teratozoospermie (OAT) auraient des causes le plus souvent multifactorielles avec une composante génétique moindre que dans d'autres phénotypes (Cavallini, 2006). On s'attend en particulier à ce que les causes génétiques soient prépondérantes dans les atteintes les plus graves. C'est le cas notamment dans les tératozoospermies sévères regroupant le plus souvent des phénotypes purs et homogènes et dont le défaut en cause a probablement une origine uniquement génétique, monofactorielle et monogénique.

L'abondance de gènes potentiellement candidats rend l'identification des mutations responsables difficile et complexe. Malgré cela, la découverte de nouveaux gènes impliqués dans l'infertilité masculine demeure primordiale pour une meilleure compréhension de sa physiopathologie et l'identification des acteurs intervenant aux différentes étapes de la spermatogénèse. Ces données peuvent permettre d'aborder de nouvelles stratégies de prise en charge diagnostique et thérapeutique des patients infertiles. En effet, il existe un réel besoin de nouvelles stratégies thérapeutiques car les techniques d'aide médicale à la procréation (AMP), qui permettent une fécondation avec un sperme déficient soulèvent le problème de la transmission du facteur génétique causal à la descendance ainsi que le risque d'initier des grossesses avec des gamètes présentant des anomalies épigénétiques qui pourraient augmenter les risques pathologiques à long terme.

C'est dans ce contexte que s'intègre mon travail de thèse dont les objectifs ont été d'approfondir les connaissances des mécanismes génétiques et moléculaires à l'origine de l'infertilité masculine et en particulier de la tératozoospermie. Ce travail permettra, nous

l'espérons, d'envisager de nouvelles perspectives pour l'amélioration de la prise en charge diagnostique et thérapeutique des patients infertiles.

PHÉNOTYPE 1:

**LES SPERMATOZOÏDES
MACROCEPHALE**

Article 1

Identification of a new recurrent aurora kinase C mutation in both European and African men with macrozoospermia.

Ben Khelifa M*, Coutton C*, Blum MG, Abada F, Harbuz R, Zouari R, Guichet A, May-Panloup P, Mitchell V, Rollet J, Triki C, Merdassi G, Vialard F, Koscinski I, Viville S, Keskes L, Soulie JP, Rives N, Dorphin B, Lestrade F, Hesters L, Poirot C, Benzacken B, Jouk PS, Satre V, Hennebicq S, Arnoult C, Lunardi J, Ray PF

* Co-premiers auteurs

Human Reproduction, 27:3337-46, November 2012

Contexte et objectifs

Le premier phénotype d'infertilité masculine étudié dans ce travail est la macrozoospermie. Ce phénotype a été décrit pour la première fois en 1977 (Nistal *et al.*, 1977) et est caractérisé par une tératozoospermie avec 100 % de spermatozoïdes présentant un volume de la tête augmenté par un facteur de 3 à 4 et plusieurs flagelles (Escalier, 1983).

En 2007, notre équipe a identifié une mutation récurrente (c.144delC) dans le gène de l'Aurora kinase C (*AURKC*) chez 14 patients d'origine Nord-Africaine avec 100 % de spermatozoïdes macrocéphales (Dieterich *et al.*, 2007). Par la suite, notre équipe a pu objectiver par cytométrie en flux la présence d'une tétraploïdie constante dans les spermatozoïdes macrocéphales (Dieterich *et al.*, 2009) démontrant que la mutation homozygote c.144delC du gène *AURKC* entraînait un blocage méiotique responsable du phénotype. En 2011, une nouvelle mutation associée à la mutation récurrente c.144delC et affectant le site accepteur d'épissage de l'exon 5 du gène *AURKC* (c.436-2A>G) a pu être identifiée chez deux frères hétérozygotes composites (Ben Khelifa *et al.*, 2011). Cependant, ces nouvelles mutations restent exceptionnelles puisque seules 2 mutations différentes de la mutation c.144delC ont pu être alors mises en évidence. La prévalence de la mutation c.144delC dans la population générale Nord-Africaine est étonnamment élevée et est estimée à 1/50 soit une fréquence de la maladie dans la population maghrébine d'environ 1/10000. Inversement, cette mutation n'a jamais été retrouvée chez des patients d'origine européenne (Harbuz *et al.*, 2009).

Dans ce contexte, l'objectif de mon travail sur ce phénotype a été d'étudier 87 patients d'origine Nord-Africaine et Européenne présentant une macrozoospermie afin de rechercher de nouvelles mutations dans le gène *AURKC*, d'identifier s'il existe une distribution géographique spécifique de ces mutations, d'établir une corrélation entre le génotype et les paramètres spermatiques observés et de proposer une stratégie diagnostique et de prise en charge adaptée aux données génétiques.

Identification of a new recurrent Aurora kinase C mutation in both European and African men with macrozoospermia

Mariem Ben Khelifa^{1,2,3,4,†}, Charles Coutton^{1,2,5,†}, Michael G.B. Blum⁶, Farid Abada^{1,2,3}, Radu Harbuz^{1,2,3}, Raoudha Zouari⁷, Agnès Guichet⁸, Pascale May-Panloup⁹, Valérie Mitchell¹⁰, Jacques Rollet¹¹, Chema Triki¹², Ghaya Merdassi¹³, François Vialard^{14,15}, Isabelle Koscinski¹⁶, Stéphane Viville^{17,18}, Leila Keskes¹⁹, Jean Pierre Soulie²⁰, Nathalie Rives²¹, Béatrice Dorphin²², Florence Lestrade²³, Laeticia Hesters²⁴, Catherine Poirot²⁵, Brigitte Benzacken²⁶, Pierre-Simon Jouk^{1,2}, Véronique Satre^{1,2,5}, Sylviane Hennebicq^{1,2,5}, Christophe Arnoult^{1,2}, Joël Lunardi^{2,3}, and Pierre F. Ray^{1,2,3,*}

¹Laboratoire AGIM, CNRS FRE3405, Equipe 'Génétique, Infertilité et Thérapeutiques', La Tronche F-38700, France ²Université Joseph Fourier, Grenoble F-38000, France ³Pôle de Biologie, DBTP, UM de Biochimie et Génétique Moléculaire, CHU de Grenoble, Grenoble F-38000, France

⁴Laboratoire de Génomique Biomédicale et Oncogénétique LRI/IPT05, Institut Pasteur of Tunis, Tunis, Tunisia ⁵Département de Génétique et Procréation, CHU de Grenoble, cedex 9, Grenoble F-38043, France ⁶Centre National de la Recherche Scientifique, Laboratoire TIMC-IMAG UMR 5525, Université Joseph Fourier, Grenoble 38041, France ⁷Clinique de la Reproduction les Jasmins, 23, Av. Louis BRAILLE, 1002, Tunis, Tunisia ⁸CHU d'Angers, Service de génétique médicale, France ⁹Laboratoire de Biologie de Reproduction, Centre Hospitalier Universitaire, Angers cedex 9 49933, France ¹⁰Institut de biologie de la reproduction – Spermiologie – CECOS, EA 4008 CHRU-59035, Lille Cedex, France

¹¹Clinique du Val d'Ouest, Écully, France ¹²CMRDP, 5, rue Ibn Hazem, 1002, Tunis Beledère, Tunisia ¹³Unité de Procréation Médicalement Assistée, Hôpital Aziza Othmana, Tunis, Tunisia ¹⁴Department of Reproductive Biology, Centre Hospitalier Poissy Saint Germain, Poissy, France ¹⁵EA2493, UVSQ, Versailles, France ¹⁶Service de Biologie de la Reproduction, Centre Hospitalier Universitaire, Strasbourg F 67000 France

¹⁷Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de Santé et de Recherche Médicale (INSERM) U964/Centre National de Recherche Scientifique (CNRS) UMR 1704/Université de Strasbourg, Illkirch 67404, France ¹⁸Centre Hospitalier Universitaire, Strasbourg F 67000 France ¹⁹Laboratoire d'Histologie, Faculté de Médecine de Sfax, Sfax, Tunisia ²⁰Centre de FIV Saint Roch, Montpellier, France ²¹Laboratoire de Biologie de la Reproduction – CECOS, CHU – Hôpitaux de Rouen, EA 4308 'Gaméto-génèse et Qualité du gamète', IRIB, Université de Rouen, Rouen, France ²²Centre d'AMP 74, CHI d'Annemasse-Bonnevillle, Bonneville, France ²³Centre d'AMP Sainte Croix, Metz, France ²⁴Service de gynécologie-obstétrique et médecine de la reproduction, AP-HP, Hôpital Antoine-Béclère, Clamart, France ²⁵UF de Biologie de la Reproduction, Hôpital Pitié Salpêtrière, APHP, Paris, France ²⁶Histologie Embryologie Cytogénétique CECOS, CHU Jean Verdier, APHP, Inserm, U676, Paris 75019, France

*Correspondence address. UM de Biochimie et Génétique Moléculaire, Institut de Biologie et Pathologie, CHU de Grenoble, 38 043 Grenoble cedex 9, France. Tel: +33-476-765-573; Fax: +33-476-765-837; E-mail: pray@chu-grenoble.fr

Submitted on May 10, 2012; resubmitted on June 20, 2012; accepted on July 11, 2012

STUDY QUESTION: Can we identify new sequence variants in the aurora kinase C gene (*AURKC*) of patients with macrozoospermia and establish a genotype–phenotype correlation?

SUMMARY ANSWER: We identified a new non-sense mutation, p.Y248*, that represents 13% of all mutant alleles. There was no difference in the phenotype of individuals carrying this new mutation versus the initially described and main mutation c.144delC.

† These authors contributed equally to this work.

WHAT IS KNOWN ALREADY: The absence of a functional AURKC gene causes primary infertility in men by blocking the first meiotic division and leading to the production of tetraploid large-headed spermatozoa. We previously demonstrated that most affected men were of North African origin and carried a homozygous truncating mutation (c.144delC).

STUDY DESIGN, SIZE, DURATION: This is a retrospective study carried out on patients consulting for infertility and described as having >5% large-headed spermatozoa. A total of 87 patients are presented here, 43 patients were published previously and 44 are new patients recruited between January 2008 and December 2011.

PARTICIPANTS/MATERIALS, SETTING, METHODS: All patients consulted for primary infertility in fertility clinics in France ($n = 44$), Tunisia ($n = 30$), Morocco ($n = 9$) or Algeria ($n = 4$). Sperm analysis was carried out in the recruiting fertility clinics and all molecular analyses were performed at Grenoble teaching hospital. DNA was extracted from blood or saliva and the seven AURKC exons were sequenced. RT-PCR was carried out on transcripts extracted from leukocytes from one patient homozygous for p.Y248*. Microsatellite analysis was performed on all p.Y248* patients to evaluate the age of this new mutation.

MAIN RESULTS AND THE ROLE OF CHANCE: We identified a new non-sense mutation, p.Y248*, in 10 unrelated individuals of European ($n = 4$) and North African origin ($n = 6$). We show that this new variant represents 13% of all mutant alleles and that the initially described c.144delC variant accounts for almost all of the remaining mutated alleles (85.5%). No mutated transcripts could be detected by RT-PCR suggesting a specific degradation of the mutant transcripts by non-sense mediated mRNA decay. A rare variant located in the 3' untranslated region was found to strictly co-segregate with p.Y248*, demonstrating a founding effect. Microsatellite analysis confirmed this linkage and allowed us to estimate a mutational age of between 925 and 1325 years, predating the c.144delC variant predicted by the same method to have arisen 250–650 years ago. Patients with no identified AURKC mutation ($n = 15$) have significantly improved parameters in terms of vitality and concentration of normal spermatozoa, and a decreased rate of spermatozoa with a large head and multiple flagella ($P < 0.001$).

LIMITATIONS, REASONS FOR CAUTION: Despite adherence to the World Health Organization guidelines, large variations in most characteristic sperm parameters were observed, even for patients with the same homozygous mutation. We believe that is mainly related to inter-laboratory variability in sperm parameter scoring. This prevented us from establishing clear-cut values to indicate a need for molecular analysis of patients with macrozoospermia.

WIDER IMPLICATIONS OF THE FINDINGS: This study confirms yet again the importance of AURKC mutations in the aetiology of macrozoospermia. Although a large majority of patients are of North African origin, we have now identified European patients carrying a new non-sense mutation indicating that a diagnosis of absence of a functional AURKC gene should not be ruled out for non-Magrebian individuals. Indirect evidence indicates that AURKC might be playing a role in the meiotic spindle assembly checkpoint (SAC) during meiosis. We postulate that heterozygous men might have a more relaxed SAC leading to a more abundant sperm production and a reproductive advantage. This could be the reason for the rapid accumulation of the two AURKC mutations we observe in North African individuals.

STUDY FUNDING/COMPETING INTEREST(S): None of the authors have any competing interest. This work is part of the project 'Identification and Characterization of Genes Involved in Infertility (ICG2I)' funded by the programme GENOPAT 2009 from the French Research Agency (ANR).

Key words: male infertility / spermatogenesis / gene mutations / genetic diagnosis / aurora kinase C

Introduction

Men presenting with a primary infertility characterized by the presence of the ejaculate of 100% morphologically abnormal spermatozoa with a majority of large-headed, multi-tailed gametes have been reported in the scientific literature for over three decades (Nistal et al., 1977; German et al., 1981; Escalier, 1983; Escalier, 2002). This sperm abnormality is listed in Online Mendelian Inheritance in Man (#243060) and referred to as 'male infertility with large headed multiflagellar spermatozoa', 'macronuclear spermatozoa' or 'macrozoospermia'. Studies of Feulgen-stained preparations (German et al., 1981), spermatocyte C-banding (Pieters et al., 1998) and fluorescence *in situ* hybridization (FISH) analysis (In't Veld et al., 1997; Pieters et al., 1998; Viville et al., 2000; Benzacken et al., 2001; Devillard et al., 2002; Guthäuser et al., 2006; Mateu et al., 2006; Perrin et al., 2008; Chelli et al., 2010) soon demonstrated that these spermatozoa had a large excess of genetic material, suggestive of a meiotic dysfunction. The description of affected

siblings and the high incidence of cases from consanguineous couples were suggestive of a genetic cause with an autosomal recessive inheritance. We therefore carried out a whole genome scan on a small series of North African patients which allowed us to identify a shared region of homozygosity in 19q13 which was present in a large majority of the tested patients. The Aurora Kinase C gene (AURKC), reported to be expressed preferentially in the testis and involved in chromosomal segregation and cytokinesis, was localised in the centre of this candidate region and thus appeared as a good candidate. The same homozygous mutation, c.144delC, was identified in all 14 patients analysed, suggesting a role for AURKC in genesis of the macrozoospermia phenotype (Dieterich et al., 2007). We subsequently analysed 18 additional patients with a pure phenotype: 17 were c.144delC homozygous and one was a compound heterozygote with the recurrent mutation and p.C229Y, a new missense mutation (Dieterich et al., 2009). A heterozygous splicing mutation was also identified in two affected brothers who also carried the c.144delC mutation (Ben Khelifa, et al., 2011). Previous

FISH studies had highlighted the presence of a large chromosomal excess in these abnormal spermatozoa with a heterogeneous population of haploid, diploid and tetraploid gametes. Analysis of the DNA content of four patients by flow cytometry after propidium iodine staining demonstrated that all analysed spermatozoa were in fact tetraploid, indicating that meiosis was blocked before the completion of the first division. These results, discordant with data from the previous FISH studies, suggested that FISH analyses on such chromosomally abnormal gametes led to a large underestimate of the number of chromosomes/chromatids present in the analysed spermatozoa, probably owing to the frequent overlapping of signals (Dieterich *et al.*, 2009).

Aurora kinases A and B (AURKA, B) are ubiquitous cell cycle regulatory serine/threonine kinases which are essential to the successful execution of mitotic cell division by ensuring the formation of a bipolar spindle prior to chromosome segregation (Bischoff *et al.*, 2002). AURKC shares a high amino-acid sequence identity with AURKB but it is expressed predominantly in male germ cells (Bernard *et al.*, 1998; Tang *et al.*, 2001; Tang *et al.*, 2006). It has been demonstrated that AURKC and AURKB share the same substrates as both phosphorylate *in vitro* the centromeric histone Centromere Protein-A (CENP-A) and Borealin (Slattery *et al.*, 2008). An abnormal cell division was observed *in vitro* upon the depletion of AURKB or overexpression of AURKB and AURKC mutant proteins (Tatsuka *et al.*, 1998; Honda *et al.*, 2003). In each case large multinucleated cells accumulated, reminiscent of the large-headed spermatozoa observed in macrozoospermia. AURKC could rescue the AURKB-silenced multinucleation phenotype, suggesting that its function can overlap with and complement AURKB during mitosis (Sasai *et al.*, 2004). Interestingly, AURKC has also been described as being highly expressed in early human preimplantation embryos (Avo Santos *et al.*, 2011). It has been suggested that AURKC is likely involved in chromosome segregation in the first few embryonic divisions and speculated that it could be related to the high aneuploidy rate observed in preimplantation embryos (Avo Santos *et al.*, 2011). The importance of *Aurkc* during preimplantation development has been confirmed in mice as it was demonstrated that *Aurkc* endogenous expression alone in *Aurkb*-deficient mice could sustain preimplantation development up to the late blastocyst stage (Fernandez-Miranda *et al.*, 2011). Furthermore, the authors demonstrated that the timing of the developmental arrest of the *Aurkb* knock-out mice coincides with the switching off of *Aurkc* suggesting again that *Aurkc* could replace *Aurkb* during these first few meiotic divisions (Fernandez-Miranda *et al.*, 2011). The opposite is, however, not true as *Aurkb* is expressed in male germ cells (Tang *et al.*, 2006) but the absence of *Aurkc* leads to macrozoospermia by provoking meiosis I arrest (Dieterich *et al.*, 2007; Dieterich *et al.*, 2009). Interestingly, the specificity of AURKC in male and female meiosis seems to be reversed in man and mouse. Homozygous male knockout mice produce abnormal spermatozoa but are fertile and produce litters of reduced size, whereas all homozygous mutated men are sterile (Kimmens *et al.*, 2007). Conversely, women with homozygous AURKC mutations are fertile (Dieterich *et al.*, 2009), whereas the repression of *Aurkc* in mouse oocytes was shown to cause cytokinesis failure in meiosis I, resulting in the production of large polyploid oocytes, a pattern similar to AURKC-deficient human spermatozoa (Yang *et al.*, 2010).

Here we present the genetic analysis of a large cohort of macrozoospermic patients. We identified a second ancestral mutation

presenting with a geographic range of Europe to North Africa. We assessed the effect of this mutation at the mRNA level. In the light of AURKC function in spermatogenesis, we discuss the possible mechanisms that might have contributed to the spreading and accumulation of these two deleterious mutations over time.

Materials and Methods

Patient information

Forty-four patients were recruited between January 2008 and December 2011. Patients were included when they were described to have >5% 'large-headed' spermatozoa.

To obtain meaningful epidemiological data we also included data from two of our previous publications: 41 patients (out of 62 patients described) with >5% large-headed spermatozoa who were presented in Dieterich *et al.* (2009) and the two patients who were described in Ben Khelifa *et al.* (2011).

All 87 patients consulted for primary infertility in fertility clinics in France ($n = 44$), Tunisia ($n = 30$), Morocco (9) or Algeria (4). A large majority of the French patients originated from North Africa (30/44), while the other 14 French patients had no known ascendants from North Africa. The patients were unrelated apart from four sets of two brothers recruited in France ($n = 1$ set), Algeria ($n = 1$) and Tunisia ($n = 2$). A total of 83 index cases (166 alleles) are therefore presented in Fig. 1. None of the patients had chromosomal abnormalities detected by karyotype analysis.

Control DNAs were extracted from blood of anonymous French donors originating from North Africa ($n = 100$) and Europe ($n = 100$). All patients, family members and anonymous donors gave their written informed consent, and all national laws and regulations were respected. Ethical approval was obtained from Grenoble CHU review board.

Sperm analysis

Sperm analysis was carried out in the source laboratories during the course of the routine biological examination of the patient, according to World Health Organization (WHO, 1999) guidelines. Small variations in protocol might occur between the different laboratories.

Molecular analysis

DNA extraction

DNA was extracted from blood or saliva. Blood DNA extraction was carried out from 5 to 10 ml of frozen EDTA blood using the quick guanidium chloride extraction procedure (Jeanpierre, 1987). Saliva was collected with a Oragene DNA Self-Collection Kit (DNAgentech, Canada) and DNA extraction was performed following the manufacturer's recommendations.

RNA extraction

Nucleated cells were isolated from whole blood using ficoll® 400 (Sigma-Aldrich Corp., St. Louis, MO, USA) following the manufacturer's protocol. RNA extraction was carried out on the isolated white blood cells using Macherey Nagel NucleoSpin® RNA II columns (Macherey Nagel, Hoerdt, France) using the manufacturer's protocol.

Mutation detection

The seven AURKC exons and intron/exon boundaries were amplified by PCR and sequenced, as described previously (Dieterich *et al.*, 2007). All analyses were carried out using the BigDye Terminator v3.1 sequencing kit and an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA).

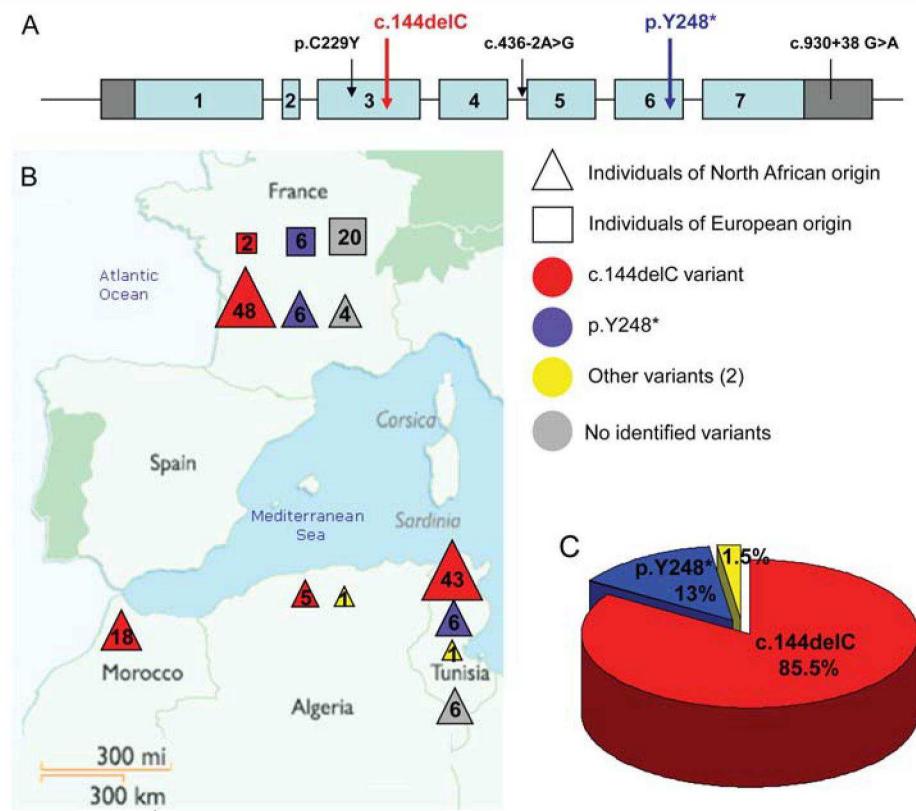


Figure 1 Schematic representation of the aurora kinase C (AURKC) gene (A); country of patient's recruitment and patient's ethnico-geographic origin (B) and overall frequency of the identified mutations (C). (A) Exons are indicated as a light blue rectangle, untranslated regions as a grey rectangle, the localisation of the identified mutations are shown by an arrow and the p.Y248* associated polymorphism located in the 3' untranslated region (UTR) is indicated by a simple line. (B) The position of the boxes on the map indicates the country where the patients were recruited, the shape of box indicates the ethnico-geographic origin of the patients (European or North African) and numbers indicate the number of alleles identified in index cases (excluding brothers). The number of recruited patients reflects the recruiting activity of participating centres and can in no way give an indication of the frequency of the phenotype in the different countries. (C) Overall frequency of the identified mutations in positive index cases (n = 83).

RT-PCR

Reverse transcription was carried out with 5 µl of extracted RNA (~500 ng). Hybridization of the oligo dT was performed by incubating for 5 min at 65°C and quenching on ice with the following mix: 5 µl of RNA, 3 µl of poly T oligo primers (dT)12–18 (10 mM, Pharmacia), 3 µl of the four dNTPs (0.5 mM, Roche diagnostics) and 2.2 µl of H₂O. Reverse Transcription then was carried out for 30 min at 55°C after the addition of 4 µl of 5× buffer, 0.5 µl RNase inhibitor and 0.5 µl of Transcripter Reverse transcriptase (Roche Diagnostics). Two microlitres of the obtained cDNA mix was used for the subsequent PCR. AURKC 5' primer was located on exon 4 (5' CAATATCCTGCCCTGTATACT 3') and the 3' primer on exon 6 (5' TCATTCTGGCGGAAGT 3'). Two microlitres of the reversed transcribed RNA was amplified with these primers (40 cycles) at an elongation temperature of 58°C. Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) primers were used as control. Thirty-five cycles of PCR amplification were carried out at an elongation temperature of 60°C with the following GAPDH primers (5' to 3'): GAGTCAACGGATTGGTCGT and TTGATTTGGAGGGATCTCG.

High-resolution melting

High-resolution melting (HRM) analysis was performed with the LightCycler 480 (Roche), using the LightCycler 480 High-Resolution

Melting Master kit as described in Harbuz et al. (2010). Results were analysed with the Gene Scanning software (Roche).*

Statistical analysis

The different sperm parameters (Table I) were compared in different groups using a two-tailed t-test. A value of P < 0.05 was considered significant. Analyses were carried out with the GraphPad software.

Results

Here we present new sequence data from 44 patients with macrozoospermia and analyse these together with our previous data from an additional 43 patients.

AURKC sequencing and HRM analysis

In a previous study (Dieterich et al., 2009) we reported that 31/41 patients presenting with >5% large-headed spermatozoa were c.144delC homozygous, one was a compound heterozygote carrying c.144delC with a false-sense variant: p.C229Y and 9 (classified as atypical) were non-deleted. No variants had been detected in

19 patients with <5% large-headed spermatozoa (Dieterich *et al.*, 2009), as they do not fulfil the criteria for this study (i.e. >5% large-headed spermatozoa) these 19 patients are not included here. In a subsequent work we described two brothers who were compound

heterozygotes carrying c.144delC with a splicing variant: c.436–2A>G (Ben Khelifa *et al.*, 2011). In these two studies we therefore had recruited and analysed a total of 43 patients with macrozoospermia, of whom 34 carried two AURKC mutated alleles.

A total of 44 patients were recruited between January 2008 and December 2011 who we had not described previously. The AURKC exon 3 was sequenced in all 44 patients. The remaining 6 exons and intron boundaries were sequenced for all who did not present a c.144delC homozygous mutation. No mutations were identified in 6 patients and the remaining 38 carried two mutations. A new non-sense mutation: p.Y248*; c.744C>G (NM_001015878) was identified in 11 patients (Fig. 2). Overall 27 men were c.144delC homozygous, 9 were p.Y248* homozygous (including 2 brothers) and 2 were compound heterozygotes for these two mutations. An additional variant located in AURKC 3'UTR (c.930+38G>A) was homozygous in all p.Y248* homozygous patients and heterozygous in the two compound heterozygotes (Fig. 2).

When all patients are pooled, we have a total of 72 patients with a mutation out of 87 (83%) or 68 out of 83 probands (82%). The c.144delC variant represents 85.5% of the mutated alleles (taking into account only one proband in familial cases), the p.Y248* represents 13% of the alleles and two familial mutations account for the remaining 1.5% (Fig. 1). Probands were recruited equally in France (43) and in North Africa (40). All patients recruited in North Africa originated from North Africa and the majority of the French patients

Table 1 Semen parameters of patients according to aurora kinase C (AURKC) genotype.

Patients	AURKC mutation (n = 72)	No mutation (n = 15)
Sperm volume (ml)	2.5 (0.1–5.5)	3.7 (1.8–8.3)
Sperm concentration (10 ⁶ per ml)	13.1 (0.01–98.7)	11.3 (0.01–50)
Round cells (10 ⁶ per ml)	4.9 (0–27.95)	5.1 (0–28.8)
Motility A + B, 1 h	19.5 (0–70)	15 (0–40)
Sperm vitality	24.0 (0–66)*	56.7 (10–84)
Normal spermatozoa	0.1 (0–1)*	5.1 (0–19)
Large-headed spermatozoa	79.6 (34–100)*	35.2 (5–75)
Multiflagellar spermatozoa	37.9 (7–100)*	7.8 (0–28)
Multiple anomalies index	3.0 (2.95–4.1)	2.6 (1.36–3.2)

Values are expressed as the mean with the lower and higher values between brackets. Values are expressed as a percentage, unless specified otherwise.

*Significant P < 0.001.

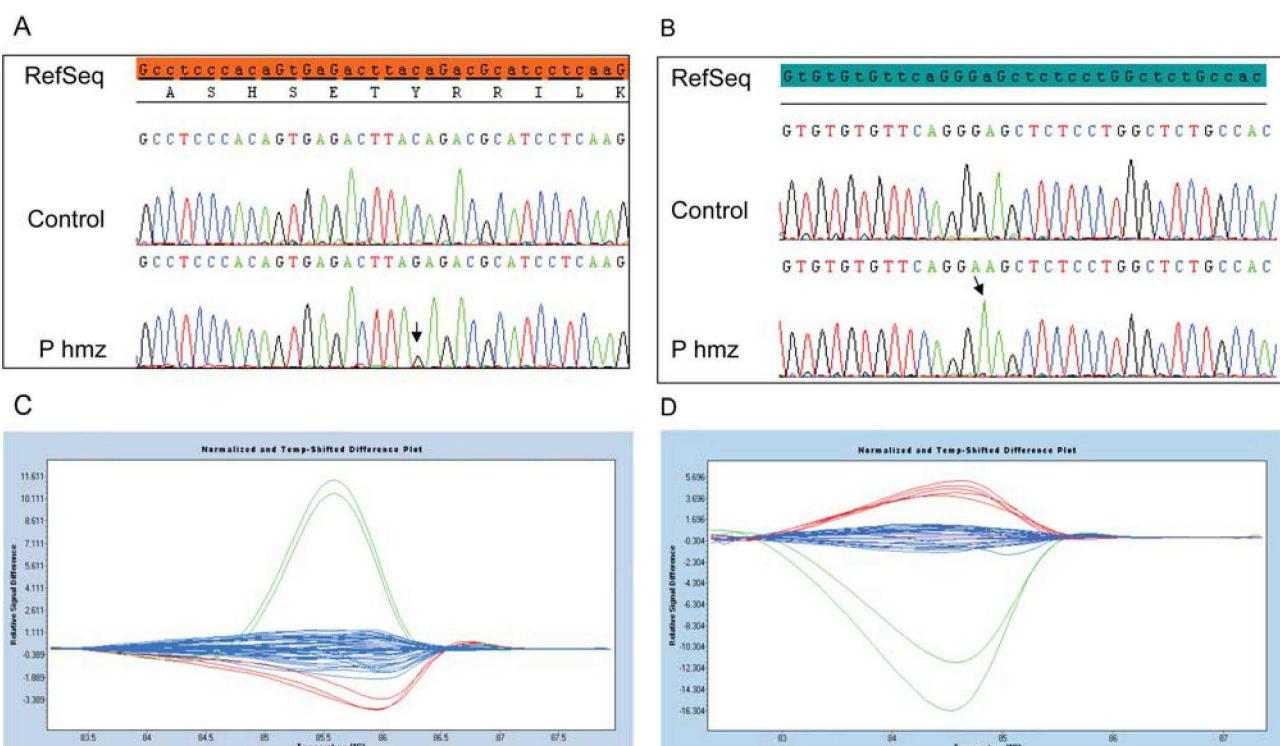


Figure 2 Electropherogram and HRM analysis of AURKC exons 6 and 7 with the 3'UTR. **(A)** Electropherogram showing part of AURKC exon 6 for a wild-type subject (Control) and a p.Y248* homozygous patient (P hmz). **(B)** Electropherogram showing part of AURKC 3' UTR from the same individuals showing the presence (or absence) of the c.930+38G>A variant. **(C)** and **(D)** HRM analysis of AURKC exon 6 and 7 (with 3'UTR) respectively, of control subjects (blue), heterozygous (green) and homozygous (red) patients.

were also of North African origin (67%). The frequency of positive diagnoses was much higher in probands of North African origin (93%) compared with probands of European origin (29%) (Fig. 1).

The p.Y248* was identified in six homozygous probands of North African origin and in four probands of European origin including two homozygotes and two compound heterozygotes who also carried the c.144delC variant. These individuals had no known relative of North African descent. These two compound heterozygotes were the only two patients of European origin identified as carrying the c.144delC variant.

To exclude the possibility that the identified p.Y248* and its associated c.930+38G>A variant may be common in the studied populations, we performed an HRM of AURKC exon 6 and 7 in 100 individuals of French origin and 100 individuals of North African origin. Homozygous and heterozygous patients were passed in triplicate and are shown in red and green, respectively, in Fig. 2. Each profile (homozygous, heterozygous and non-mutated) was clearly distinct (Fig. 2C and D). There was no abnormal exon 6 profile from any of the 200 tested DNAs, confirming that the p.Y248* mutation is not frequent in either European or North African individuals. HRM of exon 7 allowed us to identify three c.930+8G>A heterozygotes among the European and five among the North African control subjects indicating a frequency of heterozygosity of ~4% or an allelic frequency of 2%. UCSC genome bioinformatics presents a pool of 1097 individuals with a heterozygosity rate of 4.1%, concordant with our findings. These results indicate that c.930+38G>A is a rare variant and exclude a chance association of c.930+38G>A with p.Y248*.

Transcript analysis

As a truncating mutation, p.Y248* is expected to have a severe impact on the protein function. A residual activity could however result from a shortened peptide. We therefore carried out RT-PCR on transcripts extracted from leukocytes from one p.Y248* homozygous patient. Control amplification with a house keeping gene (GAPDH) was positive from the patient and control subjects, whereas AURKC amplification could only be obtained from the controls (Fig. 3). This indicates that AURKC RNA is not present in p.Y248* patients, suggestive of mRNA decay associated with this mutation.

Sperm analysis

The sperm parameters of the patients carrying c.144delC were compared with those carrying p.Y248* (including the two heterozygous individuals). There was no difference between the two groups. Data from both groups were therefore pooled and compared with data from the patients without a mutation (Table I).

All patients with a mutation had a large majority of macrozoosperms (average of 80%) and 37% of the spermatozoa were reported as multiflagellar. There was no difference in sperm volume, spermatozoa, round cells concentration, mobility or multiple anomaly index. There was, however, a significant increase in the proportion of large-headed and multiflagellar spermatozoa in patients with a mutation compared with those with no mutation, and there was a decrease in the frequency of typical spermatozoa and sperm vitality in the individuals with a mutation.

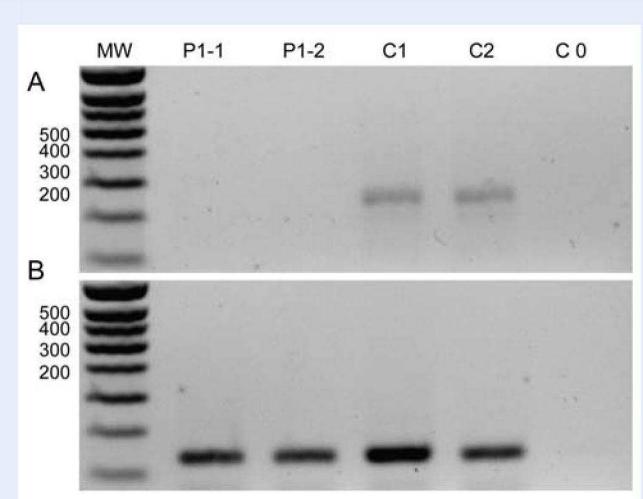


Figure 3 Evidence of AURKC mRNA decay in patients with the p.Y248* variant. RT-PCR analysis of patient P1, (A) Y248* homozygote (P1-1 and P1-2) and control subjects from the general population (C1 and C2). Electrophoresis showing the RT-PCR amplification of AURKC exons 4–6. C1 and C2 controls yield a normal fragment of 329 bp, whereas patient P1 shows no amplification. There is no amplification from the RT-negative control (C0). (B) Electropherogram showing the amplification of the same cDNAs with GAPDH primers. Bands of equivalent intensity are obtained from all samples except the RT-negative control (C0).

Dating of the mutations

To compute the age of p.Y248*, we used the number of ancestral haplotypes at the different microsatellite markers that are linked to the locus of interest (Fig. 4). There are 9, 14, 10 and 5 (among 18) ancestral haplotypes between the locus of interest and D19S210, D19S214, D19S218 and D19S890, respectively. The probability to observe a non-ancestral haplotype is given by $(1 - c)^g$ where g is the allele age in generations and c is the crossover rate (Slatkin and Rannala, 2000). The crossover rate in the region of chromosome 19 between D19S210 and D19S890 is 3.1 cM/MB as provided by the genetic map of Marshfield (Broman et al., 1998). The number of ancestral haplotypes is a binomial random variable with $n = 18$ and $P = (1 - c)^g$. Multiplying the four binomial distributions corresponding to the four microsatellite loci, we calculate a pseudo-likelihood function which is maximal at $g = 37$ generations.

Labuda et al. (1996) have shown that the dating arising from such genetic clock methods is biased downwards because this method does not account for the demographic expansion of the population. To account for an exponential expansion of rate r per generation, the mathematical expression: $(1/r) \ln[c e^r / (e^r - 1)]$ should be added to the estimates provided by genetic clock methods. Assuming that an upper bound for the expansion rate is given by the expansion $r = 0.4$, as was calculated for the Ashkenazy Jewish populations (Risch et al., 1995), we found that the age of the mutation is in the range of 37–53 generations. Taking 25 years for the generation time, the mutation would have then arisen some 925–1325 years ago.

We then estimated the age of the c.144delC mutation using the same method. For that we used the haplotypes of 14 individuals published previously (Dieterich et al., 2007). There are 27, 20 and

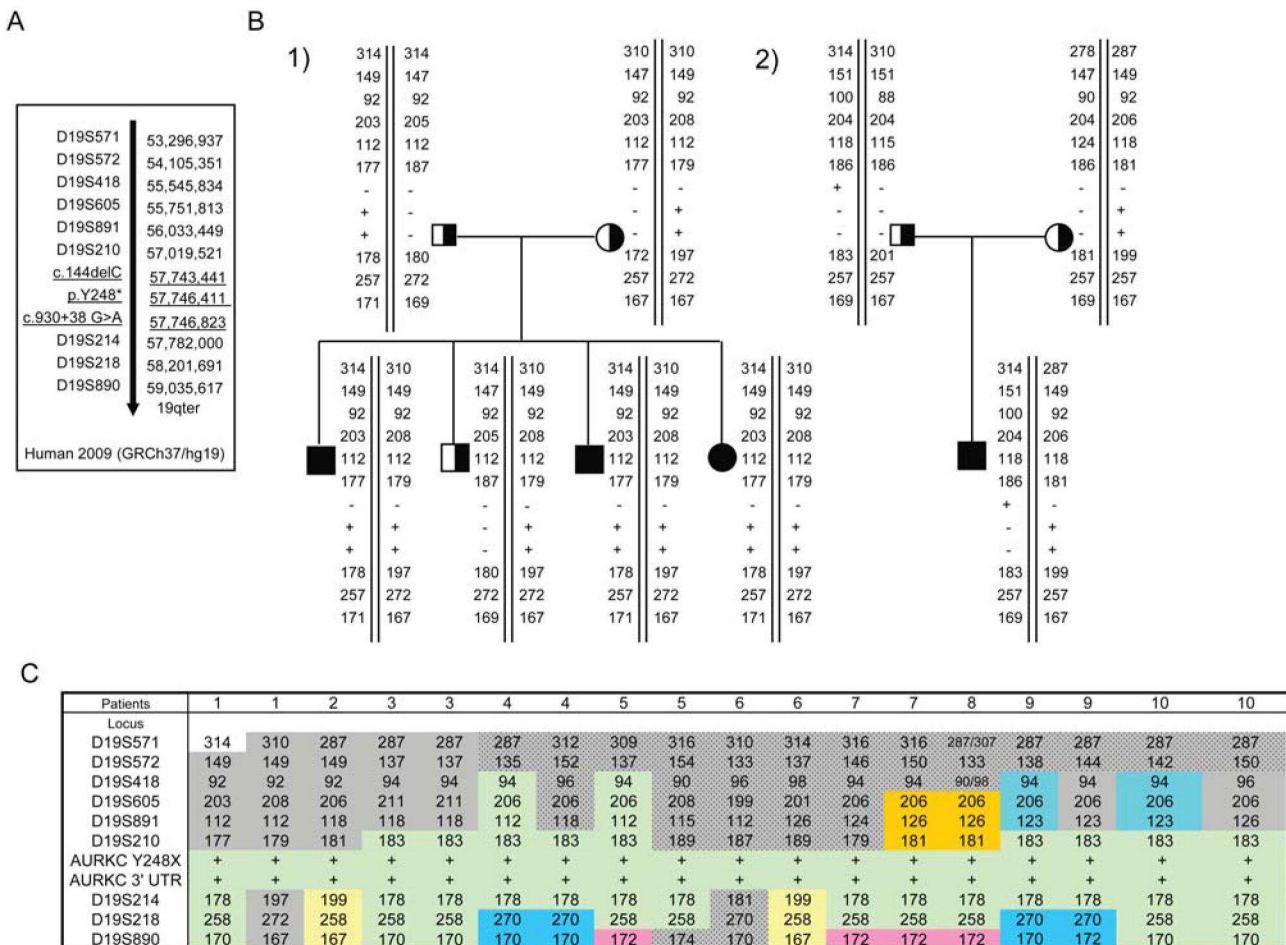


Figure 4 The haplotype of p.Y248*. **(A)** Reference and location of the studied microsatellite markers and AURKC variants. **(B)** Family trees of the first two identified index cases. The two brothers from family 1 are homozygous for the p.Y248* variant, whereas proband from family 2 is a compound heterozygote carrying p.Y248* and c.144delC. **(C)** Haplotypes from 10 unrelated patients carrying the p.Y248* variant. All individuals are homozygous for p.Y248* except for patients 2 (proband from family 2) and 8 who are compound heterozygotes for c.144delC. Each column corresponds to one allele. A familial study could only be performed for patients 1 and 2 (from panel B). For all the other patients, permutations of the allelic values were obtained to try to identify identical haplotypes between patients. The grey area corresponds to regions with no shared haplotypes and the spotted grey area corresponds to regions of allelic random attribution. The green area corresponds to the core-shared haplotype. Other common haplotypes, defined by the presence of at least two adjacent markers in two separate individuals, are indicated with other colours.

15 (among 28) ancestral haplotypes between the locus of interest and D19S214, D19S218 and D19S890, respectively. After using the adjustment of Labuda *et al.* (1996), this results in a mutation occurring 10–26 generations (250–650 years) ago.

Discussion

We presented data from a total of 87 patients (44 new patients and 43 reported previously) presenting with macrozoospermia and have identified a new mutation, p.Y248*, in 10 unrelated individuals of both Maghrebian and European origin. We observed that this variant accounts for 13% of the mutant alleles and that 85% of the remaining alleles bear the initially described c.144delC mutation. We show that p.Y248* is also a founding mutation that largely predates the c.144delC variant. As a nonsense mutation, we anticipated that p.Y248* would have a severe effect on the protein. We wanted to

assess whether p.Y248* transcripts could lead to the production of a truncated protein which could potentially retain some functionality, or if the mutated transcript was subjected to nonsense-mediated mRNA decay (NMD) resulting in the absence of the protein (Maquat, 2004). Although AURKC is preferentially expressed in the testis, it is also weakly expressed in several tissues including the lung, ovary or skeletal muscle (Yan *et al.*, 2005) and we had also noticed that AURKC transcripts were present in blood leukocytes. We carried out RT-PCR on leukocytes from one of the p.Y248* homozygous patients. Contrary to what was observed in a fertile control, we could not detect any AURKC transcript in our patient, demonstrating that the p.Y248* mutation was subjected to NMD, at least in leukocytes (Fig. 3). We had previously observed that the recurrent c.144delC mutated transcript, which contains several premature stop codons, was also subject to NMD (Ben Khelifa *et al.*, 2011). From what we observed in blood we can conclude that both recurrent

mutations are very likely to have the same effect and both lead to a total absence of the protein.

The values measured during the routine spermogram and spermocytogram were compared between patients with c.144delC, Y248* and patients without a mutation (Table I). Not surprisingly, as we demonstrated that both c.144delC and p.Y248* resulted in the absence of the protein, there was no phenotypic difference between these two groups of patients. Patients with no identified AURKC mutation ($n = 15$) have significantly improved parameters in terms of vitality and concentration of normal spermatozoa, and a decrease in spermatozoa with a large head and multiple flagella (Table I). Patients came from different centres and despite a common observance to the WHO guidelines (WHO, 1999) important variations in scoring were observed, even in genotypically identical individuals. In our experience, patients with a mutation never present any typical spermatozoa and have a large majority of large-headed spermatozoa. We believe that inter-laboratory variability in scoring of the most relevant morphological features, such as the concentration of large-headed and multi-flagellar spermatozoa, is responsible for some of the overlap observed between patients with and without mutations. It is perhaps the presence of $>1\%$ normal spermatozoa that is the most discriminant parameter as this was never reported in individuals with a mutation. The absence of normal spermatozoa was observed in a few individuals with no mutation but this was only observed in patients with very severe oligozoospermia who more likely suffered from severe oligoteratozoospermia rather than pure macrozoospermia. As a routine diagnostic strategy in men, we propose the sequencing of AURKC exons 3 and 6 for all patients presenting with a high proportion of large-headed spermatozoa. In the absence of a mutation, the sequencing of the remaining exons could be proposed only to individuals with a sperm concentration >1 million and described as having $<1\%$ normal spermatozoa.

We have estimated here the age of the two AURKC recurrent mutations. We calculated that p.Y248* arose 925–1325 years ago and therefore largely predates c.144delC, dated as having occurred 250–650 years ago. The fact that p.Y248* is more ancient than c.144delC is concordant with the wider geographical coverage of p.Y248* that is found in individuals of both European and North African origin. We had calculated previously that the c.144delC variant was more ancient than reported here (Dieterich et al., 2007). We had used a simplified method and considered that the smallest conserved region spanned from the mutation to D19S214, not taking into account the conserved sequences centromeric to AURKC. We believe that the method used here is likely to give a much better estimate of the age of the mutations.

We had previously measured that 1 in 50 individuals originating from North Africa carried a copy of c.144delC (Dieterich et al., 2009). It seems surprising that a pathological variant having a negative effect on reproduction reached such a high frequency in a relatively short period. It is also surprising to identify a second ancestral founding allele with a wide geographic coverage.

Research carried out on aurora B and C, as summarised in the introduction, demonstrated that AURKC has an essential role in meiotic interkinesis (the equivalent of cytokinesis in mitosis) and also that AURKC participates in the control of microtubule–kinetochore attachment, verifying the bi-orientation of the tensions preceding chromosome segregation, thus ensuring the production of euploid

gametes. During meiosis, germ cells with misaligned chromosomes or abnormal microtubule tension likely to result in missegregation, are blocked in prometaphase. This control, called spindle assembly checkpoint (SAC), is essential in mitosis to limit the production of aneuploid cells, which are likely to transform into malignant cells, and in meiosis to limit the conception of chromosomally abnormal gametes and embryos. There is now an increasing body of evidence indicating that the action of AURKC in microtubule–kinetochore attachment is an integral part of the SAC. Lane et al. (2010) studied the effect of ZM447439, a general inhibitor of aurora kinases on oocyte meiosis and they showed that while a continuous exposure to ZM447439 logically led to a meiosis I block, a short exposure at the prometaphase stage accelerated polar body extrusion. They also demonstrated that this treatment allowed an overriding of the SAC induced by nocodazole treatment, therefore indicating that the accelerated meiosis was linked to a speedier, and thus relaxed, SAC provoked by the inactivation of the aurora kinases. They observed a 15% increase in aneuploid oocytes in the ZM447439-treated group therefore supporting a role for an alteration in the checkpoint mechanism.

The involvement of AURKC in the SAC might explain some intriguing features concerning large-headed spermatozoa. First, to our knowledge, macrozoospermia is the only sperm phenotype associated with meiotic arrest. All other genetic alterations leading to meiotic arrest result in azoospermia (following appropriate activation of the SAC). We can take, for example, chromosome Y microdeletions or the effect of Dmc1 deficiency in mice (Bannister et al., 2007). The macrozoospermia phenotype caused by the absence of AURKC therefore constitutes an *in vivo* confirmation of the predominant role of AURKC in the SAC, as in the absence of AURKC interkinesis is blocked but spermiogenesis is pursued leading to the production of tetraploid flagellated spermatozoa. Second, we are surprised to find two deleterious mutations which have reached a relatively high frequency (c.144delC heterozygous frequency is 1/50 in the North African population (Dieterich et al., 2009)) over a relatively short period of time. We can wonder if the reduced amount of AURKC protein in heterozygous men could have an effect similar to that of ZM447439 in the oocyte: the SAC could be slightly relaxed leading to speedier and thus more abundant spermatogenesis. We therefore would expect heterozygotes to have an increased sperm count, probably accompanied by a small increase in aneuploidy. This could entail a small reproductive advantage that could explain the rapid propagation of AURKC mutations. We intend to verify this theory but we are having difficulties in recruiting heterozygous donors.

Most genetic diseases have manifestations that are dependent on the effect of the mutation(s) on the protein. Cystic fibrosis is a classic example; more than 1500 mutations have been identified so far leading to symptoms ranging from typical manifestation with severe lung and pancreatic impairment to male infertility only, owing to atresia of the vas deferens. In macrozoospermia we have detected only four mutations: the two recurrent mutations described here, a missense mutation identified in only one patient and a splicing variant resulting in loss of an exon identified in two brothers. All patients with a mutation present with the most severe form of the disease: 0% normal spermatozoa with close to 100% large-headed spermatozoa and we have not identified any mutations in patients presenting with a mosaic of normal and large-headed spermatozoa. There can be three explanations for these two observations (I) ‘milder’

mutations are infrequent and we have not yet sampled the right patients, (2) AURKC protein sequence is extremely sensitive to change and all/most mutations give a severe phenotype but this does not explain why so few variants have been identified so far and 3) 'milder' mutations result in a distinct phenotype, thus explaining why we identified so few variants and why we did not identify any AURKC mutations in mosaic macrozoospermia. We favour this last hypothesis as we can imagine that less severe mutations (which would not result in protein loss by mRNA decay) might have an effect on microtubule–kinetochore attachment and/or on interkinesis without blocking the SAC activity of the protein. If that were the case these less severe mutations would lead to azoospermia, a phenotype considered, paradoxically, to be more severe than macrozoospermia. We therefore would not be surprised to identify 'milder' AURKC variants in non-obstructive azoospermia patients. We plan to sequence the exome of a cohort of azoospermia patients, and will be particularly attentive to the presence of AURKC variants that could reinforce this hypothesis.

This study confirms yet again the importance of AURKC mutations in the aetiology of macrozoospermia. Although a large majority of patients are of North African origin, we have now identified European patients carrying mainly a new non-sense mutation indicating that AURKC diagnosis should not be ruled out for non-Magrebian individuals. Indirect evidence concerning AURKC function in meiosis raises the possibility that the rapid accumulation of two AURKC mutations could be the consequence of a reproductive advantage for heterozygous individuals. We also postulate that other AURKC molecular alterations might induce a different and unexpected phenotype, namely azoospermia.

Acknowledgements

This work was part of the ICG2I project funded by the programme GENOPAT 2009 from the French Research Agency (ANR). We thank our patients and control individuals for their participation.

Authors' roles

M.B.K., C.C., F.A and R.H. carried out and interpreted all the molecular work. M.B., P.S.J., C.A. and J.L. and PR realised the analysis and interpretation of data; R.Z., A.G., P.M.P., V.M., J.R., C.T., G.M., F.V., I.K., S.V., L.K., J.P.S., N.R., B.D., F.L., L.H., C.P., B.B. and S.H. supervised all the patient's information and contributed to the acquisition of data. All authors revised the manuscript and approved its final version. P.F.R. designed the overall study, supervised all molecular laboratory work, wrote the main draft of the manuscript, had full access to all of the data and takes responsibility for the integrity of the data and its accuracy.

Funding

This work is part of the project 'Identification and Characterization of Genes Involved in Infertility (ICG2I)' funded by the programme GENOPAT 2009 from the French Research Agency (ANR).

Conflict of interest

None declared.

References

- Avo Santos M, van de Werken C, de Vries M, Jahr H, Vromans MJ, Laven JS, Fauser BC, Kops GJ, Lens SM, Baart EB. A role for Aurora C in the chromosomal passenger complex during human preimplantation embryo development. *Hum Reprod* 2011; **26**:1868–1881.
- Bannister LA, Pezza RJ, Donaldson JR, de Rooij DG, Schimenti KJ, Camerini-Otero RD, Schimenti JC. A dominant, recombination-defective allele of Dmc1 causing male-specific sterility. *PLoS Biol* 2007; **5**:e105.
- Ben Khelifa M, Zouari R, Harbuz R, Halouani L, Arnoult C, Lunardi J, Ray PF. A new AURKC mutation causing macrozoospermia: implications for human spermatogenesis and clinical diagnosis. *Mol Hum Reprod* 2011; **17**:762–768.
- Benzacken B, Gavelle FM, Martin-Pont B, Dupuy O, Lievre N, Hugues JN, Wolf JP. Familial sperm ploidy induced by genetic spermatogenesis failure: case report. *Hum Reprod* 2001; **16**:2646–2651.
- Bernard M, Sanseau P, Henry C, Couturier A, Prigent C. Cloning of STK13, a third human protein kinase related to *Drosophila* aurora and budding yeast *Ipl1* that maps on chromosome 19q13.3-ter. *Genomics* 1998; **53**:406–409.
- Bischoff FZ, Sinacori MK, Dang DD, Marquez-Do D, Horne C, Lewis DE, Simpson JL. Cell-free fetal DNA and intact fetal cells in maternal blood circulation: implications for first and second trimester non-invasive prenatal diagnosis. *Hum Reprod Update* 2002; **8**:493–500.
- Broman KW, Murray JC, Sheffield VC, White RL, Weber JL. Comprehensive human genetic maps: individual and sex-specific variation in recombination. *Am J Hum Genet* 1998; **63**:861–869.
- Chelli MH, Albert M, Ray PF, Guthausen B, Izard V, Hammoud I, Selva J, Vialard F. Can intracytoplasmic morphologically selected sperm injection be used to select normal-sized sperm heads in infertile patients with macrocephalic sperm head syndrome? *Fertil Steril* 2010; **93**:1347 e1–1347 e5.
- Devillard F, Metzler-Guillemain C, Pelletier R, DeRobertis C, Bergues U, Hennebicq S, Guichaoua M, Sele B, Rousseaux S. Polyploidy in large-headed sperm: FISH study of three cases. *Hum Reprod* 2002; **17**:1292–1298.
- Dieterich K, Soto Rifo R, Faure AK, Hennebicq S, Ben Amar B, Zahi M, Perrin J, Martinez D, Sele B, Jouk PS et al. Homozygous mutation of AURKC yields large-headed polyploid spermatozoa and causes male infertility. *Nat Genet* 2007; **39**:661–665.
- Dieterich K, Zouari R, Harbuz R, Vialard F, Martinez D, Bellayou H, Prisant N, Zoghmar A, Guichaoua MR, Koscienski I et al. The Aurora Kinase C c.144delC mutation causes meiosis I arrest in men and is frequent in the North African population. *Hum Mol Genet* 2009; **18**:1301–1309.
- Escalier D. Human spermatozoa with large heads and multiple flagella: a quantitative ultrastructural study of 6 cases. *Biol Cell* 1983; **48**:65–74.
- Escalier D. Genetic approach to male meiotic division deficiency: the human macronuclear spermatozoa. *Mol Hum Reprod* 2002; **8**:1–7.
- Fernandez-Miranda G, Trakala M, Martin J, Escobar B, Gonzalez A, Ghyselinck NB, Ortega S, Canamero M, Perez de Castro I, Malumbres M. Genetic disruption of aurora B uncovers an essential role for aurora C during early mammalian development. *Development* 2011; **138**:2661–2672.

- German J, Rasch EM, Huang CY, MacLeod J, Imperato-McGinley J. Human infertility due to production of multiple-tailed spermatozoa with excessive amounts of DNA. *Am Soc Hum Genet* 1981;33:64A.
- Guthauser B, Vialard F, Dakouane M, Izard V, Albert M, Selva J. Chromosomal analysis of spermatozoa with normal-sized heads in two infertile patients with macrocephalic sperm head syndrome. *Fertil Steril* 2006;85:750 e5–750 e7.
- Harbuz R, Lespinasse J, Boulet S, Francannet C, Creveaux I, Benkhelifa M, Jouk PS, Lunardi J, Ray PF. Identification of new FOXP3 mutations and prenatal diagnosis of IPEX syndrome. *Prenat Diagn* 2010;30:1072–1078.
- Honda R, Korner R, Nigg EA. Exploring the functional interactions between Aurora B, INCENP, and survivin in mitosis. *Mol Biol Cell* 2003;14:3325–3341.
- In't Veld PA, Broekmans FJ, de France HF, Pearson PL, Pieters MH, van Kooij RJ. Intracytoplasmic sperm injection (ICSI) and chromosomally abnormal spermatozoa. *Hum Reprod* 1997;12:752–754.
- Jeanpierre M. A rapid method for the purification of DNA from blood. *Nucleic Acids Res* 1987;15:9611.
- Kimmins S, Crosio C, Kotaja N, Hirayama J, Monaco L, Hoog C, van Duin M, Gossen JA, Sassone-Corsi P. Differential functions of the aurora-B and aurora-C kinases in Mammalian spermatogenesis. *Mol Endocrinol* 2007;21:726–739.
- Labuda M, Labuda D, Korab-Laskowska M, Cole DE, Zietkiewicz E, Weissenbach J, Popowska E, Pronicka E, Root AW, Glorieux FH. Linkage disequilibrium analysis in young populations: pseudo-vitamin D-deficiency rickets and the founder effect in French Canadians. *Am J Hum Genet* 1996;59:633–643.
- Lane SL, Chang HY, Jennings PC, Jones KT. The Aurora kinase inhibitor ZM447439 accelerates first meiosis in mouse oocytes by overriding the spindle assembly checkpoint. *Reproduction* 2010;140:521–530.
- Maquat LE. Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol* 2004;5:89–99.
- Mateu E, Rodrigo L, Prados N, Gil-Salom M, Remohi J, Pellicer A, Rubio C. High incidence of chromosomal abnormalities in large-headed and multiple-tailed spermatozoa. *J Androl* 2006;27:6–10.
- Nistal M, Paniagua R, Herruzo A. Multi-tailed spermatozoa in a case with asthenospermia and teratospermia. *Virchows Arch B Cell Pathol* 1977;26:111–118.
- Perrin A, Morel F, Moy L, Colleu D, Amice V, Braekeleer MD. Study of aneuploidy in large-headed, multiple-tailed spermatozoa: case report and review of the literature. *Fertil Steril* 2008;90:1201.e13–1201.e7.
- Pieters MH, Speed RM, de Boer P, Vreeburg JT, Dohle G, In't Veld PA. Evidence of disturbed meiosis in a man referred for intracytoplasmic sperm injection. *Lancet* 1998;351:957.
- Risch N, de Leon D, Ozelius L, Kramer P, Almasy L, Singer B, Fahn S, Breakefield X, Bressman S. Genetic analysis of idiopathic torsion dystonia in Ashkenazi Jews and their recent descent from a small founder population. *Nat Genet* 1995;9:152–159.
- Sasai K, Katayama H, Stenoien DL, Fujii S, Honda R, Kimura M, Okano Y, Tatsuka M, Suzuki F, Nigg EA et al. Aurora-C kinase is a novel chromosomal passenger protein that can complement Aurora-B kinase function in mitotic cells. *Cell Motil Cytoskeleton* 2004;59:249–263.
- Slatkin M, Rannala B. Estimating allele age. *Annu Rev Genomics Hum Genet* 2000;1:225–249.
- Slattery SD, Moore RV, Brinkley BR, Hall RM. Aurora-C and Aurora-B share phosphorylation and regulation of CENP-A and Borealin during mitosis. *Cell Cycle* 2008;7:787–795.
- Tang CJ, Chuang CK, Hu HM, Tang TK. The zinc finger domain of Tzfp binds to the tbs motif located at the upstream flanking region of the Aiel (aurora-C) kinase gene. *J Biol Chem* 2001;276:19631–19639.
- Tang CJ, Lin CY, Tang TK. Dynamic localization and functional implications of Aurora-C kinase during male mouse meiosis. *Dev Biol* 2006;290:398–410.
- Tatsuka M, Katayama H, Ota T, Tanaka T, Odashima S, Suzuki F, Terada Y. Multinuclearity and increased ploidy caused by overexpression of the aurora- and IplI-like midbody-associated protein mitotic kinase in human cancer cells. *Cancer Res* 1998;58:4811–4816.
- Viville S, Molland R, Bach ML, Falquet C, Gerlinger P, Warter S. Do morphological anomalies reflect chromosomal aneuploidies? case report. *Hum Reprod* 2000;15:2563–2566.
- WHO. World Health Organization WHO laboratory manual for the examination of human semen and sperm–cervical mucus interaction, 1999.
- Yan X, Cao L, Li Q, Wu Y, Zhang H, Saiyin H, Liu X, Zhang X, Shi Q, Yu L. Aurora C is directly associated with Survivin and required for cytokinesis. *Genes Cells* 2005;10:617–626.
- Yang KT, Li SK, Chang CC, Tang CJ, Lin YN, Lee SC, Tang TK. Aurora-C kinase deficiency causes cytokinesis failure in meiosis I and production of large polyploid oocytes in mice. *Mol Biol Cell* 2010;21:2371–2383.

Principaux résultats

L’analyse moléculaire complète des sept exons du gène AURKC chez 87 patients dont 83 cas-index présentant un phénotype de macrozoospermie a permis l’identification de mutations pathogènes dans le gène AURKC chez 82% des patients analysés (68/83). La mutation c.144delC est retrouvée dans plus de 85% des allèles mutés confirmant que cette mutation est l’évènement génétique prépondérant pour ce phénotype.

On observe des différences significatives du pourcentage de spermatozoïdes macrocéphales et multiflagellés dans l’éjaculat des patients mutés (79.6% et 37.9% respectivement) par rapport aux patients non mutés (7.8% et 35.2%), indiquant logiquement que les patients présentant le phénotype le plus pur sont plus fréquemment mutés.

Une nouvelle mutation récurrente, p.Y248*, a été retrouvée chez 10 patients non-apparentés (8 à l’état homozygote et 2 à l’état hétérozygote composite). Les analyses d’expression de l’ARNm de AURKC chez les patients homozygotes pour cette mutation montrent la dégradation totale du transcrit anormal par nonsense-mediated mRNA decay. De manière logique, nous n’observons pas de différence phénotypique entre les patients homozygotes pour c.144delC, les homozygotes pour la mutation p.Y248* et les hétérozygotes composites porteurs de ces deux mutations.

Cette nouvelle mutation s’avère être la mutation majoritairement retrouvée chez les patients d’origine européenne. Un haplotype commun chez les patients porteurs de la mutation ainsi que la co-ségrégation systématique d’un marqueur intragénique rare dans la région 3’UTR du gène avec cette nouvelle mutation suggèrent un effet fondateur. Par l’analyse de différents microsatellites, nous avons pu estimer que la mutation p.Y248* est apparue il y a 925-1325 années alors que la mutation c.114delC avait été estimée être apparue il y a 250-650 années.

Discussion et Perspectives

Cette étude confirme que le phénotype de macrozoospermie est causé très majoritairement par des mutations dans le gène *AURKC* quel que soit l’origine des patients concernés. Deux mutations récurrentes (p.Y248* et c.144delC) représentent plus de 98% des

mutations identifiées. Ces résultats suggèrent de séquencer les exons 3 et 6 du gène *AURKC* en premier intention chez tous les patients avec des spermatozoïdes macrocéphales. Un résultat positif contre-indiquant formellement l'ICSI en raison du contenu tétraploïde de spermatozoïdes macrocéphales.

Aucune mutation du gène *AURKC* n'est retrouvée dans les formes partielles de macrozoospermie. Il est donc probable que d'autres gènes ou des origines non génétiques (environnementales, toxiques) puissent expliquer ces formes atypiques.

Cette étude met aussi en lumière l'étonnant constat que seulement 4 mutations différentes ont été identifiées dans le gène *AURKC*, toutes responsables d'une macrozoospermie complète. Plusieurs hypothèses sont envisageables pour expliquer cette observation : (i) un nombre insuffisant de patients a été encore étudié, (ii) les mutations moins délétères permettent à la protéine une fonction résiduelle suffisante pour conserver sa fonction normale, (iii) les autres mutations sont responsables d'un phénotype différent.

Enfin, l'identification de ces deux mutations récurrentes ancestrales et leur maintien au cours de l'évolution malgré leur effet délétère sur la reproduction masculine à l'état homozygote, pose la question d'un potentiel avantage sélectif procuré par l'haplo-insuffisance d'*AURKC* dont le mécanisme reste à préciser.

PHÉNOTYPE 2:

GLOBOZOOSPERMIE

Article 2

MLPA and sequence analysis of DPY19L2 reveals point mutations causing globozoospermia.

Coutton C, Zouari R, Abada F, Ben Khelifa M, Merdassi G, Triki C, Escalier D, Hesters L, Mitchell V, Levy R, Sermondade N, Boitrelle F, Vialard F, Satre V, Hennebicq S, Jouk PS, Arnoult C, Lunardi J, Ray PF.

Human Reproduction, 27:2549-58, August 2012

Contexte et objectifs

La seconde partie de mon travail a porté sur la globozoospermie, un phénotype rare d'infertilité masculine qui se caractérise par la présence de spermatozoïdes avec une tête ronde, dépourvue d'acrosome. On distingue la globozoospermie totale ou partielle en fonction du taux de spermatozoïdes dépourvus d'acrosome (Dam, Feenstra *et al.*, 2007).

Des cas précédemment décrits de globozoospermie au sein d'une même fratrie orientaient vers une origine génétique de cette tératozoospermie. Cette hypothèse a pu être rapidement confortée par la description d'une mutation familiale dans le gène *SPATA16* (Dam, Koscinski *et al.*, 2007). Cependant aucune autre mutation sur le gène *SPATA16* n'a pu être identifiée sur une large cohorte de patients globozoocéphales (Dam, Koscinski *et al.*, 2007). En 2011, à partir d'une cohorte de 20 patients présentant un tableau clinique de globozoocéphalie complète, notre équipe a démontré que la délétion homozygote du gène *DPY19L2* était la principale cause de globozoospermie (Harbuz *et al.*, 2011).

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 *et al.*, 2006). La présence de ces séquences hautement homologues a jusqu'à présent grandement compliqué l'amplification spécifique et le séquençage de *DPY19L2*. Ainsi, le premier objectif de mon travail sur ce phénotype a été de caractériser le spectre mutationnel complet du gène *DPY19L2* à partir d'une cohorte de 34 patients globozoospermiques et d'estimer la fréquence des altérations de ce gène.

MLPA and sequence analysis of DPY19L2 reveals point mutations causing globozoospermia

Charles Coutton^{1,2,3,4}, Raoudha Zouari⁵, Farid Abada^{1,2,4},
Mariem Ben Khelifa^{1,2,4}, Ghaya Merdassi⁶, Chema Triki⁷,
Denise Escalier⁸, Laetitia Hesters⁹, Valérie Mitchell¹⁰, Rachel Levy¹¹,
Nathalie Sermondade¹¹, Florence Boitrelle¹², François Vialard¹²,
Véronique Satre^{1,2,3}, Sylviane Hennebicq^{1,2,3}, Pierre-Simon Jouk^{2,3},
Christophe Arnoult^{1,2}, Joël Lunardi^{2,4,13}, and Pierre F. Ray^{1,2,4,*}

¹Laboratoire AGIM, CNRS FRE3405, Equipe 'Génétique, Infertilité et Thérapeutiques', La Tronche F-38700, France ²Université Joseph Fourier, Grenoble F-38000, France ³Département de Génétique et Procréation, CHU de Grenoble, Grenoble cedex 9 F-38043, France ⁴UM de Biochimie et Génétique Moléculaire, Département de Biochimie Pharmacologie et Toxicologie, Pole de Biologie, CHU de Grenoble, Grenoble, F-38000, France ⁵Clinique de la Reproduction les Jasmins, 23, Av. Louis BRAILLE, 1002 Tunis, Tunisia ⁶Unité de Procréation Médicalement Assistée, Hôpital Aziza Othmana, Tunis, Tunisia ⁷CMRDP, 5, rue Ibn Hazem, 1002 Tunis Beledère, Tunisia ⁸INSERM U654, Hôpital Armand Trousseau, Paris F-75571, France ⁹APHP, Departments of Biology and Genetics of Reproduction, Antoine Béclère Hospital, Clamart, France ¹⁰Institut de biologie de la reproduction – Spermologie – CECOS, EA 40308 CHRU- 59035 Lille Cedex, France, and EA 4308 spermatogenesis and quality of the male gamete ¹¹Histologie Embryologie Cytogénétique CECOS, CHU Jean Verdier, 93143 Bondy, France ¹²Department of Reproductive Biology, Centre Hospitalier Poissy Saint Germain, Poissy, France ¹³Grenoble Institut des Neurosciences, INSERM U.836, F-38000 Grenoble, France

*Correspondence address. UM de Biochimie et Génétique Moléculaire, Institut de Biologie et Pathologie, CHU de Grenoble, 38 043 Grenoble cedex 9, France. Tel: +33-476-765-573; Fax: +33-476-765-837; E-mail: pray@chu-grenoble.fr

Submitted on February 24, 2012; resubmitted on March 29, 2012; accepted on April 10, 2012

STUDY QUESTION: Do DPY19L2 heterozygous deletions and point mutations account for some cases of globozoospermia?

SUMMARY ANSWER: Two DPY19L2 heterozygous deletions and three point mutations were identified, thus further confirming that genetic alterations of the DPY19L2 gene are the main cause of globozoospermia and indicating that DPY19L2 molecular diagnostics should not be stopped in the absence of a homozygous gene deletion.

WHAT IS KNOWN ALREADY: Globozoospermia is a rare phenotype of primary male infertility characterized by the production of a majority of round-headed spermatozoa without acrosome. We demonstrated previously that most cases in man were caused by a recurrent homozygous deletion of the totality of the DPY19L2 gene, preventing sperm head elongation and acrosome formation. In mammals, DPY19L2 has three paralogs of yet unknown function and one highly homologous pseudogene showing >95% sequence identity with DPY19L2. Specific amplification and sequencing of DPY19L2 have so far been hampered by the presence of this pseudogene which has greatly complicated specific amplification and sequencing.

STUDY DESIGN, SIZE, DURATION: In this cohort study, 34 patients presenting with globozoospermia were recruited during routine infertility treatment in infertility centers in France and Tunisia between January 2008 and December 2011. The molecular variants identified in patients were screened in 200 individuals from the general population to exclude frequent non-pathological polymorphisms.

PARTICIPANTS/MATERIALS, SETTING, METHODS: We developed a Multiplex Ligation-dependent Probe Amplification test to detect the presence of heterozygous deletions and identified the conditions to specifically amplify and sequence the 22 exons and intronic boundaries of the DPY19L2 gene. The pathogenicity of the identified mutations and their action on the protein were evaluated in silico.

MAIN RESULTS AND THE ROLE OF CHANCE: There were 23 patients who were homozygous for the DPY19L2 deletion (67.6%). Only eight of the eleven non-homozygously deleted patients could be sequenced due to poor DNA quality of three patients. Two patients were compound heterozygous carrying one DPY19L2 deleted allele associated respectively with a nonsense (p.Q342*) and a missense mutation (p.R290H). One patient was homozygous for p.M358K, another missense mutation affecting a highly conserved amino acid. Due to

the localization of this mutation and the physicochemical properties of the substituted amino acids, we believe that this variant is likely to disrupt one of the protein transmembrane domains and destabilize the protein. Overall, 84% of the fully analysed patients ($n = 31$) had a molecular alteration of *DPY19L2*. There was no clear phenotypic difference between the homozygous deleted individual, patients carrying a point mutation and undiagnosed patients.

LIMITATIONS, REASONS FOR CAUTION: Globally poor fertilization rates are observed after intracytoplasmic sperm injection of round spermatozoa. Further work is needed to assess whether *DPY19L2* mutated patients present a better or worse prognostic than the non-diagnosed patients. Evaluation of the potential benefit of treatment with a calcium ionophore, described to improve fertilization, should be evaluated in these two groups.

WIDER IMPLICATIONS OF THE FINDINGS: In previous work, deletions of *DPY19L2* had only been identified in North African patients. Here we have identified *DPY19L2* deletions and point mutations in European patients, indicating that globozoospermia caused by a molecular defect of *DPY19L2* can be expected in individuals from any ethnic background.

STUDY FUNDING/COMPETING INTEREST(S): None of the authors have any competing interest. This work is part of the project 'Identification and Characterization of Genes Involved in Infertility (ICG2I)' funded by the program GENOPAT 2009 from the French Research Agency (ANR).

Key words: *DPY19L2* / globozoospermia / point mutations / MLPA / sequence analysis

Introduction

Approximately 15% of all couples are confronted with infertility (Evers, 2002; Gurunath et al., 2011). A male factor is believed to be present in nearly half the cases, often manifested by a qualitative and/or quantitative defect of sperm parameters. Globozoospermia (MIM #613958), characterized by the presence in the ejaculate of a majority of round-headed spermatozoa devoid of acrosome, is a rare (incidence $<0.1\%$) but severe disorder in male infertility (Dam et al., 2007a). In a previous work, we demonstrated that a recurrent homozygous deletion of the *DPY19L2* gene was found in a large majority of globozoospermia patients. This 200 kb deletion encompasses the totality of *DPY19L2* coding sequence without infringing on other surrounding genes (Harbuz et al., 2011). The mechanism leading to the deletion of *DPY19L2* is based on the non-allelic homologous recombination (NAHR) between two highly homologous sequences, or low-copy repeats (LCR), surrounding the breakpoint regions. *DPY19L2* deletion which has been described to account for 75% (Harbuz et al., 2011) and 19% (Koscinski et al., 2011) of studied patients is now recognized as the mutation responsible for most cases of globozoospermia. As a NAHR caused by the presence of a highly homologous LCR, this mutational event is expected to be recurrent. Analysis of public databases compiling data from comparative genomic hybridization experiments in the general population showed the presence of heterozygous deletions centred around *DPY19L2* at a frequency of about 1/220, implying a theoretical disease frequency close to 1/200 000, which is concordant with the rarity of the phenotype. *DPY19L2* is a transmembrane protein (Harbuz et al., 2011) of yet unknown function. Previous work has indicated that DPY-19, a *DPY19L2* ortholog in *Caenorhabditis elegans*, is involved in the establishment of cell polarity in the worm (Honigberg and Kenyon, 2000), a function coherent with the failure to achieve sperm head elongation that is observed in our patients.

Other genetic causes are, however, involved in globozoospermia. A mutation in *SPATA16*, encoding a protein present in the Golgi vesicles of spermatids which serve to fill the acrosome, was described to be responsible for a familial case of globozoospermia (Dam et al., 2007b). No other *SPATA16* mutations could, however, be identified

in a large cohort of globozoospermic patients (Dam et al., 2007b). Recently, using a candidate genes strategy, heterozygous mutations in *ZPBPI* were described in patients presenting with abnormal sperm head morphology, but their involvement in the disease has not been formally demonstrated (Yatsenko et al., 2012). Similarly, a homozygous missense *PICK1* mutation was identified in a Chinese family (Liu et al., 2010), however, with only one familial case it is also difficult to make formal conclusions. Several knockout mice have also been described to present a globozoospermia-like phenotype when the following genes are mutated: *Csnk2a2* (Xu et al., 1999), *Hrb* (Kang-Decker et al., 2001), *Gopc* (Yao et al., 2002), *Pick1* (Liu et al., 2010), *Hsp90b1* (Audouard and Christians, 2011), *Vps54* (Paiardi et al., 2011) and *Zpbp1* (Lin et al., 2007). Only a few mutations have been identified in the human orthologs of these genes, but with no formal proof of causality. Overall, these data suggest that despite a common phenotype, globozoospermia is somewhat phenotypically heterogeneous and genetically heterogeneous but with a predominance of cases involving *DPY19L2*. The full involvement of *DPY19L2* has, however, not been investigated as *DPY19L2* non-deleted patients have not been sequenced to identify *DPY19L2* point mutations.

In order to fully assess the involvement of *DPY19L2* in globozoospermia, we analyzed a large cohort of globozoospermia patients. When a positive diagnosis was not obtained (absence of a homozygous deletion), a sequence analysis of *DPY19L2* 22 exons was carried out to search for potentially deleterious nucleotide variants. We identified three novel point mutations in *DPY19L2* confirming the prevalence of *DPY19L2* mutational events in globozoospermia (Fig. 1). Ultimately, these new findings might help improve our understanding of the structure-function of the protein and its role in acrosome formation and sperm head elongation.

Materials and Methods

Patient and control individuals

We included in this study the 20 patients described in Harbuz et al. (2011). We had reported that 15 out of 20 patients with globozoospermia had a

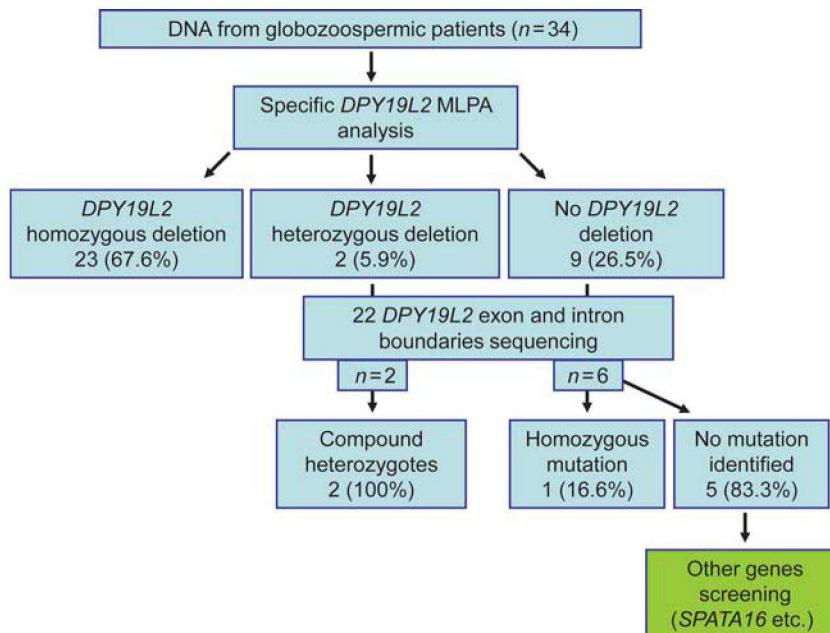


Figure 1 Schematic representation of the molecular investigation carried out on patients with globozoospermia.

homozygous deletion of the *DPY19L2* region. Multiplex Ligation-dependent Probe Amplification (MLPA) was carried out here on the five non-deleted patients. Unfortunately, we ran out of DNA for three patients and could not obtain any additional biological material. Sequence analysis could therefore only be carried for the two remaining patients.

An additional 14 unrelated and unpublished globozoospermic patients were included. These patients were analysed by MLPA and genomic sequencing of all exons and intron-exon junctions was carried out on all who did not carry a homozygous deletion.

Overall, 24 patients consulted for infertility in Tunisia and were of Tunisian origin, and of the 10 who consulted in France, 6 were of North African origin.

A total of 200 anonymous individuals were screened by high-resolution melting (HRM). All had agreed to donate their DNA for research purposes. The fertility of these individuals was not documented. All were French citizens. Half (100) were of North African origin (Algeria, Morocco and Tunisia) and half were of European origin. All patients and anonymous donors gave their written informed consent, and all national laws and regulations were respected.

Sperm analysis

Sperm analysis was carried out in the source laboratories during the course of the routine biological examination of the patient, according to World Health Organization (WHO) guidelines (WHO, 1999). Small protocol variations might be observed between the different laboratories. The different parameters were compared between the different groups (Table I) using a two-tailed *t*-test.

Molecular analyses

DNA extraction

DNA was extracted from blood or saliva. Blood DNA extraction was carried out from 5 to 10 ml of frozen EDTA blood using the quick guanidium chloride extraction procedure (Jeanpierre, 1987). Saliva was collected with an Oragene DNA Self-Collection Kit (DNAgenotech,

Canada) and DNA extraction was performed using the manufacturer's recommendations.

Mutation detection

The 22 *DPY19L2* exons and intronic boundaries were amplified as indicated in Table II. Sequencing analyses were carried out using the BigDye Terminator v3.1 sequencing kit and an ABI PRISM 3130 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) using protocols described in Ben Khelifa *et al.* (2011).

The nomenclature of the identified variants was established according to Human Genome Variation Society (HGVS) as indicated in: www.hgvs.org/rec.html.

Sequence numbering referred to NP_776173.3 for the protein sequence and to DPY19L2-001 ENST00000324472 for the cDNA sequence.

HRM analysis

HRM analysis was performed with the LightCycler 480 (Roche), using the LightCycler 480 HRM master kit. Results were analyzed with the Gene scanning software (Roche) as described in Harbuz *et al.* (2010).

MLPA analysis

The design of the MLPA probes, MLPA reaction and data analysis were performed according to the recommendation of the MRC-Holland synthetic protocol (www.mlpa.com).

For this study, three synthetic MLPA probes specific for exons 1, 17 and 22 of *DPY19L2* were designed (Table III). Because of the high homology between *DPY19L2* and the other *DPY19L* paralogs and pseudogenes, MLPA probes were designed in order to match specific *DPY19L2* single nucleotide mismatches at the ligation site (Schouten *et al.*, 2002). In addition, three MLPA control probes specific to the *OCRL1* gene were included to serve as control probes for copy number quantification. Information about sequences and ligation sites of these control probes can be obtained in Coutton *et al.* (2010). The comparative height of the control probes

Table I Semen parameters measured according to the patient's genotype.

Patients	Homozygous deleted (n = 23)	Point mutation carriers (n = 3)	No identified mutation (n = 5)
Sperm volume (ml)	3.5 (1.7–7.5)	5.6 (3.7–7.5)	3.2 (0.9–6.3)
Nb spz ($\times 10^6$ per ml)	60 (0.6–108)	37.1 (7.5–53)	20 (14–25)
Round cells ($\times 10^6$ cells)	3.1 (0–16)	0.1 (0–0.2)	0.8 (0.4–1.4)
Motility A + B, 1 h	27 (1–40)	27 (20–34)	35 (20–50)
Vitality	60.2 (48–81)	59 (42–76)	62 (47–76)
Normal spermatozoa	0	0	2 (0–6)
Globozoospermes	90.2 (29–100)	71 (53–89)	63 (12–100)
Rolled flagella	24 (8–38)	26 (26–27)	20 (0–41)
Intermediate piece angulation	12.4 (1–26)	13 (8–19)	19 (0–34)
Flagella of irregular calibre	14.2 (0–42)	15 (0–31)	4.3 (1–16)
Shortened flagella	2.8 (0–16)	0.7 (0–1)	1 (0–2)
Absence of flagella	1 (0–4)	0	0
Multiple flagella	0.2 (0–2)	0.7 (0–1)	1 (0–2)
Multiple heads	0.4 (0–4)	0.7 (0–1)	3.8 (0–10)
Multiple Anomalies Index	2.7 (2.1–3)	2.9 (2.3–3.4)	2.8 (1.9–3.3)

Values are expressed as the mean with the lower and higher values between brackets. Values are expressed in percents, unless specified otherwise.

Table II DNA sequences of the 22 DPY19L2 primers pairs and respective melting temperatures (Tm).

Exon	Primer sequences (5' → 3')	Tm (°C)
	Forward	Reverse
1	GGCCAACCTTCTTCTACTCGGAC	ATTCACAGTCGCCATGACG
2	GCTGTTCACATATGAG	AAAGCAGCTATTAAATGAC
3	GAAACAGTGCAGTTGACCAG	ATTCAGGTGTGCCATAC
4	TGGCCATTATTCACACACTAAGG	GCGAGAAGTGATTAGGAAGTCTT
5	ATAGTCAAGATTGCCCTACA	TAATATCAAACACGCAGT
6	ATGACTTGAGATAGAA	AACTATATAATCACTCATTA
7	TAAGGCAAGAGATTTCATGT	GTAAGCTGAGATTTCGACA
8	GCCTTCGTTTATAATCG	GGTAGTTATTGCTGCTAC
9	GCATACATTACCTACAT	AGTTCTTTAGTACTTTAAG
10	CCAAAGAGGAGGTACCGTATAA	GCCATCCATCTTTTAATTCTG
11	AACCTCCTCAAGTGACTTAG	TTGGCCAAGAGTCATT
12	GAAGGTTAATTGAAGCTAGA	ATTAGCCTGCAGAAAATGGT
13	AGAACCTTCATGTTAA	TTCTTCTGGCTATT
14	CTTAGAGGGATGTCAAATAT	TCCAAGTGGCTAGATTATC
15	CCGGTGTACCTACAATGTTAAT	AATGAAAAATTGTAAGTAACC
16	TTTAAACTTGAGTTGGTTACA	GGCATCTATAGTATGACCGTCC
17	GCTCAGCAGGCTAGAAACTG	GCCACCATCGGAACAC
18	AATTAGTCAGCAAAGCCACA	TAGACATTGATAAATTATTGC
19	GGGTTTAATTGATTGACATT	AATTATTGTTGACCCCTACG
20	CAGAGGCAACAGGTACGTAT	ACCCTTAGAACTGTGAAGATTA
21	AGGGTTAATCACTCACTAGTT	TATATTCTGAAAAGTGTGAA
22	GTGTCAGTTATAAAGCTTG	ATTGTCTCTAGACAGCAATACAT

Table III DNA sequences of the DPY19L2 specific MLPA home-made probes.

Exon	MLPA probes (5' → 3')		Total length (nt)
	Forward	Reverse	
1	CAGCTCCGGGAAAGGTGCAG	GAACTGCAGCGCGCGGTTC	96
17	GGTGATCAGCAGCAAGGTTACCTTGTCAATTAT	CCTCAAACATAATTAACACAAGTCTACCAGAGA	124
22	CTAACCCCTCCCTATGTAGCGTCTGCTCG	AAGACGCCAGGCCTTACTTCACCAACAGTATT	118

Specific DPY19L2 single nucleotide polymorphisms (SNP) at the 3' ligation site are underlined. Only the hybridizing sequences (left and right oligos) are represented. Universal primers or stuffer are not represented. Total probe length represents the total length of amplification product (nt: nucleotides) = forward universal primers + left oligo sequence + right oligo sequence + reverse universal primers ± stuffer (if used).

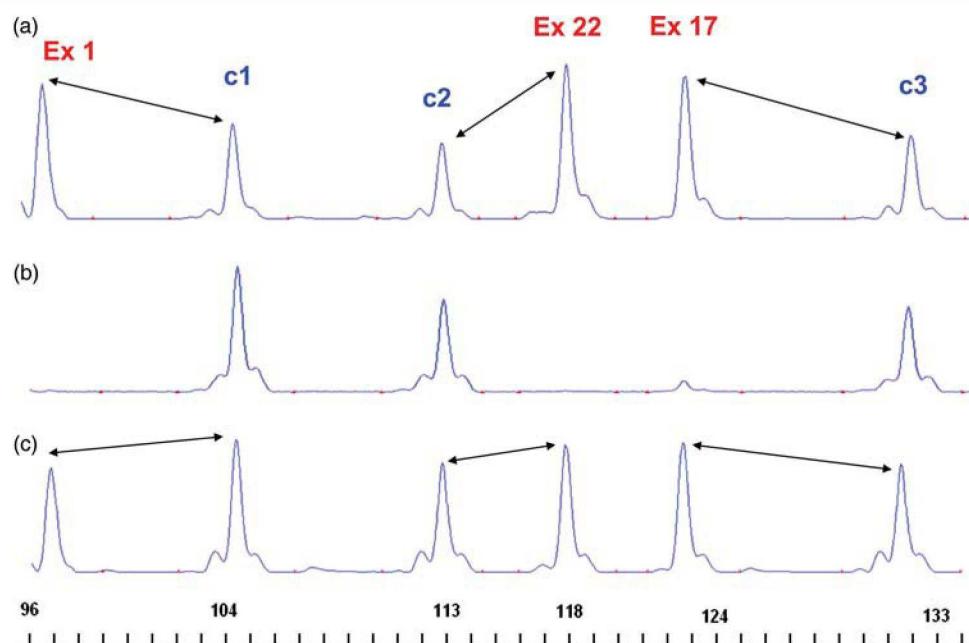


Figure 2 MLPA analysis of DPY19L2 exon 1, 17, 22. Probes set contain three control probes for normalization purposes (c1, c2, c3) and three specific DPY19L2 probes (X1, X17 and X22) corresponding to exon 1, 17 and 22 of DPY19L2. (a) MLPA profile of a normal (control) DNA sample. (b) MLPA profile of homozygous deleted patient showing a total absence of DPY19L2 exon probes signals. (c) MLPA profile of heterozygous deleted patient showing a halving of the specific DPY19L2 probes signals (the variation is normalized with the three control probes).

with target probes allows access to the number of target copies present in the sampled DNA (Fig. 2).

In-silico analyses of sequence variants and prediction of protein conformation

The pathogenicity of the identified false sense variants was evaluated using the MutPred (<http://mutpred.mutdb.org/>) and MuStab (<http://bioinfo.ggc.org/mustab/>) webserver as recommended by Thusberg *et al.* (2011). The potential effect of these variants on RNA splicing was assessed with <http://www.umd.be/HSF/> (Yeo and Burge, 2004; Desmet *et al.*, 2009).

Protein alignments [DPY19L2 paralogs and orthologs (Fig. 3b)] were realized with CLC Sequence viewer 6 (<http://mac.softpedia.com/get/Math-Scientific/CLC-Free-Workbench.shtml>).

Prediction of transmembrane helices in proteins was performed using TMHMM server v.2 based on a hidden Markov model (<http://www.cbs.dtu.dk/services/TMHMM/>).

Results

MLPA analysis

Of the 34 analysed patients, 23 (67.6%) had a homozygous DPY19L2 deletion, 2 were heterozygous (5.9%) and 9 were non-deleted (26.4%; Fig. 1). For illustration of the MLPA technique, see Fig. 2.

DPY19L2 sequencing and HRM analysis

Due to lack of DNA, full sequence analysis could not be performed on three patients. A total of six deletion-negative and two heterozygous patients were analysed by DNA sequencing (Fig. 1). All 22 DPY19L2 exons and intron boundaries were sequenced for these eight patients. No mutations were identified in five patients. Gene sequencing revealed a total of three point mutations in three individuals. A point mutation was identified on the remaining allele of the two

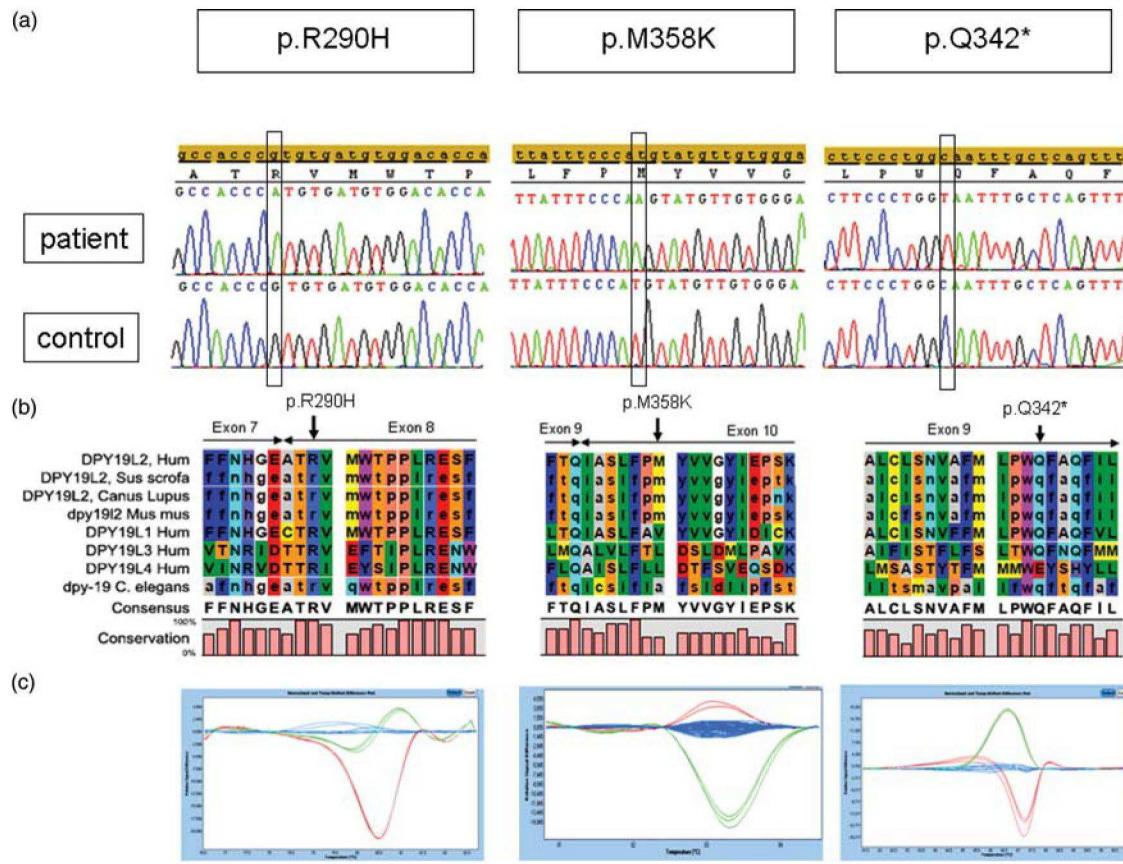


Figure 3 DPY19L2 novel point mutations. (a) Electropherogram of DPY19L2 exon 8, 10 and 9 showing the mutated sequence and sequence obtained from a control individual. Three patients carried a hemizygous missense mutation (p.R290H) in DPY19L2 exon 8, a hemizygous nonsense mutation (p.Q342*) in DPY19L2 exon 9 or a homozygous missense mutation (p.M358K) in DPY19L2 exon 10. (b) The arginin in position 290, the methionine in position 358 and the glutamine in position 342 are all conserved within species and DPY19 paralogs. Amino acid sequence alignment of exon 7, 8, 9 and 10 of the human DPY19L2 with paralogs and interspecies DPY19 sequences was realized with CLC Sequence viewer 6. (c) HRM profiles of DPY19L2 exon 8, 10 and 9 from the three patient carrying exon mutations respectively in red (replicated) and from control individuals in blue. The green lines represent artificial heterozygous DNAs. All mutations found in these patients were not present in 100 control subjects (blue lines).

heterozygotously deleted patients who were both of European origin. The first patient carried a heterozygous missense mutation in exon 8 altering the DPY19L2 869th nucleotide: c.869G>A and modifying the 290th amino acid: p.R290H. The second patient carried a heterozygous nonsense mutation in exon 9: c.1024C>T; p.Q342* (Fig. 3). A third patient, of North African origin, carried a homozygous missense mutation in exon 10: c.1073T>A; p.M358K. We did not have access to DNA from this patient's parents and could not verify that both were heterozygous for this variant to exclude the presence of a small deletion (not detected by MLPA) centred on exon 10. As the parents were first cousins, we consider that it is much more likely that the variant is indeed homozygous.

To exclude the possibility that the identified variants may be common in the studied populations, we performed an HRM of DPY19L2 exon 8 and 9 in 100 individuals of French origin and of exons 10 for 100 individuals of North African origin. The HRM technique represents an efficient way of detecting variants in amplified fragments. Three artificial heterozygous control DNA were tested by

HRM for each mutation and showed a characteristic green profile (Fig. 3c). The homozygous or hemizygous patients were also passed in triplicate and are shown in red. Each profile (homozygous, heterozygous and non-mutated) clearly showed a distinct profile (Fig. 3c). None of the control DNA samples, shown in blue, presented any abnormal profile for any of the exons tested. Moreover, M358K and Q342* were absent from the most recent database: dbSNP build (build 135, October 2011; <http://www.ncbi.nlm.nih.gov/snp/>) which regroups sequence data from several thousand individuals. One R290H allele was identified out of 4471 alleles. This low allelic frequency (0.022%) is however not surprising for a recessive trait.

Prediction of the effects of the variants on the mRNA and the protein

The nonsense mutation, p.Q342*, changed a glutamine codon (CAA) into a stop codon (TAA) at amino acid position 342 in exon 9 (Fig. 3a). This premature stop codon is expected to produce a truncated

protein missing 416 amino acids out of the 758 residues of DPY19L2. The other two mutations described here are missense mutations which cause non-conservative amino acid substitutions. Both concern highly conserved amino acids which are located in evolutionary conserved domains of the protein (Fig. 3b).

The first missense mutation (p.R290H) is an arginine to histidine substitution in exon 8 (Fig. 4a and b). The utilization of the MutPred algorithms (Li *et al.*, 2009) indicates that p.R290H is deleterious with a probability of 0.785. Furthermore, analysis with MuStab web-server (Teng *et al.*, 2010) predicts that this substitution would decrease the protein stability with a confidence of 83.6%.

The second missense mutation (p.M358K) concerns the substitution of a conserved methionine by a lysine in exon 10. For this mutation, MutPred indicates a deleterious effect with a probability of 0.766. This software also indicates a very probable loss of stability ($P = 0.0076$), the gain of a catalytic residue at M358 ($P = 0.0116$) and a gain of ubiquitination at M358 ($P = 0.0311$). MuStab prediction is

also in favour of a decreased stability with a prediction confidence of 89.6%.

False sense or silent exonic variants can also alter RNA splicing by modifying key regulatory signals. We therefore analysed the effect our two variants could have on splicing using HSF Matrice and MaxEnt software. Neither of the two variants introduced an acceptor or a donor site nor were they predicted to have any other obvious effect on splicing.

Prediction of protein conformation

To be able to predict the potential effect of the identified variants we wanted to obtain a 2D model of DPY19L2. All prediction software indicate that DPY19L2 is a multipass membrane protein. We used a TransMembrane prediction program using Hidden Markov Models (TMHMM), a membrane protein topology prediction method which discriminates between soluble and membrane domains with a

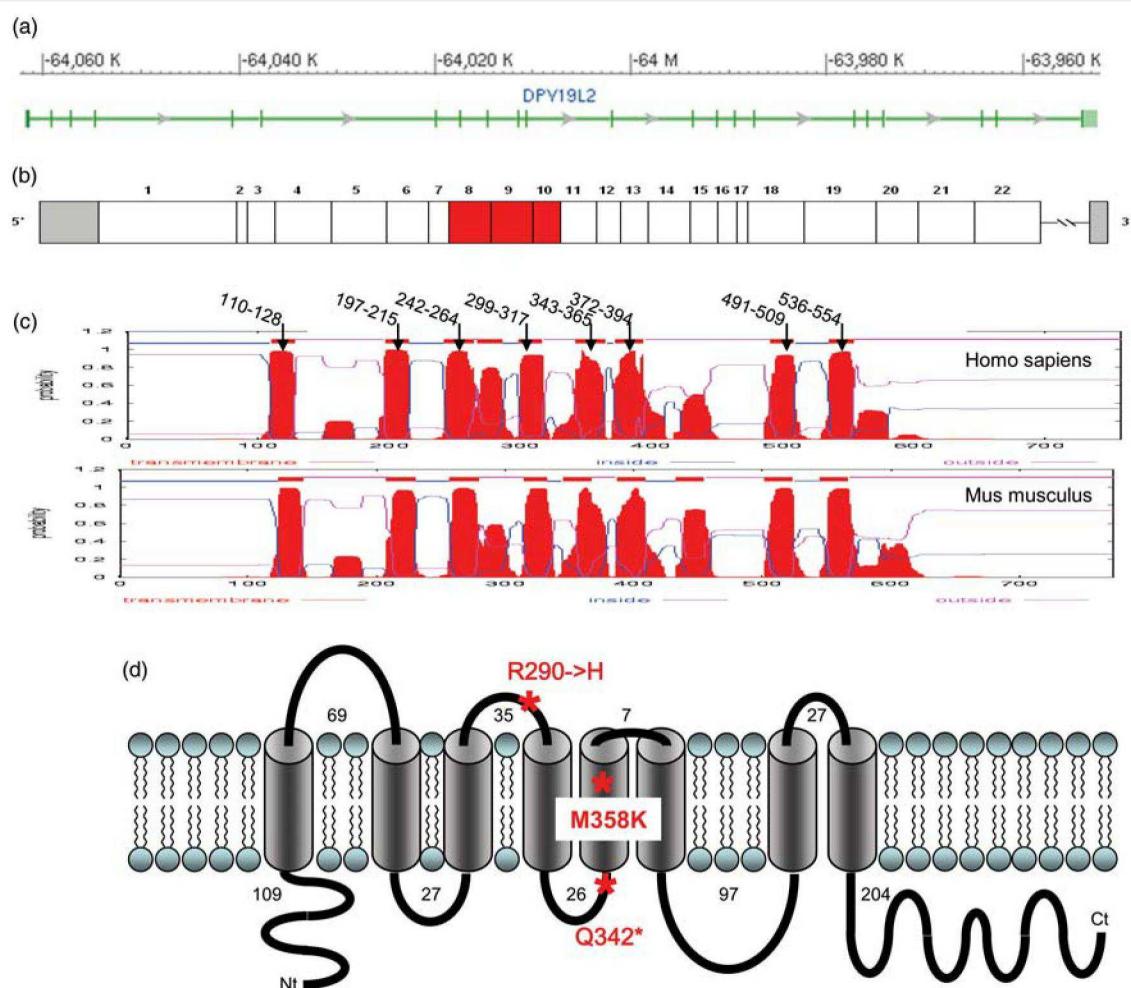


Figure 4 Representation of DPY19L2 genomic DNA, cDNA, and protein structure and localization of the identified mutations. (a) Schematic representation of the DPY19L2 gene. Physical positions in Megabase from the 12p telomere are reported in abscissa. (b) Schematic representation of the cDNA of DPY19L2 gene. The scale of the exons is respected. The three exons (8, 9, 10) where mutations were identified are highlighted in red. (c) Prediction of transmembrane helices in proteins by TMHMM server, in human (top) and mouse (bottom) based on a hidden Markov model as calculated by TMHMM Server v. 2.0. (d) Schematic representation of the DPY19L2 protein based on a model with eight transmembrane domains. The numbers indicate the length of the extra membranous domains.

specificity and sensitivity better than 99% (Kahsay et al., 2005). TMHMM indicates that the human and mice DPY19L2 proteins have nine transmembrane domains. In the human, the fourth domain is just above the threshold whereas the same sequence in mice is below and is not considered as a transmembrane domain. The opposite happens for mouse exon 7. As the protein function is preserved in both species [DPY19L2 knock-out mice perfectly mirror the human phenotype (unpublished data)], we believe that it is highly unlikely that the protein structure differs between the two species. We therefore think that it is likely that both proteins have 8 transmembrane domains as shown in Fig. 4d. We, however, cannot rule out the possibility that both proteins may have 10 transmembrane domains.

Comparison of sperm parameters

The values measured during routine spermogram and spermcytogram were compared between homozygously deleted patients, those carrying point mutations and those for whom no *DPY19L2* molecular alteration had been identified (Table I). No significant phenotypic difference was observed between these three groups. Patients came from different centres, and despite a common observance of WHO' guidelines (WHO, 1999), important scoring variations were observed even in genetically identical individuals. Additionally to the head morphological defect, we observed frequent midpiece and flagellar defects.

Discussion

In total, 23 and 2 patients out of 34 were diagnosed respectively with a homozygous and a heterozygous *DPY19L2* deletion. The detection of heterozygous individuals was realized by MLPA which allows a quantitative assessment of *DPY19L2* allelic status with more sensitivity and specificity than with other amplification-based techniques. MLPA is less sensitive to DNA fragmentation than other long PCR or other gene dosage techniques (Kozlowski et al., 2007). According to our experience, this can be particularly useful for the analysis of DNA extracted from saliva which can sometime be of substandard quality (personal experience). A long PCR amplification allowing amplification across the deleted region has been described previously (Koscinski et al., 2011). Such an approach, combined with exon-specific amplification allows the detection of heterozygous deletions. A negative result will, however, be obtained if the breakpoints fall outside the region covered by the deletion-specific primers. MLPA does not present this shortcoming and allows the detection of any deletion of *DPY19L2* whole coding the sequence, irrespectively of the localization of the breakpoints. For the diagnosis of globozoospermia, we therefore recommend initiating the molecular diagnosis by *DPY19L2* MLPA to identify all homo- and heterozygously deleted individuals.

Following MLPA analysis, *DPY19L2* sequence analysis was carried out when enough DNA was available from the non-homozygously deleted patients. We identified the first point mutations in *DPY19L2* in globozoospermia patients thereby yet confirming the predominance of *DPY19L2* molecular abnormalities in the globozoospermia phenotype. These results also further confirm that it is the absence of a

functional *DPY19L2* protein that is responsible for the globozoospermia phenotype and not an indirect effect of the deletion. The first mutation was a non-sense (stop) mutation identified in a compound heterozygous individual carrying a heterozygous deletion. We wanted to assess whether a truncated protein was produced or if the aberrant mRNA was degraded by non-sense mediated mRNA decay. Since *DPY19L2* presents a testis restricted expression, we realized an RT-PCR on fresh sperm cells from fertile control individuals but did not obtain any amplification (data not shown). As there was no clinical rationale for conducting a testis biopsy, we could not obtain any mRNA from this mutated patient and could not assess whether the transcript was degraded or not. We believe that the presence of a shortened protein lacking more than half of its sequence would be at least as detrimental to the cell as the total absence of that protein. Irrespective of the molecular physiopathology, we therefore consider that this mutation provokes a complete loss of function. We then wanted to understand the action of the two amino acid substitutions on the protein function. Both variants involve amino acids that are highly conserved throughout evolution and concern amino acids with very different physicochemical properties. We used several software analyses to evaluate the effect these variants have at the protein level and on mRNA splicing. Both variants were predicted to have a highly deleterious effect on the protein but no obvious effect on mRNA splicing. We therefore wanted to better evaluate the localization of these variants on the protein domains. Comparing the prediction of the localization of the transmembrane domains between man and mouse, we postulate that *DPY19L2* has 8 transmembrane domains (Fig. 4d). Based on this prediction, we see that p.R290H is located on an extra-membrane domain. Physicochemical changes resulting from this substitution could modify an interaction site between *DPY19L2* and an essential protein partner. The second missense mutation (p.M358K) is predicted to be located in the fifth transmembrane domain. It substitutes a non-polar (M methionine) with a polar amino acid (K lysine). We can therefore predict that this amino acid variation is likely to disrupt a transmembrane domain and potentially the whole anchorage of *DPY19L2* in the membrane. These three novel mutations are all located in the central part of the protein (exon 8, 9 and 10), suggesting that this region is likely to have a particularly important function (Fig. 4). Transfection and expression of wild-type and mutant protein could be carried out ex vivo to better understand the action of these mutations. These mutations could be used to identify *DPY19L2* partners by subtractive pull down with wild-type and mutant peptides. This approach would be particularly suited to p.R290H, which is predicted to be located on an extra-membrane loop.

The patients presented here consulted for infertility at fertility centres in France and Tunisia. Of the 31 fully analysed patients, 22 (68%) consulted at procreation centres in Tunisia and of 10 who consulted in France, 6 were of North African descent. The four patients of European origin carried *DPY19L2* alterations: two were the compound heterozygotes mentioned previously and two were homozygously deleted. Occurrence of genomic recombinations by NAHR such as the *DPY19L2* deletion is expected to occur at a similar frequency irrespective of the genetic background of the individuals. This is indeed confirmed by the observation that this CNV was observed in different populations (Shaikh et al., 2009). The higher incidence of globozoospermia observed in North African men is therefore not likely due to

an increased allelic frequency but is more likely to be the manifestation of the high rate of intra-familial marriages specific to this population which strongly favours the emergence of recessive traits.

The values measured during routine spermogram and spermocytogram were compared between homozygously deleted patients, patients carrying point mutations and patients for whom no DPY19L2 molecular alteration had been identified (Table I). No significant phenotypic difference was observed between these three groups. Additionally to the head morphological defect that is the hallmark of globozoospermia, we observed frequent midpiece and flagellar defects that have seldom been highlighted in the description of globozoospermia. This kind of defect is logical as acrosome formation, sperm head elongation and flagellar elongation are concomitant during spermiogenesis and depend on highly interconnected processes (Kierszenbaum and Tres, 2004). We currently do not have any clear genotype, phenotype (IVF success rate) correlation as we obtained poor fertilization rates, but could obtain a few pregnancies after ICSI in both groups. Recent work suggests that the low fertilization rate observed with globozoospermatozoa is at least partially caused by a decrease or a defect in PLC ζ , a protein involved in the induction of calcium oscillations triggering oocyte activation (Yoon et al., 2008; Heytens et al., 2009). PLC ζ mutations have been demonstrated to cause infertility by preventing oocyte activation without inducing a globozoospermia phenotype (Kashir et al., 2012). Several centres have tried to overcome this problem using calcium ionophores which facilitate the transport of Ca²⁺ across the plasma membrane and artificially activate the oocyte. The utilization of such a chemical is not consensual as it does not mimic the physiological activation of the oocyte, but its usage does seem to improve the overall pregnancy rate (Dam et al., 2007a). Prospective studies evaluating the success rate of ICSI with or without calcium ionophore on mutated and non-mutated patients should be performed to evaluate the prognosis of these patients with different protocols.

A total of 34 globozoospermic patients are presented in this study. Overall 23 (67.6%) were homozygous for the DPY19L2 deletion. MLPA analysis indicated that two (5.9%) were heterozygous for the genomic deletion. A point mutation was identified in the non-deleted allele of these two patients. Nine individuals (26.5%) did not carry a deletion, however due to a lack of DNA, only six patients could be fully sequenced and one was found to carry a homozygous point mutation. Overall, a molecular defect of DPY19L2 was therefore identified in 26 out of 31 of the fully analysed patients (84%), confirming that the vast majority of type I globozoospermia is indeed caused by a molecular defect of DPY19L2. As a clinical diagnostic strategy, we recommend first looking for the presence of the DPY19L2 genomic deletion. For that, we advocate the use of an MLPA approach to easily determine the number of DPY19L2 alleles for each patient, independently of the localization of the breakpoint and of DNA quality. DNA sequence analysis is then required for individuals with a heterozygous deletion as a point mutation was identified in the two heterozygous patients we analyzed. We also advise sequencing the non-deleted patients as we identified one homozygous mutation out of 6 sequenced patients (16%). Due to the presence of the pseudogene, however, full sequence analysis of the 22 exons remains challenging as well as costly. As this diagnosis does not yet provide any clear prognostic or therapeutic indication, a systematic sequencing might not be practical as a routine analysis. We believe, however, that a full

DPY19L2 investigation and follow up of ICSI attempts will help gain a better comprehension of the DPY19L2 function. This could lead to a better comprehension of the physiopathology of globozoospermia and might finally translate into adequate counselling and recommendations for the patient's treatment and might even lead to the identification of new therapeutic solutions.

Acknowledgements

We thank our patients and control individuals for their participation.

Authors' roles

C.C., F.A. and M.B.K. undertook all the molecular work. R.Z., G.M., C.T., D.E., L.H., V.M., R.L., N.S., F.B., F.V., V.S. and S.H. handled the recruitment of patients, sample collection, sperm analyses and supervised the clinical aspects of the work. C.C., C.A., J.L., P.-S.J. and P.F.R. contributed to data analysis. P.F.R. designed the overall study, supervised all molecular laboratory work, had full access to all of the data in the study and takes responsibility for the integrity of the data and its accuracy. All authors contributed to the report.

Funding

This work is part of the project Identification and Characterization of Genes Involved in Infertility (ICG2I) funded by the program GENOPAT 2009 from the French Research Agency (ANR).

Conflict of interest

None declared.

References

- Audouard C, Christians E. Hsp90beta1 knockout targeted to male germline: a mouse model for globozoospermia. *Fertil Steril* 2011; **95**:1475–1477 e1471–1474.
- Ben Khelifa M, Zouari R, Harbuz R, Halouani L, Arnoult C, Lunardi J, Ray PF. A new AURKC mutation causing macrozoospermia: implications for human spermatogenesis and clinical diagnosis. *Mol Hum Reprod* 2011; **17**:762–768.
- Coutton C, Monnier N, Rendu J, Lunardi J. Development of a multiplex ligation-dependent probe amplification (MLPA) assay for quantification of the OCRL1 gene. *Clin Biochem* 2010; **43**:609–614.
- Dam AH, Feenstra I, Westphal JR, Ramos L, van Golde RJ, Kremer JA. Globozoospermia revisited. *Hum Reprod Update* 2007a; **13**:63–75.
- Dam AH, Koscinski I, Kremer JA, Moutou C, Jaeger AS, Oudakker AR, Tournaye H, Charlet N, Lagier-Tourenne C, van Bokhoven H et al. Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet* 2007b; **81**:813–820.
- Desmet FO, Hamroun D, Lalonde M, Collod-Beroud G, Claustres M, Beroud C. Human Splicing Finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 2009; **37**:e67.
- Evers JL. Female subfertility. *Lancet* 2002; **360**:151–159.
- Gurunath S, Pandian Z, Anderson RA, Bhattacharya S. Defining infertility – a systematic review of prevalence studies. *Hum Reprod Update* 2011; **17**:575–588.

- Harbuz R, Lespinasse J, Boulet S, Francannet C, Creveaux I, Benkhelifa M, Jouk PS, Lunardi J, Ray PF. Identification of new FOXP3 mutations and prenatal diagnosis of IPEX syndrome. *Prenat Diagn* 2010; **30**:1072–1078.
- Harbuz R, Zouari R, Pierre V, Ben Khelifa M, Kharouf M, Coutton C, Merdassi G, Abada F, Escoffier J, Nikas Y et al. A Recurrent Deletion of DPY19L2 Causes Infertility in Man by Blocking Sperm Head Elongation and Acrosome Formation. *Am J Hum Genet* 2011; **88**:351–361.
- Heytens E, Parrington J, Coward K, Young C, Lambrecht S, Yoon SY, Fissore RA, Hamer R, Deane CM, Ruas M et al. Reduced amounts and abnormal forms of phospholipase C zeta (PLC ζ) in spermatozoa from infertile men. *Hum Reprod* 2009; **24**:2417–2428.
- Honigberg L, Kenyon C. Establishment of left/right asymmetry in neuroblast migration by UNC-40/DCC, UNC-73/Trio and DPY-19 proteins in *C. elegans*. *Development* 2000; **127**:4655–4668.
- Jeanpierre M. A rapid method for the purification of DNA from blood. *Nucleic Acids Res* 1987; **15**:9611.
- Kahsay RY, Gao G, Liao L. An improved hidden Markov model for transmembrane protein detection and topology prediction and its applications to complete genomes. *Bioinformatics* 2005; **21**:1853–1858.
- Kang-Decker N, Mantchev GT, Juneja SC, McNiven MA, van Deursen JM. Lack of acrosome formation in Hrb-deficient mice. *Science* 2001; **294**:1531–1533.
- Kashir J, Konstantinidis M, Jones C, Lemmon B, Lee HC, Hamer R, Heindryckx B, Deane CM, De Sutter P, Fissore RA et al. A maternally inherited autosomal point mutation in human phospholipase C zeta (PLC ζ) leads to male infertility. *Hum Reprod* 2012; **27**:222–231.
- Kierszenbaum AL, Tres LL. The acrosome-acroplaxome-manchette complex and the shaping of the spermatid head. *Arch Histol Cytol* 2004; **67**:271–284.
- Koscinski I, Elinati E, Fossard C, Redin C, Muller J, Velez de la Calle J, Schmitt F, Ben Khelifa M, Ray PF, Kilani Z et al. DPY19L2 deletion as a major cause of globozoospermia. *Am J Hum Genet* 2011; **88**:344–350.
- Kozlowski P, Roberts P, Dabora S, Franz D, Bissler J, Northrup H, Au KS, Lazarus R, Domanska-Pakiela D, Kotulska K et al. Identification of 54 large deletions/duplications in TSC1 and TSC2 using MLPA, and genotype-phenotype correlations. *Hum Genet* 2007; **121**:389–400.
- Li B, Krishnan VG, Mort ME, Xin F, Kamati KK, Cooper DN, Mooney SD, Radivojac P. Automated inference of molecular mechanisms of disease from amino acid substitutions. *Bioinformatics* 2009; **25**:2744–2750.
- Lin YN, Roy A, Yan W, Burns KH, Matzuk MM. Loss of zona pellucida binding proteins in the acrosomal matrix disrupts acrosome biogenesis and sperm morphogenesis. *Mol Cell Biol* 2007; **27**:6794–6805.
- Liu G, Shi QW, Lu GX. A newly discovered mutation in PICK1 in a human with globozoospermia. *Asian J Androl* 2010; **12**:556–560.
- Paiardi C, Pasini ME, Gioria M, Berruti G. Failure of acrosome formation and globozoospermia in the wobbler mouse, a Vps54 spontaneous recessive mutant. *Spermatogenesis* 2011; **1**:52–62.
- Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. *Nucleic Acids Res* 2002; **30**:e57.
- Shaikh TH, Gai X, Perin JC, Glessner JT, Xie H, Murphy K, O'Hara R, Casalunovo T, Conlin LK, D'Arcy M et al. High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res* 2009; **19**:1682–1690.
- Teng S, Srivastava AK, Wang L. Sequence feature-based prediction of protein stability changes upon amino acid substitutions. *BMC Genomics* 2010; **11**(Suppl 2):S5.
- Thusberg J, Olatubosun A, Vihtinen M. Performance of mutation pathogenicity prediction methods on missense variants. *Hum Mutat* 2011; **32**:358–368.
- WHO. World Health Organization. *WHO Laboratory Manual for the Examination of Human Semen and Sperm–Cervical Mucus Interaction*, 4th edn. Cambridge: Cambridge University Press, 1999, 128.
- Xu X, Toselli PA, Russell LD, Seldin DC. Globozoospermia in mice lacking the casein kinase II alpha' catalytic subunit. *Nat Genet* 1999; **23**:118–121.
- Yao R, Ito C, Natsume Y, Sugitani Y, Yamanaka H, Kuretake S, Yanagida K, Sato A, Toshimori K, Noda T. Lack of acrosome formation in mice lacking a Golgi protein, GOPC. *Proc Natl Acad Sci USA* 2002; **99**:11211–11216.
- Yatsenko AN, O'Neil DS, Roy A, Arias-Mendoza PA, Chen R, Murthy LJ, Lamb DJ, Matzuk MM. Association of mutations in the zona pellucida binding protein I (ZPBPI) gene with abnormal sperm head morphology in infertile men. *Mol Hum Reprod* 2012; **18**:14–21.
- Yeo G, Burge CB. Maximum entropy modeling of short sequence motifs with applications to RNA splicing signals. *J Comput Biol* 2004; **11**:377–394.
- Yoon SY, Jellerette T, Salicioni AM, Lee HC, Yoo MS, Coward K, Parrington J, Grow D, Cibelli JB, Visconti PE et al. Human sperm devoid of PLC ζ fail to induce Ca(2+) release and are unable to initiate the first step of embryo development. *J Clin Invest* 2008; **118**:3671–3681.

Principaux résultats

Une délétion homozygote de *DPY19L2* a été retrouvée par MLPA chez 22 des 30 patients non apparentés étudiés (73.3%) et 3 nouvelles mutations ponctuelles ont été identifiées dans ce gène. Deux patients étaient hétérozygotes composites portant un allèle déleté pour *DPY19L2* associés respectivement à une mutation non-sens (p.Q342*) et à une mutation faux-sens (p.R290H). Un troisième patient était homozygote pour une autre mutation faux-sens, p.M358K, affectant un acide aminé très conservé au cours de l'évolution. En raison de la localisation des mutations faux-sens identifiées et des propriétés physico-chimiques des acides aminés substitués, il est très probable que ces variants perturbent l'un des domaines transmembranaires de la protéine *DPY19L2* et déstabilise la structure et/ou la fonction de la protéine.

Globalement, environ 84% des patients non apparentés ayant bénéficiés d'une analyse totale du gène ($n = 30$) présentent une altération moléculaire de *DPY19L2* confirmant l'implication prédominante mais non exclusive de ce gène dans le phénotype de globozoospermie. Aucune différence significative n'a été observée entre le phénotype des patients porteurs de la délétion génomique de *DPY19L2* à l'état homozygote et les patients avec une mutation ponctuelle ou les patients sans mutation de *DPY19L2*.

Discussion et Perspectives

Cette étude est la première à décrire des mutations ponctuelles de *DPY19L2* à l'origine de globozoospermie. Ces résultats indiquent que l'analyse moléculaire de *DPY19L2* ne devrait pas être limitée à la recherche de la délétion homozygote de *DPY19L2* chez les patients globozoospermiques. Cependant, l'analyse moléculaire du gène *DPY19L2* est laborieuse et compliquée en raison de la présence de nombreux pseudogènes. La délétion homozygote reste le défaut moléculaire très largement majoritaire. Ainsi, la recherche de la délétion homozygote peut être recommandée en première intention. En cas de résultat négatif, le séquençage doit se discuter et peut se justifier dans le cas d'un patient globozoospermique hétérozygote pour la délétion et chez les patients avec des formes complètes de globozoospermie.

Une nouvelle technique par MLPA a été développée permettant de détecter facilement les délétions chez les patients. Cette technique présente notamment l'avantage par rapport aux

autres approches PCR-dépendantes de s'affranchir de la position variable des points de cassure pouvant être à l'origine de faux négatifs.

Aucune différence statistiquement significative n'a pu être démontrée concernant les paramètres spermatiques entre les patients mutés pour *DPY19L2* et non mutés même si une tendance tend à prouver une augmentation du taux de spermatozoïdes globozoocéphales chez les patients mutés. Ces observations nécessitent d'être confirmées par d'autres études standardisées de corrélation génotype-phénotype.

Enfin, ces résultats ouvrent des perspectives intéressantes sur la possibilité d'identification des protéines partenaires de *DPY19L2* dans la formation de l'acrosome et *in extenso* de nouveaux gènes impliqués dans la globozoospermie. En effet, l'absence de mutation identifiée dans le gène *DPY19L2* chez une petite proportion de patients globozoospermiques dont certains avec des formes complètes suggère que d'autres gènes sont impliqués dans ce phénotype.

Article 3

Fine characterisation of a recombination hotspot at the DPY19L2 locus and resolution of the paradoxical excess of duplications over deletions in the general population.

Coutton C, Abada F, Karaouzene T, Sanlaville D, Satre V, Lunardi J, Jouk PS, Arnoult C, Thierry-Mieg N, Ray PF

Plos Genetics, 9:e1003363, March 2013

Contexte et objectifs

La délétion récurrente de *DPY19L2* est produite par le mécanisme de recombinaison homologue non allélique (NAHR) survenant entre deux séquences répétées homologues (LCR) de 28kb situées de chaque côté du gène (Harbuz *et al.*, 2011). Les modèles théoriques de NAHR prédisent que ce mécanisme génère *de novo* plus d'allèles recombinés délétés que dupliqués lors de la méiose male (Liu *et al.*, 2012). Étonnamment, dans la population générale les allèles *DPY19L2* dupliqués sont trois fois plus fréquents que les allèles délétés.

Ainsi, le second objectif de mon travail sur ce phénotype a été de caractériser précisément le mécanisme génétique et les facteurs favorisant la survenue par NAHR de la délétion homozygote récurrente emportant totalement le gène *DPY19L2*. De même, nous avons tenté de résoudre le paradoxe observé entre le modèle théorique de NAHR et la fréquence des allèles observée dans la population générale.

Fine Characterisation of a Recombination Hotspot at the DPY19L2 Locus and Resolution of the Paradoxical Excess of Duplications over Deletions in the General Population

Charles Coutton^{1,2,3}, Farid Abada^{1,2}, Thomas Karaouzene^{1,2}, Damien Sanlaville⁴, Véronique Satre^{1,3}, Joël Lunardi^{1,2}, Pierre-Simon Jouk³, Christophe Arnoult¹, Nicolas Thierry-Mieg⁵, Pierre F. Ray^{1,2,3*}

1 Equipe "Génétique, Infertilité et Thérapeutiques," Laboratoire AGIM, CNRS FRE3405, Université Joseph Fourier, Grenoble, France, **2** Biochimie et Génétique Moléculaire, DBTP, CHU Grenoble, Grenoble, France, **3** Département de Génétique et Procréation, Hôpital Couple Enfant, CHU de Grenoble, Grenoble, France, **4** Hospices Civils de Lyon, Centre de Biologie et de Pathologie Est, Service de Cytogénétique Constitutionnelle, Lyon, France, **5** Université Joseph Fourier, Centre National de la Recherche Scientifique, Laboratoire TIMC-IMAG UMR 5525, Grenoble, France

Abstract

We demonstrated previously that 75% of infertile men with round, acosomeless spermatozoa (globozoospermia) had a homozygous 200-Kb deletion removing the totality of DPY19L2. We showed that this deletion occurred by Non-Allelic Homologous Recombination (NAHR) between two homologous 28-Kb Low Copy Repeats (LCRs) located on each side of the gene. The accepted NAHR model predicts that inter-chromatid and inter-chromosome NAHR create a deleted and a duplicated recombined allele, while intra-chromatid events only generate deletions. Therefore more deletions are expected to be produced de novo. Surprisingly, array CGH data show that, in the general population, DPY19L2 duplicated alleles are approximately three times as frequent as deleted alleles. In order to shed light on this paradox, we developed a sperm-based assay to measure the de novo rates of deletions and duplications at this locus. As predicted by the NAHR model, we identified an excess of de novo deletions over duplications. We calculated that the excess of de novo deletion was compensated by evolutionary loss, whereas duplications, not subjected to selection, increased gradually. Purifying selection against sterile, homozygous deleted men may be sufficient for this compensation, but heterozygously deleted men might also suffer a small fitness penalty. The recombined alleles were sequenced to pinpoint the localisation of the breakpoints. We analysed a total of 15 homozygous deleted patients and 17 heterozygous individuals carrying either a deletion (n = 4) or a duplication (n = 13). All but two alleles fell within a 1.2-Kb region central to the 28-Kb LCR, indicating that >90% of the NAHR took place in that region. We showed that a PRDM9 13-mer recognition sequence is located right in the centre of that region. Our results therefore strengthen the link between this consensus sequence and the occurrence of NAHR.

Citation: Coutton C, Abada F, Karaouzene T, Sanlaville D, Satre V, et al. (2013) Fine Characterisation of a Recombination Hotspot at the DPY19L2 Locus and Resolution of the Paradoxical Excess of Duplications over Deletions in the General Population. PLoS Genet 9(3): e1003363. doi:10.1371/journal.pgen.1003363

Editor: Nancy B. Spinner, University of Pennsylvania, United States of America

Received: October 3, 2012; **Accepted:** January 19, 2013; **Published:** March 21, 2013

Copyright: © 2013 Coutton et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This work was supported by the program GENOPAT 2009 from the French Research Agency (ANR). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: pray@chu-grenoble.fr

Introduction

Several mechanisms have been proposed to cause genomic rearrangements, notably: Non Allelic Homologous Recombination (NAHR), Non Homologous End Joining (NHEJ), Fork Stalling and Template Switching (FoSTeS) and Break-Induced Replication (BIR) [1,2]. NAHR takes place between duplicated sequences with a high sequence identity (usually >95%) located in different genomic regions of the same chromosome [3]. These paralogous sequences or Low Copy Repeats (LCR) tend to generate polymorphic regions with deleted and duplicated alleles called Copy Number Variants (CNVs). The consensual NAHR model predicts that recombinations between LCRs located on the same chromatid result in the production of a deleted allele and a small circular molecule that will be lost by the end of the cell cycle. Recombinations between LCRs located on two distinct chromatids (whether sister-chromatids or chromatids from homologous chromosomes) result in the production of a deleted allele and a

complementary duplicated allele (Figure 1A). In consequence NAHR is expected to produce an excess of deletions over duplications. This has been verified for several NAHR hotspots using sperm typing assays: on average twice as many deletions as duplications were generated *de novo* [4]. One study however describes similar deletion and duplication frequencies at the 7q11.23, 15q11-q13 and 22q11.2 loci, suggesting a predominant inter-chromatid NAHR [5]. This study was carried out by fluorescent in situ hybridization (FISH) which allows the detection of all numerical anomalies occurring at these loci and not only the NAHR-mediated events. This could explain at least part of the discrepancy observed between the two studies, given that a recent study of the *RAII* locus suggested that complex genetic events generate an excess of duplications [6].

As NAHRs events occur at fixed LCRs they tend to be recurrent, and the recombined alleles normally share a common size defined by the distance separating the two LCRs. It is well-established that meiotic recombination events, whether resulting in crossing over or producing

Author Summary

We demonstrated previously that most men with globozoospermia, who produce only round acosomeless spermatozoa and are 100% infertile, had a homozygous deletion removing the totality of *DPY19L2*. We also showed that this deletion occurred by Non-Allelic Homologous Recombination (NAHR). NAHR results in the production of deletions and duplications of regions encompassed by two homologous sequences, normally with a higher occurrence of deletions over duplications. Analysis of public databases at the *DPY19L2* locus paradoxically revealed that, in the general population, duplications were approximately three times as frequent as deletions. Analysis of sperm DNA permits us to quantify *de novo* events that take place during male meiosis. We therefore measured the rates of *de novo* deletion and duplication in the sperm of three healthy donors. As predicted by the NAHR theoretical model and contrary to the allelic frequency observed in the general population, we identified an approximate 2-fold excess of deletions over duplications. We calculated that the measured rate of *de novo* deletion was compensated by evolutionary loss, whereas duplications, not subjected to selection, increased gradually. Purifying selection against infertile homozygously deleted men may be sufficient for this compensation, or heterozygotously deleted men may also suffer a small fitness penalty.

unbalanced alleles through NAHR, are not uniformly distributed along the human genome but occur preferentially at specific hot spots [7–9]. Myers et al. (2008) have characterized a degenerate 13 bp sequence motif (CCNCCNTNNCCNC) that is present in approximately 40% of the identified human crossover hotspots. A three nucleotide periodicity was observed within and beyond the 13-mer core, suggesting a direct interaction with a motif binding protein [10]. Subsequent work strengthened this hypothesis as it has been proposed that PRDM9, a multi-unit zinc finger binding protein expressed mainly during early meiosis in germ cells [11], specifies hotspot usage by binding specifically to this 13 bp consensus motif [12–14]. *PRDM9* was then shown to be highly polymorphic, different alleles seemingly providing preferred targeted recombination hotspots [14]. Berg et al. (2010) measured the recombination rate at ten crossover hotspots, five with a *PRDM9* recognition motif, five without a clear motif. Men with the rarer N allele showed a heavy reduction (>30-fold) at all hotspots, even at those which did not contain an obvious *PRDM9* motif [12]. Further work revealed that specific *PRDM9* alleles activated different hotspots [15]. The direct correlation between *PRDM9* recognition sequence and *PRDM9* genotype however remains elusive, indicating that the rules governing the interaction between *PRDM9* and its targeted sequences must be subtle and complex [12,15].

CNVs and other unbalanced micro recombination events are involved in the aetiology of many human pathologies such as Alpha Thalassemia, Potocki-Lupski Syndrome, Charcot-Marie Tooth, Williams-Beuren syndrome, Prader Willi/Angelman syndrome, and infertility through the production of Y-chromosome microdeletions [16]. Here we focus on the *DPY19L2* locus (12q14.2) which has recently been shown to be linked with Globozoospermia [17], a rare syndrome of male infertility [18] characterized by the presence of 100% round, acosomeless spermatozoa in the patient's ejaculate (MIM #102530). Reports of familial cases pointed to a genetic component to this pathology [19–21], and this assumption was confirmed as a homozygous mutation of *SPATA16* was identified in three siblings [22] and a homozygous missense mutation of *PICK1* was identified in a Chinese patient [23]. We demonstrated recently that *DPY19L2*

was in fact the main locus associated with globozoospermia as 15 out of 20 analysed patients presented a 200 Kb homozygous deletion removing the totality of the gene [17]. *DPY19L2* was described to have arisen, along with three other genes (*DPY19L1*, *L3* and *L4*), through the expansion and evolution of the *DPY19L* gene family from a single ortholog found in invertebrate animals [24]. We then identified *DPY19L2* point mutations and heterozygous deletions and demonstrated that 84% of the 31 globozoospermia patients analysed had a molecular alteration of *DPY19L2* [25]. Others find a slightly lower incidence of *DPY19L2* deletions in globozoospermia patients [26,27]. Comparison of the spermiogenesis between wild type and *Dpy19l2* knock out (KO) mice allowed us to demonstrate that *Dpy19l2* is expressed in the inner nuclear membrane only in the section facing the acrosome, and that it is necessary to anchor the acrosome to the nucleus. This indicates that DPY19 proteins (*DPY19L1-4* in mammals) might constitute a new family of structural transmembrane proteins of the nuclear envelope that likely participate in a function that was so far known to be only carried out by SUN proteins: constituting a bridge between the nucleoskeleton and cytoplasmic organelles and/or the cytoskeleton [28]. In our previous work we had demonstrated that *DPY19L2* was homozygously deleted in a majority of patients with globozoospermia and that this deletion occurred by NAHR between two highly homologous 28 Kb LCRs located on each side of the gene [17]. Strengthening the case for the occurrence of NAHR at the *DPY19L2* locus, heterozygous deletions and duplications have been identified in several large array CGH studies and this locus is classified as a CNV [29–33]. Surprisingly, considering that NAHR is known to generate an excess of deletions, these databases contain a large excess of duplications.

We developed a PCR assay to specifically amplify the recombinant LCRs corresponding to deleted and duplicated alleles allowing the precise localisation of the breakpoints (BP). We observed that all identified BPs clustered in the center of the LCR. We analysed this region and identified a 13-mer PRDM9 pro-recombination sequences in the middle of the hotspot. We also developed a digital PCR assay that enabled us to estimate the rates of *de novo* deletion and duplication at this locus. Contrary to the allelic frequency observed in the general population we measured an approximate 2 fold excess of deletions over duplications. We show that the negative selection against the deleted alleles could explain this apparent paradox.

Results

Estimation of the *DPY19L2* deleted and duplicated alleles' frequencies in the general population and assessment of the PCR assay's sensitivity

The *DPY19L2* CNV was analysed using array CGH data available from web servers [29–33] for a total of 6575 control individuals, mainly from the Database of Genomic Variants (<http://projects.tcag.ca/variation/>). A total of 83 gains and 26 heterozygous losses are reported for the *DPY19L2* CNV in this pool, indicating a threefold excess of duplications over deletions.

We wanted to confirm this result and exclude a potential technical bias towards duplications that could be caused by the presence on chromosome 7 of *DPY19L2P1*, a pseudogene highly homologous to *DPY19L2* [24]. To this end we re-analysed the array CGH data produced for the diagnosis of syndromic mental retardation in Grenoble and Lyon hospitals, and searched for *DPY19L2* deleted and duplicated alleles in this dataset. A total of 1699 array CGH profiles were re-analysed (see Figure S1 for illustration). We identified a total of 15 duplications and 3

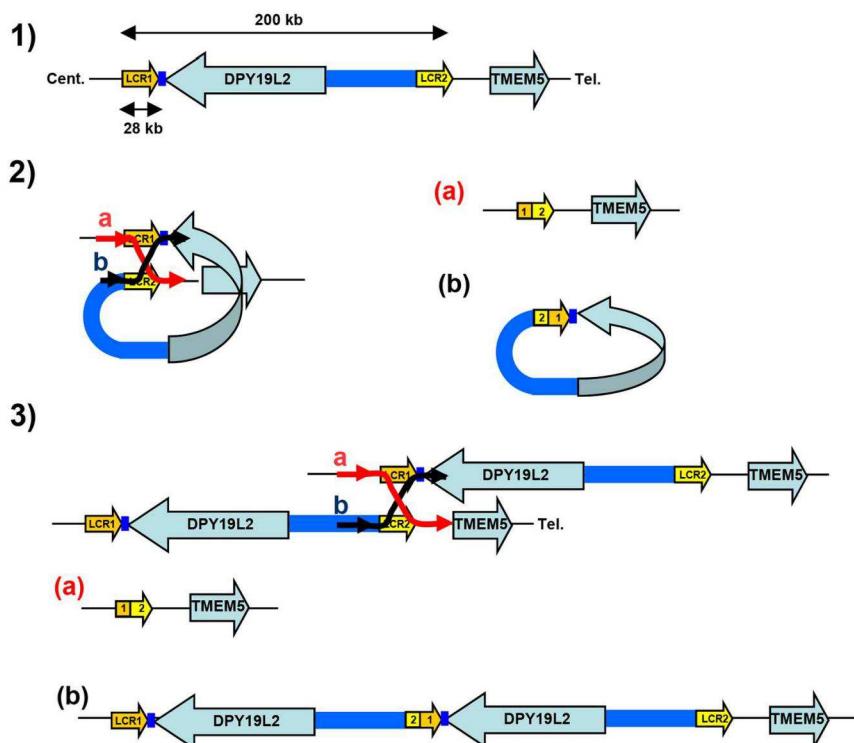
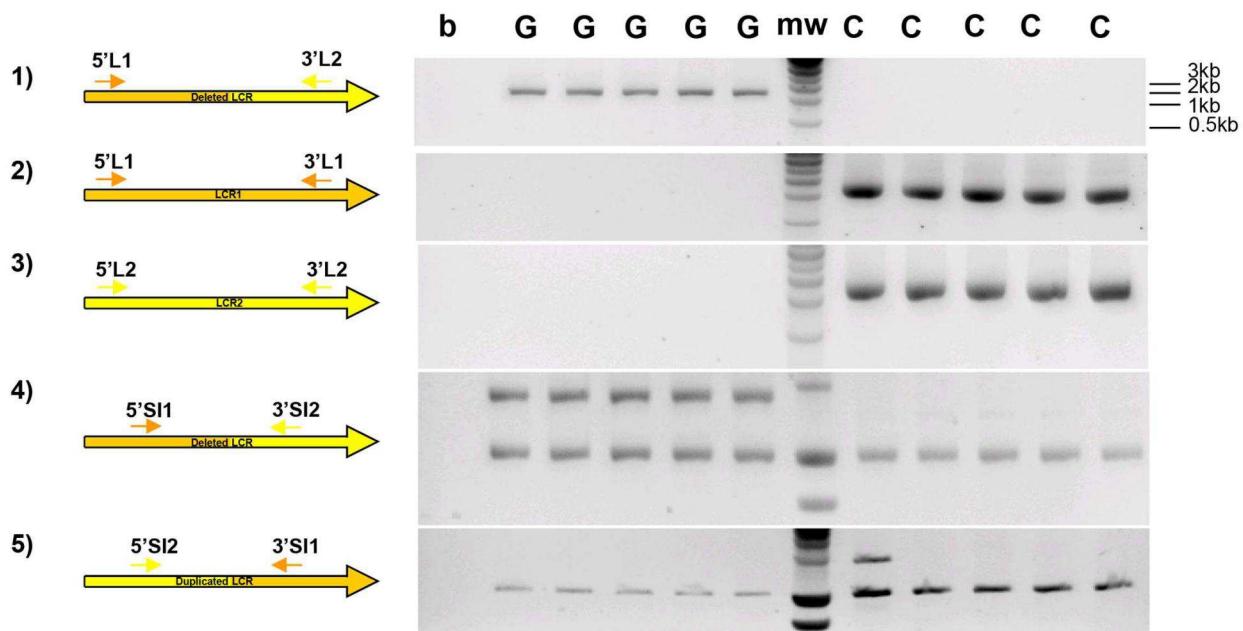
A.**B.**

Figure 1. Strategy and validation of the detection of *DPY19L2* recombinant alleles by PCR. (A) Schematic representation of NAHR at the *DPY19L2* locus. 1) LCR1 and LCR2 correspond to the centromeric and telomeric LCRs respectively. The two LCRs are separated by approximately 200 Kb and each measures 28 Kb. 2) NAHR can occur following the mis-alignment of Low Copy Repeats 1 and 2 located either on 1) the same chromatid and results in the production of a) a deleted allele with a recombinant 1-2 LCR, and b) a small circular molecule with a recombinant 2-1 LCR and the *DPY19L2* gene. This small molecule will not survive through the cell cycle. 3) NAHR can occur following the mis-alignment from two distinct chromatids (whether sister-chromatids or chromatids from homologous chromosomes). This results in the production of a) a deleted allele with a 1-2 recombinant LCR, and b) a complementary duplicated allele with a 2-1 recombinant LCR. (B) Illustration of the specificity of the LCR-specific amplification when amplifying DNA from *DPY19L2* homozygously deleted globozoospermic patients (G) and control individuals (C). 1) Primers specific to the deleted 1-2 LCR yield a 2088 nt fragment in globozoospermic patients only. 2,3) Specific amplification of LCR 1 and 2 is only obtained

from non-deleted controls. 4) Co-amplification of a control locus (bottom band) with a deleted 1-2 LCR-specific sequence. 5) Co-amplification of a control locus (bottom band) with a duplicated 2-1 LCR-specific sequence. A duplicated allele is identified in one control individual (first lane after the molecular weight markers (mw)).

doi:10.1371/journal.pgen.1003363.g001

heterozygous deletions. The recombined alleles were secondarily amplified with the long PCR primers to confirm the validity of the array CGH results. Presence of the deletion could be confirmed by our deletion-specific PCR in the three individuals putatively carrying a heterozygous deletion. DNA from 3 individuals expected to carry a duplicated allele could not be obtained. Ten out of the 12 remaining individuals putatively carrying a *DPY19L2* duplication were amplified by our duplication-specific PCR. For the two individuals that could not be amplified, the duplication was nevertheless confirmed by Multiplex Ligation-dependent Probe Amplification (MLPA). These results show that our reanalysis of the array CGH data did not yield any false positives. Overall reanalysis of these 15 individuals showed that 2 out of 15 recombinant alleles could not be detected by our PCR assay, indicating that the breakpoints of 2/15 recombined alleles fell outside of our amplified region.

We also wanted to obtain an estimation of the frequency of the deleted and duplicated alleles in the general population using our recombination-specific PCR assay. For that we designed primers that amplified a smaller sequence which could be co-amplified with an additional pair of primers (RYR2 primers) used as a positive amplification control (Figure 1B and Table S1). This duplex PCR setup controls for poor DNA quality or technical variations. We analysed 150 control individuals originating from North Africa and 150 individuals of European origin with these two duplex PCRs (for the detection of deleted and duplicated LCRs, respectively). We identified only one heterozygous deletion in an individual of North African origin and two duplications in one European and in one North African individuals.

Overall a total of 8574 individuals have been analysed, including 6575 individuals from array CGH public databases, 1699 individuals from Grenoble-Lyon array CGH data and 300 individuals analysed by recombination-specific PCR. From these cohorts we identified 30 deletions (frequency of approximately 1/290) and 100 duplications (approximate frequency 1/85) (Table S2). These values indicate that the allele frequencies of the recombined deleted and duplicated alleles are 1.7×10^{-3} (95% CI: 1.2×10^{-3} ; 2.5×10^{-3}) and 5.8×10^{-3} (95% CI: 4.7×10^{-3} ; 7.1×10^{-3}), respectively. Confidence intervals (CI) were calculated assuming a binomial model, with binom.test in R.

We note that our PCR-based assay only allows the identification of breakpoints occurring between the selected primers (1392 bp). The location of the breakpoints of each CNV detected by array CGH (an unbiased approach) located in the *DPY19L2* locus was scrutinised to establish if they were located within the LCR and hence were caused by NAHR (Table S3). This analysis shows that 87% of the deletions and 76% of the duplication fell within the LCR limits.

Overall, we believe that our PCR assay permits to identify the majority of recombinations occurring at the *DPY19L2* locus, since: 1) amplification was obtained for all 15/15 globozoospermia patients analysed, and 2) amplification was obtained for 13/15 (87%) recombined array CGH patients.

Determination of *DPY19L2* *de novo* recombination rates by digital PCR

As the previous results consistently showed an excess of duplications over deletions in the general population, we wanted to measure the rates of *de novo* duplications and deletions to verify if

the observed skew was due to the selection of duplications over deletions or if more duplications were produced *de novo*. The rate of genetic events occurring *de novo* can be measured on sperm DNA since each spermatozoon is the product of meiosis and corresponds to a new haploid genome. We first tried to develop a semi-quantitative PCR assay to directly measure the frequencies of deletions and duplications using sperm from control donors (with two copies of *DPY19L2*). The shortest fragment that could provide a reliable specific amplification and amplify the whole breakpoint area was 1392 nt long. Reliable quantitative PCR for fragments longer than 500 nt is difficult with current techniques. We therefore resorted to performing a digital PCR. First, the DNA was serially diluted and distributed in 96-well plates so that approximately 25% of the wells produced an amplicon. The appropriate quantity of sperm DNA was determined by trial experiments for each of the two PCR assays: 50 ng of sperm DNA per well (corresponding to approximately 17,000 copies of chromosome 12, assuming one haploid genome represents 3 pg of DNA) were used for the PCR specific of the *DPY19L2* deletion, and 100 ng per well ($\sim 33,000$ copies under the same assumption) were used for the duplication-specific PCR.

For example, for donor A the deletion-specific PCR produced 26 positive wells. The deletion recombination frequency λ and its 95% confidence interval were then calculated as described (see Methods), resulting in a rate of *de novo* *DPY19L2* deletion for donor A estimated at 1.9×10^{-5} (95% CI: 1.3×10^{-5} ; 2.7×10^{-5}). Similarly, the duplication-specific PCR for donor A produced 23 positive wells, but because there was twice as much starting DNA this results in a rate of *de novo* *DPY19L2* duplication estimated at 8.1×10^{-6} (95% CI: 5.3×10^{-6} ; 1.2×10^{-5}) for this donor (Table 1 and Figure 2).

When pooling the results from the three sperm donors, more robust estimates are obtained: the *de novo* *DPY19L2* deletion rate is estimated at 1.8×10^{-5} (95% CI: 1.4×10^{-5} ; 2.2×10^{-5}), while the *de novo* duplication rate is estimated at 7.7×10^{-6} (95% CI: 6.1×10^{-6} ; 9.7×10^{-6}) (Table 1). There is a significant approximately two-fold enrichment of deletions over duplications at the *DPY19L2* NAHR hotspot.

We investigated whether differential amplification efficiency between the deletion and duplication assays could explain the observed difference between deletion and duplication *de novo* rates. To this end, we performed a control experiment as described (see Methods). No significant difference in amplification efficiency was observed: the deletion-specific control PCR amplified 37 wells, and the duplication-specific PCR amplified 40 wells.

Precise localisation of the recombined allele's breakpoints

Amplification of the LCRs in the deleted alleles had not been achieved in our previous study and the breakpoint minimal region had only been narrowed down to a 15 Kb region within the LCRs (8). Here we designed and validated PCR primers that amplify a 2 Kb product in deleted individuals only (Figure 1B). We quickly realised that mapping the breakpoints was complicated by the fact that many of the nucleotides that differed between LCR1 and LCR2 in the reference sequence were in fact not specific to one or the other LCR. Since mapping the breakpoints requires markers specific to each LCR, we decided to amplify and sequence the 2 Kb breakpoint region for each LCR in 20 control individuals.

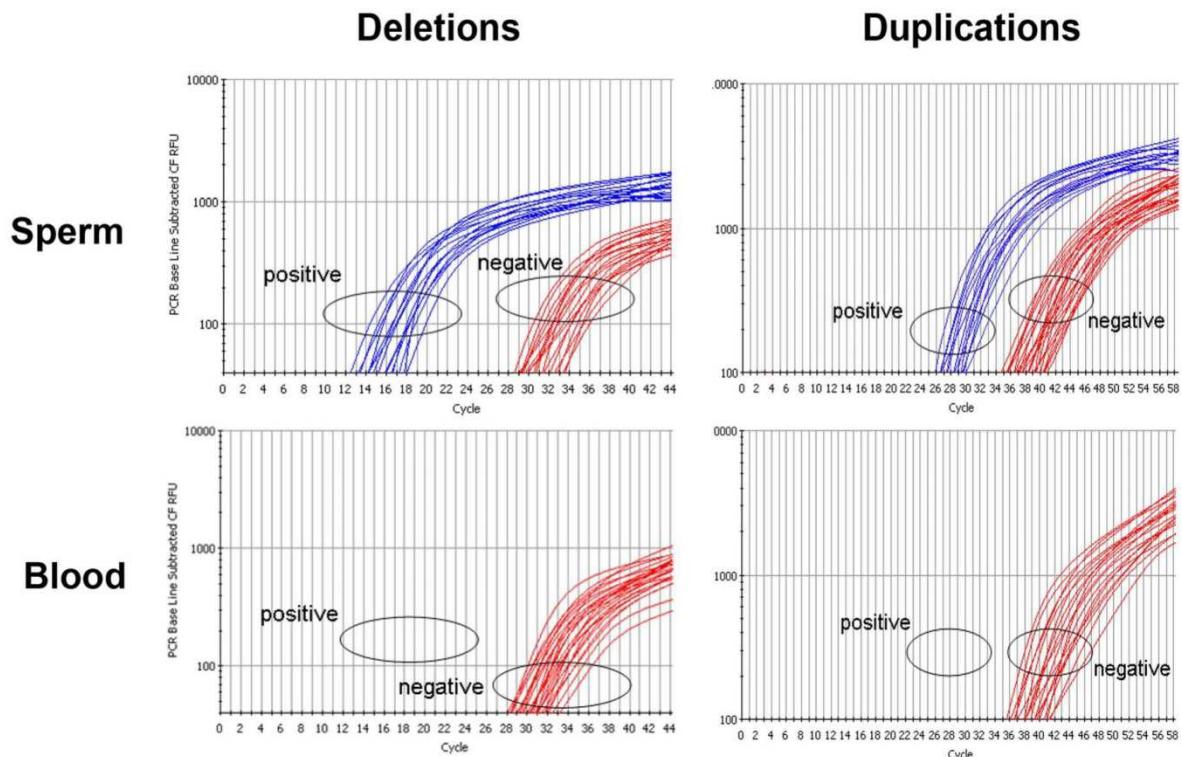
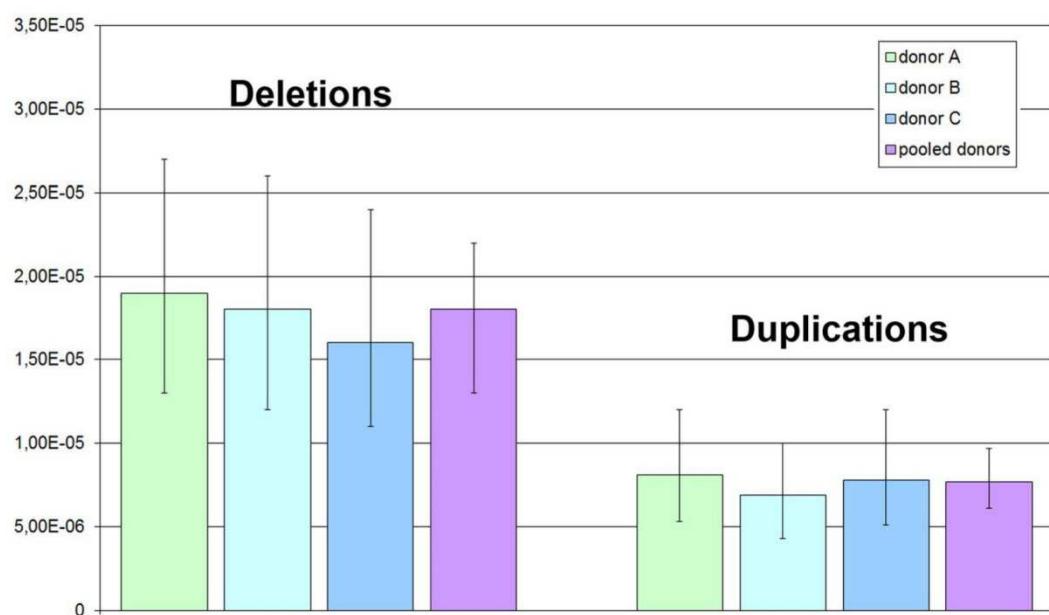
A.**B.**

Figure 2. Rate of *de novo* deletion and duplication events occurring at the DPY19L2 NAHR hotspot determined by digital PCR on sperm from 3 control donors. (A) Illustration of PCR results obtained by real time PCR. The left plots show amplification profiles obtained with primers specific to the recombinated deleted LCR, the right plots show profiles obtained with the duplication-specific primers. No amplification was observed with either pairs of primers from 200 ng of somatic (blood) DNA, indicating that the NAHR did not occur during mitosis. Sperm DNA was diluted in order to obtain a positive amplification in approximately 25% of the wells. (B) The number of positive wells allowed estimating the frequency of *de novo* deletion and duplication events in three control sperms. Error bars represent 95% CIs.

doi:10.1371/journal.pgen.1003363.g002

Table 1. Frequency of deleted and duplicated alleles in sperm from three control donors.

	Deletion				Duplication			
	donor A	donor B	donor C	Pooled	donor A	donor B	donor C	Pooled
Positive wells	26	25	23	74	23	20	22	65
Nb of recombinants	30	29	26	85	26	22	25	74
Total nb of alleles		1.6E+6		4.8E+6		3.2E+6		9.6E+6
λ	1.9E-5	1.8E-5	1.6E-5	1.8E-5	8.1E-6	6.9E-6	7.8E-6	7.7E-6
95% CI inf	1.3E-5	1.2E-5	1.1E-5	1.4E-5	5.3E-6	4.3E-6	5.1E-6	6.1E-6
95% CI sup	2.7E-5	2.6E-5	2.4E-5	2.2E-5	1.2E-5	1.0E-5	1.2E-5	9.7E-6

doi:10.1371/journal.pgen.1003363.t001

To achieve the specific amplification of LCR 1 and 2 we had to rely on the reference human genome sequence to design the primers. We had no way of confirming that the targeted LCRs were specifically amplified in control individuals, but no amplification was obtained when assaying twenty homozygous deleted patients, vouching for the specificity of the primers. We then amplified and sequenced LCR1 and 2 from a total of 20 control individuals: 10 of North African origin and 10 of European origin. Thirty-four nucleotides were indicated as specific to either LCR 1 or 2 in hg19 reference sequence but 14 of these were in fact arbitrarily found in the two LCRs (Table S4): we consider that these are non-LCR-specific single nucleotide polymorphisms (SNPs). The remaining 20 nucleotides were indeed LCR-specific: these 20 fixed markers were used to map the recombination breakpoints, and we used the 14 SNPs to establish a haplotype map of the patients' deleted alleles (Table S4).

Allele-specific amplification of the deleted LCR was carried out on 15 homozygotously deleted globozoospermia patients. Each amplification yielded a single 2088 bp product, while the PCR was negative for all the healthy controls tested ($n=20$). We sequenced all the amplicons in order to better characterize the breakpoint region. Fourteen out of the 15 patients analysed were homozygous for all markers tested. Three different breakpoints (BPs) were identified based on the presence of the 20 invariant markers. The three recombination events (BP1–3) were included in a 1153 bp maximal region (Table S4 and Figure 3). The breakpoints could not be mapped more accurately for lack of nucleotides specific to each LCR. One patient was heterozygous for markers 13 and 14, indicating that this patient was heterozygous and carried two different deleted alleles (BPs 2 and 3). If we consider that the other 14 patients carried two recombinant deleted alleles each, we have a total of 14 alleles with BP1 (between markers 17 and 18), 13 alleles with BP2 (between markers 18 and 24) and 3 with BP3 (between markers 25 and 28) (Figure 3B). The 14 identified SNPs were then used to map the different haplotypes in patients presenting the same breakpoint (Table S4). This shows the presence of a total of 7 distinct haplotypes, indicating that at least 7 recombination events are at the origin of our patients' pathology (15 patients). We also observe that 5 patients with BP2 have the same haplotype and that two groups of 3 patients with BP1 have the same haplotype, suggesting the presence of several founding deletions in our patients' population. This is not surprising as all our patients came from the same region (Tunis area) and a majority had related parents (often first cousins).

One and three deletions were identified respectively in the 300 individuals analysed by PCR and in the 1699 Grenoble-Lyon array CGH patients group. There were 3 occurrences of BP2 and 1 of BP3. Overall, including the globozoospermia patients, a total

of 34 somatic deleted alleles were examined, resulting in the detection of three different recombination breakpoints. Fourteen alleles (41.2%) had a deletion between markers 17 and 18 (BP1), 16 alleles (47.0%) between markers 18 and 24 (BP2), and 4 alleles (11.8%) were recombined between markers 25 and 28 (BP3) (Figure 4 top left).

Two and fifteen genomic duplicated alleles were detected respectively in the 300 control individuals analysed by PCR and in the Grenoble-Lyon array CGH patients. Only 12 duplicated alleles could be sequenced (for lack of DNA from 3 control subjects and because two of the subjects had breakpoints falling outside the range of the duplication-specific PCR). Seven alleles (58.3%) corresponded to the reciprocal alleles of deletion 2 (BP2) with a recombination between markers 18 and 24, and 5 alleles (41.70%) corresponded to the reciprocal alleles of deletion 3 (BP3) with a recombination between markers 25 and 28 (Figure 3B).

The position of the meiotic recombination events (deletion and duplication) obtained from three sperm donors were also characterized by DNA sequencing. A total of 74 *de novo* deleted alleles and 65 *de novo* duplicated alleles were sequenced. All recombination events (from both duplications and deletions) clustered into five breakpoints (Figure 4). Two of them are new (BP4 and BP5) i.e. not previously identified in globozoospermic patients or in the CGH control cohort. The number and percentages of deleted and duplicated breakpoints respectively are: BP1: 2 (2.7%) and 4 (6.1%); BP2: 56 (75.7%) and 38 (58.5%); BP3: 10 (13.5%) and 13 (20%); BP4: 2 (2.7%) and 3 (4.6%) and BP5: 4 (5.4%) and 7 (10.8%) (Figure 4). BP2 is by far the most frequent BP, followed by BP3, explained by the fact that these two breakpoints correspond to the largest regions. Interestingly in sperm, the distributions of the deleted and duplicated breakpoints are quite similar. This is logical as the duplicated alleles are expected to be the reciprocal alleles of some of the deleted alleles. In genomic DNA the correlation is not as good, and we note that the frequency of the deleted BP1 is particularly high. Most of the deleted alleles come from globozoospermia patients (and a few detected in CGHarray patients) most of whom were recruited in Tunis. As suggested by the shared haplotypes observed between some deleted patients (Table S4) a founder's effect is likely to account for some of the most frequent deletions, in particular BP1.

PRDM9 genotyping of the sperm donors

Sequencing of the PRDM9 ZF array was performed in the 3 sperm donors. All three donors were homozygous for the A allele which represents over 90% of the European alleles. It comprises 13 copies of the 84-bp ZF repeat that binds the 13-bp Myers recombination motif [12,14]. This result is concordant with the ethnicity of the donors.

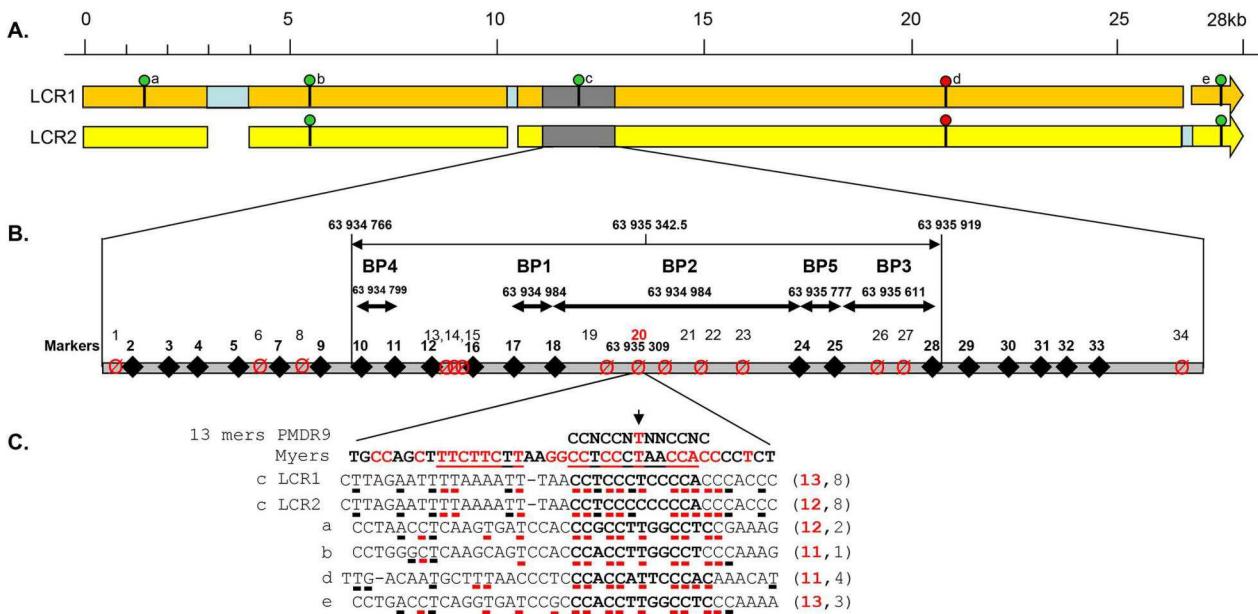


Figure 3. Details of the *DPY19L2* LCR1 and 2 and of the NAHR hotspot. (A) Detailed scaled representation of the 28.2 Kb LCR 1 (orange) and 27 Kb LCR2 (yellow). Pale blue rectangles correspond to sequences specific to one of the LCRs facing a gap in the other LCR. The presence of a 13 bp consensus PRDM9 recognition site (CCNCNTNNCCNC) on LCR1 or LCR2 is indicated by a green circle when identified on the forward DNA strand and by a red circle when identified on the reverse strand (GTGGNNAGGGTGG). The LCR arrows point toward the chromosome 12 telomere. (B) The analysed recombination region is represented in grey. The positions of LCR-specific markers (diamonds and bold numbering) and variable nucleotides (crossed circles) are represented. Details of the markers' sequences and localisations are indicated in Table S2. The five identified breakpoints (BP1–BP5) are shown as double arrows. One PRDM9 consensus sequence is localised in the centre of BP2, the central and most frequent breakpoint. (C) The central nucleotide from the consensus sequence corresponds to one of the identified SNPs (snp 20). A perfect match for the consensus sequence is present on LCR1, while the central thymine is replaced by a cytidine in LCR2. The 39 nt surrounding the 5 matches to the PRDM9 consensus sequence identified in LCR1 and 2 (sites a–e) are compared with the consensus sequence described in Myers et al [8,11]. Highly conserved nucleotides are red. For each locus the number of nucleotides identical to the consensus sequence is indicated on the right.

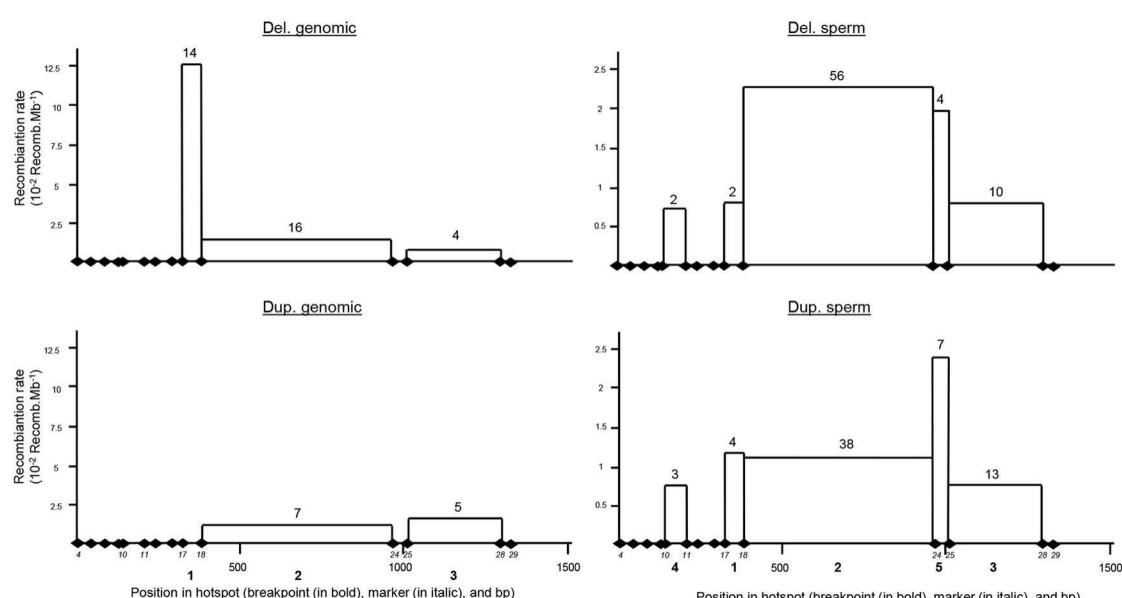


Figure 4. Distribution of deleted and duplicated breakpoints observed from somatic DNA (left two panels) and sperm DNA (right two panels). Somatic deletions were identified from sequence analysis of 15 homozygous deleted patients and two heterozygous deleted control individuals. Somatic duplications were identified from 12 positive control individuals. Data from sperm were pooled from three control donors.

Detailed analysis of LCR1 and 2

A comparison of the two LCRs is presented in Figure 3A. The illustration was produced from the results of a megablast search (<http://blast.ncbi.nlm.nih.gov/Blast.cgi?PROGRAM>). All identified recombined alleles ($n = 185$) cluster between markers 10 and 28 within a 1153 bp region. This recombination hotspot is roughly located in the middle of the 28 Kb LCR (Figure 3A). Five 13 bp PRDM9 consensus recognition sites (CCNCCNTNNCCNC) are present along the LCR (Figure 3A). One of these sites is located in the centre of the 1153 bp hotspot (less than 35 nt away from the hotspot median position) (Figure 3B). We note that the most central BP (BP2) which encompasses the 13 bp site, represents 117 out of 185 recombined alleles or 63% of the detected recombined alleles (Figure 4). Given that five PRDM9 consensus recognition sites are found within the 28 Kb LCR1 sequence, the probability that a site would occur by chance less than 35 bp away from the centre of the hotspot is $1 - (1 - 5/28000)^{70} = 0.012$.

The Thymine at the centre of the consensus recognition site (CCNCCNTNNCCNC) was present only in the reference sequence of LCR1. Sequence analysis of our control individuals showed that this nucleotide was in fact a SNP (Marker 20 in Figure 3 and Table S4) with a T allele frequently found in both LCR1 and LCR2 (Table S4). In our globozoospermia patients we observed that all patients with the BP1 (with marker 20 located after the breakpoint thus on LCR2 sequence) have the T allele, indicating the presence of a T allele on LCR2 of the original unrecombined allele (Table S4). Conversely patients with BP3 (with marker 20 located before the breakpoint thus on LCR1 sequence) have the C allele indicating the presence of a C allele on LCR1 of the original unrecombined allele. All patients with BP2 have the C allele. As marker 20 is located within the breakpoint maximal sequence we can only conclude that at least a C allele was present on either LCR1 or LCR2 of the original unrecombined allele.

We sequenced LCR1 and LCR2 of our three sperm donors and realised that all were homozygous for the C allele at both LCR1 and LCR2, suggesting that the presence of the thymine in the CCNCCNTNNCCNC consensus sequence is not necessary to initiate recombination in the *DPY19L2* LCR central region. Myers et al. [8,11] indicated that although the core 13-mer recognition sequence was associated with recombination hotspots, the recognition motif extended beyond the core sequence with preferentially associated nucleotides identified within a 39 bp sequence encompassing the PRDM9 core sequence. We therefore aligned this extended motif with the sequence of the 5 PRDM9 motifs identified within the LCR (Figure 3C). We observe a good correlation within all 5 sequences, especially for the nucleotides that had been shown to be significantly associated with hotspots (indicated in red in Figure 3C). We also observe that the sequence central to our recombination hotspot (motif c) presents the highest homology (53%) with Myers' extended recognition sequence (Figure 3C).

Discussion

It appears paradoxical that *de novo* deletions are produced twice more frequently than *de novo* duplications during meiosis, while duplicated alleles are three times more frequent than deleted alleles in the general population. We investigated whether this could be explained parsimoniously through the combined effects of selection and mutation. Men carrying a homozygous deletion of *DPY19L2* are 100% infertile, but currently there is no evidence that a heterozygous deletion of *DPY19L2* causes a phenotype or that homozygous women are affected. Additionally, the deleted

allele is rare. Under these assumptions, according to the General Selection Model (GSM), natural selection results in a decrease in the frequency of the deleted allele of approximately $q^2/2$ per generation, where q is the frequency of the deleted allele (see Methods). Given that the deleted allele has a frequency of 1.7×10^{-3} (95% CI: 1.2×10^{-3} ; 2.5×10^{-3}) in the general population according to our combined control data, the GSM predicts that this frequency decreases by 1.5×10^{-6} (95% CI: 7×10^{-7} ; 3.1×10^{-6}) per generation. Conversely, deleted alleles are produced *de novo* by NAHR at an estimated rate of 1.8×10^{-5} (95% CI: 1.4×10^{-5} ; 2.2×10^{-5}) according to our digital PCR data. Assuming the allele frequency is at an equilibrium, these two rates should balance out. In fact they are somewhat similar but the 95% confidence intervals do not overlap. However the CIs only represent the uncertainty induced by the sampling procedure, i.e. the fact that the allele frequency and recombination rate are estimated from a sample of the whole population: they do not take into account experimental biases or imperfections that may exist at various steps. In addition, the GSM is a theoretical model that assumes an infinite population size and panmixia, whereas in practice stochastic effects and population structure (including for example any potential consanguinity or local founder effects) come into play. These could result in a significantly increased impact of purifying selection on the deleted allele, so that the frequency decrease resulting from selection and the *de novo* production of deleted alleles through NAHR may in fact cancel out.

Alternatively, it is possible that heterozygously deleted men suffer a fitness penalty. This can be taken into account within the GSM, and one can calculate the relative fitness of heterozygous individuals such that the GSM-predicted decrease of the deleted allele's frequency compensates the measured NAHR-induced production of new deleted alleles. In fact, assuming women are not affected, a 98% relative fitness of heterozygous men is sufficient (see Methods). Such a small effect could have easily remained undetected, and this scenario cannot be ruled out. This potential selection could be caused by meiotic segregation distortion as was observed for the T/t mouse locus [34]. Finally we only studied the recombination rate in male germ cells and we cannot exclude the possibility that the frequency and ratio of deletion and duplication might be different in female gametes.

All in all we believe the rates are reconcilable: whether the discrepancy observed when assuming heterozygous individuals have no phenotype is due to imperfections in the data and/or to population structure which disrupts the theoretical GSM model, or whether heterozygously deleted men suffer a small fitness penalty, we propose that the frequency decrease due to purifying selection and the *de novo* production of deleted alleles through NAHR cancel out, and that the frequency of the deleted *DPY19L2* allele is today at a selection-recombination equilibrium in the population. On the other hand, to the best of our knowledge there is no evidence that the duplicated *DPY19L2* allele is either deleterious or advantageous. We therefore assume that the duplicated *DPY19L2* allele is not under selection, so its frequency can increase in the population by recurrent NAHR. This resolves the paradox.

Liu and colleagues (2011) proposed that the frequency of NAHR occurring between two paralogous LCRs was proportional to the LCR length and sequence homology but inversely proportional to the distance between the LCRs [6]. The authors logically proposed that the probability of ectopic chromosome synapsis increases with LCR length, and that ectopic synapsis is a necessary precursor to ectopic crossing-over. Here we measured that the average rate of *de novo* recombination (deletion plus duplication) by NAHR at the *DPY19L2* recombination hotspot was 2.6×10^{-5} . This rate is higher than what was measured at

other loci such as the Williams-Beuren syndrome (WBS) locus or the LCR17p locus [4]. In our case the relatively small LCR size (28 Kb) is compensated by the proximity of the repeats (200 Kb) compared with much greater distances separating the paralogous LCRs for WBS and LCR17p. *DPY19L2* LCR1 and 2 also present a very high sequence identity (98%) which could also reinforce their synapsis and recombination. Our results are in agreement with previous work suggesting that the distance separating the two LCRs, as well as their sequence homology and length are parameters likely influencing recombination frequency.

We observed that >90% of *DPY19L2* NAHR events occurred within a 1.2 Kb region located in the centre of the 28 Kb LCR, suggesting the presence of a pro-recombination sequence within this hotspot. Myers et al. (2008) have characterized a degenerate 13 bp sequence motif that is present in approximately 40% of the identified human hotspots and which constitutes a *PRDM9* recognition signal [12–14]. *PRDM9* codes for a zinc finger array which catalyses the trimethylation of the lysine 4 of histone H3 (H3K4me3) [11]. This *PRDM9*-mediated post-translational histone modification likely initiates the recruitment of the recombination initiation complex, creating a favourable chromatin environment and allowing access of SPO11 to the DNA. SPO11 then initiates the formation of double-strand breaks (DSBs) which will be repaired by homologous recombination [35]. Here we identified a hotspot of NAHR located in the centre of a 28 Kb LCR. We showed that a *PRDM9* 13-mer recognition sequence is present at the epicentre of all the identified breakpoints. We however realised that the thymine, central to the 13-mer motif (CCNCCNTNNCCNC), was a T/C SNP, each nucleotide being found arbitrarily within LCR1 or LCR2. Following this observation one can wonder if recombination events at the *DPY19L2* hotspot occur preferentially in the presence of fully matching *PRDM9* 13-mer alleles. We measured the frequency of *de novo* recombination in sperm from three donors. As it happens, sequencing revealed that all three were homozygous for the C allele on both LCR1 and LCR2. This indicates that, at this locus, the presence of the 13-mer exact match is not necessary to initiate recombination. This observation is concordant with what was described previously at different loci and confirms that *PRDM9* tropism for the 13-mer recognition site might not be very strong and/or that other mechanisms also intervene in the choosing of double strand break localization [12,15]. One explanation can come from the extended sequence surrounding the 13-mer motif. Myers and colleagues (2008) [10] described a 39 bp pro-recombination sequence encompassing the 13-mer motif. We observe a greater than 50% sequence identity for the complete 39-mer sequence, indicating that a good match to the extended motif might be at least as important as a perfect match of the core 13-mer motif.

We identified a total of 5 distinct breakpoints (BP), all localized within a 1.2 Kb region located in the centre of the 28 Kb LCR. Others have described the localization of the deletions of globozoospermia patients [22]. They described a total of 9 separate BPs in the *DPY19L2* LCR. Looking at the precise localizations of the described BPs, we noticed that the nucleotides used to delimit BPs 1–6 in that study are in fact nucleotides that we identified as SNPs (markers 19–23 and 26), which strongly questions the validity of the BP localization in that study. Reanalyzing the presented data and using LCR-specific markers only, we conclude that Elinati et al. (2012) BPs 1, 2, 4, 5, 6 fall within the boundaries of “our” BP2 and that “their” BP3 corresponds to “our” BP3. This illustrates the difficulty in precisely identifying the localization of BPs and demonstrates that this can only be achieved with a high level of confidence after confirmation

that the markers used to define the BP positions are indeed locus-specific. From our reanalysis, Elinati et al. (2012) identified deletions in 27 globozoospermia patients, 23 had our BP2, one had BP3 and one had a BP that fell just outside of our studied region. These results thus confirm the importance of the recombination hotspot described here. Two additional BPs (BP8 and 9) were also identified in Elinati's study which fell well outside of our recombination hotspot. This might constitute a second, less frequent recombination hotspot within the LCRs. We noticed that these two BPs are located 1200 bp telomeric from the 13-mer *PRDM9* site d (as indicated in Figure 3A). Thus this second putative hotspot is further away from a consensus 13-mer motif than our hotspot (the greatest distance of the BPs we identified from the 13-mer is 600 bp), but we can question again the accuracy of the positioning of these two breakpoints. Here, while analyzing the array CGH recombined patients we identified two recombinant alleles which did not fall within our studied BP area. It is possible that these recombination events are also located within this second putative hotspot.

With the *DPY19L2* locus we believe that we have a good model to study the effect of the *PRDM9* recognition site on NAHR. We plan to accurately position the yet uncharacterized BPs in relation to other *PRDM9* sites. We are also currently screening an anonymized sperm bank to identify donors that are homozygous for the central 13-mer *PRDM9* recognition T allele and/or who present rarer *PRDM9* alleles to investigate how the recombination rate is affected by both the *PRDM9* genotype and the extended *PRDM9* recognition motif. We believe that although much work remains to be done, our study illustrates and consolidates the hotspot models described previously. In a moving environment we can imagine that the central region of the LCR will have the most opportunities to synapse with its paralogous sequence. The presence of an extended *PRDM9* recognition motif in the centre of the LCR then very likely contributes to DSB and NAHR. The combination of these parameters therefore probably explains why approximately 90% of the breakpoints occurred within a few hundred nucleotides from the most centrally located *PRDM9* recognition site.

Materials and Methods

Ethics statement

All patients, family members and anonymous DNA and sperm donors gave their written informed consent, and all national laws and regulations were respected. Ethical approval was obtained from Grenoble CHU review board.

Information on patients and control individuals

We previously reported that 15 out of 20 patients with globozoospermia had a homozygous deletion of the *DPY19L2* region [17]. These patients are included in this study. All patients are unrelated apart from two who are brothers. All patients originated from North Africa (Tunisia, n = 12; Morocco, n = 2 and Algeria, n = 1).

Array CGH data from a total of 1699 control anonymous individuals were re-analysed. These analyses had been carried out as a diagnosis for syndromic mental retardation either at Grenoble or Lyon's hospital. As our aim was to identify *DPY19L2* centred CNVs in this cohort of patients and since there is no known link between *DPY19L2* and mental retardation, we believe that this cohort can serve as a control in this study. All individuals agreed to the anonymous use of their DNA in genetic studies and signed an informed consent. The fertility and ethnic origin of these individuals was not documented. All were French citizens. We

estimate that in excess of 90% of these individuals are of European origin and that the vast majority of the others are of North African origin.

There was no gender selection but this cohort contained approximately 2/3rd of males.

Array CGH results from these patients were scrutinized for the *DPY19L2* region.

Three hundred control individuals were analysed independently with recombinant *DPY19L2*-specific PCR (deleted and duplicated) to identify deleted and duplicated alleles. One hundred and fifty individuals originated from North Africa (Algeria, Morocco, and Tunisia) and 150 originated from Europe. All individuals gave their informed consent to constitute an anonymous DNA bank. Non-recombined LCR1 and 2 of twenty of these individuals were amplified and sequenced to identify LCR-specific SNPs. There was no gender selection and this cohort contained a similar number of males and females.

Lastly the *DPY19L2* CNV was also analysed from array CGH data available from web servers [29–33] for a total of 6575 control individuals, mainly from the Database of Genomic Variants (<http://projects.tcag.ca/variation/>). Most of these individuals originated from Europe (75%), Africa (18%) or Asia. Individual CNV could however not be linked to a particular individuals and its geographical origin. The location of the breakpoints of each CNV located in the *DPY19L2* locus was scrutinised to establish if they were located within the LCR and hence were caused by NAHR (Tables S2 and S3).

DNA extraction

Genomic DNA was extracted either from peripheral blood leucocytes using a guanidium chloride extraction procedure [36] or from saliva using Oragene DNA Self-Collection Kit (DNAgentech, Ottawa, Canada).

Sperm DNA was extracted from 2 ml of semen which were transferred to a 25 ml Falcon Tube (BD Biosciences). Ten ml of PBS was added, mixed gently and centrifuged at 3,000 rpm for 5 minutes. Supernatant was discarded and the pellet was resuspended again in 10 ml of PBS, mixed and centrifuged as before. Pellets were then resuspended in 1 ml digestion buffer (NTE buffer 0.5 mM NaCl, 10 mM Tris-HCl pH 7.5, 5 mM EDTA, pH 8 (100:10:1), 0.4% SDS), 25 µl of 10 mg/ml proteinase K solution (Sigma) were added and the mix was incubated overnight at 42°C with occasional mixing. Three hundred microliters of the contents of each Falcon tube were transferred into SafeLock tubes (Eppendorf). An equal volume of phenol/chloroform/isoamyl alcohol (25:24:1) was added and mixed gently until emulsified. The tube was centrifuged at 3,000 rpm for 5 minutes. We repeated this process a second time, adding an equal volume of chloroform/isoamyl alcohol. The upper aqueous layer was transferred into a clean Eppendorf tube. The aqueous layers from the two phenol/chloroform extractions were combined and an ethanol precipitation was performed: 25 µl 3M sodium acetate pH 5.4 and 1 ml 100% ethanol were added to the aqueous phase, mixed gently and centrifuged as before. The pellets were washed twice with 70% ethanol and finally resuspended in 300 µl TE buffer (10 mM Tris-Cl pH 8.0, 0.1 mM EDTA, pH 8.0 (10:1)) by incubating overnight at 50°C with gentle shaking.

Information about the sperm donors

DNA was extracted from three fertile anonymous donors of European origin with normal sperm parameters and of similar age (between 30 and 35 years old). In each case a spermogram was realised according to WHO's 2010 guidelines [37]. Sperm

concentration ranged between 60–120 × 10⁶ spz/ml, with, in each case less than 1 × 10⁶ leucocytes/ml. We therefore considered that the presence of this small percentage of leucocytes had a negligible effect on the quantification of the sperm (only) DNA and on the ensuing calculations.

A molecular analysis was carried out to determine PRDM9 ZF array genotype. PCR amplification and sequencing of the PRDM9 ZF array were performed using primers and protocols as described previously ([12]. PCR and sequencing primer sequences are listed in the Table S1.

Amplification and sequencing of the LCRs

All primers were designed to have at least their 3' nucleotide specific to the LCR of interest (Table S1). PCR primers were designed to amplify specifically LCR 1 or LCR 2, in order to perform a sequence comparison of the two LCRs. For each recombined LCR locus (resulting from deletion or duplication), two sets of specific primers were designed (Figure 1B). The external primers (long primers) were used for sequencing analysis. They were also used as an outer primer for the digital PCR that was devised to measure the rate of *de novo* recombination in sperm. The short internal primers (SI) were used in duplex with *RYR2* primers that were used as a positive amplification control. These two sets of primers were used to detect the presence of recombined alleles in the 300 control individuals. They were also used as inner primers for the digital PCR.

PCR amplification was carried out on an Applied Biosystems genAmp 2700 thermocycler. Due to the high sequence homology between the two LCRs, the use of a precise annealing temperature was critical. The same thermocycler had to be used throughout the study as small variations in block temperature could introduce discrepancies in the amplification. Both the long and short PCR cycles were preceded by a 7 minutes denaturation at 95°C and followed by a 10 minutes elongation at 72°C. The specific annealing temperature of each primer set is indicated in Table S1. Thirty-five cycles were carried out for the long PCRs, with 30 seconds of denaturation at 95°C, 30 seconds of annealing and 2 minutes of elongation at 72°C. Forty-five cycles were carried out for the short PCRs with 30 seconds of denaturation at 95°C, 20 seconds of annealing and 2 minutes of elongation at 72°C.

We performed the long and short PCRs in 1× Takara Ex Taq buffer (Takara), 250 µM dNTPs (Takara dNTP mixture), 300 nM each primer, 1 unit Takara Ex Taq (Takara) with 200 ng of somatic DNA in a total volume of 25 µl.

All sequences (native LCR 1 and 2 and deleted and duplicated LCRs) were carried out with BigDye Terminator v3.1 (Applied Biosystems Courtaboeuf, France) on an ABI 3130XL (Applied Biosystems, Courtaboeuf, France).

Oligonucleotide array CGH was performed with the Agilent 10K or 180K Human Genome CGH Microarray (Agilent Technologies, Santa Clara, CA, USA) (Hospices Civils de Lyon array CGH Platform and CHU Grenoble array CGH Platform). Extracted DNAs were labelled according to the instructions of the supplier and incubated overnight. The samples were purified and hybridised as described previously [17].

Graphical display and analysis of the data were performed with the Agilent DNA Analytics software version 4.0.81 (statistical algorithm: ADM-2, sensitivity threshold: 2.5, window: 0.5). A value of zero represents equal fluorescence intensities between sample and reference DNA. Copy-number losses shift the value to the left (≤ -1), and copy-number gains shift it to the right (≥ 0.58).

The design of the MLPA probes, MLPA reaction and data analysis were performed according to the recommendation of the

MRC-Holland synthetic protocol (www.mlpa.com) and as described in Coutton et al. (2012) [25].

Sperm assay design and digital PCR for sperm NAHR breakpoints mapping

We designed two nested LCR-specific PCRs as described in the PCR section. In addition, we designed a TaqMan dual labeled probe (Table S1) to allow the second step of the nested PCRs to be run on Biorad iCycler IQ real time PCR detection. We tested each recombinant-specific combination of primers for specificity and sensitivity on negative and positive (*DPY19L2* deleted and duplicated) control blood DNA (Figure 1B). Each of the two rearrangements was assayed on DNA extracted from three unrelated sperm donors. Each donor was confirmed to carry two copies of *DPY19L2* by MLPA analysis (data not shown). We note that our assay will not distinguish triplications of the *DPY19L2* locus, which are likely to occur at extremely low frequencies.

We performed the first LCR-specific PCR (long PCR) in 1× Takara Ex Taq buffer (Takara), 250 μM dNTPs (Takara dNTP mixture), 300 nM of each primer, 1 unit Takara Ex Taq (Takara), using sufficient copies of template DNA to give approximately 24 positive wells per 96-well plate (exact quantities determined empirically by successive dilutions) and 2.5 mM MgCl₂, in a total volume of 50 μl. Following thermal cycling we incubated 10 μl of the long PCR products with 5 μl of Exosap-IT PCR Clean-up Kit (GE Healthcare) for 15 min at 37°C to digest the long PCR primers followed by enzyme inactivation at 80°C for a further 15 min. Two μl of 10× diluted long PCR products was used as a template in the second PCR (short PCR). In the short PCR we used the same concentrations of buffer, dNTPs, primers and enzyme as in the Long PCR, but the total volume was 25 μl and we added a dual-labeled probe (final concentration 250 nM; Eurofins MWG Operon) (Table S1). To map the locations of breakpoints we re-amplified wells that we had previously identified as positive in the long PCR plate using the short primers and sequenced these amplicons.

The quantity of input sperm DNA was experimentally determined by serial dilutions to obtain approximately 24 positive breakpoint-specific amplifications per 96-well plate. The number of positive amplifications was then counted to estimate the number of recombinants in the input sperm. Each well contains a sample drawn from the input DNA without replacement, hence the number of recombinants in a given well is appropriately modeled using a hypergeometric distribution. We note that this hypergeometric distribution has often been approximated in the literature by Poisson (6, 22) or binomial (23) distributions, but although such approximations are acceptable we find no need for them in this study, as the direct calculation is simple. Indeed, using the hypergeometric distribution the probability that a well contains no recombinants is:

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

$$\frac{W!(N-W)!}{W!(N-W)!}$$

where N is the total number of copies of chromosome 12 in the input DNA (i.e. 1.6×10^6 for the deletion assay and 3.2×10^6 for the duplication assay, see Results section on digital PCR), W = N/96 is the number of copies per well, and R is the total number of recombinants. The value of R such that this probability is closest to the observed ratio of negative wells (i.e. one minus the fraction of wells that produced a positive amplification) is easily found by

tabulation. This leads to an estimation of the *de novo* recombination rate $\lambda = R/N$, and a 95% confidence interval is calculated by modeling the initial dilution to obtain the input DNA using the binomial distribution (with binom.test in the R stats package, <http://www.r-project.org>).

In order to evaluate the amplification efficiency of our duplication/deletion assays, we used as positive controls genomic DNA from one heterozygous duplicated individual and from one heterozygous deleted individuals. We believe that this type of control is more accurate than the use of cloned recombinant deleted and duplicated alleles as this reduces dilution factors. More importantly it reproduces faithfully the possible inhibitions due to the presence of the over majoritarian non-target genomic DNA or the potential amplification of homologous sequences that are present in the actual quantifying experiments.

The DNA concentration was measured by Nanodrop (ThermoScientific) and DNA quality was evaluated using an agarose gel electrophoresis (0.8%). No smear or fragments were observed. Considering that a human diploid genome represents 6 pg of DNA, we performed serial dilutions of the duplication and deletion controls to obtain a concentration of 1.5 pg/μl. One microliter of each solution was aliquoted in a 96-well plate, so that approximately 25% of the wells are expected to contain a recombinant allele (as we used heterozygous controls who carry only one copy of the deleted or duplicated alleles). The number of positive wells was then counted when amplifying deleted and duplicated DNA.

Calculations with the General Selection Model

Given a locus with two alleles (e.g. wild-type *DPY19L2* allele and deleted allele), and noting q the frequency of the minor (deleted) allele, the GSM predicts the change in allele frequency Δq at each generation given the relative fitness of each genotype. In our case the homozygous wild-type is used as a reference (fitness 1), and the homozygous deleted men are known to be 100% infertile while the deletion is considered to have no effect in women, hence the fitness of the homozygous deleted genotype is 0.5. Let W be the relative fitness of the heterozygous genotype, and p = 1-q the frequency of the wild-type allele. Note that $q \approx 1.7 \times 10^{-3}$, hence $p \approx 998 \times 10^{-3}$. The GSM therefore simplifies to: $\Delta q \approx -q[W(q-p)+p-q/2]$.

In the first scenario, heterozygous individuals are assumed to have no phenotype, hence W = 1 and the equation simplifies to:

$$\Delta q \approx -\frac{q^2}{2}$$

In the second scenario, we no longer assume W = 1 and instead wish to calculate the value of W such that the GSM-predicted Δq exactly compensates the *de novo* rate of production of deleted alleles through NAHR, i.e. $\Delta q = -1.8 \times 10^{-5}$. Turning the previous

equation around, we obtain: $W \approx \frac{p + \frac{\Delta q}{q} - \frac{q}{2}}{p - q}$. Substituting the values of p, q and Δq, this yields W = 0.99. Assuming that heterozygous women have no phenotype, we finally obtain a relative fitness of 98% for heterozygous males.

Supporting Information

Figure S1 Identification of a duplication of the *DPY19L2* locus by array CGH. Array-CGH analyses showed a 130 kb gain extending from base 63,947,732 to 64,078,229 in chromosome 12q14.2. Coordinates of variations or probes (y-axis) are based on the UCSC GRCh37/hg19 assembly. Graphical overview and analysis of the data were obtained with the Genomic Workbench

software, standard edition 6.5 (Agilent) with the following parameters: aberration algorithm ADM-2, threshold 6.0, fuzzy zero, centralisation and moving average window 0.5 Mb. The value of zero (x-axis) represents equal fluorescence intensity ratio between sample and reference DNA. Copy-number gains shift the ratio to the right (positive values). Three adjacent probes located at the *DPY19L2* locus are duplicated in the analyzed patient and the mean log₂ ratio was +0.53 according to the Alexa 5 deviation with a mirror image.

(GIF)

Table S1 Sequence of the PCR primers, position (Hg19), size of the amplified products (between brackets), hybridization temperature (Hyb.). The position of the primers is illustrated in Figure 1B. (DOC)

Table S2 Sequence of the recombination hotspot region of control subjects (10 Europeans and 10 North Africans) for the identification of LCR-specific markers and determination of the precise localisation of the breakpoints of 15 globozoospermia patients.

(DOC)

Table S3 Number and percentage (%) of recombined alleles with breakpoints located within (inside) or outside of the LCR. (DOC)

Table S4 Sequence of the recombination hotspot region of control subjects (10 Europeans and 10 North Africans) for the identification of LCR-specific markers and determination of the precise localisation of the breakpoints of 15 globozoospermia patients. The first column indicates the reference number of the identified variants as shown in Figure 2. Markers that are LCR-specific (i.e. homozygous and invariant within each LCR across all controls, but differing between LCR1 and LCR2) according to the results obtained from the 20 sequenced control individuals (columns 5–8) are indicated in larger bold lettering. Nucleotides that are not LCR-specific are considered as SNPs. The markers' sequence in each LCR according to the Hg19 reference sequence is indicated column 4. In patients, the presence of the Hg19

reference nucleotide is indicated by a cross. When a single nucleotide is detected, the patient is considered homozygous at that position. Because the rows are color-coded, with alternating grey and white rows corresponding to the Hg19 reference nucleotide for LCR1 or LCR2 respectively, and because bold crosses correspond to validated LCR-specific markers, the recombination breakpoints can be easily visualized. For each patient a vertical stretch of bold crosses in grey rows (displayed in orange rectangles) shows non-recombined genetic material coming from LCR1. This is followed by a stretch of bold crosses in white rows (displayed in yellow rectangles), which shows non-recombined DNA from LCR2. For each patient the breakpoint's localisation is inferred when his genotype shifts from LCR1- to LCR2-specific markers. Unboxed regions therefore correspond to breakpoint maximal regions. Patients 1–7 breakpoints are located between markers 17 and 18 (BP1). Patients 8–13 BPs are located between markers 18 and 24 (BP2). Patient 14 is the only heterozygous patient, with BP2 and a breakpoint between markers 25 and 28 (BP3). Patient 15 is homozygous for BP3. SNPs differing between patients with the same breakpoints are highlighted with a blue background. This indicates the presence of 3, 2 and 2 distinct haplotypes for BP1, BP2 and BP3 respectively. Overall this indicates the presence of 7 distinct haplotypes, so that the occurrence of at least 7 separate recombination events within our series of 15 patients can be inferred.

(XLS)

Acknowledgments

We thank Gaelle Vieville from Grenoble CHU and Audrey Labalme from the Hospices Civils de Lyon for their help with the array CGH analysis. We thank Daniel Turner from the Sanger institute for helpful discussions.

Author Contributions

Conceived and designed the experiments: PFR. Performed the experiments: CC FA TK. Analyzed the data: CC PFR VS JL P-SJ CA NT-M. Contributed reagents/materials/analysis tools: DS VS JL PFR. Wrote the paper: PFR CC NT-M.

References

- Gu W, Zhang F, Lupski JR (2008) Mechanisms for human genomic rearrangements. *Pathogenetics* 1: 4.
- Smith CE, Llorente B, Symington LS (2007) Template switching during break-induced replication. *Nature* 447: 102–5.
- Lupski JR, Stankiewicz P (2005) Genomic disorders: molecular mechanisms for rearrangements and conveyed phenotypes. *PLoS Genet* 1: e49. doi:10.1371/journal.pgen.0010049
- Turner DJ, Miretti M, Rajan D, Fiegler H, Carter NP et al. (2008) Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nat Genet* 40: 90–5.
- Molina O, Anton E, Vidal F, Blanco J (2011) Sperm rates of 7q11.23, 15q11q13 and 22q11.2 deletions and duplications: a FISH approach. *Hum Genet* 129: 35–44.
- Liu P, Lacaria M, Zhang F, Withers M, Hastings PJ et al. (2011) Frequency of nonallelic homologous recombination is correlated with length of homology: evidence that ectopic synapsis precedes ectopic crossing-over. *Am J Hum Genet* 89: 580–8.
- Arnheim N, Calabrese P, Tiemann-Boege I (2007) Mammalian meiotic recombination hot spots. *Annu Rev Genet* 41: 369–99.
- Paigen K, Petkov P (2010) Mammalian recombination hot spots: properties, control and evolution. *Nat Rev Genet* 11: 221–33.
- Kong A, Thorleifsson G, Gudbjartsson DF, Masson G, Sigurdsson A et al. (2010) Fine-scale recombination rate differences between sexes, populations and individuals. *Nature* 467: 1099–103.
- Myers S, Freeman C, Auton A, Donnelly P, McVean G (2008) A common sequence motif associated with recombination hot spots and genome instability in humans. *Nat Genet* 40: 1124–9.
- Hayashi K, Yoshida K, Matsui Y (2005) A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. *Nature* 438: 374–8.
- Berg IL, Neumann R, Lam KW, Sarbajna S, Odenthal-Hesse L et al. (2010) PRDM9 variation strongly influences recombination hot-spot activity and meiotic instability in humans. *Nat Genet* 42: 859–63.
- Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C et al. (2010) Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science* 327: 876–9.
- Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C et al. (2010) PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 327: 836–40.
- Berg IL, Neumann R, Sarbajna S, Odenthal-Hesse L, Butler NJ et al. (2011) Variants of the protein PRDM9 differentially regulate a set of human meiotic recombination hotspots highly active in African populations. *Proc Natl Acad Sci U S A* 108: 12378–83.
- Sasaki M, Lange J, Keeney S (2010) Genome destabilization by homologous recombination in the germ line. *Nat Rev Mol Cell Biol* 11: 182–95.
- Harbuz R, Zouari R, Pierre V, Ben Khelifa M, Kharouf M et al. (2011) A Recurrent Deletion of DPY19L2 Causes Infertility in Man by Blocking Sperm Head Elongation and Acrosome Formation. *Am J Hum Genet* 88: 351–61.
- Dam AH, Feenstra I, Westphal JR, Ramos L, van Golde RJ et al. (2007) Globozoospermia revisited. *Hum Reprod Update* 13: 63–75.
- Florke-Gerloff S, Topfer-Petersen E, Muller-Esterl W, Mansouri A, Schatz R et al. (1984) Biochemical and genetic investigation of round-headed spermatozoa in infertile men including two brothers and their father. *Andrologia* 16: 187–202.
- Kilani Z, Ismail R, Ghunaim S, Mohamed H, Hughes D et al. (2004) Evaluation and treatment of familial globozoospermia in five brothers. *Fertil Steril* 82: 1436–9.
- Nistal M, Herrero A, Sanchez-Corral F (1978) [Absolute teratozoospermia in a family. Irregular microcephalic spermatozoa without acrosome]. *Andrologia* 10: 234–40.
- Dam AH, Koscienski I, Kremer JA, Moutou C, Jaeger AS et al. (2007) Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet* 81: 813–20.
- Liu G, Shi QW, Lu GX (2010) A newly discovered mutation in PICK1 in a human with globozoospermia. *Asian J Androl* 12: 556–60.

24. Carson AR, Cheung J, Scherer SW (2006) Duplication and relocation of the functional DPY19L2 gene within low copy repeats. *BMC Genomics* 7: 45.
25. Coutton C, Zouari R, Abada F, Ben Khelifa M, Merdassi G et al. (2012) MLPA and sequence analysis of DPY19L2 reveals point mutations causing globozoospermia. *Hum Reprod* 27: 2549–58.
26. Elinati E, Kuentz P, Redin C, Jaber S, Vanden Meerschaut F et al. (2012) Globozoospermia is mainly due to DPY19L2 deletion via non-allelic homologous recombination involving two recombination hotspots. *Hum Mol Genet* 21: 3695–702.
27. Koscienski I, Elinati E, Fossard C, Redin C, Muller J et al. (2011) DPY19L2 deletion as a major cause of globozoospermia. *Am J Hum Genet* 88: 344–50.
28. Pierre V, Martinez G, Coutton C, Delaroche J, Yassine S et al. (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: 2955–65.
29. Shaikh TH, Gai X, Perin JC, Glessner JT, Xie H et al. (2009) High-resolution mapping and analysis of copy number variations in the human genome: a data resource for clinical and research applications. *Genome Res* 19: 1682–90.
30. Pinto D, Marshall C, Feuk L, Scherer SW (2007) Copy-number variation in control population cohorts. *Hum Mol Genet* 16 Spec No. 2: R168–73.
31. Itsara A, Cooper GM, Baker C, Girirajan S, Li J et al. (2009) Population analysis of large copy number variants and hotspots of human genetic disease. *Am J Hum Genet* 84: 148–61.
32. de Smith AJ, Tsalenko A, Sampas N, Scheffer A, Yamada NA et al. (2007) Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases. *Hum Mol Genet* 16: 2783–94.
33. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O et al. (2010) Origins and functional impact of copy number variation in the human genome. *Nature* 464: 704–12.
34. Schimenti J (2000) Segregation distortion of mouse t haplotypes the molecular basis emerges. *Trends Genet* 16: 240–3.
35. Grey C, Barthes P, Chauveau-Le Fric G, Langa F, Baudat F et al. (2011) Mouse PRDM9 DNA-binding specificity determines sites of histone H3 lysine 4 trimethylation for initiation of meiotic recombination. *PLoS Biol* 9: e1001176. doi:10.1371/journal.pbio.1001176
36. Jeanpierre M (1987) A rapid method for the purification of DNA from blood. *Nucleic Acids Res* 15: 9611.
37. World Health Organization (2010) WHO Laboratory Manual for the Examination and Processing of Human Semen. 5th ed. Geneva: WHO Press.

Principaux résultats

A partir d'une cohorte totale de 8574 individus contrôles, nous avons pu identifier 30 délétions et 100 duplications du gène *DPY19L2*. Ainsi la fréquence allélique de la délétion est estimée à environ 1.7×10^{-3} et celle de la duplication à environ 5.8×10^{-3} soit un ratio duplication/délétion de 3.4 Ces résultats démontrent clairement un excès de l'allèle dupliqué par rapport à l'allèle délété dans la population générale.

Paradoxalement, l'estimation du taux moyen de délétions et duplications produites *de novo* par PCR digitale sur l'ADN spermatique de 3 donneurs sains est de 1.8×10^{-5} (95% CI: 1.4×10^{-5} ; 2.2×10^{-5}) et de 7.7×10^{-6} (95% CI: 6.1×10^{-6} ; 9.7×10^{-6}) respectivement. Nous avons donc identifié un taux de délétions *de novo* supérieur à celui des duplications, comme prédit par le modèle de NAHR,

Afin de préciser les points de cassure des réarrangements (délétion et duplication) nous avons pu génotyper 20 SNPs spécifiques des LCRs télomériques et centromériques. A partir de ces données, 5 points de cassures distincts (BP1-5) ont pu être identifiés sur les 185 allèles recombinés étudiés (108 délétés et 77 dupliqués), le BP2 étant le plus fréquemment retrouvé (63%). Ces 5 points de cassures sont compris dans une région minimale de 1153 bp. L'analyse bio-informatique de cette région a mis en évidence la présence au centre de la région d'une séquence consensus de 13 nucléotides reconnue par PRDM9, une protéine à doigts de zinc qui favorise la survenue des cassures doubles brins à l'origine des processus de recombinaisons.

Discussion et Perspectives

Les modèles théoriques prédisent que, lors de la méiose, le mécanisme NAHR génère *de novo* plus d'allèles recombinés délétés que dupliqués. Comme attendu, nous avons identifié un taux de délétions *de novo* supérieur à celui des duplications. Étonnamment, dans la population générale, nous avons observé que les allèles *DPY19L2* dupliqués sont trois fois plus fréquents que les allèles délétés. Ce paradoxe peut s'expliquer par la sélection négative qui s'opère à l'encontre des hommes infertiles porteurs de la délétion homozygote entraînant la disparition de l'allèle délété au profit de l'allèle dupliqué. Egalement, nous pouvons envisager un effet délétère *a minima* de la délétion chez les hommes hétérozygotes bien que cela reste à démontrer.

Le point de cassure BP2 est le plus fréquemment observé sur l'ensemble des allèles recombinés. Ceci est dû en partie à sa taille relativement plus importante que celle des autres points de cassure et probablement du fait que le site de reconnaissance PRDM9 se trouve exactement en son milieu. On observera que chez les patients globozoospermiques le point de cassure BP1 est surreprésenté (14 allèles sur 34). Cette différence peut s'expliquer par des haplotypes communs entre ces patients tous originaires de Tunis suggérant un effet fondateur de l'allèle délété en BP1.

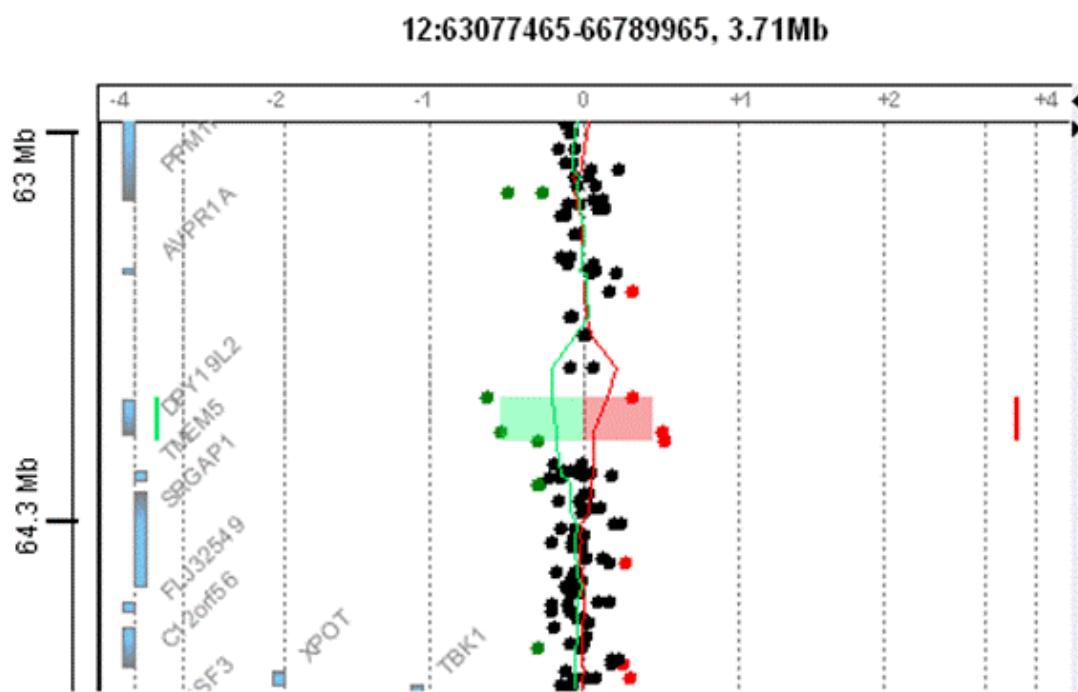
Enfin cette étude renforce l'importance des sites de reconnaissance de PRDM9 dans la survenue des recombinaisons par mécanisme NAHR. Il serait intéressant de vérifier par la suite si la présence de variations nucléotidiques au sein de ces séquences consensus peut influencer le taux global de recombinaison du locus.

Supplemental Data

Fine characterisation of a recombination hotspot at the *DPY19L2* locus and resolution of the paradoxical excess of duplications over deletions in the general population.

Coutton C, Abada F, Karaouzene T, Sanlaville D, Satre V, Lunardi J, Jouk PS, Arnoult C, Thierry-Mieg N, Ray PF

Figure S1. Identification of a duplication of the *DPY19L2* locus by array CGH.



Array-CGH analyses showed a 130 kb gain extending from base 63,947,732 to 64,078,229 in chromosome 12q14.2. Coordinates of variations or probes (y-axis) are based on the UCSC GRCh37/hg19 assembly. Graphical overview and analysis of the data were obtained with the Genomic Workbench software, standard edition 6.5 (Agilent) with the following parameters: aberration algorithm ADM-2, threshold 6.0, fuzzy zero, centralisation and moving average window 0.5 Mb. The value of zero (x-axis) represents equal fluorescence intensity ratio between sample and reference DNA. Copy-number gains shift the ratio to the right (positive values). Three adjacent probes located at the *DPY19L2* locus are duplicated in the analyzed patient and the mean log₂ ratio was +0.53 according to the Alexa 5 deviation with a mirror image.

Table S1. Details of the PCR primers.

Locus amplified	5' to 3' DNA sequence	Starting position (5') and Size of amplicon (nt)	Hyb. (°C)
1) Long deletion specific			
5'L1	TAGACTCTCTGGAAAGGTATTATCG	chr12:63,934,316	64
3'L2	CCAAGGAAATCGAAGACGCT	chr12:64131430 (2088)	
2) Long LCR1 specific			
5'L1	TAGACTCTCTGGAAAGGTATTATCG	chr12:63,934,316	55
3'L1	CAAGGAAATCGAGGGATGCC	chr12:63936360 (2087)	
3) Long LCR2 specific			
5'L2	ACTCTTCTGGAAAGAAGAGAA	chr12:64,129,361	55
3'L2	CCAAGGAAATCGAAGACGCT	chr12:64,131,430 (2089)	
4) Short Inner deletion specific			
5'SI1	ATAATCTGTAATTCCACTGCATTCA	chr12:63,934,614	60
3'SI2	GGCACAGCTGCCAGCATTC	chr12:64,131,009 (1392)	
5) Short duplication specific			
5'SI2	ATAATCTGTAATTCTACTGCATTCTA	chr12:64,129,654	60
3'SI1	GGCACAGCTGCTAGCATT	chr12:63,935,989 (1394)	
6) Long duplication specific			
5'L2	TAGACTCTCTGGAAAGAAGAGAA	chr12:64,129,358	64
3'L1	CCAAGGAAATCGAGGGATGCCGA	chr12:63,936,381 (2085)	
7) Positive control primers			
5RYRex40	TAGGCACAGAGTGAGAGGGTCAAG	chr19:38,986,748	60
3RYRex42	GCAAATTAGTCTCCTCTGGTTGG	chr19:38,987,714 (990)	
8) Dpy1912 Taqman	6FAM-ACACCATGGCACAAAGTGAGA-TAMRA	chr12:63,935,220 and chr12:64,130,260	60
9) PRDM9 ZF PCR			
PN0.6F	TGAGGTTACCTAGTCTGGCA		55
PN2.5R	ATAAGGGGTCAAGCAGACTTC		
10) PRDM9 ZF sequencing primers			
PN1.2F	TGAATCCAGGAAACACAGGC		55
PN2.4R	GCAAGTGTGTTGACCACA		

Sequence of the PCR primers, position (Hg19), size of the amplified products (between brackets), hybridization temperature (Hyb.). The position of the primers is illustrated in [Figure 1B](#).

Table S2. Number (Nb) and percentage (%) of deleted (del.) versus duplicated (dup.) allele. Total number of recombinant alleles (deleted + duplicated : total recomb.) and frequency of the different alleles in the studied populations

	Nb of individuals	Nb del. (%)	Freq. of del.	Nb dup. (%)	Freq. of dup.	Total recomb	Freq. of recomb.
Total DGV	6575	26 (24)	0,40	83 (76)	1,26	109	1,66
Home CGH cohort + PCR	1999	4 (19)	0,18	17 (81)	0,88	21	1,06
Total	8574	30 (23)	0,35	100 (77)	1,18	130	1,53

Sequence of the recombination hotspot region of control subjects (10 Europeans and 10 North Africans) for the identification of LCR-specific markers and determination of the precise localisation of the breakpoints of 15 globozoospermia patients.

Table S3. Number and percentage (%) of recombined alleles with breakpoints located within (Inside) or outside of the LCR.

	Inside LCR		Outside LCR		Total Inside LCR	Total Outside LCR	Total recomb
	Del	dup	del	dup			
Total DGV	22 (85)	61 (73)	4 (15)	22 (27)	83 (76)	26 (24)	109
Home CGH cohort + PCR	4 (100)	15 (88)	0	2 (12)	19 (90)	2 (10)	21
Total	26 (87)	76 (76)	4 (13)	24 (24)	102 (78)	28 (22)	130

Table S4. Sequence of the recombination hotspot region of control subjects (10 Europeans and 10 North Africans) for the identification of LCR-specific markers and determination of the precise localisation of the breakpoints of 15 globozoospermia patients.

Ref Fig 2	Position Hg19 Hg19	Ref Seq Hg19	Europeans		North Africans		P. 1		P. 2		P. 3		P. 4		P. 5		P. 6		P. 7		P. 8		P. 9		P. 10		P. 11		P. 12		P. 13		P. 14		P. 15	
			control		Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2	Al 1	Al 2		
	LCR1 start	63 923 419																																		
	LCR2 start	64 119 291																																		
1	LCR1	63 934 436	G	18	G	2	A	17	G	3	A	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
1	LCR2	64 129 478	A	12	A	8	G	14	A	6	G	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
2	LCR1	63 934 605	A	20	A	0	20	A	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
2	LCR2	64 129 647	G	20	G	0	20	G	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
3	LCR1	63 934 629	C	20	C	0	20	C	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
3	LCR2	64 129 671	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
4	LCR1	63 934 639	AG	20	AG	0	20	AG	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
4	LCR2	64 129 681	TA	20	TA	0	20	TA	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
5	LCR1	63 934 693	C	20	C	0	20	C	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
5	LCR2	64 129 735	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
6	LCR1	63 934 728	T	17	T	3	C	15	T	5	C																									
6	LCR2	64 129 770	C	18	C	2	T	17	C	3	C	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
7	LCR1	63 934 737	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
7	LCR2	64 129 779	C	20	C	0	20	C	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
8	LCR1	63 934 757	G	12	G	8	A	9	G	11	A	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
8	LCR2	64 129 799	A	16	A	4	G	11	A	9	G	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
9	LCR1	63 934 764	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
9	LCR2	64 129 806	C	20	C	0	20	C	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
10	LCR1	63 934 766	C	20	C	0	20	C	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
10	LCR2	64 129 808	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
11	LCR1	63 934 832	C	20	C	0	20	C	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
11	LCR2	64 129 874	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
12	LCR1	63 934 867	C	20	C	0	20	C	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
12	LCR2	64 129 909	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
13	LCR1	63 934 875	C	18	C	2	T	20	C	0		x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x				
13	LCR2	64 129 917	T	0	T	20	C	2	T	18	C	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
14	LCR1	63 934 877	C	16	C	4	A	12	C	6	A	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
14	LCR2	64 129 919	A	7	A	13	C	6	A	14	C																									
15	LCR1	63 934 902	A	8	A	12	-	10	A	10	-																									
15	LCR2	64 129 944	-	4	-	16	A	8	-	12	A	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
16	LCR1	63 934 936	A	20	A	0	20	A	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
16	LCR2	64 129 977	G	20	G	0	20	G	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
17	LCR1	63 934 955	G	20	G	0	20	G	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					
17	LCR2	64 129 996	T	20	T	0	20	T	0			x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x					

The first column indicates the reference number of the identified variants as shown in Figure 2. Markers that are LCR-specific (i.e. homozygous and invariant within each LCR across all controls, but differing between LCR1 and LCR2) according to the results obtained from the 20 sequenced control individuals (columns 5–8) are indicated in larger bold lettering. Nucleotides that are not LCR-specific are considered as SNPs. The markers' sequence in each LCR according to the Hg19 reference sequence is indicated column 4. In patients, the presence of the Hg19 reference nucleotide is indicated by a cross. When a single nucleotide is detected, the patient is considered homozygous at that position. Because the rows are color-coded, with alternating grey and white rows

corresponding to the Hg19 reference nucleotide for LCR1 or LCR2 respectively, and because bold crosses correspond to validated LCR-specific markers, the recombination breakpoints can be easily visualized. For each patient a vertical stretch of bold crosses in grey rows (displayed in orange rectangles) shows non-recombined genetic material coming from LCR1. This is followed by a stretch of bold crosses in white rows (displayed in yellow rectangles), which shows non-recombined DNA from LCR2. For each patient the breakpoint's localisation is inferred when his genotype shifts from LCR1- to LRC2-specific markers. Unboxed regions therefore correspond to breakpoint maximal regions. Patients 1–7 breakpoints are located between markers 17 and 18 (BP1). Patients 8–13 BPs are located between markers 18 and 24 (BP2). Patient 14 is the only heterozygous patient, with BP2 and a breakpoint between markers 25 and 28 (BP3). Patient 15 is homozygous for BP3. SNPs differing between patients with the same breakpoints are highlighted with a blue background. This indicates the presence of 3, 2 and 2 distinct haplotypes for BP1, BP2 and BP3 respectively. Overall this indicates the presence of 7 distinct haplotypes, so that the occurrence of at least 7 separate recombination events within our series of 15 patients can be inferred.

PHÉNOTYPE 3:

**ANOMALIES
MORPHOLOGIQUES MULTIPLES
DU FLAGELLE DU
SPERMATOZOIDE**

Article 4

Mutations in *DNAH1*, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella.

Ben Khelifa M*, Coutton C*, Zouari R, Karaouzène T, Rendu J, Bidart M, Yassine S, Pierre V, Delaroche J, Hennebicq S, Grunwald D, Escalier D, Pernet-Gallay K, Jouk PS, Thierry-Mieg N, Touré A, Arnoult C, Ray PF

* Co-premiers auteurs

The American Journal of Human Genetics, 94:95-104, January 2014

Contexte et objectifs

Les anomalies morphologiques du flagelle font partie des anomalies flagellaires primitives entraînant des altérations sévères de la mobilité (asthénozoospermie) liées à diverses anomalies structurales et/ou ultrastructurales du flagelle. La diversité des structures impliquées dans le mouvement du spermatozoïde et la variété d'anomalies ultrastructurales observées suggèrent qu'un grand nombre de gènes est impliqué dans ce groupe de pathologie (Chemes *et al.*, 1998).

L'asthénozoospermie secondaire à des anomalies flagellaires primitives s'observe très fréquemment dans les dyskinésies ciliaires primitives (DCP), un groupe de pathologies caractérisé cliniquement par des atteintes principalement pulmonaires et souvent associées à un *situs inversus*. L'asthénozoospermie observée chez ces patients est due à des anomalies ultrastructurales du flagelle généralement non identifiable en microscopie optique (Chemes and Rawe, 2003).

L'objectif de ce travail a été d'identifier par cartographie d'homoygotie une cause génétique chez une cohorte de 20 sujets présentant une asthénozoospermie secondaire à des anomalies morphologiques des flagelles sans autres signes cliniques de DCP. L'observation de l'éjaculat chez ces patients retrouve une mosaïque de spermatozoïdes avec des flagelles absents, courts, enroulés ou de calibre irrégulier. Nous avons appelé ce phénotype anomalies morphologiques multiples des flagelles (AMMF).

Mutations in *DNAH1*, which Encodes an Inner Arm Heavy Chain Dynein, Lead to Male Infertility from Multiple Morphological Abnormalities of the Sperm Flagella

Mariem Ben Khelifa,^{1,2,3,15} Charles Coutton,^{1,2,4,15} Raoudha Zouari,⁵ Thomas Karaouzène,^{1,2} John Rendu,^{1,6,7} Marie Bidart,^{1,7} Sandra Yassine,^{1,2} Virginie Pierre,^{1,2} Julie Delaroche,^{1,7} Sylviane Hennebicq,^{1,2,8} Didier Grunwald,^{1,7} Denise Escalier,⁹ Karine Pernet-Gallay,^{1,7} Pierre-Simon Jouk,^{10,11} Nicolas Thierry-Mieg,¹⁰ Aminata Touré,^{12,13,14} Christophe Arnoult,^{1,2,16} and Pierre F. Ray^{1,2,6,16,*}

Ten to fifteen percent of couples are confronted with infertility and a male factor is involved in approximately half the cases. A genetic etiology is likely in most cases yet only few genes have been formally correlated with male infertility. Homozygosity mapping was carried out on a cohort of 20 North African individuals, including 18 index cases, presenting with primary infertility resulting from impaired sperm motility caused by a mosaic of multiple morphological abnormalities of the flagella (MMAF) including absent, short, coiled, bent, and irregular flagella. Five unrelated subjects out of 18 (28%) carried a homozygous variant in *DNAH1*, which encodes an inner dynein heavy chain and is expressed in testis. RT-PCR, immunostaining, and electronic microscopy were carried out on samples from one of the subjects with a mutation located on a donor splice site. Neither the transcript nor the protein was observed in this individual, confirming the pathogenicity of this variant. A general axonemal disorganization including mislocalization of the microtubule doublets and loss of the inner dynein arms was observed. Although *DNAH1* is also expressed in other ciliated cells, infertility was the only symptom of primary ciliary dyskinesia observed in affected subjects, suggesting that *DNAH1* function in cilium is not as critical as in sperm flagellum.

Male infertility affects more than 20 million men worldwide and represents a real health concern.¹ It is a typical multifactorial disorder with a strong genetic basis and additional etiological factors such as urogenital infections, immunological or endocrine diseases, attack from reactive oxygen species (ROS), or perturbations from endocrine disruptors. To date, despite substantial efforts made to identify genes specifically involved in male infertility by many teams including ours,^{2,3} only a handful of genes have been formally correlated with human sperm defects. Male infertility caused by impaired sperm motility (asthenozoospermia) is also often observed in men with primary ciliary dyskinesia (PCD), a group of mainly autosomal-recessive disorders caused by dysfunctions of motile cilia leading primarily to respiratory infections and often to situs inversus. Recent research on PCD has been extremely prolific and allowed the identification and characterization of numerous proteins necessary for adequate axonemal molecular structure and assembly (Table S1 available online). The axoneme is a highly evolutionarily conserved structure found in motile cilia and in sperm flagella,

mainly composed of an intricate network of microtubules and dyneins. Sperm parameters have not been systematically explored and are often only scarcely described in manuscripts investigating PCD-affected individuals. Although sperm flagella and motile cilia have a similar axonemal structure based on the presence of nine peripheral microtubule doublets plus two central ones, they present several differences that might explain why PCDs are not always associated with asthenozoospermia.⁴ We note that no mutations in axonemal genes have been described as being involved exclusively in infertility without also inducing PCD.

In the present study, we analyzed 20 subjects presenting with asthenozoospermia resulting from a combination of five morphological defects of the sperm flagella (absent, short, bent, and coiled flagella and flagella of irregular width) without any of the other PCD-associated symptoms. Similar phenotypes have been previously described and named "dysplasia of the fibrous sheath," "short tails," or "stump tails."^{5–15} We propose to call this syndrome "multiple morphological anomalies of the flagella

¹Université Joseph Fourier, Grenoble 38000, France; ²Laboratoire AGIM, CNRS FRE3405, Equipe "Andrologie et Génétique," La Tronche 38700, France; ³Laboratoire de génomique Biomédicale et Oncogénétique, Institut Pasteur de Tunis, 1002 Tunis, Tunisie; ⁴CHU de Grenoble, Hôpital Couple Enfant, Département de Génétique et Procréation, Laboratoire de Génétique Chromosomique, Grenoble 38000, France; ⁵Clinique des Jasmins, 23, Av. Louis BRAILLE, 1002 Tunis, Tunisia; ⁶CHU de Grenoble, Institut de Biologie et Pathologie, Département de Biochimie, Toxicologie et Pharmacologie (DBTP), UF de Biochimie et Génétique Moléculaire, Grenoble 38000, France; ⁷INSERM, U836, Grenoble Institute of Neuroscience, La Tronche 38700, France; ⁸CHU de Grenoble, Hôpital Couple Enfant, Département de Génétique et Procréation, Laboratoire d'Aide à la Procréation – CECOS, Grenoble 38000, France; ⁹INSERM UMR_S933, Université Pierre et Marie Curie (Paris 6), Paris 75012, France; ¹⁰Université Joseph Fourier-Grenoble 1 / CNRS / TIMC-IMAG UMR 5525, Grenoble 38041, France; ¹¹CHU de Grenoble, Hôpital Couple Enfant, Département de Génétique et Procréation, Service de Génétique Clinique, Grenoble 38000, France; ¹²INSERM, U1016, Institut Cochin, Paris 75014, France; ¹³CNRS, UMR8104, Paris 75014, France; ¹⁴Université Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine, Paris 75014, France

¹⁵These authors contributed equally to this work

¹⁶These authors contributed equally to this work and are co-senior authors

*Correspondence: pray@chu-grenoble.fr

<http://dx.doi.org/10.1016/j.ajhg.2013.11.017>. ©2014 by The American Society of Human Genetics. All rights reserved.

(MMAF)," a name that provides a more accurate description of this phenotype. We carried out a SNP whole-genome scan on 20 individuals presenting with severe MMAF. The study was approved by our local ethics committee; all individuals gave their signed informed consent and national laws and regulations were respected. All individuals originated from North Africa (11 Tunisians, 7 Algerians, and 2 Libyans) and were treated in Tunis (Clinique des Jasmins, Tunis, Tunisia) for primary infertility. Twelve of the subjects were born from related parents, usually first cousins. None of the subjects were related to one another apart from three individuals (P1–P3) who were brothers. All subjects had normal somatic karyotypes. All sperm analyses were performed at least twice, in accordance with the World Health Organization recommendations.¹⁶ Subjects were recruited on the basis of the identification of >5% of at least four of the aforementioned flagellar morphological abnormalities (absent, short, coiled, bent, and irregular flagella) (Table 1). All subjects presented with severe asthenozoospermia: 11 out of 20 subjects had no (0%) motility, 8 had sperm motility <10%, and one (P6) had 35% motility. Saliva was obtained from all participants via Oragene DNA Self-Collection Kit (DNAgenotech) but only one subject (P3) agreed to donate sperm and blood samples for research use. During their medical consultation for infertility, all subjects answered a health questionnaire focused on PCD manifestations, and none indicated suffering from any of the other symptoms encountered in PCD.

Homozygosity mapping was carried out with 250K Sty1 SNP mapping arrays (Affymetrix) on DNA extracted from the 20 studied subjects' saliva samples. Common regions of homozygosity were identified with the homoSNP software. After exclusion of the centromeric regions, we identified two regions located on chromosomes 3 and 20 with a region of homozygosity > 1 Mb common to 10/20 analyzed individuals (Figure S1). In addition, 4 and 9 subjects presented with a stretch of homozygosity > 15 Mb overlapping chromosomes 20 and 3 regions, respectively. All three brothers (P1–P3) were homozygous at the chromosome 3 region, although only two of them were homozygous at the chromosome 20 region. We excluded all other regions of homozygosity because they did not fulfil the following criteria: (1) more than eight individuals including at least two of the brothers sharing a region of homozygosity > 1 Mb and (2) presence of a potential candidate gene in the region according to its expression profile and/or presumed function. Finer analysis of the chromosome 3 region showed that 15 individuals were homozygous for two smaller subregions located at chr3: 46,745,396–47,606,570 and chr3: 52,111,974–53,028,375 (UCSC Genome Browser human reference genome build hg17, Figure S1). Sixteen genes are annotated in the first subregion (Table S2), among which only one gene (*KIF9* [MIM 607910]) appeared as a good candidate; indeed, studies in the protist *Trypanosoma brucei* showed that *kif9A* (the mouse ortholog of human *KIF9*) is located in

the axoneme and that its depletion alters motility.¹⁷ The second subregion in chromosome 3 includes 28 genes (Table S2). The dynein heavy chain 1 gene (*DNAH1* [MIM 603332]) appeared as the best candidate gene because it codes for an axonemal dynein heavy chain and is expressed in various tissues including testis.¹⁸ Furthermore, asthenozoospermia was described in mice lacking *Dnahc1*, the *DNAH1* mouse ortholog (previously named *Mdhc7*).¹⁹ Finally, among the ten genes located in the selected region of chromosome 20 (chr20: 33,572,687–34,070,415), only *SPAG4* (MIM 603038) appeared as a good candidate: it was described in rat to be associated with the axoneme in elongating spermatids and epididymal sperm.²⁰ We therefore decided to sequence *KIF9* (RefSeq accession number NM_001134878.1), *DNAH1* (RefSeq NM_015512.4), and *SPAG4* (RefSeq NM_003116.1).

We sequenced the 12 exons and the intron boundaries of *SPAG4* in the 13 individuals homozygous at this locus, and the 19 exons and intron boundaries of *KIF9* in the 15 relevant individuals. We did not identify any likely pathogenic variants in these two genes. We then sequenced the 78 exons and intron boundaries of *DNAH1* in P3 (primer sequences available in Table S3). We identified one homozygous splicing mutation (c.11788–1G>A) in intron 73. The same homozygous mutation was identified in the two other brothers (Figure S2). We then sequenced *DNAH1* for the 17 remaining subjects. The same homozygous mutation (c.11788–1G>A) was identified in one additional individual (P17). We identified three other homozygous variants: another splicing mutation (c.5094+1G>A) in individual P9, a homozygous no-stop mutation disrupting the stop codon in exon 78 (c.12796T>C [p.4266Glnext*21]) in individual P8, and a homozygous missense variant in exon 23 in individual P6 (c.3877G>A [p.Asp1293Asn]). The localization of the *DNAH1* mutations is presented in Figure 1. If we consider only index cases, we identified 5 homozygous variants in 18 unrelated individuals (28%). None of these variants were detected in our control cohort of 100 individuals of North African origin. We note that the parents of the subjects could not be analyzed to confirm the transmission of the variants. We therefore cannot formally exclude the possibility that some of the identified variants may be hemizygous with a deletion on the other allele. However, depending on its size, its position, and its effect on the reading frame, a deleted allele would be at least as deleterious as the identified variants.

To evaluate the association of the variants with the pathology, we compared their frequency in our cohort with that in the Exome Variant Server (EVS) database. At the four genomic positions of interest, the EVS data are of sufficient coverage to provide genotype calls for at least 6,200 individuals, corresponding to 12,400 alleles. There were no variant nucleotides identified at positions c.5094+1, c.11788–1, or c.12796, and only one A allele was identified out of 12,460 alleles at position c.3877.

Table 1. Semen Parameters of the 20 Subjects and the 7 Subjects Carrying *DNAH1* Homozygous Variants

Semen Parameters	Average of 20 Subjects ^a	P1	P2	P3	P6	P8	P9	P17
<i>DNAH1</i> mutations	c.11788–1G>A (p.Gly3930Alafs*120)	c.11788–1G>A (p.Gly3930Alafs*120)	c.11788–1G>A (p.Gly3930Alafs*120)	c.3877G>A (p.Asp1293Asn)	c.12796 T>C (p.4266Glnext*21)	c.5094+1G>A (p.Leu1700Serfs72)	c.11788–1G>A (p.Gly3930Alafs*120)	
Consanguinity	yes	yes	yes	no	yes	yes	yes	no
Origin of the subject	Tunisia	Tunisia	Tunisia	Algeria	Algeria	Algeria	Algeria	Tunisia
Sperm volume (ml)	3.2 (1–5.5)	5	2.5	2	5	4.5	3.5	2.5
Sperm concentration 10 ⁶ /ml	22 (0–59)	45	0	2.8	57	11	53	31
Motility (A+B) 1 hr	2.5 (0–35)	0	NA	2	35	0	0.5	0
Vitality	44 (6–73)	22	NA	NA	73	61	48	NA
Normal spermatozoa	0.35 (0–6)	0	NA	0	6	0	0	0
Absent flagella	30 (8–46)	34	+	34	+	+	16	30
Short flagella	44 (16–70)	38	+	44	+	+	70	20
Coiled flagella	13 (2–32)	14	NA	14	+	+	12	32
Angulation	12 (2–19)	4	NA	2	+	+	8	6
Flagella of irregular caliber	55 (16–92)	48	NA	50	+	+	54	16
Multiple anomalies index	2.9 (1.9–3.9)	3.1	NA	2.4	NA	NA	3	2.6

Values are expressed in percents, unless specified otherwise. Abbreviations are as follows: NA, not available; plus sign, anomalies reported (>5%) but not accurately quantified.

^aValues are expressed as the mean with the lower and higher values in parentheses.

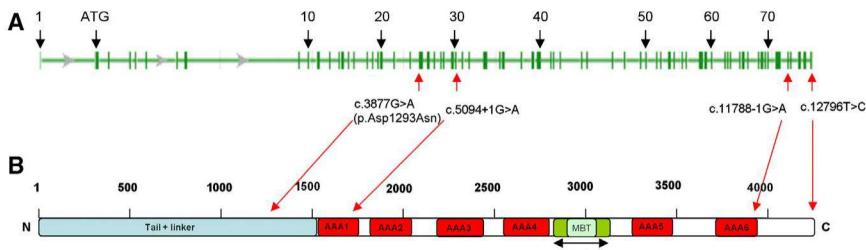


Figure 1. Location of *DNAH1* Mutations in the Intron-Exon Structure and in the Protein Representation of *DNAH1*

(A) *DNAH1* genomic structure.
(B) *DNAH1* domain map showing the location of the four identified mutations. The red boxes indicate the six known AAA-ATPase domains (AAA 1 to 6) as detected by homology (Uniprot server). The microtubule-binding domain (MBT) lies between AAA4 and AAA5. The N-terminal part of the protein binds to the intermediate, light-intermediate dynein chains. The position of the stalk and the microtubule-binding domain (MTB) are indicated.

We performed Fisher exact tests (with `fisher.test` in R) to evaluate whether each observed SNV was statistically overrepresented in our cohort of 18 unrelated individuals, compared to EVS. All four individual SNVs are significantly enriched (p values: 5.9×10^{-11} for $c.11788-1G>A$, 8.1×10^{-6} for $c.5094+1G>A$ and $c.12796T>C$, and 2.4×10^{-5} for $c.3877G>A$). We then investigated whether *DNAH1* as a whole was significantly enriched in damaging SNVs in our cohort. Overall, EVS contains five nonsense and two splice-site SNVs, all observed heterozygously in a total of ten individuals. By using the coverage data available on the EVS website, we find that 20,737 positions covering the *DNAH1* exons and intron boundaries had sufficient sequence coverage to be genotyped in 6,189 individuals on average. By contrast, counting only the two splice-site mutations as damaging, we observe 6 damaging alleles among 36 in our cohort: this represents a highly significant enrichment (Fisher exact test, p value: 3×10^{-12}). Furthermore, we note that there were no homozygous damaging variants observed in the 6,189 EVS individuals compared to 3 in our cohort of 18 (Fisher exact test, p value: 3×10^{-8}). Altogether we believe that these genetic results convincingly demonstrate that mutations in *DNAH1* are associated with MMAF.

The $c.5094+1G>A$ variant found in individual P9 affects *DNAH1* intron 31 consensus donor splice site. The abnormal splicing is predicted to cause the prolongation of exon 31 until the introduction of a nonsense codon (p.Leu1700Serfs72). The position of the next donor site was predicted by "Splice Site Prediction by Neural Network." Unfortunately we could not obtain any leukocytes from this subject to validate this prediction and observe whether this variant also led to nonsense-mediated mRNA decay (NMD). Because of the location of the variant on a consensus splice site and the unambiguous predictions of splice prediction software, we did not synthesize a minigene to verify the effect of this variant in vitro. A missense change, p.Asp1293Asn, was identified in P6. Interestingly, the Asp1293 amino acid is well conserved across species (Figure S3). This missense change is also predicted to be possibly damaging by SIFT and PolyPhen-2, two prediction softwares for nonsynonymous SNPs. It affects the N-ter of the protein (Figure 1B), known to be important for the structure of dynein arms.²¹ Variant

$p.4266Glnext*21$ found in individual P8 abolishes the stop codon in exon 78, leading to the addition of 21 codons at the 3' end of the coding sequence. The role of the C-terminal domain is uncertain, but based on the *D. discoideum* structures, it may participate in long-range allosteric communication between microtubule-binding and ATPase regions.^{22,23} The addition of 21 extra amino acids to this region is likely to disrupt these interactions.

The $c.11788-1G>A$ variant identified in four subjects (P1–P3 and P17) affects the final G nucleotide of *DNAH1* intron 73, one of the consensus splice acceptor nucleotides. The resulting abnormal splicing is predicted to recognize a new CG acceptor site located just one nucleotide further, thus shifting the reading frame and inducing a premature stop codon (p.Gly3930Alafs*120). As could be expected, P1–P3 share a common haplotype (Table S4). P17 also shares a common haplotype of 30 SNPs with P1–P3, suggesting a founder effect for this mutation. To assess the functional impact of the *DNAH1* splice acceptor site mutation $c.11788-1G>A$, we studied mRNA products isolated from control and P3 lymphocytes (primer sequences available in Table S5). RT-PCR of P3's samples yielded no product despite repeated attempts, whereas the three amplification attempts from control lymphocytes yielded the expected product (Figure 2A). RT-PCR targeting *GAPDH* (MIM 138400) and *RPLPO* (MIM 180510) confirmed the integrity of P3's RNA (Figure 2B). This suggests a specific degradation of the mutant *DNAH1* transcripts by NMD. To further validate the pathogenicity of this variant, we analyzed *DNAH1* localization in sperm from P3 by immunofluorescence and the ultrastructure of the flagella by electron microscopy. In control individuals, *DNAH1* antisera decorated the full length of the sperm flagellum (Figure 2C), suggesting a putative role in the tethering of the inner dynein arms along the entire axoneme. In contrast, in sperm from individual P3 carrying the $c.11788-1G>A$ mutation, *DNAH1* immunostaining was absent, confirming that the splicing defect results in the degradation of the transcripts by NMD (Figure 2D). We next tested the integrity of the outer and inner dynein arms by using antibodies directed against *DNAL1* and *DNAI2*, two well-established diagnostic markers of the inner and outer dynein arms, respectively. Staining with *DNAL1* was strongly reduced in the sperm of individual

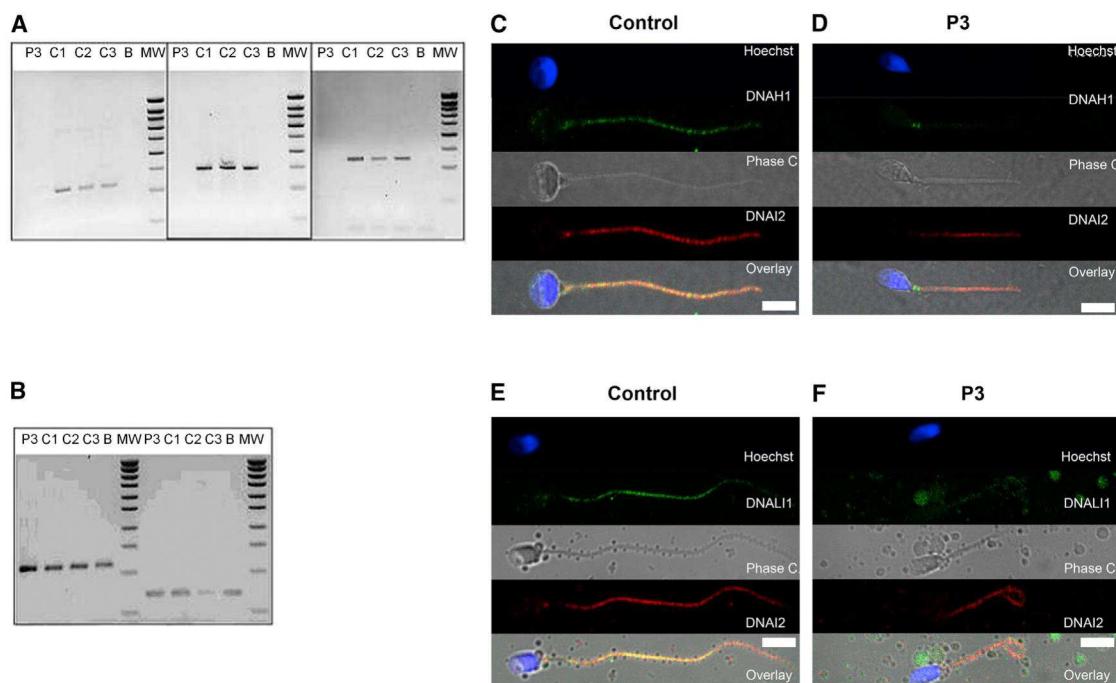


Figure 2. Analysis of P3 Carrying the c.11788–1G>A Variant Evidencing *DNAH1* mRNA Decay by RT-PCR and the Absence of *DNAH1* in Sperm by Immunolocalization

(A) RT-PCR analyses of subject P3 (c.11788–1G>A homozygote) and control individuals from the general population (C1–C3). Electrophoresis showing the RT-PCR amplification of *DNAH1* exons 30–32, 64–66, 73–75. C1, C2, and C3 yield a normal fragment of 228, 293, and 241 bp, whereas subject P3 shows no amplification. There is no amplification from the RT-negative blank control (column B).

(B) Electrophoresis showing the amplification of the same cDNAs with *GAPDH* and *RPL0* primers. Bands of equivalent intensity are obtained from all samples including P3. Reverse transcription was carried out with 500 ng of extracted RNA and oligo dT priming. Two microliters of the obtained cDNA mix was used for the subsequent PCR. PCR amplification was carried out with three couples of primers located in exons 30–32, 64–66, and 73–75 of *DNAH1* at an elongation temperature of 57°C (40 cycles), in parallel to amplification of the same samples with the control housekeeping *GAPDH* and *RPL0*, respectively, at an elongation temperature of 60°C (35 cycles). RT-PCR primers are listed in Table S4.

(C and D) Immunofluorescence staining of human spermatozoa with *DNAH1* antibodies (green) and *DNAI2* (red). *DNAH1* is observed throughout the flagellum in control sperm, whereas it is absent from P3's sperm. In both control and P3 sperm, ODA is present as witnessed by the immunostaining of *DNAI2*.

(E and F) Immunofluorescence staining of human spermatozoa with *DNAL11* antibodies (green) and *DNAI2* (red). *DNAL11*, a marker of IDA, is localized throughout flagella in control sperm, whereas it is strongly reduced in sperm from P3. No difference is noticed in the coimmunostaining of *DNAI2*. Sperm were counterstained with Hoechst 33342 (blue) as nuclei marker. White scale bars represent 5 μm. Sperm cells were washed in phosphate-buffered saline (PBS), fixed in 4% PFA for 2 min at room temperature (RT), and washed twice in PBS. Fixed spermatozoa were allowed to air-dry on poly-L-lysine coated slides followed by permeabilization with 0.5% Triton X-100. Samples were then blocked with (PBS)/1% bovine serum albumin (BSA)/2% normal goat serum (NGS) for 30 min at RT. Slides were incubated with the primary antibodies 2 hr followed by an incubation with the secondary antibodies for 45 min at RT and mounting in Dako mounting medium (Dako). Appropriate controls were performed, omitting the primary antibodies. Polyclonal mouse *DNAL11* and monoclonal mouse *DNAI2* were purchased from Abcam (UK) and Abnova Corporation (Taiwan), respectively. Polyclonal *DNAH1* antibodies were purchased from Prestige Antibodies (Sigma-Aldrich). Monoclonal mouse anti-acetylated-α-tubulin were purchased from Sigma-Aldrich. Highly cross-adsorbed secondary antibodies (Alexa Fluor 488 and Alexa Fluor 546) were obtained from Molecular Probes (Invitrogen).

P3, suggesting that inner arms were mostly absent in this individual (Figure 2F). On the other hand, the antibodies directed against *DNAI2* stained the sperm flagella in both control and individual P3, suggesting that the outer dynein arms were not affected by the absence of *DNAH1* (Figures 2E and 2F). In order to confirm that the inner arms were disorganized, we studied the ultrastructure of individual P3's sperm by transmission electron microscopy (TEM) (Figure 3). We could observe 40 doublets of microtubules in cross sections presenting a sufficient quality to observe the dynein arms. Fifteen outer dynein arms (ODA) and only 4 inner dynein arms (IDA) were observed,

confirming the complete disorganization of the IDA. Moreover, approximately one third of the microtubule doublets were malformed or absent in the observed sections. Furthermore, the central singlet of microtubules was missing (9+0) in 47% of these sections. The fibrous sheath was also strongly disorganized in 90% of the sections (Figure 3).

After complete DNA sequencing of *DNAH1*, we identified two variants altering a consensus splice site, highly likely to have a damaging effect, in 3 out of 18 unrelated individuals. The predicted effect of the other two identified variants is not as clear but the addition of 21 residues

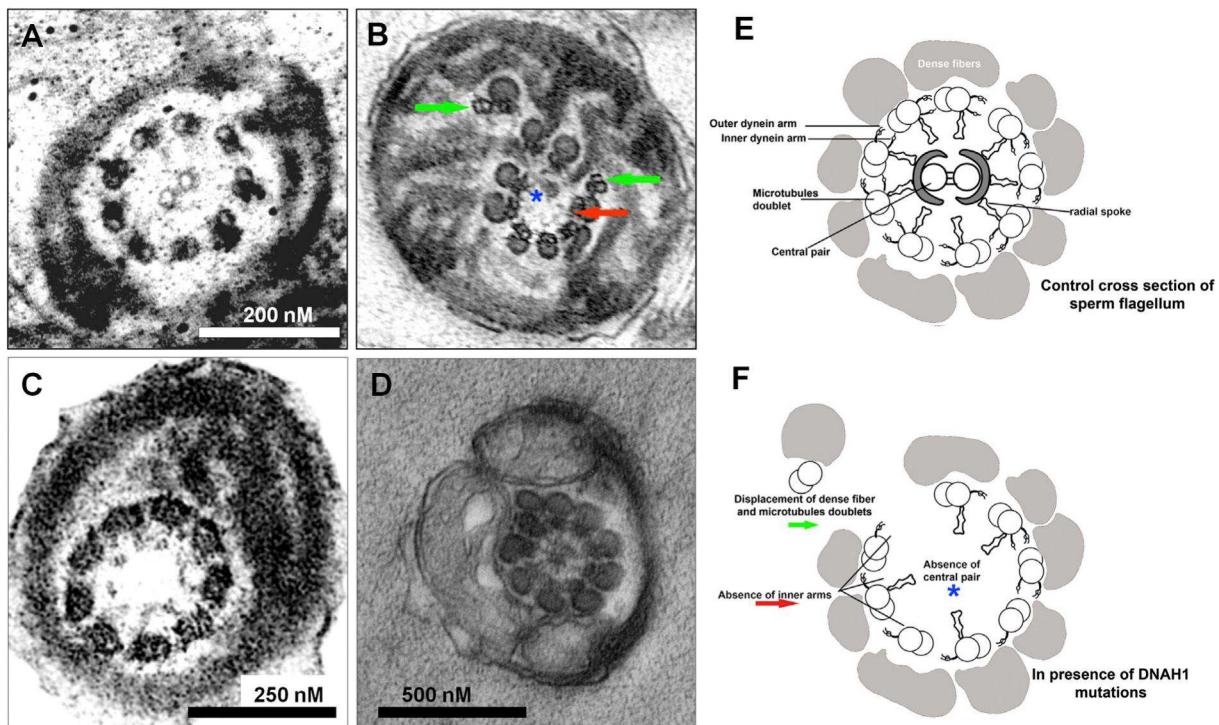


Figure 3. Electron Microscopy Analysis of Spermatozoa from P3, Carrying the c.11788–1G>A Variant, Reveals Numerous Ultrastructural Defects

- (A) EM cross-section of a flagellum from a control individual sperm sample.
 - (B) Cross-section from the individual P3 sample showing numerous defects: lack of IDA (red arrow) and axonemal disorganization with mislocalized peripheral doublets (green arrows) associated with a displacement of the central pair (blue asterisk).
 - (C) Cross-section from individual P3 sperm flagellum showing a complete absence of the central pair.
 - (D) Cross-section from individual P3 sperm flagellum showing supernumerary dense fibers with absence of mitochondrion on the right side of the mid-piece.
 - (E and F) Drawings describing the normal sperm axoneme ultrastructure with their different components (E) and different defects observed in *DNAH1* mutated subjects (F).
- Sperm cells were fixed with 2.5% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.4) during 2 hr at room temperature. Details of transmission electron microscopy technique were detailed previously.³

caused by the stop-loss variant is probably pathogenic. Interestingly, P1, P2, P3, P8, P9, and P17—who have inherited the variants predicted to have most severe effects—present 0% morphologically normal spermatozoa with a motility <2%, in contrast to P6, carrying the p.Asp1293Asn variant, who presents a milder phenotype with 35% motility and 6% morphologically normal spermatozoa (Table 1). These data therefore suggest that the c.3877G>A variant might be a hypomorphic allele, which is consistent with a single amino acid substitution in a large protein. Unfortunately we could not obtain any additional biological material from the other mutated subjects and in particular from P6 and could not assess the effects of the other variants on protein expression/localization and on the ultrastructure of the flagella. In addition, no mutations were identified in *DNAH1* in 13 subjects, suggesting that MMAF is genetically heterogeneous. We are currently sequencing the exomes of these 13 subjects in order to identify other genes involved in MMAF. We note that with the exception of P6, who carried a missense mutation and presented a milder form of the pathology, we included here only individuals with the most severe phenotypes.

We can therefore expect that individuals with intermediate asthenozoospermia and low levels of morphological anomalies could also harbor homozygous or compound heterozygous *DNAH1* mutations of moderate severity.

The data we present here are consistent with the phenotype described for *Dnahc1* knockout (KO) mice (the ortholog of *DNAH1*), which display asthenozoospermia and male infertility.¹⁹ In this model, however, no structural defects of the axoneme were observed either by optical or by transmission electronic microscopy.¹⁹ This contrasts with the strong axonemal disorganization we observed in sperm from P3 carrying the homozygous c.11788–1G>A mutation, where the inner dynein arms and the central pair of microtubules were mostly absent. In the *Dnahc1* KO, however, the authors describe that the targeted deletion did not lead to a complete disruption of the gene and resulted in a truncated protein with a preserved N terminus.¹⁹ Because the N-terminal part of the DyHCs plays a crucial role in the assembly and stabilization of the inner dynein arms, as shown in *Chlamydomonas* mutants,²⁴ it is likely that the formation of the base of the inner dynein arm is preserved and that the described

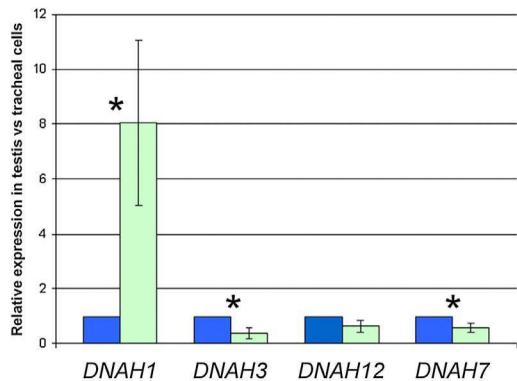


Figure 4. Relative mRNA Expression of *DNAH1*, *DNAH3*, *DNAH7*, and *DNAH12* in Testis and in Tracheal Cells

Expression of *DNAH1* mRNA (in green) in testis is significantly higher (8-fold) than in tracheal cells (in blue). Tracheal cell expression is set to 1. Expression of *DNAH3* and *DNAH7* mRNA are significantly lower in testis than in tracheal cells, whereas *DNAH12* expression is not significantly different. Data are presented as mean \pm standard deviation of three independent quantitative real-time PCR experiments. Statistical tests (paired t test) with a two-tailed p value ≤ 0.05 were considered as significant (*). Human testis and trachea cDNA were obtained from Amsbio (Abingdon). mRNA expression were assessed by qPCR with a Biorad CFX9 (Biorad). PCR primers used to amplify *DNAH1* and three other inner dynein arm heavy chain genes (*DNAH3*, *DNAH7*, and *DNAH12*) and the reference gene *ACTB* are listed in Table S5. The PCR cycle was as follows: 10 min 95°C, 1 cycle; 10 s 95°C, 30 s 58°C + fluorescence acquisition, 55 cycles. Analysis was performed with Biorad software CFX Manager v.3.0, with advanced relative quantification mode. Values for each gene were normalized to expression level of beta-actin gene (*ACTB*) via the 2- $\Delta\Delta CT$ method.²⁵ The 2- $\Delta\Delta CT$ value was set at 0 in tracheal cells, resulting in an arbitrary expression of 1.

KO animals could ensure correct axonemal biogenesis and organization. Alternatively, it is possible that the *DNAH1* role in axonemal structure is not as central in mouse as it is in human.

Apart from infertility, none of the 20 individuals declared suffering from any of the principal PCD symptoms such as an impairment of the respiratory functions. This suggests that *DNAH1* function in cilia is probably compensated by other HC dyneins. Previous phylogenetic studies indicate that *DNAH3* (MIM 603334), *DNAH7* (MIM 610061), and *DNAH12* (MIM 603340) are close paralogs to *DNAH1*, *DNAH12* being the closest.¹⁸ We therefore measured the relative expression of these four IDA heavy chains to assess whether the expression of these proteins could compensate the absence of *DNAH1* in other ciliated tissues. By using qPCR (primer sequences available in Table S6), we showed that *DNAH1* is expressed at a much higher level (8-fold) in the testis compared to the trachea (Figure 4). Conversely, the other IDA heavy chains are expressed at higher levels (*DNAH3*, *DNAH7*) or at a similar level (*DNAH12*) in control tracheal cells as in testis (Figure 4). Data available from public expression databases show a similar expression pattern to what was observed here. Moreover, these data show that whereas *DNAH3* and *DNAH7* expression is restricted to ciliated cells,

DNAH1 and *DNAH12* expression is rather atypical, because it is almost ubiquitous (EST profile viewer and GeneHub-GEPIS). We therefore note that not only are *DNAH1* and *DNAH12* closest to one another from a phylogenetic point of view, but they also share a broad expression pattern. We therefore believe that *DNAH12* is the most likely candidate for a potential functional compensation of *DNAH1* in ciliated cells. In addition and/or alternatively to this compensation, we cannot exclude the possibility that some of the affected individuals might retain expression of *DNAH1* in the ciliated cells of the trachea, although this is not expected for the most severe variants identified. Alternatively, we cannot exclude a reduction of ciliary beats, which could lead to a small decrease of cilia function in respiratory epithelium or in other ciliated tissues without pathological consequences, or at least none that have been noticed by the affected men themselves. We could not obtain nasal brushings or curette biopsies from affected individuals and therefore cannot formally exclude this possibility. Future work on *DNAH1* mutated subjects should include a thorough analysis of PCD symptoms including nasal nitric oxide measurements, video microscopy, and transcription electron microscopy. This would provide valuable information regarding the role of *DNAH1* in ciliated cells as well as indicate whether mutated men might be at risk of developing PCD symptoms, perhaps as late onset.

Inner dynein arms are organized in seven molecular complexes, viewed in electronic microscopy as globular heads arranged in 3-2-2 groups and corresponding to three different types of inner arms (IDA1 to IDA3, see Figure 5). In *Dnahc1* KO mice, electron microscopy studies indicated that one head of the IDA3 was missing, leading to a 3-2-1 globular head arrangement, suggesting that *DNAH1* is a component of IDA3. Radial spokes are present on microtubule doublets and interact with the inner arms. They allow a connection between external doublets of the microtubules and the two central microtubules. They are multiprotein complexes of more than 20 proteins. In mammals, there are three different radial spokes (RS1, RS2, and RS3) binding tightly to the inner arm bases of different IDAs. Among the different proteins involved in axonemal formation and organization, only mutations in *CCDC39* (MIM 613798) and *CCDC40* (MIM 613799) lead to a disorganization of the axonemal structure, a phenotype similar to what we observe in subjects with *DNAH1* mutations. *CCDC39* and *CCDC40* control the assembly of the dynein regulatory complex (DRC), a major regulatory node interacting with numerous axonemal structures.²⁷⁻²⁹ In the absence of DRC, RS2 anchoring is weakened, leading to the displacement or the absence of the central pair and the mislocalization of the peripheral doublets. Interestingly, in *T. thermophila*, the RS3 stalk is directly connected to the dynein d/a tail through an arc-like structure (Figure 5).²⁶ It can therefore be speculated that the absence of *DNAH1* removes the anchoring site of the radial spoke 3. As a consequence, the attachment of the two central

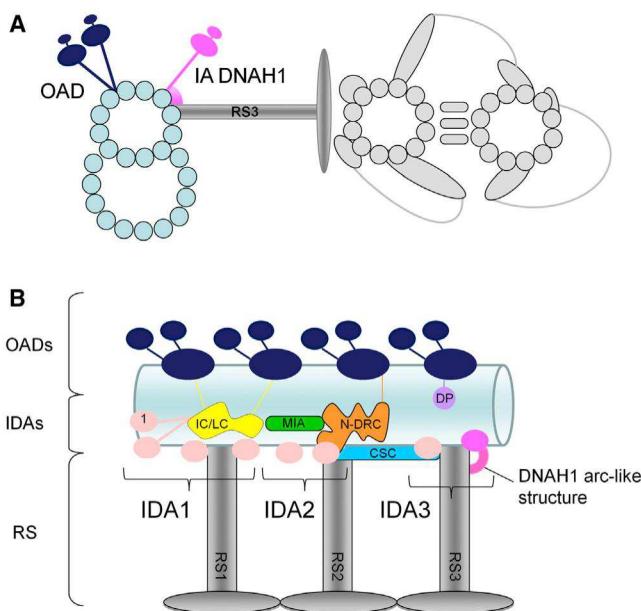


Figure 5. Proposed Schematic Model for the Location and the Function of the Inner Arm Heavy Chain DNAH1 in the Axoneme of Human Sperm Flagellum

(A) Simplified representation showing a cross-sectional view of one microtubule doublet of an axoneme surrounding the central pair complex; the viewing is from the flagellar base. Light gray, central pair complex; light blue, outer doublets; dark blue, outer arm dyneins (OAD); dark pink, inner arm dynein heavy chain DNAH1; dark gray, radial spoke 3.

(B) Longitudinal view illustrating the approximate localization on the outer doublet A-tubule of the various dyneins and regulatory structures within a single 96 nm axonemal repeat. RS3 stalk is directly connected to the DNAH1 tail through an arc-like structure. DNAH1 may therefore stabilize the RS3.²⁶ Light pink, inner arm dyneins (IAD); yellow, IC/LC, intermediate chain/light chain; orange, nexin-dynein regulatory complex (N-DRC); green, modifier of inner arms (MIA) complex; blue, calmodulin- and spoke-associated complex (CSC); purple, distal protrusion (DP).

singlet microtubules should be weakened. This scheme is in agreement with our data that show 47% of 9+0 axonemes (Figure 3).

As illustrated in Table S1, not all PCDs are associated with an infertility phenotype. For instance, subjects with mutations in *CCDC114* (MIM 615038)^{30,31} and *DNAH11* (MIM 603339)⁴ are fully fertile and could procreate spontaneously whereas subjects with *DNAAF2* (MIM 612517), *DNAH5* (MIM 603335), *DNAI1* (MIM 604366), or *HYDIN* (MIM 610812)^{4,32–34} present complete sperm immobility. Also, mutations in genes involved in the preassembly of the dynein arms *DNAAF1* (MIM 611390), *DNAAF2* (MIM 612517), *DNAAF3* (MIM 614566),^{32,35–37} and *LRRC6* (MIM 614930),³⁸ described to induce IDA loss, did not present axonemal microtubule disorganization. Consistently, IDAs are not or are only partially affected in deficient *Chlamydomonas* mutants for *ODA7* (*DNAAF1*), *ktu/pf13* (*DNAAF2*), and *pf22* (*DNAAF3*): in *ODA7* mutants, inner arms are not affected, in *ktu/pf13* mutants only IDA dynein c is missing, and in *pf22* mutants IDA dynein b and c are absent.^{27,29,36,37,39} Most importantly, IDA dynein f and

p28 remain located in the flagella. It can thus be speculated that there are several pathways for the preassembly and/or the targeting of the different mammalian IDAs, thus explaining the absence of axonemal microtubule disorganization in subjects presenting with mutations in *DNAAF1*, *DNAAF2*, *DNAAF3*, and *LRRC6*. Here we observed that *DNAL11* immunostaining was strongly reduced along the whole flagellum (Figure 2E), suggesting that *DNAL11* may be located mainly in IDA3.⁴⁰ In agreement with this result, p28, the *Chlamydomonas* ortholog of *DNAL11*, is associated with the inner dynein arm located in IDA3. Interestingly, *DNAL11* has an expression pattern similar to *DNAH1*: it presents a predominant testis expression and also a remarkable expression in nonciliated cells.⁴¹ Altogether, these facts suggest a close molecular partnership between *DNAH1* and *DNAL11*.

Mutations affecting axonemal components and/or axoneme assembly often result in PCD, which frequently includes a male infertility phenotype. In this study we describe that mutations in *DNAH1*, which codes for an axonemal component, leads to male infertility only with no other apparent PCD-associated syndromes. Our data indicate that *DNAH1* is required in spermatozoa for the formation of the inner dynein arms and that its absence is deleterious for the organization and biogenesis of the axoneme. Overall our data confirm that despite close structural similarities, sperm flagella and cilia present important divergences in axonemal organization and biogenesis.

Supplemental Data

Supplemental Data include three figures and six tables and can be found with this article online at <http://www.cell.com/AJHG/>.

Acknowledgments

This work was supported by the research grant ICG2I funded by the program GENOPAT 2009 from the French Research Agency (ANR). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. We are grateful to Frédéric Plewniak, IGBMC, Strasbourg, who developed the homoSNP software used in this study (the software is available on request to plewniak@igbmc.u-strasbg.fr).

Received: July 19, 2013

Accepted: November 18, 2013

Published: December 19, 2013

Web Resources

The URLs for data presented herein are as follows:

1000 Genomes, <http://browser.1000genomes.org>

Berkeley Drosophila Genome Project NNSplice 0.9, http://www.fruitfly.org/seq_tools/splice.html

dbSNP v.137, http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi

GeneHub-GEPIS, <http://research-public.gene.com/Research/genetech/genehub-gepis/>

NHLBI Exome Sequencing Project (ESP) Exome Variant Server, <http://evs.gs.washington.edu/EVS/>
Online Mendelian Inheritance in Man (OMIM), <http://www.omim.org/>
PolyPhen-2, <http://www.genetics.bwh.harvard.edu/pph2/>
RefSeq, <http://www.ncbi.nlm.nih.gov/RefSeq>
SIFT, http://sift.jcvi.org/www/SIFT_chr_coords_submit.html
UniGene, <http://www.ncbi.nlm.nih.gov/unigene>
UCSC Genome Browser, <http://genome.ucsc.edu>

References

- Boivin, J., Bunting, L., Collins, J.A., and Nygren, K.G. (2007). International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Hum. Reprod.* **22**, 1506–1512.
- Dieterich, K., Soto Rifo, R., Faure, A.K., Hennebicq, S., Ben Amar, B., Zahi, M., Perrin, J., Martinez, D., Sèle, B., Jouk, P.S., et al. (2007). Heterozygous mutation of AURKC yields large-headed polyploid spermatozoa and causes male infertility. *Nat. Genet.* **39**, 661–665.
- Harbuz, R., Zouari, R., Pierre, V., Ben Khelifa, M., Kharouf, M., Coutton, C., Merdassi, G., Abada, F., Escoffier, J., Nikas, Y., et al. (2011). A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. *Am. J. Hum. Genet.* **88**, 351–361.
- Schwabe, G.C., Hoffmann, K., Loges, N.T., Birker, D., Rossier, C., de Santi, M.M., Olbrich, H., Fliegauf, M., Failly, M., Liebers, U., et al. (2008). Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations. *Hum. Mutat.* **29**, 289–298.
- Rawe, V.Y., Galaverna, G.D., Acosta, A.A., Olmedo, S.B., and Chemes, H.E. (2001). Incidence of tail structure distortions associated with dysplasia of the fibrous sheath in human spermatozoa. *Hum. Reprod.* **16**, 879–886.
- Olmedo, S.B., Rawe, V.Y., Nodar, F.N., Galaverna, G.D., Acosta, A.A., and Chemes, H.E. (2000). Pregnancies established through intracytoplasmic sperm injection (ICSI) using spermatozoa with dysplasia of fibrous sheath. *Asian J. Androl.* **2**, 125–130.
- Chemes, H.E., Brugo, S., Zanchetti, F., Carrere, C., and Lavieri, J.C. (1987). Dysplasia of the fibrous sheath: an ultrastructural defect of human spermatozoa associated with sperm immotility and primary sterility. *Fertil. Steril.* **48**, 664–669.
- Neugebauer, D.C., Neuwinger, J., Jockenhövel, F., and Nieschlag, E. (1990). '9 + 0' axoneme in spermatozoa and some nasal cilia of a patient with totally immotile spermatozoa associated with thickened sheath and short midpiece. *Hum. Reprod.* **5**, 981–986.
- Escalier, D., and Albert, M. (2006). New fibrous sheath anomaly in spermatozoa of men with consanguinity. *Fertil. Steril.* **86**, e1–e9.
- Escalier, D. (2006). Arrest of flagellum morphogenesis with fibrous sheath immaturity of human spermatozoa. *Andrologia* **38**, 54–60.
- Escalier, D., Gallo, J.M., and Schrével, J. (1997). Immunohistochemical characterization of a human sperm fibrous sheath protein, its developmental expression pattern, and morphogenetic relationships with actin. *J. Histochem. Cytochem.* **45**, 909–922.
- David, G., Feneux, D., Serres, C., Escalier, D., and Jouannet, P. (1993). [A new entity of sperm pathology: peri-axonemal flagellar dyskinesia]. *Bull. Acad. Natl. Med.* **177**, 263–271, discussion 272–275.
- Stalf, T., Sánchez, R., Köhn, F.M., Schalles, U., Kleinsteiner, J., Hinz, V., Tielsch, J., Khanaga, O., Turley, H., Gips, H., et al. (1995). Pregnancy and birth after intracytoplasmic sperm injection with spermatozoa from a patient with tail stump syndrome. *Hum. Reprod.* **10**, 2112–2114.
- Moretti, E., Geminiani, M., Terzuoli, G., Renieri, T., Pasarelli, N., and Collodel, G. (2011). Two cases of sperm immotility: a mosaic of flagellar alterations related to dysplasia of the fibrous sheath and abnormalities of head-neck attachment. *Fertil. Steril.* **95**, e19–e23.
- Marmor, D., and Grob-Menendez, F. (1991). Male infertility due to asthenozoospermia and flagellar anomaly: detection in routine semen analysis. *Int. J. Androl.* **14**, 108–116.
- World Health Organization (2010). WHO Laboratory Manual for the Examination and Processing of Human Semen, fifth ed.
- Demonchy, R., Blisnick, T., Deprez, C., Toutirais, G., Loussert, C., Marande, W., Grellier, P., Bastin, P., and Kohl, L. (2009). Kinesin 9 family members perform separate functions in the trypanosome flagellum. *J. Cell Biol.* **187**, 615–622.
- Maiti, A.K., Mattéi, M.G., Jorissen, M., Volz, A., Zeigler, A., and Bouvagnet, P. (2000). Identification, tissue specific expression, and chromosomal localisation of several human dynein heavy chain genes. *Eur. J. Hum. Genet.* **8**, 923–932.
- Neesen, J., Kirschner, R., Ochs, M., Schmiedl, A., Habermann, B., Mueller, C., Holstein, A.F., Nuesslein, T., Adham, I., and Engel, W. (2001). Disruption of an inner arm dynein heavy chain gene results in asthenozoospermia and reduced ciliary beat frequency. *Hum. Mol. Genet.* **10**, 1117–1128.
- Shao, X., Tarnasky, H.A., Lee, J.P., Oko, R., and van der Hoorn, F.A. (1999). Spag4, a novel sperm protein, binds outer dense-fiber protein Odf1 and localizes to microtubules of manchette and axoneme. *Dev. Biol.* **211**, 109–123.
- Habura, A., Tikhonenko, I., Chisholm, R.L., and Koonce, M.P. (1999). Interaction mapping of a dynein heavy chain. Identification of dimerization and intermediate-chain binding domains. *J. Biol. Chem.* **274**, 15447–15453.
- Moore, D.J., Onoufriadis, A., Shoemark, A., Simpson, M.A., Zur Lage, P.I., de Castro, S.C., Bartoloni, L., Gallone, G., Petridi, S., Woppard, W.J., et al. (2013). Mutations in ZMYND10, a gene essential for proper axonemal assembly of inner and outer dynein arms in humans and flies, cause primary ciliary dyskinesia. *Am. J. Hum. Genet.* **93**, 346–356.
- Höök, P., and Vallee, R. (2012). Dynein dynamics. *Nat. Struct. Mol. Biol.* **19**, 467–469.
- Myster, S.H., Knott, J.A., Wysocki, K.M., O'Toole, E., and Porter, M.E. (1999). Domains in the 1alpha dynein heavy chain required for inner arm assembly and flagellar motility in *Chlamydomonas*. *J. Cell Biol.* **146**, 801–818.
- Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408.
- King, S.M. (2013). A solid-state control system for dynein-based ciliary/flagellar motility. *J. Cell Biol.* **201**, 173–175.
- Antony, D., Becker-Heck, A., Zariwala, M.A., Schmidt, M., Onoufriadis, A., Forouhan, M., Wilson, R., Taylor-Cox, T., Dewar, A., Jackson, C., et al.; Uk10k (2013). Mutations in CCDC39 and CCDC40 are the major cause of primary ciliary dyskinesia with axonemal disorganization and absent inner dynein arms. *Hum. Mutat.* **34**, 462–472.

28. Becker-Heck, A., Zohn, I.E., Okabe, N., Pollock, A., Lenhart, K.B., Sullivan-Brown, J., McSheene, J., Loges, N.T., Olbrich, H., Häffner, K., et al. (2011). The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat. Genet.* **43**, 79–84.
29. Merveille, A.C., Davis, E.E., Becker-Heck, A., Legendre, M., Amirav, I., Bataille, G., Belmont, J., Beydon, N., Billen, F., Clément, A., et al. (2011). CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. *Nat. Genet.* **43**, 72–78.
30. Onoufriadis, A., Paff, T., Antony, D., Shoemark, A., Micha, D., Kuyt, B., Schmidts, M., Petridi, S., Dankert-Roelse, J.E., Haarmann, E.G., et al.; UK10K (2013). Splice-site mutations in the axonemal outer dynein arm docking complex gene CCDC114 cause primary ciliary dyskinesia. *Am. J. Hum. Genet.* **92**, 88–98.
31. Knowles, M.R., Leigh, M.W., Ostrowski, L.E., Huang, L., Carson, J.L., Hazucha, M.J., Yin, W., Berg, J.S., Davis, S.D., Dell, S.D., et al.; Genetic Disorders of Mucociliary Clearance Consortium (2013). Exome sequencing identifies mutations in CCDC114 as a cause of primary ciliary dyskinesia. *Am. J. Hum. Genet.* **92**, 99–106.
32. Omran, H., Kobayashi, D., Olbrich, H., Tsukahara, T., Loges, N.T., Hagiwara, H., Zhang, Q., Leblond, G., O'Toole, E., Hara, C., et al. (2008). Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins. *Nature* **456**, 611–616.
33. Fliegauf, M., Olbrich, H., Horvath, J., Wildhaber, J.H., Zarivala, M.A., Kennedy, M., Knowles, M.R., and Omran, H. (2005). Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. *Am. J. Respir. Crit. Care Med.* **171**, 1343–1349.
34. Olbrich, H., Schmidts, M., Werner, C., Onoufriadis, A., Loges, N.T., Raidt, J., Banki, N.F., Shoemark, A., Burgoyne, T., Al Turk, S., et al.; UK10K Consortium (2012). Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry. *Am. J. Hum. Genet.* **91**, 672–684.
35. Loges, N.T., Olbrich, H., Becker-Heck, A., Häffner, K., Heer, A., Reinhard, C., Schmidts, M., Kispert, A., Zarivala, M.A., Leigh, M.W., et al. (2009). Deletions and point mutations of LRRC50 cause primary ciliary dyskinesia due to dynein arm defects. *Am. J. Hum. Genet.* **85**, 883–889.
36. Duquesnoy, P., Escudier, E., Vincensini, L., Freshour, J., Bridoux, A.M., Coste, A., Deschildre, A., de Blic, J., Legendre, M., Montantin, G., et al. (2009). Loss-of-function mutations in the human ortholog of *Chlamydomonas reinhardtii* ODA7 disrupt dynein arm assembly and cause primary ciliary dyskinesia. *Am. J. Hum. Genet.* **85**, 890–896.
37. Mitchison, H.M., Schmidts, M., Loges, N.T., Freshour, J., Dritsoula, A., Hirst, R.A., O'Callaghan, C., Blau, H., Al Dabbagh, M., Olbrich, H., et al. (2012). Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nat. Genet.* **44**, 381–389, S1–S2.
38. Kott, E., Duquesnoy, P., Copin, B., Legendre, M., Dastot-Le Moal, F., Montantin, G., Jeanson, L., Tamalet, A., Papon, J.F., Siffroi, J.P., et al. (2012). Loss-of-function mutations in LRRC6, a gene essential for proper axonemal assembly of inner and outer dynein arms, cause primary ciliary dyskinesia. *Am. J. Hum. Genet.* **91**, 958–964.
39. Freshour, J., Yokoyama, R., and Mitchell, D.R. (2007). *Chlamydomonas* flagellar outer row dynein assembly protein ODA7 interacts with both outer row and I1 inner row dyneins. *J. Biol. Chem.* **282**, 5404–5412.
40. Mazor, M., Alkrinawi, S., Chalifa-Caspi, V., Manor, E., Sheffield, V.C., Aviram, M., and Parvari, R. (2011). Primary ciliary dyskinesia caused by homozygous mutation in DNAL1, encoding dynein light chain 1. *Am. J. Hum. Genet.* **88**, 599–607.
41. Rashid, S., Breckle, R., Hupe, M., Geisler, S., Doerwald, N., and Neesen, J. (2006). The murine Dnali1 gene encodes a flagellar protein that interacts with the cytoplasmic dynein heavy chain 1. *Mol. Reprod. Dev.* **73**, 784–794.

Principaux résultats

Nous avons pu identifier par cartographie d'homozygotie cinq mutations homozygotes dans le gène *DNAH1* parmi les 18 patients non-apparentés (28%) présentant un phénotype anomalies morphologiques multiples des flagelles. Le gène *DNAH1* code pour une chaîne lourde des bras internes de dynéine exprimée dans de nombreux tissus et fortement dans le testicule. Parmi ces différentes mutations, on retrouve deux mutations situées respectivement sur des sites consensus donneurs et accepteurs d'épissage, une mutation run-on abrogeant le codon stop et une mutation faux-sens potentiellement délétère. La mutation localisée sur un site accepteur d'épissage a été retrouvée chez quatre patients dont les 3 frères inclus dans la cohorte.

Les études d'expression chez un des patients porteurs de la mutation d'épissage c.11788-1G>A ont démontré une absence totale de transcrits dans les lymphocytes. Les analyses par immunofluorescence chez ce patient ont confirmé la disparition complète de la protéine DNAH1 au niveau de l'axonème du flagelle du sperme confirmant le caractère pathogène du variant. Le co-marquage avec des protéines marquant les bras externes de dynéines DNAI2 ne montre pas de disparition de ces structures. Inversement, le co-marquage avec la protéine DNALI1 localisée dans les bras internes de dynéine a montré une diminution quasi-totale du signal. Enfin, une désorganisation générale et sévère de l'axonème incluant une disparition des doublets centraux et des bras internes de dynéine a été observée par microscopie électronique.

Discussion et Perspectives

Au final, nous avons identifié *DNAH1* comme le premier gène codant pour une dynéine axonémale responsable uniquement d'infertilité masculine. L'ensemble de ces résultats suggère que *DNAH1* est une protéine primordiale dans la biogénèse et l'organisation de l'axonème du flagelle du sperme. Des études fonctionnelles complètes chez d'autres modèles (murins ou protistes) permettraient de préciser ses interactions avec d'autres protéines axonémiales et de comprendre son rôle moteur et surtout structural au sein de l'axonème du flagelle.

Bien que *DNAH1* soit exprimé dans d'autres cellules ciliées, aucun autre symptôme relatif aux ciliopathies, à part l'infertilité, n'a été observé chez ces patients. Les analyses par

RT-qPCR réalisées montrent un profil d'expression de *DNAH1* prédominant au niveau testiculaire par rapport à d'autres tissus ciliés et aux autres dynéines axonémiales des bras internes proches phylogénétiquement de *DNAH1*. Ces données renforcent l'hypothèse que *DNAH1* a un rôle clé dans le testicule mais que son implication n'est pas essentielle dans le fonctionnement du cil d'autres tissus, ou elle est potentiellement remplacée par d'autres dynéines. Des études plus approfondies notamment au niveau de cils des cellules des voies respiratoires chez les patients mutés pour *DNAH1* seraient utiles pour répondre définitivement cette question.

Les mutations dans le gène *DNAH1* n'ont été retrouvée « que » dans environ un tiers des patients étudiés. Ceci démontre clairement l'implication d'autres gènes dans ce phénotype. De nouvelles investigations génétiques, notamment par l'intermédiaire des nouvelles approches comme le séquençage nouvelle génération devraient sans doute permettre d'identifier de nouveaux gènes en cause. Enfin, des études du gène *DNAH1* sur une plus large cohorte de patients avec un phénotype AMMF seraient intéressantes afin de mettre en évidence de nouvelles mutations et de préciser la prévalence des mutations *DNAH1* dans ce phénotype.

The American Journal of Human Genetics, Volume 94

Supplemental Data

Mutations in *DNAH1*, which Encodes an Inner Arm Heavy Chain Dynein, Lead to Male Infertility from Multiple Morphological Abnormalities of the Sperm Flagella

Mariem Ben Khelifa, Charles Coutton, Raoudha Zouari, Thomas Karaouzène, John Rendu, Marie Bidart, Sandra Yassine, Virginie Pierre, Julie Delaroche, Sylviane Hennebicq, Didier Grunwald, Denise Escalier, Karine Pernet-Gallay, Pierre-Simon Jouk, Nicolas Thierry-Mieg, Aminata Touré, Christophe Arnoult, and Pierre F. Ray

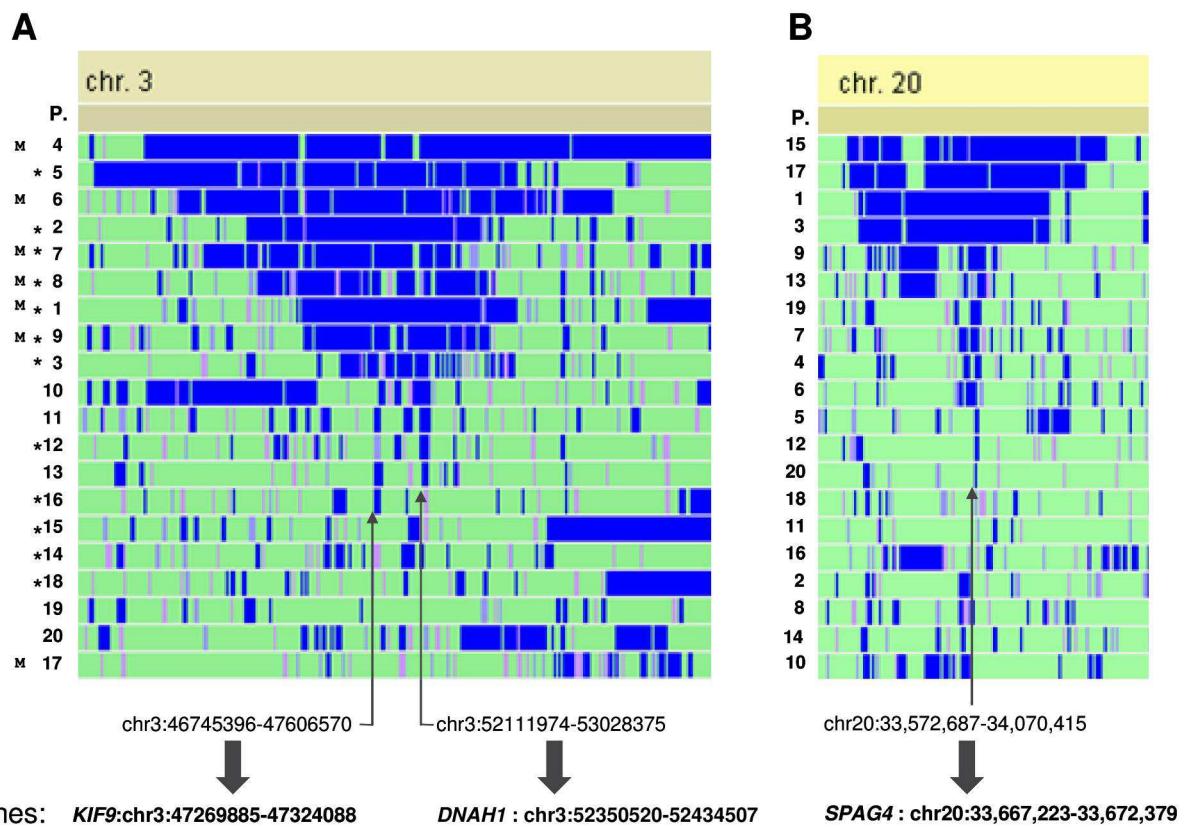


Figure S1. Localization of regions of homozygosity and identification of candidate genes. Graphic representation of SNP Array results with homoSNP software (developed by F Plewniak IGBMC- Strasbourg, software available on request to plewniak@igbmc.u-strasbg.fr). Regions of homozygosity greater than 45 SNPs are shown in blue. Coordinates of selected homozygous regions and genes are based on the UCSC server version GRCh37/hg17. Subject's reference number is indicated on the left. The presence of an asterisk (*) next to the subject's reference indicates that consanguinity had been signaled by the subject. The M indicates that a homozygous mutation was identified.

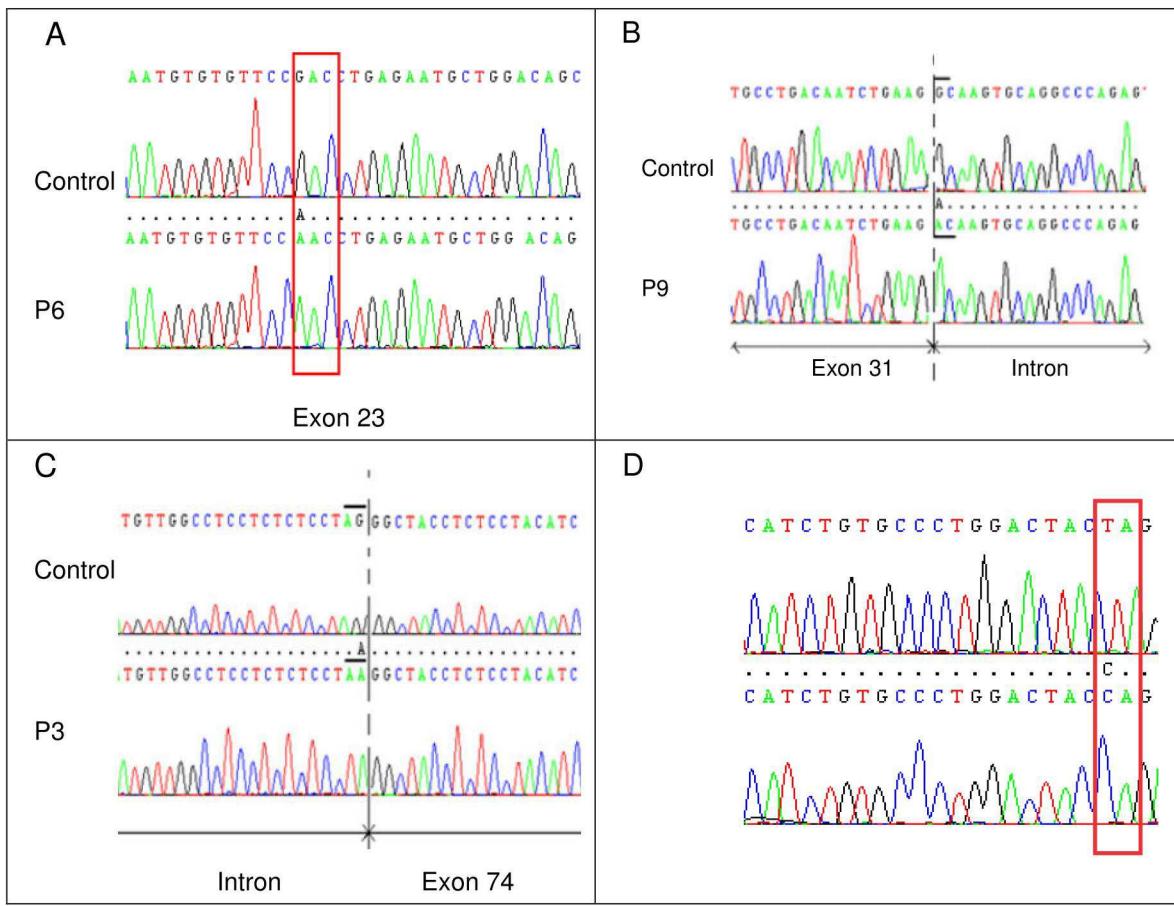


Figure S2. Electropherogram analyses of *DNAH1* mutations.

(A) Electropherogram showing part of *DNAH1* exon 23 of a control subject (Control) and the c.3877G>A homozygous individual (P6). (B) Electropherogram showing part of *DNAH1* exon 31 for a control subject and a c.5094+1G>A homozygous individual (P9). (C) Electropherogram showing part of *DNAH1* exon 74 for a control subject and a c.11788-1G>A homozygous individual (P3). (D) Electropherogram showing part of *DNAH1* exon 78 for a control subject and a c.12796 T>C homozygous individual (P8).

PCR primers used to amplify the different exons of *DNAH1* are listed in Table S2. Thirty five cycles of PCR amplification were carried out using Taq DNA polymerase (Qiagene, Courtaboeuf, France). Sequencing reactions were carried out with BigDye Terminator v3.1 (Applied Biosystems, Villebon sur Yvette, France). All sequences electrophoresis were carried out on ABI 3130XL (Applied Biosystems).

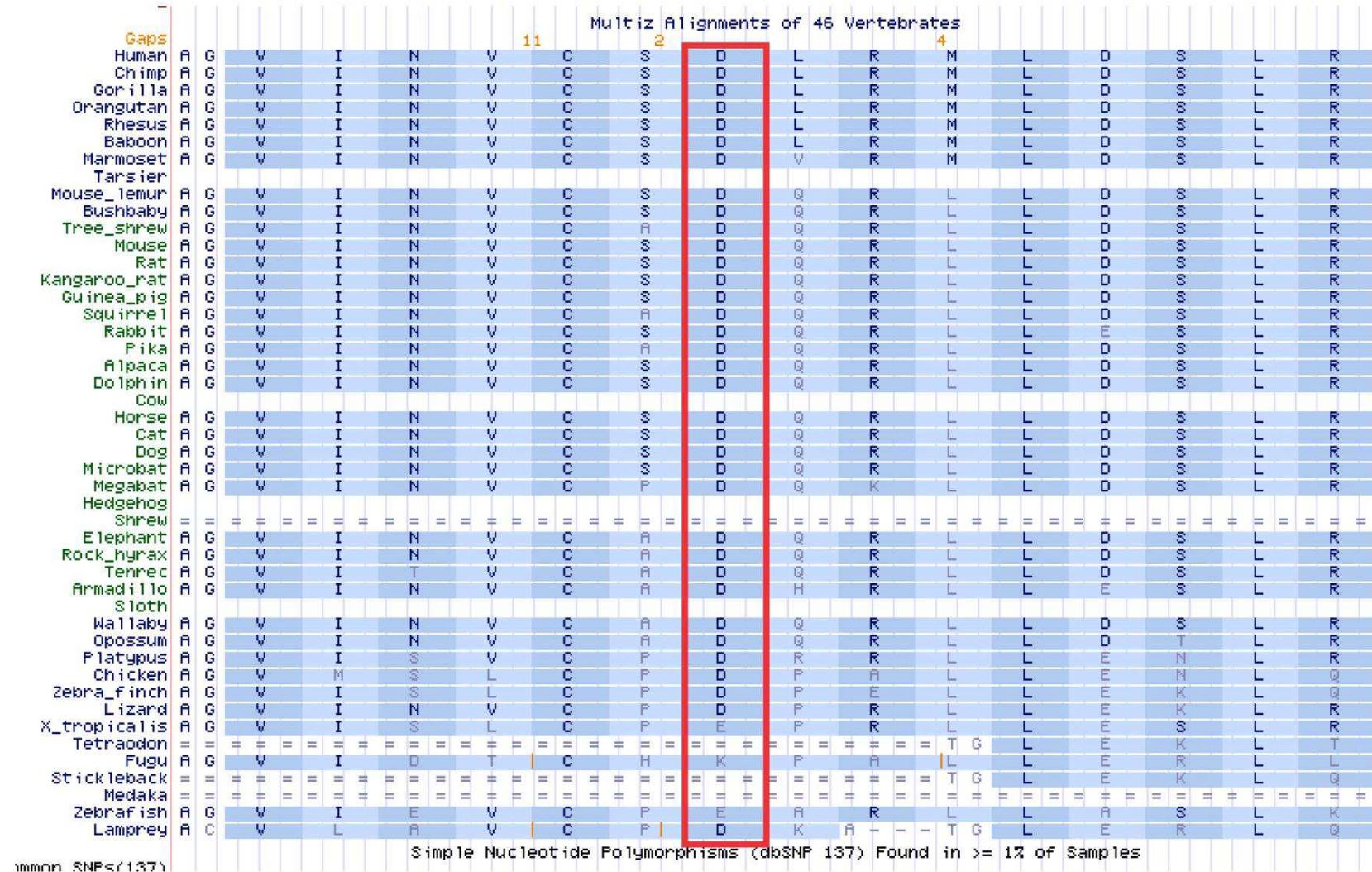


Figure S3. DNAH1 interspecies conservation at the p.Asp1293 amino-acid.

The 1293 amino acid and its corresponding counterparts in different species are framed by a black rectangle. Alignment was obtained from :

http://genome-euro.ucsc.edu/cgi-bin/hgTracks?position=chr3:52391628-52391670&db=hg19&ss=-./trash/hgSs/hgSs_genome_euro_5e1d_68d7c0.pslx+./trash/hgSs/hgSs_genome_euro_5e1d_68d7c0.fa&hgsid=194104129

	Gene name	MIM Number	Axonemal localization	Axonemal defects	Infertility phenotype	Ref
<i>Absence of ODA</i>	<i>ARMC4</i>	615408	proper targeting and anchoring of ODA	Reduced ODA	No data available	¹
	<i>CCDC114</i>	615038	DCC2, an ODA microtubule-docking complex	ODA defects	Male are fertile	^{2 3} ,
	<i>CCDC103</i>	614679	Pre-assembly of ODA	Reduced ODA	No data available	⁴
	<i>DNAH5</i>	603335	ODA component	ODA defect	sperm immobility	⁵
	<i>DNAI1</i>	604366	ODA component	ODAs are shortened or missing.	immotile spermatozoa	⁶
	<i>DNAI2</i>	605483	ODA component	ODA defects	male infertility confirmed but no sperm analysis available	⁷
	<i>DNAL1</i>	610062	ODA component	Absence of ODA	No data available	⁸
<i>Absence of ODA and IDA</i>	<i>TXND3</i>	607421	Non precisely determined	Shortened or absent ODA	No data available	⁹
	<i>CCDC103</i>	614677	dynein arm attachment factor		No data available	¹⁰
	<i>DNAAF1 (LRRC50)</i>	613193	Pre-assembly of dynein arm complexes in the cytoplasm	Absence of both ODA and IDA	Sterility for men	^{11, 12}
	<i>DNAAF2 (KTU)</i>	612518	Pre-assembly of dynein arm complexes in the cytoplasm	Absence of both ODA and IDA	Complete immotility	¹³
	<i>DNAAF3</i>	606763	Pre-assembly of dynein arm complexes in the cytoplasm	Absence of both ODA and IDA Abnormal ultrastructure.	Male infertility reported	¹⁴
	<i>DYX1C1</i>	608706	Preassembly or stability of axonemal dynein arms	Absence of both ODA and IDA	Reduced fertility to complete infertility	¹⁵
	<i>HEATR2</i>	614864	Preassembly or stability of axonemal dynein arms	Absence of ODA	No data available	¹⁶
	<i>LRRC6</i>	614930	Pre-assembly of dynein arm complexes in the cytoplasm	Absence of both outer and inner dynein arms	oligoasthenospermia.	¹⁷
	<i>RGPR</i>	300455	Involved in the transitional zone of motile cilia in airway epithelia /localized to centrioles	Absence of both ODA and IDA Abnormal central complex	No data available	¹⁸
	<i>SPAG1</i>	603395	Pre-assembly of dynein arm complexes in the cytoplasm	Absence of both ODA and IDA	No data available	¹⁹
<i>IDA and disorganization</i>	<i>ZMYND10</i>	607070	Pre-assembly of dynein arm complexes in the cytoplasm	Absence of both ODA and IDA	No human data Sperm immobility in mutant Drosophila	²⁰
	<i>CCDC39</i>	613807	Pre-assembly of IDA, DRC and Radial spokes	Absence of IDA Axonemal disorganization with mislocalized peripheral doublet Displacement or absence of the central pair	Oligoasthenospermia midpiece is narrowed flagellum is shortened	²¹

	<i>CCDC40</i>	613808	Govern the assembly of DRC and inner dynein arm complexes, but not outer dynein arm complexes	Axonemal disorganization with mislocalized peripheral doublet Displacement or absence of the central pair	No data available	22
Defective central complex	<i>HYDIN</i>	610812	C2b projection	Absence of C2b projection towards the central pair	Sperm immobility	23
	<i>RSPH4A</i>	612647	Radial spoke component	No data available	No data available	24
	<i>RSPH9</i>	612648	Radial spoke component	No data available	No data available	24
	<i>RSPH1</i>	609314	Radial spoke component	Absence of central pair	Asthenozoospermia	25
No structural defect	<i>DNAH11</i>	603339	ODA component	ultrastructure respiratory cilia is normal and outer dynein arms are intact	Male are fertile	26
	<i>DRC1 (CCDC164)</i>	615288	Nexin link component	Absence of nexin links	No data available	27

Table S1. PCD gene mutations and their consequences on axonemal ultrastructure and sperm motility. (ODA, IDA = outer and inner dynein arms respectivem; DRC = dynein regulatory complex).

Table S2 References

1. Hjeij R, Lindstrand A, Francis R, Zariwala MA, Liu X, Li Y, Damerla R, Dougherty GW, Abouhamed M, Olbrich H et al (2013) ARMC4 Mutations Cause Primary Ciliary Dyskinesia with Randomization of Left/Right Body Asymmetry. Am J Hum Genet
2. Onoufriadis A, Paff T, Antony D, Shoemark A, Micha D, Kuyt B, Schmidts M, Petridi S, Dankert-Roelse JE, Haarman EG et al (2013) Splice-site mutations in the axonemal outer dynein arm docking complex gene CCDC114 cause primary ciliary dyskinesia. Am J Hum Genet 92 (1):88-98
3. Knowles MR, Leigh MW, Ostrowski LE, Huang L, Carson JL, Hazucha MJ, Yin W, Berg JS, Davis SD, Dell SD et al (2013) Exome sequencing identifies mutations in CCDC114 as a cause of primary ciliary dyskinesia. Am J Hum Genet 92 (1):99-106
4. Panizzi JR, Becker-Heck A, Castleman VH, Al-Mutairi DA, Liu Y, Loges NT, Pathak N, Austin-Tse C, Sheridan E, Schmidts M et al (2012) CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. Nat Genet 44 (6):714-719
5. Fliegauf M, Olbrich H, Horvath J, Wildhaber JH, Zariwala MA, Kennedy M, Knowles MR, and Omran H (2005) Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. Am J Respir Crit Care Med 171 (12):1343-1349
6. Guichard C, Harricane MC, Lafitte JJ, Godard P, Zaegel M, Tack V, Lalau G, and Bouvagnet P (2001) Axonemal dynein intermediate-chain gene (DNAI1) mutations result in situs inversus and primary ciliary dyskinesia (Kartagener syndrome). Am J Hum Genet 68 (4):1030-1035

7. Loges NT, Olbrich H, Fenske L, Mussaffi H, Horvath J, Fliegauf M, Kuhl H, Baktai G, Peterffy E, Chodhari R et al (2008) DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. *Am J Hum Genet* 83 (5):547-558
8. Mazor M, Alkrinawi S, Chalifa-Caspi V, Manor E, Sheffield VC, Aviram M, and Parvari R (2011) Primary ciliary dyskinesia caused by homozygous mutation in DNAL1, encoding dynein light chain 1. *Am J Hum Genet* 88 (5):599-607
9. Duriez B, Duquesnoy P, Escudier E, Bridoux AM, Escalier D, Rayet I, Marcos E, Vojtek AM, Bercher JF, and Amselem S (2007) A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia. *Proc Natl Acad Sci U S A* 104 (9):3336-3341
10. Panizzi JR, Becker-Heck A, Castleman VH, Al-Mutairi DA, Liu Y, Loges NT, Pathak N, Austin-Tse C, Sheridan E, Schmidts M et al (2012) CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. *Nat Genet* 44 (6):714-719
11. Loges NT, Olbrich H, Becker-Heck A, Haffner K, Heer A, Reinhard C, Schmidts M, Kispert A, Zariwala MA, Leigh MW et al (2009) Deletions and point mutations of LRRC50 cause primary ciliary dyskinesia due to dynein arm defects. *Am J Hum Genet* 85 (6):883-889
12. Duquesnoy P, Escudier E, Vincensini L, Freshour J, Bridoux AM, Coste A, Deschildre A, de BJ, Legendre M, Montantin G et al (2009) Loss-of-function mutations in the human ortholog of Chlamydomonas reinhardtii ODA7 disrupt dynein arm assembly and cause primary ciliary dyskinesia. *Am J Hum Genet* 85 (6):890-896
13. Omran H, Kobayashi D, Olbrich H, Tsukahara T, Loges NT, Hagiwara H, Zhang Q, Leblond G, O'Toole E, Hara C et al (2008) Ktu/PF13 is required for cytoplasmic pre-assembly of axonemal dyneins. *Nature* 456 (7222):611-616
14. Mitchison HM, Schmidts M, Loges NT, Freshour J, Dritsoula A, Hirst RA, O'Callaghan C, Blau H, Dabbagh MA, Olbrich H et al (2012) Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nat Genet* 44 (4):381-3S2
15. Tarkar A, Loges NT, Slagle CE, Francis R, Dougherty GW, Tamayo JV, Shook B, Cantino M, Schwartz D, Jahnke C et al (2013) DYX1C1 is required for axonemal dynein assembly and ciliary motility. *Nat Genet* 45 (9):995-1003
16. Horani A, Druley TE, Zariwala MA, Patel AC, Levinson BT, Van Arendonk LG, Thornton KC, Giacalone JC, Albee AJ, Wilson KS et al (2012) Whole-exome capture and sequencing identifies HEATR2 mutation as a cause of primary ciliary dyskinesia. *Am J Hum Genet* 91 (4):685-693
17. Kott E, Duquesnoy P, Copin B, Legendre M, Dastot-Le MF, Montantin G, Jeanson L, Tamalet A, Papon JF, Siffroi JP et al (2012) Loss-of-Function Mutations in LRRC6, a Gene Essential for Proper Axonemal Assembly of Inner and Outer Dynein Arms, Cause Primary Ciliary Dyskinesia. *Am J Hum Genet* 91 (5):958-964
18. Moore A, Escudier E, Roger G, Tamalet A, Pelosse B, Marlin S, Clement A, Geremek M, Delaisi B, Bridoux AM et al (2006) RPGR is mutated in patients with a complex X linked phenotype combining primary ciliary dyskinesia and retinitis pigmentosa. *J Med Genet* 43 (4):326-333

19. Knowles MR, Ostrowski LE, Loges NT, Hurd T, Leigh MW, Huang L, Wolf WE, Carson JL, Hazucha MJ, Yin W et al (2013) Mutations in SPAG1 Cause Primary Ciliary Dyskinesia Associated with Defective Outer and Inner Dynein Arms. *Am J Hum Genet* 93 (4):711-720
20. Moore DJ, Onoufriadiis A, Shoemark A, Simpson MA, Zur Lage PI, de Castro SC, Bartoloni L, Gallone G, Petridi S, Woppard WJ et al (2013) Mutations in ZMYND10, a Gene Essential for Proper Axonemal Assembly of Inner and Outer Dynein Arms in Humans and Flies, Cause Primary Ciliary Dyskinesia. *Am J Hum Genet*
21. Merveille AC, Davis EE, Becker-Heck A, Legendre M, Amirav I, Bataille G, Belmont J, Beydon N, Billen F, Clement A et al (2011) CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. *Nat Genet* 43 (1):72-78
22. Becker-Heck A, Zohn IE, Okabe N, Pollock A, Lenhart KB, Sullivan-Brown J, McSheene J, Loges NT, Olbrich H, Haeffner K et al (2011) The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat Genet* 43 (1):79-84
23. Olbrich H, Schmidts M, Werner C, Onoufriadiis A, Loges NT, Raidt J, Banki NF, Shoemark A, Burgoyne T, Al TS et al (2012) Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry. *Am J Hum Genet* 91 (4):672-684
24. Castleman VH, Romio L, Chodhari R, Hirst RA, de Castro SC, Parker KA, Ybot-Gonzalez P, Emes RD, Wilson SW, Wallis C et al (2009) Mutations in radial spoke head protein genes RSPH9 and RSPH4A cause primary ciliary dyskinesia with central-microtubular-pair abnormalities. *Am J Hum Genet* 84 (2):197-209
25. Kott E, Legendre M, Copin B, Papon JF, Dastot-Le MF, Montantin G, Duquesnoy P, Piterboth W, Amram D, Bassinet L et al (2013) Loss-of-Function Mutations in RSPH1 Cause Primary Ciliary Dyskinesia with Central-Complex and Radial-Spoke Defects. *Am J Hum Genet* 93 (3):561-570
26. Schwabe GC, Hoffmann K, Loges NT, Birker D, Rossier C, de Santi MM, Olbrich H, Fliegauf M, Failly M, Liebers U et al (2008) Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations. *Hum Mutat* 29 (2):289-298
27. Wirschell M, Olbrich H, Werner C, Tritschler D, Bower R, Sale WS, Loges NT, Pennekamp P, Lindberg S, Stenram U et al (2013) The nexin-dynein regulatory complex subunit DRC1 is essential for motile cilia function in algae and humans. *Nat Genet* 45 (3):262-268

chr3:46745396-47606570	chr3:52111974-53028375	chr20:33,572,687-34,070,415
PRSS46	POC1A	ERGIC3
PRSS45	ALAS1	<u>SPAG4</u>
PRSS42	TLR9	CPNE1
MYL3	TWF2	NFS1
PTH1R	PPM1M	RBM12
CCDC12	WDR82	ROMO1
NBEAL2	MIRLET7G	RBM39
NRADDP	GLYCTK	PHF20
SETD2	<u>DNAH1</u>	SCAND1
<u>KIF9</u>	BAP1	CNBD2
KLHL18	PHF7	
PTPN23	SEMA3G	
SCAP	TNNC1	
ELP6	NISCH	
CSPG5	STAB1	
SMARCC1	NT5DC2	
	SMIM4	
	PBRM1	
	GNL3	
	GLT8D1	
	SPCS1	
	NEK4	
	ITIH1	
	ITIH3	
	ITIH4	
	MUSTN1	
	TMEM110	
	SFMBT1	

Table S2: List of the genes present in the three candidate regions.

Primer name	Primer sequence	Primer name	Primer sequence
DNAH1ex1F	CAAAAGCAGGGAAAGCAGGT	DNAH1ex41F	TCAGTCCTACTGGGCCTCAC
DNAH1ex1R	CATCCACTCTGCATCTGG	DNAH1ex41R	CCATAATTCAATTGCGGGAAC
DNAH1ex2F	TAAGACAGAGGCTGGCAAGG	DNAH1ex42F	CTGGCCCTTCTGCACTTG
DNAH1ex2R	CTGATGGGAGAAGGATGAA	DNAH1ex42R	AACCCCAATATAGGCCTTG
DNAH1ex3F	CCCTGGGAGGGTGAACATA	DNAH1ex43F	AAGGCCTATATTGGGGTTG
DNAH1ex3R	GCTTCCTTAACGTGGAGCG	DNAH1ex43R	GCAGGGGTGAGGAAGGAG
DNAH1ex4F	CTGGCTGAGGCCCTGAAG	DNAH1ex44F	GAGGACCTGGTGCCTGTG
DNAH1ex4R	CGACAGTCTCCCCACAGC	DNAH1ex44R	CGTAGCCCTGCCCTGTG
DNAH1ex5F	ACCTCACACGAGCCAAGC	DNAH1ex45F	GGCACAAAGGTCAATATGC
DNAH1ex5R	TTCCTGCCCTGTCTCAAC	DNAH1ex45R	CTTCTGCCTTGTCCATGCT
DNAH1ex6F	AGGAGCTGCTCTTCCAAT	DNAH1ex46F	GCAGCAAGAGGTCCACCTAA
DNAH1ex6R	ACAGGGCTGCTACTGGAATG	DNAH1ex46R	GCATGAACCTGAGGATGGTT
DNAH1ex7F	CACTCAGAGGGCCCTTCCT	DNAH1ex47F	AGCTCTGGCATGTGACC
DNAH1ex7R	GGAGTIGGCCCTCAGTCACAG	DNAH1ex47R	AGGATGCGGAGAGCAGTGT
DNAH1ex8F	CCCAGAAAGCTGATTTGTT	DNAH1ex48F	TGCTTGGTGGTTGAGAGAT
DNAH1ex8R	TGATGCCAGAACATAGGAAAG	DNAH1ex48R	CAGGGAGGCAGATTGAGAG
DNAH1ex9F	GGGTTCACAGGGTGTGGAG	DNAH1ex49F	ATACGGGGAGACCCCTACAGT
DNAH1ex9R	CATGGAGCCCATCTTGGTC	DNAH1ex49R	AACATGGGTCACTCAAACC
DNAH1ex10F	CAGGCAGGGCTCTGATACTGG	DNAH1ex50F	CAAGTCCCTGGCATGCTT
DNAH1ex10R	GCCAGGTTTGAGAAAAGCA	DNAH1ex50R	ATGCTCAGATGGGGTTGTT
DNAH1ex11F	TGCATAGACTCCCTGAGTGC	DNAH1ex51F	CCAGCAGACTGTGGTCATA
DNAH1ex11R	CTCCTTACTGCCCTGGGACCT	DNAH1ex51R	CCCATGTACACACCCAGCTC
DNAH1ex12F	CTCTGCCACAGTCAGCTC	DNAH1ex52F	TCTCTGGACCTCATTTGGAT
DNAH1ex12R	CTCAGGGAGCCATCAGCAG	DNAH1ex52R	TAACACAGGTCTCCCAAGG
DNAH1ex13F	ATTCACCCCCTGGCTCAT	DNAH1ex53F	ACCCAGTCCCTGTCTTCTC
DNAH1ex13R	TGCAAAACACACAGACACC	DNAH1ex53R	AGGCAGCTGACCTGTCCAC
DNAH1ex14F	GGGTGTCTGTTGTTGTGC	DNAH1ex54F	TCTCTGTCCCTGCCACAC
DNAH1ex14R	TATGTGAGGGTCCGTGTTG	DNAH1ex54R	CGAGCTTCCCTCAGAGTCAGG
DNAH1ex15F	GAAGAGCAGGTGAGGGAGG	DNAH1ex55F	TGGGGAGACTAAGATGCAGAG
DNAH1ex15R	CCCTCGGACTCTTAACACCA	DNAH1ex55R	AGGCATATGAGAGGAATGC
DNAH1ex16F	GGCACACGTGAAGTTCTCATT	DNAH1ex56F	CCACTGGGCTGAGTCTTCC
DNAH1ex16R	CTGGGGAGATTGCACAGACT	DNAH1ex56R	CGGGCCTTGTCCCTCACTC
DNAH1ex17F	CCGCATATTCTCCCAAT	DNAH1ex57F	CCTTGAGAAGCAGCTCTG
DNAH1ex17R	GAGAACTGTCCGTGCGT	DNAH1ex57R	GATTAGCTCCCAAGTACCC
DNAH1ex18F	GGAGGCCTGGACACAAGT	DNAH1ex58F	GGGGTACTTGGCGAGCTATA
DNAH1ex18R	GCATCTACATGAGCTGCTCC	DNAH1ex58R	AGGCAGGCTGGTCAGAG
DNAH1ex19F	GACTGGGAAGTGTCCAAGA	DNAH1ex59F	CAGGTGCCACAGTGGGTAG
DNAH1ex19R	CAGGGAGGAGGGGATTGG	DNAH1ex59R	GAAGGGTCTTTGGGAAGG
DNAH1ex20F	CCTGCCCTCATGCGTTT	DNAH1ex60F	GGACTTGGGAGGTCTCTGTG
DNAH1ex20R	TGGGGCTCTTGAGAGGAGTA	DNAH1ex60R	CAGCCTGGACCACACCCAC
DNAH1ex21F	CTGGTCAGACAGCATCAGGA	DNAH1ex61F	CTGCTTCAGGCCAGCTCTG
DNAH1ex21R	AGCGATGAGGGAGGACTG	DNAH1ex61R	CTCACCACTGCCCTGCTAT
DNAH1ex22F	GGACCCCTGACAGTCATATCA	DNAH1ex62F	GGCTCTCAGCCCTGCAAC
DNAH1ex22R	AGGCCTACAGCAATGTGACC	DNAH1ex62R	GCCCTGTCCAGGCCACT
DNAH1ex23F	TGGGATGAGCTATCTGCT	DNAH1ex63F	GTGATGCTCAGGCCACAGT
DNAH1ex23R	AGCCTTGTGGCAGACAGT	DNAH1ex63R	AGAGGCAGAGCTGATGCTG
DNAH1ex24F	CACGACCCGCTTCCCTAC	DNAH1ex64F	GGGGCCACCCAGGTATTA
DNAH1ex24R	TTTCTGTAAAAGGGCTCA	DNAH1ex64R	GAGCAGAAAGGTAGGGGTTT
DNAH1ex25F	ACCAGGGTACCCCCACTC	DNAH1ex65F	CCCCCTACCTCTGTCTTT
DNAH1ex25R	CTAGAGCAGGAGGGCACT	DNAH1ex65R	GGGTCAATATCTGTGAGGGTA
DNAH1ex26F	GAAGAGAAAAGAATGGGATTGG	DNAH1ex66F	GGCTCTTCCCTCTCAC
DNAH1ex26R	GTGGAGGGGGAGACAGACAG	DNAH1ex66R	CACCCCTCAGGCCAGTCTG
DNAH1ex27F	ACCCCTGATTTGACAGTGC	DNAH1ex67F	GAGGGTGGTTAGAGAGGCAC
DNAH1ex27R	GGACAGAGACAGCTTATTGC	DNAH1ex67R	TAGGGACTCATCCCCCTCGAT
DNAH1ex28F	GCCCTCATCCAAAAAGA	DNAH1ex68F	CACCAACCTCCCTCAACAG
DNAH1ex28R	AGTGTGGAAAGGCAAGAG	DNAH1ex68R	ACAAGCTCAAGGTGCCAGTC
DNAH1ex29F	GGAGTGTCCAGGCCATGT	DNAH1ex69F	CCTTGCCCCGATCTCT
DNAH1ex29R	CCCAAAGGAGAGCTAGGC	DNAH1ex69R	AGGGGCAAGAGCACTGAG
DNAH1ex30F	CCAGATTGGGCTGAAACAC	DNAH1ex70F	GGCTAGGCAGGGAGGAAG
DNAH1ex30R	AGTATGGCGCGTGAGTATG	DNAH1ex70R	GCCTCATTAGTCCCTGATGG
DNAH1ex31F	CTAGTCCCAGGCAAGTCAGC	DNAH1ex71F	GCCACCAAGATATGGGTCTCA
DNAH1ex31R	CCTGGAGTCCAAGCAACTGT	DNAH1ex71R	GGGGCAGGTAGGGTAGG
DNAH1ex32F	CTGGGGGTGAGCTCTGTT	DNAH1ex72F	GTGCAGGCCACCTACCT
DNAH1ex32R	GGAGCAGGGCAGAAGTTTC	DNAH1ex72R	CTCCCCTAAGTGCCAGTC
DNAH1ex33F	CTTGGAGAGGGACAGTGC	DNAH1ex73F	GGGGACTGGGCACCTTAGG
DNAH1ex33R	CCTCCAGAGATTCCACTCT	DNAH1ex73R	TGCTCATGGTACCCCAGAAC
DNAH1ex34F	CACTGCCCTGAAGGCTCA	DNAH1ex74F	AGCCAGGGGCTTCCATGT
DNAH1ex34R	CTGTCTCTGCCCCCTGACC	DNAH1ex74R	CACATGCATAGGGTCCTC
DNAH1ex35F	CACTGACTCCCCAAGGTCA	DNAH1ex75F	CAGCCTTGACCCCAAGTA
DNAH1ex35R	CCCTACAGGCCACTATC	DNAH1ex75R	GGCCCATAGTAACCCCTGTT

DNAH1ex36F	TCTCTGGGAGCCTCACTCTC	DNAH1ex76F	GACCCCAGCCCCACTACCTAT
DNAH1ex36R	CTCCAGAGGTCTGGGAGCAG	DNAH1ex76R	GGCTGTACCCTAGCCTGTCA
DNAH1ex37F	GCCTCCCTGATGTTCCAG	DNAH1ex77F	TCAGAAGGGAGTTTGTGCC
DNAH1ex37R	CCAGCCTGGAGGAGAGTG	DNAH1ex77R	TGGGCTGGGTTAGTCCTG
DNAH1ex38F	CCCACTTACAGGATGTGCAG	DNAH1ex78F	GCCCCTACGCTATCCTG
DNAH1ex38R	AGGGAGTAGGCACCCAGTTC	DNAH1ex78R	GCAAGCAGCTAAGGCACAG
DNAH1ex39F	CTGCCAGCCATGAGAACTG		
DNAH1ex39R	CCAGAGAGCAAGGAGAGCAG		
DNAH1ex40F	CTTGATACTGTTCCCTGTCTCAGC		
DNAH1ex40R	GTAGGACTGAGGCTCCTGGT		

Table S3. Primer sequences used for Sanger sequencing of *DNAH1* exons

Table S4. 50MB haplotype of subjects with MMAF centred around the *DNAH1* gene.

See the joined excel table.

Primer name	Primer sequence	Size of the amplicon	Tm
DNAH1_Ex30F	ACATCGAGGTGCTGTCTGTG	228 bp	57°C
DNAH1_Ex32R	TGATCATGGCGTAATCTGGA		
DNAH1_Ex64F	TGGAACTCATCAAGGTGCTG	293 bp	57°C
DNAH1_Ex66R	GGTCAGGTAGCGGTTGATGT		
DNAH1_Ex73F	CTCGGGCATCTACCACCAAG	241 bp	57°C
DNAH1_Ex75R	AATGTTTGGGTGACGTCCCT		
RPL0 Fw	GGCGACCTGGAAGTCCAAGT	148 bp	60°C
RPL0 Re	CCATCAGCACACAGCCTTC		
GAPDH 3F	GAG TCA ACG GAT TTG GTC GT	238 bp	60°C
GAPDH 3R	TTG ATT TTG GAG GGA TCT CG		

Table S5. Primer sequences used in RT-PCR and respective melting temperatures (Tm).

Oligonucléotides	Sequences 5' to 3'
ACTB_q_F	CCAACCGCGAGAAGATGA
ACTB_q_R	CCAGAGGCCGTACAGGGATAG
DNAH1_ex17F	ATGAGGAGAACGTTCCGCAAA
DNAH1_ex18R	GCAGTCCTTCAGCTGCTTCT
DNAH3_ex6F	TTTCTAACACGCTGCTGACG
DNAH3_ex7R	CCGGATCACTCTTGAGGAA
DNAH7_ex11F	ACAGCTGCAGGACCTCACTT
DNAH7_ex12R	CTTGGCACAAAATGACAGC
DNAH12_ex7F	CAGTTTGTGGACTTCACAGG
DNAH12_ex8R	TGCCTCCTTCTGGTAAAGAGA

Table S6. Primers and probes in RT-qPCR assays.

DISCUSSION GENERALE ET PERSPECTIVES

Les travaux présentés dans ce travail de thèse sur l'identification et la caractérisation des gènes de l'infertilité masculine nous encouragent à poursuivre nos efforts afin d'améliorer les connaissances sur la physiopathologie de l'infertilité et permettre d'envisager des alternatives thérapeutiques innovantes. L'essor des nouvelles techniques de séquençage et d'ingénierie génétique promet de nouvelles perspectives et apporte de nouveaux espoirs tant sur le plan du diagnostic moléculaire que de la thérapeutique.

L'ensemble des résultats obtenus seront discutés au travers d'un article de revue sur la génétique de la tératozoospermie publié dans une revue internationale. C'est cette revue qui est présentée dans ce chapitre en lieu et place de la discussion générale.

Les perspectives et les projets en continuité avec ces travaux de thèse seront présentés à la suite de l'article de revue.

Article de revue (article 5)

Teratozoospermia: spotlight on the main genetic actors in the human.

Charles Coutton, Jessica Escoffier, Guillaume Martinez, Christophe Arnoult, Pierre F. Ray

Human Reproduction Update, 2015 Apr 17. pii: dmv020.

Teratozoospermia: spotlight on the main genetic actors in the human

Charles Coutton^{1,2,3}, Jessica Escoffier^{1,2,4}, Guillaume Martinez^{1,2}, Christophe Arnoult^{1,2}, and Pierre F. Ray^{1,2,5,*}

¹Université Grenoble Alpes, Grenoble, F-38000, France ²Equipe 'Genetics Epigenetics and Therapies of Infertility' Institut Albert Bonniot, INSERM U823, La Tronche, F-38706, France ³CHU de Grenoble, UF de Génétique Chromosomique, Grenoble, F-38000, France ⁴Departments of Genetic Medicine and Development, University of Geneva Medical School, Geneva, Switzerland ⁵CHU de Grenoble, UF de Biochimie et Génétique Moléculaire, Grenoble, F-38000, France

*Correspondence address: UF de Biochimie et Génétique Moléculaire, CHU de Grenoble, 38043 Grenoble cedex 9, France.
Tel: +33-4-76-55-73; E-mail: pray@chu-grenoble.fr

Submitted on December 5, 2014; resubmitted on March 17, 2015; accepted on March 25, 2015

TABLE OF CONTENTS

- Introduction
- Methods
- Large headed multiflagellar spermatozoa
 - Aneuploidy and abnormal morphology: AURKC is the key
 - AURKC function in meiotic cells: between specificities and complementarities
 - AURKC mutations: the c.144delC reigns supreme
 - AURKC mutational status and patient management
- Globozoospermia
 - Globozoospermia and experimental models: a relevant strategy?
 - DPY19L2 mutational spectrum: toward a diagnostic strategy
 - DPY19L2 functions: new insights on the physiopathology of globozoospermia
 - DPY19L2, low copy repeats and copy number variations: from evolution to mutations
 - Clinical management of globozoospermic patients and DPY19L2: what's new?
- Multiple morphological abnormalities of the flagella
 - Cilia and flagellum: a common origin
 - Primary ciliary dyskinesia
 - Genetic investigations of MMAF
 - The central pair microtubules: the central key of MMAF phenotype?
 - PCD to MMAF phenotype: a phenotypic continuum?
 - Clinical implications of flagellar defects: return to base
- Translation of the genetic information to the clinic: the way forward

BACKGROUND: Male infertility affects >20 million men worldwide and represents a major health concern. Although multifactorial, male infertility has a strong genetic basis which has so far not been extensively studied. Recent studies of consanguineous families and of small cohorts of phenotypically homogeneous patients have however allowed the identification of a number of autosomal recessive causes of teratozoospermia. Homozygous mutations of aurora kinase C (AURKC) were first described to be responsible for most cases of macrozoospermia. Other genes defects have later been identified in spermatogenesis associated 16 (SPATA16) and dpy-19-like 2 (DPY19L2) in patients with globozoospermia and more recently in dynein, axonemal, heavy chain 1 (DNAH1) in a heterogeneous group of patients presenting with flagellar abnormalities previously described as dysplasia of the fibrous sheath or short/stump tail syndromes, which we propose to call multiple morphological abnormalities of the flagella (MMAF).

METHODS: A comprehensive review of the scientific literature available in PubMed/Medline was conducted for studies on human genetics, experimental models and physiopathology related to teratozoospermia in particular globozoospermia, large headed spermatozoa and flagellar abnormalities. The search included all articles with an English abstract available online before September 2014.

RESULTS: Molecular studies of numerous unrelated patients with globozoospermia and large-headed spermatozoa confirmed that mutations in DPY19L2 and AURKC are mainly responsible for their respective pathological phenotype. In globozoospermia, the deletion of the totality of the DPY19L2 gene represents ~81% of the pathological alleles but point mutations affecting the protein function have also been described. In macrozoospermia only two recurrent mutations were identified in AURKC, accounting for almost all the pathological alleles, raising the possibility of a putative positive selection of heterozygous individuals. The recent identification of DNAH1 mutations in a proportion of patients with MMAF is promising but emphasizes that this phenotype is genetically heterogeneous. Moreover, the identification of mutations in a dynein strengthens the emerging point of view that MMAF may be a phenotypic variation of the classical forms of primary ciliary dyskinesia. Based on data from human and animal models, the MMAF phenotype seems to be favored by defects directly or indirectly affecting the central pair of axonemal microtubules of the sperm flagella.

CONCLUSIONS: The studies described here provide valuable information regarding the genetic and molecular defects causing infertility, to improve our understanding of the physiopathology of teratozoospermia while giving a detailed characterization of specific features of spermatogenesis. Furthermore, these findings have a significant influence on the diagnostic strategy for teratozoospermic patients allowing the clinician to provide the patient with informed genetic counseling, to adopt the best course of treatment and to develop personalized medicine directly targeting the defective gene products.

Key words: male infertility / teratozoospermia / genetic diagnosis / sperm morphology / gene mutations

Introduction

Infertility, observed after 12 months of regular sexual intercourse, is a major health issue with at least 9% of couples requiring medical assistance to conceive a child (Boivin et al., 2007). It is therefore estimated to concern 72.4 million couples worldwide with 40.5 million currently seeking medical care (Boivin et al., 2007). In half of these cases, spermograms identify a reduced sperm quantity or quality, suggesting that a male factor is present, alone, or in conjunction with a female cause, in half of the concerned couples (Thonneau et al., 1991; Poongothai et al., 2009; Krausz, 2011). The diagnosis of male infertility is often merely descriptive, the etiology of the sperm defect remaining idiopathic in 30–50% of cases (Krausz, 2011; Tüttelmann et al., 2011b). One of the reasons for this lack of fundamental understanding is the heterogeneity of causal factors, as male infertility is a typical multifactorial disorder (Brugo-Olmedo et al., 2001). Genetic factors are however likely to be very frequent as it has been estimated that a genetic origin of male infertility could be found in nearly 1 in 40 men (Tüttelmann et al., 2011a). Chromosomal aberrations (mainly 47,XXY—Klinefelter syndrome), microdeletions of the Y chromosome and cystic fibrosis transmembrane conductance regulator (CTFR) mutations have been unambiguously demonstrated to be recurrent genetic causes of male infertility (Popli and Stewart, 2007; Vogt et al., 2008; Jungwirth et al., 2012). Despite this, routine screening for these well-established genetic causes (in principle justified only for patients presenting with certain specific phenotypes) results in a diagnosis in <5% of all phenotypes (Nieschlag et al., 2010). We can therefore safely assume that the bulk of the genetic causes of male infertility is still uncharacterized, likely due to the large number of genes involved (Matzuk and Lamb, 2002) and the lack of ambitious studies carried out on large cohorts of affected men.

Over the past years the genetic investigations of severe teratozoospermias has been one of the most productive areas within the topic of male infertility and recurrent mutations have been identified in three specific phenotypes, which are macrozoospermia, globozoospermia and multiple morphological abnormalities of the flagella (MMAF) (Fig. 1).

Teratozoospermia is defined as a percentage of morphologically normal spermatozoa below the lower reference limit. Cut-off values for normality varied greatly in recent decades from 50% in the first World Health Organization (WHO) classification (WHO, 1980) to 4% in the last version published in 2010 (Ombelet et al., 1995; World Health Organization 2010). This traditional definition of teratozoospermia is based on the identification of atypical sperm shapes in sperm smears thereby excluding ultrastructural defects invisible under light microscope (Kruger et al., 1986). Chemes and collaborators proposed to revisit this concept taking into consideration ultrastructural sperm anomalies underlying their functional incompetence. This new approach pushes back the limits of the definition of teratozoospermia beyond the mere description of spermatozoa morphology (Chemes and Rawe, 2003; Chemes and Alvarez Sedo, 2012). Such examination is however not compatible with the routine semen examination carried out in reproductive clinics. Teratozoospermia represents a heterogeneous group including a wide range of abnormal sperm phenotypes affecting, solely or simultaneously, the head, neck, midpiece and tail. Distinctions should also be made between specific forms with a homogenous monomorphic phenotype shared between several patients and forms with heterogeneous sperm anomalies (non-specific) randomly distributed in each individual and among different patients (Chemes and Rawe, 2003).

Oligoasthenoteratozoospermia (OAT) is one of the most common phenotypes of male infertility. It is defined by a combination of qualitative and quantitative sperm defects (Jungwirth et al., 2012). Due to a high phenotypic variability, genetic causes of OAT have not been clearly identified and remain so far largely unknown. Many genetic association studies reported numerous variants associated with OAT, the validity of which remains to be established. Many knockout mice have also been described to present with OAT (Matzuk and Lamb, 2008). No mutations have however been identified in their human orthologs with the exception of CAMK4 which was mutated in two patients with OAT (Khattri et al., 2012). A candidate gene approach also allowed the identification of a double amino-substitution in NANOS1, a gene that is apparently not required for spermatogenesis in the mouse

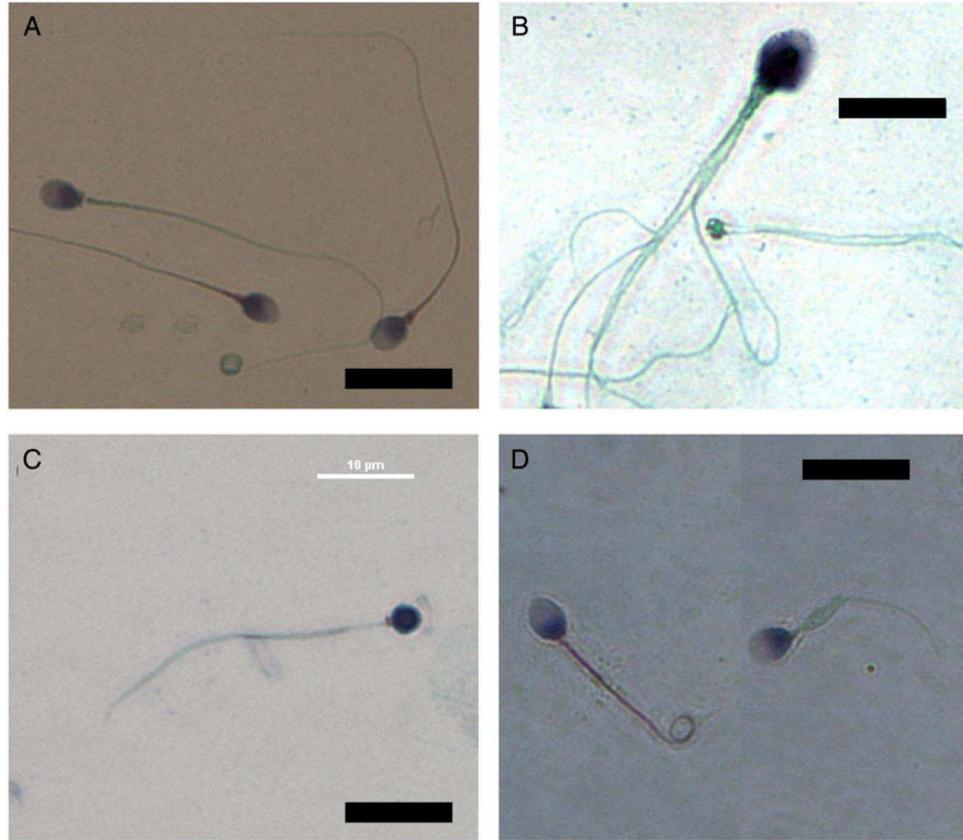


Figure 1 Light microscopy analysis of spermatozoa from different types of genetic infertility. Sperm morphology was assessed with Harris-Shorr staining; nuclei are blue and mitochondria red. **(A)** Control spermatozoa present a normal-shaped head, a typical short midpiece followed by a long principal piece. **(B)** A typical macropermic spermatozoon from a patient with a homozygous aurora kinase C (*AURKC*) c.144del mutation presents a large head and several flagella. **(C)** A typical globospermic spermatozoon from a *dpy-19-like 2 (DPY19L2)* deleted patient presents a round-shaped head and an absence of mitochondria in the midpiece. **(D)** Spermatozoa from a patient with a homozygous dynein, axonemal, heavy chain I (*DNAH1*) mutation with multiple morphological abnormalities of the flagella (MMAF) phenotype present various defects of the flagellum. Scale bars 10 µm.

(Haraguchi *et al.*, 2003), in one man with OAT (Kusz-Zamelczyk *et al.*, 2013). These mutations however, were not formally functionally validated and no further studies confirmed their implication in OAT.

Abnormalities of the sperm flagella always induce asthenozoospermia. They present highly variable morphological defects which can also affect the head and are sometimes associated with oligozoospermia. As such, some patients can be classified as a subgroup of OAT. A mosaic of different flagellar abnormalities is usually observed within the same spermogram including absent, short, bent, angulated or irregular flagella with also a high incidence of head defects. This phenotype has previously been reported as 'short tails' or 'stump tails' (Neugebauer *et al.*, 1990; Baccetti *et al.*, 1993; Stalf *et al.*, 1995; Dávila Garza and Patrizio, 2013) or dysplasia of the fibrous sheath (David *et al.*, 1993; Olmedo *et al.*, 1997, 2000; Chemes *et al.*, 1998; Rawe *et al.*, 2001; Moretti *et al.*, 2011; Ghedire *et al.*, 2014). As none of the previous terms seem to accurately define this phenotype because these patients' flagella present with a mosaic of ultrastructural flagellar defects as well as of morphological abnormalities including short tails but also, bent, curled, thick and missing flagella we previously proposed to call this heterogeneous group of subjects 'MMAF' for multiple morphological anomalies of the flagella (Ben

Khelifa *et al.*, 2014). We will use here this descriptive definition for this heterogeneous phenotype. These flagellum morphological defects are however unevenly reported or are sometimes only observed in a small fraction of patient's spermatozoa. Some gene defects have been identified which seem to induce such phenotypes including mutations in solute carrier family 26 (*SLC26A8*) or septin 12 (*SEPT12*) identified in a subset of patients with asthenozoospermia and teratozoospermia including a bent tail and/or a round head with an abnormal acrosome (Kuo *et al.*, 2012; Lin *et al.*, 2012; Dirami *et al.*, 2013). These genes were described as being involved in the organization of the septin rings forming the annulus. Other patients have been studied presenting a more severe phenotype with close to 100% abnormalities and mutations in *DNAH1*, a gene coding for an inner dynein arm protein, identified in ~28% of the patients analyzed (Ben Khelifa *et al.*, 2014).

Other phenotypes can be considered as pure teratozoospermia with 100% abnormal sperm and with a constant uniform pattern of anomalies, such as globozoospermia, large-headed multiflagellar spermatozoa or acephalic spermatozoa. To date, several genes were identified in most of these specific teratozoospermia in humans. The identification and study of these genes shed a much-needed light on the physiopathology

of teratozoospermia, as a prerequisite to improve the patient management, to provide a basis for the development of therapeutic solutions tailored to the gene defect and to provide the patients with adequate genetic counseling and expected treatment outcome. Moreover, since the investigation of gene function in human is difficult, the elucidation of the function of mutated genes and identification of possible candidate genes in animal models is crucial for the identification of factors affecting male fertility. Numerous animal models, and in particular mutant mouse models related to teratozoospermia, have been extensively described in the literature (for review Yan, 2009; De Boer et al., 2014). In this review we focus on specific teratozoospermic phenotypes for which a recurrent and reliable genetic cause has been identified in human. We provide an update of the genetic etiology of these phenotypes, summarize their underlying molecular mechanisms and detail the clinical implications arising from these findings.

Methods

A computerized literature search was conducted for all publications in PubMed/Medline until September 2014. We searched using the following MeSH or key word terms: teratozoospermia (402 records) OR globozoospermia (114 records) OR large-headed spermatozoa (35 records) OR macrozoospermia (3 records) OR macrocephalic sperm head syndrome (8 records) OR short tail sperm (187 records) OR stump tail sperm (22 records) OR dysplasia fibrous sheath (39 records) OR primary flagellar abnormalities (22 records) OR primary ciliary dyskinesia (2135 records) OR flagellum biogenesis (260 records) with a possible combination with the keywords genetics, animal models, knockout, diagnosis and ICSI. Any additional relevant articles identified from the bibliography of the initially retrieved articles and reviews were then included (24 records). Only English-language publications or articles in other languages, but with a sufficiently detailed abstract in English were included. Additionally, we have opted to focus on articles regarding specific teratozoospermia phenotypes for which a gene was formally identified in human in several non-related individuals thereby reducing the risk of erroneous diagnosis. Heterogeneous phenotypes such as OAT (249 records) or pure phenotypes but with no identified genetic cause such as 'decapitated spermatozoa' (46 records) were not included in this review. Last, we did not include studies describing the association between single nucleotide polymorphisms (SNPs) and a particular phenotype (588 records) as we believe that many of these reports may have a low level of evidence. Concerning animal models, we selected a restrictive list of knockout (KO) models among every mutant model described in the literature related to infertility and in particular teratozoospermia (1160 records) (for exhaustive review see Yan, 2009; De Boer et al., 2014). Only KO mice with a globozoospermia-like phenotype (14 records), mutant mice with inactivated aurora kinase (*Aurkc*, *Aurkb* and dynein axonemal heavy chain 1 (*Dnah1*) genes (17 records) and mice with central pair microtubule defects leading to flagellar abnormalities (10 records) are presented in this review.

Large headed multiflagellar spermatozoa

Patients with large-headed multiflagellar spermatozoa (MIM 243060), named also macrozoospermia or macrocephalic sperm head syndrome, present with a primary infertility characterized by the presence in the ejaculate of 100% abnormal spermatozoa with an oversized irregular head, abnormal midpiece and acrosome, and multiple flagella (Fig. 1B). Ultrastructural study of such spermatozoa revealed a 3-fold increase in

nuclear volume and on average 3.6 flagella for each sperm head (Escalier, 1983). This teratozoospermia is also generally associated with a severe oligoasthenozoospermia (Escalier, 1983; Benzacken et al., 2001; Devillard et al., 2002). This spectacular phenotype has been regularly described in the scientific literature since its initial description 37 years ago (Nistal et al., 1977; German et al., 1981; Escalier, 1983; In't Veld et al., 1997; Pieters et al., 1998; Kahraman et al., 1999; Benzacken et al., 2001; Devillard et al., 2002; Guthauser et al., 2006; Mateu et al., 2006; Perrin et al., 2008; Chelli et al., 2010; Brahem et al., 2012).

Aneuploidy and abnormal morphology: AURKC is the key

Several studies using Feulgen-stained preparations (German et al., 1981), spermatocyte C-banding (Pieters et al., 1998) and mainly fluorescence *in situ* hybridization (FISH) analysis showed a high rate of polyploidy and aneuploidy in spermatozoa from men with macrozoospermia. In these reports, the number of spermatozoa interpreted as haploid varied from 0 to 10.9%, as diploid from 19.8 to 60%, as triploid from 10 to 62.4% and as tetraploid from 5.1 to 36% (Yurov et al., 1996; In't Veld et al., 1997; Weissenberg et al., 1998; Viville et al., 2000; Benzacken et al., 2001; Devillard et al., 2002; Lewis-Jones et al., 2003; Vicari et al., 2003; Guthauser et al., 2006; Mateu et al., 2006; Achard et al., 2007; Perrin et al., 2008; Chelli et al., 2010; Brahem et al., 2012). Partial forms of this syndrome have been described with various percentages of large-headed spermatozoa. These 'mosaic' forms showed a percentage of euploid spermatozoa roughly correlated with their ratio of normally-sized gametes (Vicari et al., 2003; Achard et al., 2007; Brahem et al., 2012). Taken together, these observations provide evidence indicating that chromosome nondisjunction and/or cytokinesis defects occurring during the first, the second or both meiotic divisions are consistently associated with large-headed spermatozoa (Escalier et al., 1992; In't Veld et al., 1997; Weissenberg et al., 1998; Benzacken et al., 2001; Devillard et al., 2002).

A description of familial cases with consanguineous parents was suggestive of a genetic cause with an autosomal recessive inheritance. In 2007, a genome-wide low-density microsatellite analysis led to the identification of a common region of homozygosity in 7 out of 10 North African macrozoospermic patients, located in the terminal region of chromosome 19 long arm. The AURKC gene, localized in the center of this region, appeared as the ideal candidate because it was described as being expressed preferentially in male germ cells and to be involved in chromosomal segregation and cytokinesis (Fig. 2), two functions that could explain the abnormal sperm morphology and cytogenetic content of large-headed spermatozoa (Dieterich et al., 2007). Sequencing of the AURKC coding sequence allowed the identification of the same homozygous deletion (c.144delC) in all 14 patients included in the study. This mutation introduces a frameshift resulting in a premature termination of translation and yielding a truncated protein lacking its conserved kinase domain (Dieterich et al., 2007). It was later demonstrated that the mutated transcript is in fact degraded by the mechanism of nonsense mediated mRNA decay, thus indicating that these patients do not even produce the truncated protein (Ben Khelifa et al., 2011). In another study, using flow cytometry in AURKC-mutated patients, Dieterich et al. (2009) demonstrated that all spermatozoa had a homogenous tetraploid DNA content indicating that patient's germ cells undergo DNA synthesis but remain blocked without completing either of the two meiotic

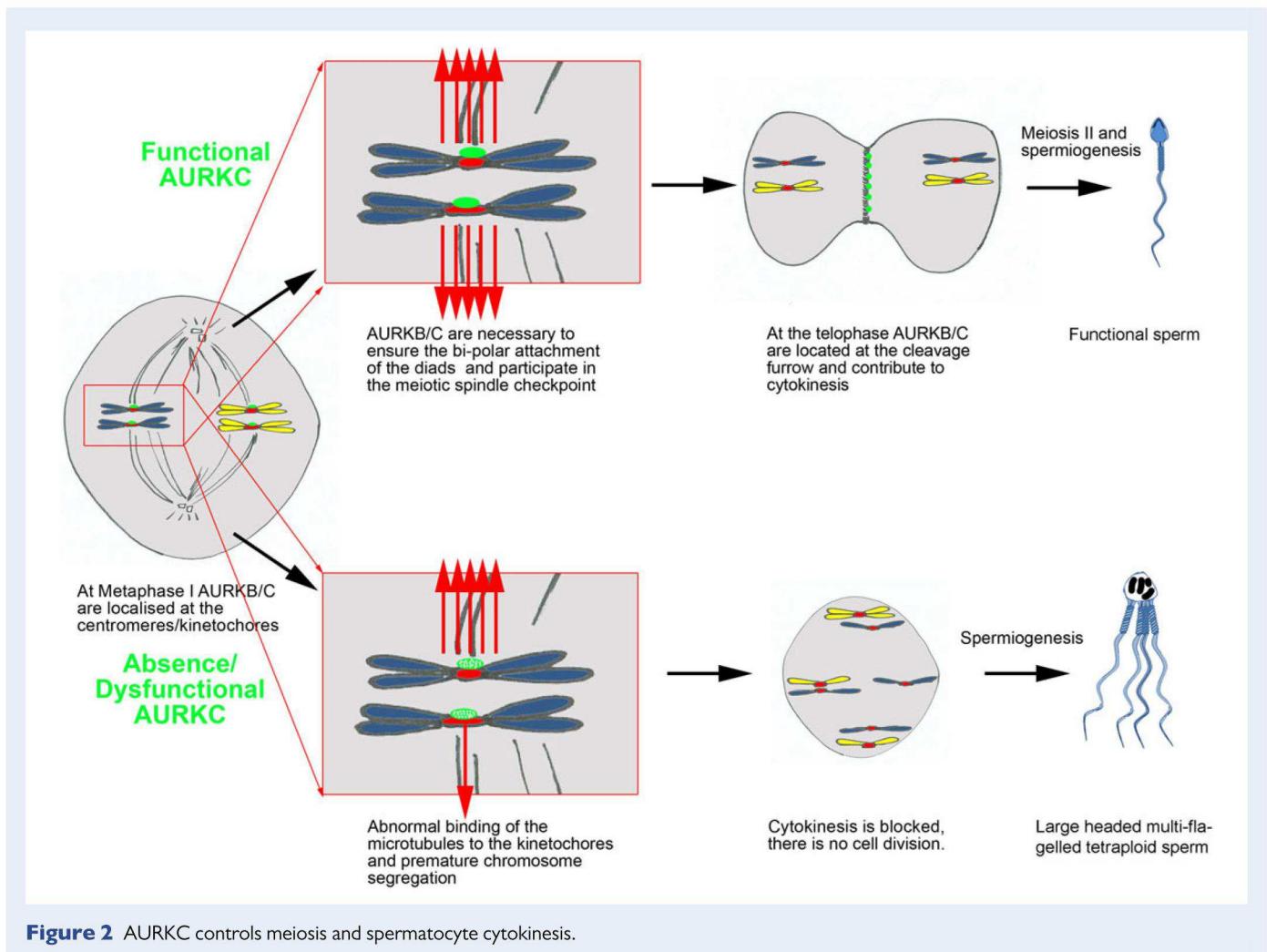


Figure 2 AURKC controls meiosis and spermatocyte cytokinesis.

divisions (Fig. 2). These results were discordant with previous FISH studies, which showed a wide heterogeneity of chromosomal abnormalities. The authors indicated that this difference is mainly due to the limits of the FISH technique and that in these abnormal gametes the superposition of FISH signals led to a great underestimate of the number of chromosomes/chromatids (Dieterich *et al.*, 2009). Surprisingly, in spite of this meiotic arrest, spermiogenesis is not blocked suggesting that meiotic checkpoint controls are abolished or inefficient in these patients. Most macrozoospermic patients however present a severely reduced sperm count concordant with the mathematical reduction in sperm number due to the absence of cell division and a partial effect of the meiotic checkpoints (Ben Khelifa *et al.*, 2011).

AURKC function in meiotic cells: between specificities and complementarities

AURKC belongs to the Aurora kinases family, which are highly evolutionary conserved serine/threonine kinases playing a key role in the control of mitosis and meiosis. In Mammals, aurora kinases also comprise two other family members, aurora kinase (AURKA) and aurora kinase B (AURKB). The three kinases share a high sequence homology in their central kinase domain and have a common ortholog in *Drosophila* and

yeast (Sunkel and Glover, 1988; Chan and Botstein, 1993; Glover *et al.*, 1995). Despite their similarities, the three Aurora kinases have distinct localizations and functions. Divergences in the N-terminal domains of these proteins confer specific protein–protein interactions abilities (Carmena and Earnshaw, 2003). AURKA and B play distinct roles during mitosis: AURKA controls centrosome maturation, bipolar assembly and chromosome separation while AURKB regulates chromosome segregation, kinetochore–microtubule interaction and cytokinesis (Goldenson and Crispino, 2014). Much less is known about AURKC which was described more recently. AURKC recently evolved from AURKB after gene duplication (Brown *et al.*, 2004), and therefore shares a high degree of homology with AURKB (73%). AURKC is expressed predominantly in testis, in particular in dividing spermatocytes, and in oocytes whereas AURKA and B are expressed ubiquitously (Bernard *et al.*, 1998; Tseng *et al.*, 1998; Assou *et al.*, 2006; Tang *et al.*, 2006; Yang *et al.*, 2010). AURKC has also been described as highly expressed in early human preimplantation embryos (Avo Santos *et al.*, 2011). There are several lines of evidence that strongly suggest that AURKC can function in place of AURKB in meiotic cells. First, AURKC function can overlap with AURKB as evidenced by the fact that AURKC could rescue AURKB-depleted somatic cells (Li *et al.*, 2004; Sasai *et al.*, 2004; Yan *et al.*, 2005). This was also supported by the fact

that AURKC is sufficient to drive mitotic progression during early embryonic divisions in preimplantation mouse embryos lacking AURKB (Fernández-Miranda et al., 2011). The study of KO mice also demonstrated a respective functional compensation of both kinases during meiosis and embryonic development (Fernández-Miranda et al., 2011; Schindler et al., 2012). Moreover, it has been reported that AURKC, like AURKB, is a component of the chromosomal passenger complex (CPC) with a similar localization and interaction with other subunits such as inner centromere protein antigens 135/155 kDa (INCENP), survivin (BIRC5) and Borealin (CDCA8) (Chen et al., 2005; Vader et al., 2008). CPC is a pivotal regulator of mitotic events including chromosome segregation and cytokinesis (Vagnarelli and Earnshaw, 2004). During chromosome alignment, CPC corrects chromosome-microtubule attachment errors generating temporarily unattached kinetochores, a condition sensed by the spindle assembly checkpoint (SAC) (Vader et al., 2008; Carmena et al., 2012; van der Waal et al., 2012; van der Horst and Lens, 2014). The SAC represents one of the main mechanisms of cell division control ensuring fidelity in chromosome segregation. It verifies whether prerequisites for chromosome segregation have been met and thereby determines whether to execute or to delay chromosome segregation (Musacchio and Hardwick, 2002; Zhou et al., 2002). In addition it has been demonstrated that Aurora kinases could contribute to recruiting different mitotic checkpoint components such as BUB1 mitotic checkpoint serine/threonine kinase (BUB1) to the kinetochore (Becker et al., 2010; van der Waal et al., 2012). Impairment of CPC functions results in chromosome segregation defects, altered SAC and cytokinesis failure. Such perturbations in mitosis can lead to the production of aneuploid/polyploid cells, which then often become malignant. By comparison, abnormalities in CPC during meiosis are expected to lead to the production of chromosomally unbalanced gametes (Sharif et al., 2010; Yang et al., 2010; van der Waal et al., 2012; van der Horst and Lens, 2014).

Many experimental studies tried to address AURKB and AURKC specific functions during meiosis. They have so far been largely hampered by the difficulty to selectively inhibit one of the two kinases (Schindler et al., 2012; Balboula and Schindler, 2014). Only one study supports a preponderant role of AURKB in mouse spermatogenesis based on the study of AURKB transgenic mice, which are shown to be subfertile due to abnormal spermatocytes, increased testicular apoptosis and spermatogenic arrest. The authors however acknowledged that the function of AURKC might also be altered in their AURKB transgenic mutants (Kimmings et al., 2007). KO strategies did not prove more useful because homozygous *Aurkc* KO mice do not perfectly mimic the typical human macrozoospermia phenotype. Unlike infertile AURKC-mutated men, KO *Aurkc* male mice are fertile but with reduced litter size. They have normal testis weight and sperm count but do produce abnormal spermatozoa. Sperm abnormalities include abnormal chromatin condensation, acrosome loss and blunted heads (Kimmings et al., 2007). It is possible that the milder phenotype observed in mouse compared with human may be due to a greater overlap of AURKB and C functions in mouse than in human spermatogenesis but this remains to be demonstrated.

The function of AURKC has been extensively studied in oocytes. In female mice, the overexpression of a dominant negative allele inhibiting the AURKC activity (AURKC-DN) in oocytes was shown to cause cytokinesis failure in meiosis I resulting in the production of large polyploid oocytes, a pattern similar to that observed in AURKC-deficient human spermatozoa (Yang et al., 2010). These results are however questionable

as AURKB might also be down-regulated in this mouse model (Schindler et al., 2012). Mice oocytes express both AURKB and C and this double expression might provide a backup ensuring the maintaining of critical functions (i.e. SAC and cytokinesis) in a transcriptionally quiescent cell (Schindler et al., 2012). *Aurkc* KO female mice however are subfertile with fewer pups per litter compared with wild type (5.5 versus 7.5 pups, on average per liter), and furthermore their oocytes show a higher incidence of chromosome misalignment and often arrest at meiosis I or later on at the I-cell stage (Schindler et al., 2012). Balboula and Schindler (2014), using a specific catalytically inactive *Aurkc* mutant (AURKC-LA), show that most AURKC-LA oocytes arrest at metaphase I suggesting that AURKC-CPC is not the sole CPC complex that regulates the SAC in female meiosis. A small percentage of oocytes proceeded through meiosis and induced cytokinesis normally but were aneuploid, indicating that AURKC-CPC is the critical CPC complex necessary to correct improper chromosome-microtubule attachments during meiosis, a central role for preventing aneuploidy (Balboula and Schindler, 2014). These data indicate that the absence of AURKC is compatible with reproduction but might have an impact during meiosis and/or embryonic development. In human the absence of AURKC is compatible with female reproduction as women with homozygous AURKC mutations were shown to be fertile (Dieterich et al., 2009). Women with homozygous AURKC mutations could however be subfertile and present an increased risk of miscarriages in relation to the production of aneuploid oocytes. Further studies on the fertility of women with homozygous AURKC mutations should be carried out to conclude on this point.

Lastly, AURKC is expressed at a very low level in other somatic cell types such as brain glial cells, lung, placenta (Yan et al., 2005; Fernández-Miranda et al., 2011) or the pineal gland where it has been implicated in circadian clock function in the rat (Price et al., 2009). Apart from infertility, homozygous AURKC-mutated patients display no other clinical features suggesting that AURKC is dispensable in somatic cells (Dieterich et al., 2007, 2009). AURKC is also aberrantly expressed in some cancer cells but its contribution to oncogenesis is not well explored and understood (Khan et al., 2011; Tsou et al., 2011; Zekri et al., 2012; Goldenson and Crispino, 2014).

AURKC mutations: the c.144delC reigns supreme

The genotype of macrozoospermic patients described in the literature is summarized in (Table I). The c.144delC deletion accounts for ~85% of the mutated alleles (Ben Khelifa et al., 2012). Other mutations have been identified: p.C229Y, a novel missense mutation in exon 6 and c.144delC (Dieterich et al., 2009), p.Y248*, a new recurrent non-sense mutation was found in 10 unrelated individuals of European and North African origin (Ben Khelifa et al., 2012) and c436-2A>G, leading to a shortened transcript devoid of exon 5 (Ben Khelifa et al., 2011). Overall, and excluding the study of Eloualid (2014) based on an unselected population of infertile men, a positive AURKC mutation diagnosis is found in between 50.8 and 100% of analyzed macrozoospermic patients (Table I).

Nearly all positive mutated AURKC patients have a typical phenotype, with close to 100% large-head spermatozoa, whereas negative patients have a lower percentage of large-headed spermatozoa (<75%) (Table II) (Dieterich et al., 2009). Positive hits are strongly correlated with the percentage of large-headed spermatozoa in the ejaculate (Table II). No

Table I Mutation status for the aurora kinase C (AURKC) gene in macrozoospermic men.

References	Number of patients studied (n)				Geographical origin (n)	AURKC mutations ^a (%; n)
	Included	Analyzed	Unrelated ^b	Newly tested		
Dieterich <i>et al.</i> (2007)	14	14	13	14	North African (14)	c.144delC +/+ 100% (13/13) p.Y248* +/+ 0% (0/13) Other 0% (0/13) No mutation 0% (0/13)
Dieterich <i>et al.</i> (2009)	62	62	61	48 ^c	North African (43), Middle East (19)	c.144delC +/+ 49.2% (30/61) p.Y248* +/+ 0% (0/61) Other ⁽¹⁾ 1.6% (1/61) No mutation 49.2% (30/61)
Ben Khelifa <i>et al.</i> (2011)	2	2	1	2	North African (2)	c.144delC +/+ 0% (0/1) p.Y248* +/+ 0% (0/1) Other ⁽²⁾ 100% (1) No mutation 0 (0%)
El Kerch <i>et al.</i> (2011)	18	18	18	18	North African (18)	c.144delC +/+ 61.1% (11/18) p.Y248* +/+ n.d Other n.d No mutation 7 (only not mutated for c.144delC)
Ben Khelifa <i>et al.</i> (2012)	87	87	83	44 ^c	North African (73), European (14)	c.144delC +/+ 67.5% (56/83) p.Y248* +/+ 9.6% (8/83) Other ⁽³⁾ 4.8% (4/83) No mutation 18.1% (15/83)
Eloualid <i>et al.</i> (2014) ^d	326 ^d	326 ^d	326 ^d	326 ^d	North African (326)	c.144delC +/+ 1.2% (4/326) ^d p.Y248* +/+ n.d Other n.d No mutation 98.8% (322/326) ^d
Ounis <i>et al.</i> (2015)	14	14	14	14	North African (14)	c.144delC +/+ 71.4% (10/14) p.Y248* +/+ 7.1% (1/14)

Continued

Table I Continued

References	Number of patients studied (n)				Geographical origin (n)	AURKC mutations ^a (%; n)
	Included	Analyzed	Unrelated ^b	Newly tested		
					Other 0 (0%) No mutation 21.5% (3/14)	

n.d: not determined; n: number of patients.

^bIndicated the number of unrelated patients among all analyzed patients.

^aMutation rate (%) is calculated only from the number of unrelated patients.

^bThis study is based on unselected population of 326 idiopathic infertile patients with various sperm parameters and phenotypes. The proportion of patients with macrozoospermia is not available.

^aThe study included 14 patients previously described in Dieterich et al. (2007).

^cThe study included 41 patients previously described in Dieterich et al. (2009) and 2 patients previously described in Ben Khelifa et al. (2011).

⁽¹⁾Patient was compound heterozygote carrying the c.144delC and p.C229Y mutations.

⁽²⁾Two brothers were compound heterozygotes carrying the c.144delC and c.436-2A>G mutations.

⁽³⁾One patient was compound heterozygote carrying the c.144delC and p.C229Y mutations, two were compound heterozygotes carrying the c.144delC and c.436-2A>G mutations and two were compound heterozygotes carrying the c.144delC and p.Y248* mutations.

mutation was generally found in patients with <70% of macrocephalic spermatozoa (Dieterich et al., 2009; Ben Khelifa et al., 2011, 2012; El Kerch et al., 2011; Eloualid et al., 2014). Furthermore, although the large majority of men with macrozoospermia present a reduced sperm count, men with extreme oligozoospermia (<0.5 M/ml) were usually not carriers of AURKC mutations.

To date Aurora kinase C remains the only gene where mutations were found in patients with a large-headed spermatozoa phenotype. The sequencing of AURKC exons 3 and 6 to search for the two recurrent mutations is appropriate as a first-line genetics test in all patients presenting with large-headed spermatozoa. The sequencing of the remaining exons should be discussed in light of the values of semen parameters. In fact, a sperm concentration >1 million, a percentage of large-head spermatozoa beyond 70% and/or a low percentage of normal spermatozoa (<1%) are in favor of a positive outcome (Table II) (Ben Khelifa et al., 2012). Other parameters could be taken into account like the presence of multiple flagella. As an example, Molinari et al. (2013) reported a case of sperm macrocephaly syndrome with 95% of large-headed spermatozoa but without tail abnormalities. AURKC gene sequencing did not reveal any mutations in the patient, suggesting that other genes may be involved in determining this atypical syndrome.

The c.144delC mutation accounts for almost all identified AURKC mutations, in particular in a Magrebian population where the prevalence of this mutation at a heterozygous state is estimated to be of 1 in 50 (Dieterich et al., 2009; Eloualid et al., 2014). Surprisingly, it makes AURKC gene alterations the most frequent defect in infertile Magrebian men before Klinefelter syndrome and Y-microdeletions (Ounis et al., 2015). It is difficult to precisely calculate the p.Y248* prevalence in North African individuals because the c.144delC is often the only mutation studied (Dieterich et al., 2009; El Kerch et al., 2011; Eloualid et al., 2014). No reliable data are currently available for European individuals due to a small number of studied patients. However, the recently identified non-sense mutation p.Y248* appears to be the main mutation in individuals of European origin (Ben Khelifa et al., 2012). Both recurrent mutations c.144delC and p.Y248* are founding mutations and ancestral haplotypes were used to compute the age of the two AURKC mutations. It was estimated that c.144delC occurred between 1350 and 1750 AD and p.Y248* between 675 and 1075 AD. The fact that p.Y248* predates

c.144delC is confirmed by the observed wider geographical distribution of p.Y248* (Ben Khelifa et al., 2012). It remains surprising that mutations with a negative effect on reproduction should reach such a high prevalence. A possible explanation lies in the assumption that the heterozygous carriers of AURKC mutations may have a selective advantage (Dieterich et al., 2007, 2009; Ben Khelifa et al., 2012; Ounis et al., 2015). No clear demonstration of this putative selective advantage has yet been provided and several hypotheses have been advanced. We previously described that AURKC was involved in the SAC. Men with AURKC heterozygous mutations could therefore have a more relaxed meiotic checkpoint permitting a faster meiosis turnover leading to an increased sperm production (Ben Khelifa et al., 2012). The produced gametes would however be expected to present a higher rate of aneuploidies likely to induce a high rate of spontaneous abortions. This hypothesis is supported by several observations of multiple miscarriages or perinatal deaths in the siblings or parents of macrozoospermic patients (Benzacken et al., 2001; Achard et al., 2007; Ounis et al., 2015).

Curiously, only four different AURKC mutations have been reported so far, and all are involved in the complete form of the phenotype. Ben Khelifa et al. (2012) propose that some mutations could affect only some functions of the AURKC protein: the microtubule–kinetochore attachment and/or cytokinesis, without altering AURKC SAC function. The severity of the defects may depend on the level of kinase inhibition; some functions of the CPC are already disturbed when the complex is only partially inhibited (e.g. correction of microtubule–kinetochore attachment), while others (e.g. its function in the SAC) may require complete inhibition (van der Horst and Lens, 2014). This would therefore lead to severe quantitative defects, and what could be considered as ‘milder’ AURKC mutations could be found in azoospermic rather than in macrozoospermic patients. This hypothesis has however not been confirmed yet.

AURKC mutational status and patient management

Although assisted reproductive technologies (ART) have revolutionized the treatment of infertile patients with severe oligoasthenoteratozoospermia, they are inefficient for patients presenting large-headed

Table II Sperm parameters of macrozoospermic men genotyped for the AURKC gene.

References	AURKC genotype (n) [§]	Sperm parameters			
		Nb spz × 10 ⁶ per ml	Large-headed spz (%)	Multiflagellar spz (%)	Normal morphology (%)
Dieterich <i>et al.</i> (2007)	Deleterious mutations (14) No mutation (0)	3.2 (0.2–7.76)	56 (34–100)	28.5 (16–50)	2 (0–5)
Dieterich <i>et al.</i> (2009)	Deleterious mutations (32) No mutation (30)	7.8 (0.4–28) 116.8 (0.01–280)	76 (34–100) 11.7 (5–75)	40.3 (20–100) 8.5 (0–28)	0.6 (0–15) 20 (0–37)
Ben Khelifa <i>et al.</i> (2011)	Deleterious mutations (2) No mutation (0)	0.85 (0.8–0.9)	100 (100)	40 (28–52)	n.a
El Kerch <i>et al.</i> (2011)	Deleterious mutations (11) No mutation (7 only not mutated for c.144delC)	n.a n.a	94.6 (71–100) 6.7 (0–38)	34.7 (4–94) 0 (0)	n.a n.a
Ben Khelifa <i>et al.</i> (2012)	Deleterious mutations (72) No mutation (15)	13.1 (0.01–98.7) 11.3 (0.01–50)	79.6 (34–100) 35.2 (5–75)	37.9 (7–100) 7.8 (0–28)	0.1 (0–1) 5.1 (0–19)
Eloualid <i>et al.</i> (2014) [§]	Deleterious mutations (4) No mutation (322) 322 (98.8%)	6 (0–17) [£] n.a	96.7 (90–100) [£] n.a	44.7 (12–90) [£] n.a	0 (0) [£] n.a
Ounis <i>et al.</i> (2015)	Deleterious mutations (11) No mutation (3)	7.51 (0.5–30.5) 13 (0.1–22.8)	99.5 (95–100) 71.7 (70–75)	+	n.a n.a

Values are expressed as the mean with the lower and the higher values between brackets if available.

n.a: not available; n: number of patients; spz: sperm cells; + indicated that the feature was present but no value was reported.

[§]Indicate the total number of patients with a deleterious mutations including all types of reported mutations in AURKC (frameshift, splicing mutations, missense) and the total number of patients with no bi-allelic mutation identified.

[£]This study is based on unselected population of 326 idiopathic infertile patients with various sperm parameters and phenotypes. The proportion of patients with macrozoospermia is not available.

[£]Average was calculated from three out of the four patients because one was azoospermic.

spermatozoa due to the high frequency of sperm chromosomal abnormalities (Perrin *et al.*, 2008). Management of patients diagnosed with macrozoospermia should begin with AURKC diagnosis. ICSI will then be formally contraindicated for all homozygous mutated patients who can have recourse to donor sperm or adoption (Ben Khelifa *et al.*, 2011). Careful selection of ‘normal-looking spermatozoa’ by motile sperm organelle morphology examination (MSOME) has previously been evaluated in several AURKC c.144delC deleted patients. FISH analyses performed on all selected spermatozoa showed that all were aneuploid, confirming that ICSI should not be attempted for AURKC-mutated patients even after a very thorough morphological selection (Chelli *et al.*, 2010). The prognosis is not as categorically unfavorable for patients not carrying AURKC mutations or with a partial/atypical phenotype. Although low, the chances of pregnancy are not negligible (Kahraman *et al.*, 1999; Achard *et al.*, 2007). For these men, sperm FISH should be carried out to assess the rate of euploid sperm and evaluate the probability of success. PGD can be proposed for those with an intermediate rate of aneuploid sperm (Kahraman *et al.*, 2004; Dieterich *et al.*, 2009). Careful scrutiny of the pregnancy and of the perinatal period is however recommended in case of success (Achard *et al.*, 2007).

Globozoospermia

First described in human in 1971, globozoospermia (MIM 613958) is a rare (incidence 0.1%) and severe form of teratozoospermia

characterized by the presence in the ejaculate of a large majority of round spermatozoa lacking the acrosome (Fig. 1C) (Sen *et al.*, 1971; Holstein *et al.*, 1973; Dam *et al.*, 2007a). Globozoospermic sperm are unable to adhere and penetrate the zona pellucida, causing primary infertility (Dam *et al.*, 2007a). The initial phenotype was divided into two subtypes: the globozoospermia type I characterized by the complete lack of acrosome and acrosomal enzymes and the globozoospermia type II characterized by a round-headed phenotype due to a residual cytoplasmic droplet surrounding the sperm head and acrosome (Anton-Lamprecht *et al.*, 1976; Singh, 1992). However, this nomenclature is confusing and was subsequently often misemployed in the literature referring to patients with a homogeneous phenotype with ~100% round-headed sperm (type I) or patients with a mosaic of normal and round-headed sperm (type II). The terms ‘total’ or ‘partial’ globozoospermia have been proposed and should be preferred to report the homogeneity of the ‘original’ type I phenotype (Lerer-Goldshtain *et al.*, 2010; Dam *et al.*, 2011) while the rarer, type II phenotype, should be referred as pseudo-globozoospermia.

Globozoospermia and experimental models: a relevant strategy?

Several familial cases of globozoospermia suggested a genetic contribution to this disorder (Kullander and Rausing, 1975; Flörke-Gerloff *et al.*, 1984; Dale *et al.*, 1994; Kilani *et al.*, 2004; Dirican *et al.*, 2008). This hypothesis was strengthened by the description in the literature

of KO mice with a globozoospermia-like phenotype lacking different genes (Table III). Almost all these genes are ubiquitously expressed and therefore do not appear as good candidates for the human globozoospermia phenotype, which only causes primary infertility. These genes encode endoplasmic reticulum proteins such as heat shock protein 90 kDa beta 1 (Hsp90b1) and glucosidase, beta (bile acid) 2 (Gba2) or vesicle trafficking-related proteins such as casein kinase 2, alpha prime polypeptide (Csnk2a2), ArfGAP with FG repeats 1 (AGFG1 or Hrb), golgi-associated PDZ and coiled-coil motif containing (Gopc), protein interacting with PRKCA 1 (Pick1), TATA element modulatory factor 1 (Tmfl), vacuolar protein sorting 54 (Vps54), small ArfGAP 2 (Smap2) and autophagy related 7 (Atg7). All regulate proacrosomal vesicle transport from the Golgi to the acrosome (Table III and Fig. 3). These proteins are implicated in a common pathway and contribute to acrosome biogenesis and sperm head organization (Fig. 3). Only zona pellucida binding protein (*Zpbp* or *Zpbp1*) and sperm acrosome associated 1 (*Spaca1*) have an expression restricted to the testis (Lin et al., 2007; Fujihara et al., 2012). *Zpbp* and *Spaca1* are integral acrosomal proteins but display different functions (Lin et al., 2007; Fujihara et al., 2012). *Zpbp* is localized in the acrosomal matrix and is involved in the binding and penetration of the sperm into the zona pellucida (Yatsenko et al., 2012) while *Spaca1* is a transmembrane protein located in the inner acrosomal membrane of spermatids and mature spermatozoa (Fig. 3) and plays an unsolved role in acrosomal morphogenesis and in sperm-egg binding and fusion (Hao et al., 2002; Fujihara et al., 2012). None of these mouse models perfectly mirrors the round-shaped acosomeless spermatozoa observed in human globozoospermia. To date, >50 genetically modified mice reported in the Mouse Genome Informatics database (<http://www.informatics.jax.org>) contain in their phenotypic description a term related to globozoospermia. This is not surprising in view of the complexity and the multitude of proteins involved in acrosome biogenesis (Alvarez Sedó et al., 2012). In spite of this large list of mouse candidate genes, only a few mutations have been identified in their human orthologs. Heterozygous missense and splicing mutations in *ZPBP* were described in patients presenting with abnormal sperm head morphology, but their involvement in the disease has not been clearly demonstrated (Yatsenko et al., 2012). Similarly, a homozygous missense mutation (G198A) in exon 13 of the *PICK1* gene was identified in a Chinese family (Liu et al., 2010). Again, the report of the identification of a relatively mild mutation without any functional validation in a single familial case remains somewhat unconvincing.

DPY19L2 mutational spectrum: toward a diagnostic strategy

Gene candidate approaches, even with strong animal model data, are often unsuccessful (Pirrello et al., 2005; Christensen et al., 2006). Homozygosity mapping using genome-wide scan analysis of a consanguineous Ashkenazi Jewish family with three globozoospermic brothers identified a homozygous mutation (c.848G>A) in *SPATA16* (spermatogenesis-associated protein 16, previously named NYD-SP12) (Dam et al., 2007b). The *SPATA16* protein localizes to the Golgi apparatus and to the proacrosomal vesicles which fuse to form the acrosome during spermiogenesis (Lu et al., 2006; Dam et al., 2007b). *SPATA16*, is highly expressed in human testis and contains a conserved tetratricopeptide repeat (TPR) domain (Xu et al., 2003) which may interact with the

GOPC and HRB proteins (Dam et al., 2007b). No mutation was subsequently found in 30 other patients with globozoospermia originating from Europe or North Africa (Dam et al., 2007b), suggesting that *SPATA16* is not the main cause of globozoospermia, which is likely genetically heterogeneous. Only recently was a new *SPATA16* mutation identified in single man from a large cohort of patients with globozoospermia (Karaca et al., 2014). In 2011, from a cohort of 20 Tunisian patients with 100% globozoospermic sperm, SNP-based genetic linkage analysis identified a common region of homozygosity in 12q14.2 including the dpy-19-like 2 (*DPY19L2*) gene strongly expressed in testis. A 200-kb homozygous deletion of the *DPY19L2* gene was found in 15 of the 20 patients analyzed (Harbuz et al., 2011). This deletion was also identified in a second study published simultaneously in 19% (4 out of 21) of the recruited globozoospermic patients (Koscinski et al., 2011). Although the deletion frequency is very disparate, both studies conclude that the *DPY19L2* deletion is the main cause of globozoospermia. Subsequently, three large studies confirmed the high prevalence of *DPY19L2* gene alterations, ranging from 60 to 83.3% of analyzed patients in cohorts of globozoospermic patients from different geographic regions and with different ethnic backgrounds (Table IV) (Coutton et al., 2012; Elinati et al., 2012; Zhu et al., 2013). A further two publications further strengthen this conclusion, reporting the presence of homozygous *DPY19L2* deletion in patients from Macedonia and Algeria (Noveski et al., 2013; Ounis et al., 2015). Homozygous deletions represent 26.7 to 73.3% of the reported *DPY19L2* mutations in the three largest studies (Table IV) (Coutton et al., 2012; Elinati et al., 2012; Zhu et al., 2013). This heterogeneous deletion rate could be explained by a possible founder effect in some studies with a geographic-based recruitment accentuating the deletion prevalence (Harbuz et al., 2011; Coutton et al., 2012; Elinati et al., 2012). In addition, inclusion parameters are different and *DPY19L2* mutations are mainly identified in type I globozoospermia with a typical morphology and a high percentage of affected spermatozoa (Table V) and the inclusion of subjects with a low level of round-headed spermatozoa might in some cases decrease diagnosis efficiency. The genotype-phenotype correlation is however not perfect, indicating that (i) some phenotypic variation exists even within *DPY19L2* deleted subjects and (ii) other gene defects likely induce a pure globozoospermia phenotype (Table V). To enable a better genotype/phenotype correlation great care should be taken to respect the international recommendation for the examination of human semen to standardize the results and avoid protocol variations between the different laboratories (World Health Organization 2010).

Homozygous and compound heterozygous point mutations further broaden the spectrum of *DPY19L2*-dependent globozoospermia (Table IV). *DPY19L2* point mutations can be either missense mutations localized mainly in the central part of the *DPY19L2* protein or nonsense/frameshift/splice-site mutations resulting in truncated proteins (Coutton et al., 2012; Elinati et al., 2012; Zhu et al., 2013). Small deletions were also reported (Table IV) indicating that exon deletions are part of the mutational spectrum of the *DPY19L2* gene (Zhu et al., 2013). A recurrent missense mutation in exon 8, p.Arg290His, was identified in several unrelated patients. This mutation changes a highly conserved arginine into a histidine and is predicted to be deleterious by multiple prediction tools (Coutton et al., 2012; Elinati et al., 2012; Zhu et al., 2013). Moreover, a recent study demonstrates that *DPY-19*, the *DPY19L2* ortholog in *Caenorhabditis elegans*, encodes a C-mannosyltransferase which is able to glycosylate several target proteins using the lipid-linked glycoside dolichol-phosphate-mannose (Dol-P-Man) as the donor substrate

Table III Mouse models with abnormalities of sperm acrosome morphogenesis and nuclear shaping.

KO mouse ID	Full name	Function	Localization and partners	Acrosome In sperm	Nuclear shape	Nuclear membrane and lamina	References
Atg7	Autophagy related 7	Golgi vesicle fusion; Protein trafficking	Cytoplasm	Present, no hook shaped	Poorly elongated <30% Globular or deformed head	Normal	Wang et al. (2014)
Csnk2a2	Casein kinase II alpha' catalytic subunit	Kinase/phosphorylation	Ubiquitous Acrosomal matrix Pick1 partner	Partially separated from the nucleus	Poorly elongated	Swelling of nuclear membranes	Xu et al. (1999) and Mannowetz et al. (2010)
Dpy19l2	Dpy-19-like 2 (<i>C. elegans</i>)	Glycosylase? Structural protein of the Inner Nuclear membrane	Inner Nuclear membrane	Absent	Round, rod shaped	Destabilization of nuclear lamina facing the acrosome. Nuclear membrane splitting off	Pierre et al. (2012)
Gba2	Glucosidase beta 2	Hydrolyzes glucosylceramide	Endoplasmic reticulum and/or plasma membrane	Partially present, disordered	Not elongated and irregular outline	?	Yildiz et al. (2006) and Walden et al. (2007)
Gopc	Golgi-associated PDZ- and coiled-coil motif-containing protein	Golgi vesicle fusion; Vesicle trafficking	Trans-Golgi Pick1 partner	Fragmented	Round or ovoid not hook shape	Normal	Yao et al. (2002) and Ito et al. (2004)
Hrb Agfg1 (official)	HIV-1 Rev binding protein ArfGAP with FG repeats 1 (official symbol)	Golgi vesicle fusion	Cytoplasmic surface of acrosome	Absent Failure of vesicle fusion	Globular	Apparently normal	Kang-Decker et al. (2001) and Kierszenbaum et al. (2004)
Hsp90b1	Heat shock protein 90b1 heat shock protein 90, beta (Grp94), member 1 (official symbol)	Endoplasmic chaperone	Endoplasmic reticulum	Partially present	Globular or deformed	?	Audouard and Christians (2011)
Pick1*	Protein interacting with C kinase 1	Golgi vesicle fusion; Protein trafficking	Trans-Golgi Gopc partner Ck2α' partner	Fragmented	Poorly elongated, not hook shape	Normal	Xiao et al. (2009)
Smap2	small ArfGAP 2 Arf GTPase-activating small GTPases protein	Golgi vesicles fusion; Protein trafficking	Trans Golgi binds to both clathrin and the clathrin assembly protein	Partially present, fragmented	Poorly elongated, not hook shape	Apparently normal	Funaki et al. (2013)
Spaca1	sperm acrosome associated 1	Structural protein of the acrosomal membrane	Inner acrosomal membrane	Partially present	Round or rod shaped	Loss of the nuclear lamina facing the acrosome	Fujihara et al. (2012)
Tmfl	TATA element modulatory factor 1	– Trafficking of Golgi-derived vesicles and/or Golgi vesicles fusion – Cytoplasm removal	Golgi-associated protein	Absent	Round or rod shaped. Nuclei are embedded in aberrant cytoplasm rest.	Apparently normal Dense nuclear lamina is present at the apical side	Lerer-Goldshtein et al. (2010)
Vps54**	Vacuolar-vesicular protein sorting 54 homolog	Vesicular sorting protein Retrograde traffic	Acrosome	Absent Failure of vesicle fusion	Poorly elongated and irregular outline	Normal	Paiardi et al. (2011)
Zppb1	Zona pellucida binding protein 1	Binding and penetration into the zona pellucida	Acrosomal matrix	Bulged	Shortened hook	Normal	Lin et al. (2007) and Yatsenko et al. (2012)

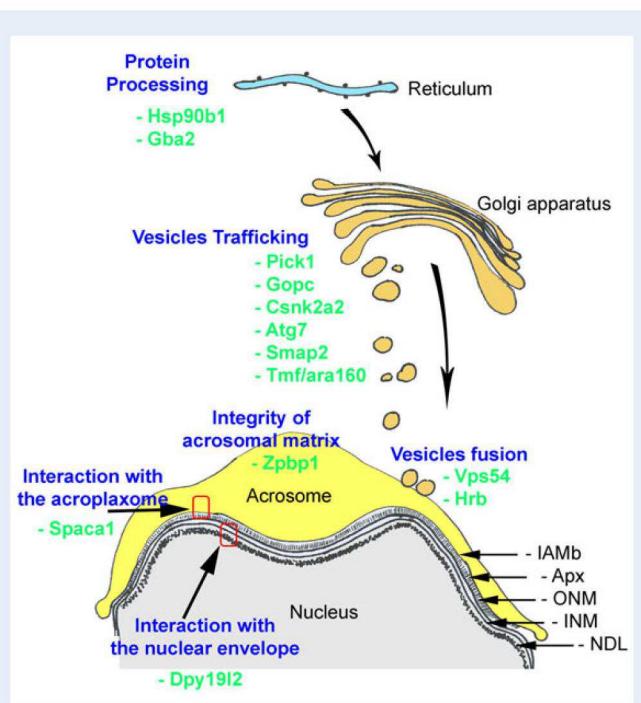


Figure 3 Known proteins involved in acrosome biogenesis whose functional absence leads to globozoospermia or globozoospermia-like phenotypes in mouse. Acrosome biogenesis includes several steps including protein processing within the reticulum, vesicle trafficking from the Golgi apparatus, vesicle fusion, interaction of the inner acrosomal membrane (IAMb) with the acroplaxome (Apx), interaction of the acroplaxome with the outer and inner nuclear membranes (ONM and INM) of the nuclear envelope and maintaining integrity of the acrosomal matrix. Numerous proteins are involved in these different steps and their absence leads to defective acrosome biogenesis and globozoospermia or globozoospermia-like phenotypes. Only the absence of sperm acrosome associated 1 (*Spaca1*) and Dpy19l2 leads to the disappearance of the nuclear dense lamina (NDL). Genes abbreviations are as follows: heat shock protein 90 kDa beta 1 (Hsp90b1); glucosidase, beta (bile acid) 2 (Gba2); protein interacting with PRKCA 1 (Pick1); golgi-associated PDZ and coiled-coil motif containing (Gopc); casein kinase 2, alpha prime polypeptide (Csnk2a2); autophagy related 7 (Atg7); small ArfGAP 2 (Smap2); TATA element modulatory factor 1; (Tmfl or ARA160); zona pellucida binding protein (Zpbp1 or Zpbp1); vacuolar protein sorting 54 (Vps54); ArfGAP with FG repeats 1 (Hrb); sperm acrosome associated 1 (*Spaca1*).

(Buettner et al., 2013). Interestingly, the amino acid corresponding to the Arg290 in *C. elegans* has been predicted to be involved in the binding of the Dol-P moiety of the substrate (Buettner et al., 2013). These data suggest that the arginine at position 290 could play a key role in DPY19L2 function. The effect of this potential glycosylation remains to be elucidated. On the whole, the fact that most reported missense mutations occur between exon 8 and 11 could indicate that this central domain of DPY19L2 may support some critical functions.

Overall, the deletion of the totality of the *DPY19L2* gene represents about 81% of the pathological alleles (Table IV). Different approaches have been described to effectively detect these *DPY19L2* deletions. First, a long-range PCR combined with exon-specific amplification is the most widely used strategy to detect both homozygous and

heterozygous deletions (Table IV) (Harbuz et al., 2011; Koscinski et al., 2011; Elinati et al., 2012; Noveski et al., 2013; Zhu et al., 2013). This cost effective strategy allows delineating with a good accuracy the breakpoints of the genomic deletion although this information is not, strictly speaking, useful for the diagnosis and clinically relevant. A strong shortcoming of this technique is that it does not detect the deletions with breakpoints falling outside the region covered by the deletion-specific primers therefore presenting a high risk of false-negative results (Coutton et al., 2012; Zhu et al., 2013). An alternative strategy is the use of a multiplex ligation-dependent probe amplification (MLPA) approach to determine the number of *DPY19L2* alleles for each patient, and independently of the localization of the breakpoint (Coutton et al., 2012). Drawbacks of MLPA are the relatively higher cost compared with PCR approaches and the required experience for the design of the MLPA specific probes to avoid cross-hybridization with *DPY19L2* pseudogenes (Zhu et al., 2013). Considering these points, we recommend the research of homozygous and heterozygous *DPY19L2* deletions as the first-line genetic analysis in globozoospermic by long-range PCR or MLPA (Coutton et al., 2012; Elinati et al., 2012; Zhu et al., 2013). Alternatively a quantitative PCR protocol could be used (Chianese et al., 2015). This approach does not rely on breakpoint localization and new robust methods are becoming available which may provide a cheap and fast diagnosis allowing the detection of all recombined alleles. In almost all cases of globozoospermic patients with a heterozygous deletion, an additional point mutation in the non-deleted allele was secondly identified (Coutton et al., 2012; Elinati et al., 2012; Zhu et al., 2013). This strongly encourages researching such a point mutation in globozoospermic patients heterozygous for the genomic deletion (Coutton et al., 2012). In the absence of a deletion (heterozygous or homozygous) the full 22 exon *DPY19L2* sequencing might not be systematically realized as a routine analysis. Similarly, the very low rate of mutations found in others genes, such as *SPATA16* or *PICK1*, questions the clinical relevance of molecular investigations of such genes (Coutton et al., 2012). A 'hot-spots' restrictive strategy based on the sequencing of the *DPY19L2* exons 8–11, in which most of the point mutations were identified, could also be envisaged. The genetic diagnosis of globozoospermia does not yet provide any clear prognostic or therapeutic indications. Nonetheless, a molecular diagnosis remains useful to provide adequate genetic counseling and to better understand the physiopathology of globozoospermia, which might help to identify novel therapeutic solutions.

DPY19L2 functions: new insights on the physiopathology of globozoospermia

DPY19L2 belongs to a new family of transmembrane proteins of the nuclear envelope including, in Mammals, fours homologous proteins: DPY19L1 to DPY19L4. Initially, inactivation of DPY-19, the only DPY19L ortholog in *C. elegans*, was shown to induce a shorter (dumpy) than wild-type phenotype (Brenner, 1974). DPY-19 has 10 predicted transmembrane domains and was involved in Q neuroblast polarization during early development (Honigberg and Kenyon, 2000). The four human homologues have 9–11 predicted transmembrane domains, suggesting a similar protein structure to DPY-19 (Carson et al., 2006). Moreover, human DPY19L proteins share a high level of sequence identity with DPY-19, consistent with a conservation of the ancestral protein function (Buettner et al., 2013). Functions of human DPY19L proteins are still poorly understood and in particular for

Table IV Mutation status for the DPY19L2 gene in globozoospermic men.

References	Number of patients				Geographical origin (n)	DPY19L2 mutations ^a (%; n)	DPY19L2 deletion detection method	Point mutations identified
	Included	Analyzed	Unrelated?	Newly tested				
Harbuz <i>et al.</i> (2011)	20	20	19	20	North African (19) and Eastern Europe (1)	Hm deletion 73.7% (14/19) Hm point mutations n.d Compound Htz n.d No mutation 26.3% (5/19)	LR-PCR	None
Koscinski <i>et al.</i> (2011)	28	28	21	28	North African (7), Middle East (6), European (12), unknown (3)	Hm deletion 19% (4/21) Hm point mutations n.d Compound Htz n.d No mutation 81% (17/21)	LR-PCR	None
Coutton <i>et al.</i> (2012)	34	31	30	14 ^{\$}	North African (30), European (4)	Hm deletion 73.3% (22/30) Hm point mutations 3.3% (1/30) Compound Htz 6.7% (2/30) No mutation 16.7% (5/30)	MLPA	p.R290H; p.Q342*; p.M358K
Elinati <i>et al.</i> (2012)	54	54	54	33 [£]	North African (n.a), Middle East (n.a), European (n.a)	Hm deletion 46.3% (25/54) Hm point mutations 7.4% (4/54) Compound Htz 13% (7/54) No mutation 33.3% (18/54)	LR-PCR	p.R290H; p.R298C; p.Q345*; p.S395LfsX7; p.T493R; p.K680*; c.1218+1G>A; Ex5_6del; Ex5_7del
Zhu <i>et al.</i> (2013)	16	15	15	16	Asian (16)	Hm deletion 26.7% (4/15) Hm point mutations 33.3% (5/15) Compound Htz 0% (0/15) No mutation* 33.3% (5/15)	LR-PCR	c.1532delA; c.[1679delT 1681_1682delAC]; p.R290H; p.L330P
Noveski <i>et al.</i> (2013)	2	2	2	2	Eastern Europe (2)	Hm deletion 100% (2/2) Hm point mutations 0 (0%) Compound Htz 0 (0%) No mutation 0 (0%)	LR-PCR	None
Ounis <i>et al.</i> (2015)	7	5	3	7	North African (7)	Hm deletion 100% (3/3)	MLPA	None

Continued

Table IV Continued

References	Number of patients			Geographical origin (n)	DPY19L2 mutations ^a (%; n)	DPY19L2 deletion detection method	Point mutations identified
	Included	Analyzed	Unrelated ^b	Newly tested			
					Hm point mutations 0 (0%) Compound Htz 0 (0%) No mutation 0 (0%)		

n.d: not determined; n.a: not available; n: number of patients; spz: sperm cells; MLPA: multiplex ligation-dependent probes amplification; LR-PCR: long-range PCR.
^bIndicated the number of unrelated patients among all analyzed patients.
^aThe study included 20 patients previously described in Harbuz et al. (2011).
^cThe study included 21 patients previously described in Koscinski et al. (2011).
^dMutation rate (%) is calculated only from the number of unrelated patients.
^eOne additional patient (1/15, 6.7%) had a heterozygous deletion in one allele but with no mutation identified in the non-deleted allele.

Table V Sperm parameters of globozoospermic men genotyped for the DPY19L2 gene.

References	DPY19L2 genotype (n) [§]	Sperm parameters			
		Nb spz × 10 ⁶ /ml	Total sperm count (10 ⁶)	Normozoospermic [¶] % (n)	Round-headed spz (%)
Harbuz et al. (2011)	Deleterious mutations (15)	62.8 (0.02–154)	194.6 (0.09–665)	86.7 (13)	99.5 (94–100)
	No mutation (5)	9.7 (0.04–25.2)	31.2 (0.12–68)	40 (2)	89.8 (64–100)
Koscinski et al. (2011)	Deleterious mutations (8)	83.4 (13.2–223)	n.a	100 (8)	100 (100)
	No mutation (20)	38.7 (0.35–109)	n.a	30 (6)	97.6 (84–100)
Coutton et al. (2012)	Deleterious mutations (26)	57.3 (0.6–108)	210 (n.a)	n.a	88 (29–100)
	No mutation (5)	20 (14–25)	64 (n.a)	n.a	63 (12–100)
Elinati et al. (2012)	Deleterious mutations (36)	n.a	n.a	n.a	n.a
	No mutation (18)	n.a	n.a	n.a	n.a
Zhu et al. (2013)	Deleterious mutations (9)	46.7 (0.8–90)	183.6 (n.a)	n.a	100 (100)
	No mutation (5)	20 (14–25)	60 (n.a)	n.a	100 (100)
Noveski et al. (2013)	Deleterious mutations (2)	n.a	n.a	n.a	100 (100)
	No mutation (0)				
Ounis et al. (2015)	Deleterious mutations (5)	64.6 (16–171)	n.a	100 (5)	99 (95–100)
	No mutation (0)				

Values are expressed as the mean with the lower and the higher values between brackets if available. n.a: not available; n: number of patients; spz: sperm cells;

[§]Indicate the number of patients with a deleterious mutations including all types of reported mutations in DPY19L2 (homozygous deletions, point mutations, compound heterozygotes) and the number of patients with no bi-allelic mutation identified.

[¶]Indicated the total number and percentage of patients normozoospermic with a sperm concentration upper to the 5th centiles of World Health Organization 2010 values (15 × 10⁶/ml (CI 12–16)).

DPY19L4 for which nothing is known. It has been demonstrated that down-regulation of Dpy19l1 and Dpy19l3 during neurogenesis of mouse embryos leads to strong neuron migration anomalies (Watanabe et al., 2011). DPY19L3 has also been associated with bipolar disorder (Smith et al., 2009). These results strengthen the possibility of a conserved function of the different DPY19L proteins. Interestingly, apart from DPY19L2, which presents a predominant expression in the testis, DPY19L proteins show a relatively ubiquitous pattern of expression (Carson et al., 2006).

After its identification as the main gene involved in globozoospermia, study of Dpy19l2 KO mice showed that homozygous animals faithfully

reproduced the human phenotype with a complete male infertility, 100% of acrosomeless round-headed spermatozoa and manchette abnormalities (Table III). Study of this model led to the conclusion that DPY19L2 is a transmembrane protein located in the inner nuclear membrane (Fig. 3) and that it is necessary to anchor the acrosome to the nucleus. In the absence of Dpy19l2 the forming acrosome slowly separates from the nucleus before being removed from the sperm with the cytoplasm (Pierre et al., 2012). In addition to its structural function during acrosome biogenesis, the C-mannosyltransferase function of the ancestral protein DPY-19 has raised the hypothesis that DPY19L2 may have a function in glycosylation of sperm proteins (Buettnner et al.,

2013). Several potential partner proteins have been proposed to interact with DPY19L2 during spermiogenesis, in particular the SUN proteins belonging to the LINC (linker of nucleoskeleton and cytoskeleton) complexes in association with KASH proteins (Pierre et al., 2012). SUN proteins appeared as the best candidate due to their abundance during spermiogenesis and their role in nuclear shaping during sperm head formation (Göb et al., 2010; Frohnert et al., 2011). This hypothesis has however not been confirmed and the molecular partners of DPY19L2 remain to be characterized (Yassine et al., 2015b). Comparative testicular transcriptome studies of wild type and globozoospermic Dpy19l2 KO mice were also realized to attempt to identify Dpy19l2 molecular partners but no conclusive result was obtained (Karaouzène et al., 2013).

DPY19L2, low copy repeats and copy number variations: from evolution to mutations

Copy number variations (CNV) refer to gains by duplication or losses by deletion of genetic material greater than 1 kb. They can be pathogenic and cause Mendelian traits or be associated with complex diseases but can also represent benign polymorphic variants. CNVs and in particular gene duplications are involved as a predominant mechanism driving gene and genome evolution (reviewed in Zhang et al., 2009). Such a mechanism was involved in the DPY19L gene family evolution throughout the vertebrate lineage. The DPY19L family genes, encompassing four genes (DPY19L1, 2, 3 and 4) and six pseudogenes, derived from a common ancestor homologous to the *C. elegans* dpy-19 gene owing to multiple gene duplications and pseudogenizations throughout evolution (Carson et al., 2006). It is estimated that the duplication that generated DPY19L2 arose between 173 and 360 million years ago prior to mammalian divergence.

Recurrent CNVs arise mostly by homologous recombination between repetitive DNA sequences. This process is called non-allelic homologous recombination (NAHR) and is well-known in human diseases to be responsible for many recurrent genomic syndromes (Stankiewicz and Lupski, 2010). The DPY19L2 locus is conducive to recurrent deletions and duplications by NAHR due to the presence of two homologous 28-kb low copy repeats (LCRs) located on each side of the gene (Harbuz et al., 2011; Koscinski et al., 2011). Such a recombination during the meiosis will result in the production of recombinant gametes having either a duplication or a deletion of the DPY19L2 locus. Sequencing analysis of the NAHR breakpoints at the DPY19L2 locus demonstrated that the recombination events occurred preferentially in the vicinity of a 13 nucleotide recognition motif for the PRDM9 protein localized in the center of the 28-kb LCRs (Coutton et al., 2013). PRDM9, a zinc finger protein that binds to DNA sequences, is a major trans-regulator of meiotic recombination hotspots facilitating double strand breaks, a prerequisite for NAHR (Baudat et al., 2010; Myers et al., 2010; Parvanov et al., 2010). An additional putative recombination hotspot within the 28-kb LCRs might constitute a second, less frequent recombination site (Elinati et al., 2012).

The NAHR mechanism favors deletions over duplications because the inter-chromatid and inter-chromosome NAHRs create a deleted and a duplicated recombined allele, while intra-chromatid events only generate deletions (Liu et al., 2012). More gametes carrying a deletion are thus expected to be produced *de novo* and this was verified for several NAHR hotspots (Turner et al., 2008) and in particular for the DPY19L2 locus (Coutton et al., 2013). This is paradoxical regarding the overrepresentation of the duplicated DPY19L2 allele in the general

population. In fact, the frequency of the DPY19L2 duplication and heterozygous deletion in the general population is estimated to be 1/85 and 1/290, respectively. This can be explained by the loss of deleted alleles occurring during selection against sterile, homozygous deleted men, whereas duplications carriers are not subjected to selection. Also, heterozygous deleted men might also suffer a small fitness penalty, expediting the loss of the deleted allele (Coutton et al., 2013).

Clinical management of globozoospermic patients and DPY19L2: what's new?

The relationship between sperm chromosomal abnormalities and sperm morphology has been widely explored. A link has been clearly established in particular for certain types of morphologically abnormal spermatozoa, such as large-headed multiflagellar spermatozoa, but remains controversial for many others (Sun et al., 2006). Globozoospermia is no exception and contradictory results were obtained. Some studies report an increased frequency in sperm aneuploidy compared with normal sperm cells while others do not observe any difference (reviewed in Dam et al., 2007a; Perrin et al., 2013). Finally, even observed, the slightly increased aneuploidy rate in round-headed sperm is comparable to that commonly found in other types of infertility (Perrin et al., 2013) and therefore does not counter indicate the treatment of patients by ICSI (Kuentz et al., 2013). Despite this, many studies using ICSI for patients with globozoospermia reported low fertilization, pregnancy and live birth rates, overall estimated at 38, 20 and 14%, respectively (reviewed in Dávila Garza and Patrizio, 2013). A first well-established explanation for the low efficiency of ICSI when using globozoospermic sperm is the reduction or absence of the sperm factor, a testis-specific phospholipase (phospholipase C zeta (PLCζ)) involved in the oocyte activation (Yoon et al., 2008; Heytens et al., 2009; Kashir et al., 2010; Amdani et al., 2013; Escoffier et al., 2015). PLCζ localizes to the inner acrosomal membrane and the nuclear theca in the post-acrosomal region of human sperm which is in accordance with its vanishing in Dpy19l2-KO globozoospermic sperm (Escoffier et al., 2015). The absence of oocyte activation can be overcome by artificial oocyte activation (AOA) using Ca²⁺ ionophore (Rybouchkin et al., 1997; Tejera et al., 2008; Kyono et al., 2009; Taylor et al., 2010). In case of globozoospermia, AOA combined with ICSI improved the efficiency of fertilization, pregnancy and live birth rates compared with conventional ICSI and this, regardless of the genotype (DPY19L2 related or not) of the patients (Kuentz et al., 2013). However, the live birth rate per transfer remains lower in globozoospermic patients than in other infertile patients in the same age group (Palermo et al., 2009; Kuentz et al., 2013). This may be due to sperm DNA damage related to defective chromatin condensation and the DNA fragmentation described in globozoospermic sperm cells (Dam et al., 2007a; Brahem et al., 2011; Perrin et al., 2013; Yassine et al., 2015a). Abnormal chromatin condensation in globozoospermia could be linked to an altered replacement of histones by protamines (Blanchard et al., 1990; Carrell et al., 1999). Yassine et al. (2015a) using the *dpy19l2* KO mouse models confirmed that the nuclear invasion by protamines during the last stage of compaction is defective. Moreover, the absence of protamine increases dramatically the susceptibility to DNA breaks. Altogether, epigenetics defects and DNA fragmentation in globozoospermic sperm nuclei impair the developmental potential of embryos generated by ICSI using *Dpy19l2*-dependent spermatozoa (Yassine et al., 2015a). In summary,

globozoospermic patients with *DPY19L2* gene alterations combine a low oocyte activation rate and a poor development of embryos generated by ICSI, thus explaining the consistently disappointingly low pregnancy rate obtained with these patients.

Mutational status in globozoospermic patients treated with ICSI was generally unknown or not investigated. Few studies reported successful fertilization or pregnancies in globozoospermic patients with well documented *DPY19L2* or *SPATA16* mutations (Dam et al., 2007b; Harbuz et al., 2011; Kuentz et al., 2013; Karaca et al., 2014). Due to this small number of cases, no evidence of a difference in pregnancy success rate regarding the mutational status of globozoospermic patients has currently been demonstrated. Further studies taking into account genetic mutations may influence the therapeutic strategy and facilitate improving the management of patients diagnosed with globozoospermia. To date, only the type of globozoospermia is known to have an impact on the therapeutic approach as intracytoplasmic morphologically selected sperm injection technique (IMSI) has proven effective only for patients with a partial globozoospermia (Sermondade et al., 2011; Kuentz et al., 2013).

Multiple morphological abnormalities of the flagella

Morphological abnormalities of the sperm flagella leading to asthenozoospermia have been reported regularly since 1984 (Escalier and David, 1984). Chemes and colleagues carried out much of the early work on this phenotype and extensively studied the ultrastructure of the sperm flagella of affected men (Chemes et al., 1987). They observed recurrent abnormalities of the fibrous sheath (FS) which defines the principal piece surrounding the axoneme and the outer dense fibers and consists of two longitudinal columns connected by circumferential ribs (Eddy et al., 2003) (Fig. 4). Escalier had previously identified similar flagellar defects and had described that abnormalities of the peri-axonemal structures (including FS, outer dense fibers, mitochondrial sheath) were always associated with axonemal defects (Escalier and David, 1984). She argued that these peri-axonemal defects might be secondary to the axonemal defects (Escalier and Serres, 1985). Further ultrastructural studies among patients with sperm flagella abnormalities showed a wide range of different peri-axonemal and axonemal defects (Rawe et al., 2001; Chemes and Rawe, 2003).

Cilia and flagellum: a common origin

Although the sperm flagellum and motile cilia (found in the epithelial cells of the airways, the Fallopian tubes, the choroid plexus and the brain ventricles) are specialized for a particular function, they share common structural elements as the central cytoskeletal structure called the axoneme, which is highly conserved throughout evolution. The axoneme contains 9 outer microtubules doublets and 2 central singlets (9+2 structure), several axonemal dyneins, radial spokes, nexin links and many other components that drive and regulate ciliary or flagellar motility. All these microtubule-associated proteins are attached along the axoneme in a regular 96 nm long repeated formation (Inaba, 2003; Satir and Christensen, 2008). The axonemal dyneins are multi-component proteins organized along each microtubule doublet as inner and outer rows (Fig. 4B and C) and comprise motor multiprotein complexes which produce the cilia/flagella beating force (Wickstead and Gull, 2007). Dynein arms are composed of heavy chains (HC),

intermediate chains (IC), light-intermediate chains (LIC) and light chains (LC) and numerous regulatory proteins (Bisgrove and Yost, 2006). The structure of the inner dynein arms is more complex than the structure of the outer dynein arms, which, in mammals, contains only two different HCs repeated all along the axoneme. In mammals, the inner arms are organized in seven molecular complexes, viewed in electron microscopy as globular heads arranged in 3-2-2 groups, thus corresponding to 3 types of inner dynein arms (IDA): IDA1 to IDA3 (Vernon et al., 2005). The mammalian organization was however much less investigated than in *Chlamydomonas* (Bui et al., 2008), and the subunit composition of each inner arm remains to be determined. Although sperm flagella and motile cilia have a similar microtubule structure based on the presence of a 9+2 axoneme, they present several differences. Motile cilia are constituted almost exclusively by the axoneme, whereas mammalian sperm flagella are divided into three principal parts characterized by the presence of additional accessory structures surrounding the axoneme: the outer dense fiber, the mitochondrial sheath, the fibrous sheath (Fig. 4). Recent proteomic analysis identified over 700 proteins that are exclusively localized in the human sperm flagella, of which many are axonemal proteins (Baker et al., 2013). These data highlight the complexity of the flagella structure and biogenesis, suggesting that many genes could be linked to flagellar abnormalities and/or asthenozoospermia.

Primary ciliary dyskinesia

Primary ciliary dyskinesia (PCD) is a multisystemic disorder caused by motility defects of motile cilia and flagella (Afzelius and Eliasson, 1983; Munro et al., 1994). PCD is mainly characterized by recurrent respiratory tract infections with varying symptoms ranging from chronic rhinosinusitis to bronchiectasis and male infertility due to sperm immotility (Ibañez-Tallón et al., 2003). Female subfertility is less common and is caused by dysmotile Fallopian tubes cilia (Lyons et al., 2006). PCD is also associated in 50% of cases with *situs inversus* due to dysfunction of motile embryonic node cilia perturbing organ laterality (Zariwala et al., 1993). More rarely, hydrocephalus arises as a consequence of ependymal cilia dysmotility leading to a blockage of the cerebrospinal fluid flow (Ibañez-Tallón et al., 2002; Kosaki et al., 2004). PCD has an estimated incidence of 1 per 15 000 births (Knowles et al., 2013a) but it may be underestimated due to diagnostic failure (Boon et al., 2013).

A high prevalence of PCD was observed in different genetic isolates and consanguineous populations, a fact consistent with a mainly autosomal recessive mode of inheritance (Jeganathan et al., 2004; O'Callaghan et al., 2010). PCD is genetically heterogeneous (Horani et al., 2014) and research on PCD allowed the characterization of numerous proteins necessary for adequate axonemal molecular structure and assembly. Cilia and flagella are remarkably well conserved throughout evolution. Different experimental models were used to decipher the molecular composition and function of these organelles (Ostrowski et al., 2011). The first model was the green alga *Chlamydomonas* in which *DNAI1*, the first gene to be associated with PCD, was identified (Penarun et al., 1999). The function of the genes involved in PCD was also efficiently assessed in other multicellular organisms, such as zebrafish, *Xenopus*, *Caenorhabditis elegans* or *Drosophila*, and protists, such as *Paramecium*, *Tetrahymena*, *Trypanosoma* or *Leishmania*, each bringing specific advantages to the study of cilium biology (review in Vincensini et al., 2011). Thereafter, various studies using a candidate gene approach

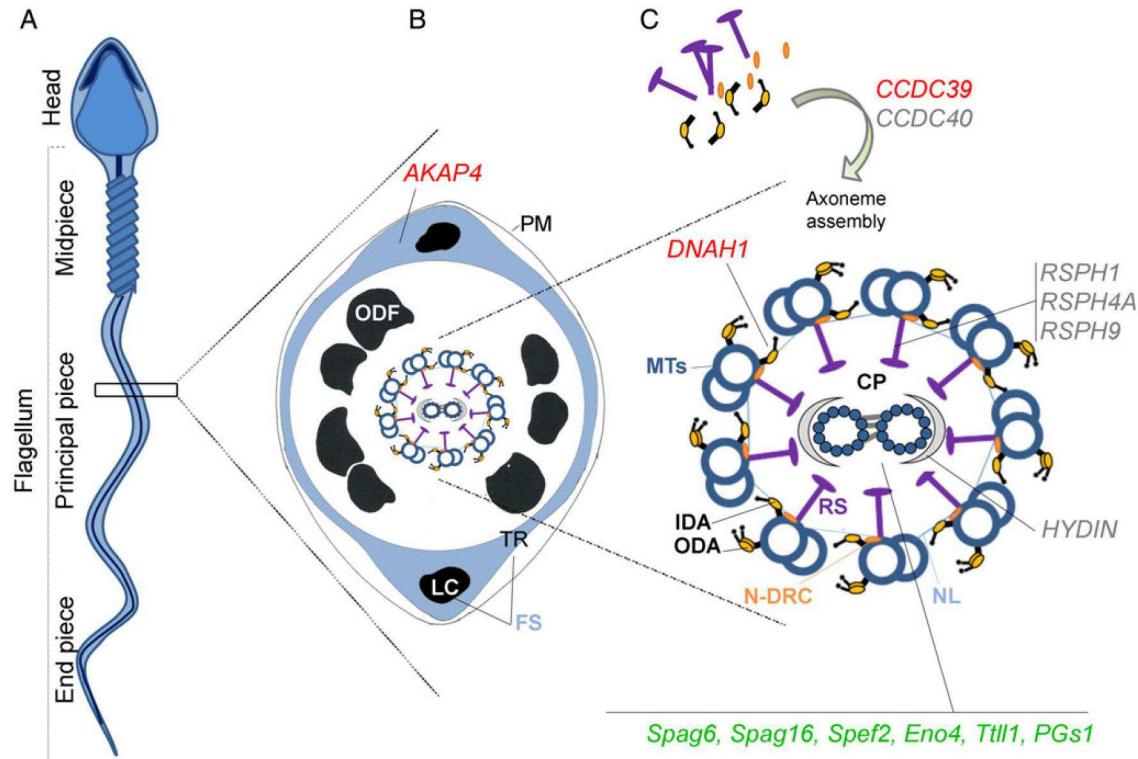


Figure 4 Structure of sperm flagellum and localization of proteins whose functional absence leads to a defective flagellar structure and MMAF-like phenotype. **(A)** Mammalian sperm flagellum is structurally divided into three areas: midpiece, principal piece and end piece. **(B)** Schematic cross section through a representative segment of the principal piece showing the plasma membrane (PM) surrounding 7 outer dense fibers (ODF, there are 9 ODF in the mid-piece). The fibrous sheath (FS) is composed of two longitudinal columns (LC) which are connected by transverse ribs (TR). Within the ODF are the components of the axoneme. **(C)** The axoneme is enlarged and the offset shows: the 9 outer microtubule doublets of the axoneme (MTs) with associated inner dynein arms (IDA), outer dynein arms (ODA), radial spokes (RS), nexin-dynein regulator complex (N-DRC), nexin links (NL) and the central pair of microtubule doublets (CP). Projections are represented on the CP (gray bow) but not detailed. Genes formally identified for the MMAF phenotype in human are reported in red. Candidate genes for the MMAF phenotype identified in animal models and in humans are reported in green and gray, respectively. Genes abbreviations are as follows: A kinase (PRKA) anchor protein 4 (AKAP4); coiled-coil domain containing 39 (CCDC39); coiled-coil domain containing 40 (CCDC40); radial spoke head 1 homolog (RSPH1); radial spoke head 4 homolog A (RSPH4A); radial spoke head 9 homolog RSPH9; HYDIN, axonemal central pair apparatus protein (HYDIN); sperm associated antigen 6 (Spag6); Sperm associated antigen 16 (Spag16); Sperm flagellar 2 (Speg2); Enolase 4 (Eno4); Tubulin tyrosine ligase-like 1 (Ttll1); Tubulin polyglutamylase complex subunit 1 (Psg1).

mainly based on data from *Chlamydomonas* have resulted in the identification of many mutated genes in PCD patients. Different KO mice or inbred dogs deficient in various axonemal components have also been successfully described as mammalian models for PCD (review in Inaba, 2011). Mutations in 28 genes leading to various ultrastructural defects have so far been described to cause PCD (Table VI) and account for the genetic etiology of ~70% of affected individuals (Knowles et al., 2014). A male infertility phenotype has often been described as part of the clinical symptoms but this particular aspect of the physiopathology of the PCD is not systematically explored and is often only scarcely described in scientific reports (Table VI). The sperm of infertile male patients with PCD are usually immotile and present various ultrastructural defects of sperm flagella such as missing dynein arms, microtubular translocations, and lack of radial spokes. Nevertheless, in most cases the sperm flagellum appears morphologically normal under direct light microscopy (Table VI) thereby excluding most of these genes as a genetic cause of teratozoospermia or MMAF.

Genetic investigations of MMAF

A genetic origin of MMAF was strongly suspected based on a family clustering reported in at least 20% of patients (Chemes and Alvarez Sedo, 2012). In 2005, Bacetti et al. (Bacetti et al., 2005a) first reported a partial deletion in the A kinase (PRKA) anchor protein 3 (AKAP3) and A kinase (PRKA) anchor protein 4 (AKAP4) genes in a patient presenting with short-tail spermatozoa. Ultrastructural sperm evaluation showed MMAF and an altered axonemal structure lacking dynein arms and microtubular doublets including the central pair. A kinase (PRKA) anchor protein 3 (AKAP3) and A kinase (PRKA) anchor protein 3 (AKAP4) encode two A-kinase anchoring proteins interacting with the regulatory subunits of cAMP-dependent protein kinase A. AKAP3 and AKAP4 are the most abundant structural proteins of the fibrous sheath. AKAP3 is involved in organizing the basic structure of the fibrous sheath while AKAP4 has a major role in completing fibrous sheath assembly (Brown et al., 2003). Akap4 KO mice are infertile with half-length flagella with a

Table VI Primary ciliary dyskinesia gene mutations and their consequences for axonemal ultrastructure and sperm phenotype.

Gene ID	Full name	MIM number	Axonemal localization	Axonemal defects	Infertility phenotype	References
ARMC4	Armadillo repeat containing 4	615408	Assembly of ODA	Reduced outer dynein arms	No data available	Hjeij et al. (2013) and Onoufriadis et al. (2014a)
C21orf59	Chromosome 21 open reading frame 59	615494	Dynein arm assembly	Absence of both outer and inner dynein arm components	No data available	Austin-Tse et al. (2013)
CCDC103	Coiled-coil domain containing 103	614677	Assembly of ODA	Reduced outer dynein arms	No data available	Panizzi et al. (2012)
CCDC114	Coiled-coil domain containing 114	615038	DCC2, an ODA microtubule-docking complex	Ciliary outer dynein arm defects	Male are fertile	Knowles et al. (2013b) and Onoufriadis et al. (2013)
CCDC151	Coiled-coil domain containing 151	615956	ODA docking complex	Outer dynein arm defect	No data available	Hjeij et al. (2014) and Jerber et al. (2014)
CCDC164 (DRC1)	Coiled-coil domain containing 164	615288	Nexin-dynein regulatory complex	Absence of nexin links, disruption of the nexin-dynein regulatory complex	No data available	Wirschell et al. (2013)
CCDC39	Coiled-coil domain containing 39	613798	Assembly of IDA, N-DRC and Radial spokes	<ul style="list-style-type: none"> - absence of inner dynein arms and nexin links - axonemal disorganization with mislocalized peripheral doublet - displacement or absence of the central pair 	Oligoasthenozoospermia midpiece is narrowed flagellum is shortened	Merveille et al. (2011), Blanchon et al. (2012) and Antony et al. (2013)
CCDC40	Coiled-coil domain containing 40	613799	Govern the assembly of N-DRC and inner dynein arm complexes, but not outer dynein arm complexes	<ul style="list-style-type: none"> - reduction of inner dynein arms - axonemal disorganization with mislocalized peripheral doublet, displacement or absence of the central pair - abnormal radial spokes and nexin links 	No data available	Becker-Heck et al. 2011, Blanchon et al. (2012) and Antony et al. (2013)
CCDC65	Coiled-coil domain containing 65	611088	Assembly of the N-DRC	Reduction in inner dynein arms and nexin links	No data available	Austin-Tse et al. (2013) and Horani et al. (2013b)
DNAAF1 (LRRK50)	Dynein, axonemal, assembly factor 1	613190	Assembly of dynein arm complexes in the cytoplasm	Absence of both outer and inner dynein arms	Male infertility reported but no details were provided	Duquesnoy et al. (2009) and Loges et al. (2009)
DNAAF2 (KTU)	Dynein, axonemal, assembly factor 2	612517	Assembly of dynein arm complexes in the cytoplasm	Absence of both outer and inner dynein arms	Asthenozoospermia	Omran et al. (2008)
DNAAF3	Dynein, axonemal, assembly factor 3	614566	Assembly of dynein arm complexes in the cytoplasm	Absence of both outer and inner dynein arms	Male infertility reported but no details were provided	Mitchison et al. (2012)
DNAH11	Dynein, axonemal, heavy chain 11	603339	Outer dynein arm heavy chain	Ultrastructure respiratory cilia is normal and outer dynein arms are intact	Male are fertile	Bartoloni et al. (2002), Schwabe et al. (2008), Knowles et al. (2012) and Lucas et al. (2012)
DNAH5	Dynein, axonemal, heavy chain 5	603335	Outer dynein arm heavy chain	Outer dynein arm defect	Asthenozoospermia	Olbrich et al. (2002), Fliegauf et al. (2005), Hornef et al. (2006) and Failly et al. (2009)
DNAI1	Dynein, axonemal, intermediate chain 1	604366	Outer dynein arm intermediate chain	The outer dynein arms are shortened or missing.	Asthenozoospermia	Pennarun et al. (1999), Guichard et al. (2001); Zariwala et al. (2001)
DNAI2	Dynein, axonemal, intermediate chain 2	605483	Outer dynein arm intermediate chain	Outer dynein arm defect	Male infertility reported but no details were provided	Loges et al. (2008)

DNAL1	Dynein, axonemal, light chain 1	610062	Outer dynein arm light chain	The outer dynein arms are shortened or missing	No data available	Mazor et al. (2011)
DYXIC1 (DNAAF4)	Dyslexia susceptibility 1 candidate 1	608706	Assembly of dynein arm complexes in the cytoplasm	Absence of both outer and inner dynein arms	Asthenozoospermia	Tarkar et al. (2013)
HEATR2	HEAT repeat containing 2	614864	Assembly or stability of axonemal dynein arms	Absence of outer dynein arms and partial lack of inner dynein arms	Male infertility reported but no details were provided	Horani et al. (2012)
HYDIN	HYDIN, axonemal central pair apparatus protein	610812	C2b projection	Lack the C2b projection of the central pair	Asthenozoospermia	Olbrich et al. (2012)
LRRC6	Leucine rich repeat containing 6	614930	Assembly of dynein arm complexes in the cytoplasm	Absence of both outer and inner dynein arms	Asthenozoospermia	Kott et al. (2012) and Horani et al. (2013a)
RGPR	Retinitis pigmentosa GTPase regulator gene	312610	Involved in the transitional zone of motile cilia in airway epithelia/ localized to centrioles	Variable. From normal structure to lack of both dynein arms and abnormal microtubular Disorganization.	No data available	Moore et al. (2006) and Bukowy-Bierylo et al. (2013)
RSPH1	Radial spoke head 1 homolog	609314	Radial spoke component	Central pair microtubule complex and radial spoke defects	Male infertility reported but no details were provided	Kott et al. (2013), Knowles et al. (2014) and Onoufriadis et al. (2014b)
RSPH4A	Radial spoke head 4A homolog	612647	Radial spoke component	Central pair microtubule complex and radial spoke defects	Male infertility reported but no details were provided	Castleman et al. (2009) and Daniels et al. (2013)
RSPH9	Radial spoke head 9 homolog	612648	Radial spoke component	Central pair microtubule complex and radial spoke defects	Male infertility reported but no details were provided	Castleman et al. (2009)
SPAG1	Sperm associated antigen 1	603395	Assembly of dynein arm complexes in the cytoplasm	Defects of both outer and inner dynein arms	No data available	Knowles et al. (2013c)
TXND3	Thioredoxin domain containing 3	607421	Non precisely determined	Shortened or absent outer dynein arms	No data available	Duriez et al. (2007)
ZMYND10	Zinc finger, MYND-type containing 10	607070	Assembly of dynein arm complexes in the cytoplasm	Lack of outer and inner dynein arms	Male infertility reported but no details were provided	Moore et al. (2013) and Zariwala et al. (2013)

thinner principal piece and a tip sometimes curled or splayed into fine filaments. The size and integrity of mutant mitochondrial sheaths are also reduced. However, other cytoskeletal structures, namely the outer dense fibers and the axoneme, remain intact (Miki et al., 2002). Mouse models thus present strong evidence that Akap3 and 4 are involved in MMAF phenotype. Evidence of their implication in the human phenotype is however weaker. A deletion of AKAP3 and AKAP4 was described in a MMAF patient but it was only detected using conventional PCR and the genomic breakpoints were not identified (Baccetti et al., 2005a). Moreover, no quantitative analyses of DNA or mRNA were used to confirm the observed deletions and exclude a false negative PCR result. The authors then demonstrated an absence of staining in patient's spermatozoa using immunofluorescence with AKAP4 antibodies (Baccetti et al., 2005a). The absence of the AKAP4 protein in mature sperm may however be secondary to another defect disorganizing the whole sperm flagella structure (Baccetti et al., 2004). Additionally, Turner et al. (2001) did not find any mutation in these two genes in 9 MMAF patients.

More recently homozygous mutations in the *DNAH1* gene were identified in several patients with MMAF (Ben Khelifa et al., 2014). *DNAH1* encodes an axonemal inner arm dynein HC and is expressed in various tissues including the testis (Maiti et al., 2000). Homozygosity mapping was carried out on a cohort of 20 North African individuals. Four different homozygous mutations in *DNAH1*, one run-on, one missense and two splice site mutations were identified in seven patients including three brothers. Electron microscopy examination of spermatozoa revealed a general axonemal disorganization including mislocalization of the microtubule doublets, absence of the central pair in about half of the analyzed cross sections, loss of the inner arm dynein as well as severe disorganization of the fibrous sheath, the outer dense fibers and the mitochondrial sheath (Ben Khelifa et al., 2014). All these defects are hallmarks of the MMAF phenotype. Molecular and functional studies on samples from one of the mutated patients carrying the recurrent splicing mutation (c.11788-1G>A) demonstrated that the transcript and the protein were absent, confirming that the observed phenotype is due to the complete absence of *DNAH1*. This initial study thus indicates that ~70% of DFS patients are expected to bear a genetic alteration in other genes, thus confirming that, like PCD, MMAF is genetically heterogeneous and that many other genes are likely involved in this syndrome. Lastly, the patient with the less severe missense variant presented a milder phenotype with 5% motility and the presence of 6% morphologically normal spermatozoa in contrast to the other *DNAH1*-mutated patients which had 100% abnormal sperm and 0% motility (Ben Khelifa et al., 2014). It could therefore be expected that individuals with intermediate asthenozoospermia and a low level of morphological anomalies could also harbor homozygous or compound heterozygous *DNAH1* mutations of moderate severity.

The phenotype observed in *DNAH1*-mutated patients extends beyond the absence of flagellar motility, as usually observed in other axonemal component defects, and substantiates a role for this protein in the maintenance of the structural integrity of the flagella. In *Dnahc1* KO mice (the mouse ortholog of *DNAH1*, previously called *Mdhc7*), 'rapid-freeze deep-etch electron microscopy' studies indicated that one head of the IDA3 was missing, leading to a 3-2-1 globular head arrangement and suggesting that *DNAH1* is one of the components of the IDA3 (Vernon et al., 2005). In *Tetrahymena thermophila*, IDA3 corresponds to the dynein d/a which is directly connected to the radial spoke RS3 through an arc-like

structure (King, 2013). In mammals, the three radial spokes (RS1, RS2, and RS3) are multiprotein complexes allowing a connection between the external microtubule doublets and the two central microtubules thus stabilizing the axoneme (Pigino et al., 2011). Therefore, the severe axonemal disorganization observed in *DNAH1*-mutated patients may occur as a result of the abrogation of RS3 anchoring, leading to the absence of the central pair and the mislocalization of the peripheral doublets (Ben Khelifa et al., 2014). KO male mice lacking *Dnah1* are infertile but only present with asthenozoospermia, without the sperm morphological abnormalities observed in human. Aside from the fact that there might be some divergences in flagellar biogenesis between mouse and human, an incomplete disruption of the targeted dynein was reported in the mouse model, which may explain this phenotypic difference (Neesen et al., 2001; Ben Khelifa et al., 2014).

The central pair microtubules: the central key of MMAF phenotype?

The central pair microtubules (CP) are composed of two microtubule singlets, named C1 and C2, which are structurally and biochemically distinct. Structural studies in *Chlamydomonas* showed that the C1 tubule has two long projections (1a and 1b) and two short projections (1c and 1d). The C2 tubule has three short projections termed 2a, 2b and 2c. The CP contains at least 23 distinct proteins and some of these proteins are uniquely associated with either C1 or C2, indicating that the two microtubules may be functionally specialized. Evidence supports the fact that the CP projections interact with the RS heads and modulate dynein activity (reviewed in Wargo and Smith, 2003).

It is worth noting that the lack of the CPs leading to an abnormal '9+0' configuration of the axoneme is the main defect observed in most cases associated with MMAF (Chemes and Rawe, 2003). Such a defect was also observed in patients with *DNAH1* mutations (Ben Khelifa et al., 2014). These observations raise the hypothesis that CP disorganization might be the cornerstone leading to the MMAF phenotype. Several experimental models support the idea that some structural defects impacting the CP give rise to MMAF (Fig. 4C and Table VII). Axonemal abnormalities have been observed in spermatids from testicular biopsies underlining that the onset of the defect occurs during the late stages of spermiogenesis due to a defective assembly of cytoskeletal components of the sperm tail (Chemes and Rawe, 2010). Interestingly, in some cases, the disorganization of the microtubules has been described to increase and the sperm phenotype to worsen during the sperm transit through the epididymal duct (Sapiro et al., 2002; Sironen et al., 2011). This suggests that CP defects precede other more severe structural abnormalities by weakening all the axonemal organization making the flagella vulnerable to mechanical stress during spermiogenesis and sperm transit. This demonstrates that CP plays a major role in maintaining the global flagellum organization throughout spermiogenesis. This could explain why most other axonemal defects, not affecting the CP structure, do not lead to the MMAF phenotype.

As a consequence, all genes encoding a protein interacting directly (CP components) or indirectly with the CP apparatus, like RS, some dynein arms or different partner proteins, are henceforth good candidates for the MMAF phenotype (Fig. 4C). As an example, depletion of mouse Meig1, a protein that interacts directly with Spag16 (Table VII) leads to shorter flagella, absence of the normal '9+2' axoneme arrangement and peri-axonemal disorganization (Zhang et al., 2004; Salzberg et al.,

Table VII Animal models with central pair of microtubule doublet defects presenting with a DFS/MMAF-like phenotype.

Inactivated genes ID	Full ID	Localization	Function	Sperm phenotype	Known protein interaction	Associated phenotype	References
<i>Spag6</i>	Sperm associated antigen 6	C1 central pair microtubule	Unknown	Truncated sperm tails. Lack of the central pair microtubules. Alterations in the fibrous sheath and/or outer dense fibers.	SPAG16 [£] , SPAG17	Hydrocephalus (50%)	Sapiro <i>et al.</i> (2000, 2002)
<i>Spag16</i>	Sperm associated antigen 16	Bridges connecting the C1 and C2 microtubules	Unknown	Sperm abnormally shaped. Lack of the central pair microtubules. Marked disorganization of the outer doublet microtubules and outer dense fibers.	SPAG16 [£] , SPAG17, MEIG1 [£]	None	Zhang <i>et al.</i> (2004)
<i>Spf2 (Kpl2)</i>	Sperm flagellar 2	C1b (?)	Unknown	Lowered sperm counts, short sperm tails. Absence of central pair and missing of peripheral doublets. Lack of organized mitochondria, outer dense fibers, and fibrous sheath structures.	IFT20	None in pig model; Hydrocephalus and sinusitis in mouse	Sironen <i>et al.</i> (2006, 2010, 2011)
<i>Eno4</i>	Enolase 4	C1b (?)	Glycolytic enzymes; ATP production	Sperm shortened, thickened and coiled. Disorganized aggregates of the fibrous sheath components. Disrupted axonemal structures. Displaced outer dense fibers in the principal piece. Defective annulus was also reported	PGAM	None	Nakamura <i>et al.</i> (2013)
<i>Ttl1 (PGs3)</i>	Tubulin tyrosine ligase-like 1	Central pair α-tubulin (?)	Axonemal α-tubulin polyglutamylation	Shortened flagella. Various axonemal abnormalities ranged from the absence of central microtubule to complete disorganization of the axonemal and peri-axonemal structures.	Unknown	Rhinosinusitis, otitis media	Ikegami <i>et al.</i> (2010) and Vogel <i>et al.</i> (2010)
<i>PGs1 (Gtrgeo22)</i>	Tubulin polyglutamylase complex subunit 1	Central pair α-tubulin (?)	Axonemal α-tubulin polyglutamylation	Shortened flagella. Various axonemal abnormalities ranged from the absence of central microtubule to complete disorganization of the axonemal and peri-axonemal structures.	Unknown	Absence of intermale aggression, reduced body fat	Campbell <i>et al.</i> (2002)

? indicated a probable localization but not formally identified. £ indicated a protein partner involved in dysplasia of the fibrous sheath/multiple morphological abnormalities of the flagella (DFS/MMAF)-like phenotype. SPAG17: Sperm associated antigen 17; MEIG1: Meiosis expressed gene 1; IFT20: intraflagellar transport 20; PGAM: phosphoglycerate mutase.

2010; Teves et al., 2013). In humans, several studies reported mutations in the coiled-coil domain containing 39 (*CCDC39*), coiled-coil domain containing 40 (*CCDC40*), radial spoke head 1 homolog (*RSPH1*), radial spoke head 4 homolog A (*RSPH4A*), radial spoke head 9 homolog *RSPH9* and HYDIN, axonemal central pair apparatus protein (*HYDIN*) genes leading to CP defects. However, few phenotypic descriptions or structural studies have been carried out on sperm samples from mutated patients (Table VI). Only Merveille et al. (2011) reported that the sperm flagellum was shortened in patients with a mutation in the *CCDC39* gene, a feature consistent with the MMAF phenotype and our hypothesis.

PCD to MMAF phenotype: a phenotypic continuum?

DNAH1 is the first inner arm dynein gene involved in a human pathology and the first axonemal gene responsible for a male infertility phenotype without any of the other symptoms usually observed in PCD. These novel mutations in the axonemal dynein *DNAH1* strengthen the emerging point of view that MMAF may be a phenotypic variation of the classical forms of PCD with a continuum of clinical manifestations ranging from infertile PCD patients to MMAF patients with no or low noise PCD manifestations. Indeed, many genes are likely to have a specific function in the sperm flagellum with potentially several paralogs that could carry out a similar function in other ciliated tissues, thus explaining the absence of a more severe PCD phenotype. For instance, absence of respiratory symptoms in *DNAH1*-mutated patients may be explained by other dyneins phylogenetically close to *DNAH1* which may compensate the absence of *DNAH1* in other motile cilia (Ben Khelifa et al., 2014). This hypothesis has also been mentioned in a reverse situation where PCD patients with mutations in the *CCDC114*, a gene encoding an outer dynein arm docking complex, had no fertility problems. The authors proposed that *CCDC114* function could be partially replaced by *CCDC63*, a homologous protein of *CCDC114* with an expression restricted to the testis (Onoufriadi et al., 2013). Unfortunately, no nasal brushings or biopsies from *DNAH1*-mutated individuals could be obtained to search for infra-clinical manifestations by exploring the patient's ciliary function in respiratory epithelium. In addition, alternative splicing is widespread in mammals and splice variants often exhibit tissue-specific expression patterns (Yu et al., 2014). Therefore some mutations may affect alternative variants specifically expressed in testis while having no effect in other tissues.

Overall, these observations support the fact that important differences exist between the axonemal assembly/function of respiratory cilia and sperm flagella that explain why PCDs are not always associated with primary flagellar defects and vice versa. This was supported by mutations in *DNAH5*, an outer dynein arm, which are associated with the absence of outer dynein arms in respiratory cilia while a normal distribution of ODA heavy chain was reported in the sperm flagellum of mutated patients (Fliegauf et al., 2005). Mutations in *DNAH11*, another outer dynein arm, although associated with a normal axonemal ultrastructure, led to a PCD phenotype also with normal male fertility (Bartoloni et al., 2002; Schwabe et al., 2008). In many mouse models with a MMAF-like phenotype, the structural disorder appears to be specific to sperm tail development, because no effects on the architecture of cilia in the respiratory or female reproductive tract have been observed (Pilder et al., 1997; Fossella et al., 2000; Campbell et al., 2002; Miki et al.,

2002; Sapiro et al., 2002; Sironen et al., 2006, 2011; Zhang et al., 2006; Lessard et al., 2007; Lee et al., 2008; Salzberg et al., 2010; Nakamura et al., 2013). Despite this normal axonemal organization, respiratory symptoms or lower tracheal ciliary beat frequency and/or hydrocephalus were described (Sapiro et al., 2002; Lee et al., 2008; Sironen et al., 2011). These findings highlight the duality of function of various proteins, involved in sperm flagellum biogenesis and in the motility of tracheal and ependymal cilia. Further studies are needed to elucidate the mechanisms underlying differences between cilia and flagella and to decipher the specific proteome of these two closely related organelles.

Clinical implications of flagellar defects: return to base

In dividing cells, the centrioles form the core of the centrosome which is the primary microtubule-organizing center in animal cells and it is involved in numerous functions such as the organization of the mitotic/meiotic spindle (reviewed in Bettencourt-Dias and Glover, 2007). During flagellum biogenesis, the sperm centrioles migrate to the spermatid periphery and form the basal bodies from which the flagellar axoneme originates (Chemes, 2012). Centrioles are embedded in a matrix of proteins known as the pericentriolar material that supports motor proteins such as dyneins or kinesins (Schatten and Sun, 2009). Moreover, during fertilization in human, the sperm centrosome organizes the sperm aster which is essential to unite the sperm and the oocyte pronuclei (Sathananthan et al., 1991). The sperm centrosome also controls the syngamy and the first mitotic divisions after fertilization (Schatten and Sun, 2009).

Primary flagellar abnormalities have been associated with an elevated frequency of gonosomal disomies and diploidies (Lewis-Jones et al., 2003; Baccetti et al., 2005b; Rives et al., 2005; Collodel and Moretti, 2006; Ghedir et al., 2014). In the light of the above, chromosomal abnormalities observed in MMAF patients may be linked to the common components shared between the sperm centrosome and the flagella. Therefore, it could be speculated that defects in some of these centrosome-associated proteins may disturb both flagellum formation and spindle assembly during sperm meiosis, resulting in nondisjunction errors and spermatozoa aneuploidy (Rives et al., 2005). No chromosomal abnormalities were however detected in one patient with short-tailed spermatozoa (Viville et al., 2000) pointing out that all flagellar defects do not impact the centrosomal function.

Sperm aneuploidy could impede ICSI outcomes in patients with flagellar defects. Nevertheless several studies described the successful application of ICSI to treat male infertility due to MMAF or PCD (for review Chemes and Rawe, 2003; Chemes and Alvarez-Sedo, 2012; Dávila Garza and Patrizio, 2013). The most recent review estimated the mean overall fertilization, pregnancy and live birth rates in MMAF patients at 63, 57 and 43%, respectively (Dávila Garza and Patrizio, 2013). In PCD patients, this review found fertilization and pregnancy rates ranging from 55 to 65% and from 35 to 45%, respectively, depending with the sperm source (ejaculated or testicular). The overall live birth rate was estimated at 39%.

Interestingly although the results compare favorably with the overall results obtained after ICSI, the rate of ICSI success seems to be influenced by the type of ultrastructural flagellar defects carried by the patients (Mitchell et al., 2006; Fauque et al., 2009). Mitchell et al. (2006) reported a lower implantation and clinical pregnancy rate in patients without axonemal central structures (i.e. '9+0' axoneme).

Likewise, Fauque et al. (2009) reported a slower kinetics of early embryo cleavage and a lower implantation rate when a central pair of singlet microtubules was missing. These data question the possibility of a link between some specific axonemal structural defects and abnormal embryonic development in human. Since axonemal structures and sperm asters come from the sperm centrosome, it is possible that some of the described cases of fertilization failure and abnormal embryonic development reported in MMAF patients might be caused by defects in centrosomal or pericentrosomal proteins (Sathananthan, 1994; Van Blerkom, 1996; Chemes, 2012). The use of heterologous ICSI systems (human-bovine, human-rabbit) with sperm from infertile men with MMAF supported this hypothesis, showing a lower rate of sperm aster formation (Rawe et al., 2002; Terada et al., 2004). Interestingly, IVF experiments in *Dnah1* KO mice demonstrated a retarded rate of early embryo development (Neesen et al., 2001). Unfortunately, in patients with *DNAH1* mutations no data relative to ICSI attempts were available (Ben Khelifa et al., 2014). Lastly, although there has been no report of an abnormal birth from patients with MMAF or PCD, the risk of genetic defects other than infertility linked to the sperm structural defects should be discussed during genetic counseling (Sha et al., 2014). We believe that genetics data combined with careful analysis using electron microscopy will, in time, permit us to optimize the course of treatment for MMAF patients.

Translation of the genetic information to the clinic: the way forward

ART can circumvent infertility, and IVF with ICSI is possible for most patients regardless of the etiology of the defects. Unfortunately the availability of these technologies does not guarantee success and almost half the couples who seek reproductive assistance fail to achieve a pregnancy. The improvement of ART procedures is therefore particularly relevant and this strongly relies on a better understanding of spermatogenesis and of the molecular physiopathology of infertility. This mission involves the identification of genes responsible for male infertility. To date, a very short list of genes was identified which is in sharp contrast with the fact that several hundreds of genes (probably well in excess of a thousand) are estimated to be involved in spermatogenesis and male reproduction. Although genetic causes of human male fertility remain largely unknown, genetics investigations of some specific teratozoospermic phenotypes have yielded fruitful results. The key to success in the identification of infertility genes was largely based on the study of clinically and genetically homogenous phenotypes in small cohorts with consanguineous subjects and large families. This led in particular to the identification of AURKC and DPY19L2 and to the development of diagnoses for both genes, which are now recommended for patients presenting macrozoospermia and globozoospermia, respectively. Additional efforts are needed to identify genes in others syndromes and more generally in all phenotypes of male infertility in which many genes remain to be identified in human. The recent identification of *DNAH1* in patients with MMAF is promising although a positive diagnosis is achieved for only ~28% of the patients. This indicates that MMAF is genetically heterogeneous and that many other genes are likely involved in this syndrome.

Homozygosity mapping has so far permitted the identification of the three main genes discussed in this review. This strategy can only be

successful when studying large families or consanguineous patients presenting with a very homogeneous phenotypes. Due to the great genetic heterogeneity of infertility, this technique has now reached its limits and microarray analyses are now replaced by the fast evolving technology of next generation sequencing. Sequencing of all the coding sequences, or exome sequencing, can now be realized in a few weeks at a cost of 400–1000 €. Whole exome sequencing will be available on the next day (in a few years time) for probably less than that cost. This technical revolution is what was needed to unlock the secrets of male infertility. Exomic/genomic sequencing of large cohorts of patients is now necessary to identify the thousands of mutations involved in infertility and the hundreds of genes necessary to achieve efficient spermatogenesis. The identification of variants in candidate genes will be fast, but long and complex analyses will be necessary to confirm the pathogenic effect of the identified variants and more so to understand the function of the identified proteins. We can be confident that the number of genes and mutations identified in the pathological context of infertility will increase exponentially over the next few years.

The next step will be to exploit these data to the benefit of the patients. Genotype/phenotype correlation will allow us to provide the best advice and care to patients and to answer many questions: is the risk of aneuploidy increased and is it therefore advisable to do ICSI? Or do we recommend performing PGD? Can we expect to find some sperm in this patient's testis and is it worth doing a biopsy? Such prognoses are already available for a handful of patients: ICSI is contraindicated to patients with a homozygous AURKC mutation and a testicular biopsy should not be attempted for men with an AZFa or b deletion (Krausz et al., 2014). The discovery of DPY19L2 provided a better insight of the molecular pathogeny of globozoospermia. A recent work showed that testicular sperm present similar DNA defects to epididymal sperm thus suggesting that testicular biopsy is not relevant in this case (Yassine et al., 2015a). Moreover, the demonstration that the sperm factor (PLCzeta) is absent in globozoospermic sperm might justify the use of AOA in this pathology (Escoffier et al., 2015). With an increased list of genes associated with infertility we can expect that a better prognosis and improved advice and treatments will soon be available to many more patients.

We are also convinced that another revolution is brewing. We are currently measuring the limitation of the whole IVF and ICSI strategy that is proposed to most infertile patients, especially those bearing the most severe defects. Alternative treatment strategies will only be possible with an in-depth comprehension of all aspects of spermatogenesis and of the physiopathology of sperm defects. The basis of this comprehension stems from the identification of the genes involved and of their function. We believe that male infertility might be among the pathologies that are best suited for targeted protein therapy and we are convinced that restoration of a functional spermatogenesis will be possible by reintroducing a deficient or missing protein as: (i) treatment success can be measured easily and objectively by a mere spermogram, (ii) in man spermatogenesis lasts ~70 days. This corresponds to the necessary treatment duration to obtain functional gametes, which can then be cryopreserved and used at a later date to initiate any number of pregnancies. This relatively short treatment is in sharp contrast with life-long supplementations needed for most other genetic diseases. We thus believe the identification of the genetic causes of infertility will permit us to develop therapies tailored to restore specifically the identified defects. Different options are possible, ranging from direct protein or mRNA injection into

the rete testis (Ogawa et al., 1997) to injection of expression vectors via lipid shuttles or a virus vector such as adeno-associated viruses (AAV). As the targeted cells produce the gametes that will be used to achieve a pregnancy it is crucial that no genetic material is inserted into the genome, to preclude the transmission of potential deleterious effects to the next generations. At the moment even AAV vectors, which are considered as non-integrating and have been used successfully on post mitotic tissues in many clinical trials, have been found integrating the genome of the target cells at a frequency of 10^{-4} – 10^{-5} (Kaufmann et al., 2013). For the time being, this low level of insertion precludes the utilization of viral vectors in the context of infertility, but we can hope that fully non-integrative vector will soon be available. Protein therapy is also possible ex vivo in testis culture systems but it is labor intensive and therefore less amenable to transposition to many different specific genetic defects. KO mice that have so far been used to study the physiopathology of abnormal spermatogenesis should now be used to evaluate the feasibility of different treatment options. The efficient delivery of the replacement protein will be subordinate to the nature and the localization of the defective native protein and the mode of delivery will have to be adjusted to ensure specificity. We however firmly believe that the restricted localization of the targeted cells (in the testis) and the limited duration of treatment clearly favor the development of innovative therapies applied to male infertility.

In conclusion, now is a very exciting time in the field of the genetics of infertility. We have so far seen only a small part of the tip of the iceberg but we are confident that the rest will come to light in the foreseeable future. We believe that this work will be the cornerstone of the next revolution in ART allowing the clinician not just to bypass the abnormalities by forceful ICSI but to propose new therapeutic strategies allowing the selective repair of the broken parts.

Authors' roles

Conception and organization of manuscript: C.C., C.A. and P.F.R.; literature search and analysis of evidence: C.C., J.E. and G.M.; writing and editing: C.C., J.E., G.M., C.A. and P.F.R.

Funding

Part of this work was funded by the Agence Nationale de la Recherche (ANR Genopat 2009, projet ICG2I)

Conflict of interest

None declared.

References

- Achard V, Paulmyer-Lacroix O, Mercier G, Porcu G, Saïas-Magnan J, Metzler-Guillemain C, Guichaoua MR. Reproductive failure in patients with various percentages of macronuclear spermatozoa: high level of aneuploid and polyploid spermatozoa. *J Androl* 2007;28:600–606.
- Afzelius BA, Eliasson R. Male and female infertility problems in the immotile-cilia syndrome. *Eur J Respir Dis Suppl* 1983;127:144–147.
- Alvarez Sedó C, Rawe VY, Chemes HE. Acrosomal biogenesis in human globozoospermia: immunocytochemical, ultrastructural and proteomic studies. *Hum Reprod* 2012;27:1912–1921.
- Amdani SN, Jones C, Coward K. Phospholipase C zeta (PLCζ): oocyte activation and clinical links to male factor infertility. *Adv Biol Regul* 2013;53:292–308.
- Anton-Lamprecht I, Kotzur B, Schopf E. Round-headed human spermatozoa. *Fertil Steril* 1976;27:685–693.
- Antony D, Becker-Heck A, Zariwala MA, Schmidts M, Onoufriadias A, Forouhan M, Wilson R, Taylor-Cox T, Dewar A, Jackson C et al. Mutations in CCDC39 and CCDC40 are the major cause of primary ciliary dyskinesia with axonemal disorganization and absent inner dynein arms. *Hum Mutat* 2013;34:462–472.
- Assou S, Anahory T, Pantesco V, Le Carroux T, Pellestor F, Klein B, Reyftmann L, Dechaud H, De Vos J, Hamamah S. The human cumulus—oocyte complex gene-expression profile. *Hum Reprod* 2006;21:1705–1719.
- Audouard C, Christians E. Hsp90β1 knockout targeted to male germline: a mouse model for globozoospermia. *Fertil Steril* 2011;95:1475–1477.e1–4.
- Austin-Tse C, Halbritter J, Zariwala MA, Gilberti RM, Gee HY, Hellman N, Pathak N, Liu Y, Panizzi JR, Patel-King RS et al. Zebrafish ciliopathy screen plus human mutational analysis identifies C21orf59 and CCDC65 defects as causing primary ciliary dyskinesia. *Am J Hum Genet* 2013;93:672–686.
- Avo Santos M, van de Werken C, de Vries M, Jahr H, Vromans MJM, Laven JSE, Fauzer BC, Kops GJ, Lens SM, Baart EB. A role for Aurora C in the chromosomal passenger complex during human preimplantation embryo development. *Hum Reprod* 2011;26:1868–1881.
- Baccetti B, Burrini AG, Capitani S, Collodel G, Moretti E, Piomboni P, Renieri T. Notulae seminologicae. 2. The 'short tail' and 'stump' defect in human spermatozoa. *Andrologia* 1993;25:331–335.
- Baccetti B, Bruni E, Gambera L, Moretti E, Piomboni P. An ultrastructural and immunocytochemical study of a rare genetic sperm tail defect that causes infertility in humans. *Fertil Steril* 2004;82:463–468.
- Baccetti B, Collodel G, Estenoz M, Manca D, Moretti E, Piomboni P. Gene deletions in an infertile man with sperm fibrous sheath dysplasia. *Hum Reprod* 2005a;20:2790–2794.
- Baccetti B, Collodel G, Gambera L, Moretti E, Serafini F, Piomboni P. Fluorescence in situ hybridization and molecular studies in infertile men with dysplasia of the fibrous sheath. *Fertil Steril* 2005b;84:123–129.
- Baker MA, Naumovski N, Hetherington L, Weinberg A, Velkov T, Aitken RJ. Head and flagella subcompartmental proteomic analysis of human spermatozoa. *Proteomics* 2013;13:61–74.
- Balboul AZ, Schindler K. Selective disruption of aurora C kinase reveals distinct functions from aurora B kinase during meiosis in mouse oocytes. *PLoS Genet* 2014;10:e1004194.
- Bartoloni L, Blouin J-L, Pan Y, Gehrig C, Maiti AK, Scamuffa N, Rossier C, Jorissen M, Armento M, Meeks M et al. Mutations in the DNAH11 (axonemal heavy chain dynein type II) gene cause one form of situs inversus totalis and most likely primary ciliary dyskinesia. *Proc Natl Acad Sci USA* 2002;99:10282–10286.
- Baudat F, Buard J, Grey C, Fledel-Alon A, Ober C, Przeworski M, Coop G, de Massy B. PRDM9 is a major determinant of meiotic recombination hotspots in humans and mice. *Science* 2010;327:836–840.
- Becker M, Stolz A, Ertych N, Bastians H. Centromere localization of INCENP-Aurora B is sufficient to support spindle checkpoint function. *Cell Cycle Georget Tex* 2010;9:1360–1372.
- Becker-Heck A, Zohn IE, Okabe N, Pollock A, Lenhart KB, Sullivan-Brown J, McSheene J, Loges NT, Olbrich H, Haefner K et al. The coiled-coil domain containing protein CCDC40 is essential for motile cilia function and left-right axis formation. *Nat Genet* 2011;43:79–84.
- Ben Khelifa M, Zouari R, Harbuz R, Halouani L, Arnoult C, Lunardi J, Ray PF. A new AURKC mutation causing macrozoospermia: implications for human spermatogenesis and clinical diagnosis. *Mol Hum Reprod* 2011;17:762–768.
- Ben Khelifa M, Coutton C, Blum MGB, Abada F, Harbuz R, Zouari R, Guichet A, May-Panloup P, Mitchell V, Rollet J et al. Identification of a new recurrent aurora kinase C mutation in both European and African men with macrozoospermia. *Hum Reprod* 2012;27:3337–3346.
- Ben Khelifa M, Coutton C, Zouari R, Karaouzène T, Rendu J, Bidart M, Yassine S, Pierre V, Delaroche J, Hennebicq S et al. Mutations in DNAH1, which encodes an inner arm heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. *Am J Hum Genet* 2014;94:95–104.
- Benzacken B, Gavelle FM, Martin-Pont B, Dupuy O, Lièvre N, Hugues JN, Wolf JP. Familial sperm polyploidy induced by genetic spermatogenesis failure: case report. *Hum Reprod* 2001;16:2646–2651.

- Bernard M, Sanseau P, Henry C, Couturier A, Prigent C. Cloning of STK13, a third human protein kinase related to *Drosophila* aurora and budding yeast *Ipl1* that maps on chromosome 19q13.3-ter. *Genomics* 1998; **53**:406–409.
- Bettencourt-Dias M, Glover DM. Centrosome biogenesis and function: centrosomics brings new understanding. *Nat Rev Mol Cell Biol* 2007; **8**:451–463.
- Bisgrove BW, Yost HJ. The roles of cilia in developmental disorders and disease. *Development* 2006; **133**:4131–4143.
- Blanchard Y, Lescoat D, Le Lannou D. Anomalous distribution of nuclear basic proteins in round-headed human spermatozoa. *Andrologia* 1990; **22**:549–555.
- Blanchon S, Legendre M, Copin B, Duquesnoy P, Montantin G, Kott E, Dastot F, Jeanson L, Cachanado M, Rousseau A et al. Delineation of CCDC39/CCDC40 mutation spectrum and associated phenotypes in primary ciliary dyskinesia. *J Med Genet* 2012; **49**:410–416.
- Boivin J, Bunting L, Collins JA, Nygren KG. International estimates of infertility prevalence and treatment-seeking: potential need and demand for infertility medical care. *Hum Reprod* 2007; **22**:1506–1512.
- Boon M, Jorissen M, Proesmans M, De Boeck K. Primary ciliary dyskinesia, an orphan disease. *Eur J Pediatr* 2013; **172**:151–162.
- Brahem S, Mehdi M, Elgezal H, Saad A. Analysis of sperm aneuploidies and DNA fragmentation in patients with globozoospermia or with abnormal acrosomes. *Urology* 2011; **77**:1343–1348.
- Brahem S, Mehdi M, Elgezal H, Saad A. Study of aneuploidy rate and sperm DNA fragmentation in large-headed, multiple-tailed spermatozoa. *Andrologia* 2012; **44**:130–135.
- Brenner S. The genetics of *Caenorhabditis elegans*. *Genetics* 1974; **77**:71–94.
- Brown PR, Mikl K, Harper DB, Eddy EM. A-kinase anchoring protein 4 binding proteins in the fibrous sheath of the sperm flagellum. *Biol Reprod* 2003; **68**:2241–2248.
- Brown JR, Koretke KK, Birkeland ML, Sanseau P, Patrick DR. Evolutionary relationships of Aurora kinases: implications for model organism studies and the development of anti-cancer drugs. *BMC Evol Biol* 2004; **4**:39.
- Brugo-Olmedo S, Chillik C, Kopelman S. Definition and causes of infertility. *Reprod Biomed Online* 2001; **2**:41–53.
- Buettner FFR, Ashikov A, Tiemann B, Lehle L, Bakker H. *C. elegans* DPY-19 is a C-mannosyltransferase glycosylating thrombospondin repeats. *Mol Cell* 2013; **50**:295–302.
- Bui KH, Sakakibara H, Movassagh T, Oiwa K, Ishikawa T. Molecular architecture of inner dynein arms *in situ* in *Chlamydomonas reinhardtii* flagella. *J Cell Biol* 2008; **183**:923–932.
- Bukowy-Bieryło Z, Ziętkiewicz E, Loges NT, Wittmer M, Geremek M, Olbrich H, Fliegauf M, Voelkel K, Rutkiewicz E, Rutland J et al. RPGR mutations might cause reduced orientation of respiratory cilia. *Pediatr Pulmonol* 2013; **48**:352–363.
- Campbell PK, Waymire KG, Heier RL, Sharer C, Day DE, Reimann H, Jaje JM, Friedrich GA, Burmeister M, Bartness TJ et al. Mutation of a novel gene results in abnormal development of spermatid flagella, loss of intermale aggression and reduced body fat in mice. *Genetics* 2002; **162**:307.
- Carmena M, Earnshaw WC. The cellular geography of aurora kinases. *Nat Rev Mol Cell Biol* 2003; **4**:842–854.
- Carmena M, Wheelock M, Funabiki H, Earnshaw WC. The chromosomal passenger complex (CPC): from easy rider to the godfather of mitosis. *Nat Rev Mol Cell Biol* 2012; **13**:789–803.
- Carrell DT, Emery BR, Liu L. Characterization of aneuploidy rates, protamine levels, ultrastructure, and functional ability of round-headed sperm from two siblings and implications for intracytoplasmic sperm injection. *Fertil Steril* 1999; **71**:511–516.
- Carson AR, Cheung J, Scherer SW. Duplication and relocation of the functional DPY19L2 gene within low copy repeats. *BMC Genomics* 2006; **7**:45.
- Castleman VH, Romio L, Chodhari R, Hirst RA, de Castro SCP, Parker KA, Ybot-Gonzalez P, Emes RD, Wilson SW, Wallis C et al. Mutations in radial spoke head protein genes RSPH9 and RSPH4A cause primary ciliary dyskinesia with central-microtubular-pair abnormalities. *Am J Hum Genet* 2009; **84**:197–209.
- Chan CS, Botstein D. Isolation and characterization of chromosome-gain and increase-in-ploidy mutants in yeast. *Genetics* 1993; **135**:677–691.
- Chelli MH, Albert M, Ray PF, Guthausen B, Izard V, Hammoud I, Selva J, Vialard F. Can intracytoplasmic morphologically selected sperm injection be used to select normal-sized sperm heads in infertile patients with macrocephalic sperm head syndrome? *Fertil Steril* 2010; **93**:1347.e1–5.
- Chemes HE. Sperm centrioles and their dual role in flagellogenesis and cell cycle of the zygote. In: Schatten H. (ed) *The Centrosome*. Humana Press, 2012, 33–48. Available at: http://link.springer.com/gate2.inist.fr/chapter/10.1007/978-1-62703-035-9_2. Accessed June 20, 2014.
- Chemes HE, Alvarez-Sedo C. Tales of the tail and sperm head aches: changing concepts on the prognostic significance of sperm pathologies affecting the head, neck and tail. *Asian J Androl* 2012; **14**:14–23.
- Chemes EH, Rawe YV. Sperm pathology: a step beyond descriptive morphology. Origin, characterization and fertility potential of abnormal sperm phenotypes in infertile men. *Hum Reprod Update* 2003; **9**:405–428.
- Chemes HE, Rawe YV. The making of abnormal spermatozoa: cellular and molecular mechanisms underlying pathological spermiogenesis. *Cell Tissue Res* 2010; **341**:349–357.
- Chemes HE, Brugo S, Zanchetti F, Carrere C, Lavieri JC. Dysplasia of the fibrous sheath: an ultrastructural defect of human spermatozoa associated with sperm immotility and primary sterility. *Fertil Steril* 1987; **48**:664–669.
- Chemes HE, Olmedo SB, Carrere C, Osés R, Carizza C, Leisner M, Blaquier J. Ultrastructural pathology of the sperm flagellum: association between flagellar pathology and fertility prognosis in severely asthenozoospermic men. *Hum Reprod* 1998; **13**:2521–2526.
- Chen H-L, Tang C-JC, Chen C-Y, Tang TK. Overexpression of an Aurora-C kinase-deficient mutant disrupts the Aurora-B/INCENP complex and induces polyploidy. *J Biomed Sci* 2005; **12**:297–310.
- Chianese C, Fino MG, Riera Escamilla A, López Rodrigo O, Vinci S, Guarducci E, Daguina F, Muratori M, Tamburrino L, Lo Giacco D et al. Comprehensive investigation in patients affected by sperm macrocephaly and globozoospermia. *Andrology* 2015 Mar 6. doi: 10.1111/andr.12016 (epub ahead of print).
- Christensen GL, Ivanov IP, Atkins JF, Campbell B, Carrell DT. Identification of polymorphisms in the Hrb, GOPC, and Csnk2a2 genes in two men with globozoospermia. *J Androl* 2006; **27**:11–15.
- Collodel G, Moretti E. Sperm morphology and aneuploidies: defects of supposed genetic origin. *Andrologia* 2006; **38**:208–215.
- Coutton C, Zouari R, Abada F, Ben Khelifa M, Merdassi G, Triki C, Escalier D, Hesters L, Mitchell V, Levy R et al. MLPA and sequence analysis of DPY19L2 reveals point mutations causing globozoospermia. *Hum Reprod* 2012; **27**:2549–2558.
- Coutton C, Abada F, Karaouzene T, Sanlaville D, Satre V, Lunardi J, Jouk P-S, Arnoult C, Thierry-Mieg N, Ray PF. Fine characterisation of a recombination hotspot at the DPY19L2 locus and resolution of the paradoxical excess of duplications over deletions in the general population. *PLoS Genet* 2013; **9**:e1003363.
- Dale B, Iaccarino M, Fortunato A, Gragnaniello G, Kyozuka K, Tosti E. A morphological and functional study of fusibility in round-headed spermatozoa in the human. *Fertil Steril* 1994; **61**:336–340.
- Dam AHDM, Feenstra I, Westphal JR, Ramos L, van Golde RJT, Kremer JAM. Globozoospermia revisited. *Hum Reprod Update* 2007; **13**:63–75.
- Dam AHDM, Koscienski I, Kremer JAM, Moutou C, Jaeger A-S, Oudakker AR, Tournaye H, Charlet N, Lagier-Tourenne C, van Bokhoven H et al. Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet* 2007b; **81**:813–820.
- Dam AH, Ramos L, Dijkman HB, Woestenenk R, Robben H, van den Hoven L, Kremer JA. Morphology of partial globozoospermia. *J Androl* 2011; **32**:199–206.
- Daniels MLA, Leigh MW, Davis SD, Armstrong MC, Carson JL, Hazucha M, Dell SD, Eriksson M, Collins FS, Knowles MR et al. Founder mutation in RSPH4A identified in patients of Hispanic descent with primary ciliary dyskinesia. *Hum Mutat* 2013; **34**:1352–1356.
- David G, Feneux D, Serres C, Escalier D, Jouannet P. A new entity of sperm pathology: peri-axonemal flagellar dyskinesia. *Bull Académie Natl Médecine* 1993; **177**:263–271; discussion 272–275.
- Dávila Garza SA, Patrizio P. Reproductive outcomes in patients with male infertility because of Klinefelter's syndrome, Kartagener's syndrome, round-head sperm, dysplasia fibrous sheath, and "stump" tail sperm: an updated literature review. *Curr Opin Obstet Gynecol* 2013; **25**:229–246.
- De Boer P, de Vries M, Ramos L. A mutation study of sperm head shape and motility in the mouse: lessons for the clinic. *Andrology* 2014 Dec 16. doi: 10.1111/andr.300 (Epub ahead of print).
- Devillard F, Metzler-Guillemain C, Pelletier R, DeRobertis C, Bergues U, Hennebicq S, Guichaoua M, Sèle B, Rousseaux S. Polyploidy in large-headed sperm: FISH study of three cases. *Hum Reprod* 2002; **17**:1292–1298.
- Dieterich K, Soto Rifo R, Faure AK, Hennebicq S, Ben Amar B, Zahri M, Perrin J, Martinez D, Sèle B, Jouk P-S et al. Homozygous mutation of AURKC yields

- large-headed polyploid spermatozoa and causes male infertility. *Nat Genet* 2007; **39**:661–665.
- Dieterich K, Zouari R, Harbuz R, Vialard F, Martinez D, Bellayou H, Prisant N, Zoghmar A, Guichaoua MR, Koscienski I et al. The Aurora Kinase C c.144delC mutation causes meiosis I arrest in men and is frequent in the North African population. *Hum Mol Genet* 2009; **18**:1301–1309.
- Dirami T, Rode B, Jollivet M, Da Silva N, Escalier D, Gaitch N, Norez C, Tuffery P, Wolf J-P, Becq F et al. Missense mutations in SLC2A8, encoding a sperm-specific activator of CFTR, are associated with human asthenozoospermia. *Am J Hum Genet* 2013; **92**:760–766.
- Dirican EK, Isik A, Vicdan K, Sozen E, Suludere Z. Clinical pregnancies and livebirths achieved by intracytoplasmic injection of round headed acrosomeless spermatozoa with and without oocyte activation in familial globozoospermia: case report. *Asian J Androl* 2008; **10**:332–336.
- Duquesnoy P, Escudier E, Vincensini L, Freshour J, Bridoux A-M, Coste A, Deschildre A, de Blic J, Legendre M, Montantin G et al. Loss-of-function mutations in the human ortholog of Chlamydomonas reinhardtii ODA7 disrupt dynein arm assembly and cause primary ciliary dyskinesia. *Am J Hum Genet* 2009; **85**:890–896.
- Deriez B, Duquesnoy P, Escudier E, Bridoux A-M, Escalier D, Rayet I, Marcos E, Vojtek A-M, Bercher J-F, Amselem S. A common variant in combination with a nonsense mutation in a member of the thioredoxin family causes primary ciliary dyskinesia. *Proc Natl Acad Sci USA* 2007; **104**:3336–3341.
- Eddy EM, Toshimori K, O'Brien DA. Fibrous sheath of mammalian spermatozoa. *Microsc Res Tech* 2003; **61**:103–115.
- Elinati E, Kuentz P, Redin C, Jaber S, Vanden Meerschaut F, Makarian J, Koscienski I, Nasr-Esfahani MH, Demiroli A, Gurgan T et al. Globozoospermia is mainly due to DPY19L2 deletion via non-allelic homologous recombination involving two recombination hotspots. *Hum Mol Genet* 2012; **21**:3695–3702.
- El Kerch F, Lamzouri A, Laarabi FZ, Zahri M, Ben Amar B, Sefiani A. Confirmation of the high prevalence in Morocco of the homozygous mutation c.144delC in the aurora kinase C gene (AURKC) in the teratozoospermia with large-headed spermatozoa. *J Gynecol Obstet Biol Reprod* 2011; **40**:329–333.
- Eloualid A, Rouba H, Rhaissi H, Barakat A, Louanjli N, Bashamboo A, McElreavey K. Prevalence of the Aurora kinase C c.144delC mutation in infertile Moroccan men. *Fertil Steril* 2014; **101**:1086–1090.
- Escalier D. Human spermatozoa with large heads and multiple flagella: a quantitative ultrastructural study of 6 cases. *Biol Cell* 1983; **48**:65–74.
- Escalier D, David G. Pathology of the cytoskeleton of the human sperm flagellum: axonemal and peri-axonemal anomalies. *Biol Cell* 1984; **50**:37–52.
- Escalier D, Serres C. Aberrant distribution of the peri-axonemal structures in the human spermatozoon: possible role of the axoneme in the spatial organization of the flagellar components. *Biol Cell* 1985; **53**:239–250.
- Escalier D, Bermúdez D, Gallo JM, Viellefond A, Schrével J. Cytoplasmic events in human meiotic arrest as revealed by immunolabelling of spermatocyte proacrosin. *Differentiation* 1992; **51**:233–243.
- Escoffier J, Yassine S, Lee HC, Martinez G, Delaroche J, Coutton C, Karaouzène T, Zouari R, Metzler-Guillemain C, Pernet-Gallay K et al. Subcellular localization of phospholipase C ζ in human sperm and its absence in DPY19L2-deficient sperm are consistent with its role in oocyte activation. *Mol Hum Reprod* 2015; **21**:157–168.
- Failli M, Bartoloni L, Letourneau A, Munoz A, Falconnet E, Rossier C, de Santi MM, Santamaría F, Sacco O, DeLozier-Blanchet CD et al. Mutations in DNAH5 account for only 15% of a non-preselected cohort of patients with primary ciliary dyskinesia. *J Med Genet* 2009; **46**:281–286.
- Fauque P, Albert M, Serres C, Viallon V, Davy C, Epelboin S, Chalas C, Jouannet P, Patrat C. From ultrastructural flagellar sperm defects to the health of babies conceived by ICSI. *Reprod Biomed Online* 2009; **19**:326–336.
- Fernández-Miranda G, Trakala M, Martín J, Escobar B, González A, Ghyselinck NB, Ortega S, Cañamero M, Pérez de Castro I, Malumbres M. Genetic disruption of aurora B uncovers an essential role for aurora C during early mammalian development. *Development* 2011; **138**:2661–2672.
- Fliegauf M, Olbrich H, Horvath J, Wildhaber JH, Zariwala MA, Kennedy M, Knowles MR, Omran H. Mislocalization of DNAH5 and DNAH9 in respiratory cells from patients with primary ciliary dyskinesia. *Am J Respir Crit Care Med* 2005; **171**:1343–1349.
- Förkle-Gerloff S, Töpfer-Petersen E, Müller-Esterl W, Mansouri A, Schatz R, Schirren C, Schill W, Engel W. Biochemical and genetic investigation of round-headed spermatozoa in infertile men including two brothers and their father. *Andrologia* 1984; **16**:187–202.
- Fossella J, Samant SA, Silver LM, King SM, Vaughan KT, Olds-Clarke P, Johnson KA, Mikami A, Vallee RB, Pilder SH. An axonemal dynein at the Hybrid Sterility 6 locus: implications for t haplotype-specific male sterility and the evolution of species barriers. *Mamm Genome* 2000; **11**:8–15.
- Frohnert C, Schweizer S, Hoyer-Fender S. SPAG4L/SPAG4L-2 are testis-specific SUN domain proteins restricted to the apical nuclear envelope of round spermatids facing the acrosome. *Mol Hum Reprod* 2011; **17**:207–218.
- Fujihara Y, Satoh Y, Inoue N, Isotani A, Ikawa M, Okabe M. SPACA1-deficient male mice are infertile with abnormally shaped sperm heads reminiscent of globozoospermia. *Development* 2012; **139**:3583–3589.
- Funaki T, Kon S, Tanabe K, Natsume W, Sato S, Shimizu T, Yoshida N, Wong WF, Ogura A, Ogawa T et al. The Arf GAP SMAP2 is necessary for organized vesicle budding from the trans-Golgi network and subsequent acrosome formation in spermiogenesis. *Mol Biol Cell* 2013; **24**:2633–2644.
- German J, Rasch EM, Huang CY, MacLeod J, Imperato-McGinley J. Human infertility due to production of multiple-tailed spermatozoa with excessive amounts of DNA. *Am Soc Hum Genet* 1981; **33**:64A.
- Ghedir H, Mehri A, Mehdi M, Brahem S, Saad A, Ibala-Romdhane S. Meiotic segregation and sperm DNA fragmentation in Tunisian men with dysplasia of the fibrous sheath (DFS) associated with head abnormalities. *J Assist Reprod Genet* 2014; **31**:1167–1174.
- Glover DM, Leibowitz MH, McLean DA, Parry H. Mutations in aurora prevent centrosome separation leading to the formation of monopolar spindles. *Cell* 1995; **81**:95–105.
- Göb E, Schmitt J, Benavente R, Alsheimer M. Mammalian sperm head formation involves different polarization of two novel LINC complexes. *PLoS One* 2010; **5**:e10272.
- Goldenson B, Crispino JD. The aurora kinases in cell cycle and leukemia. *Oncogene* 2014; **34**:537–545.
- Guichard C, Harricane MC, Lafitte JJ, Godard P, Zaegel M, Tack V, Lalau G, Bouvagnet P. Axonemal dynein intermediate-chain gene (DNAII) mutations result in situs inversus and primary ciliary dyskinesia (Kartagener syndrome). *Am J Hum Genet* 2001; **68**:1030–1035.
- Guthausen B, Vialard F, Dakouane M, Izard V, Albert M, Selva J. Chromosomal analysis of spermatozoa with normal-sized heads in two infertile patients with macrocephalic sperm head syndrome. *Fertil Steril* 2006; **85**:e5–750.e7.
- Hao Z, Wolkowicz MJ, Shetty J, Klotz K, Bolling L, Sen B, Westbrook VA, Coonrod S, Flickinger CJ, Herr JC. SAMP32, a testis-specific, isoantigenic sperm acrosomal membrane-associated protein. *Biol Reprod* 2002; **66**:735–744.
- Haraguchi S, Tsuda M, Kitajima S, Sasaoka Y, Nomura-Kitabayashid A, Kurokawa K, Saga Y. nanosI: a mouse nanos gene expressed in the central nervous system is dispensable for normal development. *Mech Dev* 2003; **120**:721–731.
- Harbuz R, Zouari R, Pierre V, Ben Khelifa M, Kharouf M, Coutton C, Merdassi G, Abada F, Escoffier J, Nikas Y et al. A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. *Am J Hum Genet* 2011; **88**:351–361.
- Heytens E, Parrington J, Coward K, Young C, Lambrecht S, Yoon S-Y, Fissore RA, Hamer R, Deane CM, Ruas M et al. Reduced amounts and abnormal forms of phospholipase C zeta (PLC ζ) in spermatozoa from infertile men. *Hum Reprod* 2009; **24**:2417–2428.
- Hjeij R, Lindstrand A, Francis R, Zariwala MA, Liu X, Li Y, Damerla R, Dougherty GW, Abouhamad M, Olbrich H et al. ARMC4 mutations cause primary ciliary dyskinesia with randomization of left/right body asymmetry. *Am J Hum Genet* 2013; **93**:357–367.
- Hjeij R, Onoufriadi A, Watson CM, Slagle CE, Klena NT, Dougherty GW, Kurkwiak M, Loges NT, Diggle CP, Morante NFC et al. CCDC151 mutations cause primary ciliary dyskinesia by disruption of the outer dynein arm docking complex formation. *Am J Hum Genet* 2014; **95**:257–274.
- Holstein AF, Schirren C, Schirren CG. Human spermatids and spermatozoa lacking acrosomes. *J Reprod Fertil* 1973; **35**:489–491.
- Honigberg L, Kenyon C. Establishment of left/right asymmetry in neuroblast migration by UNC-40/DCC, UNC-73/Trio and DPY-19 proteins in *C. elegans*. *Development* 2000; **127**:4655–4668.
- Horani A, Druley TE, Zariwala MA, Patel AC, Levinson BT, Van Arendonk LG, Thornton KC, Giacalone JC, Albee AJ, Wilson KS et al. Whole-exome capture and sequencing identifies HEATR2 mutation as a cause of primary ciliary dyskinesia. *Am J Hum Genet* 2012; **91**:685–693.
- Horani A, Ferkol TW, Shoseyov D, Wasserman MG, Oren YS, Kerem B, Amirav I, Cohen-Cymberknob M, Dutcher SK, Brody SL et al. LRRC6 mutation causes primary ciliary dyskinesia with dynein arm defects. *PLoS One* 2013a; **8**:e59436.

- Horani A, Brody SL, Ferkol TW, Shoseyov D, Wasserman MG, Ta-shma A, Wilson KS, Bayly PV, Amirav I, Cohen-Cymberknob M et al. CCDC65 mutation causes primary ciliary dyskinesia with normal ultrastructure and hyperkinetic cilia. *PLoS One* 2013; **8**:e72299.
- Horani A, Brody SL, Ferkol TW. Picking up speed: advances in the genetics of primary ciliary dyskinesia. *Pediatr Res* 2014; **75**:158–164.
- Hornef N, Olbrich H, Horvath J, Zariwala MA, Fliegauf M, Loges NT, Wildhaber J, Noone PG, Kennedy M, Antonarakis SE et al. DNAH5 mutations are a common cause of primary ciliary dyskinesia with outer dynein arm defects. *Am J Respir Crit Care Med* 2006; **174**:120–126.
- Ibañez-Tallon I, Gorokhova S, Heintz N. Loss of function of axonemal dynein Mdnah5 causes primary ciliary dyskinesia and hydrocephalus. *Hum Mol Genet* 2002; **11**:715–721.
- Ibañez-Tallon I, Heintz N, Omran H. To beat or not to beat: roles of cilia in development and disease. *Hum Mol Genet* 2003; **12**(Spec No 1):R27–R35.
- Ikegami K, Sato S, Nakamura K, Ostrowski LE, Setou M. Tubulin polyglutamylated is essential for airway ciliary function through the regulation of beating asymmetry. *Proc Natl Acad Sci USA* 2010; **107**:10490–10495.
- Int' Veld PA, Broekmans FJ, de France HF, Pearson PL, Pieters MH, van Kooij RJ. Intracytoplasmic sperm injection (ICSI) and chromosomally abnormal spermatozoa. *Hum Reprod* 1997; **12**:752–754.
- Inaba K. Molecular architecture of the sperm flagella: molecules for motility and signaling. *Zoolog Sci* 2003; **20**:1043–1056.
- Inaba K. Sperm flagella: comparative and phylogenetic perspectives of protein components. *Mol Hum Reprod* 2011; **17**:524–538.
- Ito C, Suzuki-Toyota F, Maekawa M, Toyama Y, Yao R, Noda T, Toshimori K. Failure to assemble the peri-nuclear structures in GOPC deficient spermatids as found in round-headed spermatozoa. *Arch Histol Cytol* 2004; **67**:349–360.
- Jeganathan D, Chodhari R, Meeks M, Faeroe O, Smyth D, Nielsen K, Amirav I, Luder AS, Bisgaard H, Gardiner RM et al. Loci for primary ciliary dyskinesia map to chromosome 16p12.1-12.2 and 15q13.1-15.1 in Faroe Islands and Israeli Druze genetic isolates. *J Med Genet* 2004; **41**:233–240.
- Jerber J, Baas D, Soulavie F, Chhin B, Cortier E, Vesque C, Thomas J, Durand B. The coiled-coil domain containing protein CCDC151 is required for the function of IFT-dependent motile cilia in animals. *Hum Mol Genet* 2014; **23**:563–577.
- Jungwirth A, Giwercman A, Tournaye H, Diemer T, Kopa Z, Dohle G, Krausz C, European Association of Urology Working Group on Male Infertility. European Association of Urology guidelines on Male Infertility: the 2012 update. *Eur Urol* 2012; **62**:324–332.
- Kahraman S, Akarsu C, Cengiz G, Dirican K, Sözen E, Can B, Güven C, Vanderzwalmen P. Fertility of ejaculated and testicular megalohed spermatozoa with intracytoplasmic sperm injection. *Hum Reprod* 1999; **14**:726–730.
- Kahraman S, Sertetyl S, Findikli N, Kumtepe Y, Oncu N, Melil S, Unal S, Yelke H, Vanderzwalmen P. Effect of PGD on implantation and ongoing pregnancy rates in cases with predominantly macrocephalic spermatozoa. *Reprod Biomed Online* 2004; **9**:79–85.
- Kang-Decker N, Mantchev GT, Juneja SC, McNiven MA, van Deursen JM. Lack of acrosome formation in Hrb-deficient mice. *Science* 2001; **294**:1531–1533.
- Karaca N, Yilmaz R, Kanten GE, Kervancioglu E, Solakoglu S, Kervancioglu ME. First successful pregnancy in a globozoospermic patient having homozygous mutation in SPATA16. *Fertil Steril* 2014; **102**:103–107.
- Karaouzène T, El Atifi M, Issartel J-P, Grepillat M, Coutton C, Martinez D, Arnoult C, Ray P. Comparative testicular transcriptome of wild type and globozoospermic Dpy19L2 knock out mice. *Basic Clin Androl* 2013; **23**. Available at: <http://hal.univ-grenoble-alpes.fr/inserm-00880088/>. Accessed May 7, 2014.
- Kashir J, Heindryckx B, Jones C, Sutter PD, Parrington J, Coward K. Oocyte activation, phospholipase C zeta and human infertility. *Hum Reprod Update* 2010; **16**:690–703.
- Kaufmann KB, Büning H, Galy A, Schambach A, Grez M. Gene therapy on the move. *EMBO Mol Med* 2013; **5**:1642–1661.
- Khan J, Ezan F, Crémel J-Y, Fautrel A, Gilot D, Lambert M, Benaud C, Troadec M-B, Prigent C. Overexpression of active Aurora-C kinase results in cell transformation and tumour formation. *PLoS One* 2011; **6**:e26512.
- Khattari A, Reddy VP, Pandey RK, Sudhakar DVS, Gupta NJ, Chakravarty BN, Deenadayal M, Singh L, Thangaraj K. Novel mutations in calcium/calmodulin-dependent protein kinase IV (CAMK4) gene in infertile men. *Int J Androl* 2012; **35**:810–818.
- Kierszenbaum AL, Tres LL, Rivkin E, Kang-Decker N, van Deursen JMA. The acroplaxome is the docking site of Golgi-derived myosin Va/Rab27a/b-containing proacrosomal vesicles in wild-type and Hrb mutant mouse spermatids. *Biol Reprod* 2004; **70**:1400–1410.
- Kilani Z, Ismail R, Ghunaim S, Mohamed H, Hughes D, Brewis I, Barratt CLR. Evaluation and treatment of familial globozoospermia in five brothers. *Fertil Steril* 2004; **82**:1436–1439.
- Kimmins S, Crosio C, Kotaja N, Hirayama J, Monaco L, Höög C, van Duin M, Gossen JA, Sassone-Corsi P. Differential functions of the Aurora-B and Aurora-C kinases in mammalian spermatogenesis. *Mol Endocrinol* 2007; **21**:726–739.
- King SM. A solid-state control system for dynein-based ciliary/flagellar motility. *J Cell Biol* 2013; **201**:173–175.
- Knowles MR, Leigh MW, Carson JL, Davis SD, Dell SD, Ferkol TW, Olivier KN, Sagel SD, Rosenfeld M, Burns KA et al. Mutations of DNAH11 in patients with primary ciliary dyskinesia with normal ciliary ultrastructure. *Thorax* 2012; **67**:433–441.
- Knowles MR, Daniels LA, Davis SD, Zariwala MA, Leigh MW. Primary ciliary dyskinesia. Recent advances in diagnostics, genetics, and characterization of clinical disease. *Am J Respir Crit Care Med* 2013a; **188**:913–922.
- Knowles MR, Leigh MW, Ostrowski LE, Huang L, Carson JL, Hazucha MJ, Yin W, Berg JS, Davis SD, Dell SD et al. Exome sequencing identifies mutations in CCDC114 as a cause of primary ciliary dyskinesia. *Am J Hum Genet* 2013b; **92**:99–106.
- Knowles MR, Ostrowski LE, Loges NT, Hurd T, Leigh MW, Huang L, Wolf WE, Carson JL, Hazucha MJ, Yin W et al. Mutations in SPAG1 cause primary ciliary dyskinesia associated with defective outer and inner dynein arms. *Am J Hum Genet* 2013c; **93**:711–720.
- Knowles MR, Ostrowski LE, Leigh MW, Sears PR, Davis SD, Wolf WE, Hazucha MJ, Carson JL, Olivier KN, Sagel SD et al. Mutations in RSPH1 cause primary ciliary dyskinesia with a unique clinical and ciliary phenotype. *Am J Respir Crit Care Med* 2014; **189**:707–717.
- Kosaki K, Ikeda K, Miyakoshi K, Ueno M, Kosaki R, Takahashi D, Tanaka M, Torikata C, Yoshimura Y, Takahashi T. Absent inner dynein arms in a fetus with familial hydrocephalus-situs abnormality. *Am J Med Genet A* 2004; **129A**:308–311.
- Koscinski I, Elinati E, Fossard C, Redin C, Muller J, Velez de la Calle J, Schmitt F, Ben Khelifa M, Ray PF, Ray P et al. DPY19L2 deletion as a major cause of globozoospermia. *Am J Hum Genet* 2011; **88**:344–350.
- Kott E, Duquesnoy P, Copin B, Legendre M, Dastot-Le Moal F, Montantin G, Jeanson L, Tamalet A, Papon J-F, Siffroi J-P et al. Loss-of-function mutations in LRRC6, a gene essential for proper axonemal assembly of inner and outer dynein arms, cause primary ciliary dyskinesia. *Am J Hum Genet* 2012; **91**:958–964.
- Kott E, Legendre M, Copin B, Papon J-F, Dastot-Le Moal F, Montantin G, Duquesnoy P, Piterborth W, Amram D, Bassinet L et al. Loss-of-function mutations in RSPH1 cause primary ciliary dyskinesia with central-complex and radial-spoke defects. *Am J Hum Genet* 2013; **93**:561–570.
- Krausz C. Male infertility: pathogenesis and clinical diagnosis. *Best Pract Res Clin Endocrinol Metab* 2011; **25**:271–285.
- Krausz C, Hoeflsloot L, Simoni M, Tüttelmann F, European Academy of Andrology, European Molecular Genetics Quality Network. EAA/EMQN best practice guidelines for molecular diagnosis of Y-chromosomal microdeletions: state-of-the-art 2013. *Andrology* 2014; **2**:5–19.
- Kruger TF, Menkveld R, Stander FS, Lombard CJ, Van der Merwe JP, van Zyl JA, Smith K. Sperm morphologic features as a prognostic factor in in vitro fertilization. *Fertil Steril* 1986; **46**:1118–1123.
- Kuentz P, Vanden Meerschaut F, Elinati E, Nasr-Esfahani MH, Gurgan T, Iqbal N, Carré-Pigeon F, Brugnon F, Gitlin SA, Velez de la Calle J et al. Assisted oocyte activation overcomes fertilization failure in globozoospermic patients regardless of the DPY19L2 status. *Hum Reprod* 2013; **28**:1054–1061.
- Kullander S, Rausing A. On round-headed human spermatozoa. *Int J Fertil* 1975; **20**:33–40.
- Kuo Y-C, Lin Y-H, Chen H-I, Wang Y-Y, Chiou Y-W, Lin H-H, Pan H-A, Wu C-M, Su S-M, Hsu C-C et al. SEPT12 mutations cause male infertility with defective sperm annulus. *Hum Mutat* 2012; **33**:710–719.
- Kusz-Zamelczyk K, Sajek M, Spik A, Glazar R, Jędrzejczak P, Łatos-Bieleńska A, Kotecki M, Pawełczyk L, Jaruzelska J. Mutations of NANOS1, a human homologue of the Drosophila morphogen, are associated with a lack of germ cells in testes or severe oligo-astheno-teratozoospermia. *J Med Genet* 2013; **50**:187–193.
- Kyono K, Nakajo Y, Nishinaka C, Hattori H, Kyoya T, Ishikawa T, Abe H, Araki Y. A birth from the transfer of a single vitrified-warmed blastocyst using intracytoplasmic sperm injection with calcium ionophore oocyte activation in a globozoospermic patient. *Fertil Steril* 2009; **91**:931.e7–11.

- Lee L, Campagna DR, Pinkus JL, Mulhern H, Wyatt TA, Sisson JH, Pavlik JA, Pinkus GS, Fleming MD. Primary ciliary dyskinesia in mice lacking the novel ciliary protein Pcdp1. *Mol Cell Biol* 2008;28:949–957.
- Lerer-Goldstein T, Bel S, Shpungin S, Pery E, Motro B, Goldstein RS, Bar-Sheshet SI, Breitbart H, Nir U. TMF/ARA160: a key regulator of sperm development. *Dev Biol* 2010;348:12–21.
- Lessard C, Lothrop H, Schimenti JC, Handel MA. Mutagenesis-generated mouse models of human infertility with abnormal sperm. *Hum Reprod* 2007;22:159–166.
- Lewis-Jones I, Aziz N, Seshadri S, Douglas A, Howard P. Sperm chromosomal abnormalities are linked to sperm morphologic deformities. *Fertil Steril* 2003;79:212–215.
- Li X, Sakashita G, Matsuzaki H, Sugimoto K, Kimura K, Hanaoka F, Taniguchi H, Furukawa K, Urano T. Direct association with inner centromere protein (INCENP) activates the novel chromosomal passenger protein, Aurora-C. *J Biol Chem* 2004;279:47201–47211.
- Lin Y-N, Roy A, Yan W, Burns KH, Matzuk MM. Loss of zona pellucida binding proteins in the acrosomal matrix disrupts acrosome biogenesis and sperm morphogenesis. *Mol Cell Biol* 2007;27:6794–6805.
- Lin Y-H, Wang Y-Y, Chen H-I, Kuo Y-C, Chiou Y-W, Lin H-H, Wu C-M, Hsu C-C, Chiang H-S, Kuo P-L. SEPTIN12 genetic variants confer susceptibility to teratozoospermia. *PLoS One* 2012;7:e34011.
- Liu G, Shi Q-W, Lu G-X. A newly discovered mutation in PICK1 in a human with globozoospermia. *Asian J Androl* 2010;12:556–560.
- Liu P, Carvalho CMB, Hastings PJ, Lupski JR. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev* 2012;22:211–220.
- Loges NT, Olbrich H, Fenske L, Mussaffi H, Horvath J, Fliegauf M, Kuhl H, Bakta G, Peterfy E, Chodhari R et al. DNAI2 mutations cause primary ciliary dyskinesia with defects in the outer dynein arm. *Am J Hum Genet* 2008;83:547–558.
- Loges NT, Olbrich H, Becker-Heck A, Häffner K, Heer A, Reinhard C, Schmidts M, Kispert A, Zariwala MA, Leigh MW et al. Deletions and point mutations of LRRC50 cause primary ciliary dyskinesia due to dynein arm defects. *Am J Hum Genet* 2009;85:883–889.
- Lu L, Lin M, Xu M, Zhou Z-M, Sha J-H. Gene functional research using polyethylenimine-mediated in vivo gene transfection into mouse spermatogenic cells. *Asian J Androl* 2006;8:53–59.
- Lucas JS, Adam EC, Goggin PM, Jackson CL, Powles-Glover N, Patel SH, Humphreys J, Fray MD, Falconett E, Blouin J-L et al. Static respiratory cilia associated with mutations in Dnahc11/DNAH11: a mouse model of PCD. *Hum Mutat* 2012;33:495–503.
- Lyons RA, Saridogan E, Djahanbakhch O. The reproductive significance of human Fallopian tube cilia. *Hum Reprod Update* 2006;12:363–372.
- Maiti AK, Mattéi MG, Jorissen M, Volz A, Zeigler A, Bouvagnet P. Identification, tissue specific expression, and chromosomal localisation of several human dynein heavy chain genes. *Eur J Hum Genet* 2000;8:923–932.
- Mannowetz N, Kartarius S, Wennemuth G, Montenarh M. Protein kinase CK2 and new binding partners during spermatogenesis. *Cell Mol Life Sci* 2010;67:3905–3913.
- Mateu E, Rodrigo L, Prados N, Gil-Salom M, Remohí J, Pellicer A, Rubio C. High incidence of chromosomal abnormalities in large-headed and multiple-tailed spermatozoa. *J Androl* 2006;27:6–10.
- Matzuk MM, Lamb DJ. Genetic dissection of mammalian fertility pathways. *Nat Cell Biol* 2002;4(Suppl):s41–s49.
- Matzuk MM, Lamb DJ. The biology of infertility: research advances and clinical challenges. *Nat Med* 2008;14:1197–1213.
- Mazor M, Alkrinawi S, Chalifa-Caspi V, Manor E, Sheffield VC, Aviram M, Parvari R. Primary ciliary dyskinesia caused by homozygous mutation in DNAL1, encoding dynein light chain 1. *Am J Hum Genet* 2011;88:599–607.
- Merveille A-C, Davis EE, Becker-Heck A, Legendre M, Amirav I, Bataille G, Belmont J, Beydon N, Billen F, Clément A et al. CCDC39 is required for assembly of inner dynein arms and the dynein regulatory complex and for normal ciliary motility in humans and dogs. *Nat Genet* 2011;43:72–78.
- Miki K, Willis WD, Brown PR, Goulding EH, Fulcher KD, Eddy EM. Targeted disruption of the Akap4 gene causes defects in sperm flagellum and motility. *Dev Biol* 2002;248:331–342.
- Mitchell V, Rives N, Albert M, Peers M-C, Selva J, Clavier B, Escudier E, Escalier D. Outcome of ICSI with ejaculated spermatozoa in a series of men with distinct ultrastructural flagellar abnormalities. *Hum Reprod* 2006;21:2065–2074.
- Mitchison HM, Schmidts M, Loges NT, Freshour J, Dritsoula A, Hirst RA, O'Callaghan C, Blau H, Al Dabbagh M, Olbrich H et al. Mutations in axonemal dynein assembly factor DNAAF3 cause primary ciliary dyskinesia. *Nat Genet* 2012;44:381–389, SI–2.
- Molinari E, Mirabelli M, Raimondo S, Brussino A, Gennarelli G, Bongianni F, Revelli A. Sperm macrocephaly syndrome in a patient without AURKC mutations and with a history of recurrent miscarriage. *Reprod Biomed Online* 2013;26:148–156.
- Moore A, Escudier E, Roger G, Tamalet A, Pelosse B, Marlin S, Clément A, Geremek M, Delaisi B, Bridoux A-M et al. RPGR is mutated in patients with a complex X linked phenotype combining primary ciliary dyskinesia and retinitis pigmentosa. *J Med Genet* 2006;43:326–333.
- Moore DJ, Onoufriadis A, Shoemark A, Simpson MA, zur Lage PI, de Castro SC, Bartoloni L, Gallone G, Petridi S, Woollard WJ et al. Mutations in ZMYND10, a gene essential for proper axonemal assembly of inner and outer dynein arms in humans and flies, cause primary ciliary dyskinesia. *Am J Hum Genet* 2013;93:346–356.
- Moretti E, Geminiani M, Terzuoli G, Renieri T, Pasquarelli N, Collodel G. Two cases of sperm immotility: a mosaic of flagellar alterations related to dysplasia of the fibrous sheath and abnormalities of head-neck attachment. *Fertil Steril* 2011;95:1787.e19–23.
- Munro NC, Currie DC, Lindsay KS, Ryder TA, Rutman A, Dewar A, Greenstone MA, Hendry WF, Cole PJ. Fertility in men with primary ciliary dyskinesia presenting with respiratory infection. *Thorax* 1994;49:684–687.
- Musacchio A, Hardwick KG. The spindle checkpoint: structural insights into dynamic signalling. *Nat Rev Mol Cell Biol* 2002;3:731–741.
- Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS, McVean G, Donnelly P. Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. *Science* 2010;327:876–879.
- Nakamura N, Dai Q, Williams J, Goulding EH, Willis WD, Brown PR, Eddy EM. Disruption of a spermatogenic cell-specific mouse enolase 4 (eno4) gene causes sperm structural defects and male infertility. *Biol Reprod* 2013;88:90.
- Neesen J, Kirschner R, Ochs M, Schmidts M, Habermann B, Mueller C, Holstein AF, Nuesslein T, Adham I, Engel W. Disruption of an inner arm dynein heavy chain gene results in asthenozoospermia and reduced ciliary beat frequency. *Hum Mol Genet* 2001;10:1117–1128.
- Neugebauer DC, Neuwinger J, Jockenhövel F, Nieschlag E. '9 + 0' axoneme in spermatozoa and some nasal cilia of a patient with totally immotile spermatozoa associated with thickened sheath and short midpiece. *Hum Reprod* 1990;5:981–986.
- Nieschlag E, Behre HM, Nieschlag S. *Andrology: Male Reproductive Health and Dysfunction*. Springer Science & Business Media, 2010.
- Nistal M, Paniagua R, Herrero A. Multi-tailed spermatozoa in a case with asthenospermia and teratospermia. *Virchows Arch B Cell Pathol* 1977;26:111–118.
- Noveski P, Madjunkova S, Maleva I, Sotirovska V, Petanovska Z, Plaseska-Karanfilkska D. A homozygous deletion of the DPY19L gene is a cause of globozoospermia in men from the Republic of Macedonia. *Balk J Med Genet* 2013;16:73–76.
- O'Callaghan C, Chetcuti P, Moya E. High prevalence of primary ciliary dyskinesia in a British Asian population. *Arch Dis Child* 2010;95:51–52.
- Ogawa T, Aréchaga JM, Avarbock MR, Brinster RL. Transplantation of testis germinal cells into mouse seminiferous tubules. *Int J Dev Biol* 1997;41:111–122.
- Olbrich H, Häffner K, Kispert A, Völkel A, Volz A, Sasmaz G, Reinhardt R, Hennig S, Lehrach H, Konietzko N et al. Mutations in DNAH5 cause primary ciliary dyskinesia and randomization of left-right asymmetry. *Nat Genet* 2002;30:143–144.
- Olbrich H, Schmidts M, Werner C, Onoufriadis A, Loges NT, Raidt J, Banki NF, Shoemark A, Burgoyne T, Al Turki S et al. Recessive HYDIN mutations cause primary ciliary dyskinesia without randomization of left-right body asymmetry. *Am J Hum Genet* 2012;91:672–684.
- Olmedo SB, Nodar F, Chillik C, Chemes HE. Successful intracytoplasmic sperm injection with spermatozoa from a patient with dysplasia of the fibrous sheath and chronic respiratory disease. *Hum Reprod* 1997;12:1497–1499.
- Olmedo SB, Rawe VY, Nodar FN, Galaverna GD, Acosta AA, Chemes HE. Pregnancies established through intracytoplasmic sperm injection (ICSI) using spermatozoa with dysplasia of fibrous sheath. *Asian J Androl* 2000;2:125–130.
- Ombelet W, Menkveld R, Kruger TF, Steeno O. Sperm morphology assessment: historical review in relation to fertility. *Hum Reprod Update* 1995;1:543–557.
- Omran H, Kobayashi D, Olbrich H, Tsukahara T, Loges NT, Hagiwara H, Zhang Q, Leblond G, O'Toole E, Hara C et al. Ktu/PFI13 is required for cytoplasmic pre-assembly of axonemal dyneins. *Nature* 2008;456:611–616.
- Onoufriadis A, Paff T, Antony D, Shoemark A, Micha D, Kuyt B, Schmidts M, Petridi S, Dankert-Roelse JE, Haarman EG et al. Splice-site mutations in the axonemal outer

- dynein arm docking complex gene CCDC114 cause primary ciliary dyskinesia. *Am J Hum Genet* 2013; **92**:88–98.
- Onoufriadis A, Shoemark A, Munye MM, James CT, Schmidts M, Patel M, Rosser EM, Bacchelli C, Beales PL, Scambler PJ et al. Combined exome and whole-genome sequencing identifies mutations in ARMC4 as a cause of primary ciliary dyskinesia with defects in the outer dynein arm. *J Med Genet* 2014a; **51**:61–67.
- Onoufriadis A, Shoemark A, Schmidts M, Patel M, Jimenez G, Liu H, Thomas B, Dixon M, Hirst RA, Rutman A et al. Targeted NGS gene panel identifies mutations in RSPH1 causing primary ciliary dyskinesia and a common mechanism for ciliary central pair agenesis due to radial spoke defects. *Hum Mol Genet* 2014b; **23**:3362–3374.
- Ostrowski LE, Dutcher SK, Lo CW. Cilia and models for studying structure and function. *Proc Am Thorac Soc* 2011; **8**:423–429.
- Unis L, Zoghmar A, Coutton C, Rouabah L, Hachemi M, Martinez D, Martinez G, Bellil I, Khelfi D, Arnoult C et al. Mutations of the aurora kinase C gene causing macrozoospermia are the most frequent genetic cause of male infertility in Algerian men. *Asian J Androl* 2015; **17**:68–73.
- Paiardi C, Pasini ME, Gioria M, Berruti G. Failure of acrosome formation and globozoospermia in the wobbler mouse, a Vps54 spontaneous recessive mutant. *Spermatogenesis* 2011; **1**:52–62.
- Palermo GD, Neri QV, Takeuchi T, Rosenwaks Z. ICSI: where we have been and where we are going. *Semin Reprod Med* 2009; **27**:191–201.
- Panizzi JR, Becker-Heck A, Castleman VH, Al-Mutairi DA, Liu Y, Loges NT, Pathak N, Austin-Tse C, Sheridan E, Schmidts M et al. CCDC103 mutations cause primary ciliary dyskinesia by disrupting assembly of ciliary dynein arms. *Nat Genet* 2012; **44**:714–719.
- Parvanov ED, Petkov PM, Paigen K. Prdm9 controls activation of mammalian recombination hotspots. *Science* 2010; **327**:835.
- Pennarun G, Escudier E, Chapelin C, Bridoux AM, Cacheux V, Roger G, Clément A, Goossens M, Amselem S, Duriez B. Loss-of-function mutations in a human gene related to *Chlamydomonas reinhardtii* dynein IC78 result in primary ciliary dyskinesia. *Am J Hum Genet* 1999; **65**:1508–1519.
- Perrin A, Morel F, Moy L, Colleu D, Amice V, De Braekeleer M. Study of aneuploidy in large-headed, multiple-tailed spermatozoa: case report and review of the literature. *Fertil Steril* 2008; **90**:1201.e13–1201.e17.
- Perrin A, Coat C, Nguyen MH, Talagas M, Morel F, Amice J, De Braekeleer M. Molecular cytogenetic and genetic aspects of globozoospermia: a review. *Andrologia* 2013; **45**:1–9.
- Pierre V, Martinez G, Coutton C, Delaroche J, Yassine S, Novella C, Pernet-Gallay K, Hennebicq S, Ray PF, Arnoult C. Absence of Dpy19l2, a new inner nuclear membrane protein, causes globozoospermia in mice by preventing the anchoring of the acrosome to the nucleus. *Development* 2012; **139**:2955–2965.
- Pieters MH, Speed RM, de Boer P, Vreeburg JT, Dohle G, In't Veld PA. Evidence of disturbed meiosis in a man referred for intracytoplasmic sperm injection. *Lancet* 1998; **351**:957.
- Pigino G, Bui KH, Maheshwari A, Lupetti P, Diener D, Ishikawa T. Cryoelectron tomography of radial spokes in cilia and flagella. *J Cell Biol* 2011; **195**:673–687.
- Pilder SH, Olds-Clarke P, Orth JM, Jester WF, Dugan L. Hst7: a male sterility mutation perturbing sperm motility, flagellar assembly, and mitochondrial sheath differentiation. *J Androl* 1997; **18**:663–671.
- Pirrello O, Machev N, Schmidt F, Terriou P, Ménézo Y, Viville S. Search for mutations involved in human globozoospermia. *Hum Reprod* 2005; **20**:1314–1318.
- Poongothai J, Gopenath TS, Manonayaki S. Genetics of human male infertility. *Singapore Med J* 2009; **50**:336–347.
- Popli K, Stewart J. Infertility and its management in men with cystic fibrosis: review of literature and clinical practices in the UK. *Hum Fertil (Camb)* 2007; **10**:217–221.
- Price DM, Kanyo R, Steinberg N, Chik CL, Ho AK. Nocturnal activation of aurora C in rat pineal gland: its role in the norepinephrine-induced phosphorylation of histone H3 and gene expression. *Endocrinology* 2009; **150**:2334–2341.
- Rawe VY, Galaverna GD, Acosta AA, Olmedo SB, Chemes HE. Incidence of tail structure distortions associated with dysplasia of the fibrous sheath in human spermatozoa. *Hum Reprod* 2001; **16**:879–886.
- Rawe VY, Terada Y, Nakamura S, Chilik CF, Olmedo SB, Chemes HE. A pathology of the sperm centriole responsible for defective sperm aster formation, syngamy and cleavage. *Hum Reprod* 2002; **17**:2344–2349.
- Rives N, Mousset-Simeon N, Mazurier S, Mace B. Primary flagellar abnormality is associated with an increased rate of spermatozoa aneuploidy. *J Androl* 2005; **26**:61–69.
- Rybouchkin AV, Van der Straeten F, Quatacker J, De Sutter P, Dhont M. Fertilization and pregnancy after assisted oocyte activation and intracytoplasmic sperm injection in a case of round-headed sperm associated with deficient oocyte activation capacity. *Fertil Steril* 1997; **68**:1144–1147.
- Salzberg Y, Eldar T, Karminsky O-D, Itach SB-S, Pietrovski S, Don J. Meigl deficiency causes a severe defect in mouse spermatogenesis. *Dev Biol* 2010; **338**:158–167.
- Sapiro R, Tarantino LM, Velazquez F, Kiriakidou M, Hecht NB, Bucan M, Strauss JF. Sperm antigen 6 is the murine homologue of the *Chlamydomonas reinhardtii* central apparatus protein encoded by the PF16 locus. *Biol Reprod* 2000; **62**:511–518.
- Sapiro R, Kostetskii I, Olds-Clarke P, Gerton GL, Radice GL, Strauss III JF. Male infertility, impaired sperm motility, and hydrocephalus in mice deficient in sperm-associated antigen 6. *Mol Cell Biol* 2002; **22**:6298–6305.
- Sasaki K, Katayama H, Stenoien DL, Fujii S, Honda R, Kimura M, Okano Y, Tatsuka M, Suzuki F, Nigg EA et al. Aurora-C kinase is a novel chromosomal passenger protein that can complement Aurora-B kinase function in mitotic cells. *Cell Motil Cytoskeleton* 2004; **59**:249–263.
- Sathananthan AH. Functional competence of abnormal spermatozoa. *Baillière's Clin Obstet Gynaecol* 1994; **8**:141–156.
- Sathananthan AH, Kola I, Osborne J, Trounson A, Ng SC, Bongso A, Ratnam SS. Centrioles in the beginning of human development. *Proc Natl Acad Sci USA* 1991; **88**:4806–4810.
- Satir P, Christensen ST. Structure and function of mammalian cilia. *Histochem Cell Biol* 2008; **129**:687–693.
- Schatten H, Sun Q-Y. The functional significance of centrosomes in mammalian meiosis, fertilization, development, nuclear transfer, and stem cell differentiation. *Environ Mol Mutagen* 2009; **50**:620–636.
- Schindler K, Davydenko O, Fram B, Lampson MA, Schultz RM. Maternally recruited Aurora C kinase is more stable than Aurora B to support mouse oocyte maturation and early development. *Proc Natl Acad Sci USA* 2012; **109**:E2215–E2222.
- Schwabe GC, Hoffmann K, Loges NT, Birker D, Rossier C, de Santi MM, Olbrich H, Fliegauf M, Failey M, Liebers U et al. Primary ciliary dyskinesia associated with normal axoneme ultrastructure is caused by DNAH11 mutations. *Hum Mutat* 2008; **29**:289–298.
- Sen CGS, Holstein AF, Schirren C. über die Morphogenese rundköpfiger Spermatozoen des Menschen. *Andrologia* 1971; **3**:117–125.
- Sermondade N, Hafhouf E, Dupont C, Bechoua S, Palacios C, Eustache F, Poncelet C, Benacken B, Lévy R, Sifer C. Successful childbirth after intracytoplasmic morphologically selected sperm injection without assisted oocyte activation in a patient with globozoospermia. *Hum Reprod* 2011; **26**:2944–2949.
- Sha Y-W, Ding L, Li P. Management of primary ciliary dyskinesia/Kartagener's syndrome in infertile male patients and current progress in defining the underlying genetic mechanism. *Asian J Androl* 2014; **16**:101–106.
- Sharif B, Na J, Lykke-Hartmann K, McLaughlin SH, Laue E, Glover DM, Zernicka-Goetz M. The chromosome passenger complex is required for fidelity of chromosome transmission and cytokinesis in meiosis of mouse oocytes. *J Cell Sci* 2010; **123**:4292–4300.
- Singh G. Ultrastructural features of round-headed human spermatozoa. *Int J Fertil* 1992; **37**:99–102.
- Sironen A, Thomsen B, Andersson M, Ahola V, Vilkki J. An intronic insertion in KPL2 results in aberrant splicing and causes the immotile short-tail sperm defect in the pig. *Proc Natl Acad Sci USA* 2006; **103**:5006–5011.
- Sironen A, Hansen J, Thomsen B, Andersson M, Vilkki J, Toppari J, Kotaja N. Expression of SPEF2 during mouse spermatogenesis and identification of IFT20 as an interacting protein. *Biol Reprod* 2010; **82**:580–590.
- Sironen A, Kotaja N, Mulhern H, Wyatt TA, Sisson JH, Pavlik JA, Miiluniemi M, Fleming MD, Lee L. Loss of SPEF2 function in mice results in spermatogenesis defects and primary ciliary dyskinesia. *Biol Reprod* 2011; **85**:690–701.
- Smith EN, Bloss CS, Badner JA, Barrett T, Belmonte PL, Berrettini W, Byerley W, Coryell W, Craig D, Edenberg HJ et al. Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry* 2009; **14**:755–763.
- Stalf T, Sánchez R, Köhn FM, Schalles U, Kleinstein J, Hinz V, Tielsch J, Khanaga O, Turley H, Gips H. Pregnancy and birth after intracytoplasmic sperm injection with spermatozoa from a patient with tail stump syndrome. *Hum Reprod* 1995; **10**:2112–2114.
- Stankiewicz P, Lupski JR. Structural variation in the human genome and its role in disease. *Annu Rev Med* 2010; **61**:437–455.

- Sun F, Ko E, Martin RH. Is there a relationship between sperm chromosome abnormalities and sperm morphology? *Reprod Biol Endocrinol* 2006; **4**:1.
- Sunkel CE, Glover DM. Polo, a mitotic mutant of *Drosophila* displaying abnormal spindle poles. *J Cell Sci* 1988; **89**:25–38.
- Tang C-JC, Lin C-Y, Tang TK. Dynamic localization and functional implications of Aurora-C kinase during male mouse meiosis. *Dev Biol* 2006; **290**:398–410.
- Tarkar A, Loges NT, Slagle CE, Francis R, Dougherty GW, Tamayo JV, Shook B, Cantino M, Schwartz D, Jahnke C et al. DYX1C1 is required for axonemal dynein assembly and ciliary motility. *Nat Genet* 2013; **45**:995–1003.
- Taylor SL, Yoon SY, Morshed MS, Lacey DR, Jellerette T, Fissore RA, Oehninger S. Complete globozoospermia associated with PLC ζ deficiency treated with calcium ionophore and ICSI results in pregnancy. *Reprod Biomed Online* 2010; **20**:559–564.
- Tejera A, Mollá M, Muriel L, Remohí J, Pellicer A, De Pablo JL. Successful pregnancy and childbirth after intracytoplasmic sperm injection with calcium ionophore oocyte activation in a globozoospermic patient. *Fertil Steril* 2008; **90**:1202.e1–1202.e5.
- Terada Y, Nakamura S, Morita J, Tachibana M, Morito Y, Ito K, Murakami T, Yaegashi N, Okamura K. Use of mammalian eggs for assessment of human sperm function: molecular and cellular analyses of fertilization by intracytoplasmic sperm injection. *Am J Reprod Immunol* 2004; **51**:290–293.
- Teves ME, Jha KN, Song J, Nagarkatti-Gude DR, Herr JC, Foster JA, Strauss JF, Zhang Z. Germ cell-specific disruption of the MeigI gene causes impaired spermiogenesis in mice. *Andrology* 2013; **1**:37–46.
- Thonneau P, Marchand S, Tallec A, Ferial ML, Ducot B, Lansac J, Lopes P, Tabaste JM, Spira A. Incidence and main causes of infertility in a resident population (1,850,000) of three French regions (1988–1989). *Hum Reprod* 1991; **6**:811–816.
- Tseng TC, Chen SH, Hsu YP, Tang TK. Protein kinase profile of sperm and eggs: cloning and characterization of two novel testis-specific protein kinases (AIE1, AIE2) related to yeast and fly chromosome segregation regulators. *DNA Cell Biol* 1998; **17**:823–833.
- Tsou J-H, Chang K-C, Chang-Liao P-Y, Yang S-T, Lee C-T, Chen Y-P, Lee Y-C, Lin B-W, Lee J-C, Shen M-R et al. Aberrantly expressed AURKC enhances the transformation and tumourigenicity of epithelial cells. *J Pathol* 2011; **225**:243–254.
- Turner RM, Musse MP, Mandal A, Klotz K, Jayes FC, Herr JC, Gerton GL, Moss SB, Chemes HE. Molecular genetic analysis of two human sperm fibrous sheath proteins, AKAP4 and AKAP3, in men with dysplasia of the fibrous sheath. *J Androl* 2001; **22**:302–315.
- Turner DJ, Miretti M, Rajan D, Fiegler H, Carter NP, Blayney ML, Beck S, Hurles ME. Germline rates of de novo meiotic deletions and duplications causing several genomic disorders. *Nat Genet* 2008; **40**:90–95.
- Tüttelmann F, Simoni M, Kliesch S, Ledig S, Dworniczak B, Wieacker P, Röpke A. Copy number variants in patients with severe oligozoospermia and Sertoli-cell-only syndrome. *PLoS One* 2011a; **6**:e19426.
- Tüttelmann F, Werny F, Cooper TG, Kliesch S, Simoni M, Nieschlag E. Clinical experience with azoospermia: aetiology and chances for spermatozoa detection upon biopsy. *Int J Androl* 2011b; **34**:291–298.
- Vader G, Maia AF, Lens SM. The chromosomal passenger complex and the spindle assembly checkpoint: kinetochore-microtubule error correction and beyond. *Cell Div* 2008; **3**:10.
- Vagnarelli P, Earnshaw WC. Chromosomal passengers: the four-dimensional regulation of mitotic events. *Chromosoma* 2004; **113**:211–222.
- Van Blerkom J. Sperm centrosome dysfunction: a possible new class of male factor infertility in the human. *Mol Hum Reprod* 1996; **2**:349–354.
- Van der Horst A, Lens SMA. Cell division: control of the chromosomal passenger complex in time and space. *Chromosoma* 2014; **123**:25–42.
- Van der Waal MS, Hengeveld RCC, van der Horst A, Lens SMA. Cell division control by the Chromosomal Passenger Complex. *Exp Cell Res* 2012; **318**:1407–1420.
- Vernon GG, Neesen J, Woolley DM. Further studies on knockout mice lacking a functional dynein heavy chain (MDHC7). I. Evidence for a structural deficit in the axoneme. *Cell Motil Cytoskeleton* 2005; **61**:65–73.
- Vicari E, de Palma A, Burrello N, Longo G, Grazioso C, Barone N, Zahi M, D'Agata R, Calogero AE. Absolute polymorphic teratozoospermia in patients with oligoasthenozoospermia is associated with an elevated sperm aneuploidy rate. *J Androl* 2003; **24**:598–603.
- Vincensini L, Blisnick T, Bastin P. I001 model organisms to study cilia and flagella. *Biol Cell* 2011; **103**:109–130.
- Viville S, Mollard R, Bach ML, Falquet C, Gerlinger P, Warter S. Do morphological anomalies reflect chromosomal aneuploidies?: case report. *Hum Reprod* 2000; **15**:2563–2566.
- Vogel P, Hansen G, Fontenot G, Read R. Tubulin tyrosine ligase-like 1 deficiency results in chronic rhinosinusitis and abnormal development of spermatid flagella in mice. *Vet Pathol* 2010; **47**:703–712.
- Vogt PH, Falcao CL, Hanstein R, Zimmer J. The AZF proteins. *Int J Androl* 2008; **31**:383–394.
- Walden CM, Sandhoff R, Chuang C-C, Yildiz Y, Butters TD, Dwek RA, Platt FM, van der Spoel AC. Accumulation of glucosylceramide in murine testis, caused by inhibition of beta-glucosidase 2: implications for spermatogenesis. *J Biol Chem* 2007; **282**:32655–32664.
- Wang H, Wan H, Li X, Liu W, Chen Q, Wang Y, Yang L, Tang H, Zhang X, Duan E et al. Atg7 is required for acrosome biogenesis during spermatogenesis in mice. *Cell Res* 2014; **24**:852–869.
- Wargo MJ, Smith EF. Asymmetry of the central apparatus defines the location of active microtubule sliding in *Chlamydomonas* flagella. *Proc Natl Acad Sci USA* 2003; **100**:137–142.
- Watanabe K, Takebayashi H, Bepari AK, Esumi S, Yanagawa Y, Tamamaki N. Dpy19II, a multi-transmembrane protein, regulates the radial migration of glutamatergic neurons in the developing cerebral cortex. *Development* 2011; **138**:4979–4990.
- Weissenberg R, Aviram A, Golani R, Lewin LM, Levron J, Madgar I, Dor J, Barkai G, Goldman B. Concurrent use of flow cytometry and fluorescence in-situ hybridization techniques for detecting faulty meiosis in a human sperm sample. *Mol Hum Reprod* 1998; **4**:61–66.
- Wickstead B, Gull K. Dyneins across eukaryotes: a comparative genomic analysis. *Traffic* 2007; **8**:1708–1721.
- Wirschl M, Olbrich H, Werner C, Tritschler D, Bower R, Sale VS, Loges NT, Pennekamp P, Lindberg S, Stenram U et al. The nexin-dynein regulatory complex subunit DRC1 is essential for motile cilia function in algae and humans. *Nat Genet* 2013; **45**:262–268.
- World Health Organization. *WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction*, 1st edn. Singapore: Press Concern, 1980.
- World Health Organization. *WHO Laboratory Manual for the Examination and Processing of Human Semen*, 5th edn. Cambridge: Cambridge University Press, 2010, 272 p.
- Xiao N, Kam C, Shen C, Jin W, Wang J, Lee KM, Jiang L, Xia J. PICK1 deficiency causes male infertility in mice by disrupting acrosome formation. *J Clin Invest* 2009; **119**:802–812.
- Xu X, Toselli PA, Russell LD, Seldin DC. Globozoospermia in mice lacking the casein kinase II alpha' catalytic subunit. *Nat Genet* 1999; **23**:118–121.
- Xu M, Xiao J, Chen J, Li J, Yin L, Zhu H, Zhou Z, Sha J. Identification and characterization of a novel human testis-specific Golgi protein, NYD-SP12. *Mol Hum Reprod* 2003; **9**:9–17.
- Yan W. Male infertility caused by spermiogenic defects: lessons from gene knockouts. *Mol Cell Endocrinol* 2009; **306**:24–32.
- Yan X, Cao L, Li Q, Wu Y, Zhang H, Saiyin H, Liu X, Zhang X, Shi Q, Yu L. Aurora C is directly associated with Survivin and required for cytokinesis. *Genes Cells* 2005; **10**:617–626.
- Yang K-T, Li S-K, Chang C-C, Tang C-JC, Lin Y-N, Lee S-C, Tang TK. Aurora-C kinase deficiency causes cytokinesis failure in meiosis I and production of large polyploid oocytes in mice. *Mol Biol Cell* 2010; **21**:2371–2383.
- Yao R, Ito C, Natsume Y, Sugitani Y, Yamanaka H, Kuretake S, Yanagida K, Sato A, Toshimori K, Noda T. Lack of acrosome formation in mice lacking a Golgi protein, GOPC. *Proc Natl Acad Sci USA* 2002; **99**:11211–11216.
- Yassine S, Escoffier J, Martinez G, Coutton C, Karaouzène T, Zouari R, Ravanat J-L, Metzler-Guillemain C, Fissore R, Hennebicq S et al. Dpy19I2-deficient globozoospermic sperm display altered genome packaging and DNA damage that compromises the initiation of embryo development. *Mol Hum Reprod* 2015a; **21**:169–185.
- Yassine S, Escoffier J, Nahed RA, Pierre V, Karaouzene T, Ray PF, Arnoult C. Dynamics of Sun5 localization during spermatogenesis in wild type and Dpy19I2 knock-out mice indicates that Sun5 is not involved in acrosome attachment to the nuclear envelope. *PLoS One* 2015b; **10**:e0118698.
- Yatsenko AN, O'Neil DS, Roy A, Arias-Mendoza PA, Chen R, Murthy LJ, Lamb DJ, Matzuk MM. Association of mutations in the zona pellucida binding protein 1 (ZPBPI) gene with abnormal sperm head morphology in infertile men. *Mol Hum Reprod* 2012; **18**:14–21.
- Yildiz Y, Matern H, Thompson B, Allegood JC, Warren RL, Ramirez DMO, Hammer RE, Hamra FK, Matern S, Russell DW. Mutation of beta-glucosidase 2

- causes glycolipid storage disease and impaired male fertility. *J Clin Invest* 2006; **116**:2985–2994.
- Yoon S-Y, Jellerette T, Salicioni AM, Lee HC, Yoo M-S, Coward K, Parrington J, Grow D, Cibelli JB, Visconti PE et al. Human sperm devoid of PLC ζ fail to induce Ca(2+) release and are unable to initiate the first step of embryo development. *J Clin Invest* 2008; **118**:3671–3681.
- Yu Y, Fuscoe JC, Zhao C, Guo C, Jia M, Qing T, Bannon DL, Lancashire L, Bao W, Du T et al. A rat RNA-Seq transcriptomic BodyMap across 11 organs and 4 developmental stages. *Nat Commun* 2014; **5**:3230.
- Yurov YB, Saias MJ, Vorsanova SG, Erny R, Soloviev IV, Sharonin VO, Guichaoua MR, Luciani JM. Rapid chromosomal analysis of germ-line cells by FISH: an investigation of an infertile male with large-headed spermatozoa. *Mol Hum Reprod* 1996; **2**:665–668.
- Zariwala MA, Knowles MR, Leigh MW. Primary ciliary dyskinesia. In: Pagon RA, Adam MP, Ardinger HH, Bird TD, Dolan CR, Fong C-T, Smith RJ, Stephens K (eds). GeneReviews®. Seattle, WA: University of Washington, Seattle, 1993. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1122/>. Accessed August 20, 2014.
- Zariwala MA, Noone PG, Sannuti A, Minnix S, Zhou Z, Leigh MW, Hazucha M, Carson JL, Knowles MR. Germline mutations in an intermediate chain dynein cause primary ciliary dyskinesia. *Am J Respir Cell Mol Biol* 2001; **25**:577–583.
- Zariwala MA, Gee HY, Kurkowiak M, Al-Mutairi DA, Leigh MW, Hurd TW, Hjeij R, Dell SD, Chaki M, Dougherty GW et al. ZMYND10 is mutated in primary ciliary dyskinesia and interacts with LRRC6. *Am J Hum Genet* 2013; **93**:336–345.
- Zekri A, Lesan V, Ghaffari SH, Tabrizi MH, Modarressi MH. Gene amplification and overexpression of Aurora-C in breast and prostate cancer cell lines. *Oncol Res* 2012; **20**:241–250.
- Zhang Z, Kostetskii I, Moss SB, Jones BH, Ho C, Wang H, Kishida T, Gerton GL, Radice GL, Strauss JF. Haploinsufficiency for the murine orthologue of Chlamydomonas PF20 disrupts spermatogenesis. *Proc Natl Acad Sci USA* 2004; **101**:12946–12951.
- Zhang Z, Kostetskii I, Tang W, Haig-Ladewig L, Sapiro R, Wei Z, Patel AM, Bennett J, Gerton GL, Moss SB et al. Deficiency of SPAG16L causes male infertility associated with impaired sperm motility. *Biol Reprod* 2006; **74**:751–759.
- Zhang F, Gu W, Hurles ME, Lupski JR. Copy number variation in human health, disease, and evolution. *Annu Rev Genomics Hum Genet* 2009; **10**:451–481.
- Zhou J, Yao J, Joshi HC. Attachment and tension in the spindle assembly checkpoint. *J Cell Sci* 2002; **115**:3547–3555.
- Zhu F, Gong F, Lin G, Lu G. DPY19L2 gene mutations are a major cause of globozoospermia: identification of three novel point mutations. *Mol Hum Reprod* 2013; **19**:395–404.

Conclusion et Perspectives

Comme cela a été démontré de ce travail, la macrozoospermie et la globozoospermie sont des illustrations parfaites de maladies monogéniques avec respectivement, pour les formes complètes, une contribution largement majoritaire des gènes *AURKC* et *DPY19L2* (Dieterich *et al.*, 2007, 2009 ; Harbuz *et al.*, 2011). Cependant, ces exemples de phénotypes d'infertilité masculine « monogénique » ne signifient pas qu'un seul gène soit la cause de chaque phénotype. C'est le cas des anomalies flagellaires pour lesquelles l'implication de *DNAH1* n'a été retrouvée « que » dans près de 30% des cas étudiés (Ben Khelifa *et al.*, 2014). Cela suggère évidemment que d'autres gènes peuvent être impliqués dans un phénotype « *a priori* » identique. Nous avons donc décidé de poursuivre les investigations de nouveaux gènes par WES à partir de cohortes de plusieurs dizaines de patients (apparentés ou non) présentant des phénotypes identiques de térazoospermie et chez lesquels les investigations dans les gènes déjà identifiés (*AURKC*, *DPY19L2* et *DNAH1*) se sont révélées négatives. Les résultats sont en cours d'analyse, mais les données préliminaires sont extrêmement prometteuses. Nous prévoyons également d'étendre ces études à d'autres phénotypes d'infertilité comme l'azoospermie ou encore l'infertilité féminine.

Actuellement la principale difficulté n'est plus l'identification des gènes mais la validation des mutations identifiées. Les modèles cellulaires, *a priori* les plus simples d'utilisation, sont difficilement applicables dans le cadre de l'infertilité masculine, les cellules spermatogéniques étant difficilement cultivables. Les modèles animaux semblent finalement les plus adaptés à l'étude des gènes de la spermatogénèse et en particulier ceux impliqués dans les anomalies flagellaires. Les flagelles et les cils sont remarquablement bien conservés au cours de l'évolution. Différents modèles expérimentaux ont ainsi été utilisés pour améliorer les connaissances sur ces organelles (Ostrowski *et al.*, 2011). On peut citer par exemple l'utilisation possible de certains protistes flagellés comme *Chlamydomonas reinhardtii*, *Paramecium*, *Tetrahymena* and *Trypanosoma* ou *Leishmania* (Vincensini *et al.*, 2011). *Chlamydomonas* est en particulier largement utilisé dans la validation des gènes impliqués dans les ciliopathies comme les DCP (Ostrowski *et al.*, 2011). D'autres organismes multicellulaires ont également été utilisés efficacement pour l'étude des gènes de ciliopathies comme le Zebrafish, *Xenopus*, *Caenorhabditis elegans* ou *Drosophila*. Ainsi certain de ces modèles pourraient être également efficacement employés pour l'étude de protéines du

flagelle spermatique (Inaba, 2011). Certains modèles mammifères et en particulier la souris ont été également largement utilisés pour l'étude des gènes impliqués dans les désordres de la spermatogénèse (de Boer *et al.*, 2014). La difficulté reste cependant la complexité technique pour obtenir des lignées stables d'animaux mutants. Récemment, la découverte du système CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) chez *Streptococcus pyogenes*, agissant comme mécanisme de défense contre les virus et l'insertion d'ADN étranger, a révolutionné les pratiques pour la création d'animaux modèles génétiquement modifiés (Hsu *et al.*, 2014). Ce système utilise une endonucléase bactérienne, Cas9 qui, complexée à un ARN guide non codant (ARNg) complémentaire de la séquence-cible, clive l'ADN de façon spécifique en amont d'une séquence PAM (Protospacer Adaptor Motif) localisée sur l'ADN génomique (Gaj *et al.*, 2013) (figure 30). Par cette technique, il est désormais possible d'obtenir plus rapidement et « simplement » des souris knock-out ou knock-in pour les différents gènes à valider. Il faut toutefois garder à l'esprit que le modèle murin n'est pas un modèle parfait et que malgré la ressemblance entre la spermatogénèse murine et humaine, elle n'en demeure pas moins différente. Un exemple probant de cette différence pourrait être le modèle KO de *DNAH1* qui présente un phénotype d'asthénozoospermie stricte bien différent de la tératozoospermie sévère observée chez l'homme (Neesen *et al.*, 2001 ; Ben Khelifa *et al.*, 2014).

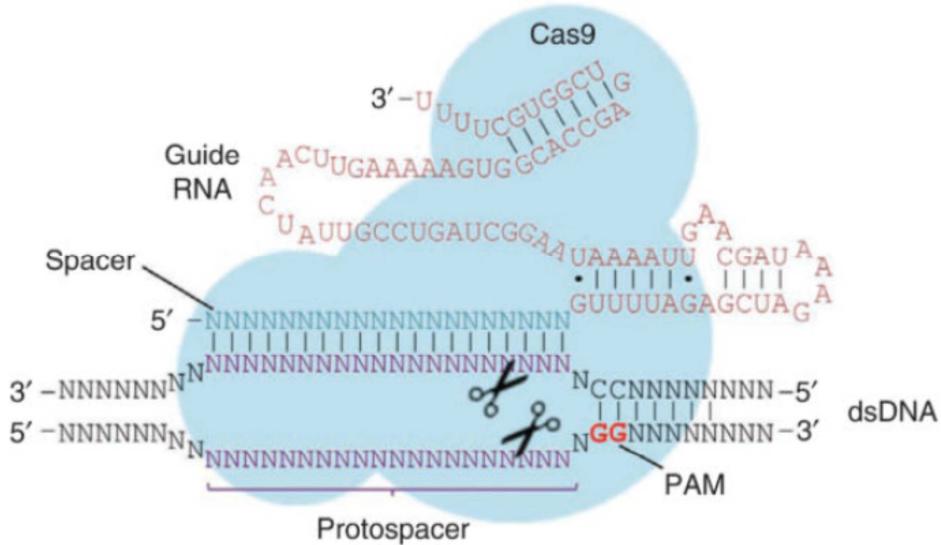


Figure 30. Principe général du système CRISPR/Cas9. D'après (Mali *et al.*, 2013).

Ainsi, dans une volonté de valider rapidement nos mutations et gènes candidats, nous sommes en train de mettre en place un schéma de validation systématique de nos gènes sur des

modèles murins reposant sur la technologie CRISPR/Cas9 (collaboration Pr Serge Nef, Department of Genetic Medicine and Development, Geneva, Switzerland) ainsi que sur des modèles trypanosomes (collaboration Dr Mélanie Bonhivers, Laboratoire de Microbiologie Fondamentale et Pathogénicité UMR-CNRS 5234, Bordeaux, France). Notre stratégie d'identification et de caractérisation des gènes de l'infertilité masculine est résumée figure 31.

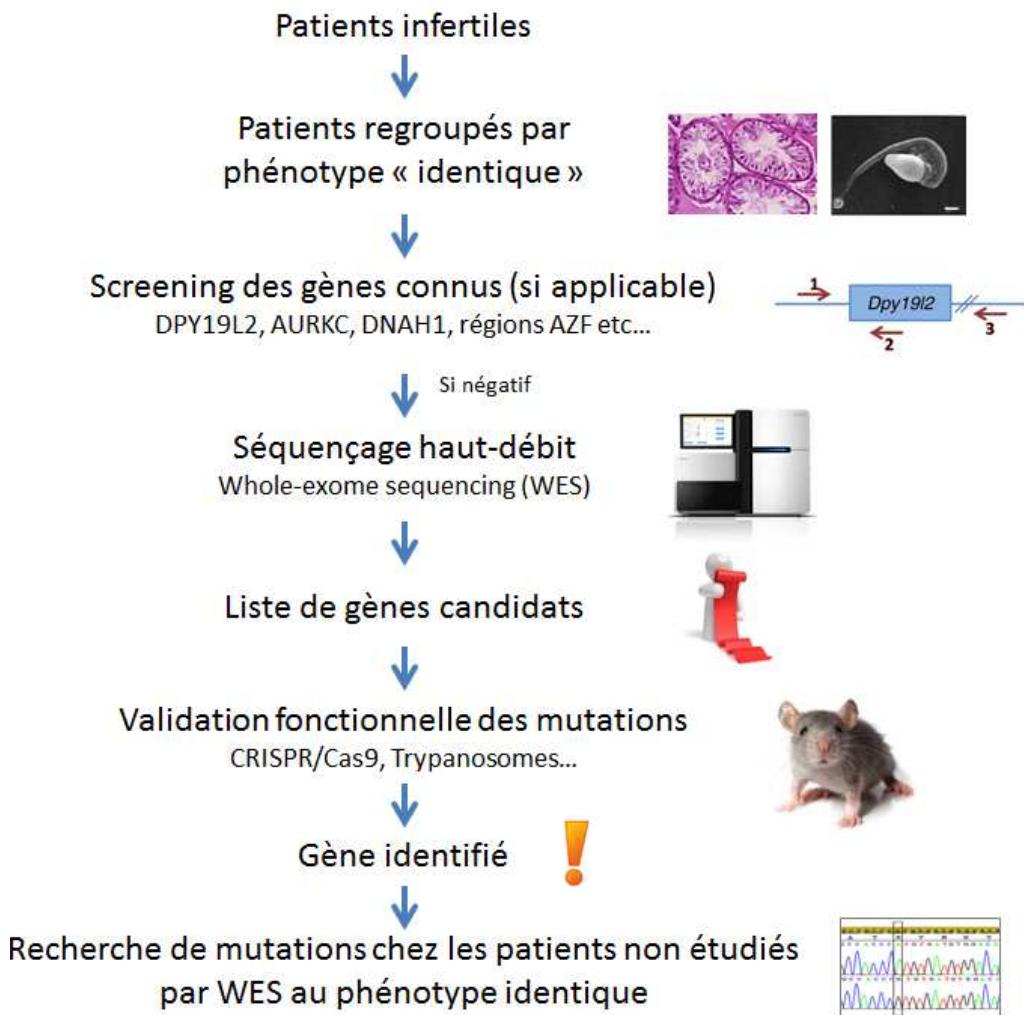


Figure 31. Stratégie générale d'identification et de caractérisation des gènes de l'infertilité masculine appliquée dans notre laboratoire.

L'objectif final de l'identification de nouveaux gènes impliqués dans l'infertilité est l'amélioration de la prise en charge des patients (figure 32). Tout d'abord, ces connaissances permettront de préciser le diagnostic ainsi que le conseil génétique et l'information donnée au patient sur les possibilités de sa prise en charge. En cas de mauvais pronostic, comme par

exemple dans le cas de patients avec une macrozoospermie et mutés pour *AURKC*, les solutions d'AMP sont contre-indiquées en raison du risque chromosomal et des alternatives comme le recours au don ou à l'adoption doivent être obligatoirement privilégiées. Dans le cas contraire, des solutions d'AMP seront alors envisageables. De même, la cause génétique identifiée peut permettre de préciser ou de rassurer quant à un éventuel risque chez l'enfant à naître de survenue de maladies (chromosomiques, géniques, épigénétiques) directement liées au défaut génétique « véhiculé » par le sperme déficient. Enfin, au-delà de la prise en charge avec les méthodes d'AMP actuellement disponibles, la connaissance du gène en cause permettrait d'envisager un traitement personnalisé ciblant le défaut moléculaire.

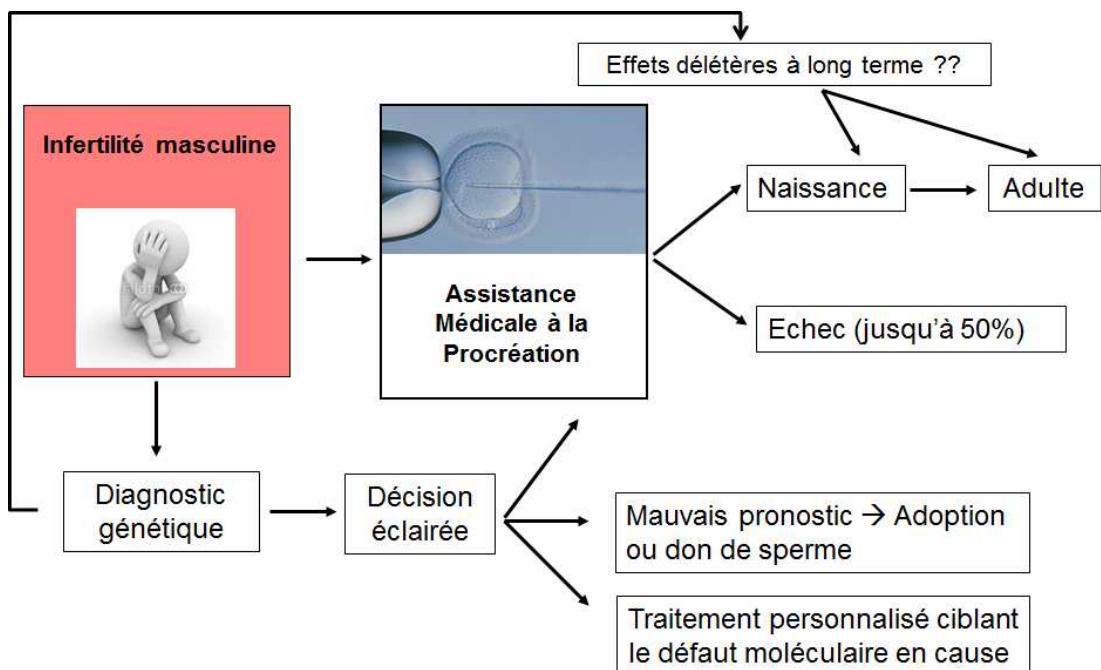


Figure 32. Importance du diagnostic génétique dans la prise en charge du patient infertile.

L'infertilité masculine est un excellent modèle pour le développement de thérapies correctrices innovantes : car elle touche un seul organe, un traitement court peut être envisagé uniquement basé sur le temps de la spermatogénèse (74 jours chez l'homme), enfin l'effet du traitement est facilement mesurable (spermogramme). Ces avantages contrastent nettement avec d'autres maladies génétiques multi-systémiques et pour lesquelles la thérapie devra être conduite durant toute la vie de l'individu atteint. La principale limite de la thérapie de l'infertilité masculine sera de trouver des molécules capables de cibler spécifiquement le

testicule et de diffuser à travers la barrière-hémato testiculaire à des concentrations suffisantes pour corriger le déficit. Une alternative pourrait être de réaliser des injections intra-testiculaires (Ogawa *et al.*, 1997) mais cela semble cependant difficilement réalisable et répétable même sur une courte période. Des stratégies par thérapie génique pourraient être aussi une alternative (figure 33). Elles requièrent l'utilisation de différents vecteurs capables de diffuser dans l'organe cible afin d'introduire une copie « normale » du gène muté en cause dans la pathologie, ou bien d'en modifier l'expression (Kaufmann *et al.*, 2013). Les vecteurs principalement utilisés sont les vecteurs viraux (ex : rétrovirus, adénovirus). L'utilisation de ces vecteurs viraux pose toutefois la question d'un potentiel risque tumoral dû à la possible intégration aléatoire du génome viral dans le génome de la cellule cible ou dans le contexte de l'infertilité, de la modification du génome des cellules germinales avec le risque que cela comporte pour la descendance. L'utilisation des vecteurs viraux non-intégratifs (c'est-à-dire ne s'intégrant pas dans le génome de la cellule hôte comme par exemple les AAVs (adeno-associated virus)) ou de vecteurs non-viraux (ex: liposomes, ADN nu) apparaissent comme des solutions alternatives bien que leur efficacité et/ou leur innocuité à long terme nécessitent d'être encore évaluées (Misra, 2013).

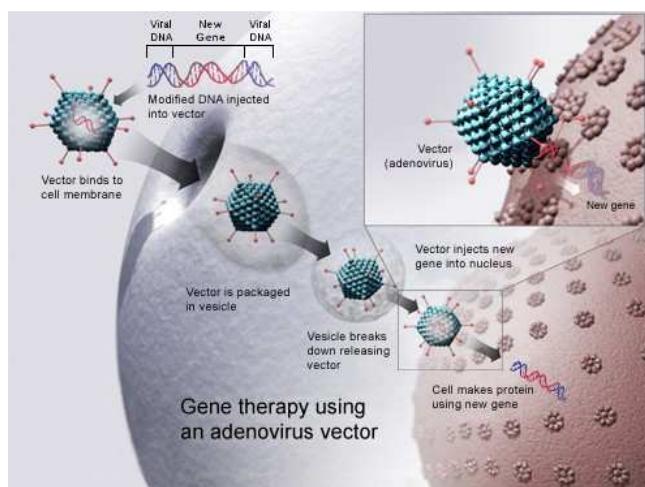


Figure 33. Illustration de la thérapie génique par utilisation de vecteurs viraux. D'après Misra, 2013.

Nous envisageons prochainement de tester l'une de ces stratégies au travers d'un essai clinique chez deux patients azoospermiques chez qui nous avons identifié et validé la mutation causale afin de restaurer temporairement une spermatogénèse normale.

REFERENCES

Adelman CA, Petrini JHJ. ZIP4H (TEX11) deficiency in the mouse impairs meiotic double strand break repair and the regulation of crossing over. *PLoS Genet* 2008;4:e1000042.

Akhmanova A, Mausset-Bonnefont A-L, van Cappellen W, Keijzer N, Hoogenraad CC, Stepanova T, Drabek K, van der Wees J, Mommaas M, Onderwater J, et al. The microtubule plus-end-tracking protein CLIP-170 associates with the spermatid manchette and is essential for spermatogenesis. *Genes Dev* 2005;19:2501–2515.

Amann RP. The cycle of the seminiferous epithelium in humans: a need to revisit? *J Androl* 2008;29:469–487.

Anakwe OO, Gerton GL. Acrosome biogenesis begins during meiosis: evidence from the synthesis and distribution of an acrosomal glycoprotein, acrogranin, during guinea pig spermatogenesis. *Biol Reprod* 1990;42:317–328.

Arnedo N, Templado C, Sánchez-Blanque Y, Rajmil O, Nogués C. Sperm aneuploidy in fathers of Klinefelter's syndrome offspring assessed by multicolour fluorescent in situ hybridization using probes for chromosomes 6, 13, 18, 21, 22, X and Y. *Hum Reprod* 2006;21:524–528.

Aston KI. Genetic susceptibility to male infertility: news from genome-wide association studies. *Andrology* 2014;2:315–321.

Aston KI, Conrad DF. A review of genome-wide approaches to study the genetic basis for spermatogenic defects. *Methods Mol Biol* 2013;927:397–410.

Aston KI, Krausz C, Laface I, Ruiz-Castané E, Carrell DT. Evaluation of 172 candidate polymorphisms for association with oligozoospermia or azoospermia in a large cohort of men of European descent. *Hum Reprod* 2010;25:1383–1397.

Aston KI, Punj V, Liu L, Carrell DT. Genome-wide sperm deoxyribonucleic acid methylation is altered in some men with abnormal chromatin packaging or poor in vitro fertilization embryogenesis. *Fertil Steril* 2012;97:285–292.

Avenarius MR, Hildebrand MS, Zhang Y, Meyer NC, Smith LLH, Kahrizi K, Najmabadi H, Smith RJH. Human male infertility caused by mutations in the CATSPER1 channel protein. *Am J Hum Genet* 2009;84:505–510.

Ayhan Ö, Balkan M, Guven A, Hazan R, Atar M, Tok A, Tolun A. Truncating mutations in TAF4B and ZMYND15 causing recessive azoospermia. *J Med Genet* 2014;51:239–244.

Baccetti B, Collodel G, Estenoz M, Manca D, Moretti E, Piomboni P. Gene deletions in an infertile man with sperm fibrous sheath dysplasia. *Hum Reprod* 2005;20:2790–2794.

Bailey JL. Factors regulating sperm capacitation. *Syst Biol Reprod Med* 2010;56:334–348.

Baker MA, Naumovski N, Hetherington L, Weinberg A, Velkov T, Aitken RJ. Head and flagella subcompartmental proteomic analysis of human spermatozoa. *Proteomics* 2013;13:61–74.

- Balasch J, Gratacós E. Delayed childbearing: effects on fertility and the outcome of pregnancy. *Fetal Diagn Ther* 2011;29:263–273.
- Ballow D, Meistrich ML, Matzuk M, Rajkovic A. Sohlh1 is essential for spermatogonial differentiation. *Dev Biol* 2006;294:161–167.
- Baltz JM, Williams PO, Cone RA. Dense fibers protect mammalian sperm against damage. *Biol Reprod* 1990;43:485–491.
- Bamshad MJ, Ng SB, Bigham AW, Tabor HK, Emond MJ, Nickerson DA, Shendure J. Exome sequencing as a tool for Mendelian disease gene discovery. *Nat Rev Genet* 2011;12:745–755.
- Barroso G, Valdespin C, Vega E, Kershenovich R, Avila R, Avendaño C, Oehninger S. Developmental sperm contributions: fertilization and beyond. *Fertil Steril* 2009;92:835–848.
- Bashamboo A, Ferraz-de-Souza B, Lourenço D, Lin L, Sebire NJ, Montjean D, Bignon-Topalovic J, Mandelbaum J, Siffroi J-P, Christin-Maitre S, et al. Human male infertility associated with mutations in NR5A1 encoding steroidogenic factor 1. *Am J Hum Genet* 2010;87:505–512.
- Benet J, Oliver-Bonet M, Cifuentes P, Templado C, Navarro J. Segregation of chromosomes in sperm of reciprocal translocation carriers: a review. *Cytogenet Genome Res* 2005;111:281–290.
- Berg JS, Evans JP, Leigh MW, Omran H, Bizon C, Mane K, Knowles MR, Weck KE, Zariwala MA. Next generation massively parallel sequencing of targeted exomes to identify genetic mutations in primary ciliary dyskinesia: implications for application to clinical testing. *Genet Med Off J Am Coll Med Genet* 2011;13:218–229.
- Berkovits BD, Wolgemuth DJ. The role of the double bromodomain-containing BET genes during mammalian spermatogenesis. *Curr Top Dev Biol* 2013;102:293–326.
- Berruti G, Paiardi C. Acrosome biogenesis: Revisiting old questions to yield new insights. *Spermatogenesis* 2011;1:95–98.
- Bojesen A, Gravholt CH. Morbidity and mortality in Klinefelter syndrome (47,XXY). *Acta Paediatr* 2011;100:807–813.
- Borg CL, Wolski KM, Gibbs GM, O'Bryan MK. Phenotyping male infertility in the mouse: how to get the most out of a “non-performer.” *Hum Reprod Update* 2010;16:205–224.
- Bower R, Tritschler D, Vanderwaal K, Perrone CA, Mueller J, Fox L, Sale WS, Porter ME. The N-DRC forms a conserved biochemical complex that maintains outer doublet alignment and limits microtubule sliding in motile axonemes. *Mol Biol Cell* 2013;24:1134–1152.
- Bredrup C, Saunier S, Oud MM, Fiskerstrand T, Hoischen A, Brackman D, Leh SM, Midtbø M, Filhol E, Bole-Feysot C, et al. Ciliopathies with skeletal anomalies and renal insufficiency due to mutations in the IFT-A gene WDR19. *Am J Hum Genet* 2011;89:634–643.

- Brown PR, Miki K, Harper DB, Eddy EM. A-kinase anchoring protein 4 binding proteins in the fibrous sheath of the sperm flagellum. *Biol Reprod* 2003;68:2241–2248.
- Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schübeler D, Stadler MB, Peters AHFM. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat Struct Mol Biol* 2010;17:679–687.
- Buck C, Balasubramanian R, Crowley WF. Isolated Gonadotropin-Releasing Hormone (GnRH) Deficiency. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, Bean LJ, Bird TD, Dolan CR, Fong C-T, Smith RJ, et al. (eds) GeneReviews®. Seattle (WA): University of Washington, Seattle, 1993. Available at: <http://www.ncbi.nlm.nih.gov/books/NBK1334/>. Accessed June 4, 2015.
- Bui KH, Sakakibara H, Movassagh T, Oiwa K, Ishikawa T. Molecular architecture of inner dynein arms in situ in *Chlamydomonas reinhardtii* flagella. *J Cell Biol* 2008;183:923–932.
- Burgoyne PS, Mahadevaiah SK, Turner JMA. The management of DNA double-strand breaks in mitotic G2, and in mammalian meiosis viewed from a mitotic G2 perspective. *Bioessays* 2007;29:974–986.
- Cappallo-Obermann H, Schulze W, Jastrow H, Baukloh V, Spiess A-N. Highly purified spermatozoal RNA obtained by a novel method indicates an unusual 28S/18S rRNA ratio and suggests impaired ribosome assembly. *Mol Hum Reprod* 2011;17:669–678.
- Carbajal-González BI, Heuser T, Fu X, Lin J, Smith BW, Mitchell DR, Nicastro D. Conserved structural motifs in the central pair complex of eukaryotic flagella. *Cytoskeleton (Hoboken)* 2013;70:101–120.
- Carlson AE, Westenbroek RE, Quill T, Ren D, Clapham DE, Hille B, Garbers DL, Babcock DF. CatSper1 required for evoked Ca²⁺ entry and control of flagellar function in sperm. *Proc Natl Acad Sci USA* 2003;100:14864–14868.
- Carson AR, Cheung J, Scherer SW. Duplication and relocation of the functional DPY19L2 gene within low copy repeats. *BMC Genomics* 2006;7:45.
- Carrell DT. The clinical implementation of sperm chromosome aneuploidy testing: pitfalls and promises. *J Androl* 2008;29:124–133.
- Carrell DT. Epigenetics of the male gamete. *Fertil Steril* 2012;97:267–274.
- Carrell DT, Aston KI. The search for SNPs, CNVs, and epigenetic variants associated with the complex disease of male infertility. *Syst Biol Reprod Med* 2011;57:17–26.
- Carter AP. Crystal clear insights into how the dynein motor moves. *J Cell Sci* 2013;126:705–713.
- Cavallini G. Male idiopathic oligoasthenoteratozoospermia. *Asian J Androl* 2006;8:143–157.

- Chemes EH, Rawe YV. Sperm pathology: a step beyond descriptive morphology. Origin, characterization and fertility potential of abnormal sperm phenotypes in infertile men. *Hum Reprod Update* 2003;9:405–428.
- Chemes HE, Olmedo SB, Carrere C, Osés R, Carizza C, Leisner M, Blaquier J. Ultrastructural pathology of the sperm flagellum: association between flagellar pathology and fertility prognosis in severely asthenozoospermic men. *Hum Reprod Oxf Engl* 1998;13:2521–2526.
- Cheng CY, Wong EW, Lie PP, Li MW, Mruk DD, Yan HH, Mok K-W, Mannu J, Mathur PP, Lui W-Y, et al. Regulation of blood-testis barrier dynamics by desmosome, gap junction, hemidesmosome and polarity proteins: An unexpected turn of events. *Spermatogenesis* 2011;1:105–115.
- Choi Y, Jeon S, Choi M, Lee M, Park M, Lee DR, Jun K-Y, Kwon Y, Lee O-H, Song S-H, et al. Mutations in SOHLH1 gene associate with nonobstructive azoospermia. *Hum Mutat* 2010;31:788–793.
- Clermont Y. The cycle of the seminiferous epithelium in man. *Am J Anat* 1963;112:35–51.
- Comazzetto S, Di Giacomo M, Rasmussen KD, Much C, Azzi C, Perlas E, Morgan M, O’Carroll D. Oligoasthenoteratozoospermia and infertility in mice deficient for miR-34b/c and miR-449 loci. *PLoS Genet* 2014;10:e1004597.
- Cortessis VK, Thomas DC, Levine AJ, Breton CV, Mack TM, Siegmund KD, Haile RW, Laird PW. Environmental epigenetics: prospects for studying epigenetic mediation of exposure-response relationships. *Hum Genet* 2012;131:1565–1589.
- Cortez D, Marin R, Toledo-Flores D, Froidevaux L, Liechti A, Waters PD, Grützner F, Kaessmann H. Origins and functional evolution of Y chromosomes across mammals. *Nature* 2014;508:488–493.
- Costello S, Michelangeli F, Nash K, Lefievre L, Morris J, Machado-Oliveira G, Barratt C, Kirkman-Brown J, Publicover S. Ca²⁺-stores in sperm: their identities and functions. *Reproduction* 2009;138:425–437.
- Coutton C, Satre V, Arnoult C, Ray P. [Genetics of male infertility: the new players]. *Med Sci (Paris)* 2012;28:497–502.
- Dada R, Kumar M, Jesudasan R, Fernández JL, Gosálvez J, Agarwal A. Epigenetics and its role in male infertility. *J Assist Reprod Genet* 2012;29:213–223.
- Dam AHDM, Feenstra I, Westphal JR, Ramos L, van Golde RJT, Kremer JAM. Globozoospermia revisited. *Hum Reprod Update* 2007;13:63–75.
- Dam AHDM, Koscinski I, Kremer JAM, Moutou C, Jaeger A-S, Oudakker AR, Tournaye H, Charlet N, Lagier-Tourenne C, van Bokhoven H, et al. Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet* 2007;81:813–820.
- De Boer P, de Vries M, Ramos L. A mutation study of sperm head shape and motility in the mouse: lessons for the clinic. *Andrology* 2015;3:174–202.

- Dieterich K, Soto Rifo R, Faure AK, Hennebicq S, Ben Amar B, Zahi M, Perrin J, Martinez D, Sèle B, Jouk P-S, et al. Homozygous mutation of AURKC yields large-headed polyploid spermatozoa and causes male infertility. *Nat Genet* 2007;39:661–665.
- Dieterich K, Zouari R, Harbuz R, Vialard F, Martinez D, Bellayou H, Prisant N, Zoghmar A, Guichaoua MR, Koscienski I, et al. The Aurora Kinase C c.144delC mutation causes meiosis I arrest in men and is frequent in the North African population. *Hum Mol Genet* 2009;18:1301–1309.
- Dirami T, Rode B, Jollivet M, Da Silva N, Escalier D, Gaitch N, Norez C, Tuffery P, Wolf J-P, Becq F, et al. Missense mutations in SLC26A8, encoding a sperm-specific activator of CFTR, are associated with human asthenozoospermia. *Am J Hum Genet* 2013;92:760–766.
- Dymek EE, Heuser T, Nicastro D, Smith EF. The CSC is required for complete radial spoke assembly and wild-type ciliary motility. *Mol Biol Cell* 2011;22:2520–2531.
- Eddy EM. The scaffold role of the fibrous sheath. *Soc Reprod Fertil Suppl* 2007;65:45–62.
- El Hajj N, Zechner U, Schneider E, Tresch A, Gromoll J, Hahn T, Schorsch M, Haaf T. Methylation status of imprinted genes and repetitive elements in sperm DNA from infertile males. *Sex Dev* 2011;5:60–69.
- El Inati E, Muller J, Viville S. Autosomal mutations and human spermatogenic failure. *Biochim Biophys Acta* 2012;1822:1873–1879.
- Elliott DJ, Cooke HJ. The molecular genetics of male infertility. *Bioessays* 1997;19:801–809.
- Escalier D. Human spermatozoa with large heads and multiple flagella: a quantitative ultrastructural study of 6 cases. *Biol Cell Auspices Eur Cell Biol Organ* 1983;48:65–74.
- Escoffier J, Yassine S, Lee HC, Martinez G, Delaroche J, Coutton C, Karaouzène T, Zouari R, Metzler-Guillemain C, Pernet-Gallay K, et al. Subcellular localization of phospholipase C ζ in human sperm and its absence in DPY19L2-deficient sperm are consistent with its role in oocyte activation. *Mol Hum Reprod* 2015;21:157–168.
- Falender AE, Freiman RN, Geles KG, Lo KC, Hwang K, Lamb DJ, Morris PL, Tjian R, Richards JS. Maintenance of spermatogenesis requires TAF4b, a gonad-specific subunit of TFIID. *Genes Dev* 2005;19:794–803.
- Ferlin A, Rocca MS, Vinanzi C, Ghezzi M, Di Nisio A, Foresta C. Mutational screening of NR5A1 gene encoding steroidogenic factor 1 (SF-1) in cryptorchidism and male factor infertility and functional analysis of seven undescribed mutations. *Fertil Steril* 2015.
- Fernández-González R, Moreira PN, Pérez-Crespo M, Sánchez-Martín M, Ramírez MA, Pericuesta E, Bilbao A, Bermejo-Alvarez P, de Dios Hourcade J, de Fonseca FR, et al. Long-term effects of mouse intracytoplasmic sperm injection with DNA-fragmented sperm on health and behavior of adult offspring. *Biol Reprod* 2008;78:761–772.

Fliegauf M, Benzing T, Omran H. When cilia go bad: cilia defects and ciliopathies. *Nat Rev Mol Cell Biol* 2007;8:880–893.

Gaj T, Gersbach CA, Barbas CF. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol* 2013;31:397–405.

Gaucher J, Boussouar F, Montellier E, Curtet S, Buchou T, Bertrand S, Hery P, Jounier S, Depaux A, Vitte A-L, et al. Bromodomain-dependent stage-specific male genome programming by Brdt. *EMBO J* 2012;31:3809–3820.

Gaucher J, Reynoird N, Montellier E, Boussouar F, Rousseaux S, Khochbin S. From meiosis to postmeiotic events: the secrets of histone disappearance. *FEBS J* 2010;277:599–604.

Gekas J, Thepot F, Turleau C, Siffroi JP, Dadoune JP, Briault S, Rio M, Bourouillou G, Carré-Pigeon F, Wasels R, et al. Chromosomal factors of infertility in candidate couples for ICSI: an equal risk of constitutional aberrations in women and men. *Hum Reprod* 2001;16:82–90.

Gilissen C, Hoischen A, Brunner HG, Veltman JA. Disease gene identification strategies for exome sequencing. *Eur J Hum Genet* 2012;20:490–497.

Govin J, Caron C, Lestrat C, Rousseaux S, Khochbin S. The role of histones in chromatin remodelling during mammalian spermiogenesis. *Eur J Biochem* 2004;271:3459–3469.

Griswold MD. Interactions between germ cells and Sertoli cells in the testis. *Biol Reprod* 1995;52:211–216.

Grunewald S, Paasch U, Glander H-J, Anderegg U. Mature human spermatozoa do not transcribe novel RNA. *Andrologia* 2005;37:69–71.

Habura A, Tikhonenko I, Chisholm RL, Koonce MP. Interaction mapping of a dynein heavy chain. Identification of dimerization and intermediate-chain binding domains. *J Biol Chem* 1999;274:15447–15453.

Hammoud SS, Nix DA, Hammoud AO, Gibson M, Cairns BR, Carrell DT. Genome-wide analysis identifies changes in histone retention and epigenetic modifications at developmental and imprinted gene loci in the sperm of infertile men. *Hum Reprod* 2011;26:2558–2569.

Handyside AH. Molecular origin of female meiotic aneuploidies. *Biochim Biophys Acta* 2012;1822:1913–1920.

Haraguchi S, Tsuda M, Kitajima S, Sasaoka Y, Nomura-Kitabayashid A, Kurokawa K, Saga Y. nanos1: a mouse nanos gene expressed in the central nervous system is dispensable for normal development. *Mech Dev* 2003;120:721–731.

Harbuz R, Zouari R, Dieterich K, Nikas Y, Lunardi J, Hennebicq S, Ray P-F. [Function of aurora kinase C (AURKC) in human reproduction]. *Gynécologie Obstétrique Fertil* 2009;37:546–551.

Harbuz R, Zouari R, Pierre V, Ben Khelifa M, Kharouf M, Coutton C, Merdassi G, Abada F, Escoffier J, Nikas Y, et al. A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. *Am J Hum Genet* 2011;88:351–361.

Van der Heijden GW, Dieker JW, Derijck AAHA, Muller S, Berden JHM, Braat DDM, van der Vlag J, de Boer P. Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech Dev* 2005;122:1008–1022.

Hennebicq S, Pelletier R, Bergues U, Rousseaux S. Risk of trisomy 21 in offspring of patients with Klinefelter's syndrome. *Lancet* 2001;357:2104–2105.

Herimo L, Pelletier R-M, Cyr DG, Smith CE. Surfing the wave, cycle, life history, and genes/proteins expressed by testicular germ cells. Part 3: developmental changes in spermatid flagellum and cytoplasmic droplet and interaction of sperm with the zona pellucida and egg plasma membrane. *Microsc Res Tech* 2010;73:320–363.

Heytens E, Parrington J, Coward K, Young C, Lambrecht S, Yoon S-Y, Fissore RA, Hamer R, Deane CM, Ruas M, et al. Reduced amounts and abnormal forms of phospholipase C zeta (PLC ζ) in spermatozoa from infertile men. *Hum Reprod* 2009;24:2417–2428.

Ho HC, Tang CY, Suarez SS. Three-dimensional structure of the Golgi apparatus in mouse spermatids: a scanning electron microscopic study. *Anat Rec* 1999;256:189–194.

Holdcraft RW, Braun RE. Hormonal regulation of spermatogenesis. *Int J Androl* 2004;27:335–342.

Hotaling J, Carrell DT. Clinical genetic testing for male factor infertility: current applications and future directions. *Andrology* 2014;2:339–350.

Hoyer-Fender S, Petersen C, Brohmann H, Rhee K, Wolgemuth DJ. Mouse Odf2 cDNAs consist of evolutionary conserved as well as highly variable sequences and encode outer dense fiber proteins of the sperm tail. *Mol Reprod Dev* 1998;51:167–175.

Hsu PD, Lander ES, Zhang F. Development and Applications of CRISPR-Cas9 for Genome Engineering. *Cell* 2014;157:1262–1278.

Hussain S, Tuorto F, Menon S, Blanco S, Cox C, Flores JV, Watt S, Kudo NR, Lyko F, Frye M. The mouse cytosine-5 RNA methyltransferase NSun2 is a component of the chromatoid body and required for testis differentiation. *Mol Cell Biol* 2013;33:1561–1570.

Huynh T, Mollard R, Trounson A. Selected genetic factors associated with male infertility. *Hum Reprod Update* 2002;8:183–198.

Hu Z, Xia Y, Guo X, Dai J, Li H, Hu H, Jiang Y, Lu F, Wu Y, Yang X, et al. A genome-wide association study in Chinese men identifies three risk loci for non-obstructive azoospermia. *Nat Genet* 2012;44:183–186.

Iguchi N, Yang S, Lamb DJ, Hecht NB. An SNP in protamine 1: a possible genetic cause of male infertility? *J Med Genet* 2006;43:382–384.

Inaba K. Molecular architecture of the sperm flagella: molecules for motility and signaling. *Zool Sci* 2003;20:1043–1056.

Inaba K. Molecular basis of sperm flagellar axonemes: structural and evolutionary aspects. *Ann N Y Acad Sci* 2007;1101:506–526.

Inaba K. Sperm flagella: comparative and phylogenetic perspectives of protein components. *Mol Hum Reprod* 2011;17:524–538.

Ishikawa H, Marshall WF. Ciliogenesis: building the cell's antenna. *Nat Rev Mol Cell Biol* 2011;12:222–234.

Ivliev AE, 't Hoen PAC, van Roon-Mom WMC, Peters DJM, Sergeeva MG. Exploring the transcriptome of ciliated cells using *in silico* dissection of human tissues. *PloS One* 2012;7:e35618.

Jenkins TG, Carrell DT. The sperm epigenome and potential implications for the developing embryo. *Reproduction* 2012;143:727–734.

Jodar M, Selvaraju S, Sendler E, Diamond MP, Krawetz SA, Reproductive Medicine Network. The presence, role and clinical use of spermatozoal RNAs. *Hum Reprod Update* 2013;19:604–624.

Jones R, James PS, Howes L, Bruckbauer A, Kleinerman D. Supramolecular organization of the sperm plasma membrane during maturation and capacitation. *Asian J Androl* 2007;9:438–444.

De Jonge C. Biological basis for human capacitation. *Hum Reprod Update* 2005;11:205–214.

Kashir J, Konstantinidis M, Jones C, Lemmon B, Lee HC, Hamer R, Heindryckx B, Deane CM, De Sutter P, Fissore RA, et al. A maternally inherited autosomal point mutation in human phospholipase C zeta (PLC ζ) leads to male infertility. *Hum Reprod* 2012;27:222–231.

Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. Role of the Dnmt3 family in *de novo* methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Hum Mol Genet* 2007;16:2272–2280.

Kaufmann KB, Büning H, Galy A, Schambach A, Grez M. Gene therapy on the move. *EMBO Mol Med* 2013;5:1642–1661.

Ben Khelifa M, Zouari R, Harbuz R, Halouani L, Arnoult C, Lunardi J, Ray PF. A new AURKC mutation causing macrozoospermia: implications for human spermatogenesis and clinical diagnosis. *Mol Hum Reprod* 2011;17:762–768.

Ben Khelifa M, Coutton C, Zouari R, Karaouzène T, Rendu J, Bidart M, Yassine S, Pierre V, Delaroche J, Hennebicq S, et al. Mutations in DNAH1, which encodes an inner arm

heavy chain dynein, lead to male infertility from multiple morphological abnormalities of the sperm flagella. *Am J Hum Genet* 2014;94:95–104.

Kichine E, Msaidie S, Bokilo AD, Ducourneau A, Navarro A, Levy N, Terriou P, Collignon P, Boetsch G, Chiaroni J, et al. Low-frequency protamine 1 gene transversions c.102G->T and c.-107G->C do not correlate with male infertility. *J Med Genet* 2008;45:255–256.

Kierszenbaum AL, Tres LL. RNA transcription and chromatin structure during meiotic and postmeiotic stages of spermatogenesis. *Fed Proc* 1978;37:2512–2516.

Kierszenbaum AL, Tres LL. The acrosome-acroplaxome-manchette complex and the shaping of the spermatid head. *Arch Histol Cytol* 2004;67:271–284.

King SM. AAA domains and organization of the dynein motor unit. *J Cell Sci* 2000;113 (Pt 14):2521–2526.

King SM. Integrated control of axonemal dynein AAA(+) motors. *J Struct Biol* 2012;179:222–228.

Kobayashi H, Hiura H, John RM, Sato A, Otsu E, Kobayashi N, Suzuki R, Suzuki F, Hayashi C, Utsunomiya T, et al. DNA methylation errors at imprinted loci after assisted conception originate in the parental sperm. *Eur J Hum Genet* 2009;17:1582–1591.

Kobayashi H, Sato A, Otsu E, Hiura H, Tomatsu C, Utsunomiya T, Sasaki H, Yaegashi N, Arima T. Aberrant DNA methylation of imprinted loci in sperm from oligospermic patients. *Hum Mol Genet* 2007;16:2542–2551.

Komada M, McLean DJ, Griswold MD, Russell LD, Soriano P. E-MAP-115, encoding a microtubule-associated protein, is a retinoic acid-inducible gene required for spermatogenesis. *Genes Dev* 2000;14:1332–1342.

Kon T, Nishiura M, Ohkura R, Toyoshima YY, Sutoh K. Distinct functions of nucleotide-binding/hydrolysis sites in the four AAA modules of cytoplasmic dynein. *Biochemistry* 2004;43:11266–11274.

Kopera IA, Bilinska B, Cheng CY, Mruk DD. Sertoli-germ cell junctions in the testis: a review of recent data. *Philos Trans R Soc Lond, B, Biol Sci* 2010;365:1593–1605.

Koscinski I, Elinati E, Fossard C, Redin C, Muller J, Velez de la Calle J, Schmitt F, Ben Khelifa M, Ray PF, Ray P, et al. DPY19L2 deletion as a major cause of globozoospermia. *Am J Hum Genet* 2011;88:344–350.

Krausz C. Male infertility: pathogenesis and clinical diagnosis. *Best Pract Res Clin Endocrinol Metab* 2011;25:271–285.

Krausz C, Forti G. Clinical aspects of male infertility. *Results Probl Cell Differ* 2000;28:1–21.

De Kretser DM, Loveland KL, Meinhardt A, Simorangkir D, Wreford N. Spermatogenesis. *Hum Reprod* 1998;13 Suppl 1:1–8.

Kuo Y-C, Lin Y-H, Chen H-I, Wang Y-Y, Chiou Y-W, Lin H-H, Pan H-A, Wu C-M, Su S-M, Hsu C-C, et al. SEPT12 mutations cause male infertility with defective sperm annulus. *Hum Mutat* 2012;33:710–719.

Kusz-Zamelczyk K, Sajek M, Spik A, Glazar R, Jędrzejczak P, Łatos-Bieleńska A, Kotecki M, Pawelczyk L, Jaruzelska J. Mutations of NANOS1, a human homologue of the *Drosophila* morphogen, are associated with a lack of germ cells in testes or severe oligo-astheno-teratozoospermia. *J Med Genet* 2013;50:187–193.

De Lamirande E, O’Flaherty C. Sperm activation: role of reactive oxygen species and kinases. *Biochim Biophys Acta* 2008;1784:106–115.

Lanfranco F, Kamischke A, Zitzmann M, Nieschlag E. Klinefelter’s syndrome. *Lancet* 2004;364:273–283.

Lindemann CB. Functional significance of the outer dense fibers of mammalian sperm examined by computer simulations with the geometric clutch model. *Cell Motil Cytoskeleton* 1996;34:258–270.

Lin J, Okada K, Raytchev M, Smith MC, Nicastro D. Structural mechanism of the dynein power stroke. *Nat Cell Biol* 2014;16:479–485.

Lin L, Achermann JC. Steroidogenic factor-1 (SF-1, Ad4BP, NR5A1) and disorders of testis development. *Sex Dev* 2008;2:200–209.

Lin Y-H, Lin Y-M, Wang Y-Y, Yu I-S, Lin Y-W, Wang Y-H, Wu C-M, Pan H-A, Chao S-C, Yen PH, et al. The expression level of septin12 is critical for spermiogenesis. *Am J Pathol* 2009;174:1857–1868.

Liu G, Shi Q-W, Lu G-X. A newly discovered mutation in PICK1 in a human with globozoospermia. *Asian J Androl* 2010;12:556–560.

Liu P, Carvalho CMB, Hastings PJ, Lupski JR. Mechanisms for recurrent and complex human genomic rearrangements. *Curr Opin Genet Dev* 2012;22:211–220.

Lohiya NK, Kothari LK, Manivannan B, Mishra PK, Pathak N. Human sperm immobilization effect of *Carica papaya* seed extracts: an in vitro study. *Asian J Androl* 2000;2:103–109.

Lourenço D, Brauner R, Lin L, De Perdigão A, Weryha G, Muresan M, Boudjenah R, Guerra-Junior G, Maciel-Guerra AT, Achermann JC, et al. Mutations in NR5A1 associated with ovarian insufficiency. *N Engl J Med* 2009;360:1200–1210.

Lourenço D, Brauner R, Rybczynska M, Nihoul-Fékété C, McElreavey K, Bashamboo A. Loss-of-function mutation in GATA4 causes anomalies of human testicular development. *Proc Natl Acad Sci USA* 2011;108:1597–1602.

Lu C, Xu M, Wang R, Qin Y, Wang Y, Wu W, Song L, Wang S, Shen H, Sha J, et al. Pathogenic variants screening in five non-obstructive azoospermia-associated genes. *Mol Hum Reprod* 2014;20:178–183.

- Makokha M, Hare M, Li M, Hays T, Barbar E. Interactions of cytoplasmic dynein light chains Tctex-1 and LC8 with the intermediate chain IC74. *Biochemistry* 2002;41:4302–4311.
- Marques CJ, Costa P, Vaz B, Carvalho F, Fernandes S, Barros A, Sousa M. Abnormal methylation of imprinted genes in human sperm is associated with oligozoospermia. *Mol Hum Reprod* 2008;14:67–74.
- Massart A, Lissens W, Tournaye H, Stouffs K. Genetic causes of spermatogenic failure. *Asian J Androl* 2012;14:40–48.
- De Mateo S, Castillo J, Estanyol JM, Ballesca JL, Oliva R. Proteomic characterization of the human sperm nucleus. *Proteomics* 2011;11:2714–2726.
- Mali P, Esvelt KM, Church GM. Cas9 as a versatile tool for engineering biology. *Nat Methods* 2013;10:957–963.
- Mazumdar M, Mikami A, Gee MA, Vallee RB. In vitro motility from recombinant dynein heavy chain. *Proc Natl Acad Sci USA* 1996;93:6552–6556.
- Mendoza-Lujambio I, Burfeind P, Dixkens C, Meinhardt A, Hoyer-Fender S, Engel W, Neesen J. The Hook1 gene is non-functional in the abnormal spermatozoon head shape (azh) mutant mouse. *Hum Mol Genet* 2002;11:1647–1658.
- Menkveld R, Holleboom CAG, Rhemrev JPT. Measurement and significance of sperm morphology. *Asian J Androl* 2011;13:59–68.
- Meola G, Cardani R. Myotonic dystrophies: An update on clinical aspects, genetic, pathology, and molecular pathomechanisms. *Biochim Biophys Acta* 2015;1852:594–606.
- Michaut M, Tomes CN, De Blas G, Yunes R, Mayorga LS. Calcium-triggered acrosomal exocytosis in human spermatozoa requires the coordinated activation of Rab3A and N-ethylmaleimide-sensitive factor. *Proc Natl Acad Sci USA* 2000;97:9996–10001.
- Miki K, Willis WD, Brown PR, Goulding EH, Fulcher KD, Eddy EM. Targeted disruption of the Akap4 gene causes defects in sperm flagellum and motility. *Dev Biol* 2002;248:331–342.
- Misra S. Human gene therapy: a brief overview of the genetic revolution. *J Assoc Physicians India* 2013;61:127–133.
- Miyamoto T, Hasuike S, Yoge L, Maduro MR, Ishikawa M, Westphal H, Lamb DJ. Azoospermia in patients heterozygous for a mutation in SYCP3. *Lancet* 2003;362:1714–1719.
- Montjean D, Ravel C, Benkhaliifa M, Cohen-Bacrie P, Berthaut I, Bashamboo A, McElreavey K. Methylation changes in mature sperm deoxyribonucleic acid from oligozoospermic men: assessment of genetic variants and assisted reproductive technology outcome. *Fertil Steril* 2013;100:1241–1247.
- Morel F, Roux C, Bresson JL. Segregation of sex chromosomes in spermatozoa of 46,XY/47,XXY men by multicolour fluorescence in-situ hybridization. *Mol Hum Reprod* 2000;6:566–570.

- Moreno RD, Alvarado CP. The mammalian acrosome as a secretory lysosome: new and old evidence. *Mol Reprod Dev* 2006;73:1430–1434.
- Moreno RD, Palomino J, Schatten G. Assembly of spermatid acrosome depends on microtubule organization during mammalian spermiogenesis. *Dev Biol* 2006;293:218–227.
- Moreno RD, Ramalho-Santos J, Sutovsky P, Chan EK, Schatten G. Vesicular traffic and golgi apparatus dynamics during mammalian spermatogenesis: implications for acrosome architecture. *Biol Reprod* 2000;63:89–98.
- Mou L, Wang Y, Li H, Huang Y, Jiang T, Huang W, Li Z, Chen J, Xie J, Liu Y, et al. A dominant-negative mutation of HSF2 associated with idiopathic azoospermia. *Hum Genet* 2013;132:159–165.
- Mruk DD, Cheng CY. Sertoli-Sertoli and Sertoli-germ cell interactions and their significance in germ cell movement in the seminiferous epithelium during spermatogenesis. *Endocr Rev* 2004;25:747–806.
- Nanassy L, Carrell DT. Abnormal methylation of the promoter of CREM is broadly associated with male factor infertility and poor sperm quality but is improved in sperm selected by density gradient centrifugation. *Fertil Steril* 2011;95:2310–2314.
- Neesen J, Kirschner R, Ochs M, Schmiedl A, Habermann B, Mueller C, Holstein AF, Nuesslein T, Adham I, Engel W. Disruption of an inner arm dynein heavy chain gene results in asthenozoospermia and reduced ciliary beat frequency. *Hum Mol Genet* 2001;10:1117–1128.
- Nicastro D, Schwartz C, Pierson J, Gaudette R, Porter ME, McIntosh JR. The molecular architecture of axonemes revealed by cryoelectron tomography. *Science* 2006;313:944–948.
- Nistal M, Paniagua R, Herruzo A. Multi-tailed spermatozoa in a case with asthenospermia and teratospermia. *Virchows Arch B Cell Pathol* 1977;26:111–118.
- Nuti F, Krausz C. Gene polymorphisms/mutations relevant to abnormal spermatogenesis. *Reprod Biomed Online* 2008;16:504–513.
- O'Donnell L, Nicholls PK, O'Bryan MK, McLachlan RI, Stanton PG. Spermiation: The process of sperm release. *Spermatogenesis* 2011;1:14–35.
- Ogawa T, Aréchaga JM, Avarbock MR, Brinster RL. Transplantation of testis germinal cells into mouse seminiferous tubules. *Int J Dev Biol* 1997;41:111–122.
- Oko RJ. Developmental expression and possible role of perinuclear theca proteins in mammalian spermatozoa. *Reprod Fertil Dev* 1995;7:777–797.
- Oko R, Sutovsky P. Biogenesis of sperm perinuclear theca and its role in sperm functional competence and fertilization. *J Reprod Immunol* 2009a;83:2–7.
- Oko R, Sutovsky P. Biogenesis of sperm perinuclear theca and its role in sperm functional competence and fertilization. *J Reprod Immunol* 2009b;83:2–7.

- Oliva A, Spira A, Multigner L. Contribution of environmental factors to the risk of male infertility. *Hum Reprod* 2001;16:1768–1776.
- Ostrowski LE, Dutcher SK, Lo CW. Cilia and models for studying structure and function. *Proc Am Thorac Soc* 2011;8:423–429.
- Ozaki-Kuroda K, Nakanishi H, Ohta H, Tanaka H, Kurihara H, Mueller S, Irie K, Ikeda W, Sakai T, Wimmer E, et al. Nectin couples cell-cell adhesion and the actin scaffold at heterotypic testicular junctions. *Curr Biol* 2002;12:1145–1150.
- Parker KL, Schimmer BP, Schedl A. Genes essential for early events in gonadal development. *Cell Mol Life Sci* 1999;55:831–838.
- Perrault I, Saunier S, Hanein S, Filhol E, Bizet AA, Collins F, Salih MAM, Gerber S, Delphin N, Bigot K, et al. Mainzer-Saldino syndrome is a ciliopathy caused by IFT140 mutations. *Am J Hum Genet* 2012;90:864–870.
- Pigino G, Bui KH, Maheshwari A, Lupetti P, Diener D, Ishikawa T. Cryoelectron tomography of radial spokes in cilia and flagella. *J Cell Biol* 2011;195:673–687.
- Pigino G, Ishikawa T. Axonemal radial spokes: 3D structure, function and assembly. *Bioarchitecture* 2012;2:50–58.
- Piomboni P, Focarelli R, Stendardi A, Ferramosca A, Zara V. The role of mitochondria in energy production for human sperm motility. *Int J Androl* 2012;35:109–124.
- Pivot-Pajot C, Caron C, Govin J, Vion A, Rousseaux S, Khochbin S. Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. *Mol Cell Biol* 2003;23:5354–5365.
- Qiu Q-M, Liu G, Li W-N, Shi Q-W, Zhu F-X, Lu G-X. [Mutation of KLHL-10 in idiopathic infertile males with azoospermia, oligospermia or asthenospermia]. *Zhonghua Nan Ke Xue* 2009;15:974–979.
- Ramasamy R, Bakırçioğlu ME, Cengiz C, Karaca E, Scovell J, Jhangiani SN, Akdemir ZC, Bainbridge M, Yu Y, Huff C, et al. Whole-exome sequencing identifies novel homozygous mutation in NPAS2 in family with nonobstructive azoospermia. *Fertil Steril* 2015.
- Rato L, Alves MG, Socorro S, Duarte AI, Cavaco JE, Oliveira PF. Metabolic regulation is important for spermatogenesis. *Nat Rev Urol* 2012;9:330–338.
- Ravel C, Berthaut I, Bresson JL, Siffroi JP, Genetics Commission of the French Federation of CECOS. Prevalence of chromosomal abnormalities in phenotypically normal and fertile adult males: large-scale survey of over 10,000 sperm donor karyotypes. *Hum Reprod* 2006;21:1484–1489.
- Ren D, Navarro B, Perez G, Jackson AC, Hsu S, Shi Q, Tilly JL, Clapham DE. A sperm ion channel required for sperm motility and male fertility. *Nature* 2001;413:603–609.
- Robinson PN, Krawitz P, Mundlos S. Strategies for exome and genome sequence data analysis in disease-gene discovery projects. *Clin Genet* 2011;80:127–132.

Rohozinski J, Lamb DJ, Bishop CE. UTP14c is a recently acquired retrogene associated with spermatogenesis and fertility in man. *Biol Reprod* 2006;74:644–651.

Rolland M, Le Moal J, Wagner V, Royère D, De Mouzon J. Decline in semen concentration and morphology in a sample of 26,609 men close to general population between 1989 and 2005 in France. *Hum Reprod* 2013;28:462–470.

Röpke A, Tewes A-C, Gromoll J, Kliesch S, Wieacker P, Tüttelmann F. Comprehensive sequence analysis of the NR5A1 gene encoding steroidogenic factor 1 in a large group of infertile males. *Eur J Hum Genet* 2013;21:1012–1015.

Rosales JL, Sarker K, Ho N, Broniewska M, Wong P, Cheng M, van der Hoorn FA, Lee K-Y. ODF1 phosphorylation by Cdk5/p35 enhances ODF1-OIP1 interaction. *Cell Physiol Biochem* 2007;20:311–318.

Rousseaux S, Caron C, Govin J, Lestrat C, Faure A-K, Khochbin S. Establishment of male-specific epigenetic information. *Gene* 2005;345:139–153.

Royal M, Mann GE, Flint AP. Strategies for reversing the trend towards subfertility in dairy cattle. *Vet J* 2000;160:53–60.

Russell LD, Bartke A, Goh JC. Postnatal development of the Sertoli cell barrier, tubular lumen, and cytoskeleton of Sertoli and myoid cells in the rat, and their relationship to tubular fluid secretion and flow. *Am J Anat* 1989;184:179–189.

Sasaki H, Matsui Y. Epigenetic events in mammalian germ-cell development: reprogramming and beyond. *Nat Rev Genet* 2008;9:129–140.

Satir P, Christensen ST. Structure and function of mammalian cilia. *Histochem Cell Biol* 2008;129:687–693.

Sato Y, Jinam T, Iwamoto T, Yamauchi A, Imoto I, Inoue I, Tajima A. Replication study and meta-analysis of human nonobstructive azoospermia in Japanese populations. *Biol Reprod* 2013;88:87.

Schatten H, Sun Q-Y. The functional significance of centrosomes in mammalian meiosis, fertilization, development, nuclear transfer, and stem cell differentiation. *Environ Mol Mutagen* 2009;50:620–636.

Schmidts M, Frank V, Eisenberger T, Al Turki S, Bizet AA, Antony D, Rix S, Decker C, Bachmann N, Bald M, et al. Combined NGS approaches identify mutations in the intraflagellar transport gene IFT140 in skeletal ciliopathies with early progressive kidney Disease. *Hum Mutat* 2013;34:714–724.

Sendler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R, Krawetz SA. Stability, delivery and functions of human sperm RNAs at fertilization. *Nucleic Acids Res* 2013;41:4104–4117.

Sharpe RM, McKinnell C, Kivlin C, Fisher JS. Proliferation and functional maturation of Sertoli cells, and their relevance to disorders of testis function in adulthood. *Reproduction* 2003;125:769–784.

Singleton AB. Exome sequencing: a transformative technology. *Lancet Neurol* 2011;10:942–946.

Sofikitis N, Giotitsas N, Tsounapi P, Baltogiannis D, Giannakis D, Pardalidis N. Hormonal regulation of spermatogenesis and spermiogenesis. *J Steroid Biochem Mol Biol* 2008;109:323–330.

Song X, Zhao Y, Cai Q, Zhang Y, Niu Y. Association of the Glutathione S-transferases M1 and T1 polymorphism with male infertility: a meta-analysis. *J Assist Reprod Genet* 2013;30:131–141.

Stouffs K, Lissens W, Tournaye H, Van Steirteghem A, Liebaers I. SYCP3 mutations are uncommon in patients with azoospermia. *Fertil Steril* 2005;84:1019–1020.

Takada S, Wilkerson CG, Wakabayashi K, Kamiya R, Witman GB. The outer dynein arm-docking complex: composition and characterization of a subunit (oda1) necessary for outer arm assembly. *Mol Biol Cell* 2002;13:1015–1029.

Takasaki N, Tachibana K, Ogasawara S, Matsuzaki H, Hagiuda J, Ishikawa H, Mochida K, Inoue K, Ogonuki N, Ogura A, et al. A heterozygous mutation of GALNTL5 affects male infertility with impairment of sperm motility. *Proc Natl Acad Sci USA* 2014;111:1120–1125.

Tiepolo L, Zuffardi O. Localization of factors controlling spermatogenesis in the nonfluorescent portion of the human Y chromosome long arm. *Hum Genet* 1976;34:119–124.

Tomes CN. The proteins of exocytosis: lessons from the sperm model. *Biochem J* 2015;465:359–370.

Turner RM, Musse MP, Mandal A, Klotz K, Jayes FC, Herr JC, Gerton GL, Moss SB, Chemes HE. Molecular genetic analysis of two human sperm fibrous sheath proteins, AKAP4 and AKAP3, in men with dysplasia of the fibrous sheath. *J Androl* 2001;22:302–315.

Tüttelmann F, Laan M, Grigorova M, Punab M, Söber S, Gromoll J. Combined effects of the variants FSHB -211G>T and FSHR 2039A>G on male reproductive parameters. *J Clin Endocrinol Metab* 2012;97:3639–3647.

Tüttelmann F, Simoni M, Kliesch S, Ledig S, Dworniczak B, Wieacker P, Röpke A. Copy number variants in patients with severe oligozoospermia and Sertoli-cell-only syndrome. *PLoS ONE* 2011;6:e19426.

Tynan SH, Gee MA, Vallee RB. Distinct but overlapping sites within the cytoplasmic dynein heavy chain for dimerization and for intermediate chain and light intermediate chain binding. *J Biol Chem* 2000;275:32769–32774.

Vincensini L, Blisnick T, Bastin P. 1001 model organisms to study cilia and flagella. *Biol Cell* 2011;103:109–130.

Vogt PH. AZF deletions and Y chromosomal haplogroups: history and update based on sequence. *Hum Reprod Update* 2005;11:319–336.

- Vorona E, Zitzmann M, Gromoll J, Schüring AN, Nieschlag E. Clinical, endocrinological, and epigenetic features of the 46,XX male syndrome, compared with 47,XXY Klinefelter patients. *J Clin Endocrinol Metab* 2007;92:3458–3465.
- Wang G, Ying Z, Jin X, Tu N, Zhang Y, Phillips M, Moskophidis D, Mivechi NF. Essential requirement for both hsf1 and hsf2 transcriptional activity in spermatogenesis and male fertility. *Genesis* 2004;38:66–80.
- Wang S, Zheng H, Esaki Y, Kelly F, Yan W. Cullin3 is a KLHL10-interacting protein preferentially expressed during late spermiogenesis. *Biol Reprod* 2006;74:102–108.
- Ward WS. The structure of the sleeping genome: implications of sperm DNA organization for somatic cells. *J Cell Biochem* 1994;55:77–82.
- Wargo MJ, Smith EF. Asymmetry of the central apparatus defines the location of active microtubule sliding in Chlamydomonas flagella. *Proc Natl Acad Sci USA* 2003;100:137–142.
- Wei B, Xu Z, Ruan J, Zhu M, Jin K, Zhou D, Xu Z, Hu Q, Wang Q, Wang Z. MTHFR 677C>T and 1298A>C polymorphisms and male infertility risk: a meta-analysis. *Mol Biol Rep* 2012;39:1997–2002.
- Wistuba J, Stukenborg JB, Luetjens CM. Mammalian Spermatogenesis. *Funct Dev Embryol* 2007;1:99–117.
- Wu H, Maciejewski MW, Marintchev A, Benashski SE, Mullen GP, King SM. Solution structure of a dynein motor domain associated light chain. *Nat Struct Biol* 2000;7:575–579.
- Yang F, Gell K, van der Heijden GW, Eckardt S, Leu NA, Page DC, Benavente R, Her C, Höög C, McLaughlin KJ, et al. Meiotic failure in male mice lacking an X-linked factor. *Genes Dev* 2008;22:682–691.
- Yang P, Diener DR, Yang C, Kohno T, Pazour GJ, Dienes JM, Agrin NS, King SM, Sale WS, Kamiya R, et al. Radial spoke proteins of Chlamydomonas flagella. *J Cell Sci* 2006;119:1165–1174.
- Yan W, Ma L, Burns KH, Matzuk MM. Haploinsufficiency of kelch-like protein homolog 10 causes infertility in male mice. *Proc Natl Acad Sci USA* 2004;101:7793–7798.
- Yan W, Si Y, Slaymaker S, Li J, Zheng H, Young DL, Aslanian A, Saunders L, Verdin E, Charo IF. Zmynd15 encodes a histone deacetylase-dependent transcriptional repressor essential for spermiogenesis and male fertility. *J Biol Chem* 2010;285:31418–31426.
- Yatsenko AN, Georgiadis AP, Röpke A, Berman AJ, Jaffe T, Olszewska M, Westernströer B, Sanfilippo J, Kurpisz M, Rajkovic A, et al. X-linked TEX11 mutations, meiotic arrest, and azoospermia in infertile men. *N Engl J Med* 2015;372:2097–2107.
- Yatsenko AN, O’Neil DS, Roy A, Arias-Mendoza PA, Chen R, Murthy LJ, Lamb DJ, Matzuk MM. Association of mutations in the zona pellucida binding protein 1 (ZPBP1) gene with abnormal sperm head morphology in infertile men. *Mol Hum Reprod* 2012;18:14–21.

Yatsenko AN, Roy A, Chen R, Ma L, Murthy LJ, Yan W, Lamb DJ, Matzuk MM. Non-invasive genetic diagnosis of male infertility using spermatozoal RNA: KLHL10 mutations in oligozoospermic patients impair homodimerization. *Hum Mol Genet* 2006;15:3411–3419.

Yatsenko AN, Yatsenko SA, Weedin JW, Lawrence AE, Patel A, Peacock S, Matzuk MM, Lamb DJ, Cheung SW, Lipshultz LI. Comprehensive 5-year study of cytogenetic aberrations in 668 infertile men. *J Urol* 2010;183:1636–1642.

Yoshinaga K, Toshimori K. Organization and modifications of sperm acrosomal molecules during spermatogenesis and epididymal maturation. *Microsc Res Tech* 2003;61:39–45.

Yuan L, Liu JG, Zhao J, Brundell E, Daneholt B, Höög C. The murine SCP3 gene is required for synaptonemal complex assembly, chromosome synapsis, and male fertility. *Mol Cell* 2000;5:73–83.

Yu J, Chen Z, Ni Y, Li Z. CFTR mutations in men with congenital bilateral absence of the vas deferens (CBAVD): a systemic review and meta-analysis. *Hum Reprod* 2012;27:25–35.

Zare-Abdollahi D, Safari S, Mirfakhraie R, Movafagh A, Bastami M, Azimzadeh P, Salsabili N, Ebrahimizadeh W, Salami S, Omrani MD. Mutational screening of the NR5A1 in azoospermia. *Andrologia* 2015;47:395–401.

Zhou J, Du Y-R, Qin W-H, Hu Y-G, Huang Y-N, Bao L, Han D, Mansouri A, Xu G-L. RIM-BP3 is a manchette-associated protein essential for spermiogenesis. *Development* 2009;136:373–382.

ANNEXES

Liste et résumés des publications sur la tératozoospermie non incluses dans la thèse.

Article S1

Harbuz R, Zouari R, Pierre V, Ben Khelifa M, Kharouf M, Coutton C, Merdassi G, Abada F, Escoffier J, Nikas Y, *et al.* A recurrent deletion of DPY19L2 causes infertility in man by blocking sperm head elongation and acrosome formation. *Am J Hum Genet* 2011;88:351–361.

Article S2

Coutton C, Satre V, Arnoult C, Ray P. [Genetics of male infertility: the new players]. *Médecine Sci MS* 2012;28:497–502.

Article S3

Pierre V, Martinez G, Coutton C, Delaroche J, Yassine S, Novella C, Pernet-Gallay K, Hennebicq S, Ray PF, Arnoult C. Absence of Dpy19l2, a new inner nuclear membrane protein, causes globozoospermia in mice by preventing the anchoring of the acrosome to the nucleus. *Dev Camb Engl* 2012;139:2955–2965.

Article S4

Karaouzène T, El Atifi M, Issartel JP, Grepillat M, Coutton C, Martinez D, Arnoult C, Ray PF. Comparative testicular transcriptome of wild type and globozoospermic Dpy19l2 knock out mice. *Basic Clin Androl* 2013; 3;23:7

Article S5

Ounis L, Zoghmar A, Coutton C, Rouabah L, Hachemi M, Martinez D, Martinez G, Bellil I, Khelifi D, Arnoult C, *et al.* Mutations of the aurora kinase C gene causing macrozoospermia are the most frequent genetic cause of male infertility in Algerian men. *Asian J Androl* 2015;17:68-73

Article S6

Escoffier J, Yassine S, Lee HC, Martinez G, Delaroche J, Coutton C, Karaouzène T, Zouari R, Metzler-Guillemain C, Pernet-Gallay K, *et al.* Subcellular localization of phospholipase C ζ in human sperm and its absence in DPY19L2-deficient sperm are consistent with its role in oocyte activation. *Mol Hum Reprod* 2015;21:157-68

Article S7

Yassine S, Escoffier J, Martinez G, Coutton C, Karaouzène T, Zouari R, Ravanat J-L, Metzler-Guillemain C, Fissore R, Hennebicq S, *et al.* Dpy19l2-deficient globozoospermic sperm display altered genome packaging and DNA damage that compromises the initiation of embryo development. *Mol Hum Reprod* 2015;21:169-85

A Recurrent Deletion of *DPY19L2* Causes Infertility in Man by Blocking Sperm Head Elongation and Acrosome Formation

Radu Harbuz,^{1,2,3} Raoudha Zouari,⁴ Virginie Pierre,^{2,5} Mariem Ben Khelifa,^{1,2,3} Mahmoud Kharouf,⁴ Charles Coutton,^{1,2,3} Ghaya Merdassi,⁶ Farid Abada,^{1,2,3} Jessica Escoffier,⁵ Yorgos Nikas,⁷ François Vialard,⁸ Isabelle Koscinski,^{9,10} Chema Triki,¹¹ Nathalie Sermondade,¹² Thérèse Schweitzer,¹³ Amel Zhioua,⁶ Fethi Zhioua,⁶ Habib Latrous,⁴ Lazhar Halouani,⁴ Marrakchi Ouafi,⁴ Mounir Makni,⁴ Pierre-Simon Jouk,^{1,2} Bernard Sèle,¹ Sylviane Hennebicq,¹ Véronique Satre,^{1,2} Stéphane Viville,^{9,10} Christophe Arnoult,^{2,5} Joël Lunardi,^{1,3,5} and Pierre F. Ray^{1,2,3,*}

An increasing number of couples require medical assistance to achieve a pregnancy, and more than 2% of the births in Western countries now result from assisted reproductive technologies. To identify genetic variants responsible for male infertility, we performed a whole-genome SNP scan on patients presenting with total globozoospermia, a primary infertility phenotype characterized by the presence of 100% round acrosomeless spermatozoa in the ejaculate. This strategy allowed us to identify in most patients (15/20) a 200 kb homozygous deletion encompassing only *DPY19L2*, which is highly expressed in the testis. Although there was no known function for *DPY19L2* in humans, previous work indicated that its ortholog in *C. elegans* is involved in cell polarity. In man, the *DPY19L2* region has been described as a copy-number variant (CNV) found to be duplicated and heterozygously deleted in healthy individuals. We show here that the breakpoints of the deletions are located on a highly homologous 28 kb low copy repeat (LCR) sequence present on each side of *DPY19L2*, indicating that the identified deletions were probably produced by nonallelic homologous recombination (NAHR) between these two regions. We demonstrate that patients with globozoospermia have a homozygous deletion of *DPY19L2*, thus indicating that *DPY19L2* is necessary in men for sperm head elongation and acrosome formation. A molecular diagnosis can now be proposed to affected men; the presence of the deletion confirms the diagnosis of globozoospermia and assigns a poor prognosis for the success of in vitro fertilization.

The increasing incidence of infertile couples, potentially caused by a general deterioration of sperm parameters, is becoming a major concern worldwide.¹ Although environmental or infectious causes play an important role in infertility, genetic defects are also believed to be frequently involved in the pathological process.² Several hundred genes are believed to be involved in spermatogenesis, yet very few have so far been directly connected with male infertility in human. A better understanding of gametogenesis through the identification of genes involved in infertility can help to decrease this worrying trend. We investigated patients with globozoospermia (MIM 102530) to identify genes involved in this infertility syndrome. Globozoospermia is a rare phenotype of primary male infertility characterized by the production of a majority of round-headed spermatozoa without acrosome (Figure 1, for review see³). The phenotype was described over 30 years ago, and familial cases pointed to a genetic component for this defect.^{4–7} Three brothers affected with total globozoospermia were recently

analyzed and found to carry a homozygous mutation of *SPATA16* (MIM 609856).⁸ The testicular expression of *SPATA16* and its intracellular localization in the acrosome-building Golgi vesicles in the spermatids correlated well with the observed phenotype.^{9,10} However, no *SPATA16* mutations were detected in 29 other affected men, thus suggesting that *SPATA16* was not the main locus associated with globozoospermia.⁸ We recently demonstrated that the strategy of using whole-genome homozygosity mapping applied to infertile patients from the same ethno-geographical background presenting with a specific morphologic anomaly of the sperm could lead to the localization and identification of genes involved in spermatogenesis.¹¹ We demonstrated that *AURKC* (MIM 603495) deficiency led to male infertility due to the production of large-headed, multiflagellar, polyploid spermatozoa (MIM 24306).^{12,13} Here, we applied the same genetic strategy to a cohort of mainly Tunisian patients presenting with total globozoospermia, and we were able to demonstrate that 15 of the 20 patients that we analyzed

¹Faculté de Médecine-Pharmacie de Grenoble, Université Joseph Fourier, Domaine de la Merci, 38706 Grenoble, France; ²Laboratoire AGIM, FRE 3405 CNRS - UJF, Equipe Génétique Infertilité et Thérapeutique (GIT), Domaine de la Merci, 38706 Grenoble, France; ³UM de Biochimie et Génétique Moléculaire, DBTP, CHU de Grenoble, 38700 Grenoble, France; ⁴Clinique de la Reproduction les Jasmins, 23, Av. Louis BRAILLE, 1002 Tunis, Tunisia; ⁵Grenoble Institut des Neurosciences, INSERM U836, BP170, 38042 Grenoble, France; ⁶Unité de Procréation Médicalement Assistée, Hôpital Aziza Othmana, 1004 Tunis, Tunisia; ⁷Athens Innovative Microscopy Skra 36, Voula 16673, Greece; ⁸Department of Reproductive Biology, Centre Hospitalier Poissy Saint Germain, 78300 Poissy, France; ⁹Service de Biologie de la Reproduction, Centre Hospitalier Universitaire, Strasbourg F 67000 France; ¹⁰Institut de Génétique et de Biologie Moléculaire et Cellulaire (IGBMC), Institut National de Santé et de Recherche Médicale (INSERM) U964/Centre National de Recherche Scientifique (CNRS) UMR 1704/Université de Strasbourg, 67404 Illkirch, France; ¹¹CMRDP, 5, rue Ibn Hazem, 1002 Tunis Beledère, Tunisia; ¹²Histologie Embryologie Cytogénétique CECOS, CHU Jean Verdier, 93143 BONDY, France; ¹³Centre d'AMP, Hôpital Maternité de Metz, 1/5 Place Sainte Croix, 57045 Metz, France

*Correspondence: pray@chu-grenoble.fr

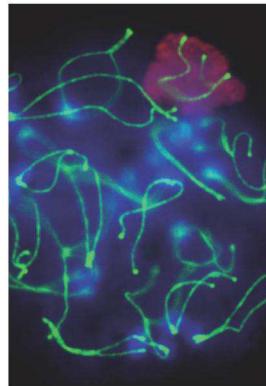
DOI 10.1016/j.ajhg.2011.02.007. ©2011 by The American Society of Human Genetics. All rights reserved.

► L'infertilité concerne de 10 % à 15 % des couples ayant un désir d'enfant et une composante masculine est retrouvée dans près de la moitié des cas. Dans une proportion importante de cas, une base génétique chromosomique ou génique connue, parfois transmise par les parents, est en cause. L'utilisation des nouvelles techniques de génotypage pangénomique a récemment permis d'impliquer de nouveaux gènes dans des phénotypes spécifiques d'infertilité masculine. Pour les patients concernés, ces avancées permettent de réaliser un diagnostic, d'affiner le pronostic et donc de mieux orienter la prise en charge. À terme, ce travail permettra de mieux comprendre les mécanismes moléculaires de la spermatogénèse et donc de proposer des solutions thérapeutiques qui pourront être applicables au plus grand nombre. ◀

Génétique de l'infertilité masculine

Les nouveaux acteurs

Charles Coutton^{1,2,3}, Véronique Satre^{1,2,3}, Christophe Arnoult^{1,2}, Pierre Ray^{1,2,4}



¹ Équipe génétique, infertilité et thérapeutiques, laboratoire AGIM (*ageing, imaging and modeling*), CNRS FRE3405, La Tronche, 38700, France ;

² Université Joseph Fourier, Grenoble, 38000, France ;

³ Centre hospitalo-universitaire (CHU) de Grenoble, unité fonctionnelle (UF) de génétique chromosomique, Grenoble, 38000, France ;

⁴ CHU de Grenoble, UF de biochimie et génétique moléculaire, Grenoble, 38000, France. pray@chu-grenoble.fr

d'azoospermies et d'oligozoospermies (respectivement absence totale de spermatozoïdes dans le sperme ou quantité anormalement réduite) à des anomalies chromosomiques ou à des mutations géniques touchant des gènes intervenant dans la production ou la fonction de cellules germinales. Par ailleurs, 30 % des infertilités restent inexpliquées et près de 40 % ont des causes incertaines. Ainsi, l'infertilité masculine d'origine génétique pourrait concerter près de 1 homme sur 40 [3]. Parmi les causes génétiques actuellement bien établies, on trouve les anomalies chromosomiques, les microdélétions du chromosome Y et les mutations du gène *CFTR* (*cystic fibrosis transmembrane conductance regulator*).

Anomalies chromosomiques

Des anomalies du nombre ou de la structure des autosomes, mais surtout des gonomosomes, peuvent être impliquées. Chez les patients infertiles, le pourcentage d'anomalies chromosomiques observées sur le caryotype fait à partir de cellules sanguines s'échelonne entre 2 et 8 %. Il peut atteindre 15 % chez les patients azoospermiques, soit 10 à 20 fois la fréquence retrouvée dans la population générale [4]. Le syndrome de Klinefelter (47, XXY) est la cause la plus fréquente d'hypogonadisme et d'infertilité chez l'homme. Sa prévalence est 50 fois plus élevée chez les patients infertiles azoospermiques (14 %) que dans la population générale (0,2 %) [5].

Les microdélétions du chromosome Y

Les microdélétions du chromosome Y représentent également une cause fréquente d'anomalies génétiques chez les hommes infertiles. Chez les patients présentant une azoospermie non obstructive, leur

Génétique et infertilité masculine : quelques chiffres

Selon l'OMS (organisation mondiale de la santé), on définit l'infertilité comme l'incapacité pour un couple de procréer après deux ans de rapports sexuels non protégés [1]. Il est généralement admis qu'environ 15 % des couples sont confrontés à des problèmes d'infertilité. Selon une étude publiée en 2007, la prévalence de l'infertilité serait toutefois plus proche de 9 %, ce qui concerne tout de même près de 70 millions de couples dans le monde [2]. L'infertilité est un problème majeur de santé publique et représente un enjeu médical et scientifique important.

Dans environ 50 % des cas, elle implique soit exclusivement l'homme, soit les deux membres du couple. Les causes d'infertilité masculine sont nombreuses et multifactorielles. On distingue des causes sécrétaires, les plus fréquentes, accompagnées d'un défaut de la spermatogenèse, et des causes excrétoires empêchant l'excration des spermatozoïdes. Un bilan clinique complet permet d'attribuer environ 30 % des cas



Development 139, 2955-2965 (2012) doi:10.1242/dev.077982
 © 2012. Published by The Company of Biologists Ltd

Absence of Dpy19l2, a new inner nuclear membrane protein, causes globozoospermia in mice by preventing the anchoring of the acrosome to the nucleus

Virginie Pierre^{1,2}, Guillaume Martinez^{1,2,3}, Charles Coutton^{1,2,4}, Julie Delaroche^{1,5}, Sandra Yassine^{1,2}, Caroline Novella^{1,2}, Karin Pernet-Gallay^{1,5}, Sylviane Hennebicq^{1,2,3}, Pierre F. Ray^{1,2,6,*} and Christophe Arnoult^{1,2,*,‡}

SUMMARY

Sperm-head elongation and acrosome formation, which take place during the last stages of spermatogenesis, are essential to produce competent spermatozoa that are able to cross the oocyte zona pellucida and to achieve fertilization. During acrosome biogenesis, acrosome attachment and spreading over the nucleus are still poorly understood and to date no proteins have been described to link the acrosome to the nucleus. We recently demonstrated that a deletion of *DPY19L2*, a gene coding for an uncharacterized protein, was responsible for a majority of cases of type I globozoospermia, a rare cause of male infertility that is characterized by the exclusive production of round-headed acrosomeless spermatozoa. Here, using *Dpy19l2* knockout mice, we describe the cellular function of the *Dpy19l2* protein. We demonstrate that the protein is expressed predominantly in spermatids with a very specific localization restricted to the inner nuclear membrane facing the acrosomal vesicle. We show that the absence of *Dpy19l2* leads to the destabilization of both the nuclear dense lamina (NDL) and the junction between the acroplaxome and the nuclear envelope. Consequently, the acrosome and the manchette fail to be linked to the nucleus leading to the disruption of vesicular trafficking, failure of sperm nuclear shaping and eventually to the elimination of the unbound acrosomal vesicle. Finally, we show for the first time that *Dpy19l3* proteins are also located in the inner nuclear envelope, therefore implying that the *Dpy19* proteins constitute a new family of structural transmembrane proteins of the nuclear envelope.

KEY WORDS: *Dpy19l2*, Acrosome biogenesis, Globozoospermia, Nuclear envelope, Nuclear lamina, Spermiogenesis, Mouse

INTRODUCTION

Spermatozoon is doubtless one of the most specialized cells in mammals and the description and understanding of the molecular aspects of spermiogenesis represents a very challenging task. The acrosome, a giant vesicle of secretion tightly bound to the nucleus via the acroplaxome (a network of proteins including keratin 5 and β-actin) (Kierszenbaum et al., 2003), is a highly specialized organelle found only in sperm. The molecular basis of acrosome biogenesis, and particularly its attachment and spreading over the nucleus, are poorly understood at molecular level; to date, there has been no report of any protein of the nuclear envelope (NE) anchoring the acroplaxome to the nuclear dense lamina (NDL) (Kierszenbaum et al., 2011). Knockout animal models presenting with spermiogenesis defects are interesting tools with which to characterize new actors of spermatid differentiation. Among teratozoospermia, globozoospermia is characterized by the production of round-headed acrosomeless spermatozoa and mouse strains presenting with such a defect represent very valuable models to decipher acrosome biogenesis. Globozoospermia was first described in human (Schirren

et al., 1971), and familial cases rapidly pointed to a genetic pathogenesis. In recent years *SPATA16* (Dam et al., 2007b) and *DPY19L2* were described to be involved in globozoospermia (Harbuz et al., 2011; Koscinski et al., 2011). *SPATA16* located in the Golgi apparatus is likely to be involved in vesicular trafficking necessary for acrosome biogenesis but its precise function remains uncharacterized (Lu et al., 2006; Xu et al., 2003). *DPY19L2* is a testis-specific member of an uncharacterized gene family, including four genes in mammals: *DPY19L1* to *L4* (Carson et al., 2006). The inactivation of *DPY19*, the ortholog of *DPY19L2* in *C. elegans*, was shown to block neuroblasts migration during the worm organogenesis (Honigberg and Kenyon, 2000). Its cellular localization and its physiological role have, however, remained elusive so far. We therefore decided to study *Dpy19l2* knockout mice (*Dpy19l2*^{-/-}) to unravel the function of *Dpy19l2*. In mice, a total of eight genes has been described to trigger globozoospermia. The first group of four proteins – Pick1 (Xiao et al., 2009), Gopc (Yao et al., 2002), Vps54 (Paiardi et al., 2011) and Hrb (Kang-Decker et al., 2001), as *SPATA16* – controls Golgi vesicles fusion necessary for acrosome formation. The second set of globozoospermia-inducing proteins comprises Zpbp1 (Lin et al., 2007), Ck2α' (Xu et al., 1999), Hsp90b1 (Audouard and Christians, 2011) and Gba2 (Yildiz et al., 2006), which have more diverse cellular localizations and functions: Zpbp1 and Ck2α' are proteins of the acrosomal matrix, whereas Hsp90b1 is expressed in the reticulum and Gba2 is expressed in both germ and Sertoli cells. Table S1 in the supplementary material summarizes the main characteristics of these mouse mutants. None perfectly mirrors the round-shaped acrosomeless spermatozoa observed in human type I globozoospermia, suggesting that *Dpy19l2* has an original cellular function during spermiogenesis.

¹Université Joseph Fourier, Grenoble F-38000, France. ²Equipe 'Génétique, Infertilité et Thérapeutiques' Laboratoire AGIM, CNRS FRE3405, La Tronche F-38700, France.
³CHU de Grenoble, Centre d'AMP-CECOS, BP217, Grenoble cedex 9 F-38043, France. ⁴CHU de Grenoble, UF de Génétique Chromosomique, Grenoble F-38000, France. ⁵Grenoble Institut des Neurosciences, INSERM U.836, Grenoble F-38000, France. ⁶CHU de Grenoble, UF de Biochimie et Génétique Moléculaire, Grenoble F-38000, France.

*Joint co-authors

†Author for correspondence (christophe.arnoult@ujf-grenoble.fr)



RESEARCH ARTICLE

Open Access

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

Thomas Karaouzène^{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



Open Access

ORIGINAL ARTICLE

Male Fertility

Mutations of the aurora kinase C gene causing macrozoospermia are the most frequent genetic cause of male infertility in Algerian men

Leyla Ounis^{1,2}, Abdelali Zoghmar², Charles Coutton^{3,4,5}, Leila Rouabah¹, Maroua Hachemi¹, Delphine Martinez⁶, Guillaume Martinez^{3,4}, Ines Bellil¹, Douadi Khelifi¹, Christophe Arnoult^{3,4}, Julien Faure^{3,6,7}, Sebti Benbouhedja², Abdelkader Rouabah¹, Pierre F Ray^{3,4,6}

Klinefelter syndrome and Y-chromosomal microdeletion analyses were once the only two genetic tests offered to infertile men. Analyses of aurora kinase C (*AURKC*) and *DPY19L2* are now recommended for patients presenting macrozoospermia and globozoospermia, respectively, two rare forms of teratozoospermia particularly frequent among North African men. We carried out genetic analyses on Algerian patients, to evaluate the prevalence of these syndromes in this population and to compare it with the expected frequency of Klinefelter syndrome and Y-microdeletions. We carried out a retrospective study on 599 consecutive patients consulting for couple infertility at the assisted reproduction unit of the Ibn Rochd Clinique, Constantine, Algeria. Abnormal sperm parameters were observed in 404 men. Fourteen and seven men had typical macrozoospermia and globozoospermia profiles, respectively. Molecular diagnosis was carried out for these patients, for the *AURKC* and *DPY19L2* genes. Eleven men with macrozoospermia had a homozygous *AURKC* mutation (79%), corresponding to 2.7% of all patients with abnormal spermograms. All the men with globozoospermia studied ($n = 5$), corresponding to 1.2% of all infertile men, presented a homozygous *DPY19L2* deletion. By comparison, we would expect 1.6% of the patients in this cohort to have Klinefelter syndrome and 0.23% to have Y-microdeletion. Our findings thus indicate that *AURKC* mutations are more frequent than Klinefelter syndrome and constitute the leading genetic cause of infertility in North African men. Furthermore, we estimate that *AURKC* and *DPY19L2* molecular defects are 10 and 5 times more frequent, respectively, than Y-microdeletions.

Asian Journal of Andrology (2015) **17**, 68–73; doi: 10.4103/1008-682X.136441; published online: 22 August 2014

Keywords: aurora kinase C; *DPY19L2*; globozoospermia; intracytoplasmic sperm injection; infertility; macrozoospermia

INTRODUCTION

At least 70 million couples worldwide have infertility problems.¹ Male infertility has diverse, often multifactorial causes, resulting in quantitative and/or qualitative sperm defects in 61% of cases.² A large proportion of male infertility cases are caused by genetic defects, but few genes have been formally shown to be associated with spermatogenic defects since the discovery of microdeletions in the 1970s.³ Aurora kinase C (*AURKC*)⁴ and *DPY19L2*⁵ are the two principal genes implicated in sperm defects. They have been found to be mutated in large numbers of unrelated individuals presenting macrozoospermia and globozoospermia, respectively. The abnormal spermatozoa of men with these conditions display typical morphological abnormalities (Figure 1).

Patients with macrozoospermia produce ejaculates containing mostly large-headed multiflagellar polyploid spermatozoa. The c.144delC deletion has been identified in most North African patients with this condition.⁶ Further studies led to the identification of rare

familial mutations⁷ and of a nonsense mutation, p.Y248*, in both European and North African men.⁸ An overall carriage frequency for these defects of 1/50 has been established for individuals from the general population of the Maghreb.⁶ We have shown that the large-headed spermatozoa of *AURKC*-deficient patients are tetraploid, indicating that meiosis cannot be completed without a functional *AURKC* protein.⁶ This confirms that men with homozygous *AURKC* mutations are unlikely to be able to father children. A positive genetic diagnosis thus provides a formal contra-indication for *in vitro* fertilization by intracytoplasmic sperm injection (ICSI).⁶

Globozoospermia is another severe sperm defect leading both to primary infertility and to poor fertilization and embryo growth following ICSI. This phenotype is characterized by the production of small, round-headed spermatozoa with no acrosome. Mutations of *SPATA16*⁹ and *PICK1*¹⁰ have been identified in individuals with a family history of the globozoospermia, suggesting a role of these genes in the phenotype, although they cannot be the main cause of

¹Laboratoire de Biologie Moléculaire et Cellulaire, Faculté des Sciences de la Nature et de la Vie, Université Constantine I, Algérie; ²Clinique Ibn Rochd, Centre de Médecine de la Reproduction, Constantine, Algérie; ³Université Grenoble Alpes, Grenoble, France; ⁴Equipe Génétique Epigénétique et Thérapies de l'Infertilité, CNRS, AGIM, Grenoble, France; ⁵Laboratoire de Génétique Chromosomique, CHU Grenoble, Grenoble, France; ⁶Laboratoire de Biochimie et Génétique Moléculaire, CHU Grenoble, Grenoble, France; ⁷Equipe Muscle et Pathologies, INSERM U836, Grenoble Institut des Neurosciences, Grenoble, France.

Correspondence: Dr. PF Ray (pray@chu-grenoble.fr)

Received: 20 January 2014; Revised: 07 March 2014; Accepted: 14 May 2014

Subcellular localization of phospholipase C ζ in human sperm and its absence in DPY19L2-deficient sperm are consistent with its role in oocyte activation

Jessica Escoffier^{1,2,†}, Sandra Yassine^{1,2,†}, Hoi Chang Lee³, Guillaume Martinez^{1,2}, Julie Delaroche^{1,4}, Charles Coutton^{1,2,5}, Thomas Karaouzène^{1,2}, Raoudha Zouari⁶, Catherine Metzler-Guillemain^{7,8}, Karin Pernet-Gallay^{1,4}, Sylviane Hennebicq^{1,9}, Pierre F. Ray^{1,2,10}, Rafael Fissore³, and Christophe Arnoult^{1,2,*}

¹Université Grenoble Alpes, Grenoble F-38000, France ²Equipe 'Andrologie, Génétique et Cancer' Laboratoire AGIM, CNRS FRE3405, La Tronche F-38700, France ³Department of Veterinary and Animal Sciences, University of Massachusetts, Amherst, 661 North Pleasant Street, Amherst, MA 01003, USA ⁴Grenoble Institut des Neurosciences, INSERM U.836, F-38000 Grenoble, France ⁵CHU de Grenoble, UF de Génétique Chromosomique, Grenoble F-38000, France ⁶Clinique des Jasmins, 23, Av. Louis BRAILLE, 1002 Tunis, Tunisia ⁷Aix-Marseille Université-Inserm UMR 910, Génétique Médicale et Génomique Fonctionnelle, I3385 Marseille Cedex 5, France ⁸APHM Hôpital La Conception, Gynépôle, Laboratoire de Biologie de la Reproduction - CECOS, I3385 Marseille Cedex 5, France ⁹CHU de Grenoble, Centre d'AMP-CECOS, BP217, Grenoble Cedex 9 F-38043, France ¹⁰CHU de Grenoble, UF de Biochimie et Génétique Moléculaire, Grenoble F-38000, France

*Correspondence address. AGIM, Equipe 'Andrologie, Génétique et Cancer', Faculté de Médecine, 38700 La Tronche, France.
Tel: +33-4-76-63-74-08; E-mail christophe.arnoult@ujf-grenoble.fr

Submitted on August 28, 2014; resubmitted on October 13, 2014; accepted on October 24, 2014

ABSTRACT: We recently identified the DPY19L2 gene as the main genetic cause of human globozoospermia (70%) and described that *Dpy19l2* knockout (KO) mice faithfully reproduce the human phenotype of globozoospermia making it an excellent model to characterize the molecular physiopathology of globozoospermia. Recent case studies on non-genetically characterized men with globozoospermia showed that phospholipase C, zeta (PLC ζ), the sperm factor thought to induce the Ca²⁺ oscillations at fertilization, was absent from their sperm, explaining the poor fertilization potential of these spermatozoa. Since 30% of globozoospermic men remain genetically uncharacterized, the absence of PLC ζ in DPY19L2 globozoospermic men remains to be formally established. Moreover, the precise localization of PLC ζ and the reasons underlying its loss during spermatogenesis in globozoospermic patients are still not understood. Herein, we show that PLC ζ is absent, or its presence highly reduced, in human and mouse sperm with DPY19L2-associated globozoospermia. As a consequence, fertilization with sperm from *Dpy19l2* KO mice failed to initiate Ca²⁺ oscillations and injected oocytes remained arrested at the metaphase II stage, although a few human oocytes injected with DPY19L2-defective sperm showed formation of 2-pronuclei embryos. We report for the first time the subcellular localization of PLC ζ in control human sperm, which is along the inner acrosomal membrane and in the perinuclear theca, in the area corresponding to the equatorial region. Because these cellular components are absent in globozoospermic sperm, the loss of PLC ζ in globozoospermic sperm is thus consistent and reinforces the role of PLC ζ as an oocyte activation factor necessary for oocyte activation. In our companion article, we showed that chromatin compaction during spermiogenesis in *Dpy19l2* KO mouse is defective and leads to sperm DNA damage. Together, these defects explain the poor fertilization potential of DPY19L2-globozoospermic sperm and the compromised developmental potential of embryos obtained using sperm from patients with a deletion of the DPY19L2 gene.

Key words: male infertility / globozoospermia / DPY19L2 / phospholipase C zeta / acrosome / ICSI

† Shared first authorship.

Dpy19l2-deficient globozoospermic sperm display altered genome packaging and DNA damage that compromises the initiation of embryo development

Sandra Yassine^{1,2,†}, Jessica Escoffier^{1,2,†}, Guillaume Martinez^{1,2}, Charles Coutton^{1,2,3}, Thomas Karaouzène^{1,2}, Raoudha Zouari⁴, Jean-Luc Ravanat^{1,5}, Catherine Metzler-Guillemain^{6,7}, Hoi Chang Lee⁸, Rafael Fissore⁸, Sylviane Hennebicq^{1,9}, Pierre F. Ray^{1,2,10}, and Christophe Arnoult^{1,2,*}

¹Université Grenoble Alpes, Grenoble, F-38000, France ²Equipe 'Andrologie, Génétique et Cancer' Laboratoire AGIM, CNRS FRE3405, La Tronche, F-38700, France ³CHU de Grenoble, UF de Génétique Chromosomique, Grenoble, F-38000, France ⁴Clinique des Jasmins, 23, Av. Louis BRAILLE, 1002 Tunis, Tunisia ⁵Laboratoire Lésions des Acides Nucléiques, CEA, INAC-SCIB, F-38000 Grenoble, France

⁶Aix-Marseille Université-Inserm UMR 910, Génétique médicale et Génomique Fonctionnelle, 13385 Marseille Cedex 5, France

⁷APHM Hôpital La Conception, Gynépôle, Laboratoire de Biologie de la Reproduction – CECOS, 13385 Marseille Cedex 5, France

⁸Department of Veterinary and Animal Sciences, University of Massachusetts, Amherst, 661 North Pleasant Street, Amherst, MA 01003, USA

⁹CHU de Grenoble, Centre d'AMP-CECOS, BP217, Grenoble Cedex 9, F-38043, France ¹⁰CHU de Grenoble, UF de Biochimie et Génétique Moléculaire, Grenoble, F-38000, France

*Correspondence address: AGIM, Equipe 'Andrologie, Génétique et Cancer', Faculté de Médecine, 38700 La Tronche, France.
Tel: +33-4-76-63-74-08; E-mail: christophe.arnoult@ujf-grenoble.fr

Submitted on August 28, 2014; resubmitted on October 13, 2014; accepted on October 24, 2014

ABSTRACT: We recently identified the DPY19L2 gene as the main genetic cause of human globozoospermia. Non-genetically characterized cases of globozoospermia were associated with DNA alterations, suggesting that DPY19L2-dependent globozoospermia may be associated with poor DNA quality. However the origins of such defects have not yet been characterized and the consequences on the quality of embryos generated with globozoospermic sperm remain to be determined. Using the mouse model lacking Dpy19l2, we compared several key steps of nuclear compaction. We show that the kinetics of appearance and disappearance of the histone H4 acetylation waves and of transition proteins are defective. More importantly, the nuclear invasion by protamines does not occur. As a consequence, we showed that globozoospermic sperm presented with poor sperm chromatin compaction and sperm DNA integrity breakdown. We next assessed the developmental consequences of using such faulty sperm by performing ICSI. We showed in the companion article that oocyte activation (OA) with globozoospermic sperm is very poor and due to the absence of phospholipase C ζ ; therefore artificial OA (AOA) was used to bypass defective OA. Herein, we evaluated the developmental potential of embryos generated by ICSI + AOA in mice. We demonstrate that although OA was fully rescued, preimplantation development was impaired when using globozoospermic sperm. In human, a small number of embryos could be generated with sperm from DPY19L2-deleted patients in the absence of AOA and these embryos also showed a poor developmental potential. In conclusion, we show that chromatin compaction during spermiogenesis in *Dpy19l2* KO mouse is defective and leads to sperm DNA damage. Most of the DNA breaks were already present when the sperm reached the epididymis, indicating that they occurred inside the testis. This result thus suggests that testicular sperm extraction in Dpy19l2-dependent globozoospermia is not recommended. These defects may largely explain the poor embryonic development of most mouse and human embryos obtained with globozoospermic sperm.

Key words: male infertility / globozoospermia / Dpy19l2 / DNA compaction / protamine

† Shared first authorship.