



HAL
open science

Genetic analyses of a large cohort of infertile patients with globozoospermia, DPY19L2 still the main actor, GGN confirmed as a guest player

Tristan Celse, Caroline Cazin, Flore Mietton, Guillaume Martinez, Delphine Martinez, Nicolas Thierry-Mieg, Amandine Septier, Catherine Metzler-Guillemain, Julie Beurois, Antoine Clergeau, et al.

► To cite this version:

Tristan Celse, Caroline Cazin, Flore Mietton, Guillaume Martinez, Delphine Martinez, et al.. Genetic analyses of a large cohort of infertile patients with globozoospermia, DPY19L2 still the main actor, GGN confirmed as a guest player. *Human Genetics*, 2021, *Molecular Genetics of Male Infertility*, 140 (1), pp.43-57. 10.1007/s00439-020-02229-0 . hal-03025179

HAL Id: hal-03025179

<https://hal.science/hal-03025179v1>

Submitted on 11 Mar 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Genetic analyses of a large cohort of infertile patients with globozoospermia, *DPY19L2* still the main actor, *GGN* confirmed as a guest player

Tristan Celse^{1,2} · Caroline Cazin^{1,2} · Flore Mietton² · Guillaume Martinez^{1,3} · Delphine Martinez² · Nicolas Thierry-Mieg⁴ · Amandine Septier⁴ · Catherine Guillemain^{5,6} · Julie Beurois¹ · Antoine Clergeau⁷ · Selima Fourati Ben Mustapha⁸ · Mahmoud Kharouf⁸ · Abdelali Zoghmar⁹ · Ahmed Chargui¹⁰ · Aline Papaxanthos¹¹ · Béatrice Dorphin¹² · Bernard Foliguet¹³ · Chema Triki¹⁴ · Christophe Sifer¹⁵ · Dominique Lauton¹⁶ · Gérard Tachdjian¹⁷ · Gilles Schuler¹⁸ · Hervé Lejeune¹⁹ · Jacques Puechberty²⁰ · Julien Bessonnat²¹ · Laurent Pasquier²² · Lionel Mery²³ · Marine Poulain²⁴ · Myriam Chaabouni⁸ · Nathalie Sermondade²⁵ · Rosalie Cabry²⁶ · Sebti Benbouhadja⁹ · Ségolène Veau²⁷ · Cynthia Frapsauce²⁸ · Valérie Mitchell²⁹ · Vincent Achard^{30,31,32} · Veronique Satre^{1,3} · Sylviane Hennebicq^{1,21} · Raoudha Zouari⁵ · Christophe Arnoult¹ · Zine-Eddine Kherraf^{1,2} · Charles Coutton^{1,3} · Pierre F. Ray^{1,2}

Abstract

Globozoospermia is a rare phenotype of primary male infertility inducing the production of round-headed spermatozoa without acrosome. Anomalies of *DPY19L2* account for 50–70% of all cases and the entire deletion of the gene is by far the most frequent defect identified. Here, we present a large cohort of 69 patients with 20–100% of globozoospermia. Genetic analyses including multiplex ligation-dependent probe amplification, Sanger sequencing and whole-exome sequencing identified 25 subjects with a homozygous *DPY19L2* deletion (36%) and 14 carrying other *DPY19L2* defects (20%). Overall, 11 deleterious single-nucleotide variants were identified including eight novel and three already published mutations. Patients with a higher rate of round-headed spermatozoa were more often diagnosed and had a higher proportion of loss of function anomalies, highlighting a good genotype phenotype correlation. No gene defects were identified in patients carrying < 50% of globozoospermia while diagnosis efficiency rose to 77% for patients with > 50% of globozoospermia. In addition, results from whole-exome sequencing were scrutinized for 23 patients with a *DPY19L2* negative diagnosis, searching for deleterious variants in the nine other genes described to be associated with globozoospermia in human (*C2CD6*, *C7orf61*, *CCDC62*, *CCIN*, *DNAH17*, *GGN*, *PICK1*, *SPATA16*, and *ZBP1*). Only one homozygous novel truncating variant was identified in the *GGN* gene in one patient, confirming the association of *GGN* with globozoospermia. In view of these results, we propose a novel diagnostic strategy focusing on patients with at least 50% of globozoospermia and based on a classical qualitative PCR to detect *DPY19L2* homozygous deletions. In the absence of the latter, we recommend to perform whole-exome sequencing to search for defects in *DPY19L2* as well as in the other previously described candidate genes.

Introduction

Infertility is a major health concern, affecting approximately 50 million couples worldwide (Mascarenhas et al. 2012), or 12.5% of women and 10% of men (Datta et al. 2016). Among the factors leading to male infertility, alterations of spermatogenesis are the major cause including quantitative defects (azoo- or oligozoospermia) or qualitative defects (teratozoospermia or asthenozoospermia). Globozoospermia 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 without acrosome (#MIM102530). Globozoospermic sperm are thus unable to adhere and to penetrate the zona pellucida, causing primary infertility. It is important to differentiate total globozoospermia referring to patients with a homogeneous phenotype with ~100% round-headed sperm and partial globozoospermia with a variable percentage of spermatozoa of typical shape (Fig. 1). A minimum threshold of 20–50% of round-headed spermatozoa is commonly used in the literature to confirm the diagnosis of globozoospermia (Dam et al. 2007, 2011; Modarres et al. 2018; Oud et al. 2020).

Genetics analyses permitted to decipher the different factors involved in the pathogenesis of globozoospermia. Alterations of many genes encoding for proteins involved in several critical steps of acrosome biogenesis such as the vesicle transport from the Golgi to the acrosome, vesicle fusion, and membrane interaction, have been demonstrated to be responsible for globozoospermia or globozoospermia-like phenotypes (Coutton et al. 2015). As the phenotype is very severe and specific, all cases of total globozoospermia are believed to have a genetic cause. To date, mutations in *C2CD6*, *C7orf61*, *CCDC62*, *CCIN*, *DNAH17*, *GGN*, *PICK1*, *SPATA16* and *ZBP1* have all been described in globozoospermia patients, but only a handful of patients carrying variants in any of these genes have been described (Dam et al. 2007; Liu et al. 2010; Yatsenko et al. 2012; Elinati et al. 2016; Oud et al. 2020). The main genetic causes of total globozoospermia are alterations of the *DPY19L2* gene which are found in approximately two-thirds of globozoospermic patients (Coutton et al. 2015).

DPY19L2, located in 12q14.2, is predominantly expressed in the testis and encodes a transmembrane protein which is part of the DPY19 protein family. The study of wild type

and homozygous knock out (KO) *Dpy19l2* male mice demonstrated that the protein is present from the round spermatid stage onwards and that it localizes to the inner nuclear membrane, exclusively in regards to the forming acrosome (Pierre et al. 2012). In the absence of DPY19L2, the forming acrosome separates from the nucleus before being totally removed from the sperm with the cytoplasmic droplets, demonstrating that DPY19L2 is necessary to anchor the acrosome to the nucleus (Pierre et al. 2012). In addition to its structural function during acrosome biogenesis, the C-mannosyltransferase function of its ancestral protein DPY-19 has raised the hypothesis that DPY19L2 may have a function in glycosylation of sperm proteins but this remains to be demonstrated (Buettner et al. 2013). A recurrent 200-kb homozygous deletion is the most frequent event affecting *DPY19L2*, identified in a variable proportions of globozoospermic patients ranging from 19% (Koscinski et al. 2011) to 75% (Harbuz et al. 2011). This variability may be explained in part by the geographical origins of the studied cohorts of patients, the degrees of consanguinity of the studied patients but mainly by the inclusion of different proportion of patients with partial globozoospermia (Ray et al. 2017). The mechanism leading to the deletion is, however, consensual and was described to be due to a non-allelic homozygous recombination (NAHR) occurring between two homologous 28-kb low copy repeats (LCRs) located on each side of the gene (Harbuz et al. 2011; Koscinski et al. 2011; Elinati et al. 2012; Coutton et al. 2013). Many point mutations and small deletions have also been described as causal (Elinati et al. 2012; Modarres et al. 2016; Chianese et al. 2015; Shang et al. 2019; Coutton et al. 2012b; Zhu et al. 2013; Ghédir et al. 2016; Oud et al. 2020). Overall, a total of 22 deleterious variants has been described including five

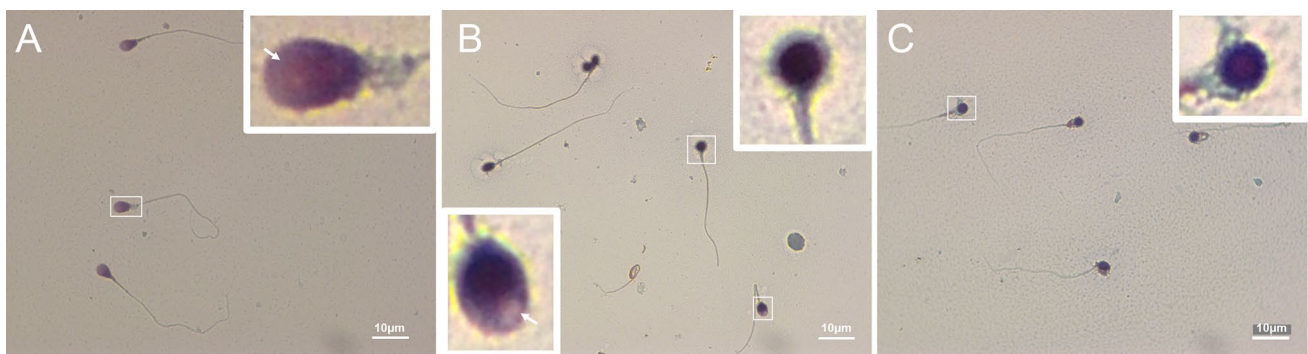


Fig. 1 Optical microscopic observations of spermatozoa. Scale=10 μ M. The acrosome is at the tip of the white arrow. **a** Semen of a control patient with only normal spermatozoa (absent from the cohort). Zoom on a spermatozoon's head of typical shape (white square in the upper right corner). The acrosome can be identified as the white halo at the front of the head. The nucleus is the black part closer to the intermediate piece. **b** Semen from a patient with partial globozoospermia. Zoom on a round headed spermatozoon

without acrosome, the shape of the head follows the shape of the nucleus (white square in the upper right corner). Zoom on a spermatozoon with an intermediate shape close to the normal spermatozoon in image a, but with a smaller irregular acrosome at the front (white square in the lower left corner). **c** Semen of a patient with total globozoospermia containing only globozoospermic spermatozoa. Zoom on a round headed spermatozoon without acrosome, similar to the one on image **b** (white square in the upper right corner)

splice variants, nine loss of function variants, a deletion of three nucleotides and seven missense variants (Fig. 2). All of them represent approximately 20% of the pathological alleles (Ray et al. 2017). Based on these data, the consensual diagnosis strategy for patients presenting a total or partial globozoospermia is to first screen for the presence of the *DPY19L2* gene deletion before searching for *DPY19L2* point mutations and then defects in other candidate genes.

Here we present the genetic results from a large cohort of 69 patients presenting with globozoospermia. We performed a first-line screening using MLPA for all patients allowing to quantify the number of *DPY19L2* alleles present in each patient. In the absence of a homozygous deletion of the whole gene, we subsequently performed either Sanger sequencing of the 22 *DPY19L2* exons and/or whole-exome

sequencing (WES). We identified eight novel homo- or hemizygous point mutations, three causal variants already reported in other publications and two heterozygous variants of unknown significance. Then we continued the analysis in 23 patients without a causal *DPY19L2* anomaly and for which we performed a whole exome sequencing allowing us to look for variants in the genes described to be associated with globozoospermia in Human (*C2CD6*, *C7orf61*, *CCDC62*, *CCIN*, *DNAH17*, *GGN*, *PICK1*, *SPATA16* and *ZPBP1*). For these nine genes analysed in 23 subjects, we only identified one deleterious variant present in a single patient. The variant was a homozygous loss of function variant affecting the *GGN* gene. We also compared the sperm parameters in our cohort of patients according to their *DPY19L2* genotype to explore a potential

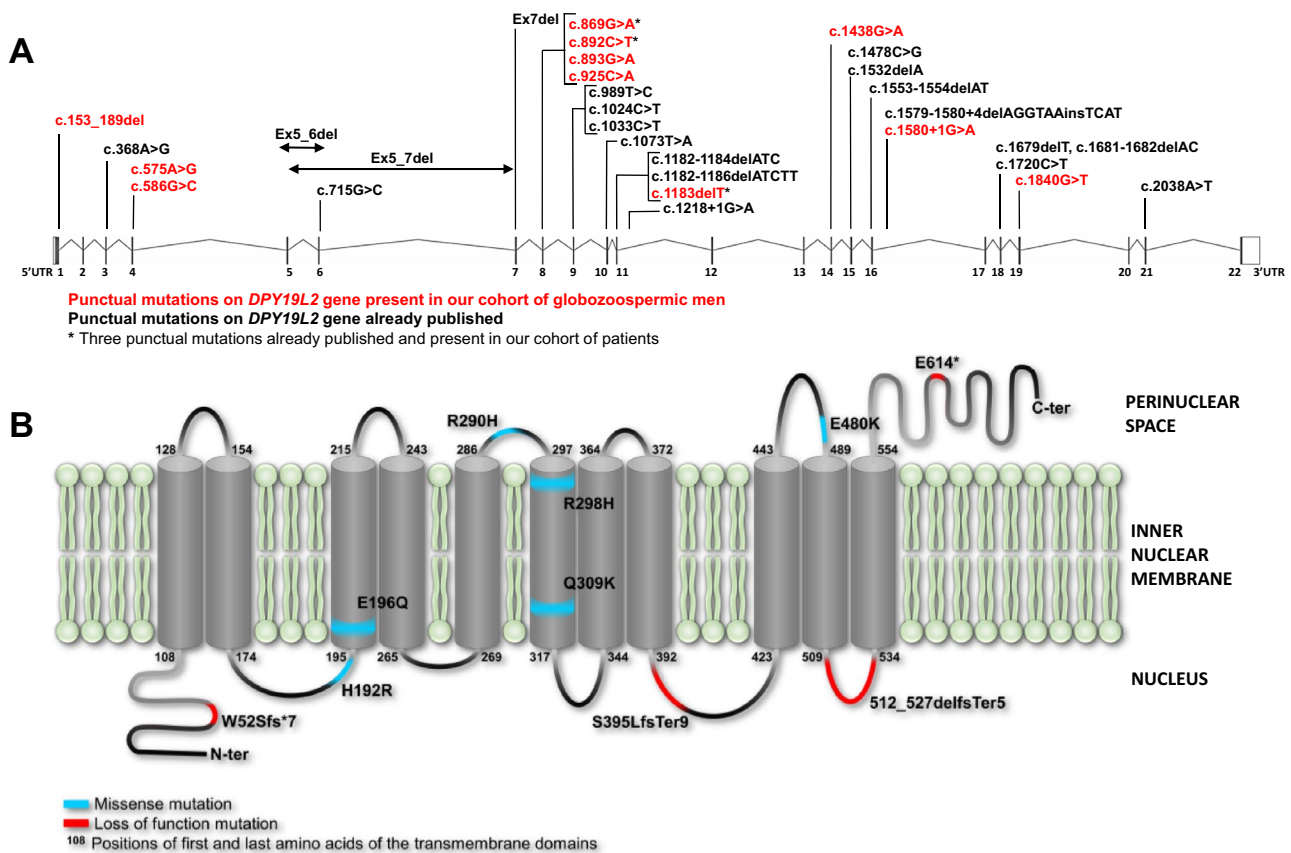


Fig. 2 Location of the point mutations found on the *DPY19L2* gene and their consequence on the *DPY19L2* protein. **a** Location of the 30 deleterious variants present on *DPY19L2* including six splice variants, 11 loss of function variants, a deletion of three nucleotides and 12 missense variants. Exons are indicated as a black box, untranslated region as a clear box, introns as a line, the localization of the identified point mutations is shown by a line and the span of the identified genomic deletions is indicated by a black line with two arrows at the end. The numbers under the boxes indicate the exon number. Point mutations identified in our present work among 69 globozoospermic patients are represented in red and bold. Point mutations in

319 already published patients are represented in black and bold. An asterisk marks the three variants present in our cohort and already present in the literature. **b** Representation of the *DPY19L2* protein and its eleven transmembrane domains with first and last transmembrane domains amino-acid (black numbers), our missense mutations (blue segment) and truncating mutations (red segment). Exon 8 contains four different missense mutations, impacting 56.5% of all published patients with a causal missense mutation on the third extramembrane domain in the perinuclear space and the sixth transmembrane domain [a Adapted from Zhu et al. (2013)]

genotype–phenotype correlation. This work permitted to identify new deleterious variants and to refine the current diagnostic strategy for globozoospermia.

Materials and methods

Patients

We recruited 73 patients, all addressed to Grenoble Hospital between 2012 and 2019 for the genetic investigation of globozoospermia. All patients had a medical consultation for infertility and a sperm analysis revealing complete or partial globozoospermia. Informed consent was obtained from all the patients participating in the study according to local protocols and the principles of the Declaration of Helsinki. The study was approved by local ethics committees, and samples were then stored in the CRB

Germethèque (certification under ISO-9001 and NF-S 96-900) following a standardized procedure or were part of the Fertithèque collection declared to the French Ministry of health (DC-2015-2580) and the French Data Protection Authority (DR-2016-392). Most patients originated from France (43/73 patients) or North Africa ($n = 26$ including 13 from Tunisia, 12 from Algeria and one from Morocco) but also from Turkey ($n = 3$), Cape-Verde ($n = 1$), Iraq ($n = 1$). All patients were unrelated and unpublished globozoospermic patients except for three Algerian patients who are brothers. Within our cohort, ten patients reported consanguinity in their family. Their geographical origin is representative of the cohort with five patients from France, three patients from North Africa (two Tunisians and one Algerian), one patient from Iraq and finally one of Turkey. We excluded patients with less than 20% of round-headed spermatozoa without acrosome and patients with a sperm concentration under 1 million/mL (Fig. 3). Overall 69 patients were analysed and are presented here.

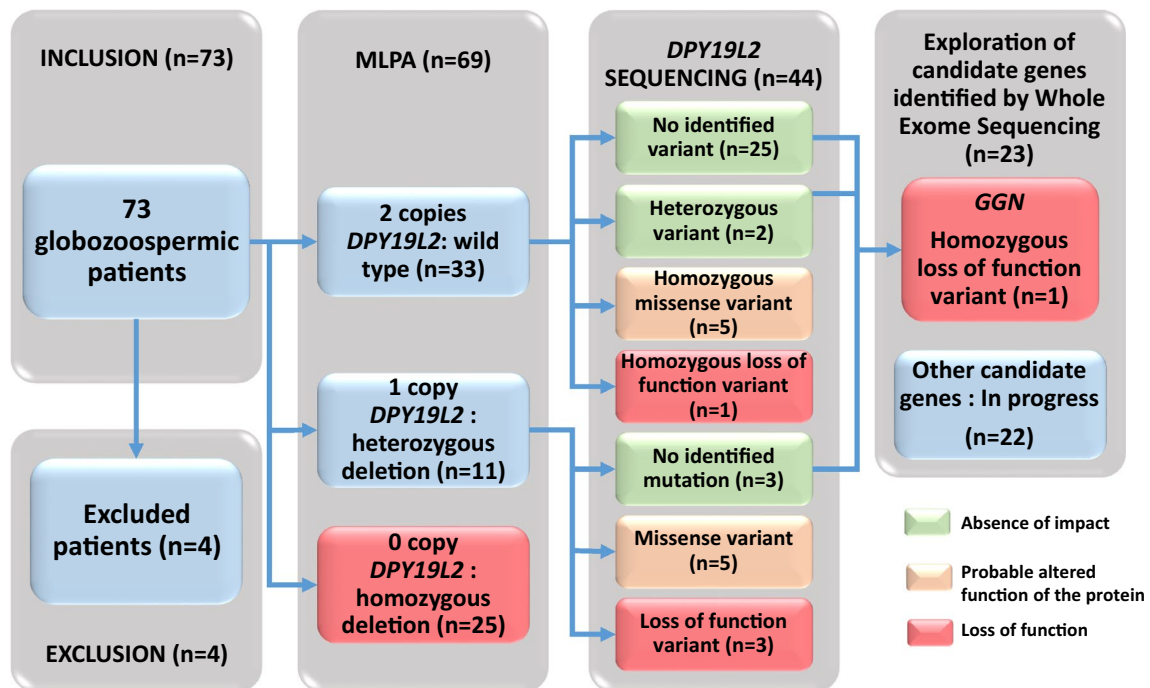


Fig. 3 Schematic representation of the molecular diagnostic investigations carried out on the cohort of 69 patients with globozoospermia. Seventy-three patients were recruited, we excluded four subjects because of a low rate of round-headed spermatozoa ($n = 1$) or a low concentration of sperm in the ejaculate ($n = 3$). We performed MLPA on 69 patients and diagnosed 25 homozygous *DPY19L2* deletions. We carried out Sanger sequencing of *DPY19L2* 22 exons in 11 heterozygous patients and diagnosed five patients with a missense variant and three patients with loss of function variant on their only allele. Among 33 patients without any *DPY19L2* deletion detected by MLPA, we detected one patient with a homozygous loss of func-

tion variant, five patients with a missense variant and two patients with a single heterozygous variant. Twenty-five patients presented no anomalies in *DPY19L2*. Among the 30 patients without any detected *DPY19L2* causal anomaly, we performed whole-exome sequencing for 23 of them and detected only one homozygous loss of function variant in *GGN* in one patient. Gene defects were only search in the genes already associated with globozoospermia in Human (*C2CD6*, *C7orf61*, *CCDC62*, *CCIN*, *DNAH17*, *GGN*, *PICK1*, *SPATA16* and *ZPBP1*). For the 22 other patients, no deleterious variants were identified in these genes and investigations are pursued to identify new globozoospermia candidate genes

Semen analysis

Sperm analysis was performed during the routine biological examination of the patients according to World Health Organization Guidelines (Cooper et al. 2010). It was carried out in different source laboratories and protocol variations cannot be excluded. Sperm parameters from the different groups of patients were compared according to their genotype, as described in Table 1. A two-tailed *t* test analyzing was employed to identify significant differences between patient groups using GraphPad Prism version 8.4.2 for Windows, GraphPad Software, San Diego, CA, USA, <https://www.graphpad.com>.

DNA extraction

DNA was extracted from blood and saliva samples. Saliva was collected with ORAGEN DNA OG-500 kits from DNA GENOTEK Inc and extraction was performed using the manufacturer's recommendation. Concerning blood samples, DNA was isolated from EDTA blood using the DNeasy Blood & Tissue Kits from QIAGEN SA (Courtaboeuf, France).

MLPA analysis

MLPA analysis uses probes specific of exons 1, 17 and 22 according to our protocol, already described by Coutton et al. (2012a, b). We realized it to investigate *DPY19L2* entire gene deletion in 69 patients.

Whole-exome sequencing and bioinformatics analyses

Whole-exome sequencing was performed for 28 out of 69 patients without *DPY19L2* homozygous deletion. Coding regions and intron/exon boundaries were sequenced after enrichment using SureSelect Human All Exon V6—from Agilent. An alignment-ready GRCh38 reference genome (including ALT, decoy and HLA) was produced using “run-gen-ref hs38DH” from Heng Li's bwa kit package (<https://github.com/lh3/bwa>). The exomes were analyzed using a bioinformatics pipeline developed in-house. The pipeline consists of two modules, both distributed under the GNU General Public License v3.0 and available on github. The first module (<https://github.com/ntm/grexome-TIMC-Prima>) takes FASTQ files as input and produces a single merged GVCF file, as follows. Adaptors are trimmed and low-quality reads filtered with fastp 0.20.0 (Chen et al. 2018), reads are aligned with BWA-MEM 0.7.17 (Li 2013), duplicates are

Table 1 Comparison of sperm parameters between groups according to the patients' genotype

Patients	Group 1: loss of function anomaly (<i>n</i> = 30)		Group 2: hemi- or homozygous missense mutation (<i>n</i> = 10)		Group 3: absence of causal anomaly (<i>n</i> = 29)	
	Mean	SD	Mean	SD	Mean	SD
Sperm volume (mL)	3.42	1.32	3.07	1.78	3.64	1.53
Nbspz ($\times 10^6$ /mL)	73.18	47.95	49.34	56.20	33.03	38.22
Vitality	68.67	14.43	62.86	16.70	53.70	16.37
Motility (A + B. 1 h)	30.62	10.94	20.13	9.78	23.50	16.86
Round cells ($\times 10^6$ /mL)	3.22	4.26	4.38	7.73	3.18	3.08
Normal spermatozoa	0.06	0.24	0.00	0.00	2.08	3.60
Intermediate piece angulation	16.00	10.19	17.33	9.78	24.04	14.16
Rolled flagella	14.07	8.81	21.17	11.98	12.12	9.10
Flagella of irregular caliber	2.93	6.42	1.00	0.89	2.77	8.54
Shortened flagella	0.93	1.34	1.80	2.71	3.15	2.80
Absence of flagella	1.43	2.06	4.00	3.90	3.12	4.08
Multiple flagella	1.29	1.58	1.67	1.37	2.12	3.57
Multiple heads	1.43	1.72	2.67	4.68	2.85	3.59
Acrosome anomaly	98.96	3.21	97.75	3.67	92.75	11.14
Round head	94.33	11.51	89.33	17.26	60.08	29.43
Microcephaly	81.79	32.62	68.67	26.66	40.77	28.72
Base anomaly	23.23	34.96	23.00	24.80	47.48	31.31
Multiple anomalies index	2.62	0.59	2.35	0.54	2.66	0.58

Values are expressed in percent, unless specified otherwise
SD standard deviation

marked using samblaster 0.1.24 files are sorted and indexed with samtools 1.9 (Li et al. 2009). SNVs and short indels are called from each BAM file using strelka 2.9.10 (Kim et al. 2018) to produce individual GVCF files. These are finally merged with merge GVCFs_strelka.pl to obtain a single multi-sample GVCF, which combines all exomes available in our laboratory. The second module (<https://github.com/ntm/grexome-TIMC-Secondary>) takes this merged GVCF as input and produces annotated analysis-ready TSV files. This is achieved by performing up to 15 streamlined tasks, including the following. Low-quality variant calls (DP < 10, GQX < 20, or less than 15% of reads supporting the ALT allele) are discarded. Variant Effect Predictor v92 (McLaren et al. 2016) is used to annotate the variants and predict their impact, allowing to filter low-impact variants and/or prioritize high-impact ones (e.g. stop-gain or frameshift variants). Gene expression data from the Genotype-Tissue Expression project (GTEx v7) are added. Variants with a minor allele frequency greater than 1% in gnomAD v2.0, 3% in 1000 Genomes Project phase 3, or 5% in NHLBI ESP6500 are filtered. Variants are also compared to those obtained from 250 exomes of healthy control individuals or of patients presenting a clearly different phenotype. Because all variants result from the same bioinformatics pipeline, this allows to filter artifacts due to the pipeline itself, as well as genuine variants that may be missing from public databases but are actually not so rare in our cohorts. Finally, the resulting TSV files can be opened with spreadsheet software such as LibreOffice Calc or Microsoft Excel for further filtering and sorting, to identify candidate causal variants.

Candidate variants identified in *DPY19L2* and other genes were subsequently confirmed by Sanger sequencing using an Applied Biosystems 3500XL Genetic Analyzer. Analyses were performed according to the protocol described below.

***DPY19L2* Sanger sequencing**

Full *DPY19L2* Sanger sequencing was performed for 16 patients including five patients without *DPY19L2* deletion and 11 patients with a heterozygous deletion.

The 22 *DPY19L2* exons and intronic boundaries were amplified using the PCR primers described in supplementary Table 1. Sequencing reactions were performed using the BigDyeTerminator v3.1 sequencing kit (Applied Biosystems) and sequence analyses were carried out on an ABI3500XL Genetic Analyzer (Applied Biosystems). Sequences were analyzed using the Seqscape software (Applied Biosystems). The nomenclature of the identified variants was established according to Human Genome Variation Society (HGVS) (den Dunnen and Antonarakis 2000). Sequence numbering refers to ENST00000324472 for the cDNA sequence and variations or probes are based on the UCSC GRCh38/hg38 assembly.

In silico analyses of sequence variants

The pathogenicity of the identified variants was predicted using Varsome (<https://varsome.com/>) (Kopanos et al. 2019) and Polyphen (<https://genetics.bwh.harvard.edu/pph2/index.shtml>) (Adzhubei et al. 2010). The potential effect of these variants on RNA splicing was assessed with Human Splicing Finder—V3.1 (<https://www.umd.be/HSF>) (Desmet et al. 2009).

Results

A total of 73 globozoospermic men were addressed for the genetic analysis of *DPY19L2*. Four patients were excluded as they did not pass the eligibility criteria: one had less than 20% of globozoospermic spermatozoa and three had an insufficient sperm concentration (< 1 million/mL) (Fig. 3).

***DPY19L2* investigations in our cohort of 69 globozoospermic patients**

We performed MLPA analysis on 69 DNA samples extracted from globozoospermic patients: 25 carried a homozygous deletion of the whole *DPY19L2* gene (36.2%) including the three Algerian brothers, 11 had a heterozygous deletion (15.9%) and 33 patients harbored no *DPY19L2* deletion (47.8%) (Fig. 3).

Further analyses were carried out for the 44 subjects who did not carry a homozygous deletion. Sanger sequencing of *DPY19L2* 22 exons was performed for 16 patients including the first five recruited patients without *DPY19L2* deletion and all patients ($n = 11$) carrying a heterozygous deletion. Whole exome sequencing was performed on the remaining 28 non-deleted patients.

For the heterozygous deleted subjects, eight patients (73%) harbored a hemizygous deleterious mutation on their unique allele and no deleterious variants were identified in the three remaining subjects (Fig. 3). For the 33 patients without the *DPY19L2* deletion, six had a homozygous variant (6/33, 18.2%) and two had a heterozygous variant (2/33, 6.1%) (Fig. 3). These two variants correspond to a missense mutation in exon 14, c.1478C>G; p.Thr493Arg, with an uncertain impact on the protein (Polyphen score of 0.372, predicted as benign while SIFT indicated the variant as pathogenic with a score of 0.01) and the other is a synonymous mutation, c.1461G>A; p.Ala487=, affecting the last nucleotide of the exon 14 and predicted by human splicing finder (HSF) to alter the donor splice site of the intron 14 (Table 2).

Overall 14 patients out of 44 without a homozygous deletion (32%) were considered to have a positive *DPY19L2* diagnosis including eight compound heterozygous with a heterozygous deletion and a point mutation and six with a

Table 2 All point mutations identified in our cohort of 69 globozoospermic patients and their predicted impact according to ACMG classification

Patient	Genomic mutation	Type of mutation	Localization	Protein modification	ACMG classification (Richards et al. 2015)
16IF004	c.1461G>A	Synonymous (heterozygous)	Exon 14	p.Ala487=	Likely pathogenic
15IF010	c.1478C>G	Missense (heterozygous)	Exon 14	p.Thr493Arg	Uncertain significance
17IF120	c.153_189del	Frameshift	Exon 1	p.Trp52SerfsTer7	Pathogenic
16IF026	c.575A>G	Missense	Exon 4	p.His192Arg	Likely pathogenic
14IF073	c.586G>C	Missense	Exon 4	p.Glu196Gln	Likely pathogenic
19U0060	c.869G>A*	Missense	Exon 8	p.Arg290His	Likely pathogenic
15IF090					
12IF037	c.892C>T**	Missense	Exon 8	p.Arg298Cys	Likely pathogenic
14IF051					
17IF104					
19U0058	c.893G>A	Missense	Exon 8	p.Arg298His	Likely pathogenic
15IF087	c.925C>A	Missense	Exon 8	p.Gln309Lys	Likely pathogenic
14IF02	c.1183delT***	Frameshift	Exon 11	p.Ser395LeufsTer9	Pathogenic
13IF035	c.1438G>A	Missense	Exon 14	p.Glu480Lys	Uncertain significance
17IF108	c.1580+1G>A	Splice-site mutation	Intron 16	p.512_527delfsTer5	Pathogenic
14IF069	c.1840G>T	Nonsense	Exon 19	p.Glu614Ter	Pathogenic

All mutations are homozygous or hemizygous unless specified otherwise

*Coutton et al. (2012b), Elinati et al. (2012), Zhu et al. (2013)

**Elinati et al. (2012), Ghédir et al. (2016)

***Elinati et al. (2012)

homozygous variant. In total, 11 different causal variants were identified. There were four loss of function variants and seven missense mutations (Figs. 2, 3). Patient 14IF02 presents a c.1183delT; p.Ser395LeufsTer9 variant associated with a heterozygous deletion of the other allele (Table 2). This variant has already been reported in 2012 (Elinati et al. 2012), it produces a 403 amino acid truncated protein and removes the last three transmembrane domains and the C-terminal end (Fig. 2). The variant c.153_189del; p.Trp52SerfsTer7 is a homozygous frameshift mutation presents in patient 17IF120 (Table 2) which introduces a premature stop codon and leads to a truncated protein of 58 amino acids instead of 758. The mutation of patient 15IF090, c.1840G>T; p.Glu614Ter is associated with a heterozygous deletion of the entire gene (Table 2). It is also a truncating mutation which produces a protein of 613 amino acids by replacing the glutamine by a stop codon, eliminating the C-terminal domain of the protein (Fig. 2). We detected one splice site variant, c.1580+1G>A; p.512_527delfsTer5, on the only allele of patient 17IF108, abrogating the donor site at the beginning of the intron 16 (Table 2). The alternative splicing is predicted to cause skipping of exon 16 inducing a stop codon in position 517 (Fig. 2).

Concerning the missense variants, two were already reported in the literature and five are novel (Table 2; Fig. 2). The variant c.869G>A; p.Arg290His is the most frequently reported point mutation of *DPY19L2* (Coutton

et al. 2012b; Elinati et al. 2012; Zhu et al. 2013) and was found in two patients from our cohort in association with a heterozygous deletion of the other allele (Table 2). This variant affects an extramembrane domain likely located in the perinuclear space (Fig. 2). The c.892C>T; p.Arg298Cys variant is present with a homozygous status in three patients from our cohort (Table 2). It was also already described previously (Elinati et al. 2012) as a deleterious mutation affecting a conserved amino acid in DPY19L2 sixth transmembrane domain (Fig. 2). Indeed, the change of an arginine at position 298 into a cytosine is extremely rare (GnomAD: 3.99×10^{-6}) and this alteration is predicted to be deleterious by SIFT (score: 0) and Polyphen (score: 1). Interestingly, another patient (19U0058) presents a different coding variant affecting the same arginine in position 298: c.893G>A; p.Arg298His (Table 2). Three other missense variants (c.586G>C; p.His192Arg, c.575A>G; p.Glu196Gln and c.925C>A; p.Gln309Lys) are each present in one patient of our cohort, all of them in association with a heterozygous deletion of *DPY19L2* gene (Table 2). A last homozygous missense variant was detected in patient 13IF035: c.1438G>A; p.Glu480Lys (Table 2). These four missense variants are all absent of the general population according to GnomAD and have a deleterious impact on DPY19L2 protein according to SIFT (score < 0.005) and Polyphen (score > 0.94).

Overall, a causal alteration of the *DPY19L2* gene was found in 39 patients (39/69, 56.5%) including 25 patients with a homozygous deletion of the full gene (25/69, 36.2%), six carrying a homozygous deleterious point mutation (6/69, 8.7%) and eight carrying a heterozygous deletion and a hemizygous deleterious variant (8/69, 11.6%). No *DPY19L2* defects were identified in 25 subjects (25/69, 36%) and five patients harbored a single pathogenic alteration (5/69, 7.2%) including a heterozygous *DPY19L2* deletion for three patients (3/69, 4.35%) and a heterozygous point mutation for the other two (2/69, 2.9%). As the transmission of *DPY19L2*-related globozoospermia is strictly recessive these events were not considered to be responsible for the patients' phenotype and a total of 30 patients were considered to have a negative diagnosis (30/69, 43.5%) (Fig. 3).

Among the ten patients reporting inbreeding in their family, only 30% (3/10) had a homozygous deletion, 20% (2/10) had a homozygous missense mutation, and 50% (5/10) had no *DPY19L2* abnormalities.

Whole exome sequencing analysis of other genes involved in globozoospermia in human

Among the 30 patients without causal anomaly detected in *DPY19L2*, 23 had a whole exome sequencing performed (Fig. 3) and the seven others had only a Sanger sequencing of the *DPY19L2* 22 exons. Whole-exome sequencing allowed us to search for anomalies in all the known candidate genes in globozoospermic patients (*C2CD6*, *C7orf61*, *CCDC62*, *CCIN*, *DNAH17*, *GGN*, *PICK1*, *SPATA16*, and *ZPBP1*) (Dam et al. 2007; Liu et al. 2010; Yatsenko et al. 2012; ElInati et al. 2016; Oud et al. 2020). We only considered the variants affecting one of the nine candidate genes with a protein impact predicted to be deleterious. Only one homozygous frameshift variant was detected in one patient affected by a partial globozoospermia with 82% of round headed spermatozoa. This alteration is a deletion of 22 nucleotides in *GGN* exon 3, c. 416_437del, introducing a stop codon in position 147 of the gametogenetin protein, p.Leu139ArgfsTer8 (Fig. 4). It induces the production of a

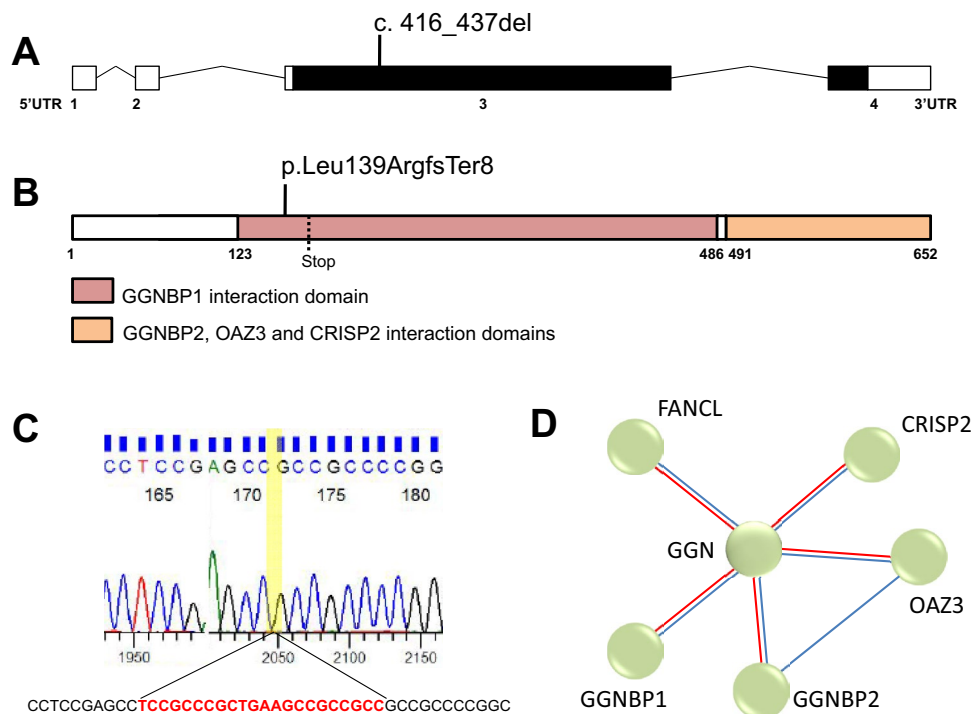


Fig. 4 Representations of the loss of function variant found in *GGN* and Sanger validation. **a** *GGN* has four exons indicated as boxes including two coding exons (exons 3 and 4) indicated as black colored boxes, introns are represented with a line, the localization of the identified point mutations is shown by a line. Numbers under the boxes depict the exons and UTR domains of the gene. **b** Representation of *GGN* protein; the anomaly is represented by a black line on amino acid 139 and the induced stop codon represented by a dotted line in position 147. Three interaction domains are localized on the protein. The GGNBP1 interaction domain from amino acid 123 to 486 and the GGNBP2 and OAZ3 interaction domain between

positions 491 and 652. The CRISP2 interaction domain overlaps the GGNBP2 and OAZ3 interaction domain and extends on the 158 C-terminal amino acids. All these domains are impacted by the loss of function mutation of *GGN*. **c** Electropherograms of Sanger sequencing showing the deletion of 22 nucleotides in patient. The deleted nucleotides are represented in red and bold on the sequence below. **d** String representation of *GGN* proteic interactions with FANCL, CRISP2, GGNBP1, GGNBP2 and OAZ3. Proteins are represented by green circles linked by strings according to the existing evidence of functional link: an experimental or biochemical data (red string), the co-mention in a Pubmed abstract (blue string)

truncated protein of 146 amino acids losing the interaction domain with GGNBP2 and OAZ3 in positions 491–652 and truncating the GGNBP1 interaction domain located in positions 123–486 (Fig. 4).

Comparison of sperm parameters

Sperm parameters were compared between three groups of patients according to their *DPY19L2* genotype. The first group is composed of 28 patients who carry bi-allelic loss of function variants (25 patients with a homozygous deletion of *DPY19L2* and four patients with homo- or hemizygous truncating mutations) for whom no functional DPY19L2 protein is expected (Fig. 3). The second group is composed of ten patients carrying a deleterious homo- or hemizygous missense variant of *DPY19L2* with an unpredictable effect on the protein expression or function (Fig. 3). The third group contains the 30 patients without bi-allelic alteration of *DPY19L2*, including patients with a unique heterozygous variant (Fig. 3).

Mean and standard deviation of all the sperm parameters for each group are presented in Table 1 and in Fig. 5. Several sperm parameters are statistically different between the first group of patients (without functional protein) and the third group of patients (without genetic alterations) (Fig. 5). The first group presents a higher proportion of round headed spermatozoa with an adjusted p value = 0.000005 and more microcephalic spermatozoa with an adjusted p value = 0.000061. Vitality and concentration of sperm are also increased in the first group with a respective adjusted p values of 0.006799 and 0.010454. The third group shows

an increased frequency of shortened flagella (adjusted p value = 0.004551). The normal spermatozoa proportion is enhanced in the third group (adjusted p value = 0.049672). There was no significant difference between these two groups concerning the other sperm parameters (Fig. 5). No statistically significant differences were identified between the second group of patients carrying a missense mutation and the two other groups (Fig. 5).

Diagnostic performances

Table 3 presents the genetic results of *DPY19L2* screening according to the rate of round-headed spermatozoa in our patients. We divided our cohort in three groups of patients: less than 50% of round headed spermatozoa, 50–89% and more than 90%. In these three groups we determined the proportion of each genetic profile (presence of a loss of function anomaly, presence of a causal missense mutation and the absence of *DPY19L2* causal anomaly) to determine the diagnostic performance in each group. Interestingly, we observed no positive diagnosis in the first group of patients with less than 50% of globozoospermia whereas a positive diagnosis was obtained for nearly 74% of patients carrying a minimum of 50% of round-headed spermatozoa (Table 3). Logically, if we compare the patients with what could be described as total globozoospermia (> 90% of globozoospermia) with those with partial globozoospermia (50–89% of globozoospermia) we observe in the first group a higher diagnosis rate (80% vs 54%) and a higher percentage of loss of function variants (63% vs 31%).

Fig. 5 Bar chart representing mean and standard deviation (SD) for each sperm parameters between the three groups of patients according to their *DPY19L2* genotype. No statistically significant differences exist between the group of patients carrying a missense mutation and the two other groups. The rate of normal spermatozoa is significantly enhanced, and the rate of globozoospermia significantly decreased, in patients without causal anomaly detected in *DPY19L2* in comparison with patients with a loss of function variant, respective p values of 0.049672 and 0.000005 (grey arrows)

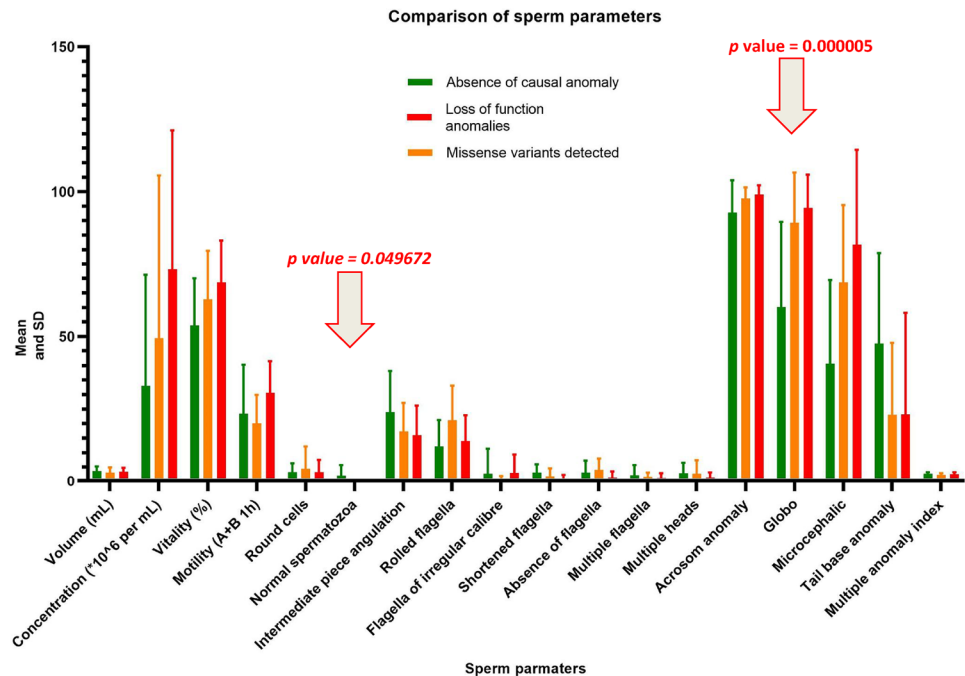


Table 3 Diagnostic performance according to the rate of round-headed spermatozoa

	Round-headed sperm		Loss of function anomalies		Missense variants		Absence of causal anomaly		Positive diagnosis (all biallelic variants)	
	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	<i>n</i>	%
<50	16	23.19	0	0.00	0	0.00	16	100.00	0	0.00
50–89	13	18.84	4	30.77	3	23.08	6	46.15	7	53.85
90–100	40	57.97	25	62.50	7	17.50	8	20.00	32	80.00
>50	53	76.81	29	54.72	10	18.87	14	26.42	39	73.58
All	69	100.00	29	42.03	10	14.49	30	43.48	39	56.52

Discussion

This study outlines the genetic diagnosis investigations performed in the largest cohort of globozoospermic patients published so far. We analyzed the *DPY19L2* gene in 69 patients from Europe, Africa and the Middle East, with a variable percentage of round headed spermatozoa ranging from 20 to 100%. We detected a *DPY19L2* homozygous deletion of the entire gene in 36.2% of our patients and causal point mutations in 20.3% (Fig. 3). We found three causal mutations already published in six patients: p.Arg290His, p.Arg298Cys and p.Ser395LeufsTer9 (Table 2) and discovered eight novel point mutations each present in one patient including five missense (p.His192Arg, p.Glu196Gln, p.Arg298His (c.893G>A), p.Gln309Lys and p.Glu480Lys) and three loss of function mutations (p.Trp52SerfsTer7, p.512_527delfsTer5 and p.Glu614Ter) (Table 2). Comparing the phenotype of different groups of patients according to their genotype allowed us to observe a genotype–phenotype correlation and led us to new recommendations in terms of diagnosis process.

Our statistical results allowed us to correlate the identification of *DPY19L2* loss of function anomalies with a higher rate of globozoospermic spermatozoa and logically with a lower proportion of normal spermatozoa in comparison with patients without causal bi-allelic alteration of *DPY19L2*. In accordance with these correlations, we obtained a higher diagnostic performance, reaching 80%, in patients with at least 90% of round headed spermatozoa (Table 3) who could be considered as total globozoospermic patients. We observed that we found *DPY19L2* bi-allelic causal anomalies exclusively in patients with more than 50% of globozoospermia with a diagnostic performance of 73.6%, whereas our diagnostic performance was null in patients with less than 50% of globozoospermic spermatozoa (Table 3). This results led us to recommend to initiate a targeted search for *DPY19L2* defects only in patients with a minimum of 50% of globozoospermic spermatozoa. Interestingly loss of function variants (in particular the complete *DPY19L2* deletion) were twice as frequent in subjects with total globozoospermia

(>90%) compared with those with partial globozoospermia (50–89%).

Surprisingly we observed a decrease in shortened flagella and an increased in microcephaly, vitality and sperm concentration in the group with loss of function variants. The difference in the rate of microcephalic sperm is probably explained by discrepancies in the semantic of different biologists: some characterizing round-headed sperm as globozocephalic, others as microcephalic. We, therefore, do not feel that the observed difference is relevant. The other differences observed could be explained by the low specificity of the recruitment, some centers prescribing the genetic analysis of *DPY19L2* for patients with low percentage of globozoospermic spermatozoa. In addition, the size of our cohort and the wide geographical origins of our patients could also contribute to the observed phenotypic heterogeneity, in opposition to studies focusing on specific populations (Harbuz et al. 2011; Zhu et al. 2013; Chianese et al. 2015). Furthermore the fact that the sperm analyses were realized in many different centers is on one hand a guarantee of the representativeness of the general population and on the other hand a weakness as it induces a great variability in the characterization of the sperm samples. Another explanation is a biased interpretation of biologists when a high rate of globozoospermic spermatozoa is observed in a semen sample, we can assume that they could then pay less attention to other anomalies such as head or flagella anomalies. Finally, two phenotypes are close to globozoospermia: acrosomal hypoplasia which can coexist with globozoospermia in the same semen sample (Chemes 2018) and pseudo-globozoospermia (Anton-Lamprecht et al. 1976; Singh 1992; Coutton et al. 2015). It could be relevant to verify that those phenotypes are not mistaken for real globozoospermia and that criteria used to characterize round-headed spermatozoa by the different centers are the same to avoid any recruitment bias.

In this large study, we found 36.2% of *DPY19L2* homozygous deletions in 69 globozoospermic patients. This value seems low, compared to what has been described in the past with an average rate of 52.3% of *DPY19L2* homozygous deletion (Ray et al. 2017). Nevertheless, the proportion of

homozygous deletion is still concordant with the literature varying from 19% (Koscinski et al. 2011) to 75% (Harbuz et al. 2011) and very close to a recent study with 35% of *DPY19L2* homozygous deletions among a large cohort of 63 patients (Alimohammadi et al. 2020). In fact several globozoospermia studies recruited only total globozoospermic patients (Zhu et al. 2013; Shang et al. 2019; Ghédér et al. 2019) and observed a much higher rate of *DPY19L2*-positive diagnosis. This is consistent with the correlation that we observed here between the high rate of globozocephalic spermatozoa and the presence of loss of function *DPY19L2* variants or deletions.

Concerning the causal point mutations found in our cohort we observed a mutational hotspot in exon 8 (Fig. 3). This latter concentrates four out of the eleven variants identified in this study concerning half of the mutation carriers and seven out of ten missense mutations carriers (Table 2). If we add our patients to the literature, the proportion of exon 8 missense mutation carriers represent approximately 56.5% of all published patients with a causal missense mutation (Coutton et al. 2012b; Elinati et al. 2012; Zhu et al. 2013; Modarres et al. 2016; Ghédér et al. 2016; Shang et al. 2019; Oud et al. 2020). In addition, four of our patients harbor a missense mutation impacting the conserved arginine in position 298 (Table 2) described to be essential for the C-mannosyltransferase activity of DPY-19, the *DPY19L2* ortholog in *Clostridium elegans* (Buettner et al. 2013). This supports the idea that the central domain of DPY19L2 protein has a critical function (Ray et al. 2017) and we could go further assuming that this critical function concern in particular the third loop of the internuclear space and the sixth transmembrane domain coded by the exon 8 (Fig. 3). We also identified a truncating mutation toward the end of the cDNA (p.Glu614Ter) indicating that the C-terminal domain of the protein also plays a critical role in protein function, perhaps permitting to anchor the acrosome to the acroplaxome.

There was no significant difference in the *DPY19L2* abnormalities found in the whole cohort and in patients declaring a familial consanguinity. Indeed, the percentages of loss of function and missense anomalies of 30% and 20%, respectively, are very close to those of the whole cohort corresponding to 36.2% and 20.3%. As could be expected, only homozygous abnormalities were found in patients with related parents. These results are to be put in perspective as all patients probably did not declare their consanguinity. The anomalies identified in our cohort and the genotype–phenotype relationship confirms the predominance of *DPY19L2* defects in globozoospermia. However, five patients carried a single heterozygous event (three with a whole deletion and two with a heterozygous variant). As heterozygous fathers are fully fertile, *DPY19L2* globozoospermia is considered to have a strict recessive inheritance. The presence of a single variant thus cannot

explain the five patient's phenotype, even if their impact on the protein is predicted to be deleterious. We did not find any other point mutation in these patients, nevertheless the sequencing technique used cannot detect the presence of deep intronic mutations or partial deletions of one or more *DPY19L2* exons.

Among the 30 patients without causal *DPY19L2* anomalies, 23 had a whole exome sequencing performed after MLPA (Fig. 3). In these patients, after the analysis of the nine genes described to be associated with human globozoospermia (*C2CD6*, *C7orf61*, *CCDC62*, *CCIN*, *DNAH17*, *GGN*, *PICK1*, *SPATA16* and *ZPBP1*) (Dam et al. 2007; Liu et al. 2010; Yatsenko et al. 2012; ElInati et al. 2016; Oud et al. 2020) we only found one homozygous deleterious variant in *GGN* in one patient with partial (82%) globozoospermia. Only one globozoospermic patient with a *GGN* defect has been described before and carried a homozygous truncating mutation in the same exon than the subject described here (Oud et al. 2020). *GGN* is located on chromosome 19 in 19q13.2 and encodes for the Gametogenin, a protein of 652 amino acids almost exclusively expressed in the testis. This protein has been detected in late pachytene spermatocytes and round spermatids before being incorporated into the principal piece of the sperm tail (Jamsai et al. 2008). *GGN* interacts with several other proteins such as FANCL (Lu and Bishop 2003), *GGNBP1*, *GGNBP2*, *OAZ3* (Zhang et al. 2005; Zhou et al. 2005) and *CRISP2* (Jamsai et al. 2008). All these genes have been related to sperm development; however, only *GGNBP1* has been described to be associated with round headed spermatozoa and the lack of acrosome (Han et al. 2020). *GGN* has been described to be related to FANCL which is implicated in double strand breaks repairs and the survival of pre-implantation embryos (Jamsai et al. 2013). *GGNBP2* was also described to be involved in cellular division (Guan et al. 2012). *OAZ3* is implicated in spermatogenesis but was described to be necessary for the formation of a rigid junction between head and tail (Tokuhiro et al. 2009). *CRISP2* is expressed in the acrosome of the spermatozoon; however, the *GGN*-*CRISP2* interaction was described to take place in the sperm tail (Jamsai et al. 2008). Thus the link between *GGN* anomalies and globozoospermia does not seem to be connected to the interactions it may have with these proteins. In contrast, *GGNBP1*, is also predominantly expressed in the testis, has been related to acrosome development and sperm head shape (Han et al. 2020). Finally the *GGN* knock out male mice are not described to have globozoospermia but present an embryonic lethality at the beginning of the pre-implantation development and *Ggn* \pm mice present defects in double strand break repairs (Jamsai et al. 2013). Therefore, although our results provide some strong evidence to link *GGN* with globozoospermia, further investigations should be performed to understand *GGN* function and clarify

its implication in globozoospermia. Its interaction with GGNBP1 seems to be the best lead to follow.

No deleterious variants were identified in the other candidate genes explored but the phenotype–genotype relationship is not clearly proved for the last published genes C2CD6, CCIN, C7orf61, and DHNA17. All these genes may have a putative location or function in acrosome but these assumptions are not clearly supported by their mutant models when available. Moreover, the published missense mutations were not validated with functional work and should be interpreted with caution. For example, missense and truncating mutations in *DNAH17* were formally associated with another severe flagellum malformations known as the MMAF phenotype in human and mouse in three recent and independent publications (Touré et al. 2020). The association of *DNAH17* with globozoospermia should, therefore, be strongly questioned. Exome sequencing data from the remaining 22 patients without any identified variant will now be fully explored to identify deleterious variants in new undescribed candidate genes.

Analysis of exome data permitted to identify *DPY19L2* deleterious variants in five patients, demonstrating that the technique is efficient to detect small gene defects in *DPY19L2* despite the presence of a highly homologous pseudogene *DPY19L2P1* (Harbuz et al. 2011). This observation brought us to balance the benefit of MLPA against a classic qualitative PCR. The only advantage of MLPA in our procedure was to detect the presence of heterozygous deletion of the entire *DPY19L2* gene but we demonstrated that this information has no interest when it comes alone and further sequencing has to be performed. Here we observed that MLPA only allowed the identification of a causal anomaly in 36.2% of patients, and those with a homozygous deletion, could be detected simply by standard PCR of several exons of the gene. We then recommend to use exome sequencing to complete the analysis of patients without a homozygous deletion. The addition of a CNV module for the detection of copy number variations (CNV) will permit to detect the whole *DPY19L2* heterozygous deletions previously detected by MLPA as well as yet undetected CNVs removing one or several exons. Naturally, as demonstrated here, WES also permits to detect *DPY19L2* point mutations and to explore other genes already involved in globozoospermia, as demonstrated here by the identification of a homozygous *GGN* mutation. It could also permit to identify new genes in patients without variants in known candidate genes. In financial terms, this strategy is also relevant as we expect an improved diagnostic performance, for a lower price. Concerning the first diagnosis step of *DPY19L2* homozygous deletion, the cost of an MLPA is around 120 euros for the laboratory, including reagents and technical time (with an estimated charged cost of 25 euros/h for an experienced technician), when a simple PCR of the same three exons is

invoiced 30 euros for an equivalent diagnostic efficiency. In the second diagnosis phase using Sanger sequencing would cost at least 70 euros per exon which accounts for approximately 1500 euros to sequence the 22 exons of *DPY19L2*, whereas a whole exome sequencing cost between 500 and 800 euros under current market conditions. Thus our previous strategy using MLPA and Sanger sequencing had an average cost around 1660 euros when the novel strategy combining a classical qualitative PCR and whole exome sequencing would cost approximately 830 euros, half the cost of the initial strategy.

In conclusion, the work presented here permitted to propose a strategy for the routine genetic investigation of globozoospermic patients: subjects with less than 50% of round-headed sperm should not be considered globozoospermic and *DPY19L2* should not be investigated. For the other patients, a classical qualitative PCR should be used to detect homozygous deletions of *DPY19L2*, and in the absence of the homozygous deletion of the gene, instead of the usual Sanger sequencing of *DPY19L2*, we recommend sequencing the entire exome, which allows a cost-effective detection of genetic defects in the *DPY19L2* gene but also in the other candidate genes described. The interest of this strategy was demonstrated by allowing the detection of a mutation in *GGN*, thus confirming its likely association with globozoospermia.

Author contributions TC, CaCa, FM, ZEK, CA, and PFR carried out and interpreted all the molecular work. TC and PFR performed all data analysis. TC and GM designed figures. TC and PR supervised all the patient's information and SFMB, MK, RZ, AZ, AC, AP, AB, AC, BF, CG, CT, CS, DL, GT, GS, HL, JP, JB, LP, LM, MP, MC, NS, RC, SB, SV, CF, SH, VM, VS, VA contributed to the acquisition of data. TC, ChCo and PFR wrote the main draft of the manuscript. PFR designed the overall study, supervised all molecular laboratory work, had full access to all of the data and takes responsibility for the integrity of the data and its accuracy.

Compliance with ethical standards

Conflict of interest The authors declare no conflict of interest.

References

- Adzhubei IA, Schmidt S, Peshkin L et al (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248–249. <https://doi.org/10.1038/nmeth0410-248>
- Alimohammadi F, Ebrahimi Nasab M, Rafeae A et al (2020) Deletion of dpy-19 like 2 (*DPY19L2*) gene is associated with total but not partial globozoospermia. *Reprod Fertil Dev*. <https://doi.org/10.1071/RD19025>
- Anton-Lamprecht I, Kotzur B, Schopf E (1976) Round-headed human spermatozoa. *Fertil Steril* 27:685–693. [https://doi.org/10.1016/s0015-0282\(16\)41900-x](https://doi.org/10.1016/s0015-0282(16)41900-x)

- Buettner FFR, Ashikov A, Tiemann B et al (2013) *C. elegans* DPY-19 is a C-mannosyltransferase glycosylating thrombospondin repeats. *Mol Cell* 50:295–302. <https://doi.org/10.1016/j.molcel.2013.03.003>
- Chemes HE (2018) Phenotypic varieties of sperm pathology: genetic abnormalities or environmental influences can result in different patterns of abnormal spermatozoa. *Anim Reprod Sci* 194:41–56. <https://doi.org/10.1016/j.anireprosci.2018.04.074>
- Chen S, Zhou Y, Chen Y, Gu J (2018) Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>
- Chianese C, Fino MG, Riera Escamilla A et al (2015) Comprehensive investigation in patients affected by sperm macrocephaly and globozoospermia. *Andrology* 3:203–212. <https://doi.org/10.1111/andr.12016>
- Cooper TG, Noonan E, von Eckardstein S et al (2010) World Health Organization reference values for human semen characteristics*. *Hum Reprod Update* 16:231–245. <https://doi.org/10.1093/humupd/dmp048>
- Coutton C, Vieville G, Satre V et al (2012a) Multiplex ligation-dependent probe amplification (MLPA) et sondes « à façon » entièrement synthétiques. Guide pratique, recommandations et expérience au CHU de Grenoble. *IRBM* 33:227–235. <https://doi.org/10.1016/j.irbm.2012.04.004>
- Coutton C, Zouari R, Abada F et al (2012b) MLPA and sequence analysis of DPY19L2 reveals point mutations causing globozoospermia. *Hum Reprod* 27:2549–2558. <https://doi.org/10.1093/humrep/des160>
- Coutton C, Abada F, Karaouzene T 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:e1003363. <https://doi.org/10.1371/journal.pgen.1003363>
- Coutton C, Escoffier J, Martinez G et al (2015) Teratozoospermia: spotlight on the main genetic actors in the human. *Hum Reprod Update* 21:455–485. <https://doi.org/10.1093/humupd/dmv020>
- Dam AHDM, Kosciński I, Kremer JAM et al (2007) Homozygous mutation in SPATA16 is associated with male infertility in human globozoospermia. *Am J Hum Genet* 81:813–820. <https://doi.org/10.1086/521314>
- Dam AH, Ramos L, Dijkman HB et al (2011) Morphology of partial globozoospermia. *J Androl* 32:199–206. <https://doi.org/10.2164/jandrol.109.009530>
- Datta J, Palmer MJ, Tanton C et al (2016) Prevalence of infertility and help seeking among 15 000 women and men. *Hum Reprod* 31:2108–2118. <https://doi.org/10.1093/humrep/dew123>
- den Dunnen JT, Antonarakis SE (2000) Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat* 15:7–12. [https://doi.org/10.1002/\(SICI\)1098-1004\(200001\)15:1%3c7::AID-HUMU4%3e3.0.CO;2-N](https://doi.org/10.1002/(SICI)1098-1004(200001)15:1%3c7::AID-HUMU4%3e3.0.CO;2-N)
- Desmet F-O, Hamroun D, Lalande M et al (2009) Human splicing finder: an online bioinformatics tool to predict splicing signals. *Nucleic Acids Res* 37:e67. <https://doi.org/10.1093/nar/gkp215>
- Elinati E, Kuentz P, Redin C et al (2012) Globozoospermia is mainly due to DPY19L2 deletion via non-allelic homologous recombination involving two recombination hotspots. *Hum Mol Genet* 21:3695–3702. <https://doi.org/10.1093/hmg/dds200>
- Elinati E, Fossard C, Okutman O et al (2016) A new mutation identified in SPATA16 in two globozoospermic patients. *J Assist Reprod Genet* 33:815–820. <https://doi.org/10.1007/s10815-016-0715-3>
- Ghédér H, Ibala-Romdhane S, Okutman O et al (2016) Identification of a new DPY19L2 mutation and a better definition of DPY19L2 deletion breakpoints leading to globozoospermia. *Mol Hum Reprod* 22:35–45. <https://doi.org/10.1093/molehr/gav061>
- Ghédér H, Braham A, Viville S et al (2019) Comparison of sperm morphology and nuclear sperm quality in SPATA16- and DPY19L2-mutated globozoospermic patients. *Andrologia* 51:e13277. <https://doi.org/10.1111/and.13277>
- Guan R, Wen X-Y, Wu J et al (2012) Knockdown of ZNF403 inhibits cell proliferation and induces G2/M arrest by modulating cell cycle mediators. *Mol Cell Biochem* 365:211–222. <https://doi.org/10.1007/s11010-012-1262-6>
- Han T, Wang L, Tang W et al (2020) GGNBP1 ensures proper spermiogenesis in response to stress in mice. *Biochem Biophys Res Commun* 525:706–713. <https://doi.org/10.1016/j.bbrc.2020.02.118>
- Harbuz R, Zouari R, Pierre V 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. <https://doi.org/10.1016/j.ajhg.2011.02.007>
- Jamsai D, Bianco DM, Smith SJ et al (2008) Characterization of gametogenetin 1 (GGN1) and its potential role in male fertility through the interaction with the ion channel regulator, cysteine-rich secretory protein 2 (CRISP2) in the sperm tail. *Reproduction* 135:751–759. <https://doi.org/10.1530/REP-07-0485>
- Jamsai D, O'Connor AE, DeBoer KD et al (2013) Loss of GGN leads to pre-implantation embryonic lethality and compromised male meiotic DNA double strand break repair in the mouse. *PLoS ONE* 8:e56955. <https://doi.org/10.1371/journal.pone.0056955>
- Kim S, Scheffler K, Halpern AL et al (2018) Strelka2: fast and accurate calling of germline and somatic variants. *Nat Methods* 15:591–594. <https://doi.org/10.1038/s41592-018-0051-x>
- Kopanos C, Tsiolkas V, Kouris A et al (2019) VarSome: the human genomic variant search engine. *Bioinformatics* 35:1978–1980. <https://doi.org/10.1093/bioinformatics/bty897>
- Kosciński I, Elinati E, Fossard C et al (2011) DPY19L2 deletion as a major cause of globozoospermia. *Am J Hum Genet* 88:344–350. <https://doi.org/10.1016/j.ajhg.2011.01.018>
- Li H (2013) Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:13033997 [q-bio]
- Li H, Handsaker B, Wysoker A et al (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25:2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liu G, Shi Q-W, Lu G-X (2010) A newly discovered mutation in PICK1 in a human with globozoospermia. *Asian J Androl* 12:556–560. <https://doi.org/10.1038/aja.2010.47>
- Lu B, Bishop CE (2003) Mouse GGN1 and GGN3, two germ cell-specific proteins from the single gene Ggn, interact with mouse POG and play a role in spermatogenesis. *J Biol Chem* 278:16289–16296. <https://doi.org/10.1074/jbc.M211023200>
- Mascarenhas MN, Flaxman SR, Boerma T et al (2012) National, regional, and global trends in infertility prevalence since 1990: a systematic analysis of 277 health surveys. *PLOS Med* 9:e1001356. <https://doi.org/10.1371/journal.pmed.1001356>
- McLaren W, Gil L, Hunt SE et al (2016) The ensembl variant effect predictor. *Genome Biol* 17:122. <https://doi.org/10.1186/s1305-016-0974-4>
- Modarres P, Tanhaei S, Tavalae M et al (2016) Assessment of DPY19L2 deletion in familial and non-familial individuals with globozoospermia and DPY19L2 genotyping. *Int J Fertil Steril* 10:196–207
- Modarres P, Tavalae M, Ghaedi K, Nasr-Esfahani MH (2018) An overview of the globozoospermia as a multigenic identified syndrome. *IJFS*. <https://doi.org/10.22074/ijfs.2019.5561>
- Oud MS, Okutman Ö, Hendricks LAJ et al (2020) Exome sequencing reveals novel causes as well as new candidate genes for human globozoospermia. *Hum Reprod* 35:240–252. <https://doi.org/10.1093/humrep/dez246>
- Pierre V, Martinez G, Coutton C 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* 139:2955–2965. <https://doi.org/10.1242/dev.077982>

- Ray PF, Toure A, Metzler-Guillemain C et al (2017) Genetic abnormalities leading to qualitative defects of sperm morphology or function: genetic abnormalities leading to qualitative sperm defects. *Clin Genet* 91:217–232. <https://doi.org/10.1111/cge.12905>
- Richards S, Aziz N, Bale S et al (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17:405–423. <https://doi.org/10.1038/gim.2015.30>
- Shang Y-L, Zhu F-X, Yan J et al (2019) Novel DPY19L2 variants in globozoospermic patients and the overcoming this male infertility. *Asian J Androl* 21:183–189. https://doi.org/10.4103/aja.aja_79_18
- Singh G (1992) Ultrastructural features of round-headed human spermatozoa. *Int J Fertil* 37:99–102
- Tokuhiro K, Isotani A, Yokota S et al (2009) OAZ-t/OAZ3 is essential for rigid connection of sperm tails to heads in mouse. *PLoS Genet* 5:e1000712. <https://doi.org/10.1371/journal.pgen.1000712>
- Touré A, Martinez G, Kherraf Z-E et al (2020) The genetic architecture of morphological abnormalities of the sperm tail. *Hum Genet*. <https://doi.org/10.1007/s00439-020-02113-x>
- Yatsenko AN, O’Neil DS, Roy A et al (2012) Association of mutations in the zona pellucida binding protein 1 (ZPBP1) gene with abnormal sperm head morphology in infertile men. *Mol Hum Reprod* 18:14–21. <https://doi.org/10.1093/molehr/gar057>
- Zhang J, Wang Y, Zhou Y et al (2005) Yeast two-hybrid screens imply that GGNBP1, GGNBP2 and OAZ3 are potential interaction partners of testicular germ cell-specific protein GGN1. *FEBS Lett* 579:559–566. <https://doi.org/10.1016/j.febslet.2004.10.112>
- Zhou Y, Zhao Q, Bishop CE et al (2005) Identification and characterization of a novel testicular germ cell-specific gene Ggnbp1. *Mol Reprod Dev* 70:301–307. <https://doi.org/10.1002/mrd.20214>
- Zhu F, Gong F, Lin G, Lu G (2013) DPY19L2 gene mutations are a major cause of globozoospermia: identification of three novel point mutations. *Mol Hum Reprod* 19:395–404. <https://doi.org/10.1093/molehr/gat018>

Affiliations

Tristan Celse^{1,2} · Caroline Cazin^{1,2} · Flore Mietton² · Guillaume Martinez^{1,3} · Delphine Martinez² · Nicolas Thierry-Mieg⁴ · Amandine Septier⁴ · Catherine Guillemain^{5,6} · Julie Beurois¹ · Antoine Clergeau⁷ · Selima Fourati Ben Mustapha⁸ · Mahmoud Kharouf⁸ · Abdelali Zoghmar⁹ · Ahmed Chargui¹⁰ · Aline Papaxanthos¹¹ · Béatrice Dorphin¹² · Bernard Foliguet¹³ · Chema Triki¹⁴ · Christophe Sifer¹⁵ · Dominique Lauton¹⁶ · Gérard Tachdjian¹⁷ · Gilles Schuler¹⁸ · Hervé Lejeune¹⁹ · Jacques Puechberty²⁰ · Julien Bessonnat²¹ · Laurent Pasquier²² · Lionel Mery²³ · Marine Poulain²⁴ · Myriam Chaabouni⁸ · Nathalie Sermondade²⁵ · Rosalie Cabry²⁶ · Sebti Benbouhadja⁹ · Ségolène Veau²⁷ · Cynthia Frapsauce²⁸ · Valérie Mitchell²⁹ · Vincent Achard^{30,31,32} · Veronique Satre^{1,3} · Sylviane Hennebicq^{1,21} · Raoudha Zouari⁵ · Christophe Arnoult¹ · Zine-Eddine Kherraf^{1,2} · Charles Coutton^{1,3} · Pierre F. Ray^{1,2}

¹ Institute for Advanced Biosciences, Team Genetics Epigenetics and Therapies of Infertility, Université Grenoble Alpes, INSERM U1209, CNRS UMR 5309, 38000 Grenoble, France

² CHU Grenoble Alpes, UM GI-DPI, 38000 Grenoble, France

³ CHU Grenoble Alpes, UM de Génétique Chromosomique, 38000 Grenoble, France

⁴ Université Grenoble Alpes, CNRS, TIMC-IMAG, 38000 Grenoble, France

⁵ Pôle Femmes-Parents-Enfants, Centre Clinico-Biologique AMP-CECOS, Plateforme Cancer et Fertilité ONCOPACA-Corse, Assistance-Publique des Hôpitaux de Marseille (AP-HM), Marseille, France

⁶ Aix Marseille University, INSERM, MMG, UMR_S 1251, Marseille, France

⁷ Diabetes Care Unit, University Hospital of Caen, Caen, France

⁸ Polyclinique les Jasmins, Centre d’Aide Médicale à la Procréation, Centre Urbain Nord, 1003 Tunis, Tunisia

⁹ Reproduction Sciences and Surgery Clinique, Ibn Rochd, Constantine, Algeria

¹⁰ Faculté de Médecine, Assistance Publique-Hôpitaux de Paris (AP-HP), Hôpital Universitaire Paris Centre, Centre Hospitalier Universitaire (CHU) Cochin, Service

d’Histologie-Embryologie-Biologie de la Reproduction, Université Paris Descartes, Sorbonne Paris Cité, Paris, France

¹¹ Department of Obstetrics, Gynecology and Reproductive Medicine, Bordeaux University Hospital, Bordeaux, France

¹² AMP74, CH Alpes Léman, Contamine-sur-Arve, France

¹³ Toxicology and Molecular Biology, Institute Jean Lamour UMR 7198 du CNRS, Université de Lorraine, 54000 Nancy, France

¹⁴ Centre d’AMP, Clinique Hannibal, Les Berges du Lac, 1053 Tunis, Tunisia

¹⁵ Service de Biologie de la Reproduction, d’Histo-Embryologie et Cytogénétique, Hôpital Jean-Verdier, Avenue du 14 Juillet, 93140 Bondy, France

¹⁶ Department of Endocrinology, Diabetes, Nutrition, Montpellier University Hospital, Montpellier, France

¹⁷ UMR 967, INSERM, Service d’Histologie Embryologie et Cytogénétique, Hôpitaux Universitaires Paris-Sud, AP-HP, Clamart, France

¹⁸ Gynécologie Obstétrique, Sallanches, France

¹⁹ Reproductive Medicine Department, Hospices Civils de Lyon, Lyon, France

- ²⁰ Département de Génétique Médicale, Maladies Rares et Médecine Personnalisée, Hôpital Arnaud de Villeneuve, CHU de Montpellier, Université Montpellier, Montpellier, France
- ²¹ CHU de Grenoble, UF de Biologie de la Procréation, 38000 Grenoble, France
- ²² Service de Génétique Clinique, CLAD Ouest, CHU Rennes, Rennes, France
- ²³ Service de Médecine de la Reproduction, CHU de Saint-Étienne, Hôpital Nord, 42055 Saint-Étienne Cedex 2, France
- ²⁴ Department of Obstetrics and Gynecology, Hôpital Foch, Université de Paris Ouest (UVSQ), Suresnes, France
- ²⁵ Service de Biologie de la Reproduction—CECOS, Hôpital Tenon, AP-HP, 75020 Paris, France
- ²⁶ Department of Obstetrics, Gynaecology and Reproductive Medicine, Picardie University Jules Verne, Amiens University Medical Centre, Amiens, France
- ²⁷ CHU, Centre d'AMP—CECOS, University Rennes, 16 Boulevard de Bulgarie, 35000 Rennes, France
- ²⁸ CHU Bretonneau, Médecine et Biologie de la Reproduction-CECOS, Tours, France
- ²⁹ EA 4308, Department of Reproductive Biology and Spermiology-CECOS Lille, University Medical Center, 59037 Lille, France
- ³⁰ CECOS-Laboratoire de Biologie de la Reproduction, Pôle de Gynécologie Obstétrique et Reproduction (Gynépôle), Assistance Publique-Hôpitaux de Marseille (AP-HM) la Conception, 13005 Marseille, France
- ³¹ Centre Clinico-Biologique d'Assistance Médicale à la Procréation, Pôle de Gynécologie Obstétrique et Reproduction (Gynépôle), Assistance Publique-Hôpitaux de Marseille (AP-HM) la Conception, 13005 Marseille, France
- ³² Faculté de Médecine, Institut Méditerranéen de Biodiversité et d'Écologie (IMBE UMR 7263), Equipe Biogénotoxicologie, Santé Humaine et Environnement, Aix Marseille Université, CNRS, IRD, Université Avignon, 27, Boulevard Jean-Moulin, 13385 Marseille Cedex 5, France