


## Article

# Whole-Genome Profile of Greek Patients with Teratozoospermia: Identification of Candidate Variants and Genes

Maria-Anna Kyrgiafani <sup>1</sup>, Themistoklis Giannoulis <sup>2</sup> , Alexia Chatziparasidou <sup>3</sup>, Nikolaos Christoforidis <sup>3</sup> and Zissis Mamuris <sup>1,\*</sup>

<sup>1</sup> Laboratory of Genetics, Comparative and Evolutionary Biology, Department of Biochemistry and Biotechnology, University of Thessaly, Viopolis, Mezourlo, 41500 Larissa, Greece

<sup>2</sup> Laboratory of Biology, Genetics and Bioinformatics, Department of Animal Sciences, University of Thessaly, Gaiopolis, 41336 Larissa, Greece

<sup>3</sup> Embryolab IVF Unit, St. 173-175 Ethnikis Antistaseos, Kalamaria, 55134 Thessaloniki, Greece

\* Correspondence: zmamur@uth.gr

**Abstract:** Male infertility is a global health problem that affects a large number of couples worldwide. It can be categorized into specific subtypes, including teratozoospermia. The present study aimed to identify new variants associated with teratozoospermia in the Greek population and to explore the role of genes on which these were identified. For this reason, whole-genome sequencing (WGS) was performed on normozoospermic and teratozoospermic individuals, and after selecting only variants found in teratozoospermic men, these were further prioritized using a wide range of tools, functional and predictive algorithms, etc. An average of 600,000 variants were identified, and of them, 61 were characterized as high impact and 153 as moderate impact. Many of these are mapped in genes previously associated with male infertility, yet others are related for the first time to teratozoospermia. Furthermore, pathway enrichment analysis and Gene ontology (GO) analyses revealed the important role of the extracellular matrix in teratozoospermia. Therefore, the present study confirms the contribution of genes studied in the past to male infertility and sheds light on new molecular mechanisms by providing a list of variants and candidate genes associated with teratozoospermia in the Greek population.

**Keywords:** teratozoospermia; male infertility; whole-genome sequencing (WGS); next-generation sequencing; Greece



**Citation:** Kyrgiafani, M.-A.; Giannoulis, T.; Chatziparasidou, A.; Christoforidis, N.; Mamuris, Z. Whole-Genome Profile of Greek Patients with Teratozoospermia: Identification of Candidate Variants and Genes. *Genes* **2022**, *13*, 1606. <https://doi.org/10.3390/genes13091606>

Academic Editor: Darren Griffin

Received: 26 July 2022

Accepted: 30 August 2022

Published: 8 September 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Infertility is defined by the *World Health Organization* (WHO) as the failure to conceive after at least 12 months of regular and unprotected sexual intercourse [1]. It is considered a major health problem that affects the couple's psychology and social life [2,3], but at the same time, it causes a significant economic burden on the health care system and patients [4]. Moreover, it is estimated that more than 186 million people are affected worldwide [5], and in half of these cases, after a thorough examination, a male cause seems to be present alone or in conjunction with female causes [6]. Based on semen analysis and defects associated with sperm quality or quantity, several subtypes of male infertility can be defined, such as asthenozoospermia, oligozoospermia, or teratozoospermia [7]. Specifically, when in semen less than 4% of spermatozoa have normal morphology, according to WHO, the sample is characterized as teratozoospermic.

Male infertility is considered a multifactorial disorder [8], and it is estimated that genetic factors are involved in 15% of cases and more [9]. Although extensive research has resulted in significant advances in the field, the identification of specific genes and mutations is a great challenge, as more than 2000 genes are required only for spermatogenesis, a crucial process for fertility [9,10]. Especially for particular subtypes of male infertility, such as teratozoospermia, detection of causal mutations that lead to specific defects in

sperm parameters remains limited [8,10]. Furthermore, the fact that assisted reproductive technologies (ART) are being used more frequently than in the past proves the progress achieved in the area of male infertility, but it has been observed that ART outcomes differ between ethnic groups probably due to different genetic factors that contribute to infertility among populations [11,12]. However, the impact of ethnic differences on male infertility is not adequately addressed in research [12]. Today, whole-genome sequencing (WGS) enables the identification of common and even rare variants, which are often not captured at Genome-Wide Association Studies (GWAS) and SNP-chips [13], observed in specific geographical populations, which have an impact on phenotypic variation for a wide range of diseases in an accurate and cost-effective way [14,15].

Thus, this study aimed to perform WGS in teratozoospermic and normozoospermic individuals of the Greek population (a) to identify and further characterize variants, including rare variants, that can contribute to the pathogenic phenotype, and (b) to highlight genes potentially linked with teratozoospermia, a specific subtype of male infertility, and to explore their role. The ultimate goal of our study was to provide a valuable reference for future male infertility research, particularly regarding teratozoospermia. In this way, the detection of mutations contributing to the infertile phenotype in ethnic minorities may be of utmost importance for unraveling the genetic basis of male infertility by identifying new or rare variants and studying the genes on which they are found, and for improving the diagnosis of teratozoospermia, as well as the chances of successful ART.

More specifically, with our study that is focused on the Greek population, as there are no other studies for the genetic background of teratozoospermia in the Balkans, we attempt to (a) provide a roadmap for future studies investigating different ART outcomes due to different genetic background among populations and maybe improve the chances for successful ART for Balkanian populations, and (b) provide preliminary information about genetic causes of male infertility specifically found on Balkan populations enabling comparison with other populations, e.g., Chinese, African, etc., in order to investigate the mechanisms causing teratozoospermia.

## 2. Materials and Methods

### 2.1. Selection of Patients and Biological Material

For this study, human blood, as well as sperm samples, were collected from volunteers in cooperation with the “Embryolab IVF Unit” (55134 Thessaloniki, Greece) for the Spermogene research program. Ethical approval was obtained from the Ethics Committee, University of Thessaly (38221 Volos, Greece), and all individuals have given their approved written informed consent.

All the volunteers recruited underwent an andrological examination and semen analysis was performed on samples derived from all of them. It should be noted that sperm samples were collected via masturbation after at least two to three days of abstinence from sexual intercourse. For semen analysis (seminogram), cell vision counting slides (Tek-Event) were used for cell counting and observation was conducted on Nikon Eclipse TS100, Nikon Eclipse E200, and Nikon Eclipse Ts2 microscopes. Semen analysis was performed according to WHO guidelines (fifth edition, 2010, <https://apps.who.int/iris/handle/10665/44261> (accessed on 26 June 2022)) that include information on assessing semen volume, sperm count, motility, morphology, etc. These reference values proposed in this edition were used to define normozoospermic and teratozoospermic phenotypes.

Moreover, the inclusion criterion for this study was the Greek ethnicity as according to research, infertility exerts racial differences [16,17] and the ART outcomes differ between populations [11,12], but there is limited information regarding specific variations and polymorphisms associated with male infertility for ethnic groups found in the Balkans, and especially for the Greek population and teratozoospermia. Therefore, place of birth and relevant data were collected from the volunteers through the questionnaire which was required to fill in along with the consent form. Demographic information on the individuals enrolled in this study is presented in Table 1.

**Table 1.** Demographic information of volunteers selected to participate in this study.

Demographics	Normozoospermic ( <i>n</i> = 10)	Teratozoospermic ( <i>n</i> = 5)
Age	28–53	31–49
	Mean = 36	Mean = 38
BMI	19.5–40.4	24.8–33
	Mean = 26.97	Mean = 29.24
Smoking	30% Not Smoking, 70% Smoking	60% Not Smoking, 40% Smoking
Alcohol	100% ≤ 2 drinks/week	80% ≤ 2 drinks/week
Drug Use	75% No	80% No
Nationality	Greek	Greek

## 2.2. Sample Preparation and Whole-Genome Sequencing

Genomic DNA was extracted from blood samples of teratozoospermic and normozoospermic individuals using the PureLink Genomic DNA Mini Kit (Invitrogen, Waltham, MA, USA—Catalog number: K182002) according to the manufacturer's instructions. DNA quality and quantity were assessed by agarose gel electrophoresis and by Qubit 2.0 fluorometer using the Qubit dsDNA BR Assay Kit (Invitrogen, Waltham, MA, USA—Catalog number: Q32850), respectively. After that, three sequencing pools were created. More specifically, DNA extracted from ten normozoospermic individuals was used for the two pools (five individuals for each pool), and similarly, the third contained pooled DNAs from five teratozoospermic individuals. The DNAs were mixed equimolar for each pool in a final concentration of 100 ng/μL and a final quantity of 2 mg.

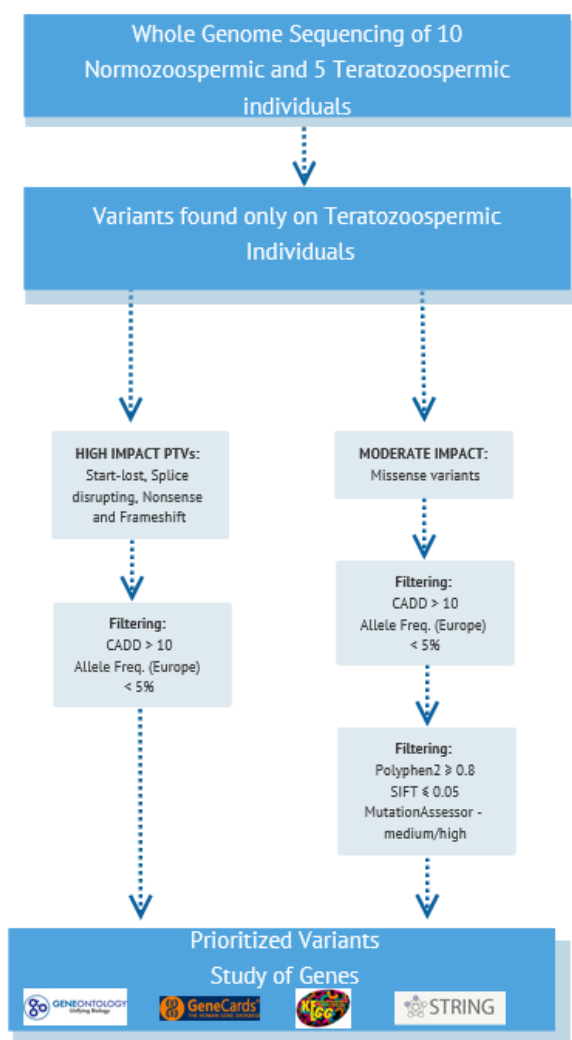
Once their preparation was completed, DNA samples were shipped to Novogene (Cambridge, UK) where 100-bp paired-end libraries were constructed and sequenced using an Illumina HiSeq 3000 in a mean sequencing coverage of 30x. For the analysis of the FASTQ files produced, at first, the quality of the reads was evaluated using FASTQC (available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>, accessed on 26 June 2022) and then low-quality reads (minimum PHRED Score: 30), as well as adapter sequences, were discarded using Trimmomatic [18]. After quality control, the reads were aligned to a human reference genome (GRCh37/hg19) retrieved from the Ensembl database [19] using the Burrows-Wheeler aligner (BWA) [20]. Duplicate reads produced by polymerase chain reaction (PCR) were marked and removed by Picard tools before further analysis. SAM files of the alignment were then converted to BAM files using SAMtools [21] and individual BAM files for the two normozoospermic pools were merged to create one file representing normozoospermic individuals using SAMtools, too. Variant calling was performed using freeBayes [22] and as an output, the results were stored in variant call format (VCF). BCFtools [21] was then used to compare the VCF files from normozoospermic and teratozoospermic individuals to detect unique variants that are present only in one of the two groups, thus, they are not shared between normozoospermic and teratozoospermic individuals. Further analysis was performed for variants found only on men diagnosed with teratozoospermia as described earlier, as the aim was to identify variants and/or polymorphisms on teratozoospermic men that have the potential to contribute to the pathogenic phenotype and may cause teratozoospermia in the Greek population. Moreover, as they exist only on patients, they could be used for the development of teratozoospermia biomarkers in the Greek population for effective diagnosis as well as for ART outcome improvement.

Following the detection of unique variants for teratozoospermic individuals, annotation was performed using the VEP tool (available at <https://www.ensembl.org/Tools/VEP>, accessed on 26 June 2022), provided by the Ensembl database. Furthermore, a list of databases, software tools, and additional prediction algorithms were used to retrieve more biological information and prevent bias in the filtering and prioritization of variants that were performed afterward, since it is common for potentially disease-relevant mutations to be ignored due to inadequate annotations [23]. Among them, the Single Nucleotide Poly-

morphism Database (dbSNP) [24], 1000 Genomes Project [25], and Genome Aggregation Database (gnomAD) [26] were used to obtain information about allele frequencies and the identification of rare/novel variants. Further analysis was carried out to predict the effect of variants on protein's functionality or their potential pathogenic effect using Polymorphism Phenotyping v2 (PolyPhen2) [27], Sorting Intolerant From Tolerant (SIFT) [28], Combined Annotation Dependent Depletion (CADD) [29], and MutationAssessor [30].

### 2.3. Variant Prioritization

In the present study, the variants found only in teratozoospermic individuals were filtered to prioritize those that are more likely to be involved in the occurrence of this specific subtype of male infertility in the Greek population and contribute to the unique genomic profile of Greek patients with teratozoospermia. The corresponding genes that the variants are mapped on were also studied to explore their role. Thus, the prioritization was performed as follows (Figure 1):



**Figure 1.** The pipeline used to filter variants found in Greek patients with teratozoospermia. At first, high and moderate impact variants were selected and then, bioinformatics tools were used to assess their impact on protein structure, function, and conservation. Genes on which the prioritized variants were found were further evaluated using several other databases (Gene Ontology; <http://geneontology.org/>, accessed on 26 June 2022, GeneCards; <https://www.genecards.org/>, accessed on 26 June 2022, KEGG; <https://www.genome.jp/kegg/>, accessed on 26 June 2022 and STRING; <https://string-db.org/>, accessed on 26 June 2022).

(a): High Impact variants [31]. At first, protein-truncating variants (PTVs) that lead to a truncated protein or its complete absence and thus, can have a severe consequence on protein function [32], as they have been associated with the causing of several diseases according to studies [33], were selected. The PTVs that were prioritized in this study included nonsense single-nucleotide variants (SNVs), frameshift insertions or deletions (indels), splice-disrupting variants, and start-loss SNVs. Common variants with an allele frequency  $> 0.05$  according to data retrieved from the 1000 Genomes Project [25] for the European population as well as from the gnomAD database [26] for Non-Finnish Europeans were excluded. This filter was also applied because the aim was to identify rare variants. According to studies, rare variants can help to shed light on common diseases as they are not found frequently in the population, and thus, there is a greater possibility to be associated with pathogenic phenotypes [13]. It should also be noted that information about allele frequencies was obtained only for the geographical regions mentioned above, as this study is strictly focused on the Greek population. In addition, as an extra filter, a CADD Score  $> 10$  was used to further prioritize the above variants. Any variant with a CADD Score  $> 10$  is considered to be in the top 10% of the human genome's likely functional and harmful variants.

(b): Moderate Impact Variants [31]. Moderate impact variants, including inframe indels, missense and protein-altering variants, though non-disruptive, have the potential to affect protein's effectiveness [31]. Of these, we selected to study missense variants as they have been more often associated with complex diseases [34]. In addition, missense variants that can affect protein functionality and which have a low frequency in a population may represent the "missing link" for explaining inheritable diseases [35]. Therefore, missense variants were also chosen and, to assess their effect on protein function, structure, and conservation, bioinformatics prediction tools were used. More specifically, only variants with SIFT Score [28]  $\leq 0.05$ , Polyphen2 Score [27]  $\geq 0.8$ , and with a medium or a high functional impact based on evolutionary conservation as assessed by MutationAssessor [30] were included in the analysis. The criteria regarding CADD score and allele frequency were also applied as in high-impact variants.

In all the above cases, novel variants that are not listed in the existing databases (Ensembl, 1000 Genomes and gnomAD), were also included in this study.

Furthermore, Gene Ontology (<http://www.geneontology.org/GO>, accessed on 26 June 2022), KEGG (<http://www.genome.jp/kegg/>, accessed on 26 June 2022), and GeneCards (<https://www.genecards.org/>, accessed on 26 June 2022) were used to investigate the role of the genes in which the prioritized variants were found and to identify pathways involved in teratozoospermia. STRING (<https://string-db.org/>, accessed on 26 June 2022) was also used to study interactions and correlations between proteins encoded by these genes. Genes with more than two prioritized variants were identified, as they may be more likely to be involved in the pathogenesis process and teratozoospermia.

Finally, Genotype-Tissue Expression (GTEx) database was used in an attempt to explore how the prioritized variants affect gene expression. GTEx Program is a database including information on the relationship between genetic variants and gene expression in multiple human tissues enabling among others the identification of expression quantitative trait loci (eQTLs) [36]. eQTLs are genomic loci that explain at least a fraction of the genetic variance of a gene expression phenotype and thus, can provide valuable information about the role of variants and their effect on phenotype [25].

### 3. Results

In brief, in this study, blood samples from normozoospermic and teratozoospermic individuals were used for DNA extraction. After sample preparation, whole-genome sequencing and data analysis were performed, in order to identify unique variants that are present only in teratozoospermic individuals and thus, can contribute to the pathogenic phenotype or have the potential to be used as biomarkers. Then, variant prioritization was performed by applying specific filters and selecting high- and moderate-impact variants that

have the greatest possibility to affect protein function and have a role in teratozoospermia according to bioinformatics tools. Therefore, the identified high and moderate impact variants are going to be presented in this section as well as the results of the analyses performed to investigate the role of the genes on which these were found, including GO annotation, KEGG enriched pathways, protein–protein interactions, etc.

### 3.1. Variant Calling and Annotation of WGS Data

After whole-genome sequencing, data analysis was performed. More specifically, the comparison between normozoospermic and teratozoospermic individuals to detect unique variants found only in one of the two groups revealed 617,722 variants specifically observed in teratozoospermic, while 2,342,243 variants were present only in normozoospermic men. The variants were mapped in 34,603 and 22,022 genes and characterized non-coding regions (miRNAs, lncRNAs genes, etc.) in normozoospermic and teratozoospermic males, respectively.

In the present study, only the variants identified in teratozoospermic individuals were selected for further analysis and prioritization, as the objective was to detect and investigate variants that contribute to the genomic profile of Greek patients with teratozoospermia. Of the 617,722 variants found in teratozoospermic men, 17.29% were mapped in intergenic regions, while the largest proportion (74.15%) was in intronic regions. The remaining variants were mapped in protein-coding regions and, more specifically, 0.62% and 1.07% of all variants were found in the 3' UTR and 5' UTR regions, respectively. Moreover, 0.40% of the unique teratozoospermic variants were also synonymous.

### 3.2. High Impact Variants Identification

In high impact variants, start-lost, nonsense, frameshift, and splice-disrupting variants were included. Thus, 175 variants were characterized first as high impact. Then, further filters were applied to prioritize them.

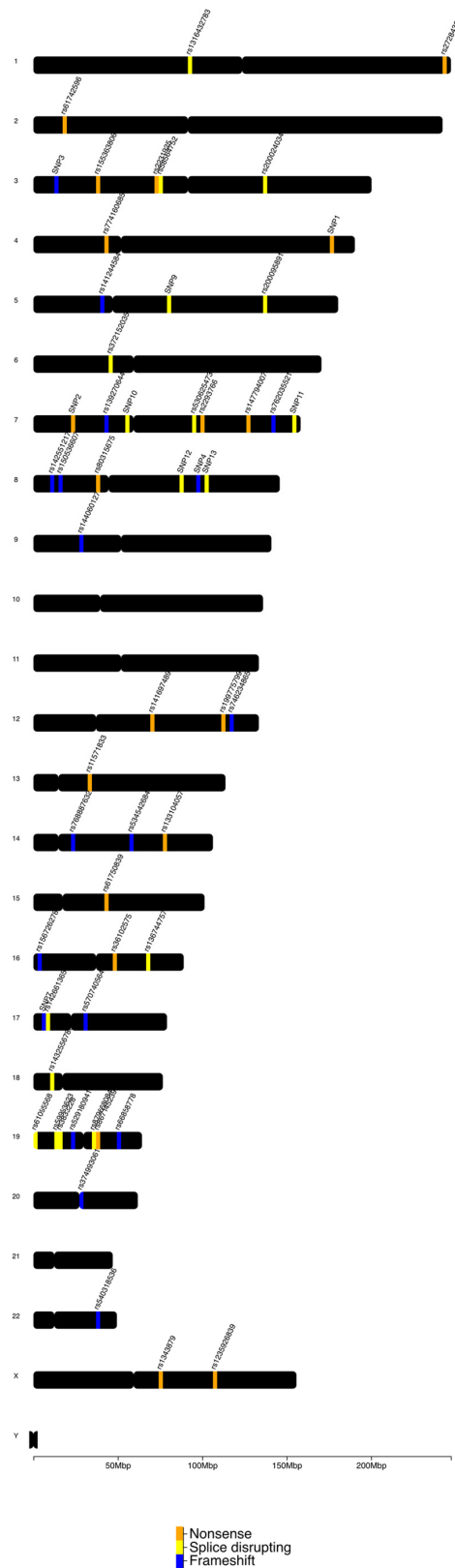
More specifically, of the 175 high-impact variants detected, 36 nonsense mutations were found among them. Variants with a CADD score less than 10, as well as variants with allele frequency in the European population greater than 5%, were removed; thus, 20 variants were identified (Table S1) and selected for further pathway enrichment analysis in the next steps. Regarding start-lost mutations, 3 variants specific for teratozoospermic individuals were detected from the WGS analysis, but after the filters referred above were applied, none of them was prioritized. However, 41 frameshift variants were found to be unique in teratozoospermic men, and after prioritization, 22 frameshift mutations (Table S2) were identified. Finally, the highest number of high impact variants were splice-disrupting mutations, since 48 were identified in teratozoospermic men. After filtering, 19 splice-disrupting variants were used for further pathway enrichment analysis (Table S3).

The prioritized variants are presented on whole-genome level in Figure 2.

### 3.3. Moderate Impact Variant Identification

Inframe indels, missense, and protein-altering variants can be characterized as moderate impact because they potentially affect protein function [31]. In the present study, among the variants found only in teratozoospermic individuals, 2141 moderate impact SNVs and indels were identified.

Then, to perform further prioritization, after applying the criteria described in the previous section (CADD Score > 10, Allele frequency (Europe) < 5%, SIFT Score ≤ 0.05, Polyphen2 Score ≥ 0.8, and medium/high impact according to MutationAssessor), 153 missense mutations were identified as the most likely to be associated with teratozoospermia in the Greek population (presented in Table S4).



**Figure 2.** Prioritized high impact variants found on teratozoospermic individuals based on their genomic location (whole-genome representation). Nonsense (orange), splice-disrupting (yellow), and frameshift variants (blue) are distinguished by color as presented. The figure was created with chromoMap [37].

### 3.4. Gene Investigation and Enrichment Analysis

After selecting all prioritized variants (total number: 214) as explained above, the next step was to evaluate the role of the genes on which these are found and to explore their potential contribution to male infertility and, particularly, teratozoospermia. More specifically, the full list of genes, as well as their description according to GeneCards [38], is presented in Table S5.

Based on these results, top genes with more than two prioritized variants from different categories (nonsense, frameshift, splice disrupting, missense) can also be identified as the accumulation of mutations may lead to a greater impact on protein function. These genes are presented in Table 2.

**Table 2.** Genes with more than two prioritized variants of different categories. Variants not found in the 1000 Genomes Project are highlighted because they were also searched in the gnomAD database. Ref; Reference allele, Obs; Observed allele.

Variant	Gene	Ref	Obs	Frequency (Europe)	Type of Mutation
rs11571833	BRCA2	A	T	0.011	Nonsense
rs766173	BRCA2	A	C	0.035	Missense
(7:55621478-55621503)	VOPPI	ACACACACACACA CACACACTCAC	CAC	-	Splice Disrupting
rs201023957	VOPPI	C	T	0.000471 (gnomAD)	Missense

In addition, KEGG enrichment analysis was performed for all the genes identified above. The extracellular matrix (ECM)—receptor interaction pathway was found to be the most gene-enriched and significant pathway, as presented in Table 3.

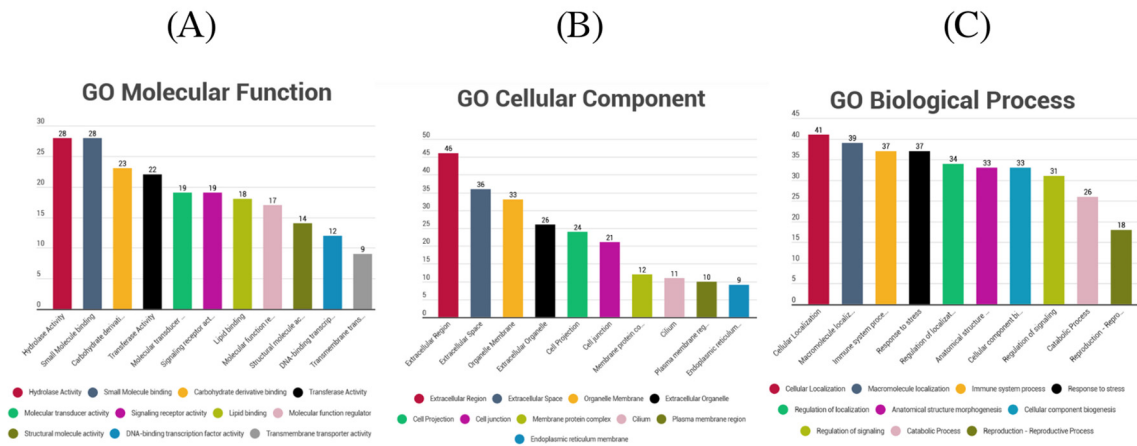
**Table 3.** Significant pathways identified after KEGG Enrichment Analysis. FDR; False discovery rate.

KEGG Enrichment Analysis				
Enrichment FDR	Number of Genes	Pathway Genes	Fold Enrichment	Pathways
0.0220	6	88	7.7	ECM—receptor interaction

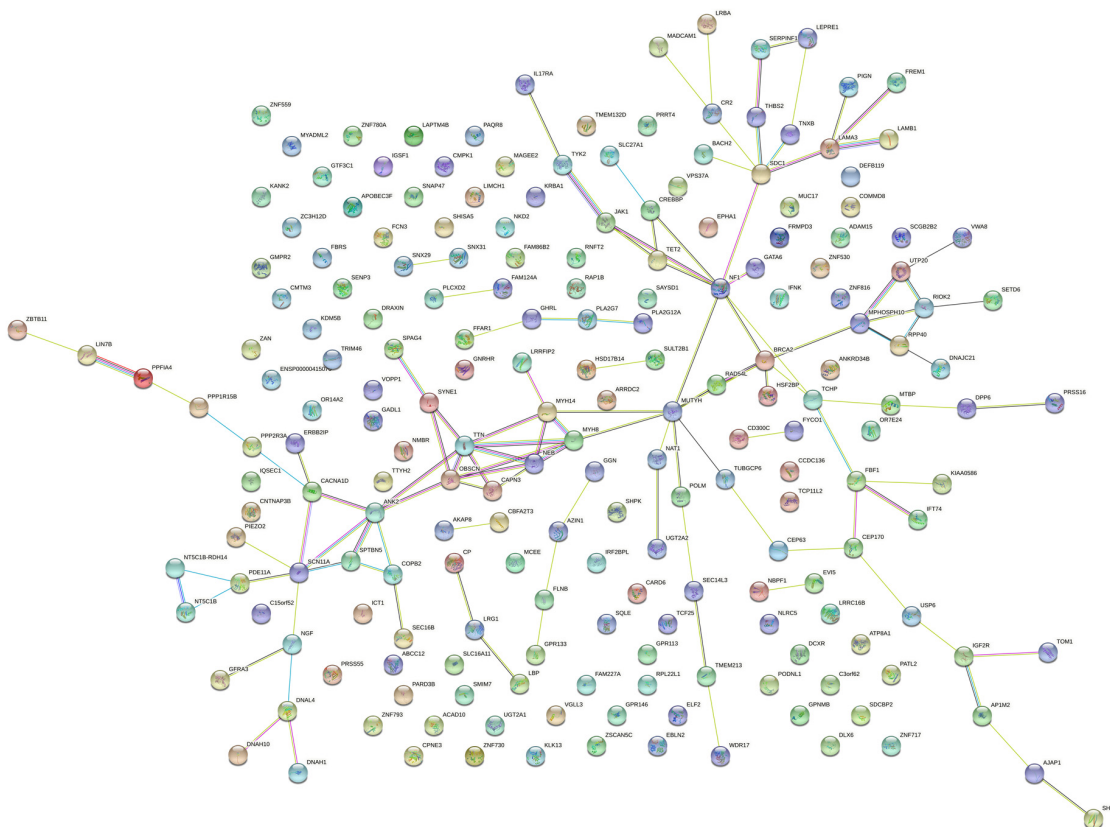
GO was also used to group genes by functional categories based on GO terms for biological processes, cellular components, and molecular functions. As our gene set was rather small, we performed gene ontology annotation instead of gene ontology enrichment in order to have reliable results and to obtain as much biological information as possible. The most important groups are presented for all three of them in Figure 3A–C. As it is observed, GO Cellular Component analysis revealed many molecules found in the extracellular region, whereas important categories identified in GO Biological Processes were morphogenesis of anatomical structures and reproduction. The most important GO terms regarding the molecular function can be considered the “molecular transducer”, “signaling receptor”, and transferases.

The STRING database was then used to study the network of the respective proteins of the genes identified above, as protein interaction networks can provide useful information on the molecular basis of complex diseases [39], such as male infertility (Figure 4). The network contained 202 nodes and 120 edges, while coiled-coil proteins were found to be enriched according to Uniprot [40] (Table 4).





**Figure 3.** Grouping of genes on which the prioritized variants were found according to GO terms for Molecular Function (A), GO terms for Cellular Component (B), and GO terms for Biological Process (C). The horizontal axis represents the GO terms and the vertical axis the number of genes that were found to be annotated on every GO term.

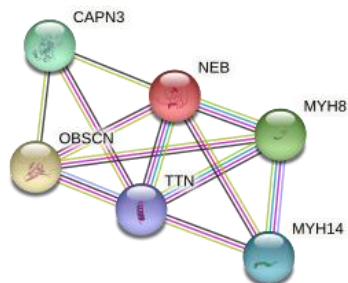


**Figure 4.** Protein–Protein Interaction (PPI) network for genes on which prioritized variants were found in teratozoospermic men. The network includes 202 nodes and 120 edges. PPI enrichment *p* value is 0.000939.

**Table 4.** Results of the functional enrichment analysis of the network according to Uniprot. The count in network indicates how many proteins in our network are annotated with the term “coiled-coil” of the total number of proteins assigned to the same term. Strength is calculated as  $\log_{10}(\text{observed}/\text{expected})$  for the number of proteins expected to be annotated with this term in a random network and finally, FDR (false discovery rate) is calculated after multiple testing correction according to Benjamini and Hochberg.

Enriched Annotated Keywords According to Uniprot			
Description term	Count in network	Strength	FDR
Coiled Coil	43 of 2139	0.290	0.0114

After that, we also used MCODE [41], a plug-in that is optimized to detect clusters in a network. More specifically, three clusters were calculated according to  $k\text{-core} = 2$ . Cluster 1 (Figure 5) contained 6 nodes and 12 edges and had the highest score among the identified clusters. The other clusters found on the network are presented in Table 5.



**Figure 5.** Cluster 1 of the PPI network identified according to MCODE. The cluster consists of 6 nodes and 12 edges.

**Table 5.** Clusters identified in the PPI network according to MCODE and their description based on Genecards.

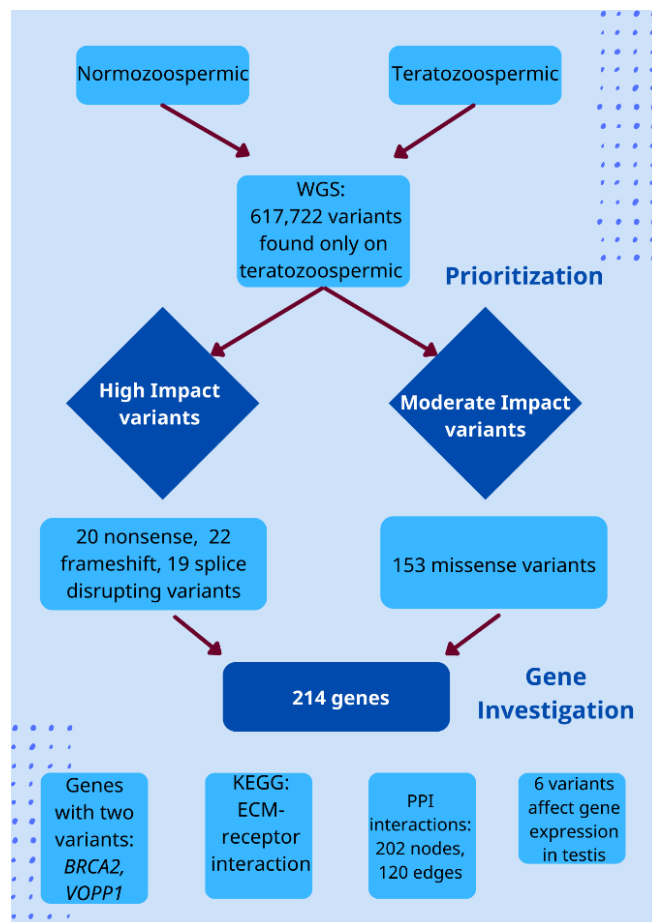
Cluster 2	
RIOK2	RIO Kinase 2
MPHOSPH10	M-Phase Phosphoprotein 10
RPP40	Ribonuclease P/MRP Subunit P40
UTP20	UTP20 Small Subunit Processome Component
Cluster 3	
NT5C1B	5'-Nucleotidase, Cytosolic 1B
NT5C1B-RDH14	NT5C1B-RDH14 Readthrough
PDE11A	Phosphodiesterase 11A

Finally, to investigate how the prioritized variants can affect gene expression, GTEx was used. It was revealed that some of them alter gene expression, particularly in the testicular tissue. The results are presented in Table 6.

**Table 6.** Prioritized variants on eQTL regions according to GTEx database.

Variant	Gene	Tissues
rs61742596	NT5C1B	Testis
rs61742596	NT5C1B-RDH14	Testis
rs61095568	MADCAM1	Testis
rs78440807	UTP20	Testis
rs113062332	FBF1	Testis
rs112138627	GGN	Testis
rs77267061	ZNF559	Testis

In brief, the main results of the present study are presented in Figure 6.



**Figure 6.** The main results of the present study. As presented, after WGS, 617,722 variants were identified only in teratozoospermic individuals. These were further prioritized to identify those with the greatest possibility to contribute to the pathogenic phenotype. Thus, 20 nonsense, 22 frameshift, and 19 splice-disrupting variants were characterized as high impact variants, being the most likely to affect protein function. A total of 153 missense variants were also identified. Further investigation of the genes on which these were found revealed that two genes carry more than one mutation in teratozoospermic individuals, and most genes play a role in the extracellular matrix–receptor interaction. In addition, many protein interactions were identified and some of the variants can affect gene.

#### 4. Discussion

In the present study, a WGS approach was used to identify and examine variants associated with clinical phenotypes of teratospermia. These may be involved in the molecular basis of teratozoospermia and have the potential to be used to improve the chances of successful ART or the diagnosis of male infertility in specific populations, since significant differences are observed and different genetic factors seem to contribute to infertility among ethnic groups [11,12]. Thus, to fill the gap regarding knowledge on polymorphisms associated with male infertility for Balkan populations, this study focused specifically on the Greek population and teratozoospermia.

Since next-generation sequencing (NGS) technologies provide massive amounts of data, a thoughtful approach should be followed to filter and select functional variants that are more likely to contribute to the pathogenic phenotype [42]. For this reason, this study focused only on protein-coding variants and after selecting variants found specifically in teratozoospermic individuals, a pipeline was developed that integrated the methodology

of Juhari et al. (2021) [43] to prioritize high- and moderate-impact variants involved in teratozoospermia in the Greek population.

#### 4.1. Genes Associated with Male Infertility

In the current study, we discovered high-impact variants specifically found in teratozoospermic men that were mapped in approximately 60 genes (Table S2). These should be of primary interest and focus in future disease-related studies as they can directly affect protein function [44]. Furthermore, the number of moderate impact or missense variants detected was higher, and these are also of great importance and could provide a list of candidate genes for future studies since they can affect protein structure and functionality leading to diseases [45].

Interestingly, many of our candidate genes have been associated in the past with abnormal spermatogenesis, sperm defects, or particular types of male infertility. Among the proteins required for successful fertilization is zonadhesin, involved in the species-specific adhesion of sperm to egg zona pellucida [46], and DCXR, also called “sperm surface protein P34H” [47]. Furthermore, *CREBP* has been suggested to play a role in azoospermia [48] and *FYCO1* was also observed recently to be involved in the regulation of the chromatoid body, which is crucial for spermatogenesis, through autophagy [49].

#### 4.2. Genes with Potential Role in Male Infertility and Teratozoospermia

For many of the highlighted genes with prioritized variants, there are indications of their involvement in male infertility, but further research is needed to investigate their exact role. Some of them are centrosomal protein 170 (*CEP170*) [50], gametogenetin (*GGN*) that is a testis-enriched gene [51], ghrelin, and obestatin, identified in human semen and the male reproductive system in general [52,53], Janus Kinase 1 (*JAK1*), found in the midpiece of spermatozoa and activated during capacitation [54,55], the lipopolysaccharide-binding protein that it is found in sperm and tail of spermatozoa [56] and tyrosine-kinase 2 (*TYK2*) that is active in human sperm [57] and plays a role in crucial signaling pathways. *TET2* is also expressed in the cytoplasm of late pachytene spermatocytes of Stage V and there are indications that its expression levels are associated with fertility status and sperm parameters [58].

In addition, several of the highlighted genes have been associated with spermatogenesis process or male fertility in other species. These genes should be of primary interest for future studies, as genes involved in spermatogenesis are highly conserved between species [59], and maybe similar molecular patterns are also involved in the presence of male infertility in humans.

For example, in mice, variants in Serine protease 55 (*Prss55*) [60,61], Centrosomal Protein 63 (*Cep63*) [62], *Ccdc136* [63], *Tom1* [64], *Maps* [65], *Ngf* [66], *Spag4* [67,68], *Pde11a* [69,70], *Atp8a1* [71,72], and *Mtbp* [73] have been associated with deficiencies in reproduction, by affecting among others meiotic recombination, spermatogenesis and sperm structure, acrosome formation, actin-based structures and p53 stability. In cattle, variants in *SEC16B* [74], *VGLL3* [75], and *DEBB119* [76] were also associated with male infertility. Finally, in boar, variants in *ADAM15* [77] and *KDM5B* [78] are associated with fertilization, sperm-egg binding and transcriptional activity in androgen receptors.

This study is also of particular interest because it draws attention to genes that have not been directly associated in the past with teratozoospermia or male infertility but are part of gene families that have members involved in the spermatogenesis process and reproduction in men. More specifically, laminins are important components of testicular basement membranes, and many laminin genes have been characterized as essential for normal testicular function [79]. Variants on laminin subunits have been also identified in the present study (*LAMA3*, *LAMB1*). Another protein family with a role in a variety of pathological mechanisms is Coiled-Coil Domain-Containing (CCDC) Proteins. Studies suggest that many members of this family, including *CCDC42*, *CCDC9*, and *CCDC87*, are required for the fertilization capacity of males and have been associated with abnormal

formation of the sperm flagella [80–82]. Coiled-coil proteins were also found to be enriched in the network produced using STRING, thus, *CCDC9B* is a very promising candidate gene for teratozoospermia. *ANK2* may also be a gene of interest, as Ankyrins are a family of proteins linking membrane and submembranous cytoskeletal proteins that play an important role in many cellular functions [83]. However, a specific gene of this family seems to be involved in reproduction as *Ankrd31*<sup>-/-</sup> mice are infertile possibly due to deregulation of the blood-epididymal barrier [84]. Other such genes are *CMTM3*, a member of the chemokine-like factor (CKLF)-like MARVEL transmembrane domain-containing family (CMTM) [85], *AKAP8*, a member of a-kinase anchoring proteins (AKAPs) [86] and *HSD17B14*, as hydroxysteroid (17-beta) dehydrogenase (HSD17B) gene family has a role in steroid hormone biosynthesis and deficiencies in such genes can even lead to sex development disorders [87]. Furthermore, *GATA6* is a candidate gene of importance, since GATA transcription factors play a crucial role in mammalian reproduction [88], whereas *KLK13* is part of a protein family involved in semen liquefaction [89] that has the potential to affect sperm quality [90].

Other candidate genes may also be of particular importance for future studies as they interact with proteins involved in reproduction. For example, *PPP1R15B* has been found to interact with Protein Phosphatase 1 which plays a key role in the spermatogenesis process and has been shown to affect sperm motility [91].

#### 4.3. The Special Case of *BRCA2*

It should also be noted that two different variants (missense and nonsense) found only in teratozoospermic individuals were identified in *BRCA2*. This gene has been associated with idiopathic cases of infertility characterized mainly by azoospermia or severe oligozoospermia [92], but in general, there are several studies linking polymorphisms in DNA repair genes with idiopathic infertility in males [93,94]. More interestingly, *BRCA2* exhibits a highly evolutionarily conserved interaction with *HSF2BP*, a testis-specific protein that is essential for mouse spermatogenesis [95]. In the present study, a missense variant characterized as moderate impact was also identified on *HSF2BP* in teratozoospermic individuals suggesting a role of this interaction in teratozoospermia that requires further investigation.

#### 4.4. The Role of Genes Associated with Cilia and Flagellum Malformation

Furthermore, the results of the present study are important as they support previous findings associating male infertility with dynein deficiencies [96–99].

Motile cilia and flagella are microtubule-based hair-like structures with a common core component: the axoneme [100]. The axoneme is an evolutionarily conserved structure from protozoa to humans, a fact underlying its significance [101,102]. In sperm flagella, axonemes have a typical “9 + 2” structure with 9 outer doublets of microtubules and a central pair of microtubules [100,103]. Several proteins are also attached to these, including axonemal dyneins classified as outer and inner dynein arms (ODAs and IDAs, respectively) [102,103]. ODAs and IDAs are multiprotein complexes consisting of light, intermediate, and heavy chain proteins [97]. In humans, there are 13 dynein axonemal heavy chain (DNAH) proteins (DNAH1–3, DNAH5–12, DNAH14, and DNAH17) [104] and mutations in many of these genes have been associated with male infertility [96,98,105–107].

In the present study, several mutations on some of the above genes specifically found in teratozoospermic men were identified. More specifically, a missense variant was identified in *DNAH1*. Mutations in *DNAH1* can cause morphological abnormalities of the sperm flagella and lead to male infertility [103,105,107]. Mutations in *DNAH10*, that exhibits testis-specific expression and causes asthenoteratozoospermia in humans and mice [96], were also identified in teratozoospermic men. Moreover, though Dynein Axonemal Light Chain 4 (*DNAL4*), on which a frameshift variant was found in teratozoospermic men, has not been recorded to be involved in teratozoospermia in humans, researchers observed that a mutation in this gene in boar affected sperm motility and caused midpiece abnormalities [108].

Furthermore, variants in teratozoospermic individuals were also identified in genes associated with primary ciliary dyskinesia (PCD), a disease characterized by many symptoms, including male infertility, due to abnormal motile cilia, or celiac disease [103]. These genes affect cilia biogenesis and structure and more specifically, in the present study variants in *FBF1*, *KIAA0586*, *ODAD1*, *GHRL*, *BACH2*, and *IFT74* were identified. Interestingly, except for its association with primary ciliary dyskinesia, intraflagellar transport protein 74 has been also associated with male infertility as for many years studies highlighted that it was an essential component for the spermatogenesis process in mice [109].

These findings suggest that maybe teratozoospermia is characterized by a set of mutations associated with cilia and flagellum malformation that act accumulative resulting in sperm deficiencies. However, the interesting fact is that these mutations identified here alter sperm morphology but do not have such a strong impact on sperm motility as sperm samples of the present study were characterized only as teratozoospermic and no combinations of male infertility subcategories were included (e.g., asthenoteratozoospermia, etc.). Thus, the identification of these mutations could be useful for distinguishing teratozoospermia from other subcategories of male infertility and providing implications for successful diagnosis. Furthermore, the study of teratozoospermic men harboring mutations that lead to sperm tail abnormalities, as usually are mutations in dyneins, is of importance because studies show that a good sperm nuclear quality in combination with mutations affecting the sperm tail are usually indicators of good embryonic development after intracytoplasmic sperm injection (ICSI) [97]. In particular, mutations in *DNAH1* have been associated in the past with a good pregnancy rate [110], suggesting that further study is required to identify more variants on dynein genes that might contribute to teratozoospermia and have the potential to improve ART outcome or prognosis.

#### 4.5. The Role of the Extracellular Matrix in Teratozoospermia

Finally, in the present study, pathway enrichment analysis showed that ECM–receptor interaction was the most significant pathway according to the genes on which prioritized variants were found specifically in Greek teratozoospermic individuals. GO analyses also identified variants in teratozoospermic men on a large number of proteins found in the extracellular region or proteins that are part of the extracellular matrix (ECM), suggesting a crucial role of the ECM in teratozoospermia.

ECM consists mainly of glycoproteins and polysaccharides but, most importantly, its components interact with a wide range of molecules, including proteases, protease inhibitors, cytokines, etc. [111,112]. Thus, this network of ECM proteins and their partners play a crucial role in the regulation of junction dynamics in testis [111,113]. Junction restructuring is required during germ cell movement in the seminiferous epithelium [112], but as this is a very complex process mediated by several mechanisms [113], it is proposed that protein deficiencies in this network can affect the spermatogenesis process and result in male infertility, as the results of this study indicate.

Although many studies are highlighting the important role of ECM in the spermatogenesis process, there are no studies directly associating the extracellular matrix with teratozoospermia. However, scientists have previously reported that abnormal basement membrane structures were observed in infertile men with aspermatogenesis [114]. Scientists have also detected in the past that disruption of pathways involved in the ECM and junction dynamics can result in male infertility because germ cells are depleted from the epithelium [112]. Taking into consideration these findings and the results presented here it can be suggested that ECM could play a crucial role in teratozoospermia which was underrated as the abnormal translocation of the germ cells across the seminiferous epithelium can also affect sperm morphology. Therefore, more research is required to identify the molecular mechanism that links ECM with teratozoospermia.

#### 4.6. Prioritized Variants' Effect on Gene Expression

Investigation of the effect of prioritized variants on gene expression based on the GTEx database for various tissues revealed that some of them affect the mRNA levels of genes on testis tissue (Table 6); more specifically, among them, NT5C1B codes for a protein called autoimmune infertility-related protein that is highly expressed in testis [115]. FBF1 is another interesting gene whose expression seems to be affected by rs113062332. This gene is associated with primary ciliary dyskinesia and in addition, studies in *Drosophila* prove that is essential for male fertility as RNAi-knockdown flies have impaired sperm flagella and are infertile [116]. GGN is a testis-enriched gene that has also been implicated to have a role in male infertility [51] too.

#### 4.7. Directions for Future Studies

Future studies aiming at the validation of these findings can be experiments of different types. More specifically, GWAS can be used to investigate if the variants found here are associated with teratozoospermia in large samples of controls–cases. RNA-sequencing experiments can also add valuable information about how the variants affect gene expression and which of the genes identified here are deregulated in teratozoospermia. Finally, functional experiments could also validate the effect of variants on protein function and provide knowledge about their specific impact on the phenotype observed, teratozoospermia.

Finally, though the variants were filtered and selected to have a low allele frequency, lower than 5%, and it is not likely to be in homozygosity, it should be investigated experimentally, e.g., PCR, if all the variants identified in this paper are found in homozygous or heterozygous state in a large sample of teratozoospermic individuals. More experiments are required to explore if these mutations are dominant or recessive and how they affect protein function potentially using animal models. This information will be extremely valuable, as in dominant mutations that cause male infertility there is a high probability that the use of assisted reproductive technologies will lead to the transfer of this allele to the next generation [117].

### 5. Conclusions

This study is the first comprehensive investigation of the genomic profile of teratozoospermic patients in the Greek population using WGS. It is of importance because it provides a roadmap for future studies, enlisting candidate genes and variants that are associated with teratozoospermia for the first time here and confirming the role of genes that have been studied in the past in male infertility. In particular, the stop-codon variants presented in the present study, that cause termination in protein production, should be further explored as they may shed light on molecular mechanisms and pathways that were previously underrated. Additionally, missense variants detected in teratozoospermic individuals should also be examined. The identification of the important role of the extracellular matrix and the process of cilia and flagellum formation as well as their direct association with teratozoospermia is very promising for future research, too.

However, the small sample size of the patients recruited is a limitation of the present study. Therefore, validation of these findings in a larger sample could provide more definitive evidence for the role of the variants and the candidate genes in teratozoospermia in the Greek population. Furthermore, assisted reproductive technology (ART) has finally expanded the opportunities for infertile couples, but previous studies have revealed the different outcomes of intracytoplasmic sperm injection (ICSI) for individuals harboring different mutations [11,110]. Thus, further experiments are required to assess and explore the impact of the variants identified in the present study on ICSI outcomes. In this way, by analyzing the genetic profile of a man with teratozoospermia, ICSI would be recommended or not based on the identification of specific mutations as it seems that these affecting only the flagellum structure may have better chances for successful ART outcomes [97]. Finally, it should be noted that synonymous variants and variants in non-coding regions were excluded in this study, but they can also provide useful information and should be

investigated for their contribution to teratozoospermia in the future as they can help to fully assess the whole-genome profile of patients.

In conclusion, the present study does not provide conclusive evidence on specific mutations, but the analysis of whole-genome data contributes to our understanding of teratozoospermia by highlighting important pathways and provides the foundation for the improvement of ART and the successful diagnosis or prognosis as a wide spectrum of variants was identified, as well as genes and pathways that had not been explored in the past acting as promising candidates for future research.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/genes13091606/s1>, Table S1: Prioritized nonsense variants found in teratozoospermic individuals. Variants not found in the 1000 Genomes Project are highlighted because they were also searched in the gnomAD database. Ref; Reference allele, Obs; Observed allele, Table S2: Prioritized frameshift variants found in teratozoospermic individuals. Variants not found in the 1000 Genomes Project are highlighted because they were also searched in the gnomAD database. Ref; Reference allele, Obs; Observed allele, Table S3: Prioritized splice-disrupting variants found in teratozoospermic individuals. Variants not found in the 1000 Genomes Project are highlighted because they were also searched in the gnomAD database. Ref; Reference allele, Obs; Observed allele, Table S4: Prioritized moderate impact variants found in teratozoospermic individuals; Table S5: Genes on which the prioritized variants (high and moderate impact) were found in teratozoospermic individuals and their description according to GeneCards. The results are presented for every group of variants studied (nonsense, frameshift, splice disrupting, and missense).

**Author Contributions:** Conceptualization, Z.M. and M.-A.K.; data analysis, M.-A.K. and T.G.; data interpretation, T.G. and M.-A.K.; data acquisition, A.C. and N.C.; writing—original draft preparation, M.-A.K.; writing—review and editing, Z.M. and T.G.; supervision, Z.M.; funding acquisition, Z.M. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was supported by the Spermogene project which is co-financed by the European Regional Development Fund of the European Union and Greek national funds through the Operational Program Competitiveness, Entrepreneurship, and Innovation, under the call RESEARCH-CREATE-INNOVATE (Grant number T1EΔK-02787).

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki and approved by the ethics committee of the Medical Faculty of the University of Thessaly (protocol code 20.04/2016 and date of approval 20 April 2016).

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The sequencing data presented in this study are available through SRA (BioProject ID PRJNA875412, <http://www.ncbi.nlm.nih.gov/bioproject/875412> (accessed on 26 June 2022)).

**Acknowledgments:** The authors wish to thank all the volunteers for their participation in this study. We would also like to thank Maria Markantoni for sample handling and support.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Zegers-Hochschild, F.; Adamson, G.D.; Dyer, S.; Racowsky, C.; de Mouzon, J.; Sokol, R.; Rienzi, L.; Sunde, A.; Schmidt, L.; Cooke, I.D.; et al. The International Glossary on Infertility and Fertility Care, 2017. *Fertil. Steril.* **2017**, *108*, 393–406. [[CrossRef](#)] [[PubMed](#)]
2. Zhao, Q.; Huangfu, C.; Li, J.; Liu, H.; Tang, N. Psychological Resilience as the Mediating Factor Between Stigma and Social Avoidance and Distress of Infertility Patients in China: A Structural Equation Modeling Analysis. *Psychol. Res. Behav. Manag.* **2022**, *15*, 391. [[CrossRef](#)] [[PubMed](#)]
3. Slade, P.; O'Neill, C.; Simpson, A.J.; Lashen, H. The Relationship between Perceived Stigma, Disclosure Patterns, Support and Distress in New Attendees at an Infertility Clinic. *Hum. Reprod.* **2007**, *22*, 2309–2317. [[CrossRef](#)]
4. Wu, A.K.; Odisho, A.Y.; Washington, S.L.; Katz, P.P.; Smith, J.F. Out-of-Pocket Fertility Patient Expense: Data from a Multicenter Prospective Infertility Cohort. *J. Urol.* **2014**, *191*, 427–432. [[CrossRef](#)] [[PubMed](#)]
5. Inhorn, M.C.; Patrizio, P. Infertility around the Globe: New Thinking on Gender, Reproductive Technologies and Global Movements in the 21st Century. *Hum. Reprod. Update* **2015**, *21*, 411–426. [[CrossRef](#)]



6. Agarwal, A.; Mulgund, A.; Hamada, A.; Chyatte, M.R. A Unique View on Male Infertility around the Globe. *Reprod. Biol. Endocrinol.* **2015**, *13*, 37. [[CrossRef](#)] [[PubMed](#)]
7. Agarwal, A.; Baskaran, S.; Parekh, N.; Cho, C.L.; Henkel, R.; Vij, S.; Arafa, M.; Panner Selvam, M.K.; Shah, R. Male Infertility. *Lancet* **2021**, *397*, 319–333. [[CrossRef](#)]
8. Kothandaraman, N.; Agarwal, A.; Abu-Elmagd, M.; Al-Qahtani, M.H. Pathogenic Landscape of Idiopathic Male Infertility: New Insight towards Its Regulatory Networks. *NPJ Genom. Med.* **2016**, *1*, 16023. [[CrossRef](#)]
9. Krausz, C.; Riera-Escamilla, A. Genetics of Male Infertility. *Nat. Rev. Urol.* **2018**, *15*, 369–384. [[CrossRef](#)]
10. De Braekeleer, M.; Nguyen, M.H.; Morel, F.; Perrin, A. Genetic Aspects of Monomorphic Teratozoospermia: A Review. *J. Assist. Reprod. Genet.* **2015**, *32*, 615. [[CrossRef](#)]
11. Spitzer, T.; Fujimoto, V. Ethnic Differences in Assisted Reproductive Technologies Outcomes. *Semin. Reprod. Med.* **2013**, *31*, 360–364. [[CrossRef](#)] [[PubMed](#)]
12. Lawrenz, B.; Coughlan, C.; Melado, L.; Fatemi, H.M. Ethnical and Sociocultural Differences Causing Infertility Are Poorly Understood—Insights from the Arabian Perspective. *J. Assist. Reprod. Genet.* **2019**, *36*, 661. [[CrossRef](#)] [[PubMed](#)]
13. Pierre, A.S.; Génin, E. How Important Are Rare Variants in Common Disease? *Brief. Funct. Genomics* **2014**, *13*, 353–361. [[CrossRef](#)]
14. Yazdani, A.; Yazdani, A.; Liu, X.; Boerwinkle, E. Identification of Rare Variants in Metabolites of the Carnitine Pathway by Whole Genome Sequencing Analysis. *Genet. Epidemiol.* **2016**, *40*, 486. [[CrossRef](#)]
15. Meienberg, J.; Bruggmann, R.; Oexle, K.; Matyas, G. Clinical Sequencing: Is WGS the Better WES? *Hum. Genet.* **2016**, *135*, 359–362. [[CrossRef](#)]
16. Chandra, A.; Copen, C.E.; Stephen, E.H. Infertility and Impaired Fecundity in the United States, 1982–2010: Data from the National Survey of Family Growth. *Natl. Health Stat. Rep.* **2013**, *67*, 1–19.
17. Skakkebaek, N.E.; Rajpert-De Meyts, E.; Buck Louis, G.M.; Toppari, J.; Andersson, A.M.; Eisenberg, M.L.; Jensen, T.K.; Jørgensen, N.; Swan, S.H.; Sapra, K.J.; et al. Male Reproductive Disorders and Fertility Trends: Influences of Environment and Genetic Susceptibility. *Physiol. Rev.* **2016**, *96*, 55–97. [[CrossRef](#)]
18. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
19. Howe, K.L.; Achuthan, P.; Allen, J.; Allen, J.; Alvarez-Jarreta, J.; Ridwan Amode, M.; Armean, I.M.; Azov, A.G.; Bennett, R.; Bhai, J.; et al. Ensembl 2021. *Nucleic Acids Res.* **2021**, *49*, D884–D891. [[CrossRef](#)]
20. Li, H.; Durbin, R. Fast and Accurate Short Read Alignment with Burrows-Wheeler Transform. *Bioinformatics* **2009**, *25*, 1754–1760. [[CrossRef](#)]
21. Danecek, P.; Bonfield, J.K.; Liddle, J.; Marshall, J.; Ohan, V.; Pollard, M.O.; Whitwham, A.; Keane, T.; McCarthy, S.A.; Davies, R.M.; et al. Twelve Years of SAMtools and BCFtools. *Gigascience* **2021**, *10*, giab008. [[CrossRef](#)] [[PubMed](#)]
22. Garrison, E.; Marth, G. Haplotype-Based Variant Detection from Short-Read Sequencing. *arXiv* **2012**, arXiv:1207.3907.
23. McCarthy, D.J.; Humburg, P.; Kanapin, A.; Rivas, M.A.; Gaulton, K.; Cazier, J.B.; Donnelly, P. Choice of Transcripts and Software Has a Large Effect on Variant Annotation. *Genome Med.* **2014**, *6*, 26. [[CrossRef](#)] [[PubMed](#)]
24. Sherry, S.T.; Ward, M.H.; Kholodov, M.; Baker, J.; Phan, L.; Smigielski, E.M.; Sirotkin, K. dbSNP: The NCBI Database of Genetic Variation. *Nucleic Acids Res.* **2001**, *29*, 308. [[CrossRef](#)] [[PubMed](#)]
25. Auton, A.; Abecasis, G.R.; Altshuler, D.M.; Durbin, R.M.; Bentley, D.R.; Chakravarti, A.; Clark, A.G.; Donnelly, P.; Eichler, E.E.; Flicek, P.; et al. A Global Reference for Human Genetic Variation. *Nature* **2015**, *526*, 68–74. [[CrossRef](#)]
26. Gudmundsson, S.; Singer-Berk, M.; Watts, N.A.; Phu, W.; Goodrich, J.K.; Solomonson, M.; Rehm, H.L.; MacArthur, D.G.; O'Donnell-Luria, A. Variant Interpretation Using Population Databases: Lessons from GnomAD. *Hum. Mutat.* **2021**, *43*, 1012–1030. [[CrossRef](#)]
27. Adzhubei, I.A.; Schmidt, S.; Peshkin, L.; Ramensky, V.E.; Gerasimova, A.; Bork, P.; Kondrashov, A.S.; Sunyaev, S.R. A Method and Server for Predicting Damaging Missense Mutations. *Nat. Methods* **2010**, *7*, 248–249. [[CrossRef](#)]
28. Vaser, R.; Adusumalli, S.; Ngak Leng, S.; Sikic, M.; Ng, P.C. SIFT Missense Predictions for Genomes. *Nat. Protoc.* **2015**, *11*, 1–9. [[CrossRef](#)]
29. Rentzsch, P.; Witten, D.; Cooper, G.M.; Shendure, J.; Kircher, M. CADD: Predicting the Deleteriousness of Variants throughout the Human Genome. *Nucleic Acids Res.* **2019**, *47*, D886–D894. [[CrossRef](#)]
30. Reva, B.; Antipin, Y.; Sander, C. Predicting the Functional Impact of Protein Mutations: Application to Cancer Genomics. *Nucleic Acids Res.* **2011**, *39*, e118. [[CrossRef](#)]
31. McLaren, W.; Pritchard, B.; Rios, D.; Chen, Y.; Flicek, P.; Cunningham, F. Deriving the Consequences of Genomic Variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics* **2010**, *26*, 2069–2070. [[CrossRef](#)]
32. Deboever, C.; Tanigawa, Y.; Lindholm, M.E.; McInnes, G.; Lavertu, A.; Ingelsson, E.; Chang, C.; Ashley, E.A.; Bustamante, C.D.; Daly, M.J.; et al. Medical Relevance of Protein-Truncating Variants across 337,205 Individuals in the UK Biobank Study. *Nat. Commun.* **2018**, *9*, 1612. [[CrossRef](#)]
33. Rivas, M.A.; Pirinen, M.; Conrad, D.F.; Lek, M.; Tsang, E.K.; Karczewski, K.J.; Maller, J.B.; Kukurba, K.R.; DeLuca, D.S.; Fromer, M.; et al. Impact of Predicted Protein-Truncating Genetic Variants on the Human Transcriptome. *Science* **2015**, *348*, 666. [[CrossRef](#)] [[PubMed](#)]
34. Kryukov, G.V.; Pennacchio, L.A.; Sunyaev, S.R. Most Rare Missense Alleles Are Deleterious in Humans: Implications for Complex Disease and Association Studies. *Am. J. Hum. Genet.* **2007**, *80*, 727–739. [[CrossRef](#)]

35. Zeng, J.; Slodkowicz, G.; James, L.C. Rare Missense Variants in the Human Cytosolic Antibody Receptor Preserve Antiviral Function. *Elife* **2019**, *8*, e48339. [[CrossRef](#)] [[PubMed](#)]
36. Lonsdale, J.; Thomas, J.; Salvatore, M.; Phillips, R.; Lo, E.; Shad, S.; Hasz, R.; Walters, G.; Garcia, F.; Young, N.; et al. The Genotype-Tissue Expression (GTEx) Project. *Nat. Genet.* **2013**, *45*, 580. [[CrossRef](#)]
37. Anand, L.; Rodriguez Lopez, C.M. ChromoMap: An R package for interactive visualization of multi-omics data and annotation of chromosomes. *BMC Bioinform.* **2019**, *23*, 33. [[CrossRef](#)]
38. Stelzer, G.; Rosen, N.; Plaschkes, I.; Zimmerman, S.; Twik, M.; Fishilevich, S.; Stein, T.I.; Nudel, R.; Lieder, I.; Mazor, Y.; et al. The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr. Protoc. Bioinforma.* **2016**, *54*, 1.30.1–1.30.33. [[CrossRef](#)] [[PubMed](#)]
39. Farooq, Q.I.A.; Shaukat, Z.; Aiman, S.; Li, C.-H. Protein-Protein Interactions: Methods, Databases, and Applications in Virus-Host Study. *World J. Virol.* **2021**, *10*, 288–300. [[CrossRef](#)]
40. Bateman, A.; Martin, M.J.; Orchard, S.; Magrane, M.; Agivetova, R.; Ahmad, S.; Alpi, E.; Bowler-Barnett, E.H.; Britto, R.; Bursteinas, B.; et al. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49*, D480–D489. [[CrossRef](#)]
41. Bader, G.D.; Hogue, C.W. An Automated Method for Finding Molecular Complexes in Large Protein Interaction Networks. *BMC Bioinform.* **2003**, *4*, 2. [[CrossRef](#)]
42. Roca, I.; Fernández-Marmiesse, A.; Gouveia, S.; Segovia, M.; Couce, M.L. Prioritization of Variants Detected by Next Generation Sequencing According to the Mutation Tolerance and Mutational Architecture of the Corresponding Genes. *Int. J. Mol. Sci.* **2018**, *19*, 1584. [[CrossRef](#)] [[PubMed](#)]
43. Juhari, W.K.W.; Noordin, K.B.A.A.; Zakaria, A.D.; Rahman, W.F.W.A.; Mokhter, W.M.M.W.M.; Hassan, M.R.A.; Sidek, A.S.M.; Zilfalil, B.A. Whole-Genome Profiles of Malay Colorectal Cancer Patients with Intact Mmr Proteins. *Genes* **2021**, *12*, 1448. [[CrossRef](#)]
44. Dashti, M.J.S.; Gamiieldien, J. A Practical Guide to Filtering and Prioritizing Genetic Variants. *Biotechniques* **2017**, *62*, 18–30. [[CrossRef](#)] [[PubMed](#)]
45. Stefl, S.; Nishi, H.; Petukh, M.; Panchenko, A.R.; Alexov, E. Molecular Mechanisms of Disease-Causing Missense Mutations. *J. Mol. Biol.* **2013**, *425*, 3919. [[CrossRef](#)] [[PubMed](#)]
46. Tardif, S.; Cormier, N. Role of Zonadhesin during Sperm–Egg Interaction: A Species-Specific Acrosomal Molecule with Multiple Functions. *Mol. Hum. Reprod.* **2011**, *17*, 661–668. [[CrossRef](#)]
47. Ebert, B.; Kisiela, M.; Maser, E. Human DCXR—Another “moonlighting Protein” Involved in Sugar Metabolism, Carbonyl Detoxification, Cell Adhesion and Male Fertility? *Biol. Rev. Camb. Philos. Soc.* **2015**, *90*, 254–278. [[CrossRef](#)]
48. Sabetian, S.; Shamsir, M.S. Systematic Analysis of Protein Interaction Network Associated with Azoospermia. *Int. J. Mol. Sci.* **2016**, *17*, 1857. [[CrossRef](#)]
49. Da Ros, M.; Lehtiniemi, T.; Olotu, O.; Fischer, D.; Zhang, F.P.; Vihinen, H.; Jokitalo, E.; Sironen, A.; Toppari, J.; Kotaja, N. FYCO1 and Autophagy Control the Integrity of the Haploid Male Germ Cell-Specific RNP Granules. *Autophagy* **2017**, *13*, 302–321. [[CrossRef](#)]
50. Omolaoye, T.S.; Omolaoye, V.A.; Kandasamy, R.K.; Hachim, M.Y.; Du Plessis, S.S. Omics and Male Infertility: Highlighting the Application of Transcriptomic Data. *Life* **2022**, *12*, 280. [[CrossRef](#)]
51. Jamsai, D.; Sarraj, M.A.; Merriner, D.J.; Drummond, A.E.; Jones, K.T.; McLachlan, R.I.; O’Byrne, M.K. GGN1 in the Testis and Ovary and Its Variance within the Australian Fertile and Infertile Male Population. *Int. J. Androl.* **2011**, *34*, 624–632. [[CrossRef](#)] [[PubMed](#)]
52. Moretti, E.; Vindigni, C.; Tripodi, S.A.; Mazzi, L.; Nuti, R.; Figura, N.; Collodel, G. Immunolocalisation of Ghrelin and Obestatin in Human Testis, Seminal Vesicles, Prostate and Spermatozoa. *Andrologia* **2014**, *46*, 979–985. [[CrossRef](#)] [[PubMed](#)]
53. Moretti, E.; Collodel, G.; Iacoponi, F.; Geminiani, M.; Pascarelli, N.A.; Campagna, S.; Franci, B.; Figura, N. Detection of Obestatin in Seminal Plasma and Its Relationship with Ghrelin and Semen Parameters. *Fertil. Steril.* **2011**, *95*, 2303–2309. [[CrossRef](#)] [[PubMed](#)]
54. Lachance, C.; Leclerc, P. Mediators of the Jak/STAT Signaling Pathway in Human Spermatozoa. *Biol. Reprod.* **2011**, *85*, 1222–1231. [[CrossRef](#)] [[PubMed](#)]
55. Sagare-Patil, V.; Modi, D. Progesterone Activates Janus Kinase 1/2 and Activators of Transcription 1 (JAK1-2/STAT1) Pathway in Human Spermatozoa. *Andrologia* **2013**, *45*, 178–186. [[CrossRef](#)] [[PubMed](#)]
56. Malm, J.; Nordahl, E.A.; Bjartell, A.; Sørensen, O.E.; Frohm, B.; Dentener, M.A.; Egesten, A. Lipopolysaccharide-Binding Protein Is Produced in the Epididymis and Associated with Spermatozoa and Prostatomes. *J. Reprod. Immunol.* **2005**, *66*, 33–43. [[CrossRef](#)]
57. D’Cruz, O.J.; Vassilev, A.O.; Uckun, F.M. Members of the Janus Kinase/Signal Transducers and Activators of Transcription (JAK/STAT) Pathway Are Present and Active in Human Sperm. *Fertil. Steril.* **2001**, *76*, 258–266. [[CrossRef](#)]
58. Ni, K.; Dansranjav, T.; Rogenhofer, N.; Oeztuerk, N.; Deuker, J.; Bergmann, M.; Schuppe, H.C.; Wagenlehner, F.; Weidner, W.; Steger, K.; et al. TET Enzymes Are Successively Expressed during Human Spermatogenesis and Their Expression Level Is Pivotal for Male Fertility. *Hum. Reprod.* **2016**, *31*, 1411–1424. [[CrossRef](#)]
59. White-Cooper, H.; Bausek, N. Evolution and Spermatogenesis. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 1465. [[CrossRef](#)]
60. Kobayashi, K.; Endo, T.; Matsumura, T.; Lu, Y.; Yu, Z.; Matzuk, M.M.; Ikawa, M. Prss55 but Not Prss51 Is Required for Male Fertility in Mice. *Biol. Reprod.* **2020**, *103*, 223–234. [[CrossRef](#)]

61. Shang, X.; Shen, C.; Liu, J.; Tang, L.; Zhang, H.; Wang, Y.; Wu, W.; Chi, J.; Zhuang, H.; Fei, J.; et al. Serine Protease PRSS55 Is Crucial for Male Mouse Fertility via Affecting Sperm Migration and Sperm-Egg Binding. *Cell. Mol. Life Sci.* **2018**, *75*, 4371–4384. [[CrossRef](#)] [[PubMed](#)]
62. Marjanović, M.; Sánchez-Huertas, C.; Terré, B.; Gómez, R.; Scheel, J.F.; Pacheco, S.; Knobel, P.A.; Martínez-Marchal, A.; Aivio, S.; Palenzuela, L.; et al. CEP63 Deficiency Promotes P53-Dependent Microcephaly and Reveals a Role for the Centrosome in Meiotic Recombination. *Nat. Commun.* **2015**, *6*, 7676. [[CrossRef](#)] [[PubMed](#)]
63. Geng, Q.; Ni, L.; Ouyang, B.; Hu, Y.; Zhao, Y.; Guo, J. A Novel Testis-Specific Gene, Ccdc136, Is Required for Acrosome Formation and Fertilization in Mice. *Reprod. Sci.* **2016**, *23*, 1387–1396. [[CrossRef](#)] [[PubMed](#)]
64. Zakrzewski, P.; Suwińska, A.; Lenartowski, R.; Rędownicz, M.J.; Buss, F.; Lenartowska, M. Myosin VI Maintains the Actin-Dependent Organization of the Tubulobulbar Complexes Required for Endocytosis during Mouse Spermiogenesis. *Biol. Reprod.* **2020**, *102*, 863. [[CrossRef](#)] [[PubMed](#)]
65. Li, M.; Zheng, J.; Li, G.; Lin, Z.; Li, D.; Liu, D.; Feng, H.; Cao, D.; Ng, E.H.Y.; Li, R.H.W.; et al. The Male Germline-Specific Protein MAPS Is Indispensable for Pachynema Progression and Fertility. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2025421118. [[CrossRef](#)] [[PubMed](#)]
66. Luo, J.; Yang, Y.; Ji, X.; He, W.; Fan, J.; Huang, Y.; Wang, Y. NGF Rescues Spermatogenesis in Azoospermic Mice. *Reprod. Sci.* **2021**, *28*, 2780–2788. [[CrossRef](#)] [[PubMed](#)]
67. Yeh, C.H.; Kuo, P.L.; Wang, Y.Y.; Chen, M.F.; Lin, D.Y.; Lai, T.H.; Chiang, H.S.; Lin, Y.H. SEPT12/SPAG4/LAMINB1 Complexes Are Required for Maintaining the Integrity of the Nuclear Envelope in Postmeiotic Male Germ Cells. *PLoS ONE* **2015**, *10*, e0120722. [[CrossRef](#)]
68. Yang, K.; Adham, I.M.; Meinhardt, A.; Hoyer-Fender, S. Ultra-Structure of the Sperm Head-to-Tail Linkage Complex in the Absence of the Spermatid-Specific LINC Component SPAG4. *Histochem. Cell Biol.* **2018**, *150*, 49–59. [[CrossRef](#)]
69. Horvath, A.; Korde, L.; Greene, M.H.; Libe, R.; Osorio, P.; Faucz, F.R.; Raffin-Sanson, M.L.; Kit, M.T.; Drori-Herishanu, L.; Patronas, Y.; et al. Functional Phosphodiesterase 11A Mutations May Modify the Risk of Familial and Bilateral Testicular Germ Cell Tumors. *Cancer Res.* **2009**, *69*, 5301–5306. [[CrossRef](#)]
70. Wayman, C.; Phillips, S.; Lunny, C.; Webb, T.; Fawcett, L.; Baxendale, R.; Burgess, G. Phosphodiesterase 11 (PDE11) Regulation of Spermatozoa Physiology. *Int. J. Impot. Res.* **2005**, *17*, 216–223. [[CrossRef](#)]
71. Yap, Y.T.; Li, Y.H.; Li, W.; Banerjee, P.; Zhang, Z. ATP8a1, an IFT27 Binding Partner, Is Dispensable for Spermatogenesis and Male Fertility. *Mol. Reprod. Dev.* **2021**, *88*, 371–375. [[CrossRef](#)] [[PubMed](#)]
72. Zhang, Y.; Liu, H.; Li, W.; Zhang, Z.; Shang, X.; Zhang, D.; Li, Y.; Zhang, S.; Liu, J.; Hess, R.A.; et al. Intraflagellar Transporter Protein (IFT27), an IFT25 Binding Partner, Is Essential for Male Fertility and Spermiogenesis in Mice. *Dev. Biol.* **2017**, *432*, 125–139. [[CrossRef](#)] [[PubMed](#)]
73. Fouchecourt, S.; Livera, G.; Messiaen, S.; Fumel, B.; Parent, A.S.; Marine, J.C.; Monget, P. Apoptosis of Sertoli Cells after Conditional Ablation of Murine Double Minute 2 (Mdm2) Gene Is P53-Dependent and Results in Male Sterility. *Cell Death Differ.* **2016**, *23*, 521–530. [[CrossRef](#)] [[PubMed](#)]
74. Tomar, A.K.; Rajak, S.K.; Aslam MK, M.; Chhikara, N.; Ojha, S.K.; Nayak, S.; Chhillar, S.; Kumaresan, A.; Yadav, S. Sub-Fertility in Crossbred Bulls: Identification of Proteomic Alterations in Spermatogenic Cells Using High Throughput Comparative Proteomics Approach. *Theriogenology* **2021**, *169*, 65–75. [[CrossRef](#)]
75. Zhao, S.; Chen, T.; Luo, X.; Chen, S.; Wang, J.; Lai, S.; Jia, X. Identification of Novel LncRNA and Differentially Expressed Genes (DEGs) of Testicular Tissues among Cattle, Yak, and Cattle-Yak Associated with Male Infertility. *Animals* **2021**, *11*, 2420. [[CrossRef](#)]
76. Zhao, W.; Mengal, K.; Yuan, M.; Quansah, E.; Li, P.; Wu, S.; Xu, C.; Yi, C.; Cai, X. Comparative RNA-Seq Analysis of Differentially Expressed Genes in the Epididymides of Yak and Cattleyak. *Curr. Genom.* **2019**, *20*, 293–305. [[CrossRef](#)]
77. Zhao, Y.; Gao, N.; Li, X.; El-Ashram, S.; Wang, Z.; Zhu, L.; Jiang, W.; Peng, X.; Zhang, C.; Chen, Y.; et al. Identifying Candidate Genes Associated with Sperm Morphology Abnormalities Using Weighted Single-Step GWAS in a Duroc Boar Population. *Theriogenology* **2020**, *141*, 9–15. [[CrossRef](#)]
78. Cui, Y.; Zhang, Y.; Wei, Z.; Gao, J.; Yu, T.; Chen, R.; Lv, X.; Pan, C. Pig KDM5B: MRNA Expression Profiles of Different Tissues and Testicular Cells and Association Analyses with Testicular Morphology Traits. *Gene* **2018**, *650*, 27–33. [[CrossRef](#)]
79. Häger, M.; Gawlik, K.; Nyström, A.; Sasaki, T.; Durbeek, M. Laminin  $\{\alpha\}$ 1 Chain Corrects Male Infertility Caused by Absence of Laminin  $\{\alpha\}$ 2 Chain. *Am. J. Pathol.* **2005**, *167*, 823–833. [[CrossRef](#)]
80. Sha, Y.; Xu, Y.; Wei, X.; Liu, W.; Mei, L.; Lin, S.; Ji, Z.; Wang, X.; Su, Z.; Qiu, P.; et al. CCDC9 Is Identified as a Novel Candidate Gene of Severe Asthenozoospermia. *Syst. Biol. Reprod. Med.* **2019**, *65*, 465–473. [[CrossRef](#)]
81. Pasek, R.C.; Malarkey, E.; Berbari, N.F.; Sharma, N.; Kesterson, R.A.; Tres, L.L.; Kierszenbaum, A.L.; Yoder, B.K. Coiled-Coil Domain Containing 42 (Ccdc42) Is Necessary for Proper Sperm Development and Male Fertility in the Mouse. *Dev. Biol.* **2016**, *412*, 208–218. [[CrossRef](#)] [[PubMed](#)]
82. Wang, T.; Yin, Q.; Ma, X.; Tong, M.H.; Zhou, Y. Ccdc87 Is Critical for Sperm Function and Male Fertility. *Biol. Reprod.* **2018**, *99*, 817–827. [[CrossRef](#)]
83. Cunha, S.R.; Mohler, P.J. Ankyrin Protein Networks in Membrane Formation and Stabilization. *J. Cell. Mol. Med.* **2009**, *13*, 4364–4376. [[CrossRef](#)] [[PubMed](#)]
84. Manfredola, F.; Martinez, G.; Coutton, C.; Rocco, D.; Reynaud, K.; Le Vern, Y.; Froment, P.; Beauclair, L.; Aubert, D.; Pierantoni, R.; et al. Ankrd31 in Sperm and Epididymal Integrity. *Front. Cell Dev. Biol.* **2021**, *9*, 741975. [[CrossRef](#)] [[PubMed](#)]

85. Liu, F.J.; Liu, X.X.; Liu, X.; Li, T.; Zhu, P.; Liu, Z.Y.; Xue, H.; Wang, W.J.; Yang, X.L.; Liu, J.; et al. Integrated Analyses of Phenotype and Quantitative Proteome of CMTM4 Deficient Mice Reveal Its Association with Male Fertility. *Mol. Cell. Proteom.* **2019**, *18*, 1070–1084. [[CrossRef](#)] [[PubMed](#)]
86. Luconi, M.; Cantini, G.; Baldi, E.; Forti, G. Role of A-Kinase Anchoring Proteins (AKAPs) in Reproduction. *Front. Biosci.* **2011**, *16*, 1315–1330. [[CrossRef](#)]
87. Audet-Walsh, É.; Bellemare, J.; Lacombe, L.; Fradet, Y.; Fradet, V.; Douville, P.; Guillemette, C.; Lévesque, É. The Impact of Germline Genetic Variations in Hydroxysteroid (17- $\beta$ ) Dehydrogenases on Prostate Cancer Outcomes After Prostatectomy. *Eur. Urol.* **2012**, *62*, 88–96. [[CrossRef](#)]
88. LaVoie, H.A. The Role of GATA in Mammalian Reproduction. *Exp. Biol. Med.* **2003**, *228*, 1282–1290. [[CrossRef](#)]
89. Emami, N.; Diamandis, E.P. Potential Role of Multiple Members of the Kallikrein-Related Peptidase Family of Serine Proteases in Activating Latent TGF  $\beta$  1 in Semen. *Biol. Chem.* **2010**, *391*, 85–95. [[CrossRef](#)]
90. Emami, N.; Scorilas, A.; Soosaipillai, A.; Earle, T.; Mullen, B.; Diamandis, E.P. Association between Kallikrein-Related Peptidases (KLKs) and Macroscopic Indicators of Semen Analysis: Their Relation to Sperm Motility. *Biol. Chem.* **2009**, *390*, 921–929. [[CrossRef](#)]
91. Fardilha, M.; Esteves, S.L.C.; Korrodi-Gregório, L.; Vintém, A.P.; Domingues, S.C.; Rebelo, S.; Morrice, N.; Cohen, P.T.W.; Da Cruz E Silva, O.A.B.; Da Cruz E Silva, E.F. Identification of the Human Testis Protein Phosphatase 1 Interactome. *Biochem. Pharmacol.* **2011**, *82*, 1403–1415. [[CrossRef](#)] [[PubMed](#)]
92. Zhoucun, A.; Zhang, S.; Yang, Y.; Ma, Y.; Zhang, W.; Lin, L. The Common Variant N372H in BRCA2 Gene May Be Associated with Idiopathic Male Infertility with Azoospermia or Severe Oligozoospermia. *Eur. J. Obstet. Gynecol. Reprod. Biol.* **2006**, *124*, 61–64. [[CrossRef](#)] [[PubMed](#)]
93. Ji, G.; Yan, L.; Liu, W.; Huang, C.; Gu, A.; Wang, X. Polymorphisms in Double-Strand Breaks Repair Genes Are Associated with Impaired Fertility in Chinese Population. *Reproduction* **2013**, *145*, 463–470. [[CrossRef](#)] [[PubMed](#)]
94. Ghasemi, H.; Khodadadi, I.; Fattahi, A.; Moghimbeigi, A.; Tavilani, H. Polymorphisms of DNA Repair Genes XRCC1 and LIG4 and Idiopathic Male Infertility. *Syst. Biol. Reprod. Med.* **2017**, *63*, 382–390. [[CrossRef](#)] [[PubMed](#)]
95. Brandsma, I.; Sato, K.; van Rossum-Fikkert, S.E.; van Vliet, N.; Sleddens, E.; Reuter, M.; Odijk, H.; van den Tempel, N.; Dekkers, D.H.W.; Bezstarosti, K.; et al. HSF2BP Interacts with a Conserved Domain of BRCA2 and Is Required for Mouse Spermatogenesis. *Cell Rep.* **2019**, *27*, 3790–3798.e7. [[CrossRef](#)] [[PubMed](#)]
96. Tu, C.; Cong, J.; Zhang, Q.; He, X.; Zheng, R.; Yang, X.; Gao, Y.; Wu, H.; Lv, M.; Gu, Y.; et al. Bi-Allelic Mutations of DNAH10 Cause Primary Male Infertility with Asthenoteratozoospermia in Humans and Mice. *Am. J. Hum. Genet.* **2021**, *108*, 1466–1477. [[CrossRef](#)]
97. Gao, Y.; Tian, S.; Sha, Y.; Zha, X.; Cheng, H.; Wang, A.; Liu, C.; Lv, M.; Ni, X.; Li, Q.; et al. Novel Bi-Allelic Variants in DNAH2 Cause Severe Asthenoteratozoospermia with Multiple Morphological Abnormalities of the Flagella. *Reprod. Biomed. Online* **2021**, *42*, 963–972. [[CrossRef](#)] [[PubMed](#)]
98. Zhu, D.; Zhang, H.; Wang, R.; Liu, X.; Jiang, Y.; Feng, T.; Liu, R.; Zhang, G. Association of DNAH11 Gene Polymorphisms with Asthenozoospermia in Northeast Chinese Patients. *Biosci. Rep.* **2019**, *39*, 20181450. [[CrossRef](#)]
99. Lu, S.; Gu, Y.; Wu, Y.; Yang, S.; Li, C.; Meng, L.; Yuan, W.; Jiang, T.; Zhang, X.; Li, Y.; et al. Bi-Allelic Variants in Human WDR63 Cause Male Infertility via Abnormal Inner Dynein Arms Assembly. *Cell Discov.* **2021**, *7*, 110. [[CrossRef](#)]
100. King, S.M. Axonemal Dynein Arms. *Cold Spring Harb. Perspect. Biol.* **2016**, *8*, a028100. [[CrossRef](#)]
101. Nielsen, M.G.; Turner, F.R.; Hutchens, J.A.; Raff, E.C. Axoneme-Specific  $\beta$ -Tubulin Specialization: A Conserved C-terminal Motif Specifies the Central Pair. *Curr. Biol.* **2001**, *11*, 529–533. [[CrossRef](#)]
102. Kollmar, M. Fine-Tuning Motile Cilia and Flagella: Evolution of the Dynein Motor Proteins from Plants to Humans at High Resolution. *Mol. Biol. Evol.* **2016**, *33*, 3249–3267. [[CrossRef](#)] [[PubMed](#)]
103. Sironen, A.; Shoemark, A.; Patel, M.; Loebinger, M.R.; Mitchison, H.M. Sperm Defects in Primary Ciliary Dyskinesia and Related Causes of Male Infertility. *Cell. Mol. Life Sci.* **2020**, *77*, 2029. [[CrossRef](#)]
104. Pazour, G.J.; Agrin, N.; Walker, B.L.; Witman, G.B. Identification of Predicted Human Outer Dynein Arm Genes: Candidates for Primary Ciliary Dyskinesia Genes. *J. Med. Genet.* **2006**, *43*, 62–73. [[CrossRef](#)] [[PubMed](#)]
105. 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. [[CrossRef](#)] [[PubMed](#)]
106. Fassad, M.R.; Shoemark, A.; Legendre, M.; Hirst, R.A.; Koll, F.; le Borgne, P.; Louis, B.; Daudvohra, F.; Patel, M.P.; Thomas, L.; et al. Mutations in Outer Dynein Arm Heavy Chain DNAH9 Cause Motile Cilia Defects and Situs Inversus. *Am. J. Hum. Genet.* **2018**, *103*, 984–994. [[CrossRef](#)]
107. Sha, Y.; Yang, X.; Mei, L.; Ji, Z.; Wang, X.; Ding, L.; Li, P.; Yang, S. DNAH1 Gene Mutations and Their Potential Association with Dysplasia of the Sperm Fibrous Sheath and Infertility in the Han Chinese Population. *Fertil. Steril.* **2017**, *107*, 1312–1318.e2. [[CrossRef](#)]
108. Wiedemann, I.; Maehlmeyer, A.; Jansen, S.; Sharifi, A.R.; Knorr, C. SNP g.1007A>G within the Porcine DNAL4 Gene Affects Sperm Motility Traits and Percentage of Midpiece Abnormalities. *Reprod. Domest. Anim.* **2018**, *53*, 401–413. [[CrossRef](#)]
109. Shi, L.; Zhou, T.; Huang, Q.; Zhang, S.; Li, W.; Zhang, L.; Hess, R.A.; Pazour, G.J.; Zhang, Z. Intraflagellar Transport Protein 74 Is Essential for Spermatogenesis and Male Fertility in Mice. *Biol. Reprod.* **2019**, *101*, 188–199. [[CrossRef](#)]

110. Wambergue, C.; Zouari, R.; Fourati Ben Mustapha, S.; Martinez, G.; Devillard, F.; Hennebicq, S.; Satre, V.; Brouillet, S.; Halouani, L.; Marrakchi, O.; et al. Patients with Multiple Morphological Abnormalities of the Sperm Flagella Due to DNAH1 Mutations Have a Good Prognosis Following Intracytoplasmic Sperm Injection. *Hum. Reprod.* **2016**, *31*, 1164–1172. [[CrossRef](#)]
111. Siu, M.K.Y.; Yan Cheng, C. Extracellular Matrix and Its Role in Spermatogenesis. *Adv. Exp. Med. Biol.* **2008**, *636*, 74. [[CrossRef](#)] [[PubMed](#)]
112. Siu, M.K.Y.; Cheng, C.Y. Dynamic Cross-Talk between Cells and the Extracellular Matrix in the Testis. *BioEssays* **2004**, *26*, 978–992. [[CrossRef](#)] [[PubMed](#)]
113. Lui, W.Y.; Lee, W.M. Regulation of Junction Dynamics in the Testis—Transcriptional and Post-Translational Regulations of Cell Junction Proteins. *Mol. Cell. Endocrinol.* **2006**, *250*, 25–35. [[CrossRef](#)]
114. Lehmann, D.; Temminck, B.; Da Rugna, D.; Leibundgut, B.; Sulmoni, A.; Müller, H. Role of Immunological Factors in Male Infertility. Immunohistochemical and Serological Evidence. *Lab. Invest.* **1987**, *57*, 21–28.
115. Sala-Newby, G.B.; Newby, A.C. Cloning of a Mouse Cytosolic 5'-Nucleotidase-I Identifies a New Gene Related to Human Autoimmune Infertility-Related Protein. *Biochim. Biophys. Acta* **2001**, *1521*, 12–18. [[CrossRef](#)]
116. Hodge, S.H.; Watts, A.; Marley, R.; Baines, R.A.; Hafen, E.; MacDougall, L.K. Twitchy, the Drosophila Orthologue of the Ciliary Gating Protein FBF1/Dyf-19, Is Required for Coordinated Locomotion and Male Fertility. *Biol. Open* **2021**, *10*, bio058531. [[CrossRef](#)]
117. Soini, S.; Ibarreta, D.; Anastasiadou, V.; Aymé, S.; Braga, S.; Cornel, M.; Coviello, D.A.; Evers-Kiebooms, G.; Geraedts, J.; Gianaroli, L.; et al. The Interface between Assisted Reproductive Technologies and Genetics: Technical, Social, Ethical and Legal Issues. *Eur. J. Hum. Genet.* **2006**, *14*, 588–645. [[CrossRef](#)] [[PubMed](#)]