DNA

MDPI

Article

Genetic Insights into Teratozoospermia: A Comprehensive Computational Study of UTR Variants in *AURKC*, *SPATA16*, and *SUN5*

Maria-Anna Kyrgiafini and Zissis Mamuris *

Laboratory of Genetics, Comparative and Evolutionary Biology, Department of Biochemistry and Biotechnology, University of Thessaly, Viopolis, Mezourlo, 41500 Larissa, Greece
* Correspondence: zmamur@uth.gr

Abstract: Teratozoospermia, a complex male fertility disorder affecting sperm morphology, has been linked to *AURKC*, *SPATA16*, and *SUN5* gene defects. However, the sheer volume of SNPs in these genes necessitates prioritization for comprehensive analysis. This study focuses on the often-overlooked untranslated region (UTR) variants in these genes, aiming to assess their association with teratozoospermia and prioritize them. We employed a multi-step filtering process, including functional significance assessment (RegulomeDB, 3DSNP v2.0, SNPinfo (FuncPred)), evaluation of gene expression impacts in testis tissue using GTEx, and assessment of miRNA binding site effects (PolymiRTS Database 3.0, miRNASNP v3). Additionally, we used SNPnexus to evaluate their conservation and association with diseases. In *AURKC*, we identified six UTR SNPs (rs11084490, rs58264281, rs35582299, rs533889458, rs2361127, rs55710619), two of which influenced gene expression in testis, while others affected the binding sites of 29 miRNAs or were located in transcription-factor binding sites. Three of these SNPs were also found to be associated with spermatogenic failure according to previous studies indicating a potential regulatory role in teratozoospermia, too. For *SPATA16*, two 3′ UTR variants, rs146640459 and rs148085657, were prioritized, with the latter impacting miRNA binding sites. In *SUN5*, three 3′ UTR variants (rs1485087675, rs762026146, rs1478197315) affected miRNA binding sites. It should be noted that none of the above variants was identified in a conserved region. Our findings shed light on the potential regulatory roles of these SNPs in teratozoospermia and lay the foundation for future research directions in this area.

Keywords: teratozoospermia; male infertility; variant; SNP; in silico; miRNA; untranslated region (UTR)

Citation: Kyrgiafini, M.-A.; Mamuris, Z. Genetic Insights into Teratozoospermia: A Comprehensive Computational Study of UTR Variants in *AURKC*, *SPATA16*, and *SUN5*. *DNA* **2023**, *3*, 148–167. https://doi.org/10.3390/dna3040013

Academic Editor: Ashis Basu

Received: 14 September 2023
Revised: 16 October 2023
Accepted: 18 October 2023
Published: 26 October 2023

1. Introduction

Infertility, defined by the World Health Organization (WHO) as "the failure to achieve a pregnancy after 12 months or more of regular unprotected sexual intercourse" remains a significant concern for couples of reproductive age. Male-related factors contribute to approximately 50% of infertility cases [1,2], underscoring the essential need for a precise evaluation of semen quality to gauge male fertilization potential [3]. While sperm count and motility stand as primary parameters assessed in semen quality analysis, the structure of spermatozoa plays an equally vital and intricate role in determining the fertilization potential of male reproductive cells [4]. Specifically, teratozoospermia is characterized by a lower percentage of normally shaped sperm compared to established reference limits. The definition of "normal" has evolved significantly over time, transitioning from 50% in 1980 [5] to 4% in the WHO classification published in 2010 [6]. This condition encompasses a spectrum of morphological deviations impacting diverse components of sperm structure, including the head, neck, midpiece, and tail [7,8]. Beyond presenting a wide array of sperm irregularities, teratozoospermia also exists across varying degrees of severity that directly influence male fertilization capacity [8]. In general, the morphological features of sperm

cells result from highly intricate cellular transformations that occur during spermatogenesis [9] and intriguingly, aberrant sperm morphology has been linked to increased indicators of sperm damage, such as DNA fragmentation [2] and overproduction of reactive oxygen species (ROS) [10,11].

Despite notable advancements in exploring teratozoospermia, a comprehensive understanding of the molecular mechanisms responsible for this male infertility condition remains elusive. In broad terms, unraveling the molecular origins of male infertility presents a substantial hurdle as more than 4000 genes are involved in the spermatogenesis process [12]. However, recently, various genes have been associated with teratozoospermia.

The Aurora Kinase C gene (*AURKC*), located on chromosome 19q13.43, encodes a member of a family of highly conserved serine/threonine kinases that are crucial for chromosome segregation during both mitosis and meiosis [13]. The two other family members, *AURKA* and *AURKCB*, are highly expressed in many cancer types and act as oncogenes [14]. Limited information is available regarding the involvement of *AURKC* in oncogenesis but it is also found to be overexpressed in certain cancers [15]. Concerning male infertility, *AURKC* is expressed in meiotic cells, and pathogenic mutations in this gene can disrupt the protein's function, leading to improper mitotic spindle formation and subsequently causing male infertility [16]. To this day, various mutations of *AURKC* have been discovered, affecting protein function and resulting in a specific form of teratozoospermia known as macrozoospermia or large-headed spermatozoa [16–21].

SPATA16 (spermatogenesis-associated protein 16), located on chromosome 3q26.31 and formerly recognized as *NYD-SP12*, exhibits high expression in the human testes, particularly during puberty, where it plays a significant role in its development [22]. It features a conserved tetratricopeptide repeat (TPR) domain, known for its role in facilitating protein–protein interactions [23]. This protein localizes within the Golgi apparatus and proacrosomal vesicles, which merge during spermiogenesis to form the acrosome [24]. Several studies have identified mutations within *SPATA16* that lead to a distinct form of teratozoospermia known as globozoospermia [25,26]. Globozoospermia is characterized by the presence of round-headed spermatozoa that lack an acrosome. Furthermore, research conducted in mice underscores the significance of *SPATA16* in the process of sperm formation [27]. Mutations in *SPATA16* have also been linked to other types of male infertility [28].

SUN5 is a gene located on chromosome 20q11.21, encoding for a transmembrane protein consisting of an N-terminal nucleoplasmic section, a coiled-coil region, a transmembrane helical domain, and a SUN domain segment [29]. It is a testis-specific gene [30] and its encoded protein localizes to the junction between the sperm head and tail [29,31]. SUN5 belongs to the family of SUN domain proteins, which play a role in tethering the centrosome to the nuclear membrane [32]. SUN5 is a relatively recent addition to the SUN family, and while limited information is available regarding its function [31], it is suggested that it may be involved in nuclear envelope reconstitution and nuclear migration [29]. Notably, studies conducted in mice have shown that $Sun5^{-/-}$ mice are infertile, and in the absence of functional SUN5, the sperm head-to-tail coupling apparatus becomes detached from the nucleus during spermatid elongation [31]. Additionally, several mutations in this gene have been identified, which are associated with acephalic spermatozoa syndrome, a severe form of teratozoospermia [29,33–35].

The aforementioned studies confirm the pivotal role of the mentioned genes in the pathogenesis of teratozoospermia and underscore the significance of specific mutations in its etiology. Nevertheless, the number of mutations associated with male infertility remains relatively limited. Single nucleotide polymorphisms (SNPs) are the most prevalent type of genetic mutation, occurring in the genome approximately every 100 to 300 base pairs [36]. While mutations in coding regions are typically linked to the development of various diseases due to alterations in the amino acid sequence, research suggests that SNPs located in non-coding regions are more likely to contribute to the pathogenesis of most genetic disorders [37]. More specifically, variants found in non-coding regions can

exert various regulatory functions within the genome, including disruption of interactions with transcription factors (TFs), microRNAs (miRNAs), and the creation or disruption of splice sites, etc. [38]. Consequently, variants in non-coding regions may impact protein function by reducing protein solubility or destabilizing protein structure [39]. Notably, SNPs in the 3′ untranslated regions (3′ UTR) are of particular significance, as they serve as primary binding sites for miRNAs. miRNAs play a crucial role in gene expression regulation and their interactions with the 3′ UTR lead to gene silencing after transcription and translation suppression [40]. Moreover, in the context of male infertility, several studies have demonstrated differential expression of miRNAs between fertile and infertile males. These miRNAs hold the potential to unveil the molecular mechanisms underlying infertility and may serve as noninvasive biomarkers for diagnosing this condition [41]. Likewise, SNPs within the 5′ untranslated region (5′ UTR) hold significant importance and can contribute to the development of various diseases. Specifically, 5′ UTRs play a pivotal role in influencing both mRNA stability and translation efficiency [42,43]. Additionally, functional elements such as the internal ribosome entry site (IRES), upstream open reading frames (uORFs), and iron-responsive element (IRE) within the 5′ UTR play a crucial role in precisely modulating protein expression in alignment with the specific requirements of the cell [42]. Consequently, SNPs have the potential to disrupt the smooth translation of mRNA or compromise the stability of the mRNA molecule, making it more susceptible to degradation. Disruption of the aforementioned functional elements can also lead to alterations in gene expression. Ultimately, the irregular gene expression resulting from mutations in the 5′ UTR can significantly contribute to the progression and manifestation of a spectrum of diseases.

Therefore, in the present day, while coding region SNPs have garnered significant attention in candidate gene studies due to their critical regulatory roles, there has been notably less emphasis on the functional analysis of non-coding SNPs [44]. The continuous evolution of SNP discovery technologies and the dynamic annotation of the genome have resulted in the accumulation of an overwhelming amount of information and a large number of SNPs that are challenging to study experimentally [45]. Consequently, computational methods are becoming increasingly indispensable in genomic research for SNP selection and the prediction of their functional consequences in disease development [46].

Today, bioinformatics tools play a crucial role in prioritizing SNPs with functional significance from the vast pool of neutral non-risk variants [47]. These tools assess the potential functional impacts of SNPs across five key levels: splicing, transcription, translation, post-translation, and protein stability. While most existing bioinformatics tools focus on evaluating SNP effects with respect to a single biological function, others offer a comprehensive analysis of SNP function by integrating various algorithms, data sources, etc. [44–46].

The objective of the present study was to analyze UTR variants in the *AURKC*, *SPATA16*, and *SUN5* genes using computational methods, given the significance of UTR variants in numerous studies and their association with various diseases. These genes are well known for their role in teratozoospermia. Therefore, the SNPs identified in their UTR were prioritized based on several criteria, including their functional significance, association with expression quantitative trait loci (eQTL) and diseases, presence within evolutionarily conserved regions, and their impact on the creation or disruption of miRNA binding sites. As a result, this study involves a rigorous process of filtering through a list of SNPs to identify SNPs that are most likely to be associated with teratozoospermia. To the best of our knowledge, this is the first comprehensive computational analysis of UTR SNPs in the *AURKC*, *SPATA16*, and *SUN5* genes. It provides a valuable foundation for future research, listing candidate variants that may be linked to teratozoospermia, thereby contributing to a deeper understanding of the molecular mechanisms underlying male infertility. Furthermore, this research can facilitate the development of biomarkers to enhance assisted reproductive technology (ART) and improve the diagnosis and prognosis of male infertility, especially teratozoospermia.

2. Materials and Methods

2.1. Retrieval of UTR SNPs in AURKC, SPATA16, and SUN5

The Ensembl genome browser (GRCh38/hg38) [48] was employed to retrieve SNPs located in the 3′ and 5′ UTRs of the *AURKC*, *SPATA16*, and *SUN5* genes. Duplicate variants and somatic mutations associated with cancer, according to the COSMIC database [49], were excluded, resulting in a final list of SNPs. These SNPs then underwent a thorough multi-step filtering process using various databases and in silico tools, as detailed in the subsequent sections and summarized in Figure 1.

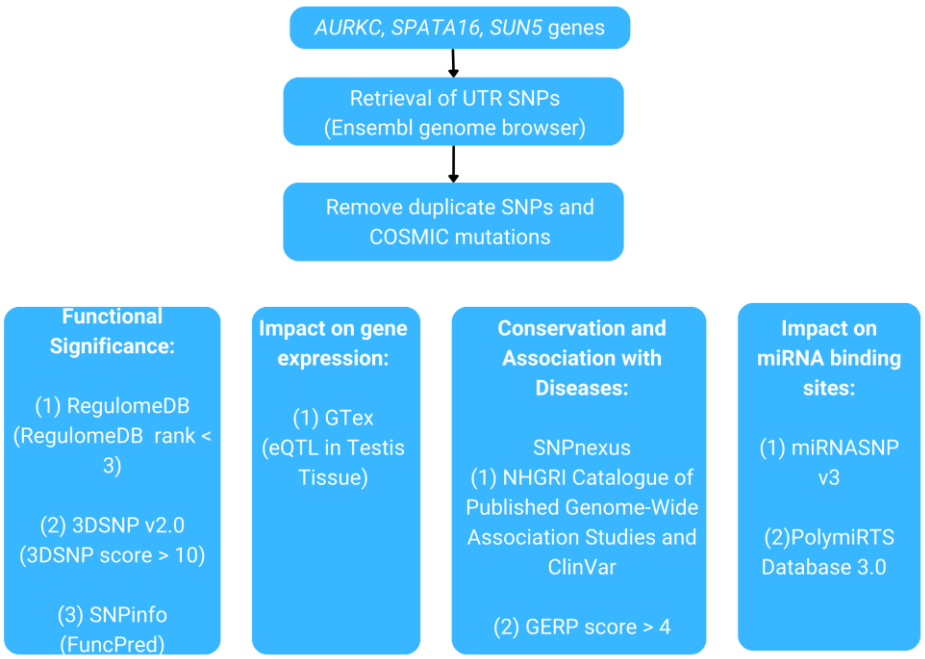


Figure 1. Workflow of the methods and tools involved in this study for UTR variant prioritization in *AURKC*, *SPATA16*, and *SUN5*.

2.2. Evaluating the Functional Significance of UTR SNPs

To gauge the functional significance of the UTR SNPs identified in *AURKC*, *SPATA16*, and *SUN5*, we combined data from three distinct databases.

Initially, we conducted functional annotation and scoring of UTR variants using RegulomeDB (https://www.regulomedb.org/regulome-search/, accessed on 13 September 2023) [50]. RegulomeDB categorizes SNPs based on the presence or absence of functional elements, such as protein binding, motifs, chromatin structure, histone modifications, and more [50]. The function of RegulomeDB is to assign scores and ranks to SNPs so that the functional SNPs can be differentiated from a broad pool. Specifically, each SNP is assigned a rank ranging from 1 to 7, with lower values indicating a higher likelihood of having a regulatory function.

We also employed 3DSNP v2.0 [51] and SNPinfo (FuncPred) [52] to assess SNP functionality. The 3DSNP v2.0 database (https://omic.tech/3dsnpv2/, accessed on 13 September 2023) is a comprehensive resource that provides information on 3D-interacting genes, enhancer and promoter states, transcription-factor binding sites, modified sequence motifs, and conservation data. It calculates a functional score for each SNP using these factors, with higher scores indicating a greater likelihood of SNP functionality [51].

SNPinfo (FuncPred) (https://snpinfo.niehs.nih.gov/snpinfo/snpfunc.html, accessed on 13 September 2023) is a web server designed to aid in the selection of SNPs for genetic association studies. It serves as a comprehensive tool for predicting SNP functionality, assessing whether SNP variants can influence transcriptional regulation by affecting the activity of transcription-factor binding sites (TFBS), or by altering splicing patterns or

efficiency through the disruption of splice sites, exonic splicing enhancers (ESE), or silencers (ESS) [52].

We prioritized SNPs with a RegulomeDB rank ranging from 1a to 3b, SNPs with a 3DSNP score greater than 10, and SNPs identified as functionally significant according to FuncPred. To obtain this information, we inputted the rs IDs of individual SNPs into the aforementioned databases.

It should be noted that the 3DSNP score is computed as the sum of scores derived from six distinct functional categories, including 3D interacting genes, enhancer state, promoter state, transcription-factor binding sites, sequence motifs altered, and conservation score. In contrast to RegulomeDB, 3DSNP employs a quantitative scoring system. The thresholds adopted in this study were based on prior related studies that combined data from RegulomeDB and 3DSNP to identify SNPs with an increased likelihood of regulatory functionality [53–56].

2.3. Impact of UTR SNPs on Gene Expression

To prioritize UTR variants that influence gene expression of *AURKC*, *SPATA16*, and *SUN5*, we leveraged the Genotype-Tissue Expression (GTEx) database [57] (https://gtexportal.org/home/index.html, accessed on 13 September 2023). The GTEx Program is a comprehensive resource that provides insights into the relationship between genetic variants and gene expression across multiple human tissues, allowing the identification of expression quantitative trait loci (eQTLs) and more [57]. For each UTR SNP on our list, we assessed its effect on gene expression by submitting the individual variants' rs IDs to the GTEx database. For filtering, we considered significant only those variants that affected gene expression in testis tissue. This focus on testicular tissue aligns with our primary objective of investigating the role of these variants in male infertility.

2.4. Association of SNPs with Diseases and Identification of SNPs in Evolutionarily Conserved Regions

To prioritize variants previously linked to male infertility or other diseases and SNPs situated within evolutionarily conserved regions, we utilized SNPnexus [58]. SNPnexus (https://www.snp-nexus.org/v4/, accessed on 13 September 2023) is a web-based variant annotation tool designed to simplify the selection and prioritization of genetic variants. It integrates data from various sources, providing information about population genetics, regulatory elements, consequences on protein, conservation, etc. [58].

To gather information regarding the association of SNPs with diseases, SNPnexus [58] integrates data from the NHGRI Catalogue of Published Genome-Wide Association Studies (https://www.ebi.ac.uk/gwas/, accessed on 13 September 2023) [59] and ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/, accessed on 13 September 2023) [60]. The former provides a curated collection of all published genome-wide association studies [59], while the latter is a publicly accessible archive containing reports on the relationships between human variations and phenotypes associated with human health [60]. Regarding conservation, SNPnexus also incorporates data from the Genomic Evolutionary Rate Profiling package (GERP++) [61]. The GERP score measures the reduction in the number of substitutions in the multi-species sequence alignment compared to the neutral expectation and practically, it is used to identify sites where mutations may have a significant impact on fitness [61,62]. Conservation scores have been widely employed in medical and population genetic studies, with variants typically considered deleterious if they possess a GERP score >4 [62].

By submitting the rs ID of each variant to SNPnexus [58], we generated a comprehensive list of UTR variants associated with diseases. Likewise, we compiled a list of UTR SNPs for each gene (*AURKC, SPATA16, SUN5*) with a GERP score exceeding 4.

2.5. Assessment of the Impact of UTR SNPs on miRNA Binding Sites

To evaluate the impact of UTR SNPs on miRNA binding sites, we employed two distinct tools: PolymiRTS Database 3.0 [63], and miRNASNP v3 [64].

PolymiRTS Database 3.0 (https://compbio.uthsc.edu/miRSNP/, accessed on 13 September 2023) is a comprehensive database that systematically identifies DNA polymorphisms within both miRNAs and miRNA target sites. This resource integrates various data sources, including hybrid sequences, crosslink experiments, and miRNA interaction ligation, to accurately pinpoint the location of non-coding SNPs [63]. Similarly, miRNASNP v3 (http://bioinfo.life.hust.edu.cn/miRNASNP/, accessed on 13 September 2023) is a database designed to predict the functional impact of SNPs located within mature miRNA, miRNA target sequences, pre-miRNA, and adjacent regions [64].

We used these tools to prioritize the UTR SNPs and identify those that either create or disrupt miRNA binding sites, thereby affecting miRNA binding affinity with target genes (*AURKC*, *SPATA16*, *SUN5*). Such changes indicate an additional regulatory role of SNPs, as miRNAs regulate gene expression by binding to messenger RNAs (mRNAs) and either preventing them from being translated into proteins or marking them for degradation. This binding significantly influences gene regulation, impacting various cellular processes. Consequently, identifying the affected miRNAs could provide valuable insights into the pathogenesis of teratozoospermia.

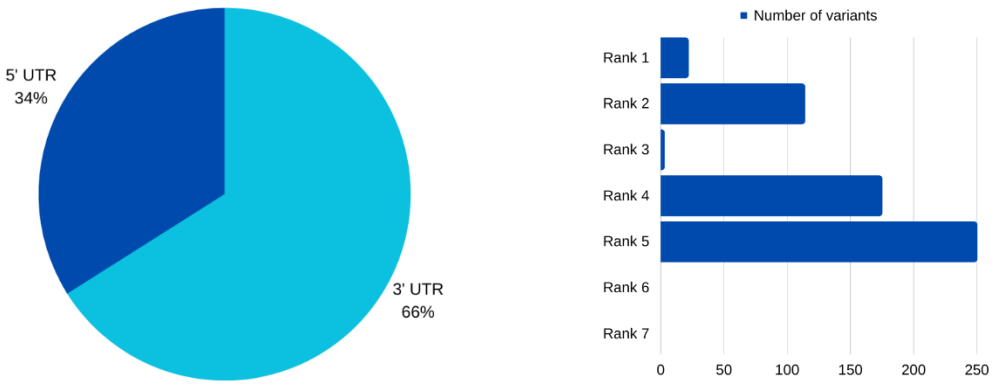
Thus, in this stage of the analysis, we submitted the rs ID of each variant to both databases, resulting in a list of all 3′ UTR SNPs in the *AURKC*, *SPATA16*, and *SUN5* genes that either created or disrupted miRNA binding sites. Additionally, we generated a list of all the affected miRNAs as output.

At this point, we want to emphasize that our selection of the above databases and tools was based on a comprehensive literature review. In particular, we curated data from various publications that employed variant prioritization methodologies. The workflow used in this study was developed following a careful selection process, focusing on databases recognized for their high quality and reliability, as indicated in previous similar publications [37,44,47,53,65]. Moreover, to ensure the reliability of our results, we sourced data from multiple databases for several parameters used in variant prioritization, such as the potential functionality of variants (FuncPred, RegulomeDB, etc.).

3. Results

3.1. Evaluation of UTR Variants in AURKC

Initially, we collected 1627 UTR variants in *AURKC* from the Ensembl genome browser [48]. Subsequently, we conducted the removal of duplicate variants and excluded somatic mutations associated with cancer according to COSMIC [49]. This process left us with 592 variants for detailed analysis. Out of this selection, 201 variants were located in the 5′ UTR, while 391 were identified in the 3′ UTR (as illustrated in Figure 2a and detailed in Table S1).



(**a**) (**b**)

Figure 2. (**a**) Total count of 3′ and 5′ UTR retrieved from the Ensembl genome browser [48] for *AURKC*. (**b**) Annotation and ranks of UTR SNPs according to RegulomeDB [50].

For the comprehensive evaluation of SNP functional significance, we employed three distinct databases: RegulomeDB [50], 3DSNP v2.0 [51], and SNPinfo (FuncPred) [52]. Among the entirety of the analyzed SNPs, a total of 564 were detected in RegulomeDB [50], and their distribution is visually represented in Figure 2b and outlined in Table S2. Of them, 139 were assigned RegulomeDB ranks ranging from 1a to 3b, elucidating their potential functional implications in gene regulation. Detailed results are provided in Table S2. For the 3DSNP v2.0 [51], only one of the UTR variants in *AURKC* obtained a 3DSNP score surpassing 10, as depicted in Table 1. Comprehensive results for all SNPs can be found in Table S3.

Table 1. List of SNPs predicted to have a functional effect according to 3DSNP v2.0 [51] for *AURKC*.

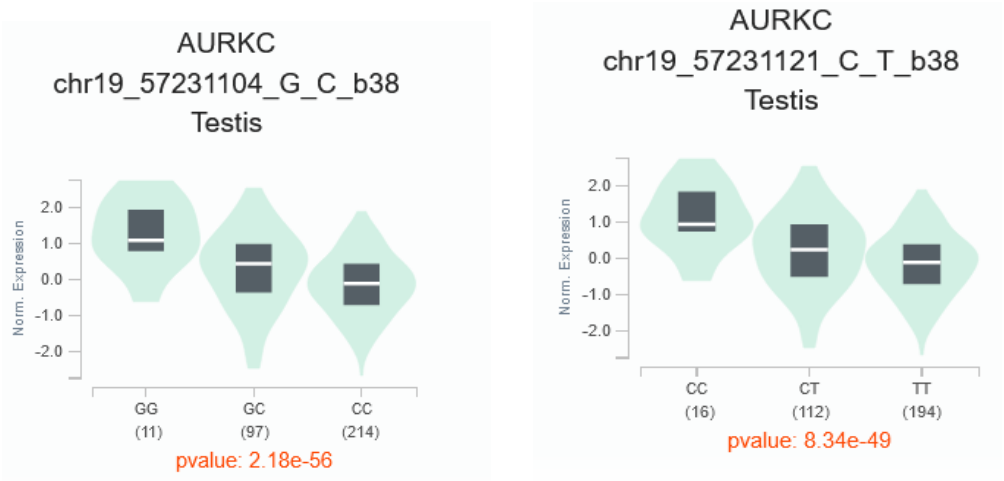
Variant ID	Chromosome Location	Score
rs533889458	19:57231016	10.02

Moreover, our analysis utilizing SNPinfo (FuncPred) [52] identified 12 SNPs with putative functional effects, affecting transcription-factor binding sites, splicing sites, and miRNA binding sites, as shown in Table 2.

Table 2. List of SNPs predicted to have a functional effect according to SNPinfo (FuncPred) [52] for *AURKC*; TFBS: transcription-factor binding site; ESE: exonic splicing enhancer; ESS: exonic splicing silencer; ✓ SNP that affects this site; -: SNP that does not affect this site.

Variant ID	Position	TFBS	Splicing (Site)	Splicing (ESE or ESS)	miRNA (miRanda)	miRNA (Sanger)
rs11084490	19:57231104	✓	-	-	-	-
rs2361127	19:57231699	✓	-	-	-	-
rs35582299	19:57235455	-	-	-	✓	✓
rs45503793	19:57232679	-	-	✓	-	-
rs45527835	19:57235250	-	-	✓	-	-
rs45555141	19:57232559	-	-	✓	-	-
rs55658999	19:57235043	-	-	✓	-	-
rs55710619	19:57232641	-	-	✓	-	-
rs55898757	19:57235249	-	-	✓	-	-
rs58264281	19:57231121	✓	✓	✓	-	-
rs61736320	19:57232163	-	-	✓	-	-
rs758098	19:57231671	✓	-	-	-	-

We additionally conducted an analysis of all UTR SNPs in *AURKC* using the GTEx database [57] to ascertain their presence in expression quantitative trait loci (eQTLs). Among the 592 variants, we identified only 16 SNPs in the GTEx database [57]. Notably, we observed a significant association between *AURKC* expression in testis and two specific SNPs, rs11084490 and rs58264281, with *p*-values of 2.18e-56 and 8.3e-49, respectively (Figure 3).



(**a**) (**b**)

Figure 3. Violin plots of (**a**) rs11084490 and (**b**) rs58264281 for single tissue eQTLs analysis through GTEx [57].

Furthermore, upon prioritizing the UTR SNPs in *AURKC* using SNPnexus [58], we discovered that none of these SNPs had previously been associated with other diseases according to the NHGRI Catalogue of Published Genome-Wide Association Studies [59]. Nevertheless, ClinVar [60] revealed that seven SNPs have previously been reported in association with fertility-related phenotypes, as detailed in Table 3. Additionally, none of the SNPs exhibited a GERP score exceeding four, as indicated in Table S4.

Table 3. SNPs in *AURKC* associated with pathogenic phenotypes according to ClinVar [60].

Variant ID	Chromosome Location	Phenotype
rs11084490	19:57231104	Spermatogenic Failure
rs121908654	19:57234985	Infertility associated with multi-tailed spermatozoa and excessive DNA
rs148940837	19:57234920	Spermatogenic Failure
rs55710619	19:57232641	Spermatogenic Failure
rs58264281	19:57231121	Spermatogenic Failure
rs886054645	19:57232112	Spermatogenic Failure
rs886054646	19:57232596	Spermatogenic Failure

In our final analysis to gauge the impact of UTR SNPs retrieved from Ensembl on miRNAs, we utilized PolymiRTS Database 3.0 [63] and miRNASNP v3 [64]. According to PolymiRTS Database 3.0 [63], one variant in *AURKC* disrupted miRNA binding sites potentially having an impact on gene regulation and pathogenesis of teratozoospermia. PolymiRTS categorizes SNPs into four classes: "D" for disruption of a conserved miRNA site, "N" for non-conserved miRNA disruption, "C" for the formation of a new miRNA site, and "O" for instances where ancestral alleles could not be determined. As detailed in Table 4, rs35582299 influenced the binding of 14 miRNAs. The G allele disrupted conserved miRNA binding sites (Function D), while the A allele created new miRNA binding sites. Furthermore, as per the findings from miRNASNP v3 [64], we identified 21 SNPs that played a role in the gain or loss of miRNA binding sites. Among these, 2 SNPs caused gain, while the remaining 19 SNPs contributed to alterations in miRNA function through both gain and loss of target sites, as detailed in Table S5.

Table 4. SNPs in *AURKC* disrupting miRNA binding sites, the affected miRNAs, and the function class of SNPs according to PolymiRTS Database 3.0 [63].

Variant ID	Chromosome Location	Allele	Affected miRNAs	Function Class
rs35582299	19:57235455	G	hsa-miR-125a-5p	D
			hsa-miR-125b-5p	D
			hsa-miR-345-3p	D
			hsa-miR-4319	D
			hsa-miR-4732-3p	D
			hsa-miR-670-5p	D
			hsa-miR-7106-3p	D
			hsa-miR-7113-3p	D
		A	hsa-miR-1200	C
			hsa-miR-3140-5p	C
			hsa-miR-378a-5p	C
			hsa-miR-516a-3p	C
			hsa-miR-516b-3p	C
			hsa-miR-7162-5p	C

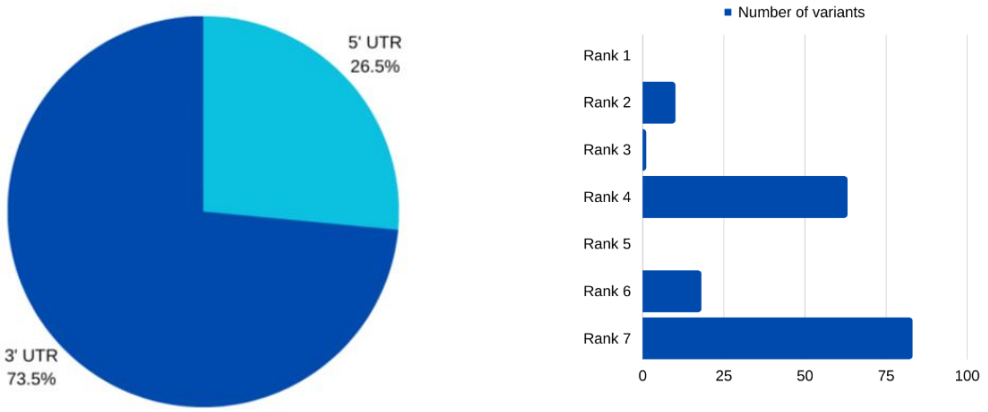
Therefore, numerous variants in *AURKC* have been identified, based on the filters outlined above, that hold potential for use in future studies. However, some of these variants can be prioritized as they are more likely to play a role in the pathogenicity of teratozoospermia because they were consistently identified in multiple analyses, as depicted in Table 5.

Table 5. List of prioritized variants in *AURKC* according to the tools used; ✓: SNP with significant effect; -: SNP with no effect; SF: spermatogenic failure.

Variant ID	Ancestral Allele	Alternative Allele	RegulomeDB Rank	3DSNP Score	SNPinfo (FuncPred)	GTex	Association with Diseases (ClinVar)	miRNAs (miRNASNP v3)	PolymiRTS 3.0	Type
rs11084490	C	A,C,T	1f	-	✓	✓	SF	-	-	5′ UTR
rs58264281	C	A,G,T	1f	-	✓	✓	SF	-	-	5′ UTR
rs35582299	G	A,C	1f	-	✓	-	-	✓	✓	3′ UTR
rs533889458	C	A	2a	10.02	-	-	-	-	-	5′ UTR
rs2361127	C	G,T	2b	-	✓	-	-	-	-	5′ UTR, 3′ UTR
rs55710619	G	A	-	-	✓	-	SF	-	-	3′ UTR

3.2. Evaluation of UTR Variants in SPATA16

According to the Ensembl genome browser [46], we initially identified a total of 365 variants in *SPATA16*. Following the removal of variants as previously outlined, we were left with 185 variants for more in-depth analysis. Among these, 49 were located in the 5′ UTR, while 136 were found in the 3′ UTR of *SPATA16* (as shown in Figure 4a and detailed in Table S6).



(**a**) (**b**)

Figure 4. (**a**) Total count of 3′ and 5′ UTR retrieved from the Ensembl genome browser [46] for *SPATA16*. (**b**) Annotation and ranks of UTR SNPs according to RegulomeDB [48].

Initially, we employed RegulomeDB [50] for the functional assessment of SNPs. Among the 185 variants, 175 were identified in RegulomeDB [50], and their rankings exhibited variation, as depicted in Figure 4b (also refer to Table S7). A total of 11 SNPs were prioritized, with 4 situated in the 3′ UTR and 7 in the 5′ UTR of *SPATA16*. The detailed results are presented in Table 6.

Table 6. List of SNPs predicted to have a functional effect according to RegulomeDB [50] for *SPATA16*.

Variant ID	Chromosome Location	Rank	Type
rs1315242177	3:173141125	2b	5′ UTR
rs1179618758	3:173141126	2b	5′ UTR
rs1380118082	3:173141127	2b	5′ UTR
rs1560137604	3:173141153	2b	5′ UTR
rs1366850067	3:173141154	2b	5′ UTR
rs1190973286	3:173141155	2b	5′ UTR
rs945530730	3:173141157	2b	5′ UTR
rs200008684	3:172913696	3a	3′ UTR
rs146640459	3:172916328	2b	3′ UTR
rs1375438765	3:172916329	2b	3′ UTR
rs1489045246	3:172916335	2b	3′ UTR

In terms of functional significance, we employed two additional tools. Among the 184 SNPs found in 3DSNP v2.0 [51] (Table S8), we identified 7 SNPs with a 3DSNP score greater exceeding 10, as detailed in Table 7. Furthermore, three variants (Table 8) were also determined to have putative functional effects according to SNPinfo (FuncPred) [52].

Table 7. List of SNPs predicted to have a functional effect according to 3DSNP v2.0 [51] for *SPATA16*.

Variant ID	Chromosome Location	Score
rs767677869	3:172916354	10.64
rs1341354857	3:172916355	10.64
rs1280211853	3:172916360	10.64
rs991346128	3:172916363	10.64
rs1330632536	3:172916364	10.64
rs1216343266	3:172916373	10.64
rs146640459	3:172916328	10.63

Table 8. List of SNPs predicted to have a functional effect according to SNPinfo (FuncPred) [52] for *SPATA16*; TFBS: transcription-factor binding site; ESE: exonic splicing enhancer; ESS: exonic splicing silencer; ✓ SNP that affects this site; -: SNP that does not affect this site.

Variant ID	Chromosome Location	TFBS	Splicing (ESE or ESS)
rs2673500	3:173141143	✓	✓
rs2673501	3:173141142	✓	✓
rs62622782	3:172913671	-	✓

Additionally, we assessed whether the UTR SNPs were present in expression quantitative trait loci (eQTLs) using the GTEx database [57]. Out of the 185 variants, only seven SNPs were identified in the GTEx database [57]. Nonetheless, our investigation did not reveal any significant correlation between these SNPs and the expression of *SPATA16* in the examined tissue. Similarly, when we used SNPnexus [58] to prioritize the SNPs, we discovered that none of them had been previously linked to diseases according to the NHGRI Catalogue of Published Genome-Wide Association Studies [59]. However, ClinVar [60] indicated that five SNPs are associated with spermatogenic failure (Table 9). Furthermore, none of the SNPs had a GERP score exceeding four, as detailed in Table S9.

Table 9. SNPs in *SPATA16* associated with pathogenic phenotypes according to ClinVar [60].

Variation ID	Chromosome Location	Phenotypes
rs115095786	3:172913722	Spermatogenic Failure
rs758141708	3:172913735	Spermatogenic Failure
rs189972919	3:173141205	Spermatogenic Failure
rs886058188	3:173141225	Spermatogenic Failure
rs566046620	3:173141226	Spermatogenic Failure

To assess the influence of UTR SNPs retrieved from Ensembl on miRNAs, we employed PolymiRTS Database 3.0 [63] and miRNASNP v3 [64]. PolymiRTS Database 3.0 [63] identified an association of one variant in *SPATA16* with miRNAs. More specifically, the function of an SNP at the PolymiRTS has been categorized into four types: "D" for disruption of a conserved miRNA site, "N" for non-conserved miRNA disruption, "C" for the formation of a new miRNA site, and "O" for instances where ancestral alleles could not be determined. As depicted in Table 10, rs148085657 affected the binding of three miRNAs, disrupting two conserved sites (Function D) and creating a new one (function C). Furthermore, according to miRNASNP v3 [64], 44 SNPs were implicated in the targeting or loss of miRNA target genes. Specifically, 3 SNPs caused gain, 3 caused loss, and 38 SNPs altered miRNA function by both gain and loss of target sites (Table S10).

Table 10. SNPs in *SPATA16* disrupting miRNA binding sites, the affected miRNAs, and the function class of SNPs according to PolymiRTS Database 3.0 [63].

Variant ID	Chromosome Location	Affected miRNAs	Function Class
rs148085657	3:172889484	hsa-miR-4267	D
		hsa-miR-5586-5p	D
		hsa-miR-5092	C

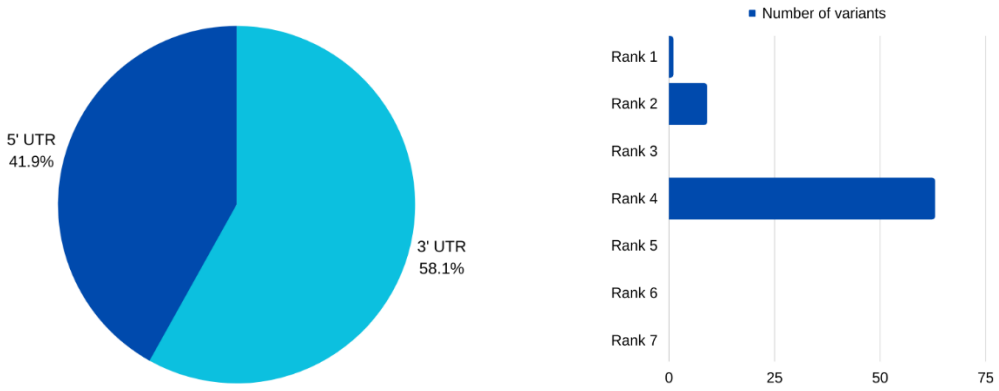
Hence, numerous variants have been prioritized based on the filters outlined above, and they hold potential for use in future studies. However, it is possible that some of these variants are more likely to play a role in the pathogenicity of teratozoospermia, as they were consistently identified in multiple analyses, as depicted in Table 11.

Table 11. List of prioritized variants in *SPATA16* according to the tools used; ✓ SNP that affects miRNA binding site; -: SNP with no effect.

Variant ID	Ancestral Allele	Alternative Allele	RegulomeDB Rank	3DSNP Score	miRNAs (miRNASNP v3)	miRNAs (PolymiRTS Database 3.0)	Type
rs146640459	C	A, G, T	2b	10.63	-	-	3′ UTR
rs148085657	G	C	-	-	✓	✓	3′ UTR

3.3. Evaluation of UTR Variants in SUN5

We initially retrieved 175 UTR variants from the Ensembl genome browser [46]. Following the elimination of duplicate variants and the exclusion of somatic mutations associated with cancer, a total of 74 variants remained for further analysis. Among these, 31 variants were located in the 5′ UTR and 43 were in the 3′ UTR (see Figure 5a and Table S11).



(**a**) (**b**)

Figure 5. (**a**) Total count of 3′ and 5′ UTR retrieved from the Ensembl genome browser [48] for SUN5. (**b**) Annotation and ranks of UTR SNPs according to RegulomeDB [50].

For the functional evaluation of SNPs, we utilized RegulomeDB [50], 3DSNP v2.0 [51], and SNPinfo (FuncPred) [52]. Among all the SNPs, 73 of them were found in RegulomeDB [50], and their distribution is depicted in Figure 5b and Table S12. A total of 10 SNPs were annotated, with 9 of them located in the 3′ UTR and 1 in the 5′ UTR of *SUN5*. These SNPs received RegulomeDB ranks ranging from 1a to 3b. The detailed results are presented in Table 12. Regarding 3DSNP v2.0 [51], none of the variants had a 3DSNP score greater than 10, as shown in Table S13. Additionally, no SNPs were identified as having putative functional effects according to SNPinfo (FuncPred) [52].

Table 12. List of SNPs predicted to have a functional effect according to RegulomeDB [50].

SNP IDs	Chromosome Location	Rank	Type
rs414191	20:32999946	1f	3′ UTR
rs1485087675	20:32983776	2b	3′ UTR
rs762026146	20:32983777	2b	3′ UTR
rs1478197315	20:32983779	2b	3′ UTR
rs1249428218	20:32983780	2b	3′ UTR
rs1294157041	20:32983780	2b	3′ UTR
rs567349892	20:32983781	2b	3′ UTR
rs1393231708	20:32983793	2b	3′ UTR
rs758511928	20:32999977	2b	3′ UTR
rs776302278	20:33004432	2b	5′ UTR

We also conducted an analysis of all UTR SNPs in *SUN5* using the GTEx database [57] to determine if they were present in expression quantitative trait loci (eQTLs). Out of the 74 variants, only two SNPs were identified in the GTEx database [57]. However, upon further investigation, no significant association was observed between these two SNPs and *SUN5* expression in testis tissue. Likewise, when we prioritized these SNPs using SNPnexus [58], we found that none of them had been previously associated with other diseases according to ClinVar [60] and the NHGRI Catalogue of Published Genome-Wide Association Studies [59]. Furthermore, none of the SNPs exhibited a GERP score greater than four, as indicated in Table S14.

Regarding the impact of variants on miRNA binding sites, only 10 out of the 74 SNPs were detected in the miRNASNP v3 [64]. Among them, three were predicted to impact miRNA binding sites, leading to the gain or loss of miRNA binding sites, as illustrated in Table 13. Additionally, it is worth noting that none of the 74 variants were identified in PolymiRTS Database 3.0 [63].

Table 13. SNPs in *SUN5* disrupting miRNA binding sites and the affected miRNAs according to the miRNASNP v3 [64].

SNP IDs	Chromosome Location	Affected miRNAs	Gain/Loss
rs1485087675	chr20:32983776	hsa-miR-7155-5p	Loss
rs762026146	chr20:32983777	hsa-miR-7162-3p	Gain
		hsa-miR-7155-5p	Loss
rs1478197315	chr20:32983779	hsa-miR-7155-5p	Loss

Therefore, the prioritized variants in SUN5 that can be used for further study as they are more likely to contribute to the pathogenicity of teratozoospermia are presented in Table 14.

Table 14. List of prioritized variants in *SUN5* according to the tools used; ✓ SNP that affects miRNA-binding site.

SNP IDs	Ancestral Allele	Alternative Allele	RegulomeDB Rank	miRNAs (miRNASNP v3)	Type
rs1485087675	G	T	2b	✓	3′ UTR
rs762026146	G	A	2b	✓	3′ UTR
rs1478197315	G	A	2b	✓	3′ UTR

4. Discussion

AURKC, *SPATA16*, and *SUN5* are pivotal genes known to play critical roles in the intricate processes of spermatogenesis and meiosis [7,8,31,66]. Numerous studies have provided evidence of the association between specific SNPs within these genes and teratozoospermia [16–21,26,28,29,33–35]. However, the sheer volume of SNPs within these genes poses a formidable challenge for comprehensive analysis. Herein, the indispensable role of bioinformatics tools comes into play, enabling the judicious selection of a limited number of prioritized variants. These selected variants hold the potential to significantly contribute to our understanding of teratozoospermia's pathogenesis and pave the way for future genetic screening endeavors. By pinpointing these key genetic factors, researchers can unravel the intricate molecular mechanisms underlying teratozoospermia, offering invaluable insights into both diagnosis and potential therapeutic strategies for this complex reproductive disorder.

Building upon this foundation, it is important to note that SNPs residing within the untranslated regions (UTRs) of genes are frequently overlooked but hold significant potential implications in the context of various pathologies [67]. In this study, we harnessed a diverse array of bioinformatics tools to thoroughly assess the impact of UTR variants within *AURKC*, *SPATA16*, and *SUN5*. The primary objective was to discern and prioritize these variants, thereby assembling a comprehensive catalog of SNPs that possess the promise of being instrumental in forthcoming investigations.

For *AURKC*, six SNPs emerged as prime candidates for potential pathogenicity, as corroborated by multiple analytical tools. Among these, three were situated within the 5′ UTR, while two resided in the 3′ UTR. Intriguingly, one SNP was characterized as impacting both the 3′ and 5′ UTRs for different transcripts of the *AURKC* gene. Notably, two of these prioritized SNPs, rs11084490 and rs58264281, were found to significantly affect *AURKC* expression in testis tissue. This observation aligns with SNPinfo (FuncPred) [52] predictions, which indicated that these variants perturbed transcription-factor binding sites. Existing evidence underscores the robust associations between nucleotide sequences within transcription-factor binding sites (TFBSs) and gene expression levels [68]. TFBS polymorphisms have garnered substantial attention, constituting 31% of trait-associated polymorphisms identified by genome-wide association studies (GWAS), underscoring their pivotal role in disease development [69]. According to ClinVar [60], these two SNPs are also linked to spermatogenic failure, particularly infertility associated with multi-tailed spermatozoa and excessive DNA, albeit being classified as benign. Nonetheless, our study findings advocate for their further exploration, given their potential regulatory role in teratozoospermia. Thus, subsequent investigations in a large sample of infertile males are suggested. Additionally, rs533889458 and rs2361127 earned prioritization based on their functional significance, with rs2361127 notably identified as a TFBS polymorphism, prompting the need for future studies elucidating its impact on *AURKC* expression levels. rs55710619 is another prioritized variant in *AURKC* that is associated with multi-tailed spermatozoa and excessive DNA and is characterized as likely benign according to ClinVar [60]. SNPinfo (FuncPred) [52] also ascribes functional significance to this variant. Special attention should be accorded to rs35582299, which exerts an impact on miRNA binding sites, as affirmed by several tools, including miRNASNP v3 [64], PolymiRTS Database 3.0 [63], and SNPinfo (FuncPred) [52]. More specifically, the above tools demonstrate that rs35582299 causes the loss or gain of sites affecting 29 miRNAs. miRNAs, small RNA molecules, are pivotal in gene expression regulation, and studies have revealed their differential expression between fertile and infertile males [41]. miRNAs fine-tune genes involved in sperm production and maturation, and dysregulation can disrupt this balance, culminating in abnormalities in sperm morphology and reduced fertility [70–72]. Thus, future investigations should delve into the list of miRNAs identified in this study as affected by SNPs, as they hold the potential to modulate *AURKC* expression. Intriguingly, none of the miRNAs that are affected by these SNPs have been previously implicated in male infertility. Similarly,

the above prioritized variants are reported for the first time as potentially involved in teratozoospermia and further exploration of their role is required.

For *SPATA16*, a gene with a crucial role in sperm production and testicular development [73], we identified two 3′ UTR variants through analysis with various tools. Among these prioritized SNPs, rs146640459 is indicated as a variant with functional significance according to 3DSNP v2.0 [51] and RegulomeDB [50]. Meanwhile, rs148085657 affects miRNA binding sites according to miRNASNP v3 [64] and PolymiRTS Database 3.0 [63]. More specifically, rs148085657 causes gain and loss of target sites, affecting six miRNAs (hsa-miR-5092, hsa-miR-205-5p, hsa-miR-5586-5p, hsa-miR-4267, hsa-miR-6512-3p, hsa-miR-6720-5p). Some of these miRNAs have been associated with different types of cancer [74–76], but none of them have shown any association with spermatogenesis or other aspects of male fertility. Similarly to *AURKC*, these two variants have not been previously associated with male infertility either.

For *SUN5*, three 3′ UTR variants were identified, all of which were characterized as having functional significance according to the RegulomeDB [50]. Simultaneously, these variants were found to impact miRNA binding sites according to miRNASNP v3 [64]. Specifically, two SNPs (rs1485087675 and rs1478197315) resulted in the loss of binding sites for the same miRNA, hsa-miR-7155-5p. The third SNP (rs762026146) not only disrupted the binding site of hsa-miR-7155-5p (resulting in target loss) but also created a binding site for hsa-miR-7162-3p. It is worth noting that there is limited research on these miRNAs, with only one publication suggesting that hsa-miR-7162-3p may play a role in the repair of endometrial stromal cell injury [77]. Furthermore, there are no available studies for the three variants prioritized and no association with male reproduction.

The present study has yielded a wealth of data and identified numerous SNPs in *AURKC*, *SPATA16*, and *SUN5* genes that hold promise for future investigations into the molecular mechanisms of teratozoospermia. To guide future research efforts, functional experiments can be designed to validate the roles of these SNPs in teratozoospermia. These experiments may focus on assessing their functional impact on mRNA–miRNA interactions and exploring how these SNPs influence the expression of *AURKC, SPATA16,* and *SUN5* genes, particularly in tissues relevant to male fertility and reproductive organs, such as the testes. Additionally, conducting large-scale GWAS studies in cohorts of individuals with and without male infertility can provide valuable insights by determining whether these SNPs are more prevalent in the affected group, thus establishing a link between these genetic variants and male infertility risk.

As the variants identified in this study may significantly contribute to teratozoospermia, it is imperative to discuss the role of UTR variants in gene regulation and their potential impact on disease pathogenesis, particularly in the context of male infertility. Variants situated within the 3′ UTRs of genes can exert profound effects on gene expression and subsequent cellular functions [78–80]. More specifically, these alterations can affect mRNA stability, thus influencing the half-life of the messenger RNA and ultimately modulating protein expression levels [81]. Additionally, disruptions in the 3′ UTR can intricately perturb post-transcriptional regulatory mechanisms, encompassing RNA processing, transport, localization, and degradation, consequently leading to dysregulation of essential cellular processes and pathways. Furthermore, they might influence mRNA localization within the cell, thereby impacting local protein synthesis and altering various cellular activities [78–81]. Similarly, variations in the 5′ UTRs can disrupt the efficiency of translation initiation, thereby affecting ribosomal binding and subsequent translation processes, ultimately resulting in variations in protein synthesis levels. Moreover, they can interfere with the binding sites for specific transcription factors or regulatory proteins, potentially influencing gene transcription and leading to dysregulation of downstream cellular processes [42,78].

As the UTR variants identified in this study are within AURKC, SPATA16, and SUN5, they may disrupt the regulatory mechanisms of these genes, potentially contributing to male infertility due to the crucial roles of the above genes in spermatogenesis. Specifically,

AURKC regulates chromosomal segregation during meiosis, ensuring the production of genetically balanced gametes [13], while SPATA16 is essential for sperm function and fertilization, participating in various processes critical for normal sperm development and function, including sperm–egg interaction and fusion [23,24]. Similarly, SUN5, belonging to the SUN domain family, is indispensable for sperm head shaping and nuclear membrane remodeling during spermatogenesis [31,34]. Thus, any UTR variants in these genes can potentially disrupt gene expression through the mechanisms described earlier, consequently impacting the intricate process of spermatogenesis. It is also worth noting that given the significant role of these genes in fertilization, UTR variants may extensively disturb the processes involved in sperm production, affecting other crucial sperm parameters such as motility or count. Therefore, further investigations, including functional experiments, are imperative to elucidate the precise mechanisms of action of the reported variants and their specific impacts on additional sperm parameters beyond morphology and teratozoospermia.

Furthermore, given the pivotal role of miRNAs in various cellular processes, it is highly promising to delve deeper into the broader miRNA interaction network involving the SNPs reported in this study. This exploration can help identify other miRNAs that may be affected by these SNPs, unveiling potential overlapping or synergistic effects on gene regulation. Notably, miRNAs altered by these SNPs may exhibit differential expression between fertile and infertile males, offering the potential for their use as biomarkers for assessing male infertility risk or as therapeutic targets. These avenues of research align with the primary goal of our study, which was to provide a prioritized list of SNPs and miRNAs to catalyze future investigations in the field of teratozoospermia.

Finally, it is important to acknowledge that while our study was an in-depth analysis employing an extensive array of bioinformatics tools and stringent criteria, it is an in silico study with inherent limitations. As such, further research is imperative to validate and expand upon the findings presented here.

5. Conclusions

This study has unveiled the potential functional significance of six UTR SNPs in *AURKC*, two UTR SNPs in *SPATA16*, and three UTR SNPs in *SUN5*, offering valuable insights into the genotype–phenotype relationship in teratozoospermia. Additionally, we have compiled a comprehensive list of miRNAs that normally target these genes, but their binding is affected due to the identified UTR SNPs. Notably, the 11 SNPs identified in this study have remained relatively unexplored to date and have not previously been associated with male infertility. Similarly, limited information is available concerning the miRNAs affected by these SNPs. Consequently, our study findings serve as a guiding map for fellow researchers, facilitating the exploration of the molecular mechanisms underpinning teratozoospermia. This research direction holds promise for advancing ART and enhancing the diagnosis and prognosis of this condition.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/dna3040013/s1, Table S1: List of total UTR variants in AURKC; Table S2: RegulomeDB scores and ranks for all UTR variants in AURKC (all) and UTR variants in AURKC with RegulomeDB rank <3 (significant); Table S3: 3DSNP scores for all UTR variants in AURKC; Table S4: GERP scores for all UTR variants in AURKC; Table S5: UTR variants in *AURKC* affecting miRNA binding sites according to miRNASNP v3 [64]; Table S6: List of all UTR variants in SPATA16; Table S7: RegulomeDB scores and ranks for all UTR variants in SPATA16; Table S8: 3DSNP scores for all UTR variants in SPATA16; Table S9: GERP scores for UTR variants in SPATA16; Table S10: Impact of UTR variants in *SPATA16* in miRNA binding sites according to miRNASNP v3 [64]; Table S11: List of total UTR variants in SUN5; Table S12: RegulomeDB scores and ranks for all UTR variants in SUN5; Table S13: 3DSNP scores for all UTR variants in SUN5; Table S14: GERP scores for UTR variants in SUN5.

Author Contributions: Conceptualization, M.-A.K. and Z.M.; methodology, M.-A.K. and Z.M.; formal analysis, M.-A.K. and Z.M.; investigation, M.-A.K. and Z.M.; writing—original draft preparation, M.-A.K.; writing—review and editing, Z.M.; visualization, M.-A.K.; supervision, Z.M.; project administration, Z.M.; funding acquisition, Z.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All the relevant data are provided in the manuscript and in the Supplementary data.

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

References

1. Pathak, U.I.; Gabrielsen, J.S.; Lipshultz, L.I. Cutting-Edge Evaluation of Male Infertility. *Urol. Clin. North Am.* **2020**, *47*, 129–138. [CrossRef] [PubMed]
2. Schulte, R.T.; Ohl, D.A.; Sigman, M.; Smith, G.D. Sperm DNA Damage in Male Infertility: Etiologies, Assays, and Outcomes. *J. Assist. Reprod. Genet.* **2010**, *27*, 3–12. [CrossRef]
3. 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] [PubMed]
4. Agarwal, A.; Tvrda, E.; Sharma, R. Relationship amongst Teratozoospermia, Seminal Oxidative Stress and Male Infertility. *Reprod. Biol. Endocrinol.* **2014**, *12*, 45. [CrossRef] [PubMed]
5. WHO. *WHO Laboratory Manual for the Examination of Human Semen and Semen-Cervical Mucus Interaction*; Cambridge University Press: Cambridge, UK, 1980.
6. WHO. *WHO Laboratory Manual for the Examination and Processing of Human Semen*; WHO: Geneva, Switzerland, 2010.
7. Coutton, C.; Escoffier, J.; Martinez, G.; Arnoult, C.; Ray, P.F. Teratozoospermia: Spotlight on the Main Genetic Actors in the Human. *Hum. Reprod. Update* **2015**, *21*, 455–485. [CrossRef]
8. Beurois, J.; Cazin, C.; Kherraf, Z.E.; Martinez, G.; Celse, T.; Touré, A.; Arnoult, C.; Ray, P.F.; Coutton, C. Genetics of Teratozoospermia: Back to the Head. *Best Pract. Res. Clin. Endocrinol. Metab.* **2020**, *34*, 101473. [CrossRef]
9. Chemes, H.E.; Rawe, V.Y. Sperm Pathology: A Step beyond Descriptive Morphology. Origin, Characterization and Fertility Potential of Abnormal Sperm Phenotypes in Infertile Men. *Hum. Reprod. Update* **2003**, *9*, 405–428. [CrossRef] [PubMed]
10. Cocuzza, M.; Sikka, S.C.; Athayde, K.S.; Agarwal, A. Clinical Relevance of Oxidative Stress and Sperm Chromatin Damage in Male Infertility: An Evidence Based Analysis. *Int. Braz. J. Urol.* **2007**, *33*, 603–621. [CrossRef]
11. Said, T.M.; Agarwal, A.; Sharma, R.K.; Mascha, E.; Sikka, S.C.; Thomas, A.J. Human Sperm Superoxide Anion Generation and Correlation with Semen Quality in Patients with Male Infertility. *Fertil. Steril.* **2004**, *82*, 871–877. [CrossRef]
12. Jan, S.Z.; Vormer, T.L.; Jongejan, A.; Röling, M.D.; Silber, S.J.; de Rooij, D.G.; Hamer, G.; Repping, S.; van Pelt, A.M.M. Unraveling Transcriptome Dynamics in Human Spermatogenesis. *Development* **2017**, *144*, 3659–3673. [CrossRef]
13. Lin, Y.S.; Su, L.J.; Yu, C.T.; Wong, F.H.; Yeh, H.H.; Chen, S.L.; Wu, J.C.; Lin, W.J.; Shiue, Y.L.; Liu, H.S.; et al. Gene Expression Profiles of the Aurora Family Kinases. *Gene Expr.* **2006**, *13*, 15–26. [CrossRef]
14. Du, R.; Huang, C.; Liu, K.; Li, X.; Dong, Z. Targeting AURKA in Cancer: Molecular Mechanisms and Opportunities for Cancer Therapy. *Mol. Cancer* **2021**, *20*, 15. [CrossRef]
15. Bejar, J.F.; DiSanza, Z.; Quartuccio, S.M. The Oncogenic Role of Meiosis-Specific Aurora Kinase C in Mitotic Cells. *Exp. Cell Res.* **2021**, *407*, 112803. [CrossRef]
16. Dieterich, K.; Soto Rifo, R.; Karen Faure, A.; Hennebicq, S.; Amar, B.B.; Zahi, M.; Perrin, J.; Martinez, D.; Sèle, B.; Jouk, P.S.; et al. Homozygous Mutation of AURKC Yields Large-Headed Polyploid Spermatozoa and Causes Male Infertility. *Nat. Genet.* **2007**, *39*, 661–665. [CrossRef] [PubMed]
17. Ben Khelifa, M.; Zouari, R.; Harbuz, R.; Halouani, L.; Arnoult, C.; Lunardi, J.; Ray, P.F. A New AURKC Mutation Causing Macrozoospermia: Implications for Human Spermatogenesis and Clinical Diagnosis. *Mol. Hum. Reprod.* **2011**, *17*, 762–768. [CrossRef]
18. Ben Khelifa, M.; Coutton, C.; Blum, M.G.B.; Abada, F.; Harbuz, R.; Zouari, R.; Guichet, A.; May-Panloup, P.; Mitchell, V.; Rollet, J.; et al. Identification of a New Recurrent Aurora Kinase C Mutation in Both European and African Men with Macrozoospermia. *Hum. Reprod.* **2012**, *27*, 3337–3346. [CrossRef] [PubMed]
19. Dieterich, K.; Zouari, R.; Harbuz, R.; Vialard, F.; Martinez, D.; Bellayou, H.; Prisant, N.; Zoghmar, A.; Guichaoua, M.R.; Koscinski, I.; et al. The Aurora Kinase C c.144delC Mutation Causes Meiosis I Arrest in Men and Is Frequent in the North African Population. *Hum. Mol. Genet.* **2009**, *18*, 1301–1309. [CrossRef]

20. Ounis, L.; Zoghmar, A.; Coutton, C.; Rouabah, L.; Hachemi, M.; Martinez, D.; Martinez, G.; Bellil, I.; Khelifi, D.; Arnoult, C.; et al. Mutations of the Aurora Kinase C Gene Causing Macrozoospermia Are the Most Frequent Genetic Cause of Male Infertility in Algerian Men. *Asian J. Androl.* **2015**, *17*, 68–73. [CrossRef]
21. Hua, J.; Wan, Y. yang Whole-Exome Sequencing Identified a Novel Mutation of AURKC in a Chinese Family with Macrozoospermia. *J. Assist. Reprod. Genet.* **2019**, *36*, 529–534. [CrossRef] [PubMed]
22. Zhang, Q.; Zhang, F.; Chen, X.H.; Wang, Y.Q.; Wang, W.Q.; Lin, A.A.; Cavalli-Sforza, L.L.; Jin, L.; Huo, R.; Sha, J.H.; et al. Rapid Evolution, Genetic Variations, and Functional Association of the Human Spermatogenesis-Related Gene NYD-SP12. *J. Mol. Evol.* **2007**, *65*, 154–161. [CrossRef]
23. Xu, M.; Xiao, J.; Chen, J.; Li, J.; Yin, L.; Zhu, H.; Zhou, Z.; Sha, J. Identification and Characterization of a Novel Human Testis-Specific Golgi Protein, NYD-SP12. *Mol. Hum. Reprod.* **2003**, *9*, 9–17. [CrossRef]
24. Lu, L.; Lin, M.; Xu, M.; Zhou, Z.M.; Sha, J.H. Gene Functional Research Using Polyethylenimine-Mediated in Vivo Gene Transfection into Mouse Spermatogenic Cells. *Asian J. Androl.* **2006**, *8*, 53–59. [CrossRef]
25. ElInati, E.; Fossard, C.; Okutman, O.; Ghédir, H.; Ibala-Romdhane, S.; Ray, P.F.; Saad, A.; Hennebicq, S.; Viville, S. A New Mutation Identified in SPATA16 in Two Globozoospermic Patients. *J. Assist. Reprod. Genet.* **2016**, *33*, 815–820. [CrossRef] [PubMed]
26. Dam, A.H.D.M.; Koscinski, I.; Kremer, J.A.M.; Moutou, C.; Jaeger, A.S.; Oudakker, A.R.; Tournaye, H.; Charlet, N.; Lagier-Tourenne, C.; Van Bokhoven, H.; et al. Homozygous Mutation in SPATA16 Is Associated with Male Infertility in Human Globozoospermia. *Am. J. Hum. Genet.* **2007**, *81*, 813–820. [CrossRef]
27. Fujihara, Y.; Oji, A.; Larasati, T.; Kojima-Kita, K.; Ikawa, M. Human Globozoospermia-Related Gene Spata16 Is Required for Sperm Formation Revealed by CRISPR/Cas9-Mediated Mouse Models. *Int. J. Mol. Sci.* **2017**, *18*, 2208. [CrossRef]
28. Behvarz, M.; Rahmani, S.A.; Siasi Torbati, E.; Danaei Mehrabad, S.; Bikhof Torbati, M. Association of CATSPER1, SPATA16 and TEX11 Genes Polymorphism with Idiopathic Azoospermia and Oligospermia Risk in Iranian Population. *BMC Med. Genomics* **2022**, *15*, 47. [CrossRef]
29. Zhu, F.; Wang, F.; Yang, X.; Zhang, J.; Wu, H.; Zhang, Z.; Zhang, Z.; He, X.; Zhou, P.; Wei, Z.; et al. Biallelic SUN5 Mutations Cause Autosomal-Recessive Acephalic Spermatozoa Syndrome. *Am. J. Hum. Genet.* **2016**, *99*, 942–949. [CrossRef] [PubMed]
30. Jiang, X.Z.; Yang, M.G.; Huang, L.H.; Li, C.Q.; Xing, X.W. SPAG4L, a Novel Nuclear Envelope Protein Involved in the Meiotic Stage of Spermatogenesis. *DNA Cell Biol.* **2011**, *30*, 875–882. [CrossRef] [PubMed]
31. Shang, Y.; Zhu, F.; Wang, L.; Ouyang, Y.C.; Dong, M.Z.; Liu, C.; Zhao, H.; Cui, X.; Ma, D.; Zhang, Z.; et al. Essential Role for SUN5 in Anchoring Sperm Head to the Tail. *eLife* **2017**, *6*, e28199. [CrossRef]
32. McGee, M.D.; Rillo, R.; Anderson, A.S.; Starr, D.A. UNC-83 Is a KASH Protein Required for Nuclear Migration and Is Recruited to the Outer Nuclear Membrane by a Physical Interaction with the SUN Protein UNC-84. *Mol. Biol. Cell* **2006**, *17*, 1790–1801. [CrossRef]
33. Xiang, M.; Wang, Y.; Wang, K.; Kong, S.; Lu, M.; Zhang, J.; Duan, Z.; Zha, X.; Shi, X.; Wang, F.; et al. Novel Mutation and Deletion in SUN5 Cause Male Infertility with Acephalic Spermatozoa Syndrome. *Reprod. Sci.* **2022**, *29*, 646–651. [CrossRef]
34. Sha, Y.W.; Xu, X.; Ji, Z.Y.; Lin, S.-B.; Wang, X.; Qiu, P.P.; Zhou, Y.; Mei, L.-B.; Su, Z.Y.; Li, L.; et al. Genetic Contribution of SUN5 Mutations to Acephalic Spermatozoa in Fujian China. *Gene* **2018**, *647*, 221–225. [CrossRef] [PubMed]
35. Liu, G.; Wang, N.; Zhang, H.; Yin, S.; Dai, H.; Lin, G.; Li, W. Novel Mutations in PMFBP1, TSGA10 and SUN5: Expanding the Spectrum of Mutations That May Cause Acephalic Spermatozoa. *Clin. Genet.* **2020**, *97*, 938–939. [CrossRef] [PubMed]
36. Brookes, A.J. The Essence of SNPs. *Gene* **1999**, *234*, 177–186. [CrossRef]
37. Shah, H.; Khan, K.; Badshah, Y.; Mahmood Ashraf, N.; Shabbir, M.; Trembley, J.H.; Afsar, T.; Abusharha, A.; Razak, S. Investigation of UTR Variants by Computational Approaches Reveal Their Functional Significance in PRKCI Gene Regulation. *Genes* **2023**, *14*, 247. [CrossRef]
38. Birney, E.; Stamatoyannopoulos, J.A.; Dutta, A.; Guigó, R.; Gingeras, T.R.; Margulies, E.H.; Weng, Z.; Snyder, M.; Dermitzakis, E.T.; Thurman, R.E.; et al. Identification and Analysis of Functional Elements in 1% of the Human Genome by the ENCODE Pilot Project. *Nature* **2007**, *447*, 799–816. [CrossRef]
39. Chasman, D.; Adams, R.M. Predicting the Functional Consequences of Non-Synonymous Single Nucleotide Polymorphisms: Structure-Based Assessment of Amino Acid Variation. *J. Mol. Biol.* **2001**, *307*, 683–706. [CrossRef]
40. Sinaei, R.; Jamebozorgi, K.; Mirshekarpour, H.; Poormasoumi, H.; Mahdizadeh, A.; Akbari, Z.; Taghizadeh, E. The Role of MiRNAs in the Diagnosis and Treatment of Male Infertility: A Review Study. *Egypt. J. Med. Hum. Genet.* **2023**, *24*, 40. [CrossRef]
41. Forouhari, S.; Mahmoudi, E.; Safdarian, E.; Beygi, Z.; Gheibihayat, S.M. MicroRNA: A Potential Diagnosis for Male Infertility. *Mini Rev. Med. Chem.* **2021**, *21*, 1226–1236. [CrossRef] [PubMed]
42. Chatterjee, S.; Berwal, S.K.; Pal, J.K. Pathological Mutations in 5′ Untranslated Regions of Human Genes. *eLS* **2010**. [CrossRef]
43. Deng, N.; Zhou, H.; Fan, H.; Yuan, Y.; Deng, N.; Zhou, H.; Fan, H.; Yuan, Y. Single Nucleotide Polymorphisms and Cancer Susceptibility. *Oncotarget* **2017**, *8*, 110635–110649. [CrossRef]
44. Kalia, N.; Sharma, A.; Kaur, M.; Kamboj, S.S.; Singh, J. A Comprehensive in Silico Analysis of Non-Synonymous and Regulatory SNPs of Human MBL2 Gene. *Springerplus* **2016**, *5*, 811. [CrossRef]
45. Bhatti, P.; Church, D.M.; Rutter, J.L.; Struewing, J.P.; Sigurdson, A.J. Candidate Single Nucleotide Polymorphism Selection Using Publicly Available Tools: A Guide for Epidemiologists. *Am. J. Epidemiol.* **2006**, *164*, 794–804. [CrossRef] [PubMed]
46. Li, L.; Wei, D. Bioinformatics Tools for Discovery and Functional Analysis of Single Nucleotide Polymorphisms. *Adv. Exp. Med. Biol.* **2015**, *827*, 287–310. [CrossRef]

47. Kashan, H.S.; Albakrye, A.M.; Elnasri, H.A.; Khaier, M.A.M. In Silico Analysis of Single Nucleotide Polymorphisms in Human GCH1 Gene. *Inform. Med. Unlocked* **2021**, *27*, 100808. [CrossRef]
48. Martin, F.J.; Amode, M.R.; Aneja, A.; Austine-Orimoloye, O.; Azov, A.G.; Barnes, I.; Becker, A.; Bennett, R.; Berry, A.; Bhai, J.; et al. Ensembl 2023. *Nucleic Acids Res.* **2023**, *51*, D933–D941. [CrossRef]
49. Tate, J.G.; Bamford, S.; Jubb, H.C.; Sondka, Z.; Beare, D.M.; Bindal, N.; Boutselakis, H.; Cole, C.G.; Creatore, C.; Dawson, E.; et al. COSMIC: The Catalogue Of Somatic Mutations In Cancer. *Nucleic Acids Res.* **2019**, *47*, D941–D947. [CrossRef] [PubMed]
50. Boyle, A.P.; Hong, E.L.; Hariharan, M.; Cheng, Y.; Schaub, M.A.; Kasowski, M.; Karczewski, K.J.; Park, J.; Hitz, B.C.; Weng, S.; et al. Annotation of Functional Variation in Personal Genomes Using RegulomeDB. *Genome Res.* **2012**, *22*, 1790–1797. [CrossRef] [PubMed]
51. Quan, C.; Ping, J.; Lu, H.; Zhou, G.; Lu, Y. 3DSNP 2.0: Update and Expansion of the Noncoding Genomic Variant Annotation Database. *Nucleic Acids Res.* **2022**, *50*, D950–D955. [CrossRef]
52. Xu, Z.; Taylor, J.A. SNPinfo: Integrating GWAS and Candidate Gene Information into Functional SNP Selection for Genetic Association Studies. *Nucleic Acids Res.* **2009**, *37*, W600. [CrossRef]
53. Shahin, M.H.; Sá, A.C.; Webb, A.; Gong, Y.; Langaee, T.; McDonough, C.W.; Riva, A.; Beitleshees, A.L.; Chapman, A.B.; Gums, J.G.; et al. Genome-Wide Prioritization and Transcriptomics Reveal Novel Signatures Associated with Thiazide Diuretics Blood Pressure Response. *Circ. Cardiovasc. Genet.* **2017**, *10*, e001404. [CrossRef] [PubMed]
54. Kyrgiafini, M.A.; Sarafidou, T.; Giannoulis, T.; Chatziparasidou, A.; Christoforidis, N.; Mamuris, Z. Gene-by-Sex Interactions: Genome-Wide Association Study Reveals Five SNPs Associated with Obesity and Overweight in a Male Population. *Genes* **2023**, *14*, 799. [CrossRef]
55. Han, Z.; Huang, H.; Gao, Y.; Huang, Q. Functional Annotation of Alzheimer's Disease Associated Loci Revealed by GWASs. *PLoS ONE* **2017**, *12*, e0179677. [CrossRef] [PubMed]
56. Singh, B.; Maiti, G.P.; Zhou, X.; Fazel-Najafabadi, M.; Bae, S.C.; Sun, C.; Terao, C.; Okada, Y.; Heng Chua, K.; Kochi, Y.; et al. Lupus Susceptibility Region Containing CDKN1B Rs34330 Mechanistically Influences Expression and Function of Multiple Target Genes, Also Linked to Proliferation and Apoptosis. *Arthritis Rheumatol.* **2021**, *73*, 2303–2313. [CrossRef]
57. 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]
58. Oscanoa, J.; Sivapalan, L.; Gadaleta, E.; Dayem Ullah, A.Z.; Lemoine, N.R.; Chelala, C. SNPnexus: A Web Server for Functional Annotation of Human Genome Sequence Variation (2020 Update). *Nucleic Acids Res.* **2020**, *48*, W185–W192. [CrossRef] [PubMed]
59. Sollis, E.; Mosaku, A.; Abid, A.; Buniello, A.; Cerezo, M.; Gil, L.; Groza, T.; Güneş, O.; Hall, P.; Hayhurst, J.; et al. The NHGRI-EBI GWAS Catalog: Knowledgebase and Deposition Resource. *Nucleic Acids Res.* **2023**, *51*, D977–D985. [CrossRef]
60. Landrum, M.J.; Lee, J.M.; Benson, M.; Brown, G.R.; Chao, C.; Chitipiralla, S.; Gu, B.; Hart, J.; Hoffman, D.; Jang, W.; et al. ClinVar: Improving Access to Variant Interpretations and Supporting Evidence. *Nucleic Acids Res.* **2018**, *46*, D1062–D1067. [CrossRef] [PubMed]
61. Davydov, E.V.; Goode, D.L.; Sirota, M.; Cooper, G.M.; Sidow, A.; Batzoglou, S. Identifying a High Fraction of the Human Genome to Be under Selective Constraint Using GERP++. *PLoS Comput. Biol.* **2010**, *6*, e1001025. [CrossRef] [PubMed]
62. Huber, C.D.; Kim, B.Y.; Lohmueller, K.E. Population Genetic Models of GERP Scores Suggest Pervasive Turnover of Constrained Sites across Mammalian Evolution. *PLoS Genet.* **2020**, *16*, e1008827. [CrossRef]
63. Bhattacharya, A.; Ziebarth, J.D.; Cui, Y. PolymiRTS Database 3.0: Linking Polymorphisms in MicroRNAs and Their Target Sites with Human Diseases and Biological Pathways. *Nucleic Acids Res.* **2014**, *42*, D86–D91. [CrossRef]
64. Liu, C.J.; Fu, X.; Xia, M.; Zhang, Q.; Gu, Z.; Guo, A.Y. MiRNASNP-v3: A Comprehensive Database for SNPs and Disease-Related Variations in MiRNAs and MiRNA Targets. *Nucleic Acids Res.* **2021**, *49*, D1276–D1281. [CrossRef]
65. Mesbah-Uddin, M.; Elango, R.; Banaganapalli, B.; Shaik, N.A.; Al-Abbasi, F.A. In-Silico Analysis of Inflammatory Bowel Disease (IBD) GWAS Loci to Novel Connections. *PLoS ONE* **2015**, *10*, e0119420. [CrossRef] [PubMed]
66. De Braekeleer, M.; Nguyen, M.H.; Morel, F.; Perrin, A. Genetic Aspects of Monomorphic Teratozoospermia: A Review. *J. Assist. Reprod. Genet.* **2015**, *32*, 615–623. [CrossRef]
67. Steri, M.; Idda, M.L.; Whalen, M.B.; Orrù, V. Genetic Variants in MRNA Untranslated Regions. *Wiley Interdiscip. Rev. RNA* **2018**, *9*, e1474. [CrossRef] [PubMed]
68. Tseng, C.C.; Wong, M.C.; Liao, W.T.; Chen, C.J.; Lee, S.C.; Yen, J.H.; Chang, S.J. Genetic Variants in Transcription Factor Binding Sites in Humans: Triggered by Natural Selection and Triggers of Diseases. *Int. J. Mol. Sci.* **2021**, *22*, 22. [CrossRef] [PubMed]
69. Nishizaki, S.S.; Ng, N.; Dong, S.; Porter, R.S.; Morterud, C.; Williams, C.; Asman, C.; Switzenberg, J.A.; Boyle, A.P. Predicting the Effects of SNPs on Transcription Factor Binding Affinity. *Bioinformatics* **2020**, *36*, 364–372. [CrossRef] [PubMed]
70. Gou, L.T.; Dai, P.; Liu, M.F. Small Noncoding RNAs and Male Infertility. *Wiley Interdiscip. Rev. RNA* **2014**, *5*, 733–745. [CrossRef]
71. Kotaja, N. MicroRNAs and Spermatogenesis. *Fertil. Steril.* **2014**, *101*, 1552–1562. [CrossRef] [PubMed]
72. Khawar, M.B.; Mehmood, R.; Roohi, N. MicroRNAs: Recent Insights towards Their Role in Male Infertility and Reproductive Cancers. *Bosn. J. Basic Med. Sci.* **2019**, *19*, 31. [CrossRef]
73. Roozbahani, G.A.; Sheidai, M.; Noormohammadi, Z.; Gourabi, H. Association Study of SPATA-16 Polymorphism with Male Infertility in Iranian Population. *Meta Gene* **2017**, *13*, 154–158. [CrossRef]

74. Jiang, M.; Zhong, T.; Zhang, W.; Xiao, Z.; Hu, G.; Zhou, H.; Kuang, H. Reduced Expression of MiR-205-5p Promotes Apoptosis and Inhibits Proliferation and Invasion in Lung Cancer A549 Cells by Upregulation of ZEB2 and Downregulation of ErbB3. *Mol. Med. Rep.* **2017**, *15*, 3231–3238. [CrossRef] [PubMed]
75. Toro, A.U.; Shukla, S.K.; Bansal, P. Micronome Revealed MiR-205-5p as Key Regulator of VEGFA During Cancer Related Angiogenesis in Hepatocellular Carcinoma. *Mol. Biotechnol.* **2023**, *65*, 1178–1186. [CrossRef] [PubMed]
76. Gupta, S.R.R.; Nagar, G.; Mittal, P.; Rana, S.; Singh, H.; Singh, R.; Singh, A.; Singh, I.K. Breast Cancer Therapeutics and Hippo Signaling Pathway: Novel MicroRNA-Gene-Protein Interaction Networks. *OMICS* **2023**, *27*, 273–280. [CrossRef] [PubMed]
77. Shi, Q.; Wang, D.; Ding, X.; Yang, X.; Zhang, Y. Exosome-Shuttled MiR-7162-3p from Human Umbilical Cord Derived Mesenchymal Stem Cells Repair Endometrial Stromal Cell Injury by Restricting APOL6. *Arch. Biochem. Biophys.* **2021**, *707*, 108887. [CrossRef] [PubMed]
78. Barrett, L.W.; Fletcher, S.; Wilton, S.D. Regulation of Eukaryotic Gene Expression by the Untranslated Gene Regions and Other Non-Coding Elements. *Cell. Mol. Life Sci.* **2012**, *69*, 3613. [CrossRef]
79. Hesketh, J. 3′ UTRs and Regulation. *eLS* **2005**. [CrossRef]
80. Zhao, W.; Blagev, D.; Pollack, J.L.; Erle, D.J. Toward a Systematic Understanding of MRNA 3′ Untranslated Regions. *Proc. Am. Thorac. Soc.* **2011**, *8*, 163. [CrossRef]
81. Hong, D.; Jeong, S. 3′UTR Diversity: Expanding Repertoire of RNA Alterations in Human MRNAs. *Mol. Cells* **2023**, *46*, 48. [CrossRef]