

Inherited defects of piRNA biogenesis cause transposon de-repression, impaired spermatogenesis, and human male infertility

Birgit Stallmeyer

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.

Clara Bühlmann

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.

Rytis Stakaitis

Division of Genetics, Oregon National Primate Research Center, Oregon Health & Science University, Portland, OR 97006, USA.; Laboratory of Molecular Neurooncology, Neuroscience Institute, Lithuanian University of Health Sciences, Kaunas 50161, Lithuania.

Ann-Kristin Dicke

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.

<https://orcid.org/0000-0002-2171-0525>

Farah Ghieh

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.

Luisa Meier

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.

Ansgar Zoch

Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, Edinburgh, UK.; Wellcome Centre for Cell Biology, School of Biological Sciences, The University of Edinburgh, UK.

David MacKenzie MacLeod

Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, Edinburgh, UK.; Wellcome Centre for Cell Biology, School of Biological Sciences, The University of Edinburgh, UK.

Johanna Steingröver

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.

Özlem Okutman

Laboratoire de Génétique Médicale LGM, Institut de génétique médicale d'Alsace IGMA, INSERM UMR 1112; Université de Strasbourg, France. Université libre de Bruxelles (ULB), Hôpital Universitaire de Bruxelles, Hôpital Erasme, Service de Gynécologie-Obstétrique, Clinique de Fertilité, Bruxelles, Belgium.

Daniela Fietz

Institute of Veterinary Anatomy, Histology and Embryology, Justus-Liebig-Universität Gießen

Adrian Pilatz

Clinic for Urology, Paediatric Urology and Andrology, Justus Liebig University Gießen, 35390 Gießen, Germany. <https://orcid.org/0000-0001-8072-1841>

Antoni Riera Escamilla

Andrology Department, Fundació Puigvert, Universitat Autònoma de Barcelona, Instituto de Investigaciones Biomédicas Sant Pau, Barcelona, 08025 Catalonia, Spain. <https://orcid.org/0000-0002-4485-3094>

Miguel Xavier

Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne NE2 4HH, UK. <https://orcid.org/0000-0003-0709-7223>

Christian Ruckert

Institute of Human Genetics, University of Münster, 48149 Münster, Germany.

Sara Di Persio

Centre of Reproductive Medicine and Andrology, University Hospital Münster, Münster 48149, Germany.

Nina Neuhaus

Centre of Reproductive Medicine and Andrology, University Hospital Münster, Münster 48149, Germany.

<https://orcid.org/0000-0003-0181-6194>

Ali Sami Gurbuz

Department of Gynecology and Obstetrics Novafertil IVF Center, Konya, Turkey.

Ahmend Şalvarci

Department of Andrology Novafertil IVF Center, Konya, Turkey.

Nicolas Le May

Laboratoire de Génétique Médicale LGM, Institut de génétique médicale d'Alsace IGMA, INSERM UMR 1112, Université de Strasbourg, France.

Kevin McEleny

Newcastle Fertility Centre, The Newcastle upon Tyne Hospitals NHS Foundation Trust, Newcastle upon Tyne, UK.

Corinna Friedrich

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.

Godfried van der Heijden

Department of Obstetrics and Gynecology, Radboud University Medical Center, Nijmegen, the Netherlands. <https://orcid.org/0000-0001-5894-7027>

Margot J. Wyrwoll

Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.; Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, Edinburgh, UK.

Sabine Kliesch

Centre of Reproductive Medicine and Andrology, University Hospital Münster, Münster 48149, Germany.

<https://orcid.org/0000-0002-7561-4870>

Joris A. Veltman

Biosciences Institute, Faculty of Medical Sciences, Newcastle University, Newcastle upon Tyne NE2 4HH, UK. <https://orcid.org/0000-0002-3218-8250>

Csilla Krausz

Andrology Department, Fundació Puigvert, Universitat Autònoma de Barcelona, Instituto de Investigaciones Biomédicas Sant Pau, Barcelona, 08025 Catalonia, Spain.; Department of Experimental and Clinical Biomedical Sciences “Mario Serio”, University of Florence, University Hospital Careggi, Viale Pieraccini 6, Florence 50139, Italy.

Stéphane Viville

Laboratoire de Génétique Médicale LGM, Institut de génétique médicale d’Alsace IGMA, INSERM UMR 1112, Université de Strasbourg, France.; Laboratoire de Diagnostic Génétique, UF3472-génétique de l’infertilité, Hôpitaux Universitaires de Strasbourg, 67000 Strasbourg, France.

Donald Conrad

Division of Genetics, Oregon National Primate Research Center, Oregon Health & Science University, Portland, OR 97006, USA. <https://orcid.org/0000-0003-3828-8970>

Donal O’Carroll

Centre for Regenerative Medicine, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh, Edinburgh, UK.; Wellcome Centre for Cell Biology, School of Biological Sciences, The University of Edinburgh, UK. <https://orcid.org/0000-0002-8626-2217>

Frank Tüttelmann (✉ frank.tuettelmann@ukmuenster.de)


Institute of Reproductive Genetics, University of Münster, 48149 Münster, Germany.
<https://orcid.org/0000-0003-2745-9965>

Article

Keywords: piRNA biogenesis, transposon, azoospermia, male infertility

Posted Date: January 9th, 2024

DOI: <https://doi.org/10.21203/rs.3.rs-3710476/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: There is **NO** Competing Interest.

Abstract

Piwi-interacting RNAs (piRNAs) are crucial for transposon silencing, germ cell maturation, and fertility in male mice. Here, we report on the genetic landscape of piRNA dysfunction in humans and present 39 infertile men carrying biallelic variants in 14 different piRNA pathway genes, including *PIWIL1*, *GTSF1*, *GPAT2*, *MAEL*, *TDRD1*, and *DDX4* as novel disease genes. The testicular phenotypes repeatedly differ from those of the respective knockout mice and range from complete germ cell loss to the production of a few morphologically abnormal spermatozoa. LINE1 expression in spermatogonia links impaired piRNA biogenesis to transposon de-silencing and serves to classify variants as functionally relevant. Furthermore, an abolished expression of not only the encoded proteins but also of additional piRNA factors reveals co-dependencies within the human pathway. These results establish the disrupted piRNA pathway as a major cause of human spermatogenic failure and provide insights into transposon silencing in human male germ cells.

Introduction

PIWI-interacting RNAs (piRNAs) represent a specific type of regulatory, single-stranded small non-coding RNAs preferentially expressed in germ cells. They are required for transposon silencing, thus safeguarding genome integrity in the fetal and adult mammalian testis, and sculpting the post-meiotic transcriptome. In contrast to mice, in which disrupted piRNA biogenesis has been tightly linked to male-specific infertility, the role of the piRNA pathway in spermatogenic failure in men remains largely unexplored.

piRNAs bind to a subclade of Argonaute proteins known as PIWI proteins (derived from P-element-induced-wimpy testis)¹. Based on their temporal expression in mice, piRNAs are classified into three distinct main categories: fetal, pre-pachytene, and pachytene piRNAs². Fetal piRNAs, which are expressed in prospermatogonia, are loaded into both PIWIL2 and PIWIL4, while pre-pachytene piRNAs, expressed in postnatal spermatogonia, are exclusively associated with PIWIL2. Finally, pachytene piRNAs bound by PIWIL1 and PIWIL2 are first produced when germ cells enter the pachytene stage of meiosis and account for approximately 95% of piRNAs in the adult testis¹.

Pachytene piRNAs originate predominantly from non-transposable element (TE) intergenic loci and regulate gene expression by controlling transcriptional degradation or translational initiation at the late and post-meiotic stages of spermatogenesis^{3–6}. On the contrary, fetal and pre-pachytene piRNAs are enriched in TE-targeting sequences and essential for their post-transcriptional silencing through the piRNA-induced silencing complex (piRISC) mediated slicer activity^{7–9}. In addition, fetal piRNAs are required for *de novo* transposon methylation¹⁰.

piRNA biogenesis can be differentiated into two pathways involving not only PIWI proteins but also Tudor domain-containing proteins (TDRDs) acting as scaffolds, along with several enzymes involved in pre-piRNA trimming and maturation¹¹ (Supplementary Fig. 1a,b). The biogenesis of pachytene piRNAs is

restricted to the primary pathway, in which long piRNA precursors are transferred from the nucleus to the cytoplasm and accumulate in perinuclear structures called nuages. Here, mature piRNAs are produced through cleavage and processing of piRNA precursors. This cleavage is independent of piRISC activity and is initiated by the endonuclease PLD6, which establishes the 5'-ends of pre-piRNAs^{12,13}. In contrast, in the fetal testis, complementary long piRNA precursors are mainly cleaved by PIWIL2-bound piRISC complexes. Here, the massive amplification of TE-derived fetal piRNAs is established in the secondary pathway (ping-pong cycle)¹.

Knockout mice for more than twenty genes related to the piRNA-pathway have been analyzed¹¹. These mice are concordantly affected by male-specific infertility, small testes, and germ cell maturation arrest at meiosis or early haploid cell stages. Furthermore, a substantial reduction in the amount of piRNAs and consequent de-repression of TEs in the fetal and/or adult testis was observed in several mouse models^{14–17}.

In men, biallelic variants in several piRNA-related genes have recently been reported to cause infertility due to spermatogenic failure leading to non-obstructive azoospermia (NOA) or cryptozoospermia, i.e., no or very few spermatozoa in their ejaculate^{18–21}. However, only biallelic variants in *PNLDC1*, *FKBP6*, and *TDRD9* have as yet been functionally linked to a reduced levels of germ cell-derived piRNAs and, thus, impaired piRNA biogenesis^{21–23}.

Here, we shed light on the impact of disrupted piRNA biogenesis on human spermatogenesis by presenting 39 infertile men carrying rare, biallelic, putative pathogenic variants in 14 different genes encoding proteins of the piRNA pathway. Interestingly, the observed testicular phenotypes repeatedly differ from those of the respective knockout mice. Furthermore, we show that the dysfunction of piRNA pathway proteins in the human adult testis not only leads to a reduced amount of pachytene piRNAs but is also associated with a gene-specific increase of transposon expression in spermatogonia. These analyses can serve as readout for the functional relevance, i.e., pathogenicity, especially of the identified missense variants.

Results

Genes encoding piRNA biogenesis proteins are frequently mutated in infertile men

To elucidate protein networks or biological pathways contributing to impaired spermatogenesis, we queried exome/genome data of >2,000 infertile men from the Male Reproductive Genomics (MERGE) study²⁴ for rare homozygous loss-of-function (LoF) variants in genes preferentially expressed in the human testis. On the 61 identified genes, we performed a gene ontology-based two-tiered hierarchical clustering of significantly enriched biological processes that showed a striking enrichment of categories associate with “piRNA processing” (Fig. 1a, Supplementary Fig. 2). Further analysis revealed that piRNA

pathway genes also contributed to the most significantly enriched individual processes (Fig. 1b,). Next, we screened the MERGE data from 2,127 infertile men with azoo-, cryptozo-, extreme or severe oligozoospermia ($< 2/ < 10$ million total sperm count; Online methods) for biallelic, high-impact variants (minor allele frequency [MAF] in gnomAD < 0.01 ; LoF or missense variants with CADD score ≥ 15) in 24 human orthologues of murine genes associated with piRNA biogenesis (Fig. 1c; Supplementary Table 1) and identified 31 men carrying variants fulfilling the selection criteria (Table 1).

Table 1

Biallelic high-impact variants identified in genes of the piRNA pathway in infertile men due to azoo-, crypto, or extreme oligozoospermia.

ID	Gene	Variant (c.)	Variant (p.)	Phenotype (semen; histology, TESE outcome)
M928	<i>DDX4</i>	[1532C > T]; [1532C > T]	[(Ala511Val)]; [(Ala511Val)]	Crypto; RsA
M2546*	<i>FKBP6</i>	[508_529dup]; [832C > T]	[(Phe177Cysfs*20)]; [(Arg278*)]	Crypto; RsA
M2548*	<i>FKBP6</i>	[508_529dup]; [508_529dup]	[(Phe177Cysfs*20)]; [(Phe177Cysfs*20)]	Crypto; RsA
M1400*	<i>FKBP6</i>	[589-2A > G];[589- 2A > G]	[(?);(?)]	Crypto; RsA
MI-0042P	<i>GPAT2</i>	[146G > A];[146G > A]	[(Arg49His)]; [(Arg49His)]	ExtOligo
M2556	<i>GPAT2</i>	[1156-1G > A]; [1156-1G > A]	[(?);(?)]	Crypto; MeiA
M13	<i>GPAT2</i>	[1130A > G]; [1954C > T]	[(His377Arg)];[(Arg652*)]	Azoo; SCO
M454	<i>GPAT2</i>	[1130A > G]; [146G > A]	[(His377Arg)];[(Arg49His)]	Azoo; SCO
17-051	<i>GPAT2</i>	[1388C > T]; [1388C > T]	[(Thr463Met)]; [(Thr463Met)]	Azoo; SCO
15-0730	<i>GPAT2</i>	[1388C > T]; [1388C > T]	[(Thr463Met)]; [(Thr463Met)]	Azoo; SCO
M690	<i>GPAT2</i>	[1879C > T]; [1879C > T]	[(Arg627Trp)]; [(Arg627Trp)]	Azoo; MeiA
M1844	<i>GPAT2</i>	[1879C > T]; [1879C > T]	[(Arg627Trp)]; [(Arg627Trp)]	Azoo; SCO
M2043	<i>GTSF1</i>	[97C > A];[97C > A]	[(His33Asn)]; [(His33Asn)]	Azoo; MeiA
M2243	<i>GTSF1</i>	[221_222del]; [221_222del]	[(Arg74Lysfs*4)]; [(Arg74Lysfs*4)]	Azoo; MeiA
M3079	<i>HENMT1</i>	[400A > T];[400A > T]	[(Ile134Leu)]; [(Ile134Leu)]	Azoo; RsA

*already described²³; already described²¹; these are also shaded in grey; a, positive TESE; Azoo, azoospermia; MeiA, meiotic (spermatocyte) arrest; RsA, round spermatid arrest; ES+, elongated spermatids present in seminiferous tubules; SCO, Sertoli cell-only; SpgA, spermatogonia arrest; Crypto, cryptozoospermia; ExtOligo, extreme oligozoospermia.

ID	Gene	Variant (c.)	Variant (p.)	Phenotype (semen; histology, TESE outcome)
M2435	<i>MAEL</i>	[799C > T];[908 + 1G > C]	[(Arg267*)];[(?)]	Azoo, MeiA
TP17	<i>MOV10L1</i>	[2179 + 3A > G]; [2179 + 3A > G]	[(?)];[(?)]	Azoo; SpgA
M1948	<i>MOV10L1</i>	[2258T > C]; [2258T > C]	[(Val753Ala)]; [(Val753Ala)]	Azoo
TP24	<i>MOV10L1</i>	[3115G > A]; [3115G > A]	[(Glu1039Lys)]; [(Glu1039Lys)]	Azoo; SCO
MI_Proband02199	<i>MOV10L1</i>	[3268G > T]; [3268G > T]	[(Val1090Phe)]; [(Val1090Phe)]	Azoo; SCO
M2006	<i>PIWIL1</i>	[688C > T];[688C > T]	[(Arg230*)];[(Arg230*)]	Azoo; RsA
TP32	<i>PIWIL2</i>	[839A > C];[839A > C]	[(Tyr280Ser)]; [(Tyr280Ser)]	Azoo; SCO
M2949	<i>PIWIL2</i>	[1697G > A]; [1697G > A]	[(Arg566His)]; [(Arg566His)]	Azoo
M2173	<i>PLD6</i>	[1A > T];[1A > T]	[(Met1?)];[(Met1?)]	Azoo; SCO
M2803	<i>PLD6</i>	[469del];[469del]	[(His157Thrfs*102)]; [(His157Thrfs*102)]	Azoo; SCO
M3274	<i>PNLDC1</i>	[790G > T]; [790G > T]	[(Val264Leu)]; [(Val264Leu)]	Crypto
M1125	<i>PNLDC1</i>	[1058A > G]; [1058A > G]	[(Tyr353Cys)]; [(Tyr353Cys)]	Crypto; ES+
M1648	<i>TDRD1</i>	[887C > A];[887C > A]	[(Ser296Tyr)]; [(Ser296Tyr)]	Azoo; MeiA
M2842	<i>TDRD9</i>	[1243G > T]; [1243G > T]	[Val415Phe]; [Val415Phe]	ExtOligo
M800	<i>TDRD9</i>	[3148dup]; [3148dup]	[(Val1050Glyfs*49)]; [(Val1050Glyfs*49)]	ExtOligo, ES+; a
M3007	<i>TDRD9</i>	[3716 + 3A > G]; [c.3716 + 3A > G]	[(?)];[(?)]	ExtOligo

*already described²³; already described²¹; these are also shaded in grey; a, positive TESE; Azoo, azoospermia; MeiA, meiotic (spermatocyte) arrest; RsA, round spermatid arrest; ES+, elongated spermatids present in seminiferous tubules; SCO, Sertoli cell-only; SpgA, spermatogonia arrest; Crypto, cryptozoospermia; ExtOligo, extreme oligozoospermia.

ID	Gene	Variant (c.)	Variant (p.)	Phenotype (semen; histology, TESE outcome)
M2442	<i>TDRD9</i>	[3826G > T]; [3826G > T]	[(Val1276Phe)]; [(Val1276Phe)]	Crypto
M2662	<i>TDRD12</i>	[287A > C];[287A > C]	[(Asp96Ala)];[(Asp96Ala)]	Azoo; SCO
M1642	<i>TDRD12</i>	[593A > G];[593A > G]	[(Asn198Ser)];[(Asn198Ser)]	Azoo; SCO
TP5	<i>TDRD12</i>	[963 + 1G > T]; [963 + 1G > T]	[(?)];[(?)]	Azoo; MeiA
M2227	<i>TDRD12</i>	[986G > A];[986G > A]	[(Trp329*)];[(Trp329*)]	Azoo; RsA
M2940	<i>TDRD12</i>	[2419C > T]; [2419C > T]	[(Arg807Cys)]; [(Arg807Cys)]	Crypto
M2317	<i>TDRD12</i>	[2432G > A]; [2432G > A]	[(Arg811Gln)]; [(Arg811Gln)]	Azoo; ES+
M2595	<i>TDRD12</i>	[3157del]; [3157del]	[(Leu1053Phefs*4)]; [(Leu1053Phefs*4)]	Azoo; ES+
*already described ²³ ; already described ²¹ ; these are also shaded in grey; a, positive TESE; Azoo, azoospermia; MeiA, meiotic (spermatocyte) arrest; RsA, round spermatid arrest; ES+, elongated spermatids present in seminiferous tubules; SCO, Sertoli cell-only; SpgA, spermatogonia arrest; Crypto, cryptozoospermia; ExtOligo, extreme oligozoospermia.				

Of these, 27 patients carried homozygous variants (11 LoF and 16 missense with CADD score ≥ 20) and four patients carried confirmed compound heterozygous variants. In total, these affected 14 different genes: *DDX4*, *FKBP6*, *GPAT2*, *GTSF1*, *HENMT1*, *PIWIL1*, *PIWIL2*, *PLD6*, *PNLDC1*, *MAEL*, *MOV10L1*, *TDRD1*, *TDRD9*, *TDRD12* (Table 1, Fig. 1c, Extended data Table 1, also including reference transcripts). The three *FKBP6* variant carriers and also one of the *TDRD12* variant carriers have been described previously^{21,23}. Detailed analysis of the exomes from each variant carrier did not reveal any other variants with a higher probability for causing the disease. In two cases, chromosomal translocations were identified (Extended data Table 2) and it cannot be excluded that they at least partially contribute to the patient's phenotype. Of the 31 affected men, 19 were azoospermic, nine were cryptozoospermic, and four had extreme oligozoospermia. Twenty-three patients underwent a testicular biopsy with the aim of sperm extraction (TESE), which was negative in 22 men, i.e., no spermatozoa could be obtained. The analyses of testicular tissue revealed phenotypes ranging from complete absence of germ cells (Sertoli cell-only, SCO, N = 7), the presence of spermatocytes (meiotic arrest, MeiA, N = 6), round spermatids (RsA, N = 7) or elongated spermatids (ES+, N = 3) as the most advanced germ cells (Supplementary Fig. 3, Extended data Table 2).

In addition, screening of exome data from three independent cohorts of infertile men identified eight further patients with homozygous high-impact variants (two LoF and six missense) in *GPAT2*, *PIWIL2*,

MOV10L1, and *TDRD12* (Table 1, Extended data Table 1), bringing the total number of variant carriers to 39. In six of the additional patients, the testicular phenotypes of SCO or spermatogenic arrest confirmed the clinically suspected non-obstructive azoospermia (Extended data Table 2). In summary, 38 distinct variants in piRNA genes were identified among 39 infertile men. Of these variants, 12 were absent from gnomAD (version v2.1.1) and 18 were extremely rare ($MAF \leq 0.0001$) (Extended data Table 1).

Variants in genes encoding components of the piRISC complex

Pachytene piRNAs have been proposed to direct PIWIL1 and PIWIL2 to cleave specific mRNAs and thereby regulate gene expression²⁵. The slicing activity of this piRNA-induced silencing complex (piRISC) requires GTSF1 as an auxiliary factor²⁶. We identified four azoospermic men with biallelic variants in genes encoding proteins essential for piRISC activity (Table 1). In *PIWIL1*, a homozygous stop-gain variant c.688C > T p.(Arg230*) localizing within the PAZ domain (Fig. 2a, Extended data Fig. 1a) was identified. The variant carrier exhibited CREM-positive, haploid, round spermatids as the most advanced germ cells in the seminiferous tubules (Fig. 2b) and TESE was negative (Extended data Table 2). Further staining demonstrated the absence of PIWIL1 in testicular germ cells, which is expressed in spermatocytes and haploid germ cells in a control with normal spermatogenesis (Fig. 2c).

Furthermore, two azoospermic men carried two different homozygous missense variants, in *PIWIL2*, c.1697G > A p.(Arg566His) and c.839A > C p.(Tyr280Ser), affecting amino acids conserved up to zebrafish (Extended data Fig. 1b). Arg566 is predicted to be a surface-accessible residue found within the linker 2 (L2) domain of PIWIL2²⁷ which bridges the PAZ and MID domains (Fig. 2a). Tyr280 is located within the structured N-terminal region of PIWIL2, which has been suggested to stabilize piRNA-target duplex conformations²⁸. Finally, two men with meiotic arrest were carriers of homozygous, high-impact variants in *GTSF1* (Fig. 2a, Extended data Fig. 1c). The frameshift variant c.221_222del p.(Arg74Lysfs*4) was predicted to result in nonsense-mediated decay (NMD), leading to abolished GTSF1 expression in the patient's spermatocytes (Fig. 2c). The missense variant c.97C > A p.(His33Asn) is located in a predicted α -helical protein domain and affects the second histidine of the highly conserved GTSF1 first zinc finger motif (Extended data Fig. 1c).

Variants in genes involved in piRNA metabolic processes

In the primary mammalian piRNA pathway, RNA helicase *MOV10L1* selectively binds to cytoplasmic piRNA precursor transcripts²⁹ and feeds them to the mitochondrial-associated endonuclease PLD6, which catalyzes the first cleavage step of piRNA processing³⁰.

We identified four azoospermic men with homozygous variants in *MOV10L1* [c.2258T > C p.(Val753Ala); c.3115G > A p.(Glu1039Lys); c.3268G > T p.(Val1090Phe); c.2179 + 3A > G p.?] (Table 1, Fig. 3a), of which two shared a testicular phenotype of SCO, one had sparse spermatogonia, and the last did not undergo a biopsy. All three missense variants affect highly conserved amino acids located in α -helical core protein domains, as predicted by AlphaFold2 (Extended data Fig. 2b). The splice region variant c.2179 + 3A > G

resulted in skipping of *MOV10L1* exon 16 (Extended data Fig. 2c) and inclusion of a premature stop codon. Two homozygous LoF variants, c.469del p.(His157Thrfs*102) and c.1A > T p.(Met1?), were identified in *PLD6* in two men with azoospermia due to SCO (Fig. 3, Extended data Fig. 2d,e).

PLD6 activity requires a GPAT protein to cleave pre-piRNAs³⁰ and in mammals, the mitochondrial-associated protein GPAT2 is crucial for primary piRNA biogenesis³¹. Biallelic variants in *GPAT2* were detected in eight men with negative TESE attempt (Extended data Table 2; Fig. 3a; Extended data Fig. 3). Five of these variant carriers share a severe testicular SCO phenotype, two had meiotic arrest, while only one presented with hypospermatogenesis leading to extreme oligozoospermia (Fig. 3b). The homozygous missense variant c.1879C > T p.(Arg627Trp) was identified in the unrelated patients M690 and M1844, and both parents of M1844 carried this variant in the heterozygous state (Extended data Fig. 3a). Patients 17-051 and 15-0730 were carriers of c.1388C > T p.(Thr463Met) (Extended data Fig. 3b). Both originate from Morocco, and a Somalier analysis indicated that they are distantly related. M13 and M454 were compound heterozygous for the missense variant c.1130A > G p.(His377Arg), which is located in the protein's GPAT/DHAPAT acetyltransferase domain and the stop-gain variant c.1954C > T p.(Arg652*) or the missense variant c.146G > A p.(Arg49His), respectively (Extended data Fig. 3c,d). In M13, p.(His377Arg) was inherited from the mother, who was not carrier of the second variant p.(Arg652*) and p.(Arg49His) was also identified in MI-0042P in the homozygous state. Finally, the splice acceptor variant c.1156-1G > A (M2556) in *GPAT2* resulted in skipping of exon 12, as confirmed by a minigene assay (Extended data Fig. 4a). Of note, GPAT2 expression was absent or strikingly reduced in the spermatocytes of the two *GPAT2* variant carriers analyzed (Fig. 3c).

Furthermore, we identified three patients with homozygous missense variants in *PNLDC1* (M3274, M1125) and *HENMT1* (M3079) (Fig. 3a) which both play a crucial role in piRNA maturation^{32,33}. The two *PNLDC1* variant carriers exhibited cryptozoospermia and, fittingly, the testicular biopsy of M112 revealed only a few tubules with elongated spermatids while TESE was negative. (Extended data Table 2). The missense variants impact two highly conserved amino acid residues, both situated in the PNLDC1 CAF domain (Extended data Fig. 4b). The homozygous variant c.400A > T p.(Ile134Leu) in *HENMT1* was identified in an azoospermic men with round spermatid arrest (Fig. 3b). The affected tyrosine residue is located in the protein's methyl-transferase domain and is conserved up to zebrafish (Extended data Fig. 4c).

Finally, we also identified biallelic variants in genes that are limited to secondary biogenesis or post piRNA maturation processes. Among these, *DDX4* encodes a germ cell-specific RNA helicase required for ribonucleoprotein remodeling during the loading of secondary piRNA intermediates onto PIWIL4¹⁷. In *DDX4*, the homozygous missense variant c.1532C > T p.(Ala511Val) was identified in an infertile men with cryptozoospermia due to predominant round spermatid arrest in the testicular tissue (M928, Fig. 3b). Alanine 511 is present in orthologous proteins up to fruit fly and is located in a highly conserved core protein region between the two predicted helicase domains (Extended data Fig. 5a). However, the cellular expression profile of DDX4 in the patient's testicular tissue remained unchanged (Supplementary Fig. 4a).

The piRNA pathway component MAEL localizes to the cytoplasm and shuttles to the nucleus in round spermatids³⁴. It may also facilitate nucleo-cytoplasmic trafficking of PIWIL4–piRNA complexes³⁵ and pachytene piRNA intermediate processing³⁶. M2435 carried the confirmed compound heterozygous stop-gain variant c.799C > T p.(Arg267*) and the splice site variant c.908 + 1G > C p.? in *MAEL*, which was shown to cause skipping of exon 9 (Extended data Fig. 5b), resulting in an in-frame deletion of 21 amino acids. No spermatozoa could be retrieved from the testicular biopsy (Extended data Table 2) showing pachytene spermatocytes and few CREM-positive haploid germ cells (Supplementary Fig. 3) in single tubules, indicating a spermatogenic arrest at stages downstream of pachytene or after completion of meiosis. Assumed degradation of mutant *MAEL* transcripts by NMD was supported by absence of MAEL-specific staining in spermatocytes and round spermatids in patient testicular sections (Fig. 3c).

Variants in the scaffold proteins encoded by the TDRD gene family

Tudor domain (TD)-containing proteins (TDRDs) play a crucial role as molecular scaffolds in piRNA biogenesis³⁷ and, in mice, several members of the TDRD gene family have been linked to piRNA biogenesis. We identified rare homozygous variants in *TDRD1*, *TDRD9*, and *TDRD12* that matched our filtering criteria.

The missense variant c.887C > A p.(Ser296Tyr) in *TDRD1* in patient (M1648) with meiotic arrest affects a highly conserved serine residue located in the first TD (Fig. 4A; Extended data Fig. 6a). In testicular tissue with complete spermatogenesis, TDRD1 is expressed in perinuclear structures within round spermatids (Supplementary Fig. 4b). Because the seminiferous tubules of M1648 lack haploid germ cells, it remains unknown whether p.(Ser296Tyr) has any effect on the expression or stability of TDRD1.

In *TDRD9*, we identified four infertile men with homozygous nucleotide substitutions predicted to affect the protein sequence: two homozygous LoF variants, c.3148dup p.(Val1050Glyfs*49) and c.3716 + 3A > G p.?, which causes skipping of *TDRD9* exon 32 (Extended data Fig. 6c), resulting in a frameshift and two missense variants, p.(Val415Phe) and p.(Val1276Phe) (Table 1, Fig. 4a). The affected valine 415 is located in the helicase domain of TDRD9, whereas valine 1276 is located in a C-terminal protein region (Fig. 4a) and both amino acids are conserved in orthologous proteins (Extended data Fig. 6b). Interestingly, haploid spermatozoa with impaired motility and abnormal morphology were observed in all four *TDRD9* variant carriers (Table 1, Extended data Table 2).

Finally, in *TDRD12*, seven men with homozygous high-impact variants (Table 1, Fig. 4a) were identified out of whom five had a negative TESE attempt. Two patients with SCO were carriers of homozygous missense variants c.287A > C p.(Asp96Ala) and c.593A > G p.(Asn198Ser), respectively, co-segregating with the disease in the families (Extended data Fig. 7a,b). TP5 with meiotic arrest carried the homozygous splice site variant c.963 + 1G > T p.?, leading to skipping of exon 9 (Extended data Fig. 7c). As a result the open reading frame is shifted resulting in the synthesis of a truncated protein, p.(Asp289Alafs*3) if the mutant transcript is not degraded by NMD.

Interestingly, in the four other *TDRD12* variant carriers, haploid germ cells (round or elongated spermatids) were detected in the seminiferous tubules (Fig. 4b, Supplementary Fig. 3) or spermatozoa were found in the ejaculate. Two of these patients were carriers of homozygous LoF variants, the stop gain variant c.986G > A p.(Trp329*) and the frameshift variant c.3157del p.(Leu1053Phefs*4) (Extended data Fig. 7d,e). The other two *TDRD12* variants result in the substitution of highly conserved arginine residues Arg807 and Arg811 (Extended data Fig. 7f) which are found on the surface of a conserved and globular structured region of TDRD12 with unknown molecular function.

Impact of identified variants on expression of piRNA pathway components

Furthermore, we aimed to explore whether the loss-of-function of one piRNA biogenesis related protein might influence the expression profile of further proteins involved in this metabolic process. To this end, we performed immunohistochemical staining for several key piRNA pathway-related proteins (Fig. 5a,b, Extended data Fig. 8, Supplementary Fig. 4–7). Indeed, a diminished expression of PIWIL1, PLD6, GPAT2, and HENMT1 in spermatocytes was observed in several variant carriers. TDRD1-specific staining was absent only in round spermatids of the PIWIL1 stop-gain variant carrier M2006, and characteristic and distinct staining of MAEL-positive structures was more diffuse in spermatocytes of the *TDRD1* and *GPAT2* variant carriers. Of note, the staining pattern of DDX4 and GTSF1 was not affected in any of the variant carriers analyzed. Collectively, these observations indicate a gene-/protein-specific impact of several of the piRNA biogenesis proteins on the expression of additional piRNA factors.

Impact of identified variants on piRNA processing and transposon silencing

In the patients with available snap frozen testicular tissue and who were not affected by total germ cell loss [M2006: PIWIL1 p.(Arg230*), M1648: TDRD1 p.(Ser296Tyr), M2595: TDRD12 p.(Leu1053Phefs*4), M2317, TDRD12 p.(Arg811Gln)], we analyzed the impact on piRNA biogenesis in germ cells and performed small-RNA sequencing. The mapped piRNA sequences were intersected with known pachytene piRNA loci detected in the human adult testis. This revealed significantly decreased amounts of piRNAs in all four patients, compared with tissue with complete spermatogenesis (Fig. 5c). Notably, the peak of piRNAs, with a length of 28–31 bases seen in the control tissue was absent in all four samples. Accordingly, approximately 50% of all piRNAs in the mutant samples had a length between 25 and 27 bases, which was the case for only 13% of piRNAs in the control tissue (Supplementary Fig. 8a).

In mice, disruption of piRNA biogenesis leads to upregulation of transposons. We, therefore, investigated the silencing of transposons in human male germ cells and performed immunohistochemical staining for LINE1 open reading frame 1 protein (LINE1 ORFp1) in the testicular sections of variant carriers. Using a monoclonal, validated antibody directed against human LINE1 ORFp1, no staining was detected in germ cells of human control samples with complete spermatogenesis. In contrast, a concordant and specific expression of LINE1 ORFp1 in spermatogonia (Fig. 5d) was observed in three *TDRD12*, two *GPAT2*, three *TDRD12*, and single *MAEL* and *HENMT1* variant carriers, while in all other cases, including carriers of

homozygous LoF variants in *PIWIL1* and *GTSF1*, no LINE1 ORFp1 was expressed. (Fig. 5d, Supplementary Fig. 8b).

In summary, in 13 variant carriers (9 biallelic LoF, 4 homozygous missense) functional data on piRNA factor expression, pachytene piRNA level and/or TE expression demonstrated the pathogenicity of the respective variants.

Discussion

The number of identified monogenic causes of male infertility due to impaired spermatogenesis is steadily increasing, and a striking number of described disease genes encode proteins with vital roles in meiosis³⁸. Through comprehensive exploration of biallelic variants in exome/genome data of > 2,000 infertile men, we provide evidence that beyond meiosis-related genes, genes encoding proteins involved in piRNA biogenesis are a major, previously underexplored contributor to human spermatogenic failure.

Biallelic LoF variants in 14 piRNA genes, including *PIWIL1*, *GTSF1*, *PLD6*, *GPAT2*, and *MAEL*, were identified in infertile men for the first time, establishing them as novel disease genes. Furthermore, no homozygous potentially pathogenic missense variants had yet been reported for *TDRD1* and *DDX4*, both of which are highly intolerant to genetic variation. Of note, the presented homozygous stop-gain variant in *PIWIL1* resolves the controversy regarding the causal link between heterozygous human *PIWIL1* variants and azoospermia^{39,40} and renders this another autosomal recessive disease gene.

Among the genes highlighted in this study, *FKBP6*²³, *PIWIL2*^{20,41}, *PNLDC1*^{21,22}, *PLD6*²¹, *HENMT1*¹⁸, *MOV10L1*⁴², *TDRD9*^{18,19,21}, and *TDRD12*²¹ were recently described in the context of piRNA pathway dysfunction and/or human male infertility and the discovery of additional variants significantly strengthens the gene-disease relationship. For several of these genes, we identified similar testicular phenotypes as previously reported: SCO in *PIWIL2*⁴¹, round spermatid arrest in *PNLDC1*²², and hypospermatogenesis in *TDRD9* variant carriers¹⁸, respectively. Interestingly, also the gene-specific patients' phenotypes observed in this study were largely consistent and severity was independent from the type of variant, indicating that the identified missense variants are indeed also loss-of-function variants on the protein level (Fig. 6a). Thus, the gene-specific testicular phenotype can be used to aid assessment of the variant's pathogenicity. In contrast, *TDRD12* variant carriers exhibited highly variable phenotypes, ranging from SCO to even a few spermatozoa in the ejaculate. In summary, with at least four different biallelic variants (including several LoF variants) identified per gene in this and other studies, *GPAT2*, *PNLDC1*, *TDRD12*, *MOV10L1*, *PLD6*, *FKBP6*, and *TDRD9* are excellent candidates to be included in the diagnostic workup of infertile men.

This study also allows a comprehensive comparison of reproductive phenotypes per gene between mice and men, which reveals only a partial gene-specific overlap (Fig. 6b). The phenotypic spectrum in humans is significantly broader, ranging from complete absence of spermatozoa to hypospermatogenesis resulting in severe oligozoospermia, compared with mice, in which round spermatid or meiotic arrest have

predominantly been described. More specifically, the round spermatid arrest observed in *Miwi* (*Piwi1*) knockout mice⁴³ was also present in the *PIWIL1* variant carrier. In contrast, SCO was the predominant testicular phenotype in human variant carriers in *PIWIL2*, *PLD6*, and *GPAT2*, whereas the corresponding knockout mice exhibited meiotic arrest^{12,14,44}. Only for the aged *PIWIL2* knockout mice, 50% of tubuli revealed loss of germ cells⁴⁵. Of note, the more severe phenotype observed in the human variant carriers was not associated with patient's age. In addition, carriers of *FKBP6* or *TDRD9* presented with round spermatid arrest or even hypospermatogenesis resulting in severe oligozoospermia, while the respective knockout mice showed meiotic arrest^{46,47}. Thus, in some instances, human spermatogenesis seems to be less stringently controlled and progresses further despite disrupted piRNA biogenesis. It remains to be determined whether the round spermatids/spermatozoa produced in some men are actually suitable for procreation.

For several of the identified variants, we demonstrate functional data linking impaired function or absence of the encoded protein to downstream cellular effects. The observed diminished expression of further piRNA factors in variant carriers was also described for *PNLDC1*²², where it has been linked to decreased expression of *MYBL1*, a testis-specific transcription factor known to regulate expression of pachytene piRNAs as well as several piRNA factor genes in mice.

In some piRNA factor knockout mice, impaired piRNA biogenesis resulted in de-repression of TEs in spermatocytes of the adult testis⁹. Surprisingly, we observed a spermatogonia-specific de-repression of LINE transposons in *GPAT2*, *TDRD12*, *FKBP6*, *HENMT1*, and *MAEL* variant carriers. On the contrary, homozygous LoF variant carriers in *PIWIL1* and *GTSF1* lacking the encoded protein, did not demonstrate TE de-repression. Thus, de-repression of transposons does not seem to be a general consequence of impaired pachytene piRNA biogenesis. This is supported by the fact that this subtype of piRNAs, which comprises over 95% of all piRNAs in the adult mouse testis¹, is depleted of transposon sequences⁴⁸ and has predominantly been linked to spermatogenic gene expression^{5,49}. Further, *Miwi* (*PIWIL1*) expression is restricted to the adult testis⁴³ and, therefore, supposed to be essential only for pachytene piRNA-mediated functions⁵. However, there are some human piRNA genes such as *GPAT2*, *FKBP6*, *TDRD12* that, when impaired, concordantly result in LINE1 de-repression in spermatogonia and the encoded proteins might also be involved in biogenesis of pre-pachytene piRNAs. These piRNAs are bound predominantly by the slicer PIWIL2, which is expressed at all stages of male germ cell maturation, and are enriched in TEs⁸ and 3'-UTR sequences⁵⁰ and are likely involved in the control of harmful transposon expression.

In conclusion, this study provides extensive data linking disrupted piRNA biogenesis to human spermatogenic failure, demonstrates that piRNA pathway genes are a major target for scrutinizing genetic causes of male infertility, and suggests that safeguarding of the genome during spermatogenesis is less stringent in men than in mice. The detailed characterization of pathogenic human variants provides insight into the molecular function of the factors involved in piRNA biogenesis and piRNA-mediated transposon silencing. This opens the possibility to investigate key protein domains and, in parallel, to assess the pathogenicity of gene variants.

Methods

Study cohorts

Four cohorts of whole exome (WES) or whole genome sequencing data (WGS) of infertile men were included in this study. The MERGE cohort includes data of 2,412 men (2,352 with WES and 60 with WGS) with various infertility phenotypes and >90% of these men were recruited at the Centre of Reproductive Medicine and Andrology (CeRA), Münster. Most men of this cohort are azoospermic (N = 1,448) or have severely reduced sperm counts: N = 454 with cryptozoospermia (sperm only identified after centrifugation of the ejaculate); N = 158 with extreme oligozoospermia (sperm count <2 million); N = 67 with severe oligozoospermic (sperm count <10 million). Numerical chromosomal aberrations such as Klinefelter syndrome (karyotype 47,XXY) and Y-chromosomal AZF-deletions are exclusion criteria. Likely pathogenic monogenic causes for the infertile phenotype were already described in about 8% of cases²⁴. Exome sequencing and bioinformatics analyses in the MERGE cohort were performed as previously described²⁴.

The Strasbourg cohort comprises 23 men diagnosed with NOA. The Barcelona cohort (BCN) comprises 235 NOA men attending the Fundació Puigvert (Barcelona)⁵¹. The Nijmegen/Newcastle cohort includes exome sequencing data of 266 infertile men, 225 affected by azoospermia and 41 by cryptozoospermia⁵².

Ethical approval

All persons gave written consent compliant with local requirements. The study protocol was approved by the local ethics committees: MERGE Münster (2010-578-f-S) and Giessen (26/11); Strasbourg (CPP 09/40 –WAC-2008-438 1W DC-2009-I 002), and Yeni Yüzyıl University, Scientific, social and noninterventional health sciences research ethics committee, Istanbul, Turkey (approval no: 2019/08); Barcelona: (2014/04c); Newcastle: (Newcastle:REC ref. 18/NE/0089), Nijmegen: (NL50495.091.14 version 5.0). All experiments were performed in accordance with relevant guidelines and regulations.

Whole exome and genome sequencing

Genomic DNA was extracted from peripheral blood leukocytes via standard methods. For exome sequencing of the MERGE and Strasbourg cohort, the samples were prepared and enrichment was carried out according to the protocol of either Agilent's SureSelectQXT Target Enrichment for Illumina Multiplexed Sequencing Featuring Transposase-Based Library Prep Technology (Agilent) or Twist Bioscience's Twist Human Core Exome. To capture libraries, Agilent's SureSelect Human All Exon Kits V4, V5 and V6 or Twist Bioscience's Human Core Exome plus RefSeq spike-in and Exome 2.0 plus comprehensive spike-in were used. For whole genome sequencing of samples from the MERGE cohort sequencing libraries were prepared according to Illumina's DNA PCR-Free library kit. For multiplexed sequencing, the libraries were index tagged using appropriate pairs of index primers. Quantity and quality of the libraries were

determined with the ThermoFisher Qubit, the Agilent TapeStation 2200 and Tecan Infinite 200 Pro Microplate reader, respectively. Sequencing was performed on the Illumina HiSeq 4000 System, the Illumina HiSeqX System, the Illumina NextSeq 500 System, the Illumina NextSeq 550 System, or the NovaSeq 6000 System, using the HiSeq 3000/4000 SBS Kit (300 cycles), the HiSeq X Ten Reagent Kit (300 cycles), the NextSeq 500 V2 High-Output Kit (300 cycles), or the NovaSeq 6000 S1 and S2 Reagent kits v1.5 (200 cycles), respectively. For the BCN cohort exome sequencing was carried out as a service by MacroGen Inc. 122 (Republic of Korea) using the 123 Agilent SureSelect_V6 enrichment and a NovaSeq 6000. Exome sequencing in the Nijmegen/Newcastle cohort was performed as previously described⁵².

Variant calling

After trimming of remaining adapter sequences and primers with Cutadapt v1.15⁵³, reads were aligned against GRCh37.p13 using BWA Mem v0.7.17⁵⁴. Base quality recalibration and variant calling were performed using the GATK toolkit v3.8⁵⁵ with haplotype caller according to the best practice recommendations. For more recent samples and whole genome samples Illumina Dragen Bio-IT platform v4.2 was used for alignment and variant calling. Both pipelines use GRCh37.7.p13 as reference genome. Resulting variants were annotated with Ensembl Variant Effect Predictor⁵⁶.

Gene Ontology analysis

WES data of infertile men from MERGE was first filtered for genes with rare (MAF£0.01 according to the gnomAD, v2.1.1) homozygous loss-of-function (LoF) variants (stop-gain, start-loss, splice acceptor, splice donor, frameshift). We then selected for genes preferentially expressed in human male germ cells according to single cell RNAseq data included in the human protein atlas (HPA)⁵⁷. Gene ontology analysis (<http://geneontology.org>)^{58–60} was performed on this gene list (for “biological processes” and “homo sapiens”) and processed with PANTHER <https://pantherdb.org/webservices/go/overrep.jsp>^{60,61} (annotation dataset: “GO biological processes complete”, test type: ‘Fisher’s Exact’, Correction: “Bonferroni”, showing results with $P < 0.05$). GO terms were then processed with Revigo⁶² (<http://revigo.irb.hr/>) using the P-value and a medium (0.7) list setting (yes to removal of obsolete GO terms, species “homo sapiens”, “SimRel” semantic similarity measure). The Revigo Table was exported and $-\log_{10}(P\text{-value})$ of representative GO terms (classed as representation: ‘null’) plotted as side-ways bar chart. Revigo tree data was processed with CirGO.py⁶³ for visualization of the 2-tiered hierarchy of GO-terms.

Screening of WES data for biallelic high impact variants

To identify potentially harmful gene variants in genes of the piRNA pathway WES data of infertile men from all cohorts included in this study were screened to identify individuals with biallelic high-impact

variants (stop-gain, start-loss, splice site and splice region, deletions and insertions as well as missense variants with CADD ≥ 15) in a total of 24 different genes of the pathway (Table S1). Only variants with a MAF ≥ 0.01 (gnomAD database, v 2.1.1) were taken into account.

To exclude presence of additional possibly pathogenic variants WES data were additionally screened for additional rare homozygous high-impact variants (LoF and missense variants with CADD ≥ 20) occurring in a list of 21 azoospermia-associated genes with at least moderate clinical validity²⁴ and 363 candidate genes associated with the Gene Ontology classification “male infertility” in the Mouse Genome Informatics Database (MGI) revealing strong expression in human male germ cells. Patients in which additional candidate variants were identified were excluded from further analysis.

Further genetic analysis

Validation of prioritized variants as well as co-segregation analyses were performed by Sanger sequencing. The regions of interest were amplified from patients genomic DNA as well as available family members with the use of primers and conditions as listed in Table S2. The PCR products were then purified and sequenced using standard protocols. For validation of variants in *GPAT2* long range PCR products using *GPAT2* specific primers that do not bind to pseudogenes *GPAT2P1* and *GPAT2P2* were amplified and used as template for subsequent nested PCR and Sanger sequencing. If a variant was found in more than one individual in MERGE (*GPAT2*: c.1879C>T in M690 and M1844 and c.1130A>G in M13 and M454) the relationship between the two mutation carriers was determined using the Somatic tool⁶⁴. In case that no parental DNA was available for analysis, biallelic occurrence of heterozygous variants was determined by long-read sequencing using long-range PCR products encompassing both genomic regions of interest amplified from variant carriers as template for library generation.

NGS library preparation and long read sequencing using the MinION system

To determine if two heterozygous variants identified in one gene of the same patient occur in cis or in trans, a long read sequencing approach was used. To this end, a long-range PCR product encompassing both variants was amplified (see Supplemental Table 2 for primer information) from patients genomic DNA using the TAKARA LA Taq® DNA Polymerase Hot-Start Version. 1 μ g of PCR products was used for subsequent preparation of MinION sequencing library. Barcoding and sequencing was carried out according to manufacturer's instructions (MinION, Oxford Nanopore Technologies). After demultiplexing of obtained reads, alignment to human reference hg19, quality control and variant calling phasing of variants on same/different alleles was determined.

Minigene assay

To determine the functional impact of splice site and splice region variants, an *in vitro* splicing assay based on a minigene construct was performed. The region of interest was amplified from genomic DNA of the respective patient as well as of a human control sample by standard PCR procedures. To analyse splice effect of variants *GPAT2* c.1156-1G>A, *MAEL* c.908+1G>C and *TDRD9* c.3716+3A>G products were cloned into pENTR™/D-TOPO® according to manufacturer's instructions. The subsequent gateway cloning was performed using Gateway™ LR Clonase™ Enzyme Mix and pDESTsplice as destination vector (pDESTsplice was a gift from Stefan Stamm (Addgene plasmid #32484)⁶⁵. To analyse *TDRD12* c.963+1G>T amplified region encompassing exon 8-10 of *TDRD12* was subcloned into pcDNA3.1 and for *MOV10L1* c.2179+3A>G into pSPL3B. A transient transfection with mutant and wildtype Minigene constructs was performed using Human Embryonic Kidney cells (HEK293T Lenti-X, Clontech Laboratories, Inc.®). Total RNA was extracted using the RNeasy Plus Mini Kit (QIAGEN®) and reverse-transcribed into cDNA with the ProtoScript® II First Strand cDNA Synthesis Kit (New England Biolabs GmbH®). Amplification of the region of interest was performed and RT-PCR products were separated on a 2% agarose gel, cut out, extracted and sequenced.

AlphaFold2 protein structure

AlphaFold2 structure predictions were obtained from EBI (<https://alphafold.ebi.ac.uk/>), except for PNLDC1 and GPAT2, which were generated with the AlphaFold2 google colab (<https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb>)^{66,67} using protein sequences encoded by the NCBI accession NM_001271862.2 for PNLDC1 and NM_001321526.1 for GPAT2. Images of protein structures were generated with Pymol (v.2.5.4, Schrödinger, LLC).

Histology and Immunohistochemical staining

Testis biopsies of patients from the MERGE cohort and control subjects were obtained from testicular sperm extraction (TESE) approaches or histological evaluation at the Department of Clinical and Surgical Andrology (University Hospital Münster, Germany) and research purposes. Biopsies were fixed in Bouin's solution overnight, washed with 70% ethanol and embedded in paraffin for routine histological evaluation. Subsequently, 5 µm sections were stained with periodic acid-Schiff (PAS) according to previously published protocols⁶⁸. Testis biopsies of patients from Giessen were processed equally but stained with hematoxylin and eosin (HE) following previously published protocols⁶⁹. Testis biopsies of patients from BCN were treated as described previously⁵¹.

For immunohistochemical analyses, 3 µm sections of testicular tissue were de-paraffinized and rehydrated as described previously⁷⁰. After rinsing with tap water (15 min, RT) heat-induced antigen retrieval was performed in HIER buffer (pH 6) or as indicated in table S3. This step was followed by cooling and washing with 1X Tris-buffered saline (TBS) before endogenous peroxidase activity was

blocked using 3% hydrogen peroxide (15 min, RT). Blocking was performed by adding 25% goat serum (#ab7481, abcam) in TBS containing 0.5% bovine serum albumin (BSA, #A9647, Merck, 30 min, RT). Sections were incubated overnight at 4°C in primary antibody solution, including 5% BSA/TBS and primary antibody as indicated in table S3. The following day, sections were washed with 1x TBS and incubated with a corresponding biotinylated secondary antibody in 5% BSA/TBS for 1h. After washing with TBS, sections were incubated with streptavidin-horseradish peroxidase (Merck – 1:500, 45 min, RT) diluted in 5% BSA/TBS. Subsequently, sections were washed with TBS and incubated with 3,3'-Diaminobenzidine tetrahydrochloride (DAB, Merck) for visualisation of antibody binding. Staining was validated by microscopical acquisition and stopped with aqua bidest. Counterstaining was conducted using Mayer's hematoxylin (Merck). Finally, sections were rinsed with tap water, dehydrated with increasing ethanol concentrations and mounted using M-GLAS® mounting medium (Merck).

Slides were evaluated and documented using a PreciPoint 08 Scanning Microscope, Olympus BX61VS Virtual Slide System Axioskop (Zeiss, Oberkochen, Germany), or an Olympus BX61 microscope with an attached Retiga 400 R camera (Olympus, Melville, NY, USA) and integrated CellSens imaging software (Olympus, Melville, NY, USA).

RNA extraction and small RNA sequencing

RNA from snap-frozen testicular tissues of three controls with full spermatogenesis and infertile men with biallelic variants in PIWIL1 (M2006), TDRD1 (M1648), TDRD12 (M2317, M2595) and FKBP6 (M2546, M2548) was extracted using Direct-zol RNA Microprep kit (Zymo Research, Cat. #R2062). The quantity and quality of the isolated RNA were assessed with Qubit RNA High Sensitivity kit (Invitrogen, Cat. #Q32852) and Agilent RNA Nano kit (Agilent, Cat. #55067-1512), respectively.

300 ng of total RNA was used for small RNA library preparation using NEXTflex Small RNA-Seq Kit v3 (PerkinElmer, Cat. #NOVA-5132-05). In addition to the manufacturer's protocol, a spike-in mix of 0.05 ng 5'-cel-miR-39-3p-3'-OH and 0.05 ng 5'-ath-159a-3'-2-OMe was added at the initial library preparation step, to check for technical errors at library preparation and sequencing steps. Sequencing was carried out at the OHSU Massively Parallel Sequencing Shared Resource facilities on Illumina NovaSeq 6000 S4 2x100 flow cell. For RNA-seq data processing and piRNA annotation sequencing reads were trimmed with Cutadapt (v.#3.0) according to instructions provided by CATS small RNAseq kit protocol (Diagenode, Cat. #C05010040, Doc. # v.2 | 09.17) or NEXTflex Small RNA-Seq Kit protocol (PerkinElmer, Cat. #NOVA-5132-05, Doc. # v.V19.01). Next, trimmed reads were aligned to reference genome (GRCh37) with Bowtie (v.#1.0.1)⁷¹ allowing only perfect matches, discarded miRNAs by selecting reads between 25 and 45 bases, and re-aligned to GRCh37 allowing one mismatch. Finally, known small non-coding RNAs, other than piRNAs, were removed from the dataset using DASHRv2 (v.#v2)⁷² and the remaining piRNA sequences were intersected with known piRNA loci detected in human adult testis³. For statistical analysis data from small RNA-seq experiments were evaluated using SciPy (ver.: 1.8.0) packages⁷³.

Shapiro-Wilk test for normality of the data and Student's T-test for expression changes in piRNA quantities of different lengths were used.

Declarations

Acknowledgments

The authors gratefully thank all patients and their family members for providing data and samples and their consent for genomic analyses. The CeRA's physicians are thanked for taking care of the patients and documenting all clinical data in the database Androbase[®]. Moreover, the authors thank Christina Burhöi for her excellent technical support.

Conflict of interest

The authors declared no conflict of interest.

Funding

FT was supported by the Interdisciplinary Centre for Clinical Research Münster (IZKF,Tüt4/011/23) and the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) sponsored Clinical Research Unit 'Male Germ Cells' (CRU326, project number 329621271). CB was supported by a grant of the MedK program of the Medical Faculty Münster. CK and ARE were funded by the Spanish Ministry of Health Instituto Carlos III-FIS (grant numbers PI17/01822 and PI20/01562). JAV was supported by an Investigator Award in Science from the Wellcome Trust (209451).

Author contributions

BS and FT conceived and designed the experiments and wrote the manuscript. BS, DC, SV and CK supervised the experiments. CB, RS, A-KD, FG, LM, JS, ÖO, NL performed the experiments. AZ, DM, ARE, MX, CR, MJW, RS, SD analyzed the data. SK, DF, AP, GV, CK, JAV, AG and AS contributed patient samples. AZ, CF, MJW, GV, NN, JAV, SV, DC and DO provided critical feedback and helped shape the research and manuscript. All authors revised and approved the final version of the manuscript.

References

1. Ozata, D. M., Gainetdinov, I., Zoch, A., O'Carroll, D. & Zamore, P. D. PIWI-interacting RNAs: small RNAs with big functions. *Nat. Rev. Genet.* **20**, 89–108 (2019).
2. Wang, X., Ramat, A., Simonelig, M. & Liu, M.-F. Emerging roles and functional mechanisms of PIWI-interacting RNAs. *Nat. Rev. Mol. Cell Biol.* **24**, 123–141 (2023).

3. Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature* **442**, 199–202 (2006).
4. Aravin, A. *et al.* A novel class of small RNAs bind to MILI protein in mouse testes. *Nature* **442**, 203–7 (2006).
5. Gou, L.-T. *et al.* Pachytene piRNAs instruct massive mRNA elimination during late spermiogenesis. *Cell Res.* **24**, 680–700 (2014).
6. Zhang, P. *et al.* MIWI and piRNA-mediated cleavage of messenger RNAs in mouse testes. *Cell Res.* **25**, 193–207 (2015).
7. De Fazio, S. *et al.* The endonuclease activity of Mili fuels piRNA amplification that silences LINE1 elements. *Nature* **480**, 259–63 (2011).
8. Aravin, A. A., Sachidanandam, R., Girard, A., Fejes-Toth, K. & Hannon, G. J. Developmentally regulated piRNA clusters implicate MILI in transposon control. *Science* **316**, 744–7 (2007).
9. Reuter, M. *et al.* Miwi catalysis is required for piRNA amplification-independent LINE1 transposon silencing. *Nature* **480**, 264–7 (2011).
10. Kuramochi-Miyagawa, S. *et al.* DNA methylation of retrotransposon genes is regulated by Piwi family members MILI and MIWI2 in murine fetal testes. *Genes Dev.* **22**, 908–17 (2008).
11. Sun, Y. H., Lee, B. & Li, X. Z. The birth of piRNAs: how mammalian piRNAs are produced, originated, and evolved. *Mamm. Genome* (2021) doi:10.1007/s00335-021-09927-8.
12. Watanabe, T. *et al.* MITOPLD is a mitochondrial protein essential for nuage formation and piRNA biogenesis in the mouse germline. *Dev. Cell* **20**, 364–75 (2011).
13. Huang, H. *et al.* piRNA-associated germline nuage formation and spermatogenesis require MitoPLD profusogenic mitochondrial-surface lipid signaling. *Dev. Cell* **20**, 376–87 (2011).
14. Shiromoto, Y. *et al.* GPAT2 is required for piRNA biogenesis, transposon silencing, and maintenance of spermatogonia in mice†. *Biol. Reprod.* **101**, 248–256 (2019).
15. Ma, L. *et al.* GASZ is essential for male meiosis and suppression of retrotransposon expression in the male germline. *PLoS Genet.* **5**, e1000635 (2009).
16. Frost, R. J. A. *et al.* MOV10L1 is necessary for protection of spermatocytes against retrotransposons by Piwi-interacting RNAs. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 11847–52 (2010).
17. Kuramochi-Miyagawa, S. *et al.* MVH in piRNA processing and gene silencing of retrotransposons. *Genes Dev.* **24**, 887–92 (2010).
18. Kherraf, Z.-E. *et al.* Whole-exome sequencing improves the diagnosis and care of men with non-obstructive azoospermia. *Am. J. Hum. Genet.* **109**, 508–517 (2022).
19. Arafat, M. *et al.* Mutation in TDRD9 causes non-obstructive azoospermia in infertile men. *J. Med. Genet.* **54**, 633–639 (2017).
20. Alhathal, N. *et al.* A genomics approach to male infertility. *Genet. Med.* **22**, 1967–1975 (2020).
21. Nagirnaja, L. *et al.* Diverse monogenic subforms of human spermatogenic failure. *Nat. Commun.* **13**, 7953 (2022).

22. Nagirnaja, L. *et al.* Variant PNLDC1, Defective piRNA Processing, and Azoospermia. *N. Engl. J. Med.* **385**, 707–719 (2021).
23. Wyrwoll, M. J. *et al.* The piRNA-pathway factor FKBP6 is essential for spermatogenesis but dispensable for control of meiotic LINE-1 expression in humans. *Am. J. Hum. Genet.* **109**, 1850–1866 (2022).
24. Wyrwoll, M. J. *et al.* Genetic Architecture of Azoospermia-Time to Advance the Standard of Care. *Eur. Urol.* **83**, 452–462 (2023).
25. Vourekas, A. *et al.* Mili and Miwi target RNA repertoire reveals piRNA biogenesis and function of Miwi in spermiogenesis. *Nat. Struct. Mol. Biol.* **19**, 773–81 (2012).
26. Arif, A. *et al.* GTSF1 accelerates target RNA cleavage by PIWI-clade Argonaute proteins. *Nature* **608**, 618–625 (2022).
27. Yamaguchi, S. *et al.* Crystal structure of Drosophila Piwi. *Nat. Commun.* **11**, 858 (2020).
28. Doxzen, K. W. & Doudna, J. A. DNA recognition by an RNA-guided bacterial Argonaute. *PLoS One* **12**, e0177097 (2017).
29. Vourekas, A. *et al.* The RNA helicase MOV10L1 binds piRNA precursors to initiate piRNA processing. *Genes Dev.* **29**, 617–29 (2015).
30. Izumi, N., Shoji, K., Suzuki, Y., Katsuma, S. & Tomari, Y. Zucchini consensus motifs determine the mechanism of pre-piRNA production. *Nature* **578**, 311–316 (2020).
31. Shiromoto, Y. *et al.* GPAT2, a mitochondrial outer membrane protein, in piRNA biogenesis in germline stem cells. *RNA* **19**, 803–10 (2013).
32. Ding, D. *et al.* PNLDC1 is essential for piRNA 3' end trimming and transposon silencing during spermatogenesis in mice. *Nat. Commun.* **8**, 819 (2017).
33. Lim, S. L. *et al.* HENMT1 and piRNA Stability Are Required for Adult Male Germ Cell Transposon Repression and to Define the Spermatogenic Program in the Mouse. *PLoS Genet.* **11**, e1005620 (2015).
34. Takebe, M., Onohara, Y. & Yokota, S. Expression of MAEL in nuage and non-nuage compartments of rat spermatogenic cells and colocalization with DDX4, DDX25 and MIWI. *Histochem. Cell Biol.* **140**, 169–81 (2013).
35. Aravin, A. A. *et al.* Cytoplasmic compartmentalization of the fetal piRNA pathway in mice. *PLoS Genet.* **5**, e1000764 (2009).
36. Castañeda, J. *et al.* Reduced pachytene piRNAs and translation underlie spermiogenic arrest in Maelstrom mutant mice. *EMBO J.* **33**, 1999–2019 (2014).
37. Gan, B., Chen, S., Liu, H., Min, J. & Liu, K. Structure and function of eTudor domain containing TDRD proteins. *Crit. Rev. Biochem. Mol. Biol.* **54**, 119–132 (2019).
38. Xie, C. *et al.* Meiotic recombination: insights into its mechanisms and its role in human reproduction with a special focus on non-obstructive azoospermia. *Hum. Reprod. Update* **28**, 763–797 (2022).

39. Gou, L.-T. *et al.* Ubiquitination-Deficient Mutations in Human Piwi Cause Male Infertility by Impairing Histone-to-Protamine Exchange during Spermiogenesis. *Cell* **169**, 1090-1104.e13 (2017).
40. Oud, M. S. *et al.* Lack of evidence for a role of PIWIL1 variants in human male infertility. *Cell* **184**, 1941–1942 (2021).
41. Wang, X. *et al.* Generation of an iPSC line (HUSTi002-A) from fibroblasts of a patient with Sertoli cell-only syndrome carrying c.731_732delAT in PIWIL2 gene. *Stem Cell Res.* **42**, 101703 (2020).
42. Ghieh, F. *et al.* Whole-exome sequencing in patients with maturation arrest: a potential additional diagnostic tool for prevention of recurrent negative testicular sperm extraction outcomes. *Hum. Reprod.* **37**, 1334–1350 (2022).
43. Deng, W. & Lin, H. miwi, a murine homolog of piwi, encodes a cytoplasmic protein essential for spermatogenesis. *Dev. Cell* **2**, 819–30 (2002).
44. Kuramochi-Miyagawa, S. *et al.* Mili, a mammalian member of piwi family gene, is essential for spermatogenesis. *Development* **131**, 839–49 (2004).
45. Vasiliauskaitė, L. *et al.* A MILI-independent piRNA biogenesis pathway empowers partial germline reprogramming. *Nat. Struct. Mol. Biol.* **24**, 604–606 (2017).
46. Xiol, J. *et al.* A role for Fkbp6 and the chaperone machinery in piRNA amplification and transposon silencing. *Mol. Cell* **47**, 970–9 (2012).
47. Shoji, M. *et al.* The TDRD9-MIWI2 complex is essential for piRNA-mediated retrotransposon silencing in the mouse male germline. *Dev. Cell* **17**, 775–87 (2009).
48. Li, X. Z. *et al.* An ancient transcription factor initiates the burst of piRNA production during early meiosis in mouse testes. *Mol. Cell* **50**, 67–81 (2013).
49. Goh, W. S. S. *et al.* piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. *Genes Dev.* **29**, 1032–44 (2015).
50. Robine, N. *et al.* A broadly conserved pathway generates 3'UTR-directed primary piRNAs. *Curr. Biol.* **19**, 2066–76 (2009).
51. Krausz, C. *et al.* Genetic dissection of spermatogenic arrest through exome analysis: clinical implications for the management of azoospermic men. *Genet. Med.* (2020) doi:10.1038/s41436-020-0907-1.
52. Oud, M. S. *et al.* A de novo paradigm for male infertility. *Nat. Commun.* (2022) doi:10.1038/s41467-021-27132-8.
53. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal* **17**, 10 (2011).
54. Li, H. & Durbin, R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* **26**, 589–595 (2010).
55. McKenna, A. *et al.* The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* **20**, 1297–1303 (2010).
56. McLaren, W. *et al.* The Ensembl Variant Effect Predictor. *Genome Biol.* **17**, 122 (2016).

57. Karlsson, M. *et al.* A single-cell type transcriptomics map of human tissues. *Sci. Adv.* **7**, (2021).
58. Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* **25**, 25–9 (2000).
59. Gene Ontology Consortium *et al.* The Gene Ontology knowledgebase in 2023. *Genetics* **224**, (2023).
60. Thomas, P. D. *et al.* PANTHER: Making genome-scale phylogenetics accessible to all. *Protein Sci.* **31**, 8–22 (2022).
61. Mi, H. *et al.* Protocol Update for large-scale genome and gene function analysis with the PANTHER classification system (v.14.0). *Nat. Protoc.* **14**, 703–721 (2019).
62. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**, e21800 (2011).
63. Kuznetsova, I., Lugmayr, A., Siira, S. J., Rackham, O. & Filipovska, A. CirGO: an alternative circular way of visualising gene ontology terms. *BMC Bioinformatics* **20**, 84 (2019).
64. Pedersen, B. S. *et al.* Somalier: rapid relatedness estimation for cancer and germline studies using efficient genome sketches. *Genome Med.* **12**, 62 (2020).
65. Kishore, S., Khanna, A. & Stamm, S. Rapid generation of splicing reporters with pSpliceExpress. *Gene* **427**, 104–10 (2008).
66. Mirdita, M. *et al.* ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).
67. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
68. Brinkworth, M. H., Weinbauer, G. F., Schlatt, S. & Nieschlag, E. Identification of male germ cells undergoing apoptosis in adult rats. *J. Reprod. Fertil.* **105**, 25–33 (1995).
69. Dicke, A.-K. *et al.* DDX3Y is likely the key spermatogenic factor in the AZFa region that contributes to human non-obstructive azoospermia. *Commun. Biol.* **6**, 350 (2023).
70. Albert, S. *et al.* Comparative marker analysis after isolation and culture of testicular cells from the immature marmoset. *Cells. Tissues. Organs* **196**, 543–54 (2012).
71. Langmead, B., Trapnell, C., Pop, M. & Salzberg, S. L. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
72. Kuksa, P. P. *et al.* DASHR 2.0: integrated database of human small non-coding RNA genes and mature products. *Bioinformatics* **35**, 1033–1039 (2019).
73. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

Figures

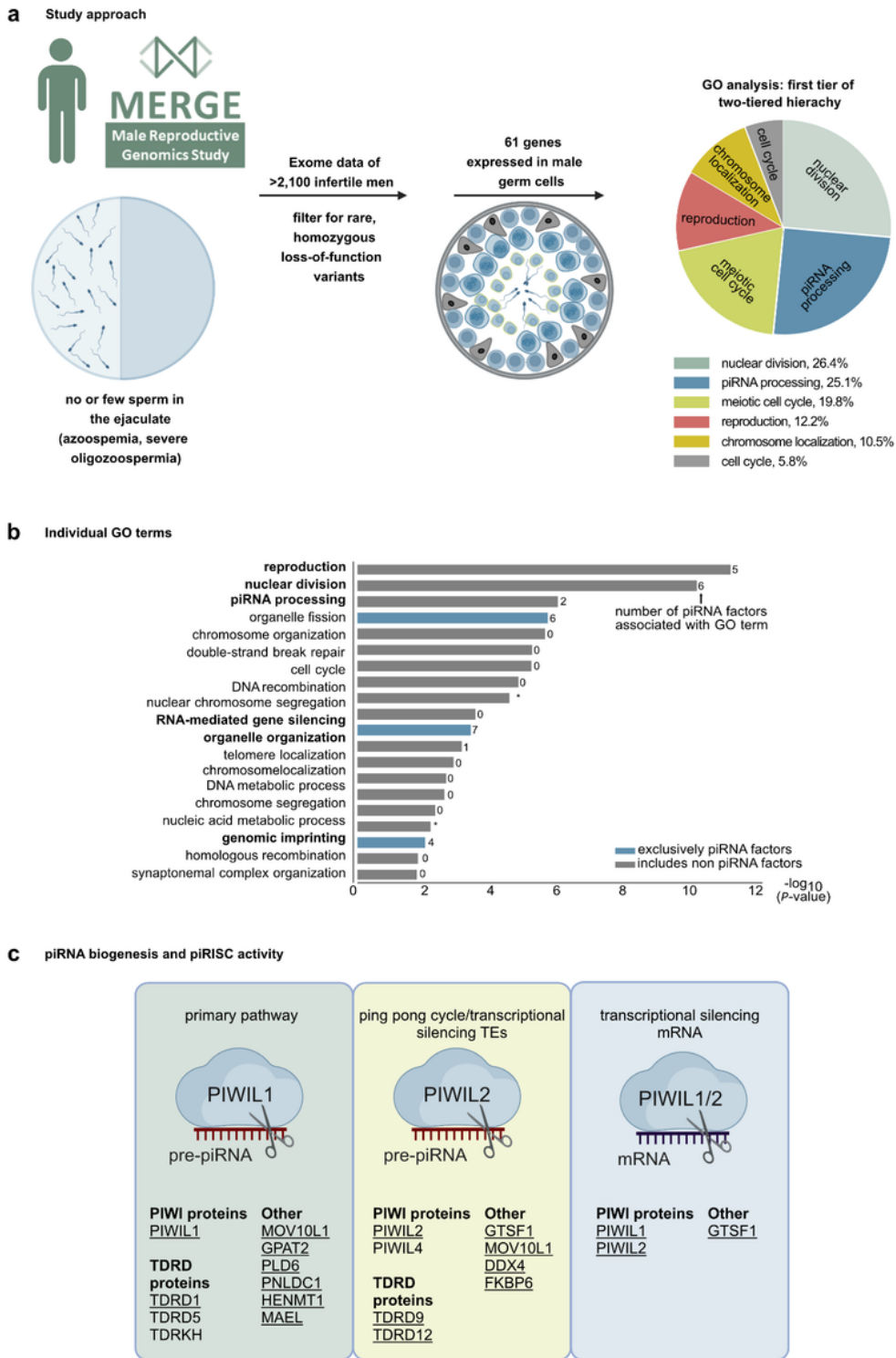


Figure 1

Genetic landscape of piRNA biogenesis-related male infertility. a. Workflow of scrutinizing biological processes related to genetically determined reduced sperm count and male infertility by gene ontology (GO) analysis. Pie chart shows first hierarchy of the two-tiered hierarchy. b. Side-by-side bar chart showing $-\log_{10}(P\text{-value})$ of individual GO terms. Number of piRNA pathway factors associated with GO term shown to the right of each bar and GO terms associated exclusively with piRNA pathway factors

highlighted in blue. c. Schematic overview on mammalian piRNA biogenesis related sub-pathways with proteins factors known to be involved from mice knockout studies. Proteins in which encoded biallelic high-impact variants were identified in infertile men are underlined.

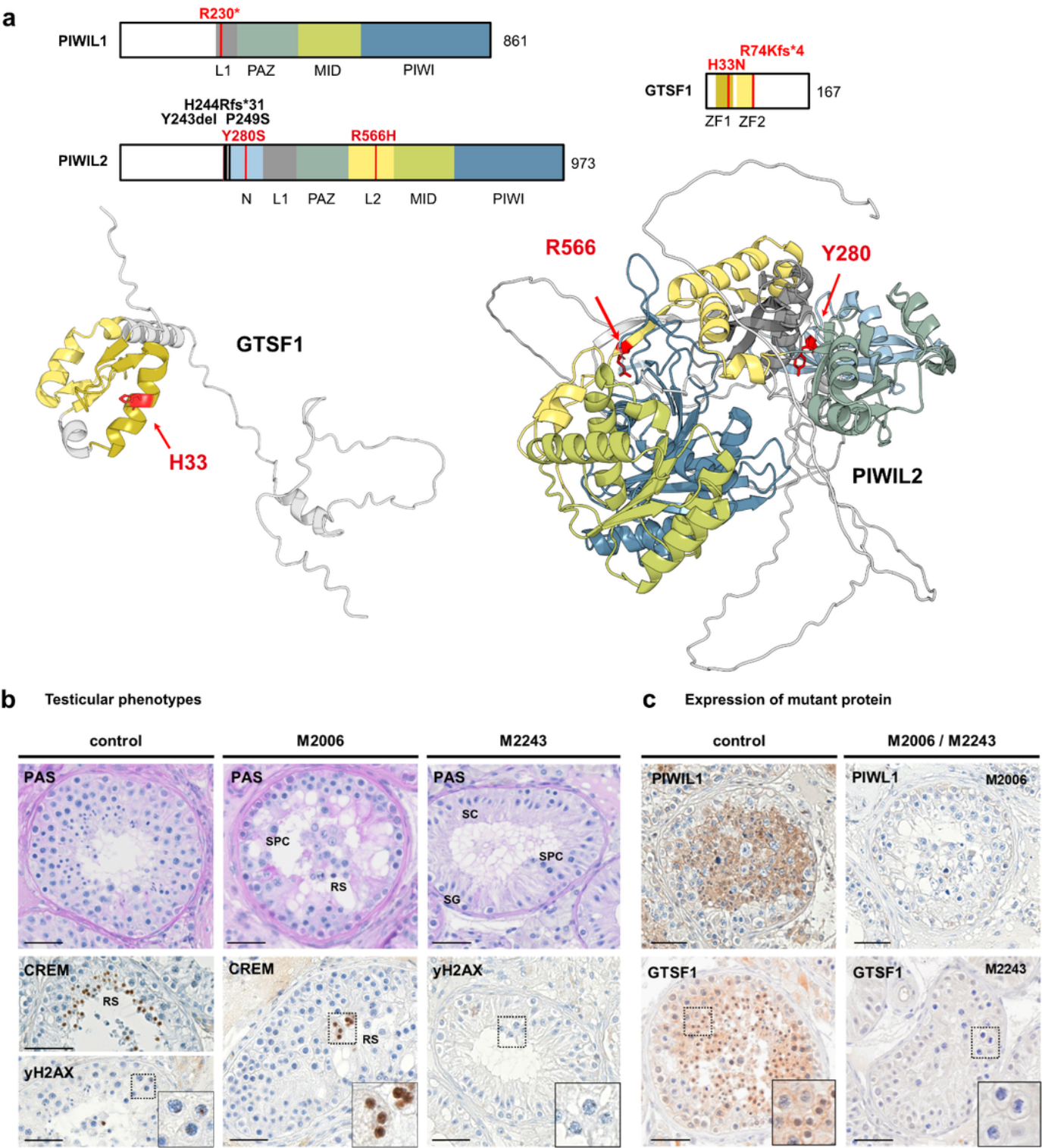
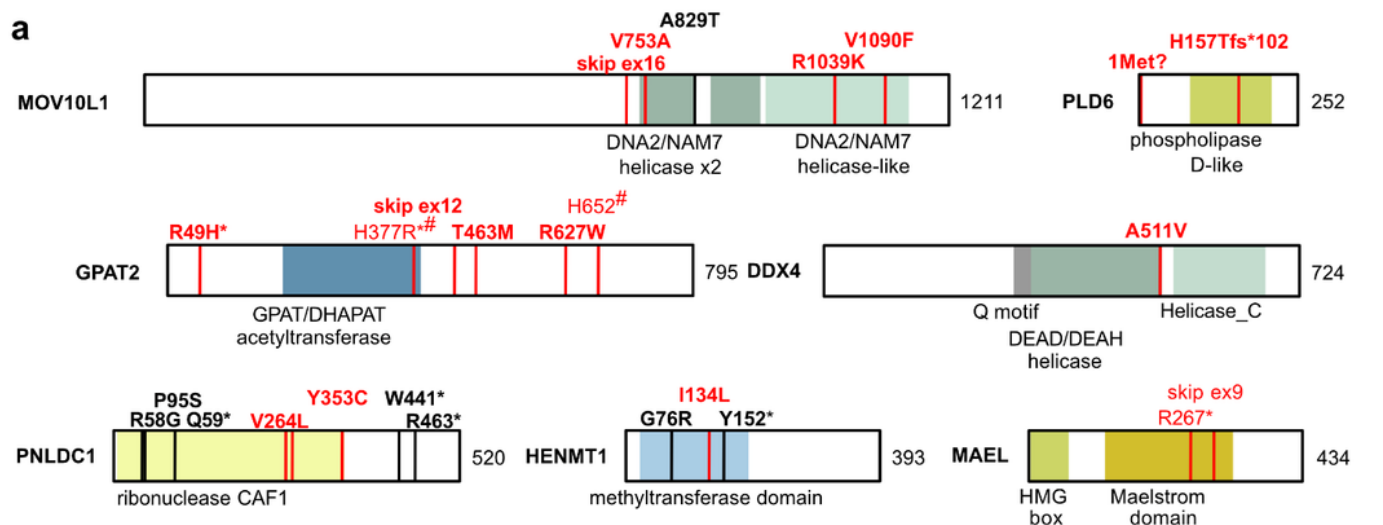
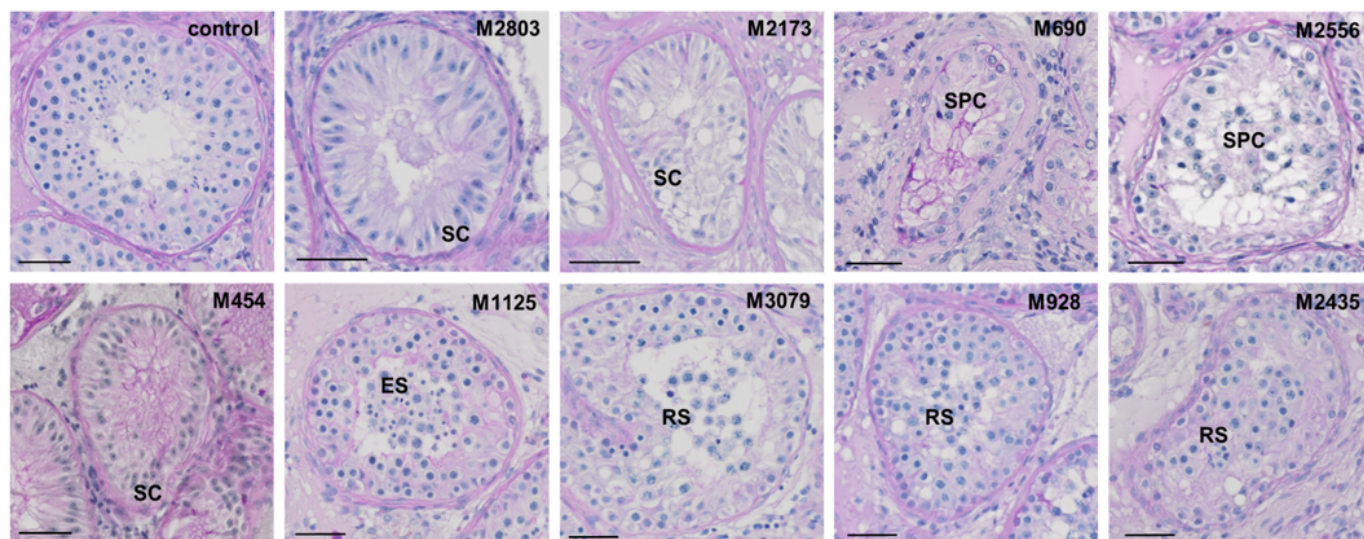


Figure 2

Homozygous variants identified in genes of the piRISC complex and associated testicular phenotypes. a. Schematic representation and AlphaFold2 structure predictions of PIWIL1, PIWIL2 and GTSF1. The schematic depicts both novel (red) and known (black) homozygous variants, with amino acids affected by new variants highlighted in the protein structure in red. N: N-terminal structured domain, PAZ: Piwi/Argonaute/Zwille domain, PIWI: Piwi-like domain; L1: linker domain 1, L2: linker domain 2, MID: middle domain, ZF: zinc finger domain. b. Periodic acid-Schiff (PAS) staining of testicular tissue of men with normal spermatogenesis (control), M2006 [PIWIL1, p.(Arg230*)] and M2243 [GTSF1, p.(Arg74Lysfs*4)] and immunohistochemical staining (IHC) for round spermatid marker protein Cyclic AMP Element Modulator (CREM) and spermatocyte marker protein γ H2AX. In M2006, round spermatids were detected as most advanced germ cells, whereas in M2243, in addition to spermatogonia, rarely pachytene spermatocytes with γ H2AX positive sex bodies but no haploid germ cells were observed. c. IHC staining for PIWIL1 and GTSF1 in controls and variant carriers demonstrating absence of PIWIL1 in M2006 due to homozygous stop-gain variant p.(Arg230*) and absence of GTSF1 in M2243 with homozygous frameshift variant p.(Arg74Lysfs*4). Scale bar = 50 μ m. SC: Sertoli cell; SPC: spermatocyte; RS: round spermatid; ES: elongated spermatid.



b Testicular phenotypes



c Expression of mutant protein

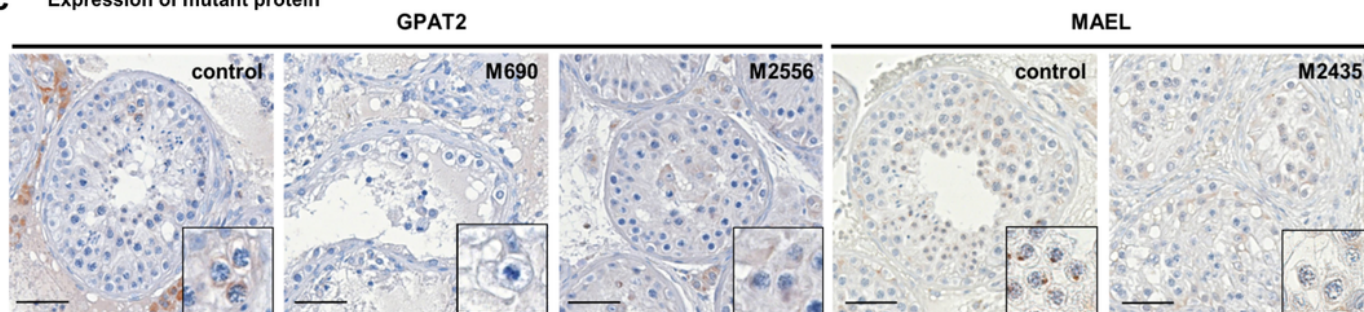


Figure 3

Biallelic variants identified in human piRNA biogenesis-associated genes. a. Localization of variants in schematic of MOV10L1, PLD6, GPAT2, PNLDC1, MAEL, DDX4, and HENMT1 structure with protein domains colored and newly identified biallelic variants (red, bold for homozygous) as well as previously described homozygous variants (black) indicated. Pairs of compound heterozygous variants are indicated by identical symbols (*,#) in superscript. b. Periodic acid-Schiff (PAS) staining of representative

testicular tissue of variant carriers demonstrating SCO in M2803 [PLD6, p.(His157Thrfs*102)], M2173 [PLD6, p.(Met1?)] and M454 [GPAT2, p.(His377Arg)/(Arg49His)] and presence of haploid germ cells (round/elongated spermatids, RS/ES) in M1125 [PNLDC1, p. (Tyr353Cys)], M3079 [HENMT1, p. (Ile134Leu)], M928 [DDX4, p. (Ala511Val)], and M2435 [MAEL, p.(Arg267*)/c.908+1G>C]. c. Immunohistochemical (IHC) staining for GPAT2 in testicular tissue with full spermatogenesis (control) and *GPAT2* variant carriers with meiotic arrest, [M690, p.(Arg627Trp)], M2556, c.1156-1G>A]. In control tissue, GPAT2 is expressed in perinuclear structures in spermatocytes and this staining pattern is absent in M690 and strikingly reduced in M2556. D. IHC for MAEL in testicular tissue with full spermatogenesis (control) and M2435 with compound heterozygous presence of two *MAEL* LoF variants [p. (Arg267*)/c.908+1G>C]. In control tissue, MAEL is expressed in perinuclear structures in spermatocytes and distinct condensed structures in round spermatids and this staining pattern is absent in the variant carrier. Scale bar = 50 µm. SC: Sertoli cell; SPC: spermatocyte; RS: round spermatid; ES: elongated spermatid.

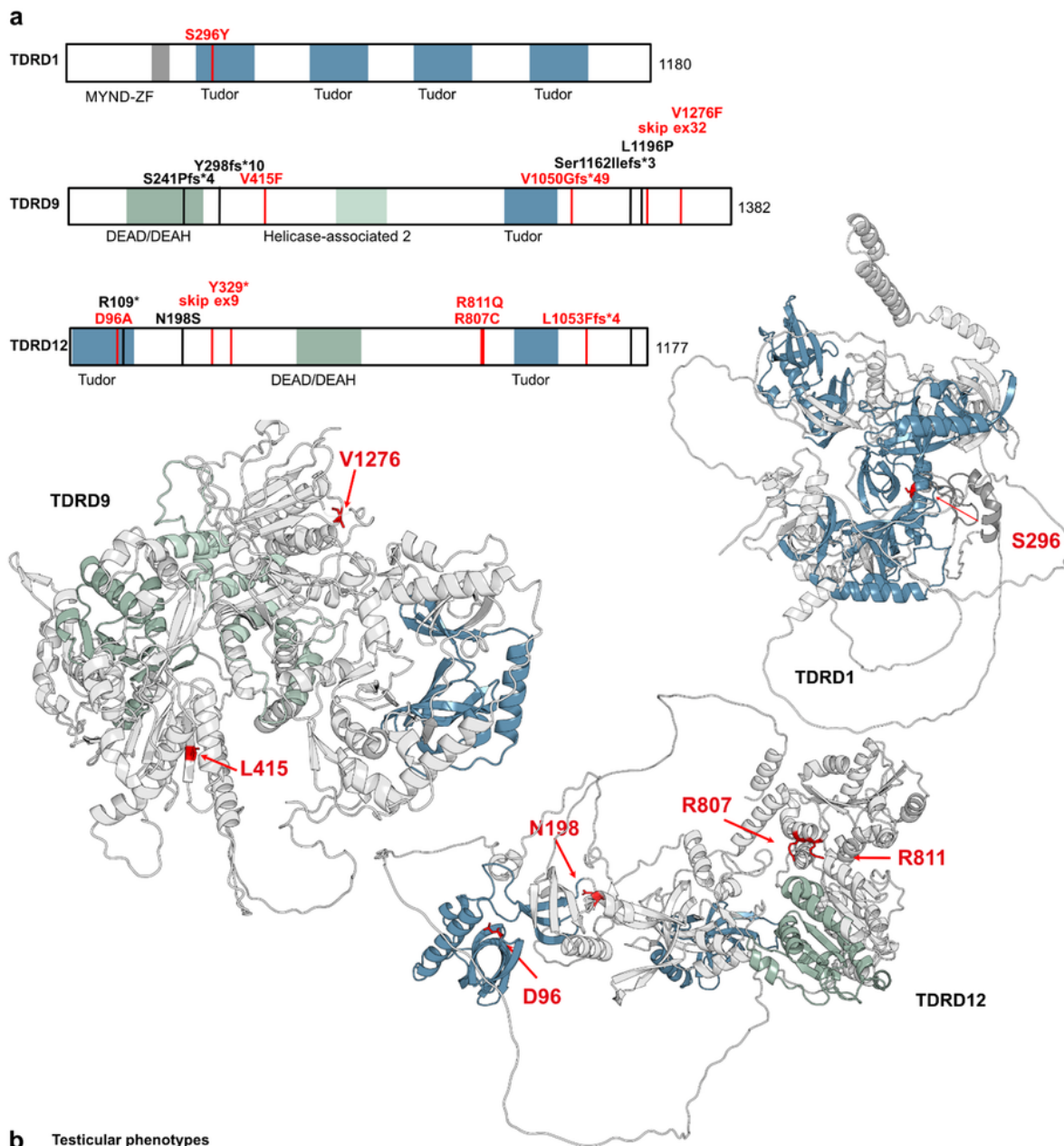


Figure 4

Homozygous high-impact variants identified in human genes of the tudor-domain containing gene family (TDRDs). a. Schematic representation and AlphaFold2 structure predictions of TDRDs. The schematic depicts both novel (red) and known (light red) homozygous variants, with amino acids affected by new variants highlighted in the protein structures in red. MYND-ZF: MYND-type zinc finger domain (grey), Tudor: tudor domain (blue), helicase domains: DEAD/DEAH (green), helicase-associated 2 (light green). b.

Periodic acid-Schiff (PAS) staining of testicular tissue of variant carriers demonstrating meiotic arrest in M1648 [TDRD1, p.(Ser296Tyr)], presence of elongated spermatids in M800 [TDRD9, p.(Val1050Glyfs*49)], SCO in M1642 [TDRD12, p.(Asn198Ser)] and round spermatid arrest in M2227 [TDRD12, p.(Trp329*)]. Scale bar = 50 μ m. SC: Sertoli cell; SG: spermatogonia; SPC: spermatocyte; RS: round spermatid; ES: elongated spermatid.

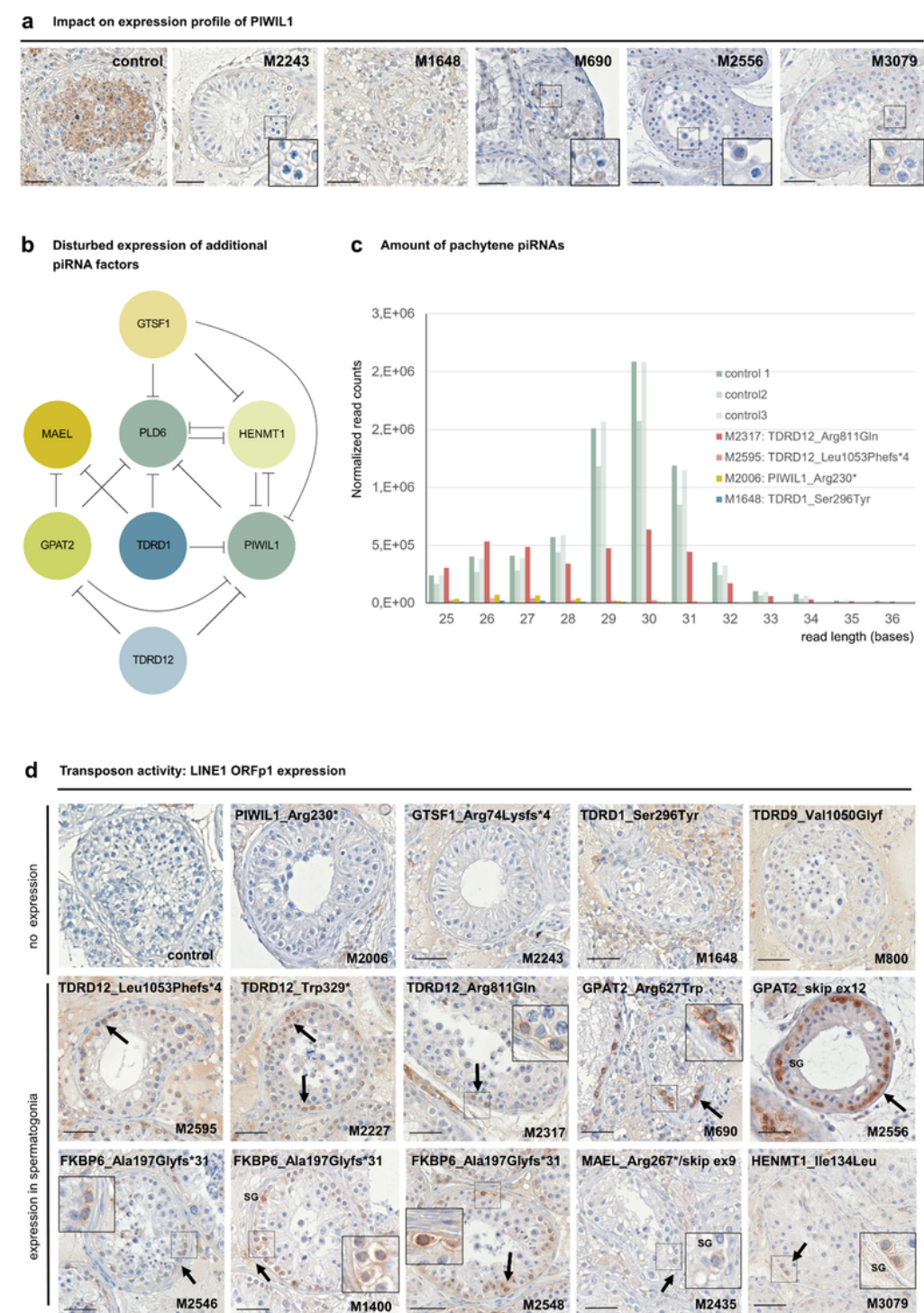
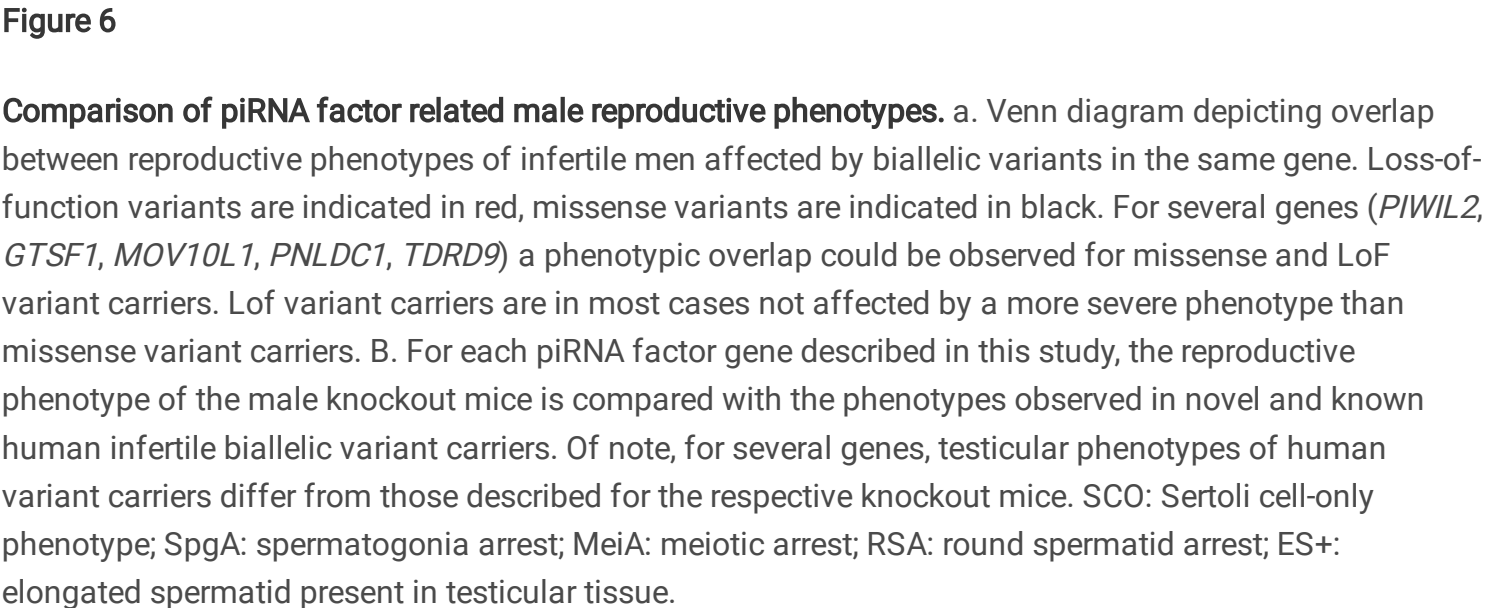


Figure 5

Functional impact of disturbed piRNA biogenesis. a. Immunohistochemical staining demonstrating diminished expression of PIWIL1 in variant carriers M2243 [GTSF1, p.(Arg74Lysfs*4)], M1648 [TDRD1, p.(Ser296Tyr)], M690 [GPAT2, p.(Arg627Trp)], M2556 (*GPAT2*, c.1156-1G>A) and M3079 [HENMT1, p.(Ile134Leu)]. b. Schematic depicting impact of loss of piRNA biogenesis protein function on expression of additional piRNA factors. Circles represent piRNA protein and inhibiting effects of loss-of- protein functions on the expression of further piRNA proteins are indicated. c. Effect of biallelic variants in genes of the piRNA pathway on biogenesis of pachytene piRNAs. RNA isolated from snap frozen testicular tissue of M2006 [PIWIL1 p.(Arg230*)], M1648 [TDRD1 p.(Ser296Tyr)], M2317 [TDRD12 p.(Arg811Gln)] and M2595 [TDRD12 p.(Leu1053Phefs*4)] revealed a major loss of pachytene piRNAs with expected lengths of 26 to 31 bases when compared with controls with complete spermatogenesis (ctrl1-3; $P < 0.001$). d. Immunohistochemical staining for LINE1 transposon specific protein LINE1 ORFp1 in testicular tissue. LINE1 ORFp1 was not detected in testicular tissue of controls with full spermatogenesis and carriers of homozygous high-impact variants in *PIWIL1*, *GTSF1*, *TDRD1*, and *TDRD9*, respectively. In contrast, all three carriers of homozygous variants in *TDRD12* (M2595, M2227, and M2317) and *FKBP6* (M2546, M2548, M1400) as well as both *GPAT2* variant carriers (M690 and M2556) revealed a concordant distinct and specific LINE1 ORFp1 staining that is restricted to spermatogonia. A similar effect was also seen for M2435 (*MAEL*) and M3079 (*HENMT1*). Scale bar = 50 μ m. SG: spermatogonia.

b Testicular phenotypes in human and mice

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Extendeddattable1.xlsx](#)
- [SupplementaryData.docx](#)
- [nrreportingsummaryNCOMMS2362310T28122023.pdf](#)
- [Extendeddata.docx](#)