OPEN ACCESS

Original research

Novel mutation leading to splice donor loss in a conserved site of *DMD* gene causes Duchenne muscular dystrophy with cryptorchidism

Jianhai Chen ,[1,2] Yangying Jia,[1,3] Jie Zhong,[4] Kun Zhang,[5] Hongzheng Dai,[6] Guanglin He,[4] Fuping Li,[7] Li Zeng,[8] Chuanzhu Fan,[9] Huayan Xu[5]

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/jmg-2024-109896).

For numbered affiliations see end of article.

Correspondence to
Dr Huayan Xu, Department of Radiology, Sichuan University, Chengdu, China;
xuhuayan89@sina.com

Received 22 January 2024
Accepted 4 April 2024
Published Online First 15 April 2024

ABSTRACT

Background As one of the most common congenital abnormalities in male births, cryptorchidism has been found to have a polygenic aetiology according to previous studies of common variants. However, little is known about genetic predisposition of rare variants for cryptorchidism, since rare variants have larger effective size on diseases than common variants.

Methods In this study, a cohort of 115 Chinese probands with cryptorchidism was analysed using whole-genome sequencing, alongside 19 parental controls and 2136 unaffected men. Additionally, CRISPR-Cas9 editing of a conserved variant was performed in a mouse model, with MRI screening used to observe the phenotype.

Results In 30 of 115 patients (26.1%), we identified four novel genes (*ARSH*, *DMD*, *MAGEA4* and *SHROOM2*) affecting at least five unrelated patients and four known genes (*USP9Y*, *UBA1*, *BCORL1* and *KDM6A*) with the candidate rare pathogenic variants affecting at least two cases. Burden tests of rare variants revealed the genome-wide significances for newly identified genes ($p<2.5\times10^{-6}$) under the Bonferroni correction. Surprisingly, novel and known genes were mainly found on X chromosome (seven on X and one on Y) and all rare X-chromosomal segregating variants exhibited a maternal inheritance rather than de novo origin. CRISPR-Cas9 mouse modelling of a splice donor loss variant in *DMD* (NC_000023.11:g.32454661C>G), which resides in a conserved site across vertebrates, replicated bilateral cryptorchidism phenotypes, confirmed by MRI at 4 and 10 weeks. The movement tests further revealed symptoms of Duchenne muscular dystrophy (DMD) in transgenic mice.

Conclusion Our results revealed the role of the *DMD* gene mutation in causing cryptorchidism. The results also suggest that maternal-X inheritance of pathogenic defects could have a predominant role in the development of cryptorchidism.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Previous GWAS (Genome-wide association studies) and pathway studies have revealed significant correlation between calponin homology domain proteins (such as *DMD* protein) and non-syndromic cryptorchidism.

WHAT THIS STUDY ADDS

⇒ Using rare variant and genetic burden analysis, we identified candidate pathogenic loci and genes related to cryptorchidism. Based on transgenic mice model and movement tests, we confirmed that a splice donor-related mutation in a conserved site of *DMD* gene may cause Duchenne muscular dystrophy (DMD) with cryptorchidism.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The study demonstrates the modelling utility of a mutation causing DMD with cryptorchidism.

INTRODUCTION

Cryptorchidism (also known as 'hidden testicle', Online Mendelian Inheritance in Man (OMIM) 219050) has a global prevalence rate of 2%–4% at full-term birth in boys.[1] Despite its relatively low prevalence, cryptorchidism is the most common birth defect involving urogenital abnormalities in newborn boys and 5.2% of DMD patients show cryptorchidism.[2] It is also the best-characterised risk factor for infertility and testicular cancer.[3] Normally, the testes spontaneously descend into the scrotum by the eighth month (33 weeks) of pregnancy, likely triggered by transient activation of the hypothalamus–pituitary–gonadal axis, leading to an increase in reproductive hormone levels.[4] In patients with cryptorchidism, one or both testes may fail to descend completely into the dependent part of the scrotum, resulting in unilateral or bilateral cryptorchidism, respectively.

Maternal risk factors before and during pregnancy, such as maternal smoking, endocrine-disrupting chemicals and gestational diabetes, have increasingly been recognised as the risk factors for congenital cryptorchidism.[5] However, the studies on genetic risk factors of cryptorchidism are still limited. Multiple aetiological features, such as familial aggregation and increased prevalence in first-degree relatives, suggest a genetic predisposition to this disease.[6] For instance, a family history of cryptorchidism can increase the risk of the condition in newborn boys by approximately threefold.[7 8] The recurrence risk has been found to be more than twice as high in brothers of cryptorchidism cases compared with healthy controls.[9] Thus, despite cryptorchidism's complex susceptibility to maternal and environmental factors, strong genetic components may also contribute to its aetiology.

Check for updates

To cite: Chen J, Jia Y, Zhong J, *et al*. *J Med Genet* 2024;**61**:741–749.

Previous microarray studies based on common variants (GWAS) have suggested a potential association of certain pathways with cryptorchidism.[10] Indeed, both population genetics theory and empirical data suggest that rare variants have much larger effect sizes than common variants,[11–14] indicating their significant role in causing genetic diseases. It is estimated that most human rare protein-altering variants (missense) are pathogenic.[15] Furthermore, disruptive mutations (non-sense) are disproportionately common causal factors in about 11.5% of human genetic diseases.[16] At the population level, these causal variants are rare in allele frequency due to strong purifying selection, which limits their accumulation across generations. Therefore, rare deleterious variants with large genetic effect sizes represent a promising area for discovering causative genes for genetic diseases,[17] especially for diseases affecting fertility, such as cryptorchidism.

The application of various genetic screening techniques, including whole-genome sequencing (WGS) and exome sequencing, has facilitated the precise diagnosis of numerous genetic diseases caused by rare variants.[18–25] For instance, an exome sequencing study on an Indian family identified a rare deleterious homozygous missense variant in the *RXFP2* gene, which causes cryptorchidism.[26] Although exome sequencing is more commonly used than WGS due to its lower cost, WGS offers a more uniform distribution of sequencing quality (including single-nucleotide variants, insertions and deletions) and a higher discovery rate of coding-region variants, approximately 3%.[27 28]

To date, a rare variants study of cryptorchidism based on WGS is still lacking. In this study, we conducted rare variants screening based on WGS data for 134 Han Chinese samples including cases (115) and family controls (19). We also used a local cohort of 2136 unaffected men for comparison. We identified candidate rare pathogenic variants based on autosomal recessive, compound heterozygous and X-linked recessive inheritance mode for cryptorchidism patients. We annotated these variants and recognised both known and novel candidate genes and variants. We also confirmed the role of a novel rare mutation in *DMD* (NC_000023.11:g.32454661C>G) in causing cryptorchidism based on transgenic mouse modelling. During follow-up examinations, two patients with mutations in the *DMD* gene were diagnosed with Duchenne muscular dystrophy (DMD). Additionally, movement assessments conducted on transgenic mice exhibited symptoms consistent with muscular dystrophy. Our study may facilitate the development of molecular diagnosis and precision medicine for patients suffering from cryptorchidism.

RESULTS

The MRI on typical bilateral and unilateral cryptorchidism and basic sequencing statistics

The diagnosis of cryptorchidism was based on a physical exam by professional paediatric urologists following standard protocols.[29] Typical bilateral cryptorchidism and unilateral cryptorchidism of two patients were shown using the MRI (figure 1A–D). In comparison to controls, the patient with bilateral cryptorchidism had two non-descended testes, while the patient with unilateral cryptorchidism had only one undescended testis (figure 1). These two patients were also diagnosed with symptoms characteristic of DMD by specialists in radiology and paediatric neuromuscular disorders.

We conducted WGS on 115 cases and 19 family controls. Among the cases, there were 21 (18.26%) bilateral cryptorchidism patients and 94 (81.74%) unilateral cryptorchidism patients. This dataset also includes three family trios of bilateral cryptorchidism, four trios of unilateral cryptorchidism, and three duos with only cryptorchidism cases and mothers. Based on previous sensitivity tests, 15× WGS depth/fold can achieve accurate SNV (single nucleotide variant) calling, while 30× is sufficient for calling indels.[30 31] In this study, the average sequencing depths were 33.03-fold for cases and 27.97-fold for controls, suggesting a balanced design of sequencing depth and quality for all samples (online supplemental table 1, figure 1F and online supplemental figure 1a). To confirm the reported relationship, we used whole-genome SNPs to estimate the relatedness between family members. We also confirmed the population ancestry of all samples based on the principal component analysis. Based on identity by state distances among family members, we confirmed that family members are genetically closer than unrelated individuals (online supplemental figure 1b). We also confirmed that all parents–son relationships are within the coefficient range of first-degree relatives (from 0.177 to 0.354). By incorporating individuals from the '1000 genomes' project, we further revealed that all newly sequenced samples are closely related to the East Asian population (online supplemental figure 1c).

The rare candidate pathogenic variants for known genes

With the keywords 'cryptorchidism variants' and 'cryptorchidism genetics', we conducted the initial screening for known genes in the PubMed literature database.[32] We retrieved 56 genes previously reported to be associated with cryptorchidism (online supplemental table 2). Most of these previous studies were based on the genetic screening of small patient samples. Thus, it is interesting to know whether these genes can be confirmed in our cohort. We found that 41 of these genes were also registered in the OMIM and Human Phenotype Ontology database (June 2022) as genes related to cryptorchidism. We also added 738 genes associated with cryptorchidism phenotypes from these databases. These 779 genes (738+41) were used as a 'pool' of known genes (online supplemental table 2). We annotated non-redundant variants using the online tool VEP (Variant Effect Predictor) in the Ensembl database[33] and then summarised both known and novel candidate genes.

We detected four known genes (*BCORL1*, *KDM6A*, *UBA1* and *USP9Y*) with candidate rare pathogenic variants present in at least two cases but absent in paternal controls (table 1 and online supplemental table 3). Six variants were not detected in all publicly available population databases (gnomAD, 1000 genomes, dbSNP, ExAC, ClinVar, etc). One variant, *BCORL1* (NC_000023.11:g.130028727G>A;NP_001171701.1:p.(G1391R)), was also classified as rare based on the allele frequency of all human populations, with the highest frequency in East Asian population (0.005439; online supplemental table 3). Although this allele frequency is relatively higher in the East Asian population than globally, it remains much lower than the global prevalence rate of cryptorchidism in boys, which is 2%–4%.

Interestingly, all known genes affecting at least two cases are on sex chromosomes. Not a single autosomal gene was associated with more than one case. This strongly suggests a predominant role of sex chromosomal genes in cryptorchidism pathogenicity. The Y-chromosome gene *USP9Y* (ubiquitin-specific peptidase 9 Y-linked) was detected in six boys with unilateral cryptorchidism. In addition, three X-chromosome genes (*BCORL1*, *UBA1* and *KDM6A*) were found to carry candidate rare pathogenic variants affecting at least two patients.

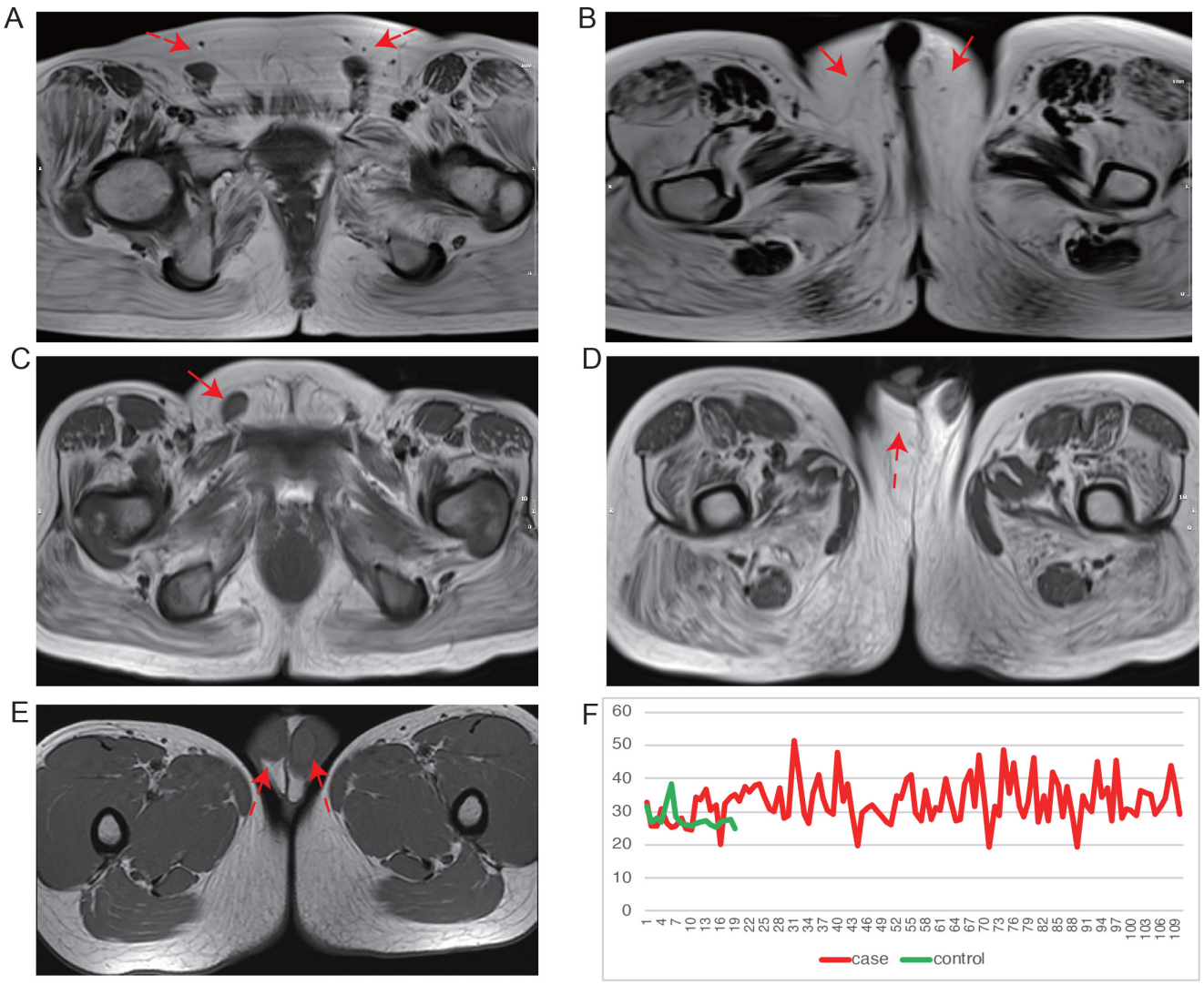

Figure 1 The MRI and sequencing depth. (A–E) The two MRI graphs for bilateral cryptorchidism (case B130). (C, D) The two MRI graphs for unilateral cryptorchidism (case U111). (E) The MRI for one control. The red arrows show the location of the testes. (F) The whole-genome sequencing overall depths for the cases (red) and control (green).

Table 1 The candidate pathogenic variants of known genes

ID	HGVS (Human Genome Variation Society)	Gene,Chr	Pred
U106	NC_000023.11:g.130028727G>A;NP_001171701.1:p.(G1391R)	*BCORL1*	6
U37	NC_000023.11:g.130028727G>A;NP_001171701.1:p.(G1391R)	*BCORL1*	6
U50	NC_000023.11:g.130028727G>A;NP_001171701.1:p.(G1391R)	*BCORL1*	6
U9	NC_000023.11:g.130028727G>A;NP_001171701.1:p.(G1391R)	*BCORL1*	6
U60	NC_000023.11:g.45069958A>C;NP_001278345.1:p.(K775T)	*KDM6A*	5
U58	NC_000023.11:g.45110152A>G;NP_001278344.1:p.(D1412G)	*KDM6A*	6
U51	NC_000023.11:g.47212835A>T;NP_003325.2:p.(Y873F)	*UBA1*	11
U60	NC_000023.11:g.47214843G>C;NP_695012.1:p.(V1031L)	*UBA1*	5
U49	NC_000024.10:g.12722148G>A;NP_004645.2:p.(E96K)	*USP9Y*	6
U50	NC_000024.10:g.12722148G>A;NP_004645.2:p.(E96K)	*USP9Y*	7
U109	NC_000024.10:g.12739592C>G;NP_004645.2:p.(S462C)	*USP9Y*	5
U73	NC_000024.10:g.12739592C>G;NP_004645.2:p.(S462C)	*USP9Y*	5
U13	NC_000024.10:g.12739592C>G;NP_004645.2:p.(S462C)	*USP9Y*	5
U71	NC_000024.10:g.12739592C>G;NP_004645.2:p.(S462C)	*USP9Y*	5

'B' and 'U' in the ID column are 'bilateral cryptorchidism' and 'unilateral cryptorchidism', respectively. The 'Pred' column shows the number of prediction methods that support deleterious effects of the variants (see online supplemental table 3 for specific methods). Note: USP9Y (underlined) is a Y chromosome gene, while all other genes reside on the X chromosome.

Table 2 The pathogenic variants in novel candidate genes

ID	HGVSc	Gene	Pred
U27	NC_000023.11:g.3024049G>C;NP_001011719.1:p.(E310D)	*ARSH*	5
U64	NC_000023.11:g.3024093G>C;NP_001011719.1:p.(G325A)	*ARSH*	13
B122	NC_000023.11:g.3033050C>G;NP_001011719.1:p.(P452A)	*ARSH*	9
U127	NC_000023.11:g.3033050C>G;NP_001011719.1:p.(P452A)	*ARSH*	9
U28	NC_000023.11:g.3033050C>G;NP_001011719.1:p.(P452A)	*ARSH*	9
U42	NC_000023.11:g.31774145C>T; LRG_199p1:p.(E2453K)	*DMD*	3
U55	NC_000023.11:g.32342234G>A;NP_004000.1:p.(R1926C)	*DMD*	6
U114	NC_000023.11:g.32346044G>C;NP_000100.2:p.(Q1821E)	*DMD*	1
U58	NC_000023.11:g.32346044G>C;NP_000100.2:p.(Q1821E)	*DMD*	1
B130	NC_000023.11:g.32454661C>G (splice donor loss variant)	*DMD*	1
U70	NC_000023.11:g.32472199A>G;NP_000100.2:p.(Y964H)	*DMD*	5
U111	NC_000023.11:g.32573783C>T; NP_000100.2:p.(D548N)	*DMD*	7
U106	NC_000023.11:g.151924064G>A;NP_001011550.1:p.(A134T)	*MAGEA4*	2
U36	NC_000023.11:g.151924064G>A;NP_001011550.1:p.(A134T)	*MAGEA4*	2
U8	NC_000023.11:g.151924064G>A;NP_001011550.1:p.(A134T)	*MAGEA4*	2
U94	NC_000023.11:g.151924064G>A;NP_001011550.1:p.(A134T)	*MAGEA4*	2
U27	NC_000023.11:g.151924194A>T;NP_001373127.1:p.(Y177F)	*MAGEA4*	4
U98	NC_000023.11:g.9786618C>A;NP_001640.1:p.(R25S)	*SHROOM2*	2
U33	NC_000023.11:g.9894590G>A;NP_001640.1:p.(D228N)	*SHROOM2*	2
U2	NC_000023.11:g.9937394C>T;NP_001307593.1:p.(P118L)	*SHROOM2*	5
U55	NC_000023.11:g.9937537G>T;NP_001307593.1:p.(A166S)	*SHROOM2*	2
B79	NC_000023.11:g.9937537G>T;NP_001307593.1:p.(A166S)	*SHROOM2*	5
B81	NC_000023.11:g.3024049G>C;NP_001011719.1:p.(E310D)	*SHROOM2*	5

'B' and 'U' in the ID column are 'bilateral cryptorchidism' and 'unilateral cryptorchidism', respectively. The 'Pred' column shows the number of prediction methods supporting the variants as deleterious (see online supplemental table 4 for specific in silico prediction algorithms). Note: all genes are on the X chromosome.

Strikingly, the two rare pathogenic variants in *BCORL1* (NC_000023.11:g.130028727G>A; NP_001171701.1:p.(G1391R)) and *USP9Y* (NC_000024.10:g.12739592C>G; NP_004645.2:p.(S462C)) were found in four unilateral patients per variant, with only the former registered in the gnomAD database. Both variants were predicted to be deleterious by multiple in silico algorithms (online supplemental table 3). In addition, the wild-type alleles of these variants showed evolutionary conservation in primate species, based on sequence alignments of primate orthologous genes from the Ensembl database (v105, online supplemental figure 2). Together, the candidate rare pathogenic variants of known genes were found to affect 12 cases, indicating a molecular diagnostic rate of 10.4% (12/115). The diagnostic rates for bilateral and unilateral cryptorchidism were 0% (0/21) and 12.77% (12/94), respectively. These candidate pathogenic variants and genotypes in cases and parents were validated using the Sanger sequencing (online supplemental figure 3).

The rare candidate pathogenic variants for novel genes

To identify candidate pathogenic variants in novel candidate genes for cryptorchidism, we excluded cases with the above-mentioned pathogenic variants in known genes. Subsequently, we ranked genes based on the burden of rare deleterious variants that occurred in cases. To increase the reliability of the newly identified genes, we focused on genes affecting at least five cases with candidate variants. We found four genes that fulfilled these rigorous requirements: *DMD*, *ARSH*, *MAGEA4* and *SHROOM2* (table 2).

Based on allele frequencies in multiple genome databases (gnomAD, 1000 genomes, dbSNP, ExAC, ClinVar, etc) and in silico predictions, all variants were classified as 'rare' based on allele frequencies in normal populations (allele frequency <0.01) and classified as 'deleterious' by at least one prediction method. The candidate pathogenic variants of these novel genes were found in 20 cases but not in paternal controls, nor in a local cohort of 2136 healthy men (online supplemental table 4). Moreover, the wild-type alleles showed a primate-wide evolutionary conservation based on the Ensembl sequence alignments of orthologous genes (online supplemental figure 4). Interestingly, a splice donor variant in *DMD* (NC_000023.11:g.32454661C>G) is located in a site highly conserved across vertebrates, ranging from mammals to birds, reptiles and fish. The cases U27 and U55 both had two candidate variants. Among the affected cases, four were bilateral and involved in genes *ARSH*, *DMD* and *SHROOM2*. The remaining 16 unilateral cases were affected by all four genes. Thus, the diagnostic rates for bilateral and unilateral cryptorchidism are 19.05% (4/21) and 17.02% (16/94), respectively. Together with the known genes, the overall diagnostic rates are 23.81% (5/21) and 28.72% (27/94) for bilateral and unilateral cryptorchidism, respectively.

Based on gnomAD records, five variants showed low allele frequencies in both the global and East Asian population, while the remaining 10 variants were newly discovered without information on allele frequency in databases (online supplemental table 4). We also found recurrent variants in multiple cases: NC_000023.11:g.32346044G>C; NP_000100.2:p.(Q1821E) (*DMD*, two patients), NC_000023.11:g.3033050C>G; NP_001011719.1:p.(P452A) (*ARSH*, three patients), NC_000023.11:g.151924064G>A; NP_001011550.1:p.(A134T) (*MAGEA4*, four patients) and NC_000023.11:g.9937537G>T; NP_001307593.1:p.(A166S) (*SHROOM2*, two patients). These candidate variants and genotypes were also validated using the Sanger sequencing for all samples (online supplemental figure 5).

Table 3 The significance levels of burden tests for novel candidate genes with rare variants

Gene	CMC	CMCFisherExact	CMCWald	Fp	Zeggini
ARSH	5.90E−20	1.19E−06	2.20E−05	4.12E−21	5.90E−20
DMD	2.01E−38	8.39E−15	3.54E−18	1.32E−52	2.01E−38
MAGEA4	3.11E−65	2.32E−22	7.75E−23	2.25E−18	1.48E−52
SHROOM2	5.41E−32	1.48E−11	6.67E−13	1.40E−46	5.41E−32

Burden tests support the statistical significance of novel candidate genes

To statistically test our newly identified candidate genes, we performed burden tests for rare variants with the RVTEST package.[34] We further used Bonferroni correction to account for multiple testing ($p=2.5\times10^{-6}$). We found that five tests supported the whole-genome significance of three genes *DMD*, *MAGEA4* and *SHROOM2* (table 3, $p<2.5\times10^{-6}$). For *ARSH*, four methods, except for the CMC (Combined and Multivariate Collapsing) Wald test, supported the genome-wide significance after Bonferroni correction ($p<2.5\times10^{-6}$).

The maternal origin of candidate variants of X-chromosome genes

We identified 20 candidate X-chromosomal variants from known and novel genes affecting at least two cases, based on variant screening results (online supplemental tables 3 and 4). There are two possibilities for the origin of these variants: de novo mutation during the early embryonic development of probands or during maternal oogenesis and the maternal X inheritance. For the de novo mutation hypothesis, we would not expect to find the variant in maternal genotypes. For the maternal X inheritance mode, we would detect heterozygotes in maternal samples.

In the seven trios with WGS data, the candidate variant NC_000023.11:g.151924064G>A; NP_001011550.1:p.(A134T) in *MAGEA4* was detected only in proband U8 (online supplemental table 4). The Sanger sequencing of these variants in trios indicated a maternal X-chromosome origin (figure 2A).

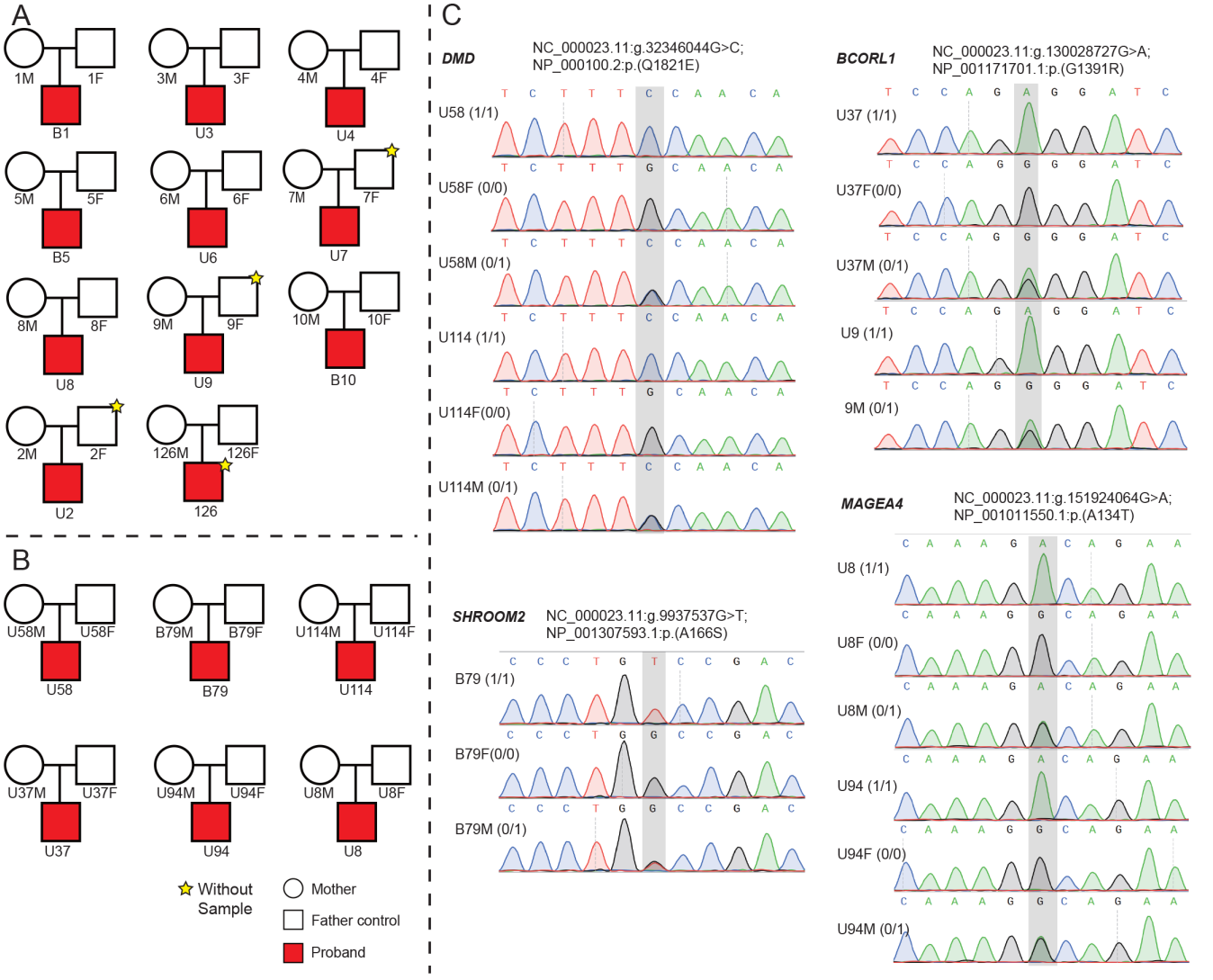

Figure 2 The pedigree information and the chromatogram of the Sanger sequencing results for cases and their parents. (A) The pedigree information for seven trios and four duos with whole-genome sequencing (WGS) data. (B) The pedigree information for six trios, with WGS data only for the cases. (C) The Sanger sequencing chromatogram for candidate variants of known and novel genes. '1/1', '0/1' and '0/0' are genotypes. U and B represent 'unilateral' and 'bilateral', respectively. 'F' and 'M' indicate 'Father' and 'Mother', respectively. '-F' and '-R' show the results of PCR sequencing from forward primer and reverse primer, respectively.

For the duo of U9 and 9 M samples, we also observed maternal heterozygotes in the X-chromosomal variant genotype, despite the unavailability of the paternal sample. For the additional trios with WGS-only probands, the Sanger sequencing of parental samples confirmed the maternal X inheritance, in which X hemizygotes, maternal heterozygotes and paternal wild types were found simultaneously in all trios (figure 2). Together, the Sanger sequencing of four variants, which affected seven cases and parents, supported the maternal X-chromosome origin exclusively, rather than de novo mutation from probands' embryogenesis or maternal germline. Future studies based on more pedigree data are needed to further evaluate the probability of de novo mutation during oogenesis of maternal germline or early development of probands.

No reliable compound heterozygous rare variants were detected

For compound heterozygous mode in autosomes, the family-based structure can facilitate tracing of the parental origin for each rare variant. We first focused on two heterozygous rare variants with disruptive effects on proteins for patients of the seven complete trio families. Then, we relaxed the criteria of variant impacts to cover the scenario of one heterozygous rare variant with disruptive effect and the other with moderate impact (sequence-altering). We did not find any gene fulfilling the requirements. Finally, we focused on the scenario of potential compound heterozygous mode, in which non-pedigree probands would carry at least two heterozygous rare variants in a gene with disruptive effects. We observed two heterozygous rare variants of *RPTN* with stop-gain and frameshift impacts in four unrelated patients (online supplemental table 5). However, the two variants had the same allele frequency in all superpopulations in the gnomAD database, suggesting that they are more likely to be inherited via the linkage disequilibrium, rather than the real compound heterozygous variants.

CRISPR-Cas9 mouse modelling of the splice donor variant in *DMD* (NC_000023.11:g.32454661C>G) and movement tests

Considering the highly conserved nature of the wild-type splice donor variant in *DMD* (NC_000023.11:g.32454661C>G, represented as 'C' in figure 3A,B), we generated transgenic mice with the mutant variant ('G') using the CRISPR-Cas9 editing technique(figure 3C). We examined cryptorchidism phenotypes in F2 male mice using MRI at weeks 4 and 10. In the *DMD* transgenic mice, we found undescended testicles that are different from the wild-type mice of the same age (figure 3D). The undescended testicles were located in the inguinal region, suggesting the involvement of the *DMD* gene in the development of cryptorchidism.

We further conducted movement tests on mutant transgenic and normal mice, using the rotarod, climbing, grip strength and treadmill running assessments (see online supplemental note and online supplemental table 6). We observed a progressive decrease in the time spent on the rotarod and an increase in the time taken to climb down a rough pole in transgenic mice at 8, 12, 16 and 20 weeks of age. Correspondingly, a gradual reduction in grip strength and treadmill running distance was also noted (figure 3E).

DISCUSSION

In this study, we conducted rare variants screening for cryptorchidism in humans. Among the known genes related to cryptorchidism, the mutation burden and rigorous filtering of rare and predicted deleterious variants support the roles of *USP9Y*, *KDM6A*, *BCORL1* and *UBA1* in the pathogenicity of cryptorchidism. *USP9Y* is one of the three genes within an azoospermia factor (AZFa) region. Previous studies have revealed that *USP9Y* showed increased transcript levels in patients suffering from cryptorchidism after a treatment with the gonadotropin-releasing hormone agonist GnRHa.[35] Exome sequencing has revealed the role of *BCORL1* in spermatogenesis.[36] *UBA1* has been reported in patients with spinal muscular atrophy 2 and cryptorchidism.[37] *KDM6A* is related to the Kabuki syndrome, which may involve hypospadias and cryptorchidism in males.[38] However, most of these studies were not designed specifically for investigating cryptorchidism-associated genes and variants, so our study expands the phenotypic spectrum of these gene defects.

More importantly, we identified four novel candidate cryptorchid genes (*DMD*, *ARSH*, *MAGEA4* and *SHROOM2*) with genome-wide significant burdens of rare mutations with both disruptive and sequence-altering impacts. The gene *MAGEA4* (melanoma antigen family A, 4) is commonly used as a marker for human spermatogonia.[39] Based on the immunohistochemistry of *MAGEA4*, the number of spermatogonia was decreased in cryptorchid testes compared with the normal testes.[40] Although poorly characterised, the gene *SHROOM2* was reported in patients with infertility.[41] Notably, the findings of the pathogenic variants of *DMD* are particularly interesting, because of the established role of *DMD* in producing the protein dystrophin critical for muscle development.[42] Clinical studies have revealed a positive correlation (OR, 2.83) between the urogenital malformation and muscular disorders, possibly due to muscular defects in the cremaster muscle or other inguinal tissues.[43] *DMD* defects have been reported in Duchenne/Becker muscular dystrophy.[44 45] Given that DMD typically manifests between the ages of 3 and 5 years, and the paediatric cases of cryptorchidism could be identified in children under the age of 1 year, ongoing monitoring and longitudinal follow-up may be valuable. Considering the dynamic process involving the muscular traction of the gubernaculum during the normal testes descendance,[46] our study suggests that muscular abnormalities may play a major role in cryptorchidism. An alternative hypothesis for future investigation could be that the mutation of *DMD* (NC_000023.11:g.32454661C>G) might be implicated in both DMD and cryptorchidism, whereas other variants identified in this study could be associated with an increased susceptibility to cryptorchidism.

As the maternal effect is higher than the paternal effect for cryptorchidism (eg, the recurrence rate is higher in maternal half-brothers than in paternal half-brothers), the possibility of maternal inheritance has long been suspected.[9 47] In this study, we revealed the genetic basis of maternal inheritance of cryptorchidism by finding the dominant role of X-chromosomal genes and their maternal origin. Indeed, based on the pedigree-based Sanger sequencing, we revealed that some X-chromosomal variants with available parental samples were inherited from maternal rather than paternal lineages. This result is consistent with the expectation of both the X-hemizygosity effect and the nature of male-specific disease. Because of X-hemizygosity in males, X-chromosomal pathogenic variants invariably affect the males, while commonly being masked by a normal allele in heterozygous females. In addition, the causative variant for a male-specific trait would affect males more heavily than females, except in case of a pleiotropic effect on both sexes. Clinically, such hereditary pattern could be used to screen those sex-related genetic disorders in male humans.

Together, among the total of eight known and novel genes affecting multiple unrelated patients of cryptorchidism, seven

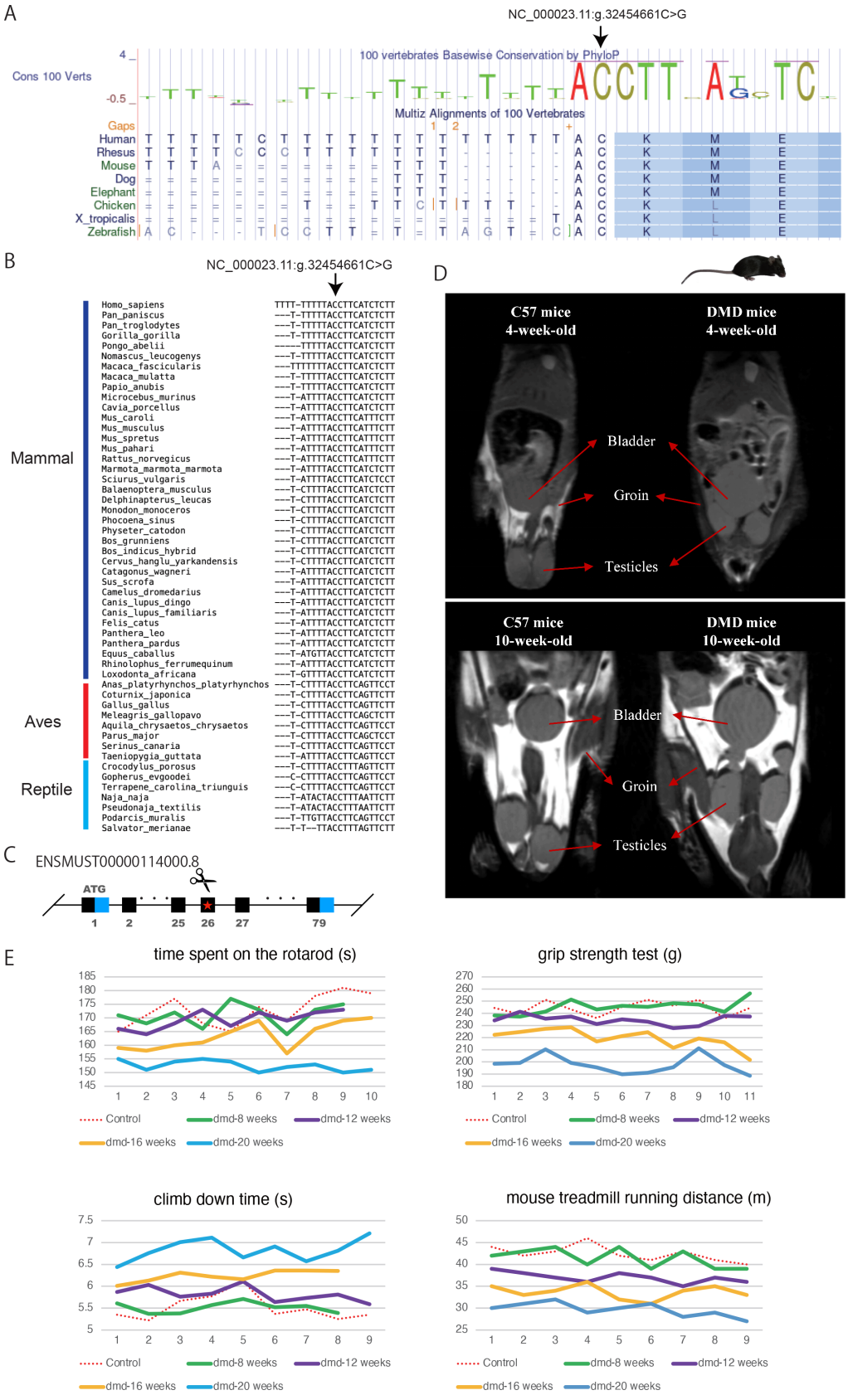

Figure 3 Evolutionary conservation of the splice donor variant in DMD (NC_000023.11:g.32454661C>G) and MRI of transgenic and normal mice. (A) Conservation of the DMD variant (NC_000023.11:g.32454661C>G) according to the UCSC (The University of California, Santa Cruz) Vertebrate Alignment. (B) Conservation based on the Amniote alignment from the Ensembl database. (C) Schematic representation of a transgenic mouse with CRISPR-Cas9 editing. (D) Position of the testicles in normal C57 mice compared to DMD transgenic mice, as determined by MRI. Undescended testicles were observed in the DMD transgenic mice at both 4 and 10 weeks, with both testicles located in the inguinal region. (E) Movement test results for transgenic and normal mice (for details, please see online supplemental note and online supplemental table 6). DMD, Duchenne muscular dystrophy.

are X-chromosomal genes and one is a Y gene. The candidate variants of X-chromosomal genes follow the expectation of maternal inheritance rather than de novo mutation. Thus, our study revealed a predominant role of sex chromosomal, and particularly the X-chromosomal, gene defects, in causing male cryptorchidism.

CONCLUSION

We identified rare pathogenic variants of four known candidate genes (*USP9Y*, *UBA1*, *BCORL1* and *KDM6A*) and four novel candidate genes (*ARSH*, *DMD*, *MAGEA4* and *SHROOM2*) in patients suffering from cryptorchidism. Considering the chromosomal distribution, seven out of eight genes are within the X-chromosome, and *USP9Y* is located on the Y-chromosome. This reveals a predominant role of X-chromosomal genes in cryptorchidism. The maternal origin of these X-chromosomal variants reflects the strong effect of X-hemizygosity on male-specific diseases. We successfully replicated cryptorchidism phenotypes and DMD in transgenic mice. These mice harbour a splice donor loss variant in *DMD* (NC_000023.11:g.32454661C>G), establishing a viable mouse model for future research and therapy studies of this condition.

MATERIALS AND METHODS

DNA samples, patient background and WGS

DNA was extracted from peripheral whole blood of 115 patients and 19 parental controls, using a local database of variants from 2136 unaffected men for subsequent comparison. The patients were diagnosed by paediatric specialists of the West China Hospital (WCH) and West China Second University Hospital (WCSUH). The parents were carefully inquired about family history and all fathers denied the existence of cryptorchidism in the family history. The WGS (150 bp paired-end) data of an insert size of 350 bp were sequenced using the DNBSEQ-T7 platform (MGI), according to the manufacturer's protocol (online supplemental table 1).

Variants calling, genotyping and annotation

Based on the high-performance computing system, we locally conducted the variant calling, genetic relationship and population ancestry using the pipeline described previously.[48–50] The rare variants were defined as alleles with frequency lower than 0.01 in all geographic human populations of gnomAD v3.1. Pathogenic variants were identified for known and novel genes (online supplemental note).

Validation of variants and genotypes using the Sanger sequencing

Following the identification of candidate rare pathogenic variants, the Sanger sequencing was conducted for all probands at first to remove false positive variants due to errors in NGS (nex generation sequencing) or calling process. For the validated variants in probands, we further conducted the Sanger sequencing for all available parental samples, including samples from additional six trios in which only probands had WGS data. The parental genotypes were further used to evaluate the origin of variants.

The C56Bl/6J transgenic mouse modelling

For the highly conserved variant, transgenic mouse was modelled using the CRISPR-Cas9 editing technique. Briefly, the process was carried out in two main stages: the in vitro stage and the in vivo stage. In the in vitro stage, the process began with the design and construction of guide RNAs (gRNAs) specific to the target DNA sequence around variant, followed by the design and construction of the donor vector, which carried the desired genetic modification. The in vivo stage involved microinjecting the designed gRNA and donor vector into fertilised eggs and then transplanting these embryos into surrogate mothers. This led to the birth of the F0 generation mice, which were subsequently screened to identify individuals with successful genetic modifications. Positive F0 generation mice were then bred to produce the F1 generation, and these offspring were also screened to confirm the presence of the genetic modification. The F2 mice were examined for phenotypes with MRI (Time Medical Systems, USA). Three transgenic mice and one control were used for movement tests (see online supplemental note).

Author affiliations
[1]Institutes for Systems Genetics, Frontiers Science Center for Disease-related Molecular Network, West China Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China
[2]Department of Ecology and Evolution, The University of Chicago, Chicago, Illinois, USA
[3]Department of Chemistry, The University of Chicago, Chicago, Illinois, USA
[4]Institute of Rare Diseases, West China Hospital, Sichuan University, Chengdu, Sichuan, People's Republic of China
[5]Department of Radiology, Key Laboratory of Birth Defects and Related Diseases of Women and Children of Ministry of Education, West China Second University Hospital, Sichuan University, Chengdu, Sichuan, China
[6]Department of Human and Molecular Genetics, Baylor College of Medicine, Houston, Texas, USA
[7]Laboratory of Molecular Translational Medicine, Center for Translational Medicine, Key Laboratory of Birth Defects and Related Diseases of Women and Children (Sichuan University), Ministry of Education, Clinical Research Center for Birth Defects of Sichuan Province, West China Second University Hospital, Sichuan University, Chengdu, Sichuan, China
[8]Department of Pediatric Surgery, West China Hospital, Sichuan University, Chengdu, Sichuan, China
[9]Department of Biological Sciences, Wayne State University, Detroit, Michigan, USA

Correction notice This article has been corrected since it was published online. Professor Chuanzhu Fan's affiliation has been amended.

Acknowledgements We acknowledge the computing support from the West China Biomedical Big Data Center and the Med-X Center for Informatics of Sichuan University. We express our gratitude to the anonymous reviewers for their insightful comments and extend special thanks to Dr Chunyu Li from West China Hospital for her invaluable discussion and advice.

Contributors HX, CF, LZ and FL supervised this work and HX is the guarantor. JC, JZ, KZ and HD designed the research. JC, JZ and GH analysed the data. JC, YJ and JZ conducted the lab work. GH provided and scrutinised the candidate rare variants screening in 2136 normal controls. CF, LZ and HX conducted all clinical works. KZ and HX conducted movement tests. JC drafted the manuscript and all authors took part in revising and improving the manuscript.

Funding This study was supported by the Short-Term Expert Fund of West China Hospital (139190032), the fifth batch of technological innovation research projects in Chengdu (2021-YF05-01331-SN), the Postdoctoral Research and Development Fund of West China Hospital of Sichuan University (2020HXBH087), Fund of Sichuan Provincial Department of Science and Technology (2021YFS0244) and Fund of Sichuan Provincial Department of Science and Technology (2021YFS0026).

Competing interests None declared.

Patient consent for publication Not applicable.

Ethics approval This study conformed with the Helsinki Declaration of 1975 (as revised in 2008) concerning Human and Animal Rights. All participating parents provided informed consent, and this study was formally approved by the ethics committees of WCH (registration number: 2021389) and WCSUH (registration number: 2021389). Participants gave informed consent to participate in the study before taking part.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

ORCID iD

Jianhai Chen http://orcid.org/0000-0003-0093-2003

REFERENCES

1 Barthold JS, González R. The epidemiology of congenital cryptorchidism, testicular ascent and orchiopexy. *J Urol* 2003;170:2396–401.
2 Askeland EJ, Arlen AM, Erickson BA, *et al*. Urological manifestations of Duchenne muscular dystrophy. *The Journal of Urology* 2013;190:S0022-5347(13)00095-5:1523–8:.
3 Ferguson L, Agoulnik AI. Testicular cancer and cryptorchidism. *Front Endocrinol* 2013;4.
4 Raivio T, Toppari J, Kaleva M, *et al*. Serum androgen bioactivity in cryptorchid and noncryptorchid boys during the postnatal reproductive hormone surge. *J Clin Endocrinol Metab* 2003;88:2597–9.
5 Elamo HP, Virtanen HE, Toppari J. Genetics of cryptorchidism and testicular regression. *Best Pract Res Clin Endocrinol Metab* 2022;36:101619.
6 Chacko JK, Barthold JS. Genetic and environmental contributors to cryptorchidism. *Pediatr Endocrinol Rev* 2009;6:476–80.
7 Elert A, Jahn K, Heidenreich A, *et al*. Population-based investigation of familial undescended testis and its association with other urogenital anomalies. *J Pediatr Urol* 2005;1:403–7.
8 Promm M, Schröder A, Neissner C, *et al*. Acquired cryptorchidism: more harm than thought. *J Pediatr Urol* 2016;12:236.
9 Schnack TH, Zdravkovic S, Myrup C, *et al*. Familial aggregation of cryptorchidism—a nationwide cohort study. *Am J Epidemiol* 2008;167:1453–7.
10 Barthold JS, Wang Y, Kolon TF, *et al*. Pathway analysis supports association of nonsyndromic cryptorchidism with genetic Loci linked to cytoskeleton-dependent functions. *Hum Reprod* 2015;30:2439–51.
11 Bloom JS, Boocock J, Treusch S, *et al*. Rare variants contribute disproportionately to quantitative trait variation in yeast. *Elife* 2019;8:e49212.
12 UK10K Consortium, Walter K, Min JL, *et al*. The Uk10K project identifies rare variants in health and disease. *Nature* 2015;526:82–90.
13 Maher MC, Uricchio LH, Torgerson DG, *et al*. Population genetics of rare variants and complex diseases. *Hum Hered* 2012;74:118–28.
14 Halvorsen M, Huh R, Oskolkov N, *et al*. Increased burden of ultra-rare structural variants localizing to boundaries of topologically associated domains in schizophrenia. *Nat Commun* 2020;11:1842.
15 Kryukov GV, Pennacchio LA, Sunyaev SR. Most rare Missense alleles are deleterious in humans: implications for complex disease and Association studies. *Am J Hum Genet* 2007;80:727–39.
16 Abrahams L, Savisaar R, Mordstein C, *et al*. Evidence in disease and non-disease contexts that nonsense mutations cause altered splicing via motif disruption. *Nucleic Acids Res* 2021;49:9665–85.
17 Ganna A, Satterstrom FK, Zekavat SM, *et al*. Quantifying the impact of rare and ultra-rare coding variation across the phenotypic spectrum. *Am J Hum Genet* 2018;102:1204–11.
18 Li J, Kong N, Han B, *et al*. Rare variants regulate expression of nearby individual genes in multiple tissues. *PLOS Genet* 2021;17:e1009596.
19 Gorlov IP, Gorlova OY, Frazier ML, *et al*. Evolutionary evidence of the effect of rare variants on disease etiology. *Clin Genet* 2011;79:199–206.
20 Wang J, Lu Y, Yan X, *et al*. Identification of novel MITF mutations in Chinese families with Waardenburg syndrome type II. *Mol Genet Genomic Med* 2021;9:e1770.
21 Bu F, Zhong M, Chen Q, *et al*. Dvpred: a disease-specific prediction tool for variant pathogenicity classification for hearing loss. *Hum Genet* 2022;141:401–11.
22 Cheng J, Zhou X, Lu Y, *et al*. Exome sequencing identifies a novel frameshift mutation of Myo6 as the cause of autosomal dominant nonsyndromic hearing loss in a Chinese family. *Ann Hum Genet* 2014;78:410–23.
23 Sun Y, Cheng J, Lu Y, *et al*. Identification of two novel Missense Wfs1 mutations, H696Y and R703H, in patients with non-syndromic low-frequency sensorineural hearing loss. *J Genet Genomics* 2011;38:71–6.
24 Cheng J, Zhu Y, He S, *et al*. Functional mutation of SMAC/DIABLO, encoding a mitochondrial proapoptotic protein, causes human progressive hearing loss Dfna64. *Am J Hum Genet* 2011;89:56–66.
25 Yuan Y, Huang D, Yu F, *et al*. A de novo Gjb2 (Connexin 26) mutation, R75W, in a Chinese pedigree with hearing loss and Palmoplantar Keratoderma. *Am J Med Genet A* 2009;149A:689–92.
26 Ayers K, Kumar R, Robevska G, *et al*. Familial bilateral cryptorchidism is caused by recessive variants in Rxfp2 *J Med Genet* 2019;56:727–33.
27 Belkadi A, Bolze A, Itan Y, *et al*. Whole-genome sequencing is more powerful than whole-exome sequencing for detecting exome variants. *Proc Natl Acad Sci USA* 2015;112:5473–8.
28 Hiatt SM, Thompson ML, Prokop JW, *et al*. Deleterious variation in Brsk2 Associates with a neurodevelopmental disorder. *Am J Hum Genet* 2019;104:701–8.
29 McDougal WS, Wein AJ, Kavoussi LR, *et al*. Campbell-Walsh Urology 11th Edition Review E-Book. Amsterdam: Elsevier Health Sciences, 2015.
30 Sun Y, Liu F, Fan C, *et al*. Characterizing sensitivity and coverage of clinical WGS as a diagnostic test for genetic disorders. *BMC Med Genomics* 2021;14:102.
31 Kishikawa T, Momozawa Y, Ozeki T, *et al*. Empirical evaluation of variant calling accuracy using ultra-deep whole-genome sequencing data. *Sci Rep* 2019;9:1784.
32 Canese K, Weis S. Pubmed: the bibliographic database. *The NCBI Handbook* 2013;2.
33 McLaren W, Gil L, Hunt SE, *et al*. The ensembl variant effect Predictor. *Genome Biol* 2016;17:122:122:.
34 Zhan X, Hu Y, Li B, *et al*. RVTESTS: an efficient and comprehensive tool for rare variant Association analysis using sequence data. *Bioinformatics* 2016;32:1423–6.
35 Gegenschatz-Schmid K, Verkauskas G, Stadler MB, *et al*. Genes located in Y-chromosomal regions important for male fertility show altered transcript levels in cryptorchidism and respond to curative hormone treatment. *Basic Clin Androl* 2019;29:8.
36 Lu C, Zhang Y, Qin Y, *et al*. Human X chromosome exome sequencing identifies bcorl1 as contributor to spermatogenesis. *J Med Genet* 2021;58:56–65.
37 Öztürk Ö, Çavdartepe BE, Bağış H. X-linked spinal muscular atrophy 2 due to a synonymous variant in the Uba1 gene in a family with novel findings from Turkey. *Mol Syndromol* 2022;13:246–53.
38 Armstrong L, Abd El Moneim A, Aleck K, *et al*. Further delineation of Kabuki syndrome in 48 well-defined new individuals. *Am J Med Genet A* 2005;132A:265–72.
39 He Z, Kokkinaki M, Jiang J, *et al*. Isolation, characterization, and culture of human spermatogonial. *Biol Reprod* 2010;82:363–72.
40 Yang S, Ping P, Ma M, *et al*. Generation of haploid spermatids with fertilization and development capacity from human spermatogonial stem cells of cryptorchid patients. *Stem Cell Reports* 2014;3:663–75.
41 Fabian-Jessing BK, Vestergaard EM, Plomp AS, *et al*. Ocular albinism with infertility and late-onset sensorineural hearing loss. *Am J Med Genet A* 2018;176:1587–93.
42 Duan D, Goemans N, Takeda S, *et al*. Duchenne muscular dystrophy. *Nat Rev Dis Primers* 2021;7:13.
43 Diniz G, Barutcuoglu M, Unalp A, *et al*. Evaluation of the relationship between urogenital abnormalities and neuromuscular disorders evaluation of the relationship between urogenital abnormalities and. *Eastern J Med* 2008;13:19–24.
44 Moser H. Duchenne muscular dystrophy: pathogenetic aspects and genetic prevention. *Hum Genet* 1984;66:17–40.
45 Emery AE, Muntoni F, Quinlivan R. Duchenne muscular dystrophy: Oxford monographs on medical G. 2015.
46 Hutson JM, Hasthorpe S, Heyns CF. Anatomical and functional aspects of testicular descent and cryptorchidism. *Endocr Rev* 1997;18:259–80.
47 Jensen MS, Toft G, Thulstrup AM, *et al*. Cryptorchidism concordance in monozygotic and dizygotic twin brothers, full brothers, and half-brothers. *Fertil Steril* 2010;93:124–9.
48 Chen J, Zhang P, Chen H, *et al*. Whole-genome sequencing identifies rare Missense variants of Wnt16 and ERVW-1 causing the systemic lupus erythematosus. *Genomics* 2022;114.
49 Oud MS, Houston BJ, Volozonoka L, *et al*. Exome sequencing reveals variants in known and novel candidate genes for severe sperm motility disorders. *Hum Reprod* 2021;36:2597–611.
50 Jia Y, Chen J, Zhong J, *et al*. Novel rare Mutation in a conserved site of PTPRB causes human Hypoplastic left heart syndrome. *Clin Genet* 2023;103:79–86.