

1 **Novel mutation leading to splice donor loss in a conserved site of *DMD***  
2 **causes cryptorchidism**

3 Jianhai Chen<sup>1,2,#</sup>, Yangying Jia<sup>1,2</sup>, Jie Zhong<sup>1</sup>, Kun Zhang<sup>3</sup>, Hongzheng Dai<sup>4</sup>,  
4 Guanglin He<sup>5</sup>, Fuping Li<sup>6,#</sup>, Li Zeng<sup>7,#</sup>, Chuanzhu Fan<sup>8,#</sup>, Huayan Xu<sup>3,#</sup>

5 <sup>1</sup>Department of Ecology and Evolution, The University of Chicago, 1101E 57th  
6 Street, Chicago, Illinois 60637, USA

7 <sup>2</sup>Institutes for Systems Genetics, Frontiers Science Center for Disease-related  
8 Molecular Network, West China Hospital, Sichuan University, Chengdu  
9 610041, P.R. China

10 <sup>3</sup>Department of Radiology, Key Laboratory of Birth Defects and Related  
11 Diseases of Women and Children of Ministry of Education, West China  
12 Second University Hospital, Sichuan University, Chengdu 610041, P.R. China

13 <sup>4</sup>Department of Molecular and Human Genetics, Baylor College of Medicine,  
14 Houston, Texas 77030, USA

15 <sup>5</sup>Institute of Rare Diseases, West China Hospital of Sichuan University,  
16 Sichuan University, Chengdu, 610044 China

17 <sup>6</sup>Laboratory of Molecular Translational Medicine, Center for Translational  
18 Medicine, Key Laboratory of Birth Defects and Related Diseases of Women  
19 and Children (Sichuan University), Ministry of Education, Clinical Research  
20 Center for Birth Defects of Sichuan Province, West China Second University  
21 Hospital, Sichuan University, Chengdu, 610041, P.R. China

22 <sup>7</sup>The Department of Pediatric Surgery, West China Hospital, Sichuan  
23 University, Chengdu 610041, China

24 <sup>8</sup>Department of Biological Sciences, Wayne State University, Detroit, Michigan,  
25 48202, USA

26 #Corresponding author: [jianhaichen@uchicago.edu](mailto:jianhaichen@uchicago.edu); [xuhuayan89@sina.com](mailto:xuhuayan89@sina.com);  
27 [zengli998@163.com](mailto:zengli998@163.com); [ex3922@wayne.edu](mailto:ex3922@wayne.edu); [lfpSnake@scu.edu.cn](mailto:lfpSnake@scu.edu.cn)

28

29 **Abstract**

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

35 **Methods** In this study, a cohort of 115 Chinese probands with cryptorchidism  
36 was analyzed using whole-genome sequencing (WGS), alongside 19 parental  
37 controls and 2136 unaffected men. Additionally, CRISPR-Cas9 editing of a  
38 conserved variant was performed in a mouse model, with MRI screening  
39 utilized to observe the phenotype.

40 **Results** In 30 of 115 patients (26.1%), we identified four novel genes (*ARSH*,  
41 *DMD*, *MAGEA4*, and *SHROOM2*) affecting at least five unrelated patients and  
42 four known genes (*USP9Y*, *UBA1*, *BCORL1*, and *KDM6A*) with the candidate  
43 rare pathogenic variants affecting at least two cases. Burden tests of rare  
44 variants revealed the genome-wide significances for newly identified genes ( $p$   
45  $< 2.5 \times 10^{-6}$ ) under the Bonferroni correction. Surprisingly, novel and known  
46 genes were mainly from X chromosome (seven on X and one on Y) and all  
47 rare X-chromosomal segregating variants exhibited a maternal inheritance  
48 rather than *de novo* origin. CRISPR-Cas9 mouse modeling of a splice donor  
49 loss variant in *DMD* (NC\_000023.11:g.32454661C>G), which resides in a  
50 conserved site across vertebrates, replicated bilateral cryptorchidism  
51 phenotypes, confirmed by Magnetic resonance imaging (MRI) at 4 and 10  
52 weeks.

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

57 **Keywords:** molecular diagnosis; WGS; rare variants; male-specific disease;  
58 male hemizygosity

## 59 **Introduction**

60 Cryptorchidism (also known as 'hidden testicle', OMIM 219050) has a global  
61 prevalence rate of 2% to 4% at full-term birth in boys <sup>1</sup>. Despite its relatively  
62 low prevalence, cryptorchidism is the most common birth defect involving  
63 urogenital abnormalities in newborn boys. It is also the best-characterized risk  
64 factor for infertility and testicular cancer <sup>2</sup>. Normally, the testes spontaneously  
65 descend into the scrotum by the eighth month (33 weeks) of pregnancy, likely  
66 triggered by transient activation of the hypothalamus-pituitary-gonadal (HPG)  
67 axis, leading to an increase in reproductive hormone levels <sup>3,4</sup>. In patients with  
68 cryptorchidism, one or both testes may fail to descend completely into the  
69 dependent part of the scrotum, resulting in unilateral or bilateral  
70 cryptorchidism, respectively.

71 Maternal risk factors before and during pregnancy, such as maternal smoking,  
72 endocrine-disrupting chemicals, and gestational diabetes, have increasingly  
73 been recognized as the risk factors for congenital cryptorchidism <sup>5</sup>. However,  
74 the studies on genetic risk factors of cryptorchidism are still limited. Multiple  
75 etiological features, such as familial aggregation and increased prevalence in

76 first-degree relatives, suggest a genetic predisposition to this disease<sup>6</sup>. For  
77 instance, a family history of cryptorchidism can increase the risk of the  
78 condition in newborn boys by approximately threefold<sup>7,8</sup>. The recurrence risk  
79 has been found to be more than twice as high in brothers of cryptorchidism  
80 cases compared to healthy controls<sup>9</sup>. Thus, despite cryptorchidism's complex  
81 susceptibility to maternal and environmental factors, strong genetic  
82 components may also contribute to its etiology.

83 Previous microarray studies based on common variants (GWAS) have  
84 suggested a potential association of certain pathways with cryptorchidism but  
85 failed to identify specific genes, likely due to the small effect size of common  
86 variants<sup>10</sup>. Indeed, both population genetics theory and empirical data  
87 suggest that rare variants have much larger effect sizes than common  
88 variants<sup>11-14</sup>, indicating their significant role in causing genetic diseases. It is  
89 estimated that most human rare protein-altering variants (missense) are  
90 pathogenic<sup>15</sup>. Furthermore, disruptive mutations (nonsense) are  
91 disproportionately common causal factors in about 11.5% of human genetic  
92 diseases<sup>16</sup>. At the population level, these causal variants are rare in allele  
93 frequency due to strong purifying selection, which limits their accumulation  
94 across generations. Therefore, rare deleterious variants with large genetic  
95 effect sizes represent a promising area for discovering causative genes for  
96 genetic diseases<sup>17</sup>, especially for diseases affecting fertility, such as  
97 cryptorchidism.

98 The application of various genetic screening techniques, including whole-  
99 genome and exome sequencing (WGS/WES), has facilitated the precise

100 diagnosis of numerous genetic diseases caused by rare variants<sup>18-25</sup>. For  
101 instance, a WES study on an Indian family identified a rare deleterious  
102 homozygous missense variant in the *RXFP2* gene, which causes  
103 cryptorchidism<sup>26</sup>. Although WES is more commonly used than WGS due to its  
104 lower cost, WGS offers a more uniform distribution of sequencing quality  
105 (including single-nucleotide variants, insertions, and deletions) and a higher  
106 discovery rate of coding-region variants, approximately 3%<sup>27 28</sup>.

107 To date, a rare variants study of cryptorchidism based on WGS is still lacking.  
108 In this study, we conducted rare variants screening based on WGS data for  
109 134 Han Chinese samples including cases (115) and family controls (19). We  
110 also used a local cohort of 2136 unaffected men for comparison. We identified  
111 candidate rare pathogenic variants based on autosomal recessive, compound  
112 heterozygous, and X-linked recessive inheritance mode for cryptorchidism  
113 patients. We annotated these variants and recognized both known and novel  
114 candidate genes and variants. We also confirmed the role of a novel rare  
115 mutation in *DMD* (NC\_000023.11:g.32454661C>G) in causing cryptorchidism  
116 based on transgenic mouse modeling. Our study may facilitate the  
117 development of molecular diagnosis and precision medicine for patients  
118 suffering from cryptorchidism.

## 119 **Results**

120 **The MRI on typical bilateral and unilateral cryptorchidism and basic**  
121 **sequencing statistics.**

122 The diagnosis of cryptorchidism was based on a physical exam by  
123 professional pediatric urologists following standard protocols <sup>29</sup>. Typical  
124 bilateral cryptorchidism and unilateral cryptorchidism of two patients were  
125 shown using the magnetic resonance imaging (MRI) (Figure 1a-1d). In  
126 comparison to controls, the patient with bilateral cryptorchidism had two non-  
127 descended testes, while the patient with unilateral cryptorchidism had only  
128 one undescended testis (Figure 1).

129 We conducted WGS on 115 cases and 19 family controls. Among the cases,  
130 there were 21 (18.26%) bilateral cryptorchidism patients and 94 (81.74%)  
131 unilateral cryptorchidism patients. This dataset also includes three family trios  
132 of bilateral cryptorchidism, four trios of unilateral cryptorchidism, and three  
133 duos with only cryptorchidism cases and mothers. Based on previous  
134 sensitivity tests, 15x WGS depth/fold can achieve accurate SNV calling, while  
135 30x is sufficient for calling indels <sup>30 31</sup>. In this study, the average sequencing  
136 depths were 33.03-fold for cases and 27.97-fold for controls, suggesting a  
137 balanced design of sequencing depth and quality for all samples  
138 (Supplementary Table 1, Figure 1f and Supplementary Figure 1a). To confirm  
139 the reported relationship, we used whole-genome SNPs to estimate the  
140 relatedness between family members. We also confirmed the population  
141 ancestry of all samples based on the principal component analysis (PCA).  
142 Based on identity by state (IBS) distances among family members, we  
143 confirmed that family members are genetically closer than unrelated  
144 individuals (Supplementary Figure 1b). We also confirmed that all parents-son  
145 relationships are within the coefficient range of first-degree relatives (from  
146 0.177 to 0.354). By incorporating individuals from the “1000 genomes” project,

147 we further revealed that all newly sequenced samples are closely related to  
148 the East Asian population (Supplementary Figure 1c).

### 149 **The Rare Candidate Pathogenic Variants for Known Genes**

150 With the keywords ‘cryptorchidism variants’ and ‘cryptorchidism genetics’, we  
151 conducted the initial screening for known genes in the PubMed literature  
152 database<sup>32</sup>. We retrieved 56 genes previously reported to be associated with  
153 cryptorchidism (Supplementary Table 2). Most of these previous studies were  
154 based on the genetic screening of small patient samples. Thus, it is  
155 interesting to know whether these genes can be confirmed in our cohort. We  
156 found that 41 of these genes were also registered in the Online Mendelian  
157 Inheritance in Man (OMIM) and Human Phenotype Ontology (HPO) database  
158 (June-2022) as genes related to cryptorchidism. We also added 738 genes  
159 associated with cryptorchidism phenotypes from these databases. These 779  
160 genes (738+41) were used as a “pool” of known genes (Supplementary Table  
161 2). We annotated non-redundant variants using the online tool VEP in the  
162 Ensembl database<sup>33</sup> and then summarized both known and novel candidate  
163 genes.

164 We detected four known genes (*BCORL1*, *KDM6A*, *UBA1*, and *USP9Y*) with  
165 candidate rare pathogenic variants present in at least two cases but absent in  
166 paternal controls (Table 1 and Supplementary Table 3). Six variants were not

167 detected in all publicly available population databases (gnomAD, 1000  
168 genomes, dbSNP, ExAC, ClinVar, etc.). One variant, *BCORL1*  
169 (NC\_000023.11:g.130028727G>A;NP\_001171701.1:p.(G1391R)), was also  
170 classified as rare based on the allele frequency of all human populations, with  
171 the highest frequency in East Asian population (0.005439; Supplementary  
172 Table 3). Although this allele frequency is relatively higher in the East Asian  
173 population than globally, it remains much lower than the global prevalence  
174 rate of cryptorchidism in boys, which is 2%-4%.

175 Interestingly, all known genes affecting at least two cases are on sex  
176 chromosomes. Not a single autosomal gene was associated with more than  
177 one case. This strongly suggests a predominant role of sex chromosomal  
178 genes in cryptorchidism pathogenicity. The Y-chromosome gene *USP9Y*  
179 (Ubiquitin Specific Peptidase 9 Y-Linked) was detected in six boys with  
180 unilateral cryptorchidism. In addition, three X-chromosome genes (*BCORL1*,  
181 *UBA1*, and *KDM6A*) were found to carry candidate rare pathogenic variants  
182 affecting at least two patients. Strikingly, the two rare pathogenic variants in  
183 *BCORL1* (NC\_000023.11:g.130028727G>A; NP\_001171701.1:p.(G1391R))  
184 and *USP9Y* (NC\_000024.10:g.12739592C>G; NP\_004645.2:p.(S462C)) were  
185 found in four unilateral patients per variant, with only the former registered in  
186 the gnomAD database. Both variants were predicted to be deleterious by  
187 multiple *in silico* algorithms (Supplementary Table 3). In addition, the wild-type



188 alleles of these variants showed evolutionary conservation in primate species,  
189 based on sequence alignments of primate orthologous genes from the  
190 Ensembl database (v105, Supplementary Figure 2). Together, the candidate  
191 rare pathogenic variants of known genes were found to affect 12 cases,  
192 indicating a molecular diagnostic rate of 10.4% (12/115). The diagnostic rates  
193 for bilateral and unilateral cryptorchidism were 0% (0/21) and 12.77% (12/94),  
194 respectively. These candidate pathogenic variants and genotypes in cases  
195 and parents were validated using the Sanger Sequencing (Supplementary  
196 Figure 3).

### 197 **The Rare Candidate Pathogenic Variants for Novel Genes**

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

205 Based on allele frequencies in multiple genome databases (gnomAD, 1000  
206 genomes, dbSNP, ExAC, ClinVar, etc.) and *in silico* predictions, all variants  
207 were classified as 'rare' based on allele frequencies in normal populations

208 (allele frequency < 0.01) and classified as 'deleterious' by at least one  
209 prediction method. The candidate pathogenic variants of these novel genes  
210 were found in 20 cases but not in paternal controls, nor in a local cohort of  
211 2136 healthy men (Supplementary Table 4). Moreover, the wild-type alleles  
212 showed a primate-wide evolutionary conservation based on the Ensembl  
213 sequence alignments of orthologous genes (Supplementary Figure 4).  
214 Interestingly, a splice donor variant in *DMD* (NC\_000023.11:g.32454661C>G)  
215 is located in a site highly conserved across vertebrates, ranging from  
216 mammals to birds, reptiles, and fish (Figure 3). The cases U27 and U55 both  
217 had two candidate variants. Among the affected cases, four were bilateral and  
218 involved in genes *ARSH*, *DMD*, and *SHROOM2*. The remaining 16 unilateral  
219 cases were affected by all four genes. Thus, the diagnostic rates for bilateral  
220 and unilateral cryptorchidism are 19.05% (4/21) and 17.02% (16/94),  
221 respectively. Together with the known genes, the overall diagnostic rates are  
222 23.81% (5/21) and 28.72% (27/94) for bilateral and unilateral cryptorchidism,  
223 respectively.

224 Based on gnomAD records, five variants showed low allele frequencies in  
225 both the global and East Asian population, while the remaining ten variants  
226 were newly discovered without information on allele frequency in databases  
227 (Supplementary Table 4). We also found recurrent variants in multiple cases:  
228 NC\_000023.11:g.32346044G>C; NP\_000100.2:p.(Q1821E) (*DMD*, two

229 patients), NC\_000023.11:g.3033050C>G; NP\_001011719.1:p.(P452A)  
230 (*ARSH*, three patients), NC\_000023.11:g.151924064G>A;  
231 NP\_001011550.1:p.(A134T) (*MAGEA4*, four patients), and  
232 NC\_000023.11:g.9937537G>T; NP\_001307593.1:p.(A166S) (*SHROOM2*, two  
233 patients). These candidate variants and genotypes were also validated using  
234 the Sanger Sequencing for all samples (Supplementary Figure 5).

### 235 **Burden Tests Support the Statistical Significance of Novel Candidate** 236 **genes**

237 To statistically test our newly identified candidate genes, we performed  
238 burden tests for rare variants with the RVTEST package<sup>34</sup>. We further used  
239 Bonferroni correction to account for multiple testing ( $p = 2.5 \times 10^{-6}$ ). We found  
240 that five tests supported the whole-genome significance of three genes *DMD*,  
241 *MAGEA4*, and *SHROOM2* (Table 3,  $p < 2.5 \times 10^{-6}$ ). For *ARSH*, four methods,  
242 except for the CMC Wald test, supported the genome-wide significance after  
243 Bonferroni correction ( $p < 2.5 \times 10^{-6}$ ).

### 244 **The Maternal Origin of Candidate Variants of X-chromosome Genes**

245 We identified 20 candidate X-chromosomal variants from known and novel  
246 genes affecting at least two cases, based on variant screening results  
247 (Supplementary Tables 3 and 4). There are two possibilities for the origin of  
248 these variants: *de novo* mutation during the early embryonic development of  
249 probands or during maternal oogenesis and the maternal X inheritance. For  
250 the *de novo* mutation hypothesis, we would not expect to find the variant in

251 maternal genotypes. For the maternal X inheritance mode, we would detect  
252 heterozygotes in maternal samples.

253 In the seven trios with WGS data, the candidate variant  
254 NC\_000023.11:g.151924064G>A; NP\_001011550.1:p.(A134T) in *MAGEA4*  
255 was detected only in proband U8 (Supplementary Table 4). The Sanger  
256 Sequencing of these variants in trios indicated a maternal X-chromosome  
257 origin (Figure 2a). For the duo of U9 and 9M samples, we also observed  
258 maternal heterozygotes in the X-chromosomal variant genotype, despite the  
259 unavailability of the paternal sample. For the additional trios with WGS-only  
260 probands, the Sanger Sequencing of parental samples confirmed the  
261 maternal X inheritance, in which X hemizygotes, maternal heterozygotes, and  
262 paternal wild types were found simultaneously in all trios (Figure 2). Together,  
263 the Sanger Sequencing of four variants, which affected seven cases and  
264 parents, supported the maternal X-chromosome origin exclusively, rather than  
265 *de novo* mutation from probands' embryogenesis or maternal germline. Future  
266 studies based on more pedigree data are needed to further evaluate the  
267 probability of *de novo* mutation during oogenesis of maternal germline or early  
268 development of probands.

### 269 **No Reliable Compound Heterozygous Rare Variants Were Detected**

270 For compound heterozygous mode in autosomes, the family-based structure  
271 can facilitate tracing of the parental origin for each rare variant. We firstly  
272 focused on two heterozygous rare variants with disruptive effects on proteins  
273 for patients of the seven complete trio families. Then, we relaxed the criteria

274 of variant impacts to cover the scenario of one heterozygous rare variant with  
275 disruptive effect and the other with moderate impact (sequence-altering). We  
276 did not find any gene fulfilling the requirements. Finally, we focused on the  
277 scenario of potential compound heterozygous mode, in which non-pedigree  
278 probands would carry at least two heterozygous rare variants in a gene with  
279 disruptive effects. We observed two heterozygous rare variants of *RPTN* with  
280 stop-gain and frameshift impacts in four unrelated patients (Supplementary  
281 Table 5). However, the two variants had the same allele frequency in all  
282 super-populations in the gnomAD database, suggesting that they are more  
283 likely to be inherited via the linkage disequilibrium, rather than the real  
284 compound heterozygous variants.

285 **CRISPR-Cas9 mouse modeling of the splice donor variant in *DMD***  
286 **(NC\_000023.11:g.32454661C>G)**

287 Considering the highly conserved nature of the wild-type splice donor variant  
288 in *DMD* (NC\_000023.11:g.32454661C>G, represented as 'C' in Figure 3a-3b),  
289 we generated transgenic mice with the mutant variant ('G') using the CRISPR-  
290 Cas9 editing technique (Figure 3c). We examined cryptorchidism phenotypes  
291 in F2 male mice using MRI at weeks 4 and 10. In the *DMD* transgenic mice,  
292 we found undescended testicles that are different from the wild-type mice of  
293 the same age (Figure 3d). The undescended testicles were located in the  
294 inguinal region, suggesting the involvement of the *DMD* gene in the  
295 development of cryptorchidism.

## 296 Discussion

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

311 More importantly, we identified four novel candidate cryptorchid genes (*DMD*,  
312 *ARSH*, *MAGEA4*, and *SHROOM2*) with genome-wide significant burdens of  
313 rare mutations with both disruptive and sequence-altering impacts. The gene  
314 *MAGEA4* (melanoma antigen family A, 4) is commonly used as a marker for  
315 human spermatogonia<sup>39</sup>. Based on the immunohistochemistry of *MAGEA4*,  
316 the number of spermatogonia was decreased in cryptorchid testes compared  
317 to the normal testes<sup>40</sup>. Although poorly characterized, the gene *SHROOM2*

318 was reported in patients with infertility<sup>41</sup>. Notably, the findings of the  
319 pathogenic variants of *DMD* are particularly interesting, because of the  
320 established role of *DMD* in producing the protein dystrophin critical for muscle  
321 development<sup>42</sup>. Clinical studies have revealed a positive correlation (OR, 2.83)  
322 between the urogenital malformation and muscular disorders, possibly due to  
323 muscular defects in the cremaster muscle or other inguinal tissues<sup>43</sup>. *DMD*  
324 defects have been reported extensively in Duchenne/Becker muscular  
325 dystrophy<sup>44 45</sup>. Considering the dynamic process involving the muscular  
326 traction of the gubernaculum during the normal testes descent<sup>46</sup>, our  
327 study suggests that muscular abnormalities may play a major role in  
328 cryptorchidism.

329 As the maternal effect is higher than the paternal effect for cryptorchidism (for  
330 example, the recurrence rate is higher in maternal half-brothers than in  
331 paternal half-brothers), the possibility of maternal inheritance has long been  
332 suspected<sup>9 47</sup>. In this study, we revealed the genetic basis of maternal  
333 inheritance of cryptorchidism by finding the dominant role of X-chromosomal  
334 genes and their maternal origin. Indeed, based on the pedigree-based Sanger  
335 sequencing, we revealed that some X-chromosomal variants with available  
336 parental samples were inherited from maternal rather than paternal lineages.  
337 This result is consistent with the expectation of both the X-hemizyosity effect  
338 and the nature of male-specific disease. Because of X-hemizyosity in males,  
339 X-chromosomal pathogenic variants invariably affect the males, while  
340 commonly being masked by a normal allele in heterozygous females. In

341 addition, the causative variant for a male-specific trait would affect males  
342 more heavily than females, except in case of a pleiotropic effect of the variant  
343 on both sexes. Clinically, such hereditary pattern could be utilized to screen  
344 those sex-related genetic disorders in male humans and to develop  
345 subsequent prevention and therapeutic measurements.

346 Together, among the total of eight known and novel genes affecting multiple  
347 unrelated patients of cryptorchidism, seven are X-chromosomal genes and  
348 one is a Y gene. The candidate variants of X-chromosomal genes follow the  
349 expectation of maternal inheritance rather than *de novo* mutation. Thus, our  
350 study revealed a predominant role of sex chromosomal, and particularly the  
351 X-chromosomal, gene defects, in causing male cryptorchidism.

## 352 **Conclusion**

353 We identified rare pathogenic variants of four known candidate genes (*USP9Y*,  
354 *UBA1*, *BCORL1*, and *KDM6A*) and four novel candidate genes (*ARSH*, *DMD*,  
355 *MAGEA4*, and *SHROOM2*) in patients suffering from cryptorchidism.

356 Considering the chromosomal distribution, seven out of eight genes are within  
357 the X-chromosome, and *USP9Y* is located on the Y-chromosome. This  
358 reveals a predominant role of X-chromosomal genes in cryptorchidism. The  
359 maternal origin of these X-chromosomal variants reflects the strong effect of  
360 X-hemizyosity on male-specific diseases. We successfully replicated



361 cryptorchidism phenotypes in transgenic mice. These mice harbor a splice  
362 donor loss variant in *DMD* (NC\_000023.11:g.32454661C>G), establishing a  
363 viable mouse model for future research and therapy studies of this condition.

## 364 **Materials and methods**

### 365 **DNA samples, patient background, and whole-genome sequencing**

366 DNA was extracted from peripheral whole blood of 115 patients and 19  
367 parental controls, using a local database of variants from 2136 unaffected  
368 men for subsequent comparison. The patients were diagnosed by pediatric  
369 specialists of the West China Hospital (WCH) and West China Second  
370 University Hospital (WCSUH). The parents were carefully inquired about  
371 family history and all fathers denied the existence of cryptorchidism in the  
372 family history. All participating parents provided informed consent, and this  
373 study was formally approved by the ethics committees of WCH (Registration  
374 number: 2021389) and WCSUH (Registration number: 2021389). The WGS  
375 (150 bp paired-end) data of an insert size of 350 bp were sequenced using  
376 the DNBSEQ-T7 platform (MGI), according to the manufacturer's protocol  
377 (Supplementary Table 1).

### 378 **Variants calling, genotyping, and annotation**

379 Based on the high-performance computing system, we locally conducted the  
380 variant calling, genetic relationship, and population ancestry using the pipeline  
381 described previously<sup>48-50</sup>. The rare variants were defined as alleles with  
382 frequency lower than 0.01 in all geographic human populations of gnomAD

383 v3.1. Pathogenic variants were identified for known and novel genes  
384 (Supplementary note).

### 385 **Validation of variants and genotypes using the Sanger sequencing**

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

### 393 **The C56BL/6J transgenic mouse modeling**

394 For the highly conserved variant, transgenic mouse was modeled using the  
395 CRISPR-Cas9 editing technique. Briefly, the process was carried out in two  
396 main stages: the in vitro stage and the in vivo stage. In the in vitro stage, the  
397 process began with the design and construction of guide RNAs (gRNAs)  
398 specific to the target DNA sequence around variant, followed by the design  
399 and construction of the donor vector, which carried the desired genetic  
400 modification. The in vivo stage involved microinjecting the designed gRNA  
401 and donor vector into fertilized eggs, and then transplanting these embryos  
402 into surrogate mothers. This led to the birth of the F0 generation mice, which  
403 were subsequently screened to identify individuals with successful genetic  
404 modifications. Positive F0 generation mice were then bred to produce the F1  
405 generation, and these offspring were also screened to confirm the presence of

406 the genetic modification. The F2 mice were examined for phenotypes with  
407 MRI (Time Medical Systems, Inc, USA).

#### 408 **Statements & Declarations**

#### 409 **Acknowledgements**

410 This study was supported by the fifth batch of technological innovation  
411 research projects in Chengdu (2021-YF05 -01331-SN), the Postdoctoral  
412 Research and Development Fund of West China Hospital of Sichuan  
413 University (2020HXBH087), the Short-Term Expert Fund of West China  
414 Hospital (139190032), Fund of Sichuan Provincial Department of Science and  
415 Technology (2021YFS0244), and Fund of Sichuan Provincial Department of  
416 Science and Technology (2021YFS0026). We also acknowledge the  
417 computing support from the West China Biomedical Big Data Center and the  
418 Med-X Center for Informatics of Sichuan University.

#### 419 **Competing Interests**

420 The authors declare no competing financial interests.

#### 421 **Author Contributions**

422 H.Y.X., C.Z.F., L.Z., and F.P.L. supervised this work. J.H.C., J.Z., K.Z., and  
423 H.Z.D., designed the research. J.H.C., J.Z., and G.L.H analyzed data. J.H.C,  
424 Y.Y.J, and J.Z. conducted lab work. G.L.H provided the candidate rare

425 variants screening in 2136 normal controls. J.H.C. drafted the manuscript and

426 all authors took part in revising and improving the manuscript.

#### 427 **Data Availability**

428 The candidate variants data are listed in the supplementary tables. The

429 variants data could be available from the corresponding author on request.

#### 430 **Ethics approval**

431 All authors declare that they have no Conflict of Interest. This study

432 conformed with the Helsinki Declaration of 1975 (as revised in 2008)

433 concerning Human and Animal Rights. All participating parents provided

434 informed consent, and this study was formally approved by the ethics

435 committees of WCH (Registration number: 2021389) and WCSUH

436 (Registration number: 2021389).

#### 437 **Consent to participate**

438 Written informed consent was obtained from the parents.

#### 439 **References**

440 1. Barthold JS, González R. The Epidemiology of Congenital Cryptorchidism, Testicular Ascent and

441 Orchiopexy. *Journal of Urology* 2003;170(6):2396-401. doi:

442 doi:10.1097/01.ju.0000095793.04232.d8

443 2. Agoulnik A, Ferguson L. Testicular Cancer and Cryptorchidism. 2013;4 doi:

444 10.3389/fendo.2013.00032

445 3. Kuiru-Hänninen T, Koskenniemi J, Dunkel L, et al. Postnatal Testicular Activity in Healthy Boys and

446 Boys With Cryptorchidism. *Frontiers in Endocrinology* 2019;10 doi:

447 10.3389/fendo.2019.00489

448 4. Raivio T, Toppari J, Kaleva M, et al. Serum Androgen Bioactivity in Cryptorchid and Noncryptorchid

449 Boys during the Postnatal Reproductive Hormone Surge. *The Journal of Clinical Endocrinology*

450 & *Metabolism* 2003;88(6):2597-99. doi: 10.1210/jc.2002-021676

- 451 5. Elamo HP, Virtanen HE, Toppari J. Genetics of cryptorchidism and testicular regression. *Best*  
452 *Practice & Research Clinical Endocrinology & Metabolism* 2022;36(1):101619. doi:  
453 <https://doi.org/10.1016/j.beem.2022.101619>
- 454 6. Chacko JK, Barthold JS. Genetic and environmental contributors to cryptorchidism. *Pediatr*  
455 *Endocrinol Rev* 2009;6(4):476-80.
- 456 7. Elert A, Jahn K, Heidenreich A, et al. Population-based investigation of familial undescended testis  
457 and its association with other urogenital anomalies. *Journal of Pediatric Urology*  
458 2005;1(6):403-07. doi: 10.1016/j.jpuro.2005.04.005
- 459 8. Promm M, Schröder A, Neissner C, et al. Acquired cryptorchidism: More harm than thought?  
460 *Journal of Pediatric Urology* 2016;12(4):236.e1-36.e6. doi: 10.1016/j.jpuro.2016.04.010
- 461 9. Schnack TH, Zdravkovic S, Myrup C, et al. Familial aggregation of cryptorchidism—a nationwide  
462 cohort study. 2008;167(12):1453-57.
- 463 10. Barthold JS, Wang Y, Kolon TF, et al. Pathway analysis supports association of nonsyndromic  
464 cryptorchidism with genetic loci linked to cytoskeleton-dependent functions. *Human*  
465 *Reproduction* 2015;30(10):2439-51. doi: 10.1093/humrep/dev180
- 466 11. Bloom JS, Boocock J, Treusch S, et al. Rare variants contribute disproportionately to quantitative  
467 trait variation in yeast. *eLife* 2019;8:e49212. doi: 10.7554/eLife.49212
- 468 12. Walter K, Min JL, Huang J, et al. The UK10K project identifies rare variants in health and disease.  
469 *Nature* 2015;526(7571):82-90. doi: 10.1038/nature14962
- 470 13. Maher MC, Uricchio LH, Torgerson DG, et al. Population Genetics of Rare Variants and Complex  
471 Diseases. *Human Heredity* 2012;74(3-4):118-28. doi: 10.1159/000346826
- 472 14. Halvorsen M, Huh R, Oskolkov N, et al. Increased burden of ultra-rare structural variants localizing  
473 to boundaries of topologically associated domains in schizophrenia. *Nature Communications*  
474 2020;11(1):1842. doi: 10.1038/s41467-020-15707-w
- 475 15. Kryukov GV, Pennacchio LA, Sunyaev SR. Most Rare Missense Alleles Are Deleterious in Humans:  
476 Implications for Complex Disease and Association Studies. *The American Journal of Human*  
477 *Genetics* 2007;80(4):727-39. doi: <https://doi.org/10.1086/513473>
- 478 16. Abrahams L, Savisaar R, Mordstein C, et al. Evidence in disease and non-disease contexts that  
479 nonsense mutations cause altered splicing via motif disruption. *Nucleic Acids Research*  
480 2021;49(17):9665-85. doi: 10.1093/nar/gkab750
- 481 17. Ganna A, Satterstrom FK, Zekavat SM, et al. Quantifying the Impact of Rare and Ultra-rare Coding  
482 Variation across the Phenotypic Spectrum. *The American Journal of Human Genetics*  
483 2018;102(6):1204-11. doi: 10.1016/j.ajhg.2018.05.002
- 484 18. Li J, Kong N, Han B, et al. Rare variants regulate expression of nearby individual genes in multiple  
485 tissues. *PLOS Genetics* 2021;17(6):e1009596. doi: 10.1371/journal.pgen.1009596

- 486 19. Gorlov I, Gorlova O, Frazier M, et al. Evolutionary evidence of the effect of rare variants on disease  
487 etiology. *Clinical Genetics* 2011;79(3):199-206. doi: <https://doi.org/10.1111/j.1399->  
488 [0004.2010.01535.x](https://doi.org/10.1111/j.1399-0004.2010.01535.x)
- 489 20. Wang J, Lu Y, Yan X, et al. Identification of novel MITF mutations in Chinese families with  
490 Waardenburg syndrome type II. *Molecular Genetics & Genomic Medicine* 2021;9(9):e1770.  
491 doi: <https://doi.org/10.1002/mgg3.1770>
- 492 21. Bu F, Zhong M, Chen Q, et al. DVPred: a disease-specific prediction tool for variant pathogenicity  
493 classification for hearing loss. *Hum Genet* 2022;141(3-4):401-11. doi: 10.1007/s00439-022-  
494 02440-1 [published Online First: 20220219]
- 495 22. Cheng J, Zhou X, Lu Y, et al. Exome sequencing identifies a novel frameshift mutation of MYO6 as  
496 the cause of autosomal dominant nonsyndromic hearing loss in a Chinese family. *Ann Hum*  
497 *Genet* 2014;78(6):410-23. doi: 10.1111/ahg.12084 [published Online First: 20140917]
- 498 23. Sun Y, Cheng J, Lu Y, et al. Identification of two novel missense WFS1 mutations, H696Y and R703H,  
499 in patients with non-syndromic low-frequency sensorineural hearing loss. *J Genet Genomics*  
500 2011;38(2):71-6. doi: 10.1016/j.jcg.2011.01.001 [published Online First: 20110223]
- 501 24. Cheng J, Zhu Y, He S, et al. Functional mutation of SMAC/DIABLO, encoding a mitochondrial  
502 proapoptotic protein, causes human progressive hearing loss DFNA64. *Am J Hum Genet*  
503 2011;89(1):56-66. doi: 10.1016/j.ajhg.2011.05.027 [published Online First: 20110630]
- 504 25. Yuan Y, Huang D, Yu F, et al. A de novo GJB2 (connexin 26) mutation, R75W, in a Chinese pedigree  
505 with hearing loss and palmoplantar keratoderma. *Am J Med Genet A* 2009;149a(4):689-92.  
506 doi: 10.1002/ajmg.a.32461
- 507 26. Ayers K, Kumar R, Robevska G, et al. Familial bilateral cryptorchidism is caused by recessive  
508 variants in RFXP2. *Journal of medical genetics* 2019;56(11):727-33.
- 509 27. Belkadi A, Bolze A, Itan Y, et al. Whole-genome sequencing is more powerful than whole-exome  
510 sequencing for detecting exome variants. *Proceedings of the National Academy of Sciences*  
511 2015;112(17):5473-78. doi: doi:10.1073/pnas.1418631112
- 512 28. Hiatt SM, Thompson ML, Prokop JW, et al. Deleterious Variation in BRSK2 Associates with a  
513 Neurodevelopmental Disorder. *The American Journal of Human Genetics* 2019;104(4):701-08.  
514 doi: 10.1016/j.ajhg.2019.02.002
- 515 29. McDougal WS, Wein AJ, Kavoussi LR, et al. Campbell-Walsh Urology 11th Edition Review E-Book:  
516 Elsevier Health Sciences 2015.
- 517 30. Sun Y, Liu F, Fan C, et al. Characterizing sensitivity and coverage of clinical WGS as a diagnostic test  
518 for genetic disorders. *BMC Medical Genomics* 2021;14(1):102. doi: 10.1186/s12920-021-  
519 00948-5

- 520 31. Kishikawa T, Momozawa Y, Ozeki T, et al. Empirical evaluation of variant calling accuracy using  
521 ultra-deep whole-genome sequencing data. *Scientific Reports* 2019;9(1):1784. doi:  
522 10.1038/s41598-018-38346-0
- 523 32. Canese K, Weis S. PubMed: the bibliographic database. *The NCBI handbook* 2013;2(1)
- 524 33. McLaren W, Gil L, Hunt SE, et al. The ensembl variant effect predictor. *Genome biology*  
525 2016;17(1):1-14.
- 526 34. Zhan X, Hu Y, Li B, et al. RVTESTS: an efficient and comprehensive tool for rare variant association  
527 analysis using sequence data. *Bioinformatics* 2016;32(9):1423-26.
- 528 35. Gegenschatz-Schmid K, Verkauskas G, Stadler MB, et al. Genes located in Y-chromosomal regions  
529 important for male fertility show altered transcript levels in cryptorchidism and respond to  
530 curative hormone treatment. *Basic and Clinical Andrology* 2019;29(1):8. doi:  
531 10.1186/s12610-019-0089-3
- 532 36. Lu C, Zhang Y, Qin Y, et al. Human X chromosome exome sequencing identifies BCORL1 as  
533 contributor to spermatogenesis. *Journal of Medical Genetics* 2021;58(1):56-65. doi:  
534 10.1136/jmedgenet-2019-106598
- 535 37. Öztürk Ö, Çavdarstepe BE, Bağış H. X-Linked Spinal Muscular Atrophy 2 due to a Synonymous  
536 Variant in the UBA1 Gene in a Family with Novel Findings from Turkey. *Molecular*  
537 *Syndromology* 2022;13(3):246-53. doi: 10.1159/000519640
- 538 38. Armstrong L, Moneim AAE, Aleck K, et al. Further delineation of Kabuki syndrome in 48 well-  
539 defined new individuals. *American Journal of Medical Genetics Part A* 2005;132A(3):265-72.  
540 doi: <https://doi.org/10.1002/ajmg.a.30340>
- 541 39. He Z, Kokkinaki M, Jiang J, et al. Isolation, Characterization, and Culture of Human  
542 Spermatogonia1. *Biology of Reproduction* 2010;82(2):363-72. doi:  
543 10.1095/biolreprod.109.078550
- 544 40. Yang S, Ping P, Ma M, et al. Generation of Haploid Spermatids with Fertilization and Development  
545 Capacity from Human Spermatogonial Stem Cells of Cryptorchid Patients. *Stem Cell Reports*  
546 2014;3(4):663-75. doi: 10.1016/j.stemcr.2014.08.004
- 547 41. Fabian-Jessing BK, Vestergaard EM, Plomp AS, et al. Ocular albinism with infertility and late-onset  
548 sensorineural hearing loss. 2018;176(7):1587-93. doi: <https://doi.org/10.1002/ajmg.a.38836>
- 549 42. Duan D, Goemans N, Takeda Si, et al. Duchenne muscular dystrophy. *Nature Reviews Disease*  
550 *Primers* 2021;7(1):13. doi: 10.1038/s41572-021-00248-3
- 551 43. Diniz G, Barutcuoglu M, Unalp A, et al. Evaluation of the relationship between urogenital  
552 abnormalities and neuromuscular disorders Evaluation of the relationship between  
553 urogenital abnormalities and. 2008;13(1-2):19-24.
- 554 44. Moser H. Duchenne muscular dystrophy: pathogenetic aspects and genetic prevention. *Human*  
555 *genetics* 1984;66(1):17-40.

- 556 45. Emery AE, Muntoni F, Quinlivan R. Duchenne muscular dystrophy: Oxford Monographs on Medical  
557 G 2015.
- 558 46. Hutson JM, Hasthorpe S, Heyns CF. Anatomical and Functional Aspects of Testicular Descent and  
559 Cryptorchidism\*. *Endocrine Reviews* 1997;18(2):259-80. doi: 10.1210/edrv.18.2.0298
- 560 47. Jensen MS, Toft G, Thulstrup AM, et al. Cryptorchidism concordance in monozygotic and dizygotic  
561 twin brothers, full brothers, and half-brothers. *Fertility and Sterility* 2010;93(1):124-29. doi:  
562 10.1016/j.fertnstert.2008.09.041
- 563 48. Chen J, Zhang P, Chen H, et al. Whole-genome sequencing identifies rare missense variants of  
564 WNT16 and ERVW-1 causing the systemic lupus erythematosus. *Genomics*  
565 2022;114(3):110332. doi: <https://doi.org/10.1016/j.ygeno.2022.110332>
- 566 49. Oud MS, Houston BJ, Volozonoka L, et al. Exome sequencing reveals variants in known and novel  
567 candidate genes for severe sperm motility disorders. *Human Reproduction* 2021;36(9):2597-  
568 611. doi: 10.1093/humrep/deab099
- 569 50. Jia Y, Chen J, Zhong J, et al. Novel rare mutation in a conserved site of PTPRB causes human  
570 hypoplastic left heart syndrome. *Clinical Genetics* 2023;103(1):79-86. doi:  
571 <https://doi.org/10.1111/cge.14234>

572

573 Table 1. The candidate pathogenic variants of known genes. “B” and “U” in ID  
574 column are “bilateral cryptorchidism” and “unilateral cryptorchidism”,  
575 respectively. The “Pred” column shows the number of prediction methods  
576 that support deleterious effects of the variants (see Supplementary Table 3  
577 for specific methods). Note: USP9Y (underlined) is a Y chromosome gene,  
578 while all other genes reside on the X chromosome.

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



U13	NC_000024.10:g.12739592C>G;NP_004645.2:p.(S462C)	<u>USP9Y</u>	5
U71	NC_000024.10:g.12739592C>G;NP_004645.2:p.(S462C)	<u>USP9Y</u>	5

579

580 Table 2. The pathogenic variants in novel candidate genes. “B” and “U” in the  
 581 ID column are “bilateral cryptorchidism” and “unilateral cryptorchidism”,  
 582 respectively. The “Pred” column shows the number of prediction methods  
 583 supporting the variants as deleterious (see Supplementary Table 4 for specific  
 584 *in silico* prediction algorithms). Note: all genes are on the X chromosome.

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

585

586 Table 3. The significance levels of burden tests for novel candidate genes  
 587 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

<i>MAGEA4</i>	3.11E-65	2.32E-22	7.75E-23	2.25E-18	1.48E-52
<i>SHROOM2</i>	5.41E-32	1.48E-11	6.67E-13	1.40E-46	5.41E-32

588

589

590

### Figures legends

591 Figure 1. The Magnetic Resonance Imaging (MRI) and Sequencing Depth. (a-  
592 e) The two magnetic resonance imaging (MRI) graphs for bilateral  
593 cryptorchidism (case B130). (c-d) The two MRI graphs for unilateral  
594 cryptorchidism (case U111). (e) The MRI for one control. The red arrows  
595 show the location of the testes. (f) The WGS overall depths for the cases (red)  
596 and control (green).

597 Figure 2. The Pedigree Information and the Chromatogram of the Sanger  
598 Sequencing Results For Cases and Their Parents. (a) The pedigree  
599 information for seven trios and four duos with WGS data. (b) The pedigree  
600 information for six trios, with WGS data only for the cases. (c) The Sanger  
601 sequencing chromatogram for candidate variants of known and novel genes.  
602 “1/1”, “0/1”, and “0/0” are genotypes. U and B represent “unilateral” and  
603 “bilateral”, respectively. “F” and “M” indicate “Father” and “Mother”,  
604 respectively. “-F” and “-R” show the results of PCR sequencing from forward  
605 primer and reverse primer, respectively.

606 Figure 3. Evolutionary Conservation of the Splice Donor Variant in *DMD*  
607 (NC\_000023.11:g.32454661C>G) and MRI of Transgenic and Normal Mice.  
608 (a) Conservation of the *DMD* variant (NC\_000023.11:g.32454661C>G)  
609 according to the UCSC Vertebrate Alignment. (b) Conservation based on the  
610 Amniote alignment from the Ensembl database. (c) Schematic representation  
611 of a transgenic mouse with CRISPR-Cas9 editing. (d) Position of the testicles  
612 in normal C57 mice compared to *DMD* transgenic mice, as determined by  
613 MRI. Undescended testicles were observed in the *DMD* transgenic mice at  
614 both 4 and 10 weeks, with both testicles located in the inguinal region.

615 Supplementary Figure 1. The Screening and Analyzing Process for Variants.  
616 (a) The bioinformatics pipeline for variant calling, filtering, and genotyping with  
617 WGS data. (b) The relationship network among all cases and their family  
618 members. Thicker lines indicate closer genetic relationships. (c) The PCA for  
619 population identity of newly sequenced individuals conducted by incorporating  
620 major populations from the “1000 genomes” project, including African (AFR),  
621 American (AMR), European (EUR), South Asian (SAS), and East Asian  
622 (Chinese Dai in Xishuanagbanna, CDX; Han Chinese in Beijing, CHB; Han  
623 Chinese South, CHS; Japanese in Tokyo, Japan, JPT; Kinh in Ho Chi Minh  
624 City, Vietnam, KHV) populations.

625 Supplementary Figure 2. The evolutionary conservation of sites for known  
626 genes across primate species, based on the Ensembl (v105) sequence  
627 alignments of orthologous genes. The arrows show the locations of variants.

628 Supplementary Figure 3. The Sanger sequencing confirmation for known  
629 genes (*UBA1*, *USP9Y*, *KDM6A*, and *BCORL1*) related to cryptorchidism.

630 Supplementary Figure 4. The evolutionary conservation of sites for novel  
631 candidate genes across primate species, based on the Ensembl (v105)  
632 sequence alignments of orthologous genes. The arrows show the loci with  
633 gene names and variant information.

634 Supplementary Figure 5. The Sanger sequencing validation of four novel  
635 candidate genes (*DMD*, *ARSH*, *MAGEA4*, and *SHROOM2*) related to  
636 cryptorchidism (B130 was not tested due to insufficient amount of DNA).





