IRIS AperTO

UNIVERSITÀ DEGLI STUDI DI TORINO

AperTO - Archivio Istituzionale Open Access dell'Università di Torino

Skewed X-chromosome inactivation in unsolved neurodevelopmental disease cases can guide re-evaluation For X-linked genes

This is a pre print version of the following article:

Original Citation:

Availability:

This version is available http://hdl.handle.net/2318/1903052 since 2023-12-07T08:03:19Z

Published version:

DOI:10.1038/s41431-023-01324-w

Terms of use:

Open Access

Anyone can freely access the full text of works made available as "Open Access". Works made available under a Creative Commons license can be used according to the terms and conditions of said license. Use of all other works requires consent of the right holder (author or publisher) if not exempted from copyright protection by the applicable law.

(Article begins on next page)

18 November 2024

1 **SKEWED X-CHROMOSOME INACTIVATION IN UNSOLVED**
2 **NEURODEVELOPMENTAL DISEASE CASES CAN GUIDE RE-EVALUATION FOR X-**
3 **LINKED GENES**
4

5 Chiara Giovenino[1], Slavica Trajkova[1], Lisa Pavinato[1], Simona Cardaropoli[2], Verdiana Pullano[1], Enza
6 Ferrero[1], Elena Sukarova-Angelovska[3], Silvia Carestiato[1], Paola Salmin[4], Antonina Rinninella[1,5],
7 Anthony Battaglia[1], Luca Bertoli[6], Antonio Fadda[6], Flavia Palermo[1], Diana Carli[2], Alessandro
8 Mussa[2], Paola Dimartino[7], Alessandro Bruselles[8], Tawfiq Froukh[9], Giorgia Mandrile[10], Barbara
9 Pasini[1,4], Silvia De Rubeis[11-13], Joseph D. Buxbaum[11-16], Tommaso Pippucci[17], Marco Tartaglia[8],
10 Marzia Rossato[6], Massimo Delledonne[6], Giovanni Battista Ferrero[2], Alfredo Brusco[1,4]
11

12 1. Department of Medical Sciences, University of Turin, 10126, Turin, Italy
13 2. Department of Public Health and Pediatrics, University of Turin, 10126, Turin, Italy
14 3. Department of Endocrinology and Genetics, University Clinic for Pediatric
15 Diseases, Faculty of Medicine, Ss. Cyril and Methodius University in Skopje, 1000 Skopje,
16 Republic of North Macedonia
17 4. Medical Genetics Unit, Città della Salute e della Scienza University Hospital, 10126, Turin,
18 Italy
19 5. Department of Biomedical and Biotechnological Sciences, Medical Genetics, University of
20 Catania, 94124, Catania, Italy
21 6. Functional Genomics Lab, Department of Biotechnology, University of Verona, 37134,
22 Verona, Italy
23 7. Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy
24 8. Genetics and Rare Diseases Research Division, Ospedale Pediatrico Bambino Gesù, IRCCS,
25 00146, Rome, Italy.
26 9. Department of Biotechnology and Genetic Engineering, Philadelphia University, Amman,
27 Jordan

28 10. Medical Genetics Unit and Thalassemia Center, San Luigi University Hospital, University
29 of Torino, Orbassano, TO, Italy.
30 11. Seaver Autism Center for Research and Treatment, Icahn School of Medicine at Mount
31 Sinai, New York, NY 10029, USA
32 12. Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029,
33 USA
34 13. The Mindich Child Health and Development Institute, Icahn School of Medicine at Mount
35 Sinai, New York, NY 10029, USA
36 14. Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY 10029,
37 USA
38 15. Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai,
39 New York, NY 10029, USA
40 16. Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY
41 10029, USA
42 17. U.O. Genetica Medica, IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna,
43 Italia
44
45 Corresponding Author: Professor Alfredo Brusco, University of Torino, Department of Medical
46 Sciences, via Santena 19, 10126, Torino, Italy. Fax: +39 011 236 5926. Email:
47 alfredo.brusco@unito.it
48
49 Running title: XCI in unsolved NDD
50
51 Keywords: chromosome X inactivation; X-linked gene; neurodevelopmental disorders; exome
52 sequencing reanalysis; breakpoint definition; Xdrop technology.
53

Abstract

Despite major advances in genome technology and analysis, >50% of patients with a neurodevelopmental disorder (NDD) remain undiagnosed after extensive evaluation. A point in case is our clinically heterogeneous cohort of NDD patients that remained undiagnosed after FRAXA testing, chromosomal microarray analysis and trio exome sequencing (ES). In this study, we explored the frequency of non-random X chromosome inactivation (XCI) in the mothers of male patients and affected females, the rationale being that skewed XCI might be masking previously discarded genetic variants found on the X chromosome. A multiplex fluorescent PCR-based assay was used to analyse the pattern of XCI after digestion with *HhaI* methylation-sensitive restriction enzyme. In families with skewed XCI, we re-evaluated trio-based ES and identified pathogenic variants and a deletion on the X chromosome. Linkage analysis and RT-PCR were used to further study the inactive X chromosome allele, and Xdrop long-DNA technology was used to define chromosome deletion boundaries. We found skewed XCI (>90%) in 16/186 (8.6%) mothers of NDD males and in 12/90 (13.3%) NDD females, far beyond the expected rate of XCI in the normal population (3.6%, OR=4.10; OR=2.51). By re-analyzing ES and clinical data, we solved 7/28 cases (25%) with skewed XCI, identifying variants in *KDM5C, PDZD4, PHF6, TAF1, OTUD5* and *ZMYM3*, and a deletion in *ATRX*. We conclude that XCI profiling is a simple assay that targets a subgroup of patients that can benefit from re-evaluation of X-linked variants, thus improving the diagnostic yield in NDD patients and identifying new X-linked disorders.

74 **INTRODUCTION**

75 Major advances in exome sequencing (ES) technologies and data analysis, along with the continuing
76 identification of new disease genes, have greatly contributed to increasing the diagnostic rate of
77 neurodevelopmental disorders (NDDs). However, despite these advances, from 50-70% of NDD
78 cases remain unsolved.(1-3) Among the reasons that make molecular diagnostics challenging are: (i)
79 mutations that create unspecific phenotypes; (ii) difficulties in interpreting variants of uncertain
80 significance (VUS)(4); (iii) newly-defined diseases that describe few patients, making it difficult to
81 draw conclusions about phenotypic expansion (4); and (iv) technical limitations of the diagnostic
82 tools used.(5) Several complementary approaches can be attempted to increase the diagnostic yield
83 of unresolved NDDs, such as transcriptome analysis and whole genome sequencing. Re-analysis of
84 ES data has proven to be the most effective, increasing the diagnostic yield by 10-15%.(6)

85 Approximately 6% of NDDs (6% in males; 6.9% females) are estimated to be X-linked
86 (7)where the pathogenic genetic lesions are can often lead to non-random X-chromosome inactivation
87 (XCI) or skewing. Although this process was identified many decades ago, the actual mechanics and
88 fine details of XCI have not yet been completely characterized. (8) Physiologically, XCI is random
89 and results in an approximate equal ratio of cells expressing either maternal or paternal X
90 chromosome genes. (9) XCI skewing is defined as 'preferential' (80:20%) or 'extreme' (90:10%),
91 and is a frequent indication of the presence of an X-linked pathogenetic variant, be it in affected
92 females (10) or in the mother of male patients. (11)

93 Females heterozygous for an X-linked pathogenetic variant are usually healthy as skewed XCI
94 favors expression of the wild-type allele, thus protecting females from the deleterious effects of the
95 variant.(11,12

96 More recently, skewed XCI has also been observed in affected females, (10) likely caused
97 again by skewing, but this time favouring expression of a deleterious allele that reaches a
98 pathogenicity threshold. Such female patients are susceptible to X-linked recessive conditions with a
99 phenotype similar to that observed in male patients. (10) For X-linked dominant conditions, which

100 may be lethal in males, skewing that selects against the deleterious allele has also been observed, thus
101 decreasing lethality. (13)
102 Of the 281 undiagnosed NNDs, 276 were informative by XCI skewing (90 females + 186
103 mothers of males). We show that XCI skewing analysis combined with targeted re-evaluation of ES
104 data and functional analyses can increase the diagnostic yield and identify novel X-linked disease
105 genes.
106
107 We exploited the analysis of XCI skewing in 276 deeply phenotyped but undiagnosed NDD
108 patients (90 females + 186 mothers of males). We show that XCI skewing analysis combined with
109 targeted re-evaluation of ES data and functional analyses can increase the diagnostic yield and
110 identify novel X-linked disease genes.
111 **MATERIAL AND METHODS**
112 *Study cohort*
113 From the patient cohort that is part of a large international collaborative study aimed at identifying
114 the genetic bases of NDDs, we selected 91 affected females, and 189 mothers of affected males with
115 negative results after trio-ES, CMA (50K Agilent) and FRAXA (see Supplemental materials and
116 methods).
117
118 *X-chromosome inactivation analysis*
119 XCI was tested using DNA extracted from whole blood using an in-house developed protocol. The
120 XCI pattern was calculated using three independent microsatellite polymorphic markers on the X
121 chromosome : (i) the CA repeat in the promoter region of the SLIT and NTRK Like Family Member
122 4 (*SLITRK4*) gene; (ii) the CAG repeat located in exon 1 of the androgen receptor (*AR*) gene(14); (iii)
123 the CA and AG tandem repeats in the first intron of Proprotein Convertase Subtilisin/Kexin Type 1
124 Inhibitor (*PCSK1N*) gene (Supplemental Figure 1; supplemental materials and methods).
125

Characterization of ATRX breakpoints by indirect sequence capture coupled with Illumina sequencing

Xdrop-based enrichment and subsequent amplification of enriched DNA was conducted at the Samplix facility as previously described(15) and subsequently sequenced in 150PE on a NovaSeq6000 (Illumina) (Supplemental materials and methods).

RNA extraction and RT-PCR

To determine which X-chromosome-derived allele (wt or variant) was expressed , we generated and amplified cDNA from total RNA extracted from patients' fresh blood (Supplemental materials and methods).

X-chromosome inactivation phasing by linkage analysis

Families with variants inherited from the mother and segregated in different subjects were analyzed by a set of markers to phase the identified variant with the active or inactive X chromosome and exclude recombination events. The following genetic markers: DXS993, DXS991, DXS986, DXS1068, DXS990, were amplified using AmpliTaq Gold DNA Polymerase (Thermo Fisher Scientific) (see Supplemental Material and Methods for details), separated by capillary electrophoresis on an ABI 3130xl DNA analyzer (Thermo Fisher Scientific) with the GeneScan 500 LYZ size standard (Thermo Fisher Scientific) and analyzed using the GeneMapper software v.4 (Thermo Fisher Scientific).

RESULTS

X chromosome inactivation assay

The HUMARA test, based on the analysis of a CAG repeat in the *AR* locus, is currently the gold standard method for XCI evaluation.(14) As HUMARA is not informative in ~21% of females (due to homozygosity or alleles of difficult interpretation), (16) we set up a fluorescent multiplex

152 methylation-sensitive PCR assay that simultaneously amplifies the *AR* and two additional
153 independent polymorphic microsatellites within *SLITRK4* and *PCSK1N* (Supplemental figure 1).
154 Firstly, we evaluated our assay in a female patient with a balanced Xq25;8q24 translocation and
155 complete XCI skewing (100:0), previously assessed by HUMARA.(17) Complete XCI skewing was
156 confirmed using the two additional informative loci (Supplemental table 2), demonstrating the
157 validity of the test.
158 To further test the assay, we evaluated the XCI pattern in four females with an NDD and four healthy
159 mothers of NDD males: all had received a previous clinical and molecular diagnosis of an X-linked
160 condition with potential skewed XCI (Supplemental Table 3). XCI skewing (>80%) was documented
161 in three affected females and three healthy mothers. Our findings were consistent with the literature
162 which describes the occurrence of XCI skewing in patients with pathogenetic variants in *NAA10*,
163 *PQBP1*, *MECP2*, and *ACSL4.* (18) Similarly, we found random XCI in a healthy mother
164 heterozygous for a pathogenic *IDS* variant, which is in line with previous observations indicating
165 occurrence of skewing in one affected female only. (18) Finally, our female patient with *DDX3X* had
166 a random XCI as reported for half of the patients affected by MRXSSB (MIM #300958). (18) No or
167 limited information was available in the literature for *HNRNPH2*, *RBM10*.
168 *XCI screening in unsolved NDD cases*
169 XCI screening was carried out on 281 individuals: 92 females with NDD and 189 healthy mothers of
170 males with NDD. Probands had been previously studied in depth by CMA/trio-based ES without
171 finding a genetic lesion. The phenotype of the patients varied: 35% (98/281) had intellectual disability
172 (ID), 41% (115/281) had autism spectrum disorder (ASD) and 24% (68/281) were complex
173 syndromic cases with facial dysmorphism (Supplemental Tables 4-5).
174 The XCI assay was informative for at least one marker in 276/281 cases (98.2%; 90/92 female
175 patients, 186/189 mothers) (Supplemental Tables 6-7).

176 We observed 28 cases with extreme skewing (>90%) i: 12/90 female patients (13.3% vs.
177 3.6%; OR=4.10; IC 95% 1.85-9.10) and 16/186 mothers (8.6% vs. 3.6%; OR=2.51; IC 95% 1.21-
178 5.19) (Table1; Supplemental Figure 2-3, Supplemental Table 2).
179
180 *Genetic analyses in XCI-skewed cases*
181 To exclude the possiblility that there were genetic causes underlying the observed skewing of XCI,
182 we sequenced the *XIST* minimal promoter in the above described 28 cases The rationale of this
183 method was to uncover rare variants which might cause epigenetic and functional differences between
184 X chromosomes in females as described by Plenge *et al*.(19) No such variants were found. Thus, we
185 decided to re-evaluate the ES data, focusing on X-chromosome variants and assessing their relevance
186 also in the light of newly available clinical data. We identified an X-linked variant consistent with
187 the phenotype in three females and five males (8/28; 28.6%) (Table 2), as detailed below.
188
189 *A novel TAF1 variant in multiple affected members of family #113*
190 The proband was a 2-year-old girl with global developmental delay and delayed psychomotor
191 development, and almost completely skewed XCI (95:5; Fig. 1A, Supplemental Table 2). She had a
192 14-year-old brother with ID, delayed speech and language development, feeding difficulties and
193 behavioral abnormalities (II.1); a second 11-year-old brother was healthy (II.2). Their mother (I.2)
194 reported she had had teaching support at school.
195 We found a c.745G>A p.(Gly249Arg) variant in Transcription initiation factor TFIID subunit
196 1 (*TAF1*), a gene associated with X-linked syndromic intellectual developmental disorder-33
197 (MRX33; MIM# 300966). MRX33 is characterized by delayed psychomotor development, ID and
198 typical facial dysmorphisms (Supplemental Table 8).20) The variant was inherited from the mother,
199 who also showed skewed XCI (90:10). The variant segregated with the affected brother but not with
200 the healthy brother (Fig.1A, B). The p.(Gly249Arg) variant was absent in the GnomAD database
201 (ver.2.1.1), and changes a highly conserved nucleotide (PhyloP= 9.37; PhastCons= 1) and amino acid

202 residue, which is maintained from vertebrates to *Drosophila melanogaster* (Fig. 1C). The change was
203 predicted to be intolerant by MetaDome(21) (Fig 1D) and deleterious by CADD (Phred: 24.7)(22).
204 Most of the reported likely pathogenic/pathogenic *TAF1* variants are missense substitutions that
205 cluster between exons 16-30, whereas p.(Gly249Arg) is located in exon 6 (Fig. 1E). However, using
206 MutScore (which takes into consideration positional clustering of variants already detected in disease-
207 associated genes and variants found in the population), we noted that the variant reached a predicted
208 pathogenicity score of 0.96 (maximum 1).(23)
209 The pathogenicity of p.(Gly249Arg) was also supported by the predicted structural damage triggered
210 by disallowed phi/psi alert in Missense 3D (24) (Supplemental Material and Methods, Supplemental
211 figure 4A).Using a series of microsatellite markers on the X chromosome (DXS993, DXS991,
212 DXS986), we analysed the segregation of the haplotype containing the c.745G>A *TAF1* variant in
213 the family. Because the haplotype also spanned the AR microsatellite, we could determine the
214 c.745G>A was located on the inactive X chromosome (X_i) in both the mother (I.2) and her daughter
215 (II.3)(Fig. 1A).
216
217 *A rare de novo PHF6 variant affecting the female proband of family NWM24*
218 The proband of family NWM24 was a 7-year-old girl, the second child of healthy parents. At birth,
219 she was small for gestational age (SGA), and presented global developmental delay, autistic behavior
220 several dysmorphic features, divergent strabismus and brachy/syndactyly. XCI was completely
221 skewed (100:0)(Fig. 1F, Supplemental Table 2, Supplemental figure 6A). We identified a *de novo*
222 c.890G>T p.(Cys297Phe) variant in PHD finger protein 6 (*PHF6;* Fig. 1G), a gene associated with
223 X-linked recessive Borjeson-Forssman-Lehmann syndrome (BFLS; MIM# 301900). We
224 reconsidered this previously missed variant because, in the meantime, *de novo* heterozygous variants
225 have been described in affected females with an overlapping but distinct phenotype including
226 characteristic facial dysmorphism, dental, finger and toe abnormalities, and linear skin pigmentation
227 (Supplemental Table 9).(25, 26) These features are present in our patient. The variant is absent in

228 GnomAD (ver 2.1.1), and changes a very conserved nucleotide (PhyloP= 9.36; PhastCons= 1) and
229 amino acid (Fig. 1H). Cys297 is located within the PHD-like zinc-binding domain where most *PHF6*
230 pathogenic/likely pathogenic variants reported in ClinVar map (MutScore= 0.949) map. Cys297 is
231 considered intolerant to change by MetaDome (Fig. 1I; PF13771; a.a. 239-330; UniProt: Q8IWS0).
232 Bioinformatic analyses predict the change to be deleterious (CADD Phred= 29.5; REVEL= 0.97;
233 Table 2). Pathogenicity of p.(Cys297Phe) was also supported by the predicted structural damage, the
234 amino acid substitution triggering a clash alert(24) (local clash score: wild type=10.47; mutant=35.67;
235 Supplemental Figure 4B).
236
237 *A KDM5C variant with variable expressivity in family #237*
238 In Family 237, we found a 10-year-old girl with moderate ID and skewed XCI (Fig. 1J). She was the
239 second of four siblings that included one affected brother (III.1) and two healthy sisters (III.3 and
240 III.4). The parents were healthy, but several male maternal relatives were reported to have ID. We
241 found a maternally inherited c.1204G>A p.(Asp402Asn) missense variant in Lysine-specific
242 demethylase 5C (*KDM5C*), a gene associated with intellectual developmental disorder, X-linked,
243 syndromic, Claes-Jensen type (MRXSCJ; MIM# 300534; fig. 1K). The variant was shared by the
244 proband's affected brother (III.1), and one of her healthy sisters (III.3). The variant was predicted to
245 be deleterious by bioinformatic analysis (CADD Phred: 29.7; REVEL: 0.866), and the affected
246 residue mapped to a region that was considered intolerant to variation by MetaDome (Fig. 1L), and
247 conserved from vertebrates to drosophila (Fig. 1M). Another variant affecting this amino acid residue
248 [c.1204G>T p.(Asp402Tyr)] was previously demonstrated to compromise KDM5C stability and
249 enzymatic activity.(27)
250 MRXSCJ is an X-linked recessive disorder, characterized by DD/ID with clinical
251 heterogeneity in affected males.(28) Recurrent features include short stature, microcephaly,
252 hyperreflexia and aggressive behavior, which were present both in the proband and her brother (III.1).
253 Females with variants in *KDM5C*, as in case III.2, have only recently been found to be associated

254 with incomplete penetrance and a variable phenotype ranging from mild to severe ID (Supplemental
255 Table 10).(28) The presence of both a male and female in this family initially led us to discard X-
256 linked genes.

257 By determining the phase of AR alleles and the *KDM5C* alleles by linkage analysis, we
258 demonstrated that the affected sister (III.2) had a preferentially active mutant allele (90%);
259 conversely, the unaffected sister (III.3) and her mother had a preferentially inactive mutant allele (Fig.
260 1J).

261

262 *A genomic ATRX deletion characterized by the Xdrop method in family #236*

263 In family 236, the proband was a 13-year-old boy with a long diagnostic odyssey (Fig. 2A). At 3
264 years of age, he presented with hypotonia, DD/ID and dysmorphisms. The phenotype was compatible
265 with mental retardation-hypotonic facies syndrome (MRXFH1, MIM# 309580); however, ES was
266 negative for an intragenic *ATRX* pathogenic variant (MIM* 300032). We found complete XCI
267 skewing (100:0) in the mother, prompting us to re-evaluate the genetic data. By visually inspecting
268 the ES reads using IGV(29), we noticed no coverage of exons 3 and 4 of *ATRX* (Fig. 2B upper panel),
269 suggesting the presence of an intragenic deletion. According to linkage analysis, the X_i chromosome
270 in the mother carried the haplotype with the deletion (Fig.2A).

271 For in-depth characterization of the deletion, we used the Indirect Sequence Capture (Xdrop
272 technology)(30,31), a powerful method for characterizing specific genomic regions. We enriched for
273 a region of ~100 kb within the *ATRX* gene, spanning the deletion. This region was subsequently
274 sequenced at high coverage using the Illumina NGS platform. The analysis identified the breakpoints
275 of the deletion (Supplemental Figure 5), with an uncertainty of 3 bp, identical on both sides of the
276 interrupted region (hg38; chrX:77,697,545-77,703,516; chrX:77,697,542-77,703,513) (Fig. 2B).
277 Remapping of the Illumina reads on the reconstructed sequence demonstrated perfect alignment,
278 without mismatches, thus confirming the correctness of the breakpoints (Figure 2B, lower panel). The

279 deletion of 5,971 bp was confirmed by Sanger sequencing using flanking PCR primers and shown to
280 be inherited from the mother (Fig. 2C, D).
281
282 *PDZD4: a possible novel NDD gene in family NWM25*
283 In family NWM25, we identified a mother of an affected boy with 90:10 XCI (Fig 2E, Supplemental
284 Table 2). Since the age of two, the son presented symptoms of DD, followed by the development of
285 kyphoscoliosis with pectus excavatum, hyperelastic skin and joints, persistent hand tremors, facial
286 dysmorphisms and polymicrogyria by brain MRI. Two maternal uncles were reported to be affected
287 by undefined ID.
288 Re-analysis of the X-chromosome variants led to the identification of a c.2190G>C
289 p.(Lys736Asn) missense variant in the PDZ domain-containing 4 gene (*PDZD4*; MIM* 300634) (Fig.
290 2F), which was inherited from the healthy mother. Lys736 is conserved in vertebrates (Fig 2G). Using
291 linkage analysis, we showed that the haplotype with p.(Lys736Asn) was located on the inactive X-
292 chromosome (Fig 2E).
293 By exploiting GeneMatcher (https://genematcher.org/), we identified a second affected 12-
294 year-old female (II.1; family TF110, Figure 2H) with a *de novo* frameshift c.10_16del
295 p.(Asn4Alafs*12) variant in *PDZD4*. She presented with an overlapping phenotype, including DD,
296 microcephaly, ID and dysmorphisms. Also in this family we observed almost complete XCI (95:5)
297 in the proband, although we could not determine if it the variant was located on the inactive X
298 chromosome.
299
300 *OTUD5 a novel recently identified gene in family #234*
301 The probands of family 234 were two brothers, aged 16 and 26 years, with mild ID. Their healthy
302 mother showed complete skewing of XCI (100:0) (Fig. 2I). We identified a missense c.1526C>T
303 p.(Pro509Leu) variant in OTU DOMAIN-CONTAINING PROTEIN 5 (*OTUD5*), a gene that has recently
304 been associated with Multiple Congenital Anomalies-Neurodevelopmental syndrome (MCAND;

305 MIM# 301056).(31) MCAND is an X-linked recessive congenital multisystemic disorder
306 characterized by poor growth, global developmental delay with impaired intellectual development
307 together with variable abnormalities of the cardiac, skeletal, and genitourinary systems. Disease
308 severity is highly variable, ranging from death in early infancy to survival into the second or third
309 decade, suggesting the variant is hypomorphic.(32)

310 We first confirmed that the c.1526C>T allele was indeed expressed in the patient's blood (II.1,
311 fig. 2J). Next, we compared the cDNA sequence of *OTUD5* from the patient's blood with the *OTUD5*
312 genomic DNA sequence (gDNA) from the mother and showed that the c.1526C>T allele was not
313 detectable, suggesting that the skewed X inactivation preferentially silenced the chromosome with
314 the variant (I.2, fig. 2J). Bioinformatic analyses predicted the variant to be likely pathogenic (Table
315 2). The substitution of leucine with proline triggers a structural damage with a local clash score of
316 33.58 versus a score of 15.21 calculated for the wild type protein (Supplemental Figure 4C).(24)

317

318 *ZMYM3: a possible novel NDD gene in family NWM127*

319 **In family NWM127, subject II.1 (Fig. 2K) is a 13-year-old male with DD, moderate ID,
320 cryptorchidism, porosis of bonesosteoporosisand dysmorphic features. He was the fourth child
321 in a family of European ancestry and had an affected sister (II.2) presenting with severe ID due
322 *de novo* tetrasomy for 15q11.2-q13.1 (MIM *608636). He was severely hypotonic in early
323 infancy and showed relevant delay in his gross motor milestones (head control at one year and
324 sitting position at five years). He never developed fine motor skills nor acquired toilet training.
325 Dysmorphic features included long face, tall forehead, thick eyebrows, deeply set eyes, broad
326 nasal tip, and low-set flashy ears with cupped formed ear lobes. Upon re-analysis of the ES data,
327 we found a c.1322G>A p.(Arg441Gln) variant in the Zinc Finger, MYM-type 3 (*ZMYM3*) gene.
328 The mother showed completely skewed XCI (100:0), and similar XCI skewing was found in
329 both the II.2 (90:10) and II.3 unaffected sister (85:15), with the mutant allele preferentially
330 inactive. Furthermore, a p.(Arg441Trp) variant was described by Philips et al. in 2014 in three**

331 **male probands with ID and several dysmorphic features shared with our proband II.1, and was**

332 **recently confirmed as a recurrent variant in a novel *ZMYM3*-associated NDD.(33,34) Other**

333 **potentially causative ES-detected variants were excluded by functional analysis [e.g., *de novo***

334 ***OSBPL8*: c.1535T>C; p.(Val512Ala)] that did not show altered protein activity (Prof. T. Balla,**

335 **Bethesda, MD, *personal communication*)SION**

336 Among the mechanisms that cause deviation from random X chromosome inactivation is

337 selection against cells expressing X chromosomes carrying a pathogenic genetic lesion. We reasoned

338 that we could take advantage of unbalanced XCI and use it as a guide for re-evaluating clinical and

339 molecular data in NDD patients in which previous genetic testing failed to make a diagnosis. To test

340 for XCI, we set up a multiplex fluorescent PCR that simultaneously analyzed the methylation status

341 of three independent polymorphic markers on the X-chromosome. This assay allowed us to increase

342 informativeness to >98%, compared to 80% using standard HUMARA.(14)

343 Analysis of 91 female NDD patients and 186 mothers of male NDD patients, previously

344 undiagnosed by CMA and trio-ES, showed a significant enrichment of subjects with extremely

345 unbalanced XCI, defined as a >90:10 XCI ratio (28/277, 10%) in line with the reults of a similar

346 study.(10) The extreme skewing of XCI suggested that some of our undiagnosed cases might be

347 attributable to a gene located on the X chromosome. Proof of principle came from the re-evaluation

348 of available trio-ES data: by focusing on X-linked coding regions, we identified likely pathogenic

349 variants in 7/28 cases, solving 25% of NDDs with skewed XCI.

350 In our original survey of 575 NDD cases, we had 28 patients with skewed XCI: nine with X-

351 linked variants classified as class 4 or 5, and 12 with class 3 variants. Taking into consideration these

352 28 cases, we estimate that X-linked genes account for 6.4-8.5% (9+28/575; 21+28/757) of the patients

353 in our survey. Our figures are in agreement with the data from a recent evaluation of the burden of

354 X-linked coding variation based on 11,044 Developmental Disorder patients, which estimated X-

355 linked causes in 6.0% of males and 6.9% of females.(7

356 We previously missed seven variants on the X chromosome for one of three reasons, namely:
357 (i) the gene was not associated with disease at the time of the analysis (*OTUD5, PDZD4,* and
358 *ZMYM3*); (ii) the variant was a structural rearrangement missed by ES (*ATRX*) or (iii) the variant was
359 overlooked because it was apparently inconsistent with X-linked segregation, since both males and
360 females were affected (*TAF1, PHF6,* and *KDM5C*). In the first category, LINKage-specific-
361 deubiquitylation-deficiency-induced embryonic defects (LINKED) syndrome was first associated
362 with pathogenic *OTUD5* variants only in 2021,(32) whereas *PDZD4* and *ZMYM3* have been at present
363 only proposed to be disease-causing genes. Variants predicted to damage protein structure or function
364 in *ZMYM3* have been identified in patients with NDD in December 2022, (33) whereas *PDZD4* has
365 been at present only proposed to be disease-causing gene. (35)the family with a microdeletion in
366 *ATRX* highlights the importance of searching for genomic rearrangements, exploiting exome data, or
367 performing genome sequencing. In this case the deletion was missed by CMA due to lack of array
368 probes in the deleted tract. The strong clinical suspicion of ID-hypotonic facies syndrome (MIM#
369 309580) prompted us to analyze the coverage of all *ATRX* exons on ES data and to finally identify
370 the deletion of exons 3-4. We also chose to locate the precise breakpoints using a novel method based
371 on the enrichment for targeted resequencing by the Xdrop technology, which combines high-
372 resolution droplet PCR (dPCR) with droplet sorting and Multiple Displacement Amplification in
373 droplets (dMDA). This approach proved to be successful in fine-mapping the deletion breakpoints,
374 narrowing them down from a large putative region of ~20 kb between exons 2 and 5. Given the
375 flexibility of this technology, we expect it to be useful when analysing other similar cases where the
376 large size of the involved region hampers the efficient use of traditional assays for the characterization
377 of structural variations at the single-base resolution. Alternatively, achieving the same results would
378 have required either genome sequencing (more expensive) or a very large sets of PCR-based assays
379 and labor/time intensive work to map the whole 20 kb region, also because the deletion maps within
380 a region rich in repeated sequences. Availability of the deletion boundaries allowed us to set up a
381 simple PCR test to follow segregation of the variant in the family.

382 The X chromosome is often underestimated in the diagnosis of female NDD patients because
383 of the common misconception that females are less susceptible to X-linked conditions.(18) Although
384 many X-linked conditions show a profound sex-linked bias, given the specific mechanism of
385 inheritance, an increasing number of X-linked diseases have been described that occur similarly in
386 both female and male patients.(18) For example, in families 113 and NWM24, we identified a
387 missense variant in *TAF1* and *PHF6* in a female; we overlooked/ignored these variants at the first ES
388 reading because inconsistent with an X-linked recessive disease. However, the literature reported
389 females with phenotypes consistent with variants in those genes. In the case of *TAF1*, completely
390 skewed XCI is consistent with other recently described cases where the phenotype, which differs in
391 females and males, is uniform within each sex.(36) XCI unbalance favors the wild-type allele in both
392 the mother (mild phenotype) and the affected daughter, leaving the pathogenic mechanism unclear.
393 We can speculate that: (i) expression of 5% of the pathogenic allele is sufficient to cause the
394 phenotype or (ii) the XCI pattern is different in affected tissues such as brain, where the pathogenic
395 allele is for some reason more expressed than in blood. In family NWM24, the phenotype associated
396 with *PHF6* is consistent with the literature that reports two females carrying the *de novo*
397 p.Cys305Phe, just a few amino acids distant from our proband's variant.(25, 37)
398 Family 237 is another example of X-linked gene complexity: three females carried a missense
399 variant in *KDM5C* but we detected skewed XCI towards the deleterious allele only in the individual
400 with the disease phenotype. Segregation analysis showed skewing towards the deleterious allele.
401 *KDM5C* is known to escape XCI and thus the role of skewing in the phenotype is not clear.(38)
402 Among the various causes of female susceptibility to X-linked conditions, XCI certainly plays
403 a key role at the penetrance level. Although the mechanism of XCI has been known for a very long
404 time, evaluating XCI's influence on phenotype remains challenging. In some cases, the presence of
405 skewed XCI in more easily explained by the selection of cells that inactivate the mutated allele,
406 expressing only the wild-type allele and gaining a selective advantage during the early stages of
407 development.(11) Typical examples are mothers heterozygous for *OTUD5*, *ATRX*, *ZMYM3* and

408 *PDZD4* variants who are protected against the deleterious effect of an X-linked pathogenetic variant
409 by skewed XCI. In females with X-linked conditions, XCI can modulate expression of the
410 phenotype;(39) it is likely that there are several mechanisms that underlie disease and skewing that
411 currently escape our understanding and are not always easily identifiable. Finally, in 20 XCI-skewed
412 cases, we could not identify any potentially causative variant. We hypothesize that the phenotype
413 might be explained by variants in coding regions not covered by exome sequencing? or by noncoding
414 variation, such as deep intronic variants that affect splicing or regulatory regions.

415 Taken together, our data conclude that XCI testing is a simple, inexpensive and productive
416 means for re-evaluating exome data from the X chromosome.

417

418 **DATA AVAILABILITY**

419 All variants have been deposited into ClinVar: SCV002583290, SCV002583291, SCV002583292,
420 SCV002583293, SCV002583294, SCV002583295, SCV002583296, SCV002583297,
421 SCV002583298, SCV002583299, SCV002583300, SCV002583301.

422

423 **REFERENCES**

424 1. Miles JH. Autism spectrum disorders--a genetics review. Genet Med. 2011;13(4):278-94.

425 2. Betancur C. Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and
426 genomic disorders and still counting. Brain Res. 2011;1380:42-77.

427 3. Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, et al. Large-Scale Exome
428 Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism.
429 Cell. 2020;180(3):568-84.e23.

430 4. Hartley T, Lemire G, Kernohan KD, Howley HE, Adams DR, Boycott KM. New Diagnostic
431 Approaches for Undiagnosed Rare Genetic Diseases. Annu Rev Genomics Hum Genet. 2020;21:351-72.

432 5. Boycott KM, Hartley T, Biesecker LG, Gibbs RA, Innes AM, Riess O, et al. A Diagnosis for All Rare
433 Genetic Diseases: The Horizon and the Next Frontiers. Cell. 2019;177(1):32-7.

434 6. Basel-Salmon L, Orenstein N, Markus-Bustani K, Ruhrman-Shahar N, Kilim Y, Magal N, et al.
435 Improved diagnostics by exome sequencing following raw data reevaluation by clinical geneticists involved
436 in the medical care of the individuals tested. Genet Med. 2019;21(6):1443-51.

437 7. Martin HC, Gardner EJ, Samocha KE, Kaplanis J, Akawi N, Sifrim A, et al. The contribution of X-
438 linked coding variation to severe developmental disorders. Nat Commun. 2021;12(1):627.

439 8. Gjaltema RAF, Schwämmle T, Kautz P, Robson M, Schöpflin R, Ravid Lustig L, et al. Distal and
440 proximal cis-regulatory elements sense X chromosome dosage and developmental state at the Xist locus. Mol
441 Cell. 2022;82(1):190-208.e17.

442 9. Harper PS. Mary Lyon and the hypothesis of random X chromosome inactivation. Hum Genet.
443 2011;130(2):169-74.

444 10. Fieremans N, Van Esch H, Holvoet M, Van Goethem G, Devriendt K, Rosello M, et al. Identification
445 of Intellectual Disability Genes in Female Patients with a Skewed X-Inactivation Pattern. Hum Mutat.
446 2016;37(8):804-11.

447 11. Giorgio E, Brussino A, Biamino E, Belligni EF, Bruselles A, Ciolfi A, et al. Exome sequencing in
448 children of women with skewed X-inactivation identifies atypical cases and complex phenotypes. Eur J
449 Paediatr Neurol. 2017;21(3):475-84.

450 12. Plenge RM, Stevenson RA, Lubs HA, Schwartz CE, Willard HF. Skewed X-chromosome inactivation
451 is a common feature of X-linked mental retardation disorders. Am J Hum Genet. 2002;71(1):168-73.
452 13. Li D, Strong A, Shen KM, Cassiman D, Van Dyck M, Linhares ND, et al. De novo loss-of-function
453 variants in X-linked MED12 are associated with Hardikar syndrome in females. Genet Med. 2021;23(4):637-
454 44.
455 14. Amos-Landgraf JM, Cottle A, Plenge RM, Friez M, Schwartz CE, Longshore J, et al. X chromosome-
456 inactivation patterns of 1,005 phenotypically unaffected females. Am J Hum Genet. 2006;79(3):493-9.
457 15. Blondal T, Gamba C, Møller Jagd L, Su L, Demirov D, Guo S, et al. Verification of CRISPR editing
458 and finding transgenic inserts by Xdrop indirect sequence capture followed by short- and long-read
459 sequencing. Methods. 2021;191:68-77.
460 16. Allen RC, Zoghbi HY, Moseley AB, Rosenblatt HM, Belmont JW. Methylation of HpaII and HhaI
461 sites near the polymorphic CAG repeat in the human androgen-receptor gene correlates with X chromosome
462 inactivation. Am J Hum Genet. 1992;51(6):1229-39.
463 17. Di Gregorio E, Bianchi FT, Schiavi A, Chiotto AM, Rolando M, Verdun di Cantogno L, et al. A de
464 novo X;8 translocation creates a PTK2-THOC2 gene fusion with THOC2 expression knockdown in a patient
465 with psychomotor retardation and congenital cerebellar hypoplasia. J Med Genet. 2013;50(8):543-51.
466 18. Migeon BR. X-linked diseases: susceptible females. Genet Med. 2020;22(7):1156-74.
467 19. Plenge RM, Hendrich BD, Schwartz C, Arena JF, Naumova A, Sapienza C, et al. A promoter mutation
468 in the XIST gene in two unrelated families with skewed X-chromosome inactivation. Nat Genet.
469 1997;17(3):353-6.
470 20. O'Rawe JA, Wu Y, Dörfel MJ, Rope AF, Au PY, Parboosingh JS, et al. TAF1 Variants Are Associated
471 with Dysmorphic Features, Intellectual Disability, and Neurological Manifestations. Am J Hum Genet.
472 2015;97(6):922-32.
473 21. Wiel L, Baakman C, Gilissen D, Veltman JA, Vriend G, Gilissen C. MetaDome: Pathogenicity
474 analysis of genetic variants through aggregation of homologous human protein domains. Hum Mutat.
475 2019;40(8):1030-8.
476 22. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for
477 estimating the relative pathogenicity of human genetic variants. Nat Genet. 2014;46(3):310-5.

478 23. Quinodoz M, Peter VG, Cisarova K, Royer-Bertrand B, Stenson PD, Cooper DN, et al. Analysis of
479 missense variants in the human genome reveals widespread gene-specific clustering and improves prediction
480 of pathogenicity. Am J Hum Genet. 2022;109(3):457-70.
481 24. Ittisoponpisan S, Islam SA, Khanna T, Alhuzimi E, David A, Sternberg MJE. Can Predicted Protein
482 3D Structures Provide Reliable Insights into whether Missense Variants Are Disease Associated? J Mol Biol.
483 2019;431(11):2197-212.
484 25. Zweier C, Kraus C, Brueton L, Cole T, Degenhardt F, Engels H, et al. A new face of Borjeson-
485 Forssman-Lehmann syndrome? De novo mutations in PHF6 in seven females with a distinct phenotype. J Med
486 Genet. 2013;50(12):838-47.
487 26. Gerber CB, Fliedner A, Bartsch O, Berland S, Dewenter M, Haug M, et al. Further characterization of
488 Borjeson-Forssman-Lehmann syndrome in females due to de novo variants in PHF6. Clin Genet.
489 2022;102(3):182-90.
490 27. Brookes E, Laurent B, Õunap K, Carroll R, Moeschler JB, Field M, et al. Mutations in the intellectual
491 disability gene KDM5C reduce protein stability and demethylase activity. Hum Mol Genet. 2015;24(10):2861-
492 72.
493 28. Carmignac V, Nambot S, Lehalle D, Callier P, Moortgat S, Benoit V, et al. Further delineation of the
494 female phenotype with KDM5C disease causing variants: 19 new individuals and review of the literature. Clin
495 Genet. 2020;98(1):43-55.
496 29. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, et al. Integrative
497 genomics viewer. Nat Biotechnol. 2011;29(1):24-6.
498 30. Grosso V, Marcolungo L, Maestri S, Alfano M, Lavezzari D, Iadarola B, et al. Characterization of.
499 Front Genet. 2021;12:743230.
500 31. Madsen EB, Höijer I, Kvist T, Ameur A, Mikkelsen MJ. Xdrop: Targeted sequencing of long DNA
501 molecules from low input samples using droplet sorting. Hum Mutat. 2020;41(9):1671-9.
502 32. Beck DB, Basar MA, Asmar AJ, Thompson JJ, Oda H, Uehara DT, et al. Linkage-specific
503 deubiquitylation by OTUD5 defines an embryonic pathway intolerant to genomic variation. Sci Adv.
504 2021;7(4).

33. Hiatt SM, Trajkova S, Sebastiano MR, Partridge EC, Abidi FE, Anderson A, et al. Deleterious, protein-altering variants in the transcriptional coregulator ZMYM3 in 27 individuals with a neurodevelopmental delay phenotype. Am J Hum Genet. 2022.

34. Philips AK, Sirén A, Avela K, Somer M, Peippo M, Ahvenainen M, et al. X-exome sequencing in Finnish families with intellectual disability--four novel mutations and two novel syndromic phenotypes. Orphanet J Rare Dis. 2014;9:49.

35. Leitão E, Schröder C, Parenti I, Dalle C, Rastetter A, Kühnel T, et al. Systematic analysis and prediction of genes associated with monogenic disorders on human chromosome X. Nat Commun. 2022;13(1):6570.

36. Cheng H, Capponi S, Wakeling E, Marchi E, Li Q, Zhao M, et al. Missense variants in TAF1 and developmental phenotypes: challenges of determining pathogenicity. Hum Mutat. 2019.

37. Wieczorek D, Bögershausen N, Beleggia F, Steiner-Haldenstätt S, Pohl E, Li Y, et al. A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. Hum Mol Genet. 2013;22(25):5121-35.

38. Tukiainen T, Villani AC, Yen A, Rivas MA, Marshall JL, Satija R, et al. Landscape of X chromosome inactivation across human tissues. Nature. 2017;550(7675):244-8.

39. Franco B, Ballabio A. X-inactivation and human disease: X-linked dominant male-lethal disorders. Curr Opin Genet Dev. 2006;16(3):254-9.

ACKNOWLEDGMENTS

We are grateful to the patients and their families for their participation in this study. We acknowledge Samplix for the technical support.

AUTHOR CONTRIBUTIONS

531 Conceptualization: C.G., S.T, A.B.; Data curation: C.G., S.T, A.B., E.S., F.P., D.C., A.M., T.F., G.M., 532 B.P., G.B.F; Formal analysis: C.G., S.T, L.P., S.C., V.P., S.C., A.R., A.B., P.D., A.B., T.P., M.T.; 533 Investigation: C.G., S.T, L.P., S.C., V.P., S.C., L.B., A.F., P.S., S.D.R., J.B.; Methodology: C.G., 534 P.S., L.B., A.F., M.R., M.D., Visualization: Writing-original draft: C.G., S.T, A.B; Writing-review 535 & editing: E.F., C.G., S.T, A.B, E.S., F.P., D.C., A.M., T.F., G.M., B.P., G.B.F, M.R., M.D.

536

537 **FUNDING SUPPORT**

538 This research received funding from the Italian Ministry for Education, University and Research 539 (Ministero dell'Istruzione, dell'Università e della Ricerca - MIUR) PRIN2020 code 20203P8C3X to 540 AB, "Associazione E.E. Rulfo per la genetica Medica" to AB, Fondazione Cassa di Risparmio di 541 Torino to AB, and Fondazione Bambino Gesù (Vite Coraggiose) to MT. Sample collection was 542 supported by the NIMH (U01MH111661 to JDB).

543

544 **ETHICS DECLARATION**

545 All individuals and families from the different institutions agreed to participate in this study and 546 signed appropriate consent forms. The Ethics Committee of Città della Salute e della Scienza 547 University Hospital (n. 0060884) and University of Skopje (n. 03-6116/7) approved this study.

548

549 **DECLARATION OF INTERESTS**

550 The authors declare no competing interests.

551

552 **Figure legends**

553 **Figure 1. Pedigree and variant analysis in the three families with XCI-skewed female cases.**

554 **A, F, J.**

555 Family trees of families 113, NWM24, and 237. We used X-Chromosome polymorphic
556 microsatellites to reconstruct the haplotypes and to phase the pathogenetic variant on the
557 inactive/active X chromosome (percentage indicated below the symbol of tested females; Xi and Xa
558 indicate the less and the most active X chromosomes). The hyphen above each symbol indicates
559 whenever DNA was available for genetic testing.

560 **B, G, K.** Sanger sequencing used to confirm the variants in *TAF1* (NM_004606.5), *PHF6*
561 (NM_01015877.2) and *KDM5C* (NM_004187.5). Representative electropherograms are shown: wild
562 type (wt); mutant hemizygous (mut); mutant heterozygous (mut/wt). **C, H, M.** Multiple sequence
563 alignment of the protein amino acid sequences in different species obtained using Marrvel software
564 for the relevant changed aminoacids (highlighted in yellow; http://marrvel.org/)(hs: *Homo sapiens*;
565 mm: *Mus musculus*; rn: *Rattus norvegicus*; xt: *Xenopus tropicalis*; dr: *Danio rerio*; dm: *Drosophila*
566 *melanogaster*). **D, I, L**. Tolerance Landscape obtained using MetaDome Web Server visualizes
567 regional tolerance to normal genetic variation (https://stuart.radboudumc.nl/metadome/). The
568 position of the missense change is indicated for each gene. The Tolerance Landscape Y-axis is
569 reported as a color scale from blue (position tolerant to variation, T), to yellow (position neutral to
570 variation, N), to red (position intolerant to variation, I). Below the X- axis, a schematic representation
571 of the known protein domains (pink). **E**. Localization of the pathogenic (red) and likely pathogenic
572 (orange) variants reported in the literature for *TAF1* gene in male (upper panel) and female cases
573 (lower panel). Our patient's variant is shown in black.

574

575 **Figure 2. Pedigrees and variants analysis in the three families with XCI skewed mothers of**
576 **affected males.**

577 **A, E, H, I, K**. Family trees of families 236, NWM25, TF110, 234 and NWM127. See legend in figure
578 2A, F, J. **B**. NGS Coverage of *ATRX* exons (schematized above) in ES data (upper panel) and with
579 Xdrop enrichment (lower panel) in the II.1 proband from family 236. Xdrop enrichment primers (blue
580 bars below) were designed 5' of the maximum estimated deletion. After enriching DNA for the
581 region, and subsequent Illumina Sequencing, we were able to precisely identify a 5,971 bp deletion
582 spanning exons 3 and 4. **C**. Sanger sequencing validation of the *ATRX* deletion in II.1 and his mother
583 (I.2) using primers flanking the deleted segment (arrows). The deletion breakpoint is shown in panel
584 **D**. **F, H.** Sanger sequencing validation of the identified variants. **J.** In family 234, we sequenced the
585 genomic region (gDNA) and the corresponding transcript (cDNA) in one of the probands (II.1) and
586 their mother (I.2). The wild-type allele only was detected in both cases in the cDNA, showing that
587 the pathogenic variant was not expressed and thus located on the inactive X-chromosome. **G.**
588 Multiple alignment of the protein amino acid sequences in different species as described in the legend
589 for figure 2 **C, H, M**.
590

Table 1- Comparison of skewed X-inactivation ratio in adult population, female NDD patients and mothers of male NDD patients.

				Percentage of population with XCI ratio of:		
	n.	**Mean**	**S.D.**	**<80:20**	**>90:10**	**>95:5**
Adult Population *	415	52:48:00	19.3	14.2%	3.6%	1.7%
Female NDD	90	68:32:00	14.2	22.2%	13.3%	6.7%
Mother NDD	186	68:32:00	12.8	18.3%	8.6%	3.8%

Notes. "Mean" indicates the mean X-chromosome inactivation (XCI) percentage within the group; S.D.: standard deviation. *Control population data obtained from [12].

Table 2. Variants found in the eight families with X-skewed females

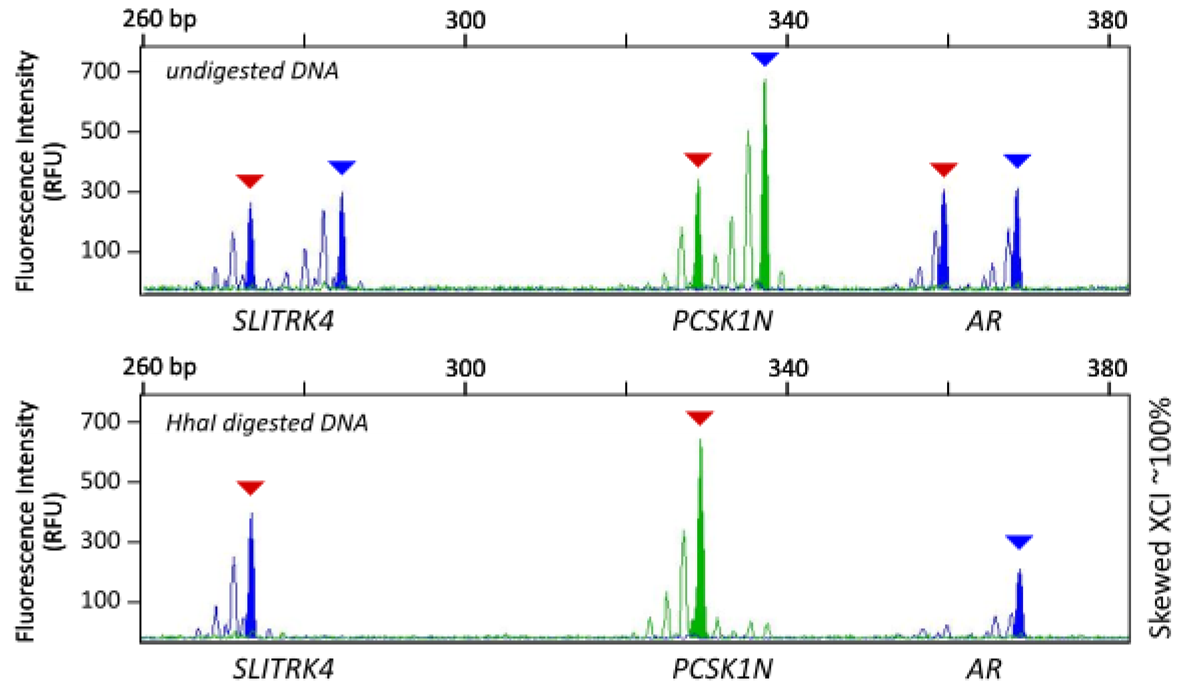
Family_ID	Sex	Diagnosis	OMIM	Inheritance	Gene	Variant c.DNA	Variant protein	SIFT DANN LRT Mutation Taster	ACMG Classification	XCI pattern	Evidence of X-skewing in the literature*	Escape gene**
113	F	MRXS33	300966	maternal	*TAF1*	NM_004606.5:c.745G>A	p.(Gly249Arg)	D 0.99 D D	Class3; PM2, PP1, PP2, PP3	100	yes	no
NWM24	F	BFLS	301900	*de novo*	*PHF6*	NM_001015877.2:c.890G>T	p.(Cys297Phe)	D 0.99 D D	Class5; PM1, PM2, PP2, PP3, PS2	90	yes	no
237	F	MRXSJ	300534	maternal	*KDM5C*	NM_004187.5:c.1204G>A	p.(Asp402Asn)	D 0.99 D D	Class4; PM2, PM5, PP2, PP1	90	yes	yes
234	M	MCAND	301056	maternal	*OTUD5*	NM_017602.4:c.1526C>T	p.(Pro509Leu)	D 0.99 N D	Class3, PM2, PP1, PP2	100	yes	no
236	M	MRXFH1	309580	maternal	*ATRX*	NC_000023.10:g.76953033_76959004del	p.(?)	/ / / /	Class5	100	yes	no
TF110	F		* 300634	*de novo*	*PDZD4*	NM_032512.5:c.10_16del	p.(Asn4Alafs*12)	/ / / /	Class4; PM2, PM4, PP3, PS2	96	N.A.	no

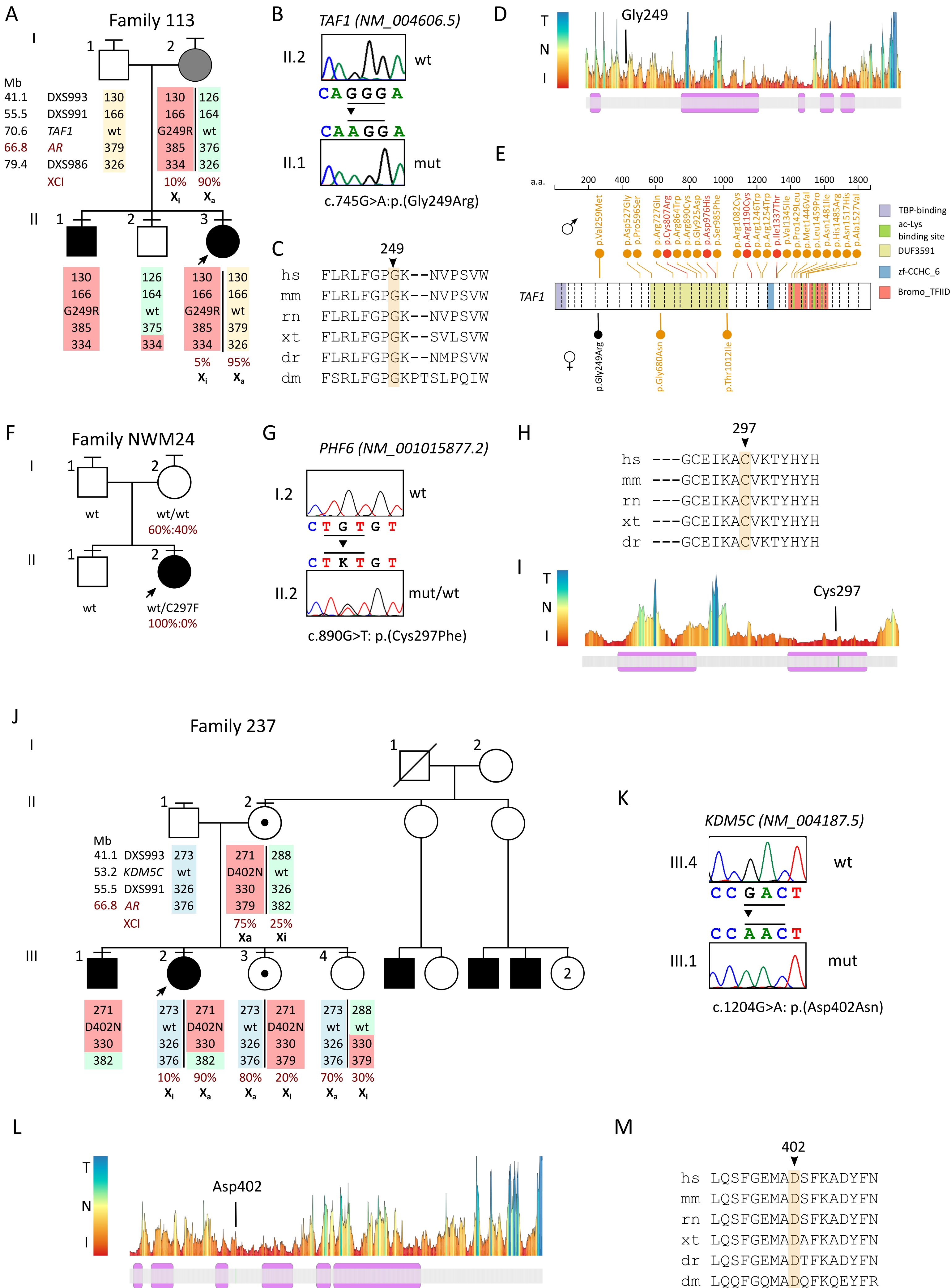
NWM25	M		* 300634	maternal	*PDZD4*	NM_032512.5:c.2190G>C	p.(Lys736Asn)	D 0.99 D D	Class3, PM2, PP3	91	N.A.	no
NWM127	M		* 300061	maternal	*ZMYM3*	NM_201599.3:c.1322G>A	p.(Arg441Gln)	T 0.99 / D	Class4, PS1, PM2	100	N.A.	no

Notes. *Data from(17); **Data from(38); D = Deleterious; N =Neutral. Variants submitted to ClinVar (SUB12130121)

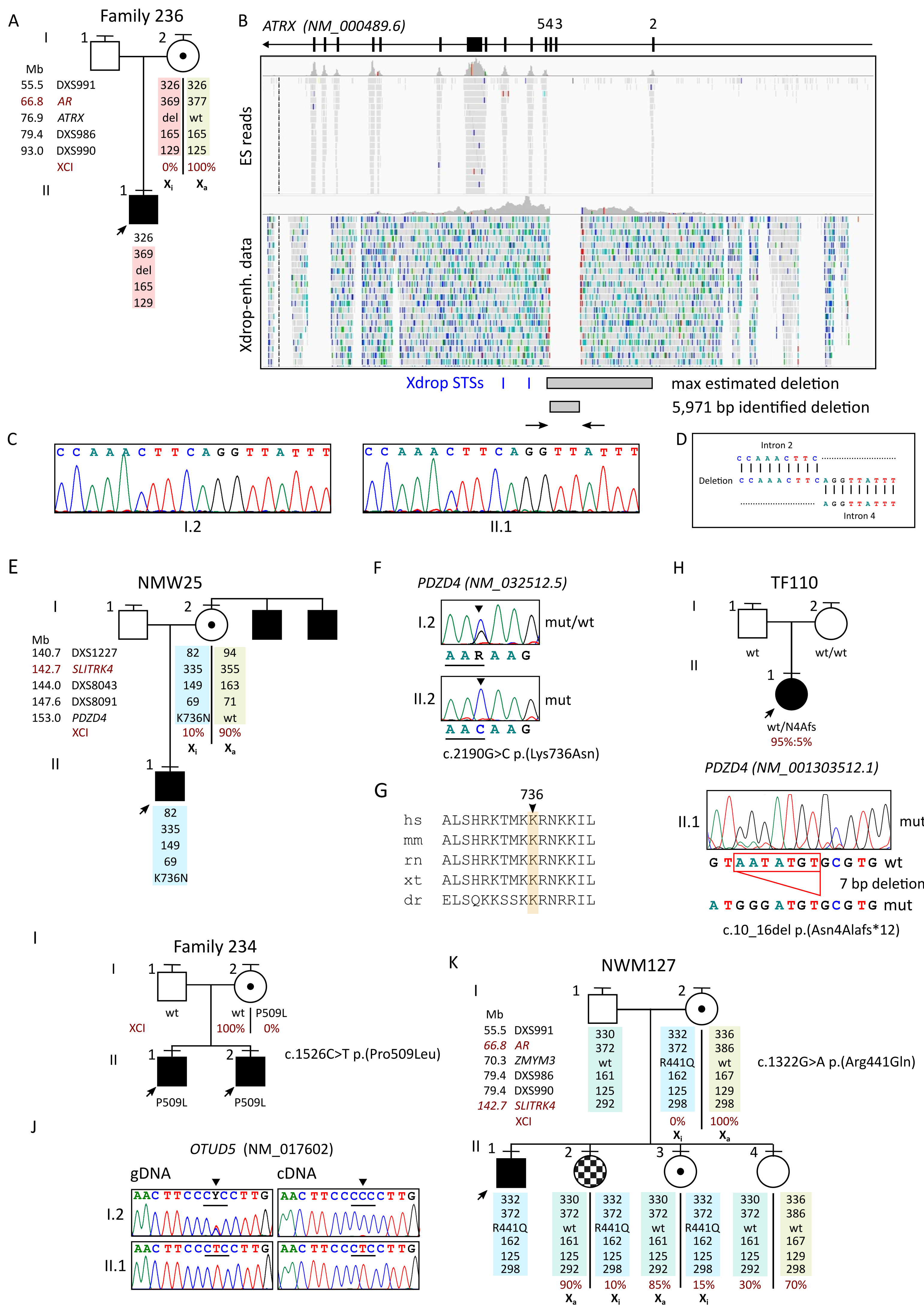
A

B

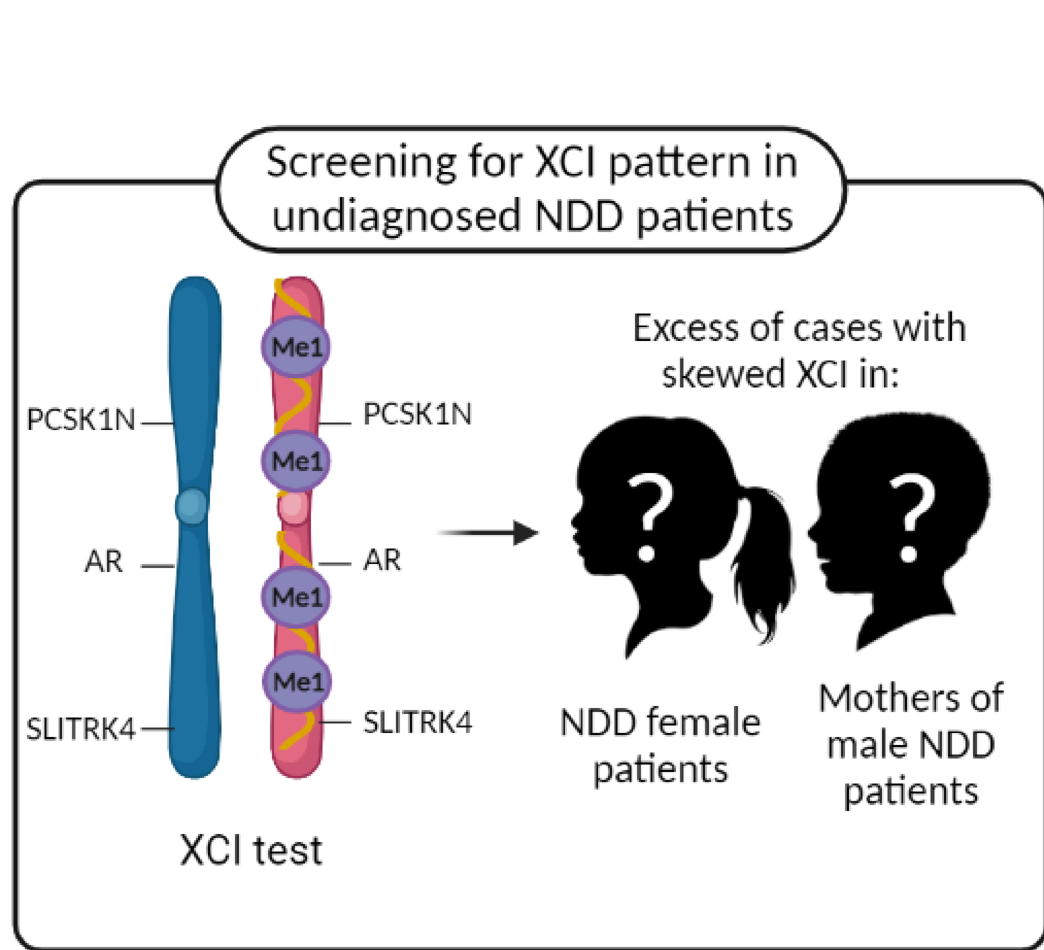




A
Family 113
I
1
2
Mb
41.1 DXS993 130 | 130 126
55.5 DXS991 166 | 166 164
70.6 TAF1 wt | G249R wt
66.8 AR 379 | 385 376
79.4 DXS986 326 | 334 326
XCI
10% 90%
Xi Xa
II
1
2
3
130 166 G249R 385 334
126 164 wt 375 334
130 166 G249R 385 334 | 130 166 wt 379 326
5% 95%
Xi Xa
B
TAF1 (NM_004606.5)
II.2
wt
CAGGGA
CAAGGA
II.1
mut
c.745G>A:p.(Gly249Arg)
C
249
hs FLRLFGPGK--NVPSVW
mm FLRLFGPGK--NVPSVW
rn FLRLFGPGK--NVPSVW
xt FLRLFGPGK--SVLSVW
dr FLRLFGPGK--NMPSVW
dm FSRLFGPGKPTSLPQIW
D
T
N
I
Gly249
E
a.a.
200 400 600 800 1000 1200 1400 1600 1800
p.Val259Met
p.Asp527Gly
p.Pro596Ser
p.Arg727Gln
p.Cys807Arg
p.Arg864Trp
p.Arg890Cys
p.Gly925Asp
p.Asp976His
p.Ser985Phe
p.Arg1082Cys
p.Arg1190Cys
p.Arg1246Trp
p.Arg1254Trp
p.Ile1337Thr
p.Val1345Ile
p.Pro1429Leu
p.Met1446Val
p.Leu1459Pro
p.Asn1481Ile
p.His1485Arg
p.Asn1517His
p.Ala1527Val
TAF1
p.Gly249Arg
p.Gly680Asn
p.Thr1012Ile
TBP-binding
ac-Lys binding site
DUF3591
zf-CCHC_6
Bromo_TFIID
F
Family NWM24
I
1
2
wt
wt/wt
60%:40%
II
1
2
wt
wt/C297F
100%:0%
G
PHF6 (NM_001015877.2)
I.2
wt
CTGTGT
CTKTGT
II.2
mut/wt
c.890G>T: p.(Cys297Phe)
H
297
hs ---GCEIKACVKTYHYH
mm ---GCEIKACVKTYHYH
rn ---GCEIKACVKTYHYH
xt ---GCEIKACVKTYHYH
dr ---GCEIKACVKTYHYH
I
T
N
I
Cys297
J
Family 237
I
1
2
II
1
2
Mb
41.1 DXS993 273 | 271 288
53.2 KDM5C wt | D402N wt
55.5 DXS991 326 | 330 326
66.8 AR 376 | 379 382
XCI
75% 25%
Xa Xi
III
1
2
3
4
2
271 D402N 330 382
273 wt 326 376 | 271 D402N 330 382
10% 90%
Xi Xa
273 wt 326 376 | 271 D402N 330 379
80% 20%
Xa Xi
273 wt 326 376 | 288 wt 330 379
70% 30%
Xa Xi
K
KDM5C (NM_004187.5)
III.4
wt
CCGACT
CCAACT
III.1
mut
c.1204G>A: p.(Asp402Asn)
L
T
N
I
Asp402
M
402
hs LQSFGEMADSFKADYFN
mm LQSFGEMADSFKADYFN
rn LQSFGEMADSFKADYFN
xt LQSFGEMADAFKADYFN
dr LQSFGEMADTFKADYFN
dm LQQFGQMADQFKQEYFR



A
Family 236
Mb
55.5 DXS991
66.8 AR
76.9 ATRX
79.4 DXS986
93.0 DXS990
XCI
326 326
369 377
del wt
165 165
129 125
0% 100%
Xi Xa
326
369
del
165
129
B
ATRX (NM_000489.6)
543
2
ES reads
Xdrop-enh. data
Xdrop STSs
max estimated deletion
5,971 bp identified deletion
C
I.2
II.1
D
Intron 2
Deletion
Intron 4
E
NMW25
Mb
140.7 DXS1227
142.7 SLITRK4
144.0 DXS8043
147.6 DXS8091
153.0 PDZD4
XCI
82 94
335 355
149 163
69 71
K736N wt
10% 90%
Xi Xa
82
335
149
69
K736N
F
PDZD4 (NM_032512.5)
I.2
mut/wt
II.2
mut
c.2190G>C p.(Lys736Asn)
G
736
hs ALSHRKTMKKRNKKIL
mm ALSHRKTMKKRNKKIL
rn ALSHRKTMKKRNKKIL
xt ALSHRKTMKKRNKKIL
dr ELSQKKSSKKRNRRIL
H
TF110
wt
wt/wt
wt/N4Afs
95%:5%
PDZD4 (NM_001303512.1)
II.1
mut
wt
7 bp deletion
mut
c.10_16del p.(Asn4Alafs*12)
I
Family 234
wt
wt P509L
XCI
100% 0%
P509L
P509L
c.1526C>T p.(Pro509Leu)
J
OTUD5 (NM_017602)
gDNA
cDNA
I.2
II.1
K
NWM127
Mb
55.5 DXS991
66.8 AR
70.3 ZMYM3
79.4 DXS986
79.4 DXS990
142.7 SLITRK4
XCI
330 332 336
372 372 386
wt R441Q wt
161 162 167
125 125 129
292 298 298
0% 100%
Xi Xa
c.1322G>A p.(Arg441Gln)
90% 10% 85% 15% 30% 70%
Xa Xi Xa Xi


Screening for XCI pattern in undiagnosed NDD patients
PCSK1N
AR
SLITRK4
Me1
Me1
Me1
Me1
PCSK1N
AR
SLITRK4
XCI test
Excess of cases with skewed XCI in:
?
?
NDD female patients
Mothers of male NDD patients
XCI modulate the expression of a pathogenetic X-linked variant generating different scenarios
Healthy mothers of male NDD patients
Cells inactivating the wild type allele are selected against
Female NDD patients
Skewing towards a deleteriuos allele that reach pathogenicity treshold
Residual expression of the deleteriuos allele resulting in a mild phenotype