An emerging female phenotype with loss of function mutations in the *Aristaless*-related homeodomain transcription factor *ARX*

Tessa Mattiske[1, 2], Ching Moey[1, 2], Lisenka E. Vissers[3], Natalie Thorne[4, 5, 6], Peter Georgeson[6,7], Madhura Bakshi [8] and Cheryl Shoubridge[1, 2].

[1] Department of Paediatrics, School of Medicine, University of Adelaide, SA 5005, Australia.
[2] Robinson Research Institute, University of Adelaide, SA 5005, Australia.
[3] Department of Human Genetics, Donders Institute for Brain, Cognition, and Behaviour, Radboud University Medical Center, 6500 HB Nijmegen, the Netherlands.
[4] Murdoch Children's Research Institute, Melbourne, Australia.
[5] University of Melbourne, Melbourne, Australia.
[6] Melbourne Genomics Health Alliance, Melbourne, Australia
[7] Victorian Life Sciences Computation Initiative, The University of Melbourne, Australia
[8]Department of Clinical Genetics, Liverpool Hospital, NSW 1871, Australia.

*Corresponding author:

A Prof Cheryl Shoubridge

Discipline of Paediatrics, Adelaide School of Medicine

Robinson Research Institute,

Faculty of Health Sciences

The University of Adelaide

Adelaide, SA 5005

AUSTRALIA

Phone: 61 8 8313 2355

Email: cheryl.shoubridge@adelaide.edu.au

This work was funded by the Australian National Health and Medical Research Council (Grant No. 1063025). CS is supported Australian Research Council (Future Fellowship FT120100086).

Abstract:

The devastating clinical presentation of X-linked lissencephaly with abnormal genitalia (XLAG) is invariably caused by loss of function mutations in the *Aristaless*-related homeobox (*ARX)* gene. Mutations in this X-chromosome gene contribute to intellectual disability (ID) with co-morbidities including seizures and movement disorders such as dystonia in affected males. The detection of affected females with mutations in *ARX* is increasing. We present a family with multiple affected individuals, including two females. Two male siblings presenting with XLAG were deceased prior to full term gestation or within the first few weeks of life. Of the two female siblings, one presented with behavioural disturbances, mild ID, a seizure disorder and complete agenesis of the corpus callosum, similar to the mother's phenotype. A novel insertion mutation in Exon 2 of *ARX* was identified, c.982delCinsTTT predicted to cause a frameshift at p.(Q328Ffs*37). Our finding is consistent with loss-of-function mutations in *ARX* causing XLAG in hemizygous males and extends the findings of ID and seizures in heterozygous females. We review the reported phenotypes of females with mutations in *ARX* and highlight the importance of screening *ARX* in male and female patients with ID, seizures and in particular with complete agenesis of the corpus callosum.

Keywords: X-linked lissencephaly-2, X-linked lissencephaly, *ARX*, Aristaless-related homeobox, Intellectual Disability, Seizure, LISX2, XLAG

Introduction:

The *Aristaless*-related homeobox gene (*ARX*; NM_139058.2; MIM# 300382) (Shoubridge, et al., 2010) is critical for early development and formation of a normal brain (Kitamura, et al., 2002), (Ohira, et al., 2002), (Stromme, et al., 2002). This paired-type homeodomain transcription factor plays a vital role in telencephalic development specifically in tangential migration and differentiation of GABAergic and cholinergic neurons (Kitamura, et al., 2002), (Colombo, et al., 2007), (Friocourt, et al., 2008), (Colasante, et al., 2009), (Lee, et al., 2014). Mutations in *ARX* result in a range of phenotypes with intellectual disability (ID) as a consistent feature. Mutations leading to loss of function of the ARX protein typically lead to brain malformation phenotypes, including X-linked lissencephaly with abnormal genitalia (XLAG, also known as LISX2) (MIM# 300215) (Kitamura, et al., 2002), (Shoubridge, et al., 2010). XLAG is a developmental disorder characterized by structural brain anomalies leading to agyria (absent cerebral grooves brain) or pachygygria (reduced folds or grooves) and agenesis of the corpus callosum. In addition, patients commonly have early-onset intractable seizures, severe psychomotor retardation, and ambiguous genitalia (Dobyns, et al., 1999; Kitamura, et al., 2002). Males are severely affected and often die within the first days or months of life.

As the *ARX* gene is on the X-chromosome, mutations in this gene typically result in families with affected males across multiple generations transmitted via (usually) asymptomatic

carrier females. However, there is an increasing prevalence of reported mutations in *ARX* that result in the severe phenotypic outcomes of XLAG in male patients that, with variable penetrance, affect females within the families resulting in a generally milder phenotype than affected males (Eksioglu, et al., 2011; Kato, et al., 2004; Kitamura, et al., 2002; Marsh, et al., 2009; Scheffer, et al., 2002; Stromme, et al., 2002). Here we report a novel mutation in *ARX* in a family ascertained by a female proband displaying a phenotype of mild learning disabilities, seizure disorder and agenesis of the corpus callosum. As part of this work we have reviewed reported phenotype of females with mutations in *ARX* and recommend screening of the *ARX* gene in female patients with suitable intellectual disability, seizure phenotypes and corpus callosum agenesis, particularly if there is evidence of X-linkage and no surviving males. *ARX* adds to a growing list of genes on the X chromosome including genes such as *USP9X*, *PHF6* and *IQSEC2* that have phenotypic effects in males and females that are distinct depending on the functional severity of the mutation (Reijnders, et al., 2016; Zarem *et al* 2016; Zweier, et al., 2013).

Materials and Methods:

Clinical description of patient and family

The proband presented to a Genetics Clinic at 10 years of age with learning difficulties, mild ID and seizures. The proband was born at term weighing 3.2kg, with no admission into the neonatal intensive care unit nor special-care nursery required. She was reported as sitting with support at 9 months of age with single words spoken at 9-10 months of age and walking at 22 months of age. Seizure onset was around 5 years of age with the first seizure classified as a prolonged generalized tonic-clonic which required intubation. At this time she was assessed and was able to draw a triangle with help, able to hop, dress herself and talk in simple sentences, knowing a few colours and numbers. No unusual grasping is reported. She was evaluated using the WISC-IV Australian at age 8 years and 10 months. She scored in the "Extremely Low" range for verbal comprehension, perceptual reasoning index, working memory index, processing speed index with a resulting full scale IQ in the "Extremely Low" range. She was assessed as operating in the Mild range of intellectual disability. She subsequently had a number of complex partial seizures, which were reasonably well controlled on Tegretol. Brain MRI revealed complete agenesis of the corpus callosum with mild dilated ventricles and colpocephaly. No lissencephaly, dysmorphic features or behaviour problems were reported.

The mother of the proband was aged 46 years old at the time of this report and presented with mild ID, seizures and mental health challenges. After admission to the public health hospital she was diagnosed with borderline personality disorder and major depressive disorder, however no neuropsychiatric assessment is currently available. She was non dysmorphic. History revealed that her first seizure was around 7-8 years of age and classified as complex

partial seizures. Brain MRI scan done at the age of 38 showed complete agenesis of the corpus callosum, with no grey matter heterotopia. Small white matter lesions were noted which comprised of tiny FLAIR (fluid-attenuated inversion recovery) hyperintensities involving the left centrum semiovale bilaterally and in the frontal region. The treating neurologist at the time felt these were specific white matter hyperintensities however the actual films are no longer available for review. In addition, EEG showed intermittent spike and wave discharge maximal over left hemispheres, which were at frequency of 2.5-3.5 Hz.

Two male offspring of II-4 were affected and died in early infancy, or were terminated during pregnancy. A maternal half-brother (III-1) of the proband was born at 36 weeks of gestation with an onset of seizures 20mins after birth. III-1 died at 26 days of age. Investigation of brain morphology identified agenesis of the corpus callosum, lissencephaly, grey matter heterotopias and bilateral optic nerve hypoplasia. Genitalia were ambiguous with labioscrotal folds, bilateral gonads and microphallus. Pelvis ultrasonography revealed the presence of a bicornuate uterus, while a male urethra was confirmed with a genitogram. Karyotype analysis showed a normal male chromosomal constitution (46,XY). Facial dysmorphic features included a large head, large anterior fontanels, low set ears, and abnormal nails. III-3 eventuated with a medical termination of pregnancy due to identification at 18 weeks gestation via ultrasonography of ventriculomegaly and abnormal genitalia. Brain malformations were confirmed at post mortem showing lissencephaly, absent corpus callosum, wide-open sylvian fissure and dilated ventricles. Facial dysmorphic features included flattened facies, mild macrocephaly and wide open sutures. Abnormal genitalia consisted of rudimentary genitalia with a small phallus. Karyotype analysis established the presence of male chromosomes.

Another pregnancy, III-5 was terminated when chorionic villus sampling revealed 45XO after ultrasound showed fetal hydrops. The remaining sibling (III-2) is a healthy female with normal intelligence. A maternal uncle's (II-1) death at one month of age was attributed to sudden infant death syndrome. The maternal grandmother (I-2) is reported to be healthy with no seizures.

Molecular analysis of *ARX* gene

The screening protocols were approved by the appropriate institution review boards and informed consent was obtained from the parents of patients. Genomic DNA from the proband was extracted from whole blood using standard techniques. Each of the five exons of *ARX* was amplified by PCR using primers designed to amplify coding and flanking non-coding sequence. The exception to this was exon 2, for which four overlapping amplicons were used to achieve robust amplification of GC-rich regions coding for three polyalanine tracts. The PCR conditions and primer sequences are described in detail previously (Tan, et al., 2013). Sequencing reactions were performed using ABI Big Dye terminator chemistry version 3.1 and purified products subjected to an automated capillary sequencing on ABI 3100 sequencer (Applied Biosystems, Foster City, CA, USA) and sequence was compared to the *ARX* reference sequence (NM_139058.2) using SeqMan module of the Lasergene DNA and protein analysis software package (DNAStar, Inc., Madison, WI).

Results

Molecular Analysis of the *ARX* variant

The diagnoses of intellectual disability and seizures in the female proband in conjunction with XLAG in her male siblings, from two different fathers, prompted evaluation of the *ARX* gene. Sequence analysis demonstrated a novel indel mutation, c.982delCinsTTT in exon 2 of the *ARX* gene, that is predicted to result in the creation of a premature stop codon, p.(Q328Ffs*37) (LOVD ID 0000128956) (http://www.lovd.nl/ARX). The child's mother presented with a similar phenotype and was confirmed to also be a symptomatic carrier of this novel *ARX* mutation. The father (II-5) and unaffected sister (III-2) do not carry the mutation (Figure 1). The *ARX* mutation was confirmed in III-4 in genomic DNA extracted from amniocytes. The amino acid affected by the mutation p.328Q is located at the start of the region containing the second nuclear localisation signal (NLS2) preceding the conserved paired-type homeodomain. The predicted stop codon arising from this variant occurs within 29 nucleotides of the exon 3-4 junction, and is likely to escape nonsense-mediated decay (NMD). Due to restricted levels of *ARX* expression in the patient derived materials we are unable to confirm if this truncated protein is produced. Despite this, even if the predicted C-terminal truncated protein was produced and subsisted at appreciable levels, the severe XLAG phenotype in affected male patients is expected with this variant resulting in a nonsense peptide being transcribed after residue p.328 and complete loss of the paired-type homeodomain residues (Figure 1C).

A methylation-specific PCR at the human FraxA and Androgen Receptor genes was performed on genomic DNA from blood as previously described (Plenge, et al., 1997). X-inactivation studies for both the proband (III-3) and the healthy sister (III-2) detected no significant deviation.

As part of initial investigations, both II-4 and III-3 were identified to carry a duplication on chromosome 5 at 5p15.5 (Chr5:10,907,608-11,363,857). This duplication contains only part of one gene, *Catenin Delta 2* (*CTNND2*). This duplication was classified as a variant of unknown significance. This duplication was later also identified in the female sibling III-2, who is healthy and has no learning issues, indicating this copy number variant is unlikely to contribute to the phenotype of the proband, her mother or affected brothers.

Discussion:

Here we report a family with a novel truncating mutation (c.982delCinsTTT/p.(Q328Ffs*37)) in *ARX*. The mutation is predicted to yield a non-functional protein product after p.328 due to the nonsense peptide prior to truncation of the protein at amino acid 364. This variant was not found in either the ExAC or 1000 genome project databases. Although this mutation may escape nonsense mediated decay, the resulting protein will have functional loss of the homeodomain and Aristaless domains. The catastrophic phenotype of XLAG reported in 2 males in this family is consistent with the predicted disruption of the ARX homeodomain function. The female proband and mother have a milder phenotype consisting of ID, seizures and agenesis of the corpus callosum. The investigation of *ARX* as a cause of the phenotype in the female proband was due largely to the distinctive phenotypic presentation and early deaths of the affected male siblings.

Our report underscores that a carrier female phenotype is likely to be under ascertained for *ARX*. This is supported by a recent examination of heterozygous females from families identified with *ARX* mutations (Marsh, et al., 2009) and examples of gender bias (92% male:

8% female) in a recent cohort of patients referred for molecular analysis of *ARX* (Marques, et al., 2015). Under ascertainment may be occurring due to several contributing factors. Patients with phenotypes such as intellectual disability and infantile spasms have been typically screened for mutations in known genes such as *STXBP1*, *CDKL5*, *KCNQ2*, *GRIN2A*, *MAGI2* and *ARX*. However, in the case of *ARX*, screening is commonly only considered in affected males. Moreover, as the majority of all mutations reported to date in *ARX* lead to expansion of the first or second polyalanine tracts, both located in exon 2, *ARX* mutation analysis is routinely limited to screening exon 2 and often using size variant analysis approaches (Marques, et al., 2015). The expanding use of next generation sequencing approaches on cohorts of individuals with intellectual disability and or seizure phenotypes are likely to provide a platform to potentially overcome some of these biases. However, even these types of approaches have constraints that need to be considered, particularly for genes with high GC content or near regions of low gene density, such as *ARX*. For example, sequence coverage across the *ARX* gene in 50 representative whole exome sequencing (WES) experiments undertaken at the Radboud University Medical Centre showed the median coverage for *ARX* was 40-fold less than the coverage of all ID genes; with the median percentages of *ARX* bases covered at 20x only being 73% compared to 97% for all ID genes (Supp. Figure S1). However, experience at this, and other centres, indicates that the coverage differences although lower for *ARX* generally, may also depend upon the region of the gene being considered (Supp. Figure S2). Exome sequencing using benchtop ion proton machines also result in poor coverage of the *ARX* gene, with mean coverage reported at 43% (Lacoste, et al., 2016). It remains to be determined how the increasing use of whole genome sequencing approaches as well as improvements to WES technologies and analysis pipelines address some of these coverage issues.

To date there have been both familial and *de novo* cases of affected females due to *ARX* mutations (Bettella, et al., 2013; Eksioglu, et al., 2011; Kato, et al., 2004; Kwong, et al., 2015; Marsh, et al., 2009; Scheffer, et al., 2002; Wallerstein, et al., 2008) (Table 1). Penetrance is variable both within and across families, with 55% of carrier females in these families presenting with a phenotype of intellectual disability to some degree with and without seizures (Table 1). Intellectual disability and/or developmental delay are prominent features in affected females across all families. A seizure phenotype was reported in 64% of the females with ID which equates to ~ 35% of all females in these families (Table 1).

Similar to the novel case we report here, affected females have been reported in families in which the male proband presents with severe XLAG or seizure phenotypes (Eksioglu, et al., 2011; Kato, et al., 2004; Marsh, et al., 2009; Scheffer, et al., 2002). In these familial cases ascertained by the male proband, the mutations are either missense mutations of residues in the homeodomain or nonsense/ deletion mutations resulting in a loss of function of the ARX homeodomain and/or aristaless domain activity (Table 2). Similarly, a smaller number of *de novo* cases also result in truncation and loss of *ARX* function (Bettella, et al., 2013; Kwong, et al., 2015; Marsh, et al., 2009; Wallerstein, et al., 2008). Across these cases there is a consistent phenotype of intellectual disability and/or developmental delay, infantile seizures and hypotonia/ dystonia/ ataxia (Table 3). The type and location of mutations in affected females is restricted compared to those contributing to affected males (Figure 2).

Brain MRI imaging is reported in approximately 35% of all females in these cases/ families, including the two females from this current report. Interestingly, of those individuals

screened, 73% had features consistent with agenesis of the corpus callosum (ACC) but only 64% of these patients were those classed as symptomatic based on intellectual disability and/or developmental delay and seizure phenotypes. Hence, there are a number of asymptomatic carrier females that do not display these key clinical features but still have the brain abnormality of ACC. There is also variable penetrance of both psychiatric and behavioural phenotypes across the symptomatic and asymptomatic females within these families. Movement disorders including ataxia and hypotonia are noted, particularly prevalent in the cases of *de novo* mutations (Table 3).

Given that the *ARX* gene is located on the X chromosome and is subject to X chromosome inactivation (XCI), the contribution of skewed X-inactivation (80:20) to the phenotype in heterozygous females is always a consideration. However, the female proband in this current report detected no skewing of X-inactivation in samples collected from blood. X-inactivation in previous studies has been inconclusive with little correlation of the affected status in heterozygous females to preferential inactivation of the normal *ARX* allele (measured in blood) (Marsh, et al., 2009). Indeed, skewing of X-inactivation that is not consistent with disease severity has been reported for other X-linked genes that were originally thought to affect males but have had affected female cases described, including mutations in genes such as *USP9X*, *MTM1*, *SLC9A6* and *MED12* (Prontera, et al., 2016; Reijnders, et al., 2016; Savarese, et al., 2016; Sinajon, et al., 2016). This highlights the complexity of X-linked inheritance. Recently, the variability and complex X-inactivation status within the brain has been demonstrated in an *Arx* knockout mouse model (Marsh, et al., 2016). Heterozygous female mice only have one functional copy of *Arx* and consistent with human data, female mice showed great variation of phenotype manifestations. These mice display striations of

radially oriented streams of *Arx* positive neurons within and emerging from the ganglionic eminence (GE) ventricular zone, which were noted to vary dramatically between embryos. This is consistent with the site of random X-inactivation suggested to occur in the ventral forebrain, followed by clonal proliferation of Arx positive or negative cells migrate radially away during early stages of development. Even this incomplete loss of *Arx* in female mice was detrimental and resulted in a change in the profile of interneurons in the adult female mice, consistent with a loss to a greater extent in hemizygous male mice.

We have identified a novel truncating mutation (c.982delCinsTTT/p.(Q328Ffs*37)) in *ARX* in a family ascertained by a female proband displaying a phenotype of mild intellectual disability, seizure disorder and agenesis of the corpus callosum, in conjunction with a phenotype of XLAG in her deceased male siblings. Review of the phenotypes of affected females with published mutations in *ARX* indicates that screening of the *ARX* gene in female patients with intellectual disability, seizure phenotypes and corpus callosum agenesis, particularly if there is evidence of X-linkage and no surviving males is warranted. Moreover, recent *de novo* mutations reported in females recommends scrutiny of *ARX* in all cases with suitable phenotypic presentation with and without other indications of X-linked inheritance. The emerging appreciation of phenotypic consequences in females with loss-of-function mutations in *ARX* will be important in counselling of affected families.

References

Bettella E, Di Rosa G, Polli R, Leonardi E, Tortorella G, Sartori S, Murgia A. 2013. Early-onset epileptic encephalopathy in a girl carrying a truncating mutation of the ARX gene: rethinking the ARX phenotype in females. Clin Genet **84**1:82-5.

Colasante G, Sessa A, Crispi S, Calogero R, Mansouri A, Collombat P, Broccoli V. 2009. Arx acts as a regional key selector gene in the ventral telencephalon mainly through its transcriptional repression activity. Dev Biol **334**:59-71.

Colombo E, Collombat P, Colasante G, Bianchi M, Long J, Mansouri A, Rubenstein JL, Broccoli V. 2007. Inactivation of Arx, the murine ortholog of the X-linked lissencephaly with ambiguous genitalia gene, leads to severe disorganization of the ventral telencephalon with impaired neuronal migration and differentiation. J Neurosci **27**:4786-98.

Dobyns WB, Berry-Kravis E, Havernick NJ, Holden KR, Viskochil D. 1999. X-linked lissencephaly with absent corpus callosum and ambiguous genitalia. Am J Med Genet **86**:331-7.

Eksioglu YZ, Pong AW, Takeoka M. 2011. A novel mutation in the aristaless domain of the ARX gene leads to Ohtahara syndrome, global developmental delay, and ambiguous genitalia in males and neuropsychiatric disorders in females. Epilepsia **52**:984-92.

Friocourt G, Kanatani S, Tabata H, Yozu M, Takahashi T, Antypa M, Raguenes O, Chelly J, Ferec C, Nakajima K, Parnavelas JG. 2008. Cell-autonomous roles of ARX in cell proliferation and neuronal migration during corticogenesis. J Neurosci **28**:5794-805.

Kato M, Das S, Petras K, Kitamura K, Morohashi K, Abuelo DN, Barr M, Bonneau D, Brady AF, Carpenter NJ, Cipero KL, Frisone F, et al. 2004. Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation. Hum Mutat **23**:147-59.

Kitamura K, Yanazawa M, Sugiyama N, Miura H, Iizuka-Kogo A, Kusaka M, Omichi K, Suzuki R, Kato-Fukui Y, Kamiirisa K, Matsuo M, Kamijo S, et al. 2002. Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. Nat Genet **32**:359-69.

Kwong AK, Ho AC, Fung CW, Wong VC. 2015. Analysis of mutations in 7 genes associated with neuronal excitability and synaptic transmission in a cohort of children with non-syndromic infantile epileptic encephalopathy. PLoS One **10**:e0126446.

Lacoste C, Desvignes JP, Salgado D, Pecheux C, Villard L, Bartoli M, Beroud C, Levy N, Badens C, Krahn M. 2016. Coverage Analysis of Lists of Genes involved in Heterogeneous Genetic Diseases following Benchtop Exome Sequencing using the Ion Proton. J Genet **95**:203-8.

Lee K, Mattiske T, Kitamura K, Gecz J, Shoubridge C. 2014. Reduced polyalanine-expanded Arx mutant protein in developing mouse subpallium alters Lmo1 transcriptional regulation. Hum Mol Genet **23**:1084-94.

Marques I, Sa MJ, Soares G, Mota Mdo C, Pinheiro C, Aguiar L, Amado M, Soares C, Calado A, Dias P, Sousa AB, Fortuna AM, et al. 2015. Unraveling the pathogenesis of ARX polyalanine tract variants using a clinical and molecular interfacing approach. Mol Genet Genomic Med **3**:203-14.

Marsh E, Fulp C, Gomez E, Nasrallah I, Minarcik J, Sudi J, Christian SL, Mancini G, Labosky P, Dobyns W, Brooks-Kayal A, Golden JA. 2009. Targeted loss of Arx results in a developmental epilepsy mouse model and recapitulates the human phenotype in heterozygous females. Brain **132**:1563-76.

Marsh ED, Nasrallah MP, Walsh C, Murray KA, Nicole Sunnen C, McCoy A, Golden JA. 2016. Developmental interneuron subtype deficits after targeted loss of Arx. BMC Neurosci **17**:35.

Ohira R, Zhang YH, Guo W, Dipple K, Shih SL, Doerr J, Huang BL, Fu LJ, Abu-Khalil A, Geschwind D, McCabe ERB. 2002. Human ARX gene: genomic characterization and expression. Molecular Genetics and Metabolism **77**:179-188.

Plenge RM, Hendrich BD, Schwartz C, Arena JF, Naumova A, Sapienza C, Winter RM, Willard HF. 1997. A promoter mutation in the XIST gene in two unrelated families with skewed X-chromosome inactivation. Nat Genet **17**:353-6.

Prontera P, Ottaviani V, Rogaia D, Isidori I, Mencarelli A, Malerba N, Cocciadiferro D, Rolph P, Stangoni G, Vulto-van Silfhout A, Merla G. 2016. A novel MED12 mutation: Evidence for a fourth phenotype. Am J Med Genet A **170**:2377-82.

Reijnders MR, Zachariadis V, Latour B, Jolly L, Mancini GM, Pfundt R, Wu KM, van Ravenswaaij-Arts CM, Veenstra-Knol HE, Anderlid BM, Wood SA, Cheung SW. 2016. De Novo Loss-of-Function Mutations in USP9X Cause a Female-Specific Recognizable Syndrome with Developmental Delay and Congenital Malformations. Am J Hum Genet **98**:373-81.

Savarese M, Musumeci O, Giugliano T, Rubegni A, Fiorillo C, Fattori F, Torella A, Battini R, Rodolico C, Pugliese A, Piluso G, Maggi L, et al. 2016. Novel findings associated with MTM1 suggest a higher number of female symptomatic carriers. Neuromuscul Disord **26**:292-9.

Scheffer IE, Wallace RH, Phillips FL, Hewson P, Reardon K, Parasivam G, Stromme P, Berkovic SF, Gecz J, Mulley JC. 2002. X-linked myoclonic epilepsy with spasticity and intellectual disability: mutation in the homeobox gene ARX. Neurology **59**:348-56.

Shoubridge C, Fullston T, Gecz J. 2010. ARX spectrum disorders: making inroads into the molecular pathology. Hum Mutat **31**:889-900.

Sinajon P, Verbaan D, So J. 2016. The expanding phenotypic spectrum of female SLC9A6 mutation carriers: a case series and review of the literature. Hum Genet **135**:841-50.

Stromme P, Mangelsdorf ME, Scheffer IE, Gecz J. 2002. Infantile spasms, dystonia, and other X-linked phenotypes caused by mutations in Aristaless related homeobox gene, ARX. Brain Dev **24**:266-8.

Tan MH, Gecz J, Shoubridge C. 2013. PCR amplification and sequence analysis of GC-rich sequences: Aristaless-related homeobox example. Methods Mol Biol **1017**:105-20.

Wallerstein R, Sugalski R, Cohn L, Jawetz R, Friez M. 2008. Expansion of the ARX spectrum. Clin Neurol Neurosurg **110**:631-4.

Zarem A, Haginoya K, Lev D, Blumkin L, Kivity S, Linder I, Shoubridge C, Palmer EE, Field M, Boyle J, Chitayat D, Gaillard WD, et al. 2016. The molecular and phenotypic spectrum of IQSEC2 related epilepsy. Epilepsia **57**:1858-1869

Zweier C, Kraus C, Brueton L, Cole T, Degenhardt F, Engels H, Gillessen-Kaesbach G, Graul-Neumann L, Horn D, Hoyer J, Just W, Rauch A, et al. 2013. A new face of Borjeson-Forssman-Lehmann syndrome? De novo mutations in PHF6 in seven females with a distinct phenotype. J Med Genet **50**:838-47.

Figure Legends

Figure 1: Identification of a c.982delCinsTTT mutation resulting in (p.(Q328Ffs*37) in ARX. A) Pedigree of immediate family members tested. Open symbols represent normal individuals, filled black circles represent females with moderate intellectual disability and seizures, filled squares represent males with Lissencephaly, hydrocephalus, profound developmental delay and ambiguous genitalia. Individual generations are numbered with Roman numerals on the left hand side of the pedigree. Individuals which were sequence confirmed to carry either the normal (N) or mutant (M) are shown B): DNA sequence electropherograms for the chrX: 25,031,130 (GRCh37/hg19 assembly) deletion of a C and insertion of TTT mutation in exon 2 of 5 of *ARX* reported in this study. A normal sequence was confirmed in unaffected father (II-5) and sister (III-2) with normal sequence shown. The mutation in the heterozygous state is shown in both the affected proband (III-4) and mother (II-4). The dotted line highlights the disrupted heterozygous trace present in females caused by the insertion (solid black underline). Additional mutation sequence change seen in the hemizygous state of an affected brother (III-3). C) Predicted functional consequence of a novel nonsense mutation in ARX. Schematic of the human ARX protein (top panel). Human ARX domains and regions are indicated above the schematic. Known functional domain are highlighted, octapeptide (OP) as horizontally hatched rectangle, nuclear localization sequences (NLS) as three black rectangles, polyalanine tracts (PA) as four white rectangles, acidic domain as vertically hatched rectangle, homeodomain as

crosshatched, and aristaless domain (OAR) as hatched. A schematic of the homeodomain and the flaking NLS domains (middle panel) with the amino acid sequence below (homeodomain sequence underlined). The amino acid change indicated by the first black arrow resulting in a nonsense peptide and a stop codon indicated by the second red arrow.

Figure 2: Identified ARX mutations in Females and Males leading to a range of phenotypes. (a) *ARX* exon structure in relation to the ARX protein. (b) Overview of ARX including the homeodomain (crosshatched) and aristaless domain (hatched) and the location of reported mutations according to their relative position at the protein level in both males and females. Differences in mutation type are indicated by a change in colour while missense mutation shown in grey and all other nonsense, insertion or deletion mutation in black. Phenotype severity is indicated on the y-axis with unaffected carrier females (below clinical threshold) shown below the dotted line.

Supplementary Figure 1: Radboud University Medical Center coverage analysis of 50 representative WES experiments of an ID gene panel consisting of 749 genes. Boxplots representing a) the median coverage of all ID genes in 101-fold, whereas for ARX this is 60-fold. b) The median percentage of bases covered at 10x for all ID gene is 100%, whereas this is 80% of bases for ARX. Similarly, the average percentage of basis covered at 20x for all ID genes is 97%, whereas this is 73% for *ARX*. (Gene Panel DG2.5.x;

https://www.radboudumc.nl/Informatievoorverwijzers/Genoomdiagnostiek/en/Pages/Intellectualdisability.aspx)

Supplementary Figure 2: Melbourne Genomics Health Alliance cohort of 250 WES samples analysed with valid coverage. a) The coverage for each sample across the 5 *ARX* exons, as well as the mean coverage for males, females and the overall means. b) The median coverage of ARX compared to the mean median coverage of all genes. Females are on the left of the grey vertical line.

Table 1: **Clinical Summary of Females with *ARX* mutations**

	Familial	This report	De novo	Total
Females	25	2	4	31
Females with ID, with and without Seizures	**11**	**2**	**4**	**17**
ID or DD	11	**2**	4	17
Seizures	5	**2**	4	11
Other clinical features	number (symptomatic : non-symptomatic)			
MRI reported	10 (5:5)	2 (2:0)	3	15
Brain malformation	9 (4:5)	2 (2:0)	3	14 (9:5)
ACC	8 (4:4)	2 (2:0)	1	11 (7:4)
Other	1 (0:1)	0	2	3 (3:0)
Movement disorder	4 (3:1)	0	4	8 (7:1)
Psychiatric features	4 (3:1)	1 (1:0)	0	5 (4:1)
Behavior disturbance	2 (2:0)	0	0	2 (2:0)

Table 2: Clinical features of females in familial cases of *ARX* mutations

Mutation cDNA / protein AA	Exon / Domain	Male Phenotype	Relationship to male proband	ID/DD	Seizures	Brain malformation	
Truncation mutations							
Exon1_2del	1, 2	XLAG	Mother	N	onset 12y (GTCS)		
Exon2_5del p.G66_C562del	2–5 HD + OAR	XLAG	Mother	N			
			Sister	ID		ACC + CVH	
c.232G>T p.E78X	2 HD + OAR	XLAG	Mother	N		ACC	
			Sister	N		ACC	
			Mother	N		ACC-p	
			Aunt	Mod ID +DD	onset 1y (GTCS)	ACC	
c.617delG p.G206Afs*119	2 HD + OAR	XLAG-HYD	Mother	N			
			Sister	ID	Unilateral ~ 9 weeks	ACC-p	
c.982delCinsTTT p.Q328Ffs*37	2 HD + OAR	XLAG	Proband	Mild ID	onset 5y	ACC	
			Mother	Mild ID	onset 7-8y	ACC	
c.1471_1472insC p.L491Pfs*41	5 OAR	OS, AG, ID	Mother	ID			An
			Mat aunt 1	N			
			Mat aunt 2	ID			Schiz
			Mat half aunt 1	ID +DD	Generalized		
			Mat half aunt 2	N			
			Mat grandmother	MD			
Missense mutations in the homeodomain							
c.998C>A p.T333N	2 HD	ACC/ AG	Mother	N		ACC-p	
			Cousin	Mod ID			
			Aunt	N			
			Aunt	Sev ID	onset 3mo	ACC-p	
c.1058C>T p.P353L	2 HD	XMESID	Mat grandmother	N		small vessel ischemic changes	
			Mat Aunt	ID			
			Mother	N			
c.1135C>A p.R379S	4 HD	ISSX	Mother	N			
			Aunt	N			
			Cousin	ID	onset 5y (absence)	N	

ACC = agenesis of the corpus callosum; ACC-p = partial ACC; / = not reported; ADD = attention deficit disorder; AG= ambiguous genitalia; CVH = cerebellar vermis hypoplasia; DD = developmental delay; HD = Homeodomain; ID = Intellectual Disability; ISSX = X-linked infantile spasm syndrome; Mat = Maternal; MD= Mild Delay; mo = months; Mod ID = Moderate ID; N=Normal; OAR = Aristaless; OS = Ohtahara Syndrome; PDD = Pervasive developmental disorder; Sev ID = Severe ID; y=years; XMESID = X-linked myoclonic epilepsy with generalized spasticity and ID

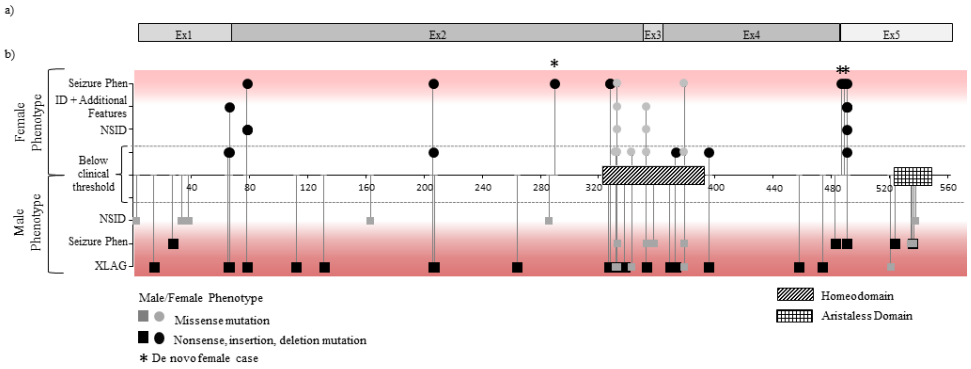
Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence for *ARX* gene [GenBank: NM_139058.2]

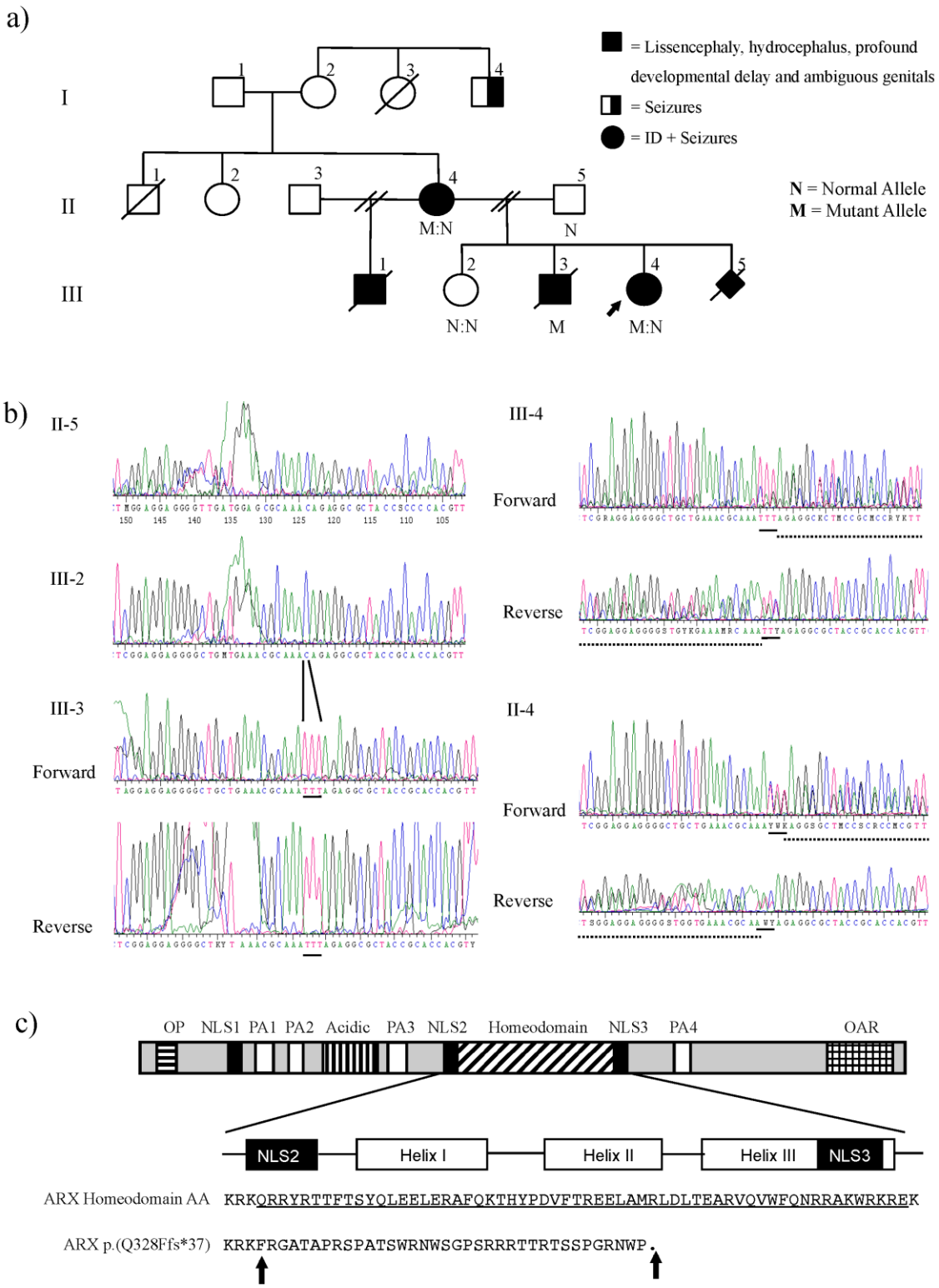
Table 3: Clinical features of females with *de novo* mutations in *ARX*

Mutation cDNA / protein AA	Exon / Domain	Mutation type	Development	ID	Seizures	Brain Malformation	Muscle Phenotype	
c.869C>A p.S290X	2 HD + OAR	NS	Sev DD		IEE onset 1m	microcephaly	dystonia	
c.1459delA p.T487Qfs*5	5 OAR	Del	GPD	ID	EIEE	N	significantly ataxic	Divergent S
c.1465delG p.A489Pfs*3	5 OAR	Del	DD	ID	ISSX (WS)	Cysts	hypotonia + torticollis	Poor visual tracki bilateral epicanth
Inversion	interruption of whole gene	Inversion	DD	ID	ISS onset in utero	ACC + HYD	Mild truncal hypotonia	Prominent forel bridge, sligl downturned corn nor

ACC-t = total agenesis of the corpus callosum; CVI = cortical visual impairment; DD = developmental delay; Del = deletion; EIEE – Early infantile epileptic encephalopathy; GPD = global psychomotor delay; HD = homeodomain; HYD = hydranencephaly; ID = Intellectual Disability; ISS = infantile spasm syndrome; ISSX = X-linked infantile spasm syndrome; N=Normal; NS = nonsense; OAR = aristaless domain; Sev Delay = Severe DD; UMN = upper motor neuron syndrome; WS = West Syndrome

Nucleotide numbering reflects cDNA numbering with +1 corresponding to the A of the ATG translation initiation codon in the reference sequence for *ARX* gene [GenBank: NM_139058.2]





Author/s:
Mattiske, T; Moey, C; Vissers, LE; Thorne, N; Georgeson, P; Bakshi, M; Shoubridge, C

Title:
An Emerging Female Phenotype with Loss-of-Function Mutations in the Aristaless-Related Homeodomain Transcription Factor ARX

Date:
2017-05-01

Citation:
Mattiske, T., Moey, C., Vissers, L. E., Thorne, N., Georgeson, P., Bakshi, M. & Shoubridge, C. (2017). An Emerging Female Phenotype with Loss-of-Function Mutations in the Aristaless-Related Homeodomain Transcription Factor ARX. HUMAN MUTATION, 38 (5), pp.548-555. https://doi.org/10.1002/humu.23190.