Published in final edited form as:

Neurogenetics. 2012 February ; 13(1): 23–29. doi:10.1007/s10048-011-0304-7.

Distinct DNA binding and transcriptional repression characteristics related to different *ARX* mutations

Ginam Cho¹, MacLean P. Nasrallah², Youngshin Lim³, and Jeffrey A Golden^{1,3}

¹Department of Pathology and Laboratory Medicine, The Children's Hospital of Philadelphia

²Neuroscience Graduate Group, Perelman School of Medicine at the University of Pennsylvania

³Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania

Abstract

Mutations in the *Aristaless-related homeobox* gene (*ARX*) are associated with a wide variety of neurologic disorders including lissencephaly, hydrocephaly, West syndrome, Partington syndrome, and X-linked intellectual disability with or without epilepsy. A genotype-phenotype correlation exists for *ARX* mutations, however the molecular basis for this association has not been investigated. To begin understanding the molecular basis for *ARX* mutations, we tested the DNA binding sequence preference and transcriptional repression activity for *Arx*, deletion mutants and mutants associated with various neurologic disorders. We found DNA binding preferences of *Arx* are influenced by the amino acid sequences adjacent to the homeodomain. Mutations in the homeodomain show a loss DNA binding activity, while the T333N and P353R homeodomain mutants still possess DNA binding activities, although less than wild type. Transcription repression activity, the primary function of *ARX*, is reduced in all mutants except the L343Q, which has no DNA binding activity and does not functionally repress *Arx* targets. These data indicate that mutations in the homeodomain result in not only a loss of DNA binding activity but also loss of transcriptional repression activity. Our results provide novel insights into the pathogenesis of *ARX* related disorders and possible directions to pursue potential therapeutic interventions.

Keywords

ARX; lissencephaly; X-linked intellectual disability and Homeodomain

INTRODUCTION

Mutations in the *Aristaless-related homeobox* gene (*ARX*) are identified in a broad spectrum of neurocognitive disorders. Patient phenotypes range from severe structural anomalies of the brain such as lissencephaly, to structurally normal appearing brains with intellectual disabilities [1,2]. Mutations in *ARX* have been associated with a variety of developmental processes including proliferation of neocortical neuronal progenitors, neuronal differentiation, and migration of interneurons [3-10]. These cellular defects are believed to underlie the various structural and functional phenotypes, observed in the patients.

ARX is a transcription factor with a paired-like homeodomain in addition to an octapeptide domain, four polyalanine tracts, three nuclear localization sequences (NLS) an acidic domain and an Aristaless domain [11-13]. It is expressed in the ventricular zone of the developing neocortex, the source of cortical excitatory neurons, and in the ganglionic eminence (GE) where cortical interneurons are derived [6,14-16]. Existing data suggest Arx controls neocortical ventricular zone cell proliferation and both radial and non-radial neuronal migration [8,9]. In addition, Arx has identified roles in olfactory bulb and pancreas development [17,18]. Both *in vitro* and *in vivo* studies have determined that Arx primarily functions as a transcriptional repressor, but that it can also act as an activator [5,18,12,19,13,20].

The homeodomain is composed of 60 highly conserved amino acids forming three alpha helices, which interact with the minor and major groove of the core sequence "TAAT" [21,22]. In addition to the alpha helices, homeodomain proteins possess an N-terminal arm that confers the DNA binding sequence specificity. Cooperative interaction with cofactors, including other homeodomain proteins, contributes to the DNA binding specificity for development-related gene regulation [23-26].

Pathogenic mutations in *ARX* have been found in multiple regions of the coding sequence and correlations have been made between genotype and phenotype [1,2]. Among the mutations identified in *ARX* are i) Expansion in the first or second polyalanine tract, ii) Frame-shift mutations resulting in a truncated protein, iii) deletions and iv) point mutations within and outside of the homeodomain. The mutations in the homeodomain allegedly disrupt DNA binding, although whether such point mutations prevent binding or result in binding off-target DNA sequences has not been well characterized.

In this study we investigated the DNA binding sites of Arx and tested the ability of several known human disease associated point mutations and deletion mutants to bind the canonical *ARX* DNA binding sequence. Furthermore, we determined whether these known point mutations, associated with human disease, maintained transcriptional repression activity. Our data provide novel insight into the functional defects associated with these human disease associated mutations.

METHODS and MATERIALS

Plasmids and Generation of mutants

Full length Arx was amplified by PCR using ArxF-E (5'-CGGAATTCCACCATGAGCAATCAGTACCAGGAAGAGGGC-3') and ArxR-MluI (5'-ACTTCAACGCGTCTACAGATCTTCTTCAGAAATAAGTTTTTGTTCGCA-3') and the product was cloned into the EcoRI and MluI sites of pCAG-IRES-EGFP. The pM-Arx constructs containing the Gal4 DNA binding domain were generated by PCR using primers ArxF-E and ArxR-Xb (5'-TGCTCTAGATTAGCACACCTCCTTCCCCGTGCTG-3'). The pTYB2-Sumo-Arx was constructed by cloning the amplified DNA fragment containing human *SUMO* and mouse *Arx* into pTYB2 (NEB). Sequences corresponding to amino acids 214-430, 286-430 of Arx were amplified by PCR using primers 286F-E (5'-CGGAATTCGAGAGGGCGGGGAGCTGTGCGC-3') and 430R-Xh (5'-CCGCTCGAGTTACGAATCAAGCGCAGGGTGATGCG-3'), digested with EcoRI and XhoI and then cloned into the EcoRI and XhoI sites of pGEX-5X-3 (Amersham Biosciences) or pET28a (Novagen) (Fig. 1a). Point mutations were introduced by a common PCR-mutagenesis protocol. The mouse sequence of Arx was used in all experiment although mutants were based on the human sequence number to preclude confusing amino acid residues number; mouse Arx is one amino acid shorter than human ARX in homeodomain region.

Protein purification

GST-Arx and *His-Arx* fusion proteins along with *SUMO-tagged Arx* were expressed in BL21 competent cells (Invitrogen) with expression induced by the addition of 0.3 mM IPTG to late logarithmic cultures (OD =0.5) for 3 h at 27°C. Cells were then harvested, resuspended in PBS (phosphate buffer saline pH 7.4) containing 0.5% triton X-100, and disrupted by sonication in the presence of a protease inhibitor-cocktail (Sigma). After centrifugation, the supernatants were applied to glutathione-Sepharose beads (Amersham Bioscience), Ni-NTA beads (Qiagen) and chitin bead (NEB), respectively. The beads were washed with PBS for glutathione beads and Ni-NTA, column buffer (20mM MES-HCl pH [6.5], 500mM NaCl) for chitin bead. Each fusion protein was eluted according to the manufacturer's instructions. The GST fusion protein was finally eluted with 30 mM glutathione in 100 mM Tris-Cl, pH 7.4. The His fusion protein was eluted with 250mM imidazole containing PBS and the Sumo fusion protein with 10 mM DTT

Electrophoretic mobility-shift assay

Electrophoretic mobility-shift assays (EMSA) were performed as previously described with slight modifications [19]. Oligonucleotides (5 pmol) were labeled by T4 polynucleotide kinase in the presence of [-32P] ATP (6000 Ci/mmol; Amersham Bioscience). Unincorporated [-32P] ATP was removed by column filtration using MicroSpin G-25 columns (Amersham Biosciences). The purified probe was annealed in Buffer H (Roche) by heating for 1 min and cooling to room temperature. One hundred nanograms of GST-Arx derivatives were incubated with 0.025 pmol of labeled-probe in total 20 l of binding buffer [20 mM HEPES pH 7.4, 50 mM KCl, 1 mM MgCl₂, 1 mM DTT, 5% glycerol, 0.05% Triton X100 and 1 g poly (dI-dC) (Roche)] for 30 min on ice. The mixture was loaded on a 5% polyacrylamide gel and electrophoresed (constant 33 mA) at 4°C in 1X TBE. Gels were dried and visualized on a phosphorimager (Molecular Dynamics). Oligonucleotides used were as follows: Lmo1: 5'- GTAATGAATTGATTAAATTAACAGGGGAGTCTGA-3', N11 (5'- GTCTGGCAGTTAATTCGGCTTAGTAGC-3'), N30 (5'- GTCTGGCCTCTAATTACGCCTAATTGC-3'), N9 (5'- GTCTGGCTCCTAATTAGTTCGATTAGC-3'), N26 (5'- GTCTGGCCTCTAATTACGCCTAATTGC-3'), N2 (5'- GTCTGGCAATTAATATATTGCATTTGC-3'), and N25 (5'- GTCTGGCCATTAATGATATCAATTAGC-3').

Cell culture, transfection and reporter assay

HEK293T cells were transfected 24 h after plating with 200 ng of luciferase reporter plasmid DNA, 50ng of pM-Arx or mutants, and 50 ng of pRL-TK-Renilla luciferase plasmid DNA (Promega) using FuGENE 6 (Roche Diagnostics, Alameda, CA). Forty-eight hours posttransfection, cell lysis and measurement of firefly and Renilla luciferase activity was performed using the Dual-Glo Luciferase Assay System (Promega) according to the manufacturer's instruction using a Veritux Microplate Luminometer (Turner BioSystems, Sunnyvale, CA). Transfections were performed in quadruplicate, and replicated in three independent experiments. The firefly luciferase activity was normalized according to the corresponding Renilla luciferase activity, and luciferase activity was reported as relative mean ($\pm$ SD).

SELEX

SELEX was performed as previously described with several modifications [27]. The double strand DNA containing the TAAT core sequences, random sequences and primer binding sites was generated by PCR using the following template (5'- TAGGCATGTGGATCCGTCTGGCN₃TAATN₁₁GCGTACCGGATCCACCTTCGAT-3')

and PCR primers (5'-TAGGCATGTGGATCCGTCTGGC-3' and 5'-ATCGAAGGTGGATCCGGTACGC-3'). The products were purified using the PCR extraction kit (Qiagen). DNA-protein binding was performed at 4°C for 30 min in reaction buffer (20mM Tris-HCl pH [7.2], 50mM KCl, 1mM EDTA, 1mM DTT, 1 g dIdC in 20 l reaction, 0.05% Triton X100 and 10% glycerol). After binding the glutathione beads (Amersham) were added to the reaction. The beads were washed 5 times with the reaction buffer without dIdC. DNA was eluted with 10mM glutathione containing reaction buffer. The eluted DNA was cleaned with the PCR extraction kit (Qiagen). The cleaned DNA was amplified with the primers for 18 cycles (annealing temperature 54°C for 30 sec) and this amplified DNA was then used for panning. This procedure was performed 6 times. The enriched DNA was subcloned into pCR2.1 vector and sequenced.

RESULTS and DISCUSSION

Arx can bind to 5'-T/cTAATT-3

To determine the DNA binding site of Arx on naked DNA, we performed SELEX (Systematic Evolution of Ligands by EXponential enrichment). As previously determined, Arx can bind to the TAAT core sequence found with most homeodomain proteins [28,19]. To characterize the specificity, we designed an oligonucleotide containing the fixed core sequence (N3**TAAT**N12; bold is core sequence) and tested the ability of the GST-Arx (286-430) protein containing the homeodomain (Fig. 1a), to bind. We performed a 6X pan for the DNA fragments to bind the GST-Arx with glutathione beads, testing DNA binding activities after each panning (Fig. 1b). The selected DNA sequence is the 5'-T/c**TAATT**-3'; 5' side of TAAT has T or C preference while the 3' side of TAAT has a T preference (Fig. 1b). These data indicate Arx requires sequence information adjacent to the 5'-TAAT-3' sequence, for DNA binding.

DNA binding preferences of Arx are variable depending on protein structure

We postulated that the adjacent amino acid sequence information near homeodomain would determine Arx DNA binding preference. To test this we performed electrophoretic mobility shift assays (EMSA) with full-length Arx protein and several deletion mutants (Fig. 2). The EMSA probes were selected from representatives found in the sequence pool determined by SELEX. We found that full-length Arx has a T/C/A preference on the 5' end and T preference on the 3' end of the TAAT core sequence (Fig. 2a and Table 1). We next considered whether Arx deletion mutants have similar DNA binding preferences. We designed two mutant constructs containing the homeodomain (located at aa 328-388). The first includes aa 286-430 and the second (aa 214-430) have a longer N-terminal sequence that contains the central acidic domain (aa 224-255) and polyalanine tract (aa 275-281). We expressed and purified these two deletion mutants, 214-430 and 286-430 (fused to GST or His-tag). Each mutant showed a unique DNA binding preference. The His-tagged 214-430 mutant protein exhibited the stronger DNA binding activity than the GST-tagged version of this same mutant (Fig. 2b-c and Table 1). In contrast, the 286-430 protein showed strong DNA binding activity when fused to GST, but lost all binding activity when fused to His-tagged (Fig. 2d-e and Table 1). Compared to full-length Arx, the 214-430 and 286-430 proteins lost their T preference on the 5' side when T was at the position 3' to the TAAT (Table 1). On the other hand, both deletion mutants maintained a T preference on the 5' side when the 3' end was a G (Table 1). GST and His tags have significant size differences; GST is 24 kD and His tag is 3kD. We predicate that the GST fusion could introduce a significant conformational change, which could cause to affect DNA binding affinity or preference. However, our result show that DNA binding affinity or preference was not altered by the fusion protein size, but a combination of the specific deletion and fusion partner. Taken together, these data indicate that DNA binding preferences are diverse depending on protein

structure influenced by both the fusion partner and the deletion. Furthermore, this implies that the Arx homeodomain's DNA binding preference could be influenced by its interaction partners including other transcription factors or transcriptional co-activators/repressors. These data lead us to reconsider the sequence information from the transcription factor binding site database, which were generated by using only the DNA binding domain.

Analysis of DNA binding activity of Arx homeodomain mutants

Seven missense mutations, leading to a single amino acid substitution in the homeodomain of ARX, have been identified in patients with intellectual disabilities [1]. Six of these mutations are correlated with structural brain malformations. The R332C/R332H/R332P, L343Q and P353R mutations are associated with X-linked lissencephaly with ambiguous genitalia (XLAG). In contrast, the T333N mutation is associated with agenesis of corpus collosum (Proud syndrome) [29,9,30]. Mutations at R332 or L343 are more severe than the T333 or P353 mutations [30]. One mutation, P353L has a milder phenotype with no structural defects but seizures and intellectual disability; this mutant has an amino acid change between conserved amino acids in the homeodomain [31]. Thus, different homeodomain point mutations can result in very different phenotypes. To understand how such difference might result from mutations in the same functional domain, we introduced these mutations into mouse *Arx* and tested their DNA binding activity and transcription repression activity. The five missense mutations associated with XLAG were all in highly conserved amino acids (Fig. 3a and b). This led us to hypothesize that all mutants have different DNA binding activity depending on the position. As expected, T333N and P353R bound DNA but with lower affinity than wild type *Arx* (Fig. 3c). These data are consistent with the DNA binding activities of mutants correlating with phenotype. Although we did not test the DNA binding activity for the P353L, we would predict this mutant would bind DNA, based on the sequence conservation at this position of the homeodomain.

Homeodomain mutants have loss of transcription repression activities

We previously identified *Lmo1*, *Ebf3* and *Shox2* as direct targets of Arx repression [19]. Given that some of the mutants have lost DNA binding activity (Fig. 3c), it would be impossible to test their transcription repression activity on these genes. Therefore, we employed a Gal4-based reporter gene assay system. This system does not require the DNA binding domain of Arx, instead, it utilizes the Gal4 DNA binding domain. We designed the Gal4 fusion protein expression constructs for wild type *Arx* or each point mutant. As shown in the previous report [13], wild type Gal4-Arx represses the activity of the Gal4 reporter. L343Q has a similar transcription repression activity to wild type. However, R332H, R332P, T333N and P353R mutants have a reduced transcription repression activity (Fig. 3d). Although T333N and P353R have similarly low DNA binding activities, our data suggests that the T333N has a more severe phenotype in patients because of reduced transcriptional repression. Taken together, our data suggest that the phenotypes associated with these mutations do not solely result from a loss of DNA binding activity but also the loss of transcription repressive activity.

In summary, we find that the DNA binding preferences of Arx and mutant forms of Arx are influenced by the evolutionally conserved amino acids in the homeodomain and its adjacent sequence information. Furthermore, we found that the phenotypic severity associated with Arx mutations is a result of both DNA binding activity and abrogation of transcriptional activity.

Acknowledgments

This work was supported by NIH grant NS46616.

Abbreviations

ARX	Aristaless-related homeobox gene
NLS	nuclear localization sequences
SELEX	Systematic Evolution of Ligands by EXponential enrichment
XLAG	X-linked lissencephaly with ambiguous genitalia
EMSA	Electrophoretic mobility-shift assays
MEME	Multiple EM for motif Elucidation

REFERENCE

- Shoubridge C, Fullston T, Gecz J. ARX spectrum disorders: making inroads into the molecular pathology. *Human Mutation*. 2010; 31(8):889–900. doi:10.1002/humu.21288. [PubMed: 20506206]
- Friocourt G. Mutations in ARX result in several defects involving GABAergic neurons. *Frontiers in cellular neuroscience*. 2010 doi:10.3389/fncel.2010.00004.
- Bonneau D, Toutain A, Laquerrière A, Marret S, Saugier-Verber P, Barthez M-A, Radi S, Biran-Mucignat V, Rodriguez D, Gélou A. X-linked lissencephaly with absent corpus callosum and ambiguous genitalia (XLAG): clinical, magnetic resonance imaging, and neuropathological findings. *Annals of neurology*. 2002; 51(3):340–349. [PubMed: 11891829]
- Colasante G, Collombat P, Raimondi V, Bonanomi D, Ferrai C, Maira M, Yoshikawa K, Mansouri A, Valtorta F, Rubenstein JLR, Broccoli V. Arx is a direct target of Dlx2 and thereby contributes to the tangential migration of GABAergic interneurons. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2008; 28(42):10674–10686. doi:10.1523/JNEUROSCI.1283-08.2008. [PubMed: 18923043]
- Colasante G, Sessa A, Crispi S, Calogero R, Mansouri A, Collombat P, Broccoli V. Arx acts as a regional key selector gene in the ventral telencephalon mainly through its transcriptional repression activity. *Developmental Biology*. 2009;1–13. doi:10.1016/j.ydbio.2009.07.014.
- Colombo E, Collombat P, Colasante G, Bianchi M, Long J, Mansouri A, Rubenstein JLR, Broccoli V. 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. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2007; 27(17):4786–4798. doi:10.1523/JNEUROSCI.0417-07.2007. [PubMed: 17460091]
- Forman MS, Squier W, Dobyns WB, Golden JA. Genotypically defined lissencephalies show distinct pathologies. *Journal of neuropathology and experimental neurology*. 2005; 64(10):847–857. [PubMed: 16215456]
- Friocourt G, Kanatani S, Tabata H, Yozu M, Takahashi T, Antypa M, Raguene O, Chelly J, Ferec C, Nakajima K, Parnavelas JG. Cell-Autonomous Roles of ARX in Cell Proliferation and Neuronal Migration during Corticogenesis. *The Journal of neuroscience: the official journal of the Society for Neuroscience*. 2008; 28(22):5794–5805. doi:10.1523/JNEUROSCI.1067-08.2008. [PubMed: 18509041]
- 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-i, Kasahara M, Yoshioka H, Ogata T, Fukuda T, Kondo I, Kato M, Dobyns WB, Yokoyama M, Morohashi K-i. Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. *Nature Genetics*. 2002; 32(3):359–369. doi:10.1038/ng1009. [PubMed: 12379852]
- Okazaki S, Ohsawa M, Kuki I, Kawawaki H, Koriyama T, Ri S, Ichiba H, Hai E, Inoue T, Nakamura H, Goto Y-I, Tomiwa K, Yamano T, Kitamura K, Itoh M. Aristaless-related homeobox gene disruption leads to abnormal distribution of GABAergic interneurons in human neocortex: evidence based on a case of X-linked lissencephaly with abnormal genitalia (XLAG). *Acta neuropathologica*. 2008; 116(4):453–462. doi:10.1007/s00401-008-0382-2. [PubMed: 18458920]

11. Lin W, Ye W, Cai L, Meng X, Ke G, Huang C, Peng Z, Yu Y, Golden JA, Tartakoff AM, Tao T. The roles of multiple importins for nuclear import of murine aristaless-related homeobox protein. *The Journal of biological chemistry*. 2009; 284(30):20428–20439. doi:10.1074/jbc.M109.004242. [PubMed: 19494118]
12. Fullenkamp AN, El-Hodiri HM. The function of the Aristaless-related homeobox (Arx) gene product as a transcriptional repressor is diminished by mutations associated with X-linked mental retardation (XLMR). *Biochemical and biophysical research communications*. 2008; 377(1):73–78. doi:10.1016/j.bbrc.2008.09.116. [PubMed: 18835247]
13. McKenzie O, Ponte I, Mangelsdorf M, Finnis M, Colasante G, Shoubridge C, Stifani S, Géczi J, Broccoli V. Aristaless-related homeobox gene, the gene responsible for West syndrome and related disorders, is a Groucho/transducin-like enhancer of split dependent transcriptional repressor. *Neuroscience*. 2007; 146(1):236–247. doi:10.1016/j.neuroscience.2007.01.038. [PubMed: 17331656]
14. Poirier K, Van Esch H, Friocourt G, Saillour Y, Bahi N, Backer S, Souil E, Castelnau-Ptakhine L, Beldjord C, Francis F, Bienvenu T, Chelly J. Neuroanatomical distribution of ARX in brain and its localisation in GABAergic neurons. *Brain research Molecular brain research*. 2004; 122(1):35–46. doi:10.1016/j.molbrainres.2003.11.021. [PubMed: 14992814]
15. Bienvenu T, Poirier K, Friocourt G, Bahi N, Beaumont D, Fauchereau F, Ben Jeema L, Zemni R, Vinet M-C, Francis F, Couvert P, Gomot M, Moraine C, van Bokhoven H, Kalscheuer V, Frints S, Gecz J, Ohzaki K, Chaabouni H, Fryns J-P, Desportes V, Beldjord C, Chelly J. ARX, a novel Prd-class-homeobox gene highly expressed in the telencephalon, is mutated in X-linked mental retardation. *Human Molecular Genetics*. 2002; 11(8):981–991. [PubMed: 11971879]
16. Miura H, Yanazawa M, Kato K, Kitamura K. Expression of a novel aristaless related homeobox gene 'Arx' in the vertebrate telencephalon, diencephalon and floor plate. *Mechanisms of development*. 1997; 65(1-2):99–109. [PubMed: 9256348]
17. Yoshihara, Si. Arx homeobox gene is essential for development of mouse olfactory system. *Development (Cambridge, England)*. 2005; 132(4):751–762. doi:10.1242/dev.01619.
18. Collombat P, Mansouri A, Hecksher-Sorensen J, Serup P, Krull J, Gradwohl G, Gruss P. Opposing actions of Arx and Pax4 in endocrine pancreas development. *Genes & development*. 2003; 17(20):2591–2603. doi:10.1101/gad.269003. [PubMed: 14561778]
19. Fulp CT, Cho G, Marsh ED, Nasrallah IM, Labosky PA, Golden JA. Identification of Arx transcriptional targets in the developing basal forebrain. *Human Molecular Genetics*. 2008; 17(23):3740–3760. doi:10.1093/hmg/ddn271. [PubMed: 18799476]
20. Seufert DW, Prescott NL, El-Hodiri HM. Xenopus aristaless-related homeobox (xARX) gene product functions as both a transcriptional activator and repressor in forebrain development. *Developmental dynamics: an official publication of the American Association of Anatomists*. 2005; 232(2):313–324. doi:10.1002/dvdy.20234. [PubMed: 15614781]
21. Gehring WJ, Qian YQ, Billeter M, Furukubo-Tokunaga K, Schier AF, Resendez-Perez D, Affolter M, Otting G, Wuthrich K. Homeodomain-DNA recognition. *Cell*. 1994; 78(2):211–223. doi: 0092-8674(94)90292-5 [pii]. [PubMed: 8044836]
22. Klemm JD, Rould MA, Aurora R, Herr W, Pabo CO. Crystal structure of the Oct-1 POU domain bound to an octamer site: DNA recognition with tethered DNA-binding modules. *Cell*. 1994; 77(1):21–32. doi:0092-8674(94)90231-3 [pii]. [PubMed: 8156594]
23. Joshi R, Passner JM, Rohs R, Jain R, Sosinsky A, Crickmore MA, Jacob V, Aggarwal AK, Honig B, Mann RS. Functional specificity of a Hox protein mediated by the recognition of minor groove structure. *Cell*. 2007; 131(3):530–543. doi:S0092-8674(07)01212-3 [pii] 10.1016/j.cell.2007.09.024. [PubMed: 17981120]
24. Bondos SE, Tan XX, Matthews KS. Physical and genetic interactions link hox function with diverse transcription factors and cell signaling proteins. *Mol Cell Proteomics*. 2006; 5(5):824–834. doi:M500256-MCP200 [pii] 10.1074/mcp.M500256-MCP200. [PubMed: 16455680]
25. Luo L, Yang X, Takihara Y, Knoetgen H, Kessel M. The cell-cycle regulator geminin inhibits Hox function through direct and polycomb-mediated interactions. *Nature*. 2004; 427(6976):749–753. doi:10.1038/nature02305 nature02305 [pii]. [PubMed: 14973489]

26. Wang N, Kim HG, Cotta CV, Wan M, Tang Y, Klug CA, Cao X. TGFbeta/BMP inhibits the bone marrow transformation capability of Hoxa9 by repressing its DNA-binding ability. *EMBO J.* 2006; 25(7):1469–1480. doi:7601037 [pii] 10.1038/sj.emboj.7601037. [PubMed: 16525506]
27. Mitsui K, Tokuzawa Y, Itoh H, Segawa K, Murakami M, Takahashi K, Maruyama M, Maeda M, Yamanaka S. The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell.* 2003; 113(5):631–642. [PubMed: 12787504]
28. Berger MF, Badis G, Gehrke AR, Talukder S, Philippakis AA, Peña-Castillo L, Alleyne TM, Mnaimneh S, Botvinnik OB, Chan ET, Khalid F, Zhang W, Newburger D, Jaeger SA, Morris QD, Bulyk ML, Hughes TR. Variation in homeodomain DNA binding revealed by high-resolution analysis of sequence preferences. *Cell.* 2008; 133(7):1266–1276. doi:10.1016/j.cell.2008.05.024. [PubMed: 18585359]
29. Kato M, Dobyns WB. X-linked lissencephaly with abnormal genitalia as a tangential migration disorder causing intractable epilepsy: proposal for a new term, “interneuronopathy”. *Journal of child neurology.* 2005; 20(4):392–397. [PubMed: 15921244]
30. Kato M, Das S, Petras K, Kitamura K, Morohashi K-i, Abuelo DN, Barr M, Bonneau D, Brady AF, Carpenter NJ, Cipero KL, Frisone F, Fukuda T, Guerrini R, Iida E, Itoh M, Lewanda AF, Nanba Y, Oka A, Proud VK, Saugier-veber P, Schelley SL, Selicorni A, Shaner R, Silengo M, Stewart F, Sugiyama N, Toyama J, Toutain A, Vargas AL, Yanazawa M, Zackai EH, Dobyns WB. Mutations of ARX are associated with striking pleiotropy and consistent genotype-phenotype correlation. *Human Mutation.* 2004; 23(2):147–159. doi:10.1002/humu.10310. [PubMed: 14722918]
31. Stromme P, Mangelsdorf ME, Shaw MA, Lower KM, Lewis SM, Bruyere H, Lutchera V, Gedeon AK, Wallace RH, Scheffer IE, Turner G, Partington M, Frints SG, Fryns JP, Sutherland GR, Mulley JC, Gecz J. Mutations in the human ortholog of *Aristaless* cause X-linked mental retardation and epilepsy. *Nat Genet.* 2002; 30(4):441–445. doi:10.1038/ng862 ng862 [pii]. [PubMed: 11889467]
32. Arnold K, Bordoli L, Kopp J, Schwede T. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics (Oxford, England).* 2006; 22(2):195–201. doi:10.1093/bioinformatics/bti770.

Fig. 1. Identification of Arx consensus binding sequence by SELEX (*Systematic Evolution of Ligands by EXponential Enrichment*). a) Schematic diagram of Arx with each functional domain represented by a unique color. The black lines show the location of the deletion mutations. b) Using the GST-286-430 protein, Arx binding sites were determined by SELEX. The EMSA data show enhancement of DNA binding activity of Arx after each panning. Lane 1 includes the probe from first pan but no Arx protein. Lanes 2-7 have Arx protein bound to the probes after each panning. The low activity in lane 5 is due to a low probe labeling efficiency. We interpret the absence of enhancement in pan 6 to be the result of enrichment saturation. Ratios were calculated by band intensity of Arx bound probe/total band intensity of Arx bound probe and unbound probe. The arrowhead and arrow indicate the Arx-DNA complex and free probe, respectively. Sequence logo for Arx-DNA binding sites was generated by MEME (Multiple EM for motif Elucidation).

Fig. 2. DNA binding preference for full-length Arx and deletion mutants. a-e). Electrophoretic mobility shift assays were performed with the indicated protein on the top of each panel. Sumo-Arx =full-length Arx with Sumo-tag. Each lane has the probes as follows; N11, N11, N30, N9, N26, N2 and N25. The sequence of each probe is given in Table 1. Arrowhead indicates the Arx/DNA complex.

Fig. 3. DNA binding and transcriptional repression for Arx homeodomain mutants. a) Cartoon of the predicted Arx homeodomain structure with arbitrary DNA, generated by SWISS-MODEL [32]. b) MEME-based sequence logo from 84 homeodomains (Reprinted from Cell, 133 /7, Marcus B. Noyes, Ryan G. Christensen, Atsuya Wakabayashi, Gary D. Stormo, Michael H. Brodsky and Scot A. Wolfe, Analysis of Homeodomain Specificities Allows the Family-wide Prediction of Preferred Recognition Sites, 1277-1289., Copyright (2008), with permission from Elsevier) is compared to the mutation sites in human ARX. c) DNA binding activities for human ARX mutants using the Lmo1 probes [19]. Arrowhead indicates the Arx-DNA complex. In this experiments, the GST-286-430 protein, which has 100% homology between the human and mouse, was used. d) Transcriptional repression activity was tested with a Gal4-based luciferase assay. The R332H, R332P and T333N mutants were statistically significant when compared to wild type ($p < 0.05$). The P353R mutant approached statistical significance ($p = 0.053$). The L343q was not significantly different from wild type ($p > 0.05$). Y-axis indicates a relative luciferase activity. Table 1. DNA binding preference of the tested mutants.

DNA binding preference of the tested mutants.

Table 1

Nucleotide position										Proteins/Binding affinity *				
Probe name	1	2	3	4	5	6	7	8	9	Sumo-Arx	GST-214-430	His-214-430	GST-286-430	His-286-430
N11										++++(100%)	-	-	+/- (7.5%)	-
N30										++(38%)	+++(+86%)	++++(100%)	++++(100%)	-
N9										+(17%)	++++(100%)	++++(97%)	++++(95%)	-
N26										+++ (75%)	+(18%)	++(16.4%)	++++(94%)	-
N2										-	-	-	+(21.7%)	-
N25										-	++(33.6%)	+++ (46%)	+++ (81.6%)	-

* Affinity ranking according to relative percentage of binding probe from the highest in each respective protein sample based on the major band intensity by Phosphorimager ImageQuant. (-) indicates background level.