Neuroanatomy and Behaviour in Mice with a Haploinsufficiency of AT-Rich Interactive Domain 1B (ARID1B) Throughout Development

Ellegood, J.¹, Petkova, S.P.²,³*, Kinman, A.¹, Qiu, L.R.⁴, Wade, A.³, Fernandes, D.¹,⁵, Lindenmaier, Z.¹,⁵, Crieghton, A.⁶, Nutter, L.⁶, Nord, A.S.²,³,⁷, Silverman, J.L.²**, and Lerch J.P.¹,⁴**

1. Mouse imaging Centre, Hospital for Sick Children, Toronto, Ontario, Canada
2. University of California, Davis, MIND Institute, School of Medicine, Department of Psychiatry and Behavioral Sciences, Sacramento, CA, USA
3. Neuroscience Graduate Group, University of California - Davis, Davis, CA, USA
4. Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neuroscience, The University of Oxford, Oxford, UK
5. Department of Medical Biophysics, University of Toronto, Toronto, Ontario, Canada
6. The Centre for Phenogenomics, Hospital for Sick Children, Toronto, Ontario, Canada
7. Department of Neurobiology, Physiology and Behavior, University of California - Davis, Davis, CA, USA

*These authors contributed equally.
** These authors contributed equally.

Running Title – Examining the Arid1b Mouse Across Development

Corresponding Author:
Jacob Ellegood
Mouse Imaging Centre (MICe)
Hospital for Sick Children
25 Orde Street
Toronto Ontario Canada
M5T 3H7
Tel: 416-813-7654 x 309546
Fax: 647-837-5832
E-Mail: jacob.ellegood@sickkids.ca
Abstract

Background - One of the causal mechanisms underlying neurodevelopmental disorders (NDDs) is chromatin modification, and the genes that regulate chromatin. AT-Rich Interactive Domain 1B (ARID1B), a chromatin modifier, has been shown to be reduced in autism spectrum disorder (ASD) and to affect rare and inherited genetic variation in a broad set of NDDs.

Methods - A novel preclinical mouse model of Arid1b deficiency was created validated to characterize and define neuroanatomical, behavioural and transcriptional phenotypes. Neuroanatomy was assess ex vivo in adult animals and in vivo longitudinally from birth to adulthood. Behavioural testing was also performed throughout development and tested all aspects of motor, learning, sociability, repetitive behaviours, seizure susceptibility and general milestones.

Results - Brains of adult Arid1b−/+ mice had a smaller cerebellum and a larger hippocampus and corpus callosum. These results stand in contrast to previously reported data highlighting losses in corpus callosum volume. In addition, a striking sex dependence was observed throughout development; males had an early emergence of this neuroanatomical phenotype at postnatal day 7, whereas females had a delayed emergence around postnatal day 40. Behaviourally, during neonatal development, Arid1b−/+ mice exhibited robust impairments in ultrasonic vocalizations (USVs) and metrics of developmental growth. As adults, Arid1b−/+ mice showed low motor skills in open field exploration and normal three chambered approach. Arid1b−/+ mice had learning and memory deficits in novel object recognition but not in visual discrimination and reversal touchscreen tasks. Social interactions in the male-female social dyad with USVs revealed social deficits on some but not all parameters. No repetitive behaviours were observed.

Limitations – The behaviour and the neuroimaging analysis were done on separate cohorts of mice, which does not allow a direct correlation between the imaging and behavioural findings.
Conclusions – This study represents a full investigation of Arid1b⁻/⁻ haploinsufficiency throughout development and highlights the importance of examining both sexes throughout development in NDDs.
Key Words

Magnetic Resonance Imaging
Coffin-Siris Syndrome
Autism
Arid1b
Mouse
Behaviour
Background

Neurodevelopmental disorders (NDDs), including autism spectrum disorder (ASD) and intellectual disability (ID), are prevalent and pervasive lifelong disorders that are behaviourally defined with numerous etiologies and currently no biological markers. The NDD behavioural phenotype is extremely heterogeneous, as are the genetics: the Simons Foundation Autism Research Initiative (SFARI) currently lists over 900 autism-relevant genes (gene.sfari.org) [1], a number that increases if all NDDs are considered. To examine the effect that individual genes have on both the brain and behaviour, several groups have begun producing genetically engineered mouse models of high confidence NDD and ASD relevant genes. Chromatin modification is one of the causal mechanisms underlying NDDs, and ARID1B is a high confidence chromatin-related gene that is reduced in ASD and affected rare and inherited genetic variation in many NDD cases [2, 3]. Genes that regulate chromatin can modify events regulating the formation of neural connections and are also critical for correct epigenetic marking with histone post translational modifications [4].

ARID1B, and genes in the chromatin modification complex, SWI/SNF, are thought to cause Coffin-Siris syndrome (CSS) [5], an exceedingly rare neurodevelopmental disorder with fewer than 200 individuals diagnosed [6]. Clinical presentation includes developmental delay, identified by difficulties with feeding, low birth weight, and general failure to thrive, along with ID, impaired motor skills and delayed or absent development of speech and language. Neuroanatomical assessments in the first CSS cases reported smaller head circumferences [5], Dandy Walker malformations (hypoplasia of the cerebellum) [7, 8], and an abnormally thin corpus callosum [9]. In one study, the patient additionally had striking deficits in the brainstem and cerebellum which included ectopic neurons throughout the medulla, inferior olive, and cerebellar white matter [9]. In a recent report of 143 patients with ARID1B mutations, an agenesis or hypoplasia of the corpus
callosum was prevalent in 43% of patients [10].

Neuroanatomical assessments of NDD mouse models have become more frequent [11-17]. However, with over 900 genes linked to NDD, characterization of the more prominent and closely linked genetic models, like Arid1b, are imperative. Characterizing the functional impact of a germline Arid1b mutation on brain development, growth, and behaviour in vivo could reveal specific and generalizable mechanisms linking chromatin biology to pathology and outcomes. Having a rigorous reproducible model system allows for testing of medicinal or genetic interventions and developing therapeutics. Three groups have recently published papers examining mouse models of ARID1B haploinsufficiency [18-20]. The authors reported several behavioural abnormalities relevant to ASD phenotypes. Neuroanatomical differences were found in the dentate gyrus and corpus callosum, consistent with a subset of human patients with CSS; however, one study only focused on areas thought to be relevant to CSS [18], while the other two studies did not specifically look for neuroanatomical alterations [19, 20] and focused on smaller microscopic differences in cortical structures. Additionally, these studies examined limited windows of time in adulthood. Thus, a full behavioural and neuroanatomical assessment across development has not yet been performed.

To address this gap, we generated a novel preclinical mouse model of Arid1b using CRISPR/Cas9 genome engineering and assayed neuroanatomical, behavioural and transcriptional phenotypes associated with Arid1b haploinsufficiency. Behavioural assays were conducted during both neonatal development and adulthood, and in vivo manganese enhanced MRI (MEMRI) was used to assess the developmental time course of neuroanatomy throughout the entire brain from birth into adulthood.
Methods

Generation of the Mice

Precision genetic engineering via the Cas9/CRISPR system was used to generate the Arid1b mutant mouse line. This allele from project TCPR0317 was generated at The Centre for Phenogenomics (TCP) by injecting Cas9 nickase (D10A) and single guide RNAs with spacer sequences of CTGCTTAGCAAGTTACCACT and GCCTGATACAGCACTTACAT targeting the 5' side and ACACTAAAGGGGTTGCTTTC and CTTGTAATCCCCCTGTAGTA targeting the 3' side of exon OTTMUSE00000314956 (exon 5) resulting in deletion of Chr17 from 5242523 to 5243410 with insertion of TT. Consistent with earlier studies, this mutation is associated with early embryonic lethality in homozygous carriers. The mouseline C57Bl/6N-Arid1b<sup>em1(IMPC)Tcp</sup> was made as part of the NorCOMM2 project funded by Genome Canada and the Ontario Genomics Institute at The Centre for Phenogenomics. It was obtained from the Canadian Mutant Mouse Repository (CMMR).

RNA-Sequencing

Transcriptomic analysis was performed as before [11]. Cerebellar tissue was micro-dissected from adult heterozygous Arid1b mice and Wild-Type (WT) littermates (postnatal day (PND)140-150). Stranded mRNA sequencing libraries were prepared using a TruSeq Stranded mRNA kit with all four samples pooled and sequenced in one lane on the Illumina HiSeq platform using a single-end 100-bp strategy. Each library was quantified and pooled before sequencing at the UC Davis DNA Technologies Core. Reads from RNA-seq were aligned to the mouse genome (mm9) using STAR (version 2.5.3a). Aligned reads mapping to genes were counted at the gene level using subreads featureCounts. The mm9 knownGene annotation track and aligned reads were
used to generate quality control information using the full RSeQC tool suite. Unaligned reads were quality checked using FastQC.

**Differential Expression Analysis**

Raw count data for all samples were used for differential expression analysis using edgeR. Genes with at least 20 reads per million in at least one sample were included for analysis, resulting in a final set of 8,942 genes for differential testing. Tagwise dispersion estimates were generated and differential expression analysis was performed using a generalized linear model with genotype as the variable for testing. Normalized expression levels were generated using the edgeR rpkm function. Normalized log2(RPKM) values were used to plot expression data for *Arid1b*. Pseudo-count values were used to plot the summary heatmap. Mouse gene ontology (GO) data was downloaded from Bioconductor (org.Mm.eg.db). We used the goseq package to test for enrichment of GO terms indicating parent:child relationships. For GO analysis, we examined down- and up-regulated genes separately for genes meeting an FDR < 0.2. For the enrichment analysis, the test set of differentially expressed genes was compared against the background set of genes expressed in our study based on minimum read-count cutoffs described above.

**Breeding**

All animals were housed in a temperature-controlled vivarium maintained on a 12-hour light-dark cycle. All procedures were approved by the Institutional Animal Care and Use Committee either at The Centre for Phenogenomics (TCP) in Toronto, Ontario, Canada, or at the University of California Davis School of Medicine in Sacramento, CA, USA, and were conducted in accordance with the Canadian Council on Animal Care Guide to Care and Use of Experimental Animals or
the National Institutes of Health Guide for the Care and Use of Laboratory Animals. Efforts were made to minimize animal suffering and the number of animals used.

Perfusions

In total, 47 fixed mouse brains were examined: 14 Arid1b<sup>+/−</sup> mice (7 females, 7 males) and 33 WT (C57Bl6/NCr, 13 females and 20 males). All mice were perfused at PND60. Mice were anesthetized with a mixture of ketamine/xylazine and intracardially perfused with 30mL of 0.1M PBS containing 0.05 U/mL heparin (Sigma) and 2mM ProHance (Bracco Diagnostics, a Gadolinium contrast agent) followed by 30 mL of 4% paraformaldehyde (PFA) containing 2mM ProHance [21, 22]. Perfusions were performed with a minipump at a rate of approximately 1 mL/min. After perfusion, mice were decapitated and the skin, lower jaw, ears, and the cartilaginous nose tip were removed. The brain and remaining skull structures were incubated in 4% PFA + 2mM ProHance overnight at 4°C then transferred to 0.1M PBS containing 2mM ProHance and 0.02% sodium azide for at least 1 month prior to MRI scanning [23].

Magnetic Resonance Imaging – Ex Vivo

A multi-channel 7.0 Tesla MRI scanner (Agilent Inc., Palo Alto, CA) was used to image the brains within their skulls. Sixteen custom-built solenoid coils were used to image the brains in parallel [22, 24]. In order to detect volumetric changes, we used the following parameters for the MRI scan: T2- weighted, 3-D fast spin-echo sequence, with a cylindrical acquisition of k-space, a TR of 350 ms, and TEs of 12 ms per echo for 6 echoes, field-of-view equaled to 20 x 20 x 25 mm<sup>3</sup> and matrix size equaled to 504 x 504 x 630. Our parameters output an image with 0.040 mm isotropic voxels [25]. The total imaging time was 14 hours.
Mice for Longitudinal MRI

Mice were bred by crossing Arid1b\(^{+/-}\) with a C57Bl6/NCrl mice to produce heterozygous offspring. The number of mouse pups in each litter was reduced to 6 to ensure equal manganese intake by pups through maternal milk. On PND2, mice were tattooed with green ink on their paws for identification and tail clipped for genotyping. Male and female mice were scanned longitudinally in two cohorts over 8 different time-points: PND4, 6, 8, 10, 17, 23, 35, and 60. Cohort 1 was scanned over the first five timepoints, and cohort two was scanned over all 8 timepoints to extend our developmental window to PND60. Cohort 1 consisted of 11 males and 11 females, 5 of which were mutants and 6 of which were WT controls of each sex. Cohort 2 consisted of 18 males and 18 females, 9 of which were mutants and 9 of which were WT controls of each sex. For the first five developmental timepoints, group sizes ranged from 12-19 mice per sex per genotype; for the final 3 timepoints group sizes ranged from 7-9 mice per sex per genotype. An incomplete group at a given timepoint was due to a loss of either the scan (corruption, movement etc.) or the animal itself during the longitudinal time-course.

Magnetic Resonance Imaging – In Vivo

Findings in the *ex vivo* experiments, and the fact that Arid1b haploinsufficiency is relevant to NDDs, highlighted the need to examine the developmental time course of the neuroanatomical differences. Therefore, a longitudinal study was performed to track the development of the neuroanatomy from birth to adulthood, similar to previous work [26, 27]. 24 hours prior to the scan, a 0.4 mmol/kg dose of 30mM manganese chloride (MnCl\(_2\)) was administered as a contrast agent. For mice that were 10 days old or younger, the dam was intraperitoneally injected with MnCl\(_2\), and pups received MnCl\(_2\) through maternal milk. Mice that were older than ten days received intraperitoneal injections of MnCl\(_2\). Up to four mice were scanned simultaneously. Custom built
3D printed holders, which allowed for anesthetic delivery and scavenging, as well as heating were used. During the scan, mice were anesthetized with 1-2% isoflurane, and respiratory rate was monitored using a self-gated signal from a modified 3D gradient echo sequence [28].

A multi-channel 7.0 Tesla MRI scanner with a 30 cm diameter bore (Bruker, Billerica, Massachusetts), equipped with 4 individual cryogenically cooled coils was used to acquire images of the mouse brains. Parameters of the scan are as follows: T1 weighted FLASH 3D gradient echo sequence, TR = 26 ms, TE = 8.250 ms, flip angle = 23°, field of view = 25 x 22 x 22 mm, with a matrix size of 334 x 294 x 294 yielding an isotropic imaging resolution of 75 µm. Imaging time was 58 minutes. After imaging, mice were transferred to a heated cage for 5-10 minutes to recover from the anesthesia, and then returned to their home cage.

**Imaging Registration and Analysis**

Image registration is necessary for quantifying the anatomical differences across images. For both the *in vivo* and *ex vivo* images the registration consisted of both linear (rigid then affine) transformations and non-linear transformations. These registrations were performed with a combination of mni_autoreg tools [29] and ANTS (advanced normalization tools) [30, 31]. A two-level registration pipeline was used on the *in vivo* images and has been detailed in previous studies [26]. Briefly, images from the same age were registered together in the first level creating a consensus average for the brain at each age. For example, all of the PND4 images, both mutant and WT, were registered together to create a PND4 average brain, all of the PND6 images are registered together to create a PND6 average brain, etc. In the second level, each age-consensus average was registered to the average of the next-oldest age (i.e. PND4 average registered to the PND6 average, PND6 to the PND8, etc.) Transformations from each level were then
concatenated so that every single image in the longitudinal dataset could be mapped to the final PND60 average. For the ex vivo image registration, a simpler one-level registration pipeline is used. After registration, all scans were resampled with the appropriate transform and averaged to create a population atlas representing the average anatomy of the study sample. The result of these registrations were deformation fields that transform images to a consensus average. Therefore, these deformations fields quantify anatomical differences between images.

As detailed in previous studies [32, 33], the Jacobian determinants of the deformation fields were computed and analyzed to measure the volume differences between subjects at every voxel. A pre-existing classified MRI atlas was warped onto the population atlas (containing 282 different segmented structures encompassing cortical lobes, large white matter structures such as the corpus callosum, ventricles, cerebellum, brain stem, and olfactory bulbs [26, 34-37]) to compute the volume of brain structures in all the input images. A linear model with a genotype predictor was used to assess the significance of genotype. The model was either fit to the volume of every structure independently (structure-wise statistics) of fit to every voxel independently (voxel-wise statistics), and multiple comparisons in this study were controlled for using the False Discovery Rate [38].

For the longitudinal analysis, a linear mixed effects model was fit to the volume of every structure with fixed effects of sex, genotype, age, sex-genotype interaction, sex-age interaction, genotype-age interaction, sex-genotype-age interaction, and whole-brain volume; and a random intercept for each mouse. To assess the significance of an effect, a reduced model (model without the particular predictor) was fit to the data and the log likelihood ratio test was used to assess whether the particular effect was significantly important in predicting the data. For example, to evaluate the genotype effect, the reduced model contained fixed effects of sex, age, and sex-age interaction, and significant log likelihood ratio test for a structure results indicate genotype had a
notable effect on the volume of that structure. To estimate direction of effect, Satterthwaite approximation [39] was used to estimate degrees-of-freedom.

**Behavioural Testing**

Behavioural experiments were performed on separate cohorts of mice at UC Davis. Cohort 1 was tested through either ultrasonic vocalizations or postnatal developmental physical milestones and neurological reflexes PND2-14, and then they were again tested through adult behaviours consisting of elevated plus maze, light↔dark exploration, open field arena exploration, novel object recognition, self-grooming, three chambered social approach, male female dyad social interactions, fear conditioning, hot plate and grip strength starting at 10 weeks of age. Cohort 2 was tested through either ultrasonic vocalizations or developmental milestones PND2-14 and were tested through touchscreen starting at 8 weeks of age. Lethal dose of PTZ served as the endpoint for both cohorts at 18 weeks of age. To minimize the carry-over effects from repeated testing, assays were performed in the order of least to the most stressful tasks.

**Developmental Milestones**

Developmental milestones were measured on PND2-14, similar to those previously described [40, 41] but elaborated based on the unique features of this line. All measures were conducted by an experimenter blind to genotype and sex. Pups were separated from the dam for a maximum time of four minutes. Pups were observed for abnormalities in eye, pinnae, whisker, and fur development each testing day. Pups were also picked up to observe limb movement for early signs of hypotonia. Physical feature measurements such as body weight and length (subtraction of total length from tip of nose to base of tail and tail length), and head width was measured using a scale (grams) and a digital caliper (mm). Developmental motor milestones to detect neonatal
hypotonia and ataxia included negative geotaxis, cliff avoidance, righting reflex and circle transverse. Negative geotaxis was tested by placing each pup, facing downwards, on a surface angled at 45° from parallel, and measuring the time for it to completely turn 180° and face up the incline. Failures to turn and climb were recorded as a maximum score of 30-s. Cliff avoidance was tested by gently placing each pup near the edge of a cardboard box and measuring the time for it to turn its head or back away from the edge. Failures to avoid the cliff was recorded as a maximum score of 30-s. Righting reflex was tested by placing each pup on its back, releasing it, and measuring the time for it to fully flip over onto four paws on each developmental day. Righting reflex, negative geotaxis and cliff aversion all test early limb coordination, vestibular system and proprioception. Circle transverse was tested by placing each pup in the center of a circle with a 5” (12.5 cm) diameter drawn on a laminated sheet of 8.5”×11” white paper and measuring the time for it to exit the circle. Circle transverse is an early metric of motor ability and ambulation; developmental delay in crawling and walking may be observed using this task. Failures to exit the circle were recorded as a maximum score of 30-s. Milestones for tracking development of fore and hind limb strength were forelimb grasp, forelimb bar holding, forelimb hang, and hindlimb hang. Forelimb grasp tests a mouse’s ability to curl its paw towards an object, a thin wooden stick. Pups typically develop this ability by PND5 and day of first presentation is noted for each pup. Starting around PND6, pups develop enough limb strength to hold themselves up by the forelimbs. Pups are placed by the forelimbs on an unfolded paper clip and time spent able to hang on independently, up to 30 seconds, is recorded. Hindlimb strength testing can be performed earlier than forelimb strength. Once pups develop some limb strength and the grasp reflex, a thin unfolded paper clip bar is placed into their paws and time able to independently hold the bar, up to 30 seconds, is recorded. This task requires high compliance and tested starting at PND8. The delay in testing from PND2 to PND8 is the result of this skill’s difficulty to assess. Subjects must grasp, comply for a set amount of time and hold the bar. Hindlimb hang is performed by placing the pup on the thin edge of a surface or container by its legs and allowing the pup to hang down
and support itself on the edge by its hindlimbs. Time spent independently hanging by its hindlimbs, up to 30 seconds, is recorded. Strength tests allow for the determination if motor deficits are due to peripheral limb weakness or centralized neural coordination dysfunction.

_Ultrasonic Vocalizations_

Pups were removed individually from the nest at random and gently placed into an isolation container (8 cm x 6 cm x 5 cm; open surface) made of glass. The isolation container was surrounded by a sound attenuating box (18 cm x 18 cm x 18 cm) made of 4 cm thick styrofoam walls. Ultrasonic vocalizations were monitored by an UltraSound Gate Condenser Microphone CM 16 (Avisoft Bioacoustics, Berlin, Germany) attached to the roof of the sound attenuating box, 7 cm above the floor. The microphone was connected via an UltraSoundGate 116 USB audio device (Avisoft Bioacoustics) to a personal computer where acoustic data was recorded with a sampling rate of 250,000 Hz by Avisoft RECORDER (version 2.97; Avisoft Bioacoustics). At the end of the 3-min recording session, both body weight and body temperature were collected. The isolation container was cleaned with 70% ethanol before the beginning of the first test session and between each new litter. Ultrasonic vocalization spectrograms were displayed using Avisoft software and calls were identified manually by a highly trained investigator blinded to genotype.

_Elevated Plus-Maze_

Conflict anxiety-like behaviour via the elevated plus maze was measured according to previously described procedures [42, 43] using a standard mouse apparatus (Med Associates, St. Albans, VT). The maze had two open arms (35.5 cm × 6 cm) and two closed arms (35.5 cm × 6 cm) radiating from a central area (6 cm × 6 cm). A 0.5 cm high lip surrounded the edges of the open arms. 20 cm high walls enclosed the closed arms. The apparatus was cleaned with 70% ethanol
before the beginning of the first test session and after each subject mouse. Room illumination was ~ 300 lux.

*Light ↔ Dark Transitions*

Light ↔ dark exploration was measured according to previously published procedures [42, 43]. Subject mice were placed in the brightly illuminated, large chamber. The smaller dark chamber (~ 5 lux) was entered by traversing a partition between the two chambers. Subject mice freely explored for 10 minutes. Time in the dark side chamber and total number of transitions between the light and dark side chambers were automatically recorded using Labview 8.5.1 software (National Instruments, Austin, TX). Room illumination was ~ 400 lux.

*Open Field*

General exploratory locomotion in a novel open field environment was assayed as previously described [42, 43]. Open field activity was considered an essential control for effects on physical activity, e.g. sedation or hyperactivity, which could confound the interpretation of results from the reciprocal interactions, self-grooming, novel object recognition, fear conditioning and social approach tasks [44, 45]. The testing room was illuminated at ~30 lux.

*Gait*

Treadmill gait analysis was performed using the DigiGait™ system (Mouse Specifics Inc., USA) [46]. Mouse paws were painted with non-toxic red food coloring to augment dark green paw tattoos that generated conflict in DigiGait™ analysis one minute prior to introduction to the walking chamber to reliably capture the entire paw surface area. Before data collection, each subject was
acclimated to the Perspex walking chamber for one minute and the treadmill was slowly accelerated to the final speed of 20 cm/s to allow mice to adjust to walking on the belt. Digital images of paw placement were recorded through a clear treadmill from the ventral plane of the animal. Mice were tested in a single session at a 20 cm/s treadmill speed maintaining a normal pace walk for WT mice. Non-performers were defined as mice who were unable to sustain walking at 20 cm/s without colliding with the posterior bumper for at least three seconds. There is no practice effect or repeated exposure, and therefore, mice were allowed retrial and retest if they were unable to adjust to walking on the belt easily. The treadmill belt and the encasing Perspex chamber were cleaned with 70% (v/v) ethanol in between tests. For each mouse, videos of ~5 second duration of all sessions were analyzed using the DigiGait™ Imaging and Analysis software v12.2 (Mouse Specifics Inc., USA). Contrast filters were determined on a mouse-by-mouse case to facilitate consistent recognition of all four paws. All analysis was conducted in a single session by experimenter blind to genotype. Stride length (distance a paw makes during a single stride) and frequency (number of strides per second to maintain pace) were automatically calculated. Data was averaged between left and right paws for fore and hind paws.

**Novel Object Recognition**

The novel object recognition (NOR) test was conducted in opaque matte white (P95 White, Tap Plastics, Sacramento, CA) open field arenas (41 cm × 41 cm × 30 cm), using methods similar to those previously described [11, 41, 43]. The experiment consisted of four sessions: a 30-min exposure to the open field arena the day before the test, a 10-min re-habituation on test day, a 10-min familiarization session and a 5-min recognition test. On day 1, each subject was habituated to a clean, empty open field arena for 30-min. 24-hr later, each subject was returned to the open field arena for another 10-min for the habituation phase. The mouse was then removed from the open field and placed in a clean temporary holding cage for approximately 2-
Two identical objects were placed in the arena. Each subject was returned to the open field in which it had been habituated and allowed to freely explore for 10-min. After the familiarization session, subjects were returned to their holding cages, which were transferred from the testing room to a nearby holding area. The open field was cleaned with 70% ethanol and let dry. One clean familiar object and one clean novel object were placed in the arena, where the two identical objects had been located during in the familiarization phase. 60-min after the end of the familiarization session, each subject was returned to the arena for a 5-min recognition test, during which time it could to freely explore the familiar object and the novel object. The familiarization session and the recognition test were recorded and scored with Ethovision XT videotracking software (Version 9.0, Noldus Information Technologies, Leesburg, VA). Object investigation was defined as time spent sniffing the object when the nose was oriented toward the object and the nose–object distance was 2-cm or less. Recognition memory was defined as spending significantly more time sniffing the novel object compared to the familiar object via a Student's paired t-test. Total time spent sniffing both objects was used as a measure of general exploration. Time spent sniffing two identical objects during the familiarization phase confirmed the lack of an innate side bias. Objects used were plastic toys: a small soft plastic orange safety cone and a hard, plastic magnetic cone with ribbed sides, as previously described [47].

**Repetitive Self-Grooming**

Spontaneous repetitive self-grooming behaviour was scored as previously described [48, 49]. Each mouse was placed individually into a standard mouse cage, (46 cm length × 23.5 cm wide × 20 cm high). Cages were empty to eliminate digging in the bedding, which is a potentially competing behaviour. The room was illuminated at ~40 lux. A front-mounted CCTV camera (Security Cameras Direct) was placed at ~1m from the cages to record the sessions. Sessions were video-taped for 20 min. The first 10-min period was habituation and was unscored. Each
subject was scored for cumulative time spent grooming all the body regions during the second 10 min of the test session. Scoring was conducted by an experimenter blind to genotype.

Social Approach

Social approach was tested in an automated three-chambered apparatus using methods similar to those previously described [50-52]. Automated Ethovision XT videotracking software (Version 9.0, Noldus Information Technologies, Leesburg, VA) and modified non-reflective materials for the chambers were employed to maximize throughput. The updated apparatus (40cm x 60cm x 23cm) was a rectangular, three-chambered box made from matte white finished acrylic (P95 White, Tap Plastics, Sacramento, CA). Opaque retractable doors (12cm x 33cm) were designed to create optimal entryways between chambers (5cm x 10cm), while providing maximal manual division of compartments. Three zones, defined using the EthoVision XT software, detected time in each chamber for each phase of the assay. Zones were defined as the annulus extending 2 cm from each novel object or novel mouse enclosure (inverted wire cup, Galaxy Cup, Kitchen Plus, http://www.kitchen-plus.com). Direction of the head, facing toward the cup enclosure, defined sniff time. A top mounted infrared sensitive camera (Ikegami ICD-49, B&H Photo, New York, NY) was positioned directly above every two 3-chambered units. Infrared lighting (Nightvisionexperts.com) provided uniform, low level illumination. The subject mouse was first contained in the center chamber for 10 minutes, then explored all three empty chambers during a 10-minute habituation session, then explored the three chambers containing a novel object in one side chamber and a novel mouse in the other side chamber. Lack of innate side preference was confirmed during the initial 10 min of habituation to the entire arena. Novel stimulus mice were 129Sv/ImJ, a relatively inactive strain, aged 10–14 weeks old, and matched to the subject mice by sex. Number of entries into the side chambers served as a within-task control for levels of general exploratory locomotion.
Male Female Social Dyad Interaction

The male-female social dyad interaction test was conducted as previously described [43, 49, 51]. Briefly, each freely moving male subject was paired for 5-min with a freely moving unfamiliar estrous WT female. A closed-circuit television camera (Panasonic, Secaucus, NJ, USA) was positioned at an angle from the Noldus PhenoTyper arena (Noldus, Leesburg, VA) for optimal video quality. An ultrasonic microphone (Avisoft UltraSoundGate condenser microphone capsule CM15; Avisoft Bioacoustics, Berlin, Germany) was mounted 20 cm above the cage. Sampling frequency for the microphone was 250 kHz, and the resolution was 16 bits. The entire apparatus was contained in a sound-attenuating environmental chamber (Lafayette Instruments, Lafayette, IN) under red light illumination (~10 lux). Duration of nose-to-nose sniffing, nose-to-anogenital sniffing, and following were scored using Noldus Observer 8.0XT event recording software (Noldus, Leesburg, VA) as previously described [53]. Ultrasonic vocalization spectrograms were displayed using Avisoft software and calls were identified manually by a highly trained investigator blinded to genotype.

Fear conditioning

Delay contextual and cued fear conditioning was conducted using an automated fear-conditioning chamber (Med Associates, St Albans, VT, USA) as previously described [43]. The conditioning chamber (32 × 25 × 23 cm³, Med Associates) was interfaced to a PC installed with VideoFreeze software (version 1.12.0.0, Med Associates) and enclosed in a sound-attenuating cubicle. Training consisted of a 2-min acclimation period followed by three tone-shock (CS–US) pairings (80 dB tone, duration 30 s; 0.5 mA footshock, duration 1 s; intershock interval 90 s) and a 2.5-min period, during which no stimuli were presented. The environment was well lit (~100 lux), with a
stainless-steel grid floor and swabbed with vanilla odor cue (prepared from vanilla extract; McCormick; 1:100 dilution). A 5-min test of contextual fear conditioning was performed 24 h after training, in the absence of the tone and footshock, but in the presence of 100 lux overhead lighting, vanilla odor and chamber cues identical to those used on the training day. Cued fear conditioning, conducted 48 h after training, was assessed in a novel environment with distinct visual, tactile and olfactory cues. Overhead lighting was turned off. The cued test consisted of a 3-min acclimation period followed by a 3-min presentation of the tone CS and a 90-s exploration period. Cumulative time spent freezing in each condition was quantified by VideoFreeze software (Med Associates).

**Touchscreen pairwise discrimination**

Pairwise visual discrimination was tested in the automated Bussey-Saksida touchscreen apparatus for mice (Campden Instruments Ltd/Lafayette Instruments, Lafayette, IL, USA), using a procedure modified from original methods described previously [52, 54-56]. The reinforcer was 20 ul of a palatable liquid nutritional supplement (Strawberry Ensure Plus, Abbott, IL, USA) diluted to 50% with water. Each session was conducted under overhead lighting (~60 lux). A standard tone cue was used to signal the delivery of the reinforcer during pre-training and acquisition. Prior to pre-training, subject mice were weighed, and placed on a restricted diet of 2–4 g of rodent chow per mouse per day, to induce a 15% weight loss. Body weight was carefully monitored throughout the experiment, to ensure that a minimum of 85% of free feeding body weight was maintained for each mouse. The pre-training consisted of four stages as previously published [52]. Stage 1 consisted of two days of habituation (20 min on day 1, and 40 min on day 2) to the chamber and the liquid diet with no images on the screen under overhead lighting (~60 lux). Stage 2 was a single 45-min session in which entering and exiting the food magazine initiates the next trial and triggers additional reward under overhead lighting. During Stage 3, subjects were trained in daily
45-min sessions during which an image (a random picture from a selection of 40 images) was presented in one of the two windows and remained on the screen until it was touched. Mice must complete 30 trials/day for two consecutive days in order to advance to the next stage. In Stage 4, subjects were trained in 45-min daily sessions in which touching the blank side of the screen was discouraged with a 5-s time-out during which the overhead lighting turned off. Completion of at least 30 trials, at an average accuracy of 80%, on two consecutive days, is required for advancement. Images used in Stages 3 and 4 were not used in the subsequent discrimination task. Only mice that completed all stages of pre-training were advanced to the pairwise visual discrimination task. Subjects were trained to discriminate between two novel images, a spider and an airplane, presented in a spatially pseudo-randomized manner in the two windows of the touchscreen. Each 45-min session consisted of unlimited number of trials separated by 15-s inter-trial intervals (ITI). Designation of the correct and incorrect images was counterbalanced across mice within each genotype. Correct responses were rewarded. Each incorrect response was followed by a correction trial in which the images were presented in an identical manner to the previous trial, until a correct response was made. Criterion was completing at least 30 trials, at an accuracy of 80% or higher, on two consecutive days. Days to reach criterion and percentage of mice reaching criterion on each day were compared between genotypes. After successful completion of the task, mice underwent task reversal in which the opposite image was now correct. Criterion was completing at least 30 trials, at an accuracy of 75% or higher, on two consecutive days. Days to reach criterion and percentage of mice reaching criterion on each day were compared between genotypes.

_Pentylenetetrazol-induced seizures_

Behavioural assessment of seizure threshold in mice was performed with injections of 80 mg/kg
of pentylenetetrazol (PTZ) as described previously [43, 57]. PTZ-induced convulsions were used to gauge susceptibility to primary generalized seizures and as a gross approximation of excitation–inhibitory balance.

**Behavioural Analysis**

Data were analyzed in Graphpad PRISM 7.0. Sexes were considered separately with genotype as the fixed factor. All significance levels were set at $p < 0.05$ and all tests were two-tailed.

**Experimental Design and Statistical Analysis**

All experiments reported in this study were designed to examine genotype effects between the $\text{Arid1b}^{+/\text{--}}$ mice and the $\text{Arid1b}^{+/\text{++}}$ wild-type control (C57Bl6/N) mice. In addition, for the longitudinal developmental comparisons, we looked at genotype, sex, age, and the interactions between those factors using a linear mixed effects model described above.

For Figure 1, Transcriptomic analysis was performed as before [11]. Cerebellar tissue was micro-dissected from $\text{Arid1b}^{+/\text{--}}$ mice and $\text{Arid1b}^{+/\text{++}}$ littermates (postnatal day (PND)140-150). For differential expression, genes with at least 20 reads per million in at least one sample were included for analysis, resulting in a final set of 8,942 genes for testing. Normalized log2(RPKM) values were used to plot expression data for $\text{Arid1b}$. Pseudo-count values were used to plot the summary heatmap.

For Figure 2, ex vivo brains samples from 47 mice were examined: 14 $\text{Arid1b}^{+/\text{--}}$ mice (7 females, 7 males) and 33 WT (C57Bl6/NCrl, 13 females and 20 males). All mice were perfused at PND60. For Figures 3 and 4, Male and female mice were scanned longitudinally in two cohorts
over 8 different time-points: PND4, 6, 8, 10, 17, 23, 35, and 60. Cohort 1 was scanned over the first five timepoints, and cohort two was scanned over all 8 timepoints to extend our developmental window to PND60. Cohort 1 consisted of 11 males and 11 females, 5 of which were mutants and 6 of which were WT controls of each sex. Cohort 2 consisted of 18 males and 18 females, 9 of which were mutants and 9 of which were WT controls of each sex. For the first five developmental timepoints, group sizes ranged from 12-19 mice per sex per genotype; for the final 3 timepoints group sizes ranged from 7-9 mice per sex per genotype. For calculating the volumes in Figures 2-4, as detailed in previous studies [32, 33], the Jacobian determinants of the deformation fields were computed and analyzed to measure the volume differences between subjects at every voxel. A linear model with a genotype predictor was used to assess the significance of genotype in the ex vivo cohort. The model was either fit to the volume of every structure independently (structure-wise statistics) or fit to every voxel independently (voxel-wise statistics), and multiple comparisons in this study were controlled for using the False Discovery Rate [38]. For the longitudinal cohorts, a linear mixed effects model was fit to the volume of every structure with fixed effects of sex, genotype, age, sex-genotype interaction, sex-age interaction, genotype-age interaction, sex-genotype-age interaction, and whole-brain volume; and a random intercept for each mouse. To assess the significance of an effect, a reduced model (model without the particular predictor) was fit to the data and the log likelihood ratio test was used to assess whether the particular effect was significantly important in predicting the data.

For the behavioural testing, Cohort 1 was sampled from 12 litters and was tested in elevated plus maze, light↔dark exploration, novel object recognition, social approach, male female social dyad interactions, self groom, contextual and cued fear conditioning, hot plate, grip strength and lethal dose PTZ. Cohort 2 was sampled from 12 litters and was tested in pairwise discrimination and reversal touchscreen task and lethal dose of PTZ. All male and female mice were used.
from every litter. In cohort 1, the order of testing was as follows with at least 48-hrs separating tasks: (1) PND3, 5, 7, 9 ultrasonic vocalizations or PND2, 4, 6, 8, 10, 12, 14 developmental milestones, (2) elevated plus-maze at 10 weeks of age, (3) light↔dark exploration task at 10 weeks of age, (4) open field locomotion at 11 weeks of age, (5) balance beam walking at 11 weeks of age, (6) spontaneous alternation at 12 weeks of age, (7) rotarod at 12 weeks of age, (8) novel object recognition at 13 weeks of age, (9) self-groom at 15 weeks of age, (10) social approach at 15 weeks of age, (11) male female social dyad interaction at 16 weeks of age, (12) grip strength at 17 weeks of age, (13) hot plate at 17 weeks of age, (14) fear conditioning at 18 weeks of age, and (15) lethal dose chemoconvulsant at 19 weeks of age. In cohort 2, ultrasonic vocalizations or developmental milestones were conducted between PND2 and 14 in separate litters, weight restriction for touchscreen assay began at 7 weeks of age, pretraining began at 9 weeks of age and touchscreen testing continued daily until completion, DigiGait was conducted at 17 weeks of age and lethal dose chemoconvulsant at 18 weeks of age. All behavioural assays in all cohorts were performed between 9 am-4 pm PST (ZT2-ZT9). For Figure 5 (A), (B), (C), Arid1b+/+ N=40, Arid1b+- N=19, for (D-I), Arid1b+/+ N=18, Arid1b+- N=29. For Figure 6 (A-B), Arid1b+/+ N=28, Arid1b+- N=25, for (C-D) Arid1b+/+ N=21, Arid1b+/+ N=19, for (E-H) Arid1b+/+ N=28, Arid1b+- N=26. † p<.09 vs Arid1b+/+ by repeated measures two-way ANOVA. For Figure 7 (A-C), Arid1b+/+ N=28, Arid1b+- N=25, for (D-F) Arid1b+/+ N=15, Arid1b+- N=17, in (G), Arid1b+/+ N=28, Arid1b+- N=25, for (G-H), Arid1b+/+ N=28, Arid1b+- N=27, for (J-L), Arid1b+/+ N=45, Arid1b+- N=40. For Figure 8 (A-B), Arid1b+/+ N=24, Arid1b+- N=15, for (C-D), Arid1b+/+ N=20, Arid1b+- N=14, for (E-F), Arid1b+/+ N=27, Arid1b+- N=23, for (G), Arid1b+/+ N=27, Arid1b+- N=21. For Figures 5-8 significance was measured by repeated measures two-way ANOVA or student’s unpaired t-Test.
Results

Model validation

We used RNA sequencing (RNA-seq) on four samples of micro-dissected cerebellum tissue from adult (~PND145) mice with heterozygous null mutation of Arid1b or their WT littermates that underwent behavioural experiments. We tested for differential expression across 8,942 genes that were robustly expressed in our samples. We found a total of 1001 differentially expressed (DE) genes in our model, with 501 genes meeting an FDR of 0.25 (p < 0.01), 235 meeting an FDR of 0.15 (p < 0.004), and 72 meeting an FDR of 0.05 (p < 0.0004). Of the total significant DE genes, 607 were downregulated and 394 were upregulated. Almost all DE genes (99.9%) had a subtle change in expression with an absolute log₂ fold change (FC) of less than 1 (Figure 1A). The most significantly DE gene was Somatostatin (log₂ FC = 0.74, p = 2.92E-09, FDR = 2.61E-05). While not the top gene, we validated that Arid1b was differentially expressed in our model (log₂ FC = -0.37, p = 0.0037, FDR = 0.14; Figure 1B). Gene set enrichment analysis of Gene Ontology (GO) terms showed enrichment for annotations associated with neuronal morphology and synaptic function in upregulated genes and annotations associated with gene regulation in downregulated genes (Figure 1C; Figure 1-1). We found differential expression of autism-relevant genes in our DE gene list that are also associated with cell adhesion (Dscam, Reln) or gene regulation (Kat2b, Kmt2c, Med13/Med13l, Tbl1xr1, Ubn2; Figure 1-1).

Hydrocephaly

There was an increased incidence of hydrocephaly in the Arid1b⁺⁻ mice compared to other mice in both the Sacramento and Toronto colonies. Prior to the beginning of this study, 667 Arid1b⁺⁻ mice were born in the Toronto lab, of which 3.1% (21) had hydrocephaly, compared to 0.2-0.3%
in other mice on the same C57BL6/N background; these mice were humanely euthanized or removed from the analysis. Of the 89 Arid1b\textsuperscript{+/−} mice born in the Sacramento lab, 55 were used and 12.7 % (7) were excluded or died due to hydrocephalus (equal rates between sexes). Of the 82 Arid1b\textsuperscript{+/+} used in this study, only one was excluded due to hydrocephalus (1.2%).

Hydrocephalus has been previously reported in Arid1b\textsuperscript{+/−} mice models at rates of 5.5% to 6.6% [18, 20]. The rates in this study were based on mice that had severe hydrocephaly and thus were excluded from the study; however, as described in more detail below, it should be noted that the ventricular system as a whole was enlarged in Arid1b\textsuperscript{+/−} mice throughout development (q=0.005), particularly the lateral ventricles (q=2.0x10\textsuperscript{-12}).

**Structural Findings – Adult**

Total brain volume of the Arid1b\textsuperscript{+/−} mutants was 2.21% smaller than their corresponding WT littermates. However, this did not reach significance (p=0.02, False Discovery Rate (FDR) q=0.37, 420 ± 14 mm\textsuperscript{3} for Arid1b\textsuperscript{+/−}; 429 ± 12 mm\textsuperscript{3} for WT, see Figure 2-1). In the 282 regions examined, 79 were significantly different at q<0.05. The main olfactory bulb was significantly smaller (-6.10%, q=0.0007), with its olfactory subregions ranging from -5.96% to -10.41% (Figure 2-1). The brainstem was also significantly smaller (-3.13%, q=0.046), due to a smaller hindbrain and pons (-5.53%, q=0.0004 and -4.99%, q=0.0005, respectively). Overall, white matter was significantly smaller (-3.71%, q=0.046). However, this difference stemmed from cerebellar and cranial nerve fibers, and not from the largest white matter tracts in the mouse brain, the fimbria and corpus callosum, commonly found altered in models of NDDs (Figure 2-1). The largest differences were in the cerebellum (-6.77%, q=0.00007), with the cerebellar cortex decreasing by -6.65% (q=0.0001) and the deep cerebellar nuclei (DCN) decreasing by -10.99% (q=4x10\textsuperscript{-8}). Relative volume differences were also measured by covarying for total brain volume. The majority of differences found in relative volume were consistent with absolute volume (Figure 2-1). However,
there were significant increases in the relative volume of the cortex \((q=0.01)\), the striatum \((q=0.006)\), and the corpus callosum in the \textit{Arid1b}\textsuperscript{+/−}\((q=0.007)\).

\textbf{Figure 2} shows a visual representation of the differences found throughout the brain in the \textit{Arid1b}\textsuperscript{+/−} mutants. The \textit{Arid1b}\textsuperscript{+/−} mutation seemed to have a stronger effect on females: 47 of the 282 regions were significantly different, while only 9 were different in the males (\textbf{Figure 2-1}). However, there was no sex by genotype interaction found in the adult mice, indicating that the mutation had similar effects in both sexes.

\textit{Structural Findings – Longitudinal}

The results of the linear mixed effects model showed that \textit{Arid1b}\textsuperscript{+/−} (genotype) had a large effect on several brain structures (\textbf{Figure 3A}, \textbf{Figure 3-1}). Across ages, there was an overall loss in total brain volume \((q=0.02)\). The strongest differences were seen in the cerebellum and hippocampus. The cerebellum as a whole was smaller in the \textit{Arid1b}\textsuperscript{+/−} mice \((q=3\times10^{-8})\); this affected all areas including the vermal regions \((q=8\times10^{-7})\), the hemispheric regions \((q=0.0004)\) with strong effects in the arbor vitae \((q=3\times10^{-11})\) and the deep cerebellar nuclei \((q=2\times10^{-9})\). Notable differences were found in the hippocampal regions, including an increase in size of the CA1 \((q=0.003)\), CA2 \((q=0.001)\), CA3 \((q=2\times10^{-6})\), and the dentate gyrus \((q=6\times10^{-7})\).

We found a significant interaction between age and genotype, indicating that \textit{Arid1b}\textsuperscript{+/−} mutants demonstrate different growth rates; mainly an increased growth rate in the hippocampus (\textbf{Figure 3B}, \(q=0.00008\)) and a decreased growth rate in the cerebellum (\textbf{Figure 3C}, \(q=0.00004\)) compared to WT. Differences in growth rate were seen in all cerebellar areas including the vermal regions \((q=0.0006)\), hemispheric regions \((q=0.03)\), the deep cerebellar nuclei (\textbf{Figure 3D}, \(q=0.0001)\), and the arbor vitae \((q=4\times10^{-9})\) (\textbf{Figure 3-1}). Cerebellar growth rates corresponded to differences in
adulthood, indicating that all areas of the cerebellum grew at a slower rate, thereby resulting in the adult differences. In addition to the cerebellar white matter, there were strong differences in white matter fiber tracts in general \( q = 0.00009 \); however, the growth rate was increased in the \textit{Arid1b}\textsuperscript{+/−} mutants.

We also observed a significant interaction between sex and genotype, and additionally several differences were found in the three-way interaction between age, sex, and genotype (Figure 3A). Figure 4A highlights the top 50 structures (based on percent difference) for both larger and smaller regions in the \textit{Arid1b}\textsuperscript{+/−} mutants versus the WT mice. Regardless of whether the difference was larger or smaller at PND60, affected structures diverged earlier in the males, with differences seen as early as PND7, while in females, the differences did not emerge until after PND40 (Figure 4B and C), thus displaying a sex differences across development.

\textit{Behavioural Findings}

\textit{Arid1b}\textsuperscript{+/−} pups were impaired on several parameters of developmental milestones including growth, reflexes, and pup ultrasonic vocalization (USVs) emissions. Figure 5A illustrates isolation-induced USVs over a developmental time course. A significant effect of age was detected on the rate of calls over time \( F(3,171) = 6.68, p \leq 0.005 \), two-way repeated measures ANOVA). Genotype effects across time were also detected \( F(1,57) = 6.67, p \leq 0.02 \), one-way repeated measures ANOVA), prominently at PND3 and 5 (\textit{Arid1b}\textsuperscript{+/−} vs. \textit{Arid1b}\textsuperscript{+/+}, Bonferroni-Sidak posthoc). Sum of calls collected were fewer in the \textit{Arid1b}\textsuperscript{+/−} (Figure 5B; \( t_{58} = 3.08, p \leq 0.005 \), student’s t-test), compared to \textit{Arid1b}\textsuperscript{+/+} littermates. Core temperature (Figure 5C; \( F(1,57) = 3.92, NS \)) was not different by genotype, a variable known to alter pup USVs [58].
Body weight, length, and head width were lower, smaller and delayed in Arid1b+/− (Figure 5D; $F_{(1, 45)} = 14.8$, $p \leq 0.005$; Figure 5E; $F_{(1, 45)} = 14.07$, $p \leq 0.005$; Figure 5F; $F_{(1, 45)} = 8.10$, $p \leq 0.01$, respectively, two-way repeated measures ANOVA), indicating atypical growth and development. Posthoc analysis revealed Arid1b+/− weighed less on PND6-12, were smaller on PND10-12 and had narrower heads on PND12 vs. Arid1b+/+ (Sidak posthoc for all analysis).

Arid1b+/− did not demonstrate hypotonia and physical core coordination deficits as they did not have longer latencies to right their bodies dorsally, but, longer latencies to reverse from an inclined position in negative geotaxis and time to traverse out of circle were detected in Arid1b+/−, respectively (Figure 5G; $F_{(1, 45)} = 3.71$, $p=0.06$; Figure 5H; $F_{(1, 45)} = 4.54$, $p \leq 0.05$; Figure 5I; $F_{(1, 45)} = 4.34$, $p \leq 0.05$, two-way repeated measures ANOVA), which indicates delayed motoric reflexes, complex limb coordination, and onset of walking skills. Clear divergence in physical metrics were detected in mid and late ranges of time points. As the pups developed, some metrics were difficult to collect with reliable accuracy at every time point, which may contribute to the incomplete penetrance of the phenotypes. However, the strong genotype effect was convincing that Arid1b+/− diverged by numerous major developmental metrics.

Adult Arid1b+/− remained smaller and weaker in measurements of size and strength. Figure 6A illustrates adult Arid1b+/− weighed less compared to Arid1b+/+. Body weight was collected across the study and all animals gained mass across all time points PND65, PND100 and PND135 (Figure 6A; $F_{(2, 102)} = 92.53$, $p \leq 0.005$, two-way repeated measures ANOVA) but less by the Arid1b+/− mice (Figure 6A; $F_{(1, 51)} = 15.22$, $p \leq 0.005$, two-way repeated measures ANOVA), compared to Arid1b+/+ sex-matched littermates. Arid1b+/− showed decreased maximum forelimb force exerted, indicating reduced muscle strength by a forelimb grip strength assay (Figure 6B;
t_{(51)} = 3.08, p \leq 0.005, \text{ student's unpaired t-test). Nuanced motoric behaviours, such as gait, balance, posture, stride and motor coordination can be measured using innovative automated treadmill systems, such as DigiGait™ or Catwalk™. Despite their smaller size, Arid1b^{+/−} did not show differences in stride length (Figure 6C; F_{(1, 42)} = 3.09, NS, two-way repeated measures ANOVA) nor stride frequency (Figure 6D; F_{(1, 43)} = 2.59, NS, two-way repeated measures ANOVA) in either fore and hind paws. While a decrease in stride length and an increase in stride frequency were expected due to their diminished size, the lack of deficit in these two major parameters indicate intact motor coordination of the limbs, and thus does not confound other complex behaviours.

Adult Arid1b^{+/−} were impaired on multiple parameters of gross motor skills using an open field novel arena assay of locomotion. Arid1b^{+/−} were hypoactive over a 30-min session. Difference for horizontal and vertical activity and total distance traversed in the novel open field, were significant, representing normal habituation to the novelty of the open field in both genotypes (Figure 6E-H; horizontal F_{(5, 260)} = 42.02, p \leq 0.005; vertical F_{(5, 260)} = 26.06, p \leq 0.005; total activity F_{(5, 260)} = 52.2, p \leq 0.005, two-way repeated measures ANOVA). Arid1b^{+/−} mice were hypoactive by horizontal and vertical activity during all 5 min binned sessions (Figure 6E; F_{(1, 52)} = 14.71, p \leq 0.005; Figure 6F; F_{(1, 52)} = 2.80, p \leq 0.005, two-way repeated measures ANOVA), by total distance traversed over time (Figure 6G; F_{(1, 52)} = 20.96, p \leq 0.005, two-way repeated measures ANOVA) and summed across the 30-minute session (Figure 6H; t_{(52)} = 4.58, p \leq 0.005), indicating a clear, corroborated motor deficit.

Behaviours relevant to ASD were assessed using two corroborating assays of social behaviour, as described previously [11, 48-50]. Sociability scores from the automated three-chambered social approach task were typical and showed significant sociability (Figure 7A; two-way repeated
measures ANOVA main effect of genotype: $F_{(1, 51)} = 3.016, p>0.05$; post-hoc novel object vs. novel mouse: $Arid1b^{+/+} p<0.001, Arid1b^{+/−} p<0.001$). Both $Arid1b^{+/−}$ and $Arid1b^{+/+}$ mice exhibited significantly more time social sniffing, which was defined as time spent within 2-cm of the wire cup, with the head facing the wire cup containing the stimulus mouse, as compared to the time spent sniffing the novel object, using the same body point detection settings, (Figure 7B; two-way repeated measures ANOVA, main effect of genotype: $F_{(1, 51)} = 0.03, p<0.005; Arid1b^{+/+} p<0.001, Arid1b^{+/−} p<0.001$). $Arid1b^{+/−}$ mice showed fewer entries into the side chambers (Figure 7C; two-way repeated measures ANOVA, main effect of genotype: $F_{(1, 51)} = 8.49, p=0.05$), indicating that reduced $Arid1b$ resulted in less general exploratory activity throughout the 3-chambered apparatus during the social approach assay. However, it did not adversely affect the outcome as both genotypes illustrated intact sociability.

Social deficits were observed on investigative and social parameters in male $Arid1b^{+/−}$ mice when compared to littermate $Arid1b^{+/+}$ during the male-female reciprocal dyad social interaction test. Figure 7D–F illustrates male-female social interaction parameters during a session of reciprocal interactions between subject male $Arid1b^{+/−}$ and $Arid1b^{+/+}$ mice paired with an unfamiliar estrous B6/NJ female. $Arid1b^{+/−}$ exhibited less time in nose-to-anogenital sniffing (Figure 7D; $t_{(30)} = 1.97, p≤0.05$, Student’s unpaired t-test) and following behaviour (Figure 7E; $t_{(30)} = 2.27, p≤0.05$, Student’s unpaired t-test). Other parameters collected during the social interaction such as nose to nose sniffing, exploration, and grooming were not significant. Once again, as adults, ultrasonic call emissions were fewer in the $Arid1b^{+/−}$ (Figure 7F; $t_{(29)} = 2.12, p≤0.005$, Student’s unpaired t-test), which may be attributed to reduced social communication but could also be the result of being smaller, with less pronounced larynx musculature [58-60].
Self-grooming is a useful measure of repetitive behaviour in models of NDDs with high repetitive behaviours. No difference was observed on repetitive self-grooming behaviour (Figure 7G; t(50) = 0.82, NS, Student’s unpaired t-test). No other reported repetitive behaviours, such as circling, jumping or back flipping, were observed during this empty cage observational assay.

Anxiety-like behaviour was assessed in the elevated plus-maze. Arid1b+/- mice spent less time on the arm and had fewer total entries calculated by the open and closed arm entries summed (Figure 7H; t(53) = 2.10, p < 0.05 and Figure 7I; t(53) = 4.61, p < 0.001, Student’s unpaired t-test). This is consistent with less motoric activity in the Arid1b+/- mice and confounds a clear interpretation of anxiety-like behaviour, although anxiety metrics indicate this to be a phenotype.

Seizures were provoked using pentylenetetrazol (PTZ; 80 mg/kg, i.p.) in Arid1b+/- and Arid1b+/- mice littermates, as previously described [57]. Latencies to generalized clonic seizure (loss of righting reflex), tonic extension, and death were collected as a preliminary characterization of seizure, subthreshold epileptiform activity, and imbalances in the excitatory/inhibitory homeostasis. Pentylenetetrazol (PTZ) is a non-competitive GABA_A antagonist leading to hyperexcitability [61, 62]. Arid1b+/- mice exhibited faster loss of righting reflexes (Figure 7J; t(83) = 2.49, p < 0.02, Student’s unpaired t-test), tonic hindlimb extension (Figure 7K; t(67) = 2.11, p < 0.04, Student’s unpaired t-test) and death (Figure 7L; t(77) = 2.56, p < 0.02, Student’s unpaired t-test).

Learning and memory was evaluated using two components, a 24-h contextual and a 48-h cued fear conditioning. High levels of freezing were observed, subsequent to the conditioned stimulus–unconditioned stimulus pairings, on the training day, across genotypes (Figure 8E; training F(1, 48) = 108.9, p < 0.001; genotype training F(1, 48) = 1.74, NS, two way repeated measures ANOVA). No
significant effects of genotype on freezing levels were observed 24-hours later during the contextual phase (Figure 8E; \( t_{(49)} = 0.21 \), NS, Student’s unpaired t-test). High levels of freezing were observed, after the auditory cue stimulus was introduced, on cued conditioning day, across genotypes indicating no impairments in cued conditioning (Figure 8F; effect of cue \( F_{(1,48)} = 205.4 \), \( p \leq 0.001 \); genotype \( F_{(1,48)} = 2.86 \), NS, two way repeated measures ANOVA); the same was observed 24-hours later during the contextual phase (Figure 8E; \( t_{(49)} = 0.21 \), NS, Student’s unpaired t-test).

Cognitive abilities as measured by the novel object recognition task were tested. All automated scores collected via Ethovision were manually confirmed by a trained blinded observer[41]. \textit{Arid1b} \(^{+/−}\) mice did not exhibit typical novel object preference (Figure 8G; \textit{Arid1b} \(^{+/+}\), \( p < 0.01 \) and \textit{Arid1b} \(^{+/−}\) \( p > 0.05 \)). Both genotypes explored the two identical objects similarly during the automated familiarization phase (data not shown).

We utilized innovative computerized touchscreen assays with high translational relevance to evaluate visual discrimination and behavioural flexibility attributable to both HPC and cortical circuitry [52]. Figure 8A-C illustrate \textit{Arid1b} \(^{+/−}\) mice required significantly fewer sessions to a stringent criterion of completing at least 30 trials, at an accuracy of 80% or higher, on two consecutive days, when compared to littermate controls. \textit{Arid1b} \(^{+/−}\) mice needed fewer sessions over the training days to learn to discriminate between two images displayed on the touchscreen (Figure 8A; \( t_{(37)} = 3.07 \), \( p \leq 0.004 \), Student’s unpaired t-test). Analysis of survival curves, i.e. percentage of mice that reached the 80% accuracy criterion on each training day, indicated that the percentage of mice that reached this criterion was significantly higher in \textit{Arid1b} \(^{+/−}\) (Figure 8B, \( X = 8.88 \), \( p \leq 0.003 \), Gehan Breslow Wilcoxon Chi Square test). During reversal, designation of the correct and incorrect stimulus was switched. No genotype effect was observed in the second,
more difficult reversal phase, as shown by sessions to reversal criteria (Figure 8C; \( t_{32} = 1.92 \), NS, Student’s unpaired t-test) and survival curves of mice reaching criterion over time (Figure 8D, \( X = 2.57 \), NS, Gehan Breslow Wilcoxon Chi Square test).
Discussion

While Arid1b<sup>+/−</sup> mice have been previously examined [18-20], we present an in-depth characterization of the behaviour and whole brain analysis through development in both sexes. We show dramatic growth and motor deficits that confirm previous anxiety and social behavioural phenotypes. The neuroanatomical phenotype encompasses the cerebellum, hippocampus and corpus callosum. Although, the corpus callosum is larger in the Arid1b<sup>+/−</sup> mice throughout development contradicting previous work and a subset of the human cases. There is also a sex dependence found in the neuroanatomy; differences in males have an early emergence.

In the original human CSS case studies, post-mortem brain samples showed a thinning, or absence, of the corpus callosum. Recent mouse studies, therefore, focussed on that region. Celen et al. reported a smaller corpus callosum in adult Arid1b<sup>+/−</sup> mice [18], confirming what was seen in a subset human patients. Shibutani et al., report no histological abnormalities in the Arid1b<sup>+/−</sup> mice [20]. In our developmental whole brain analysis, we found an increase in the corpus callosum size across development and no differences in adulthood. We also found an age by genotype interaction in the corpus callosum, indicating an increased growth rate. Another regional difference reported by Celen et al. was a decreased dentate gyrus. Again, contradicting this work, we found an increase in the size of the hippocampal formation across development. These inconsistencies with previous work could be attributable to multiple factors including background strain effects, as our mice were on a C57BL6/N background compared to the other studies (C57BL6/J), as well as interactions between age, sex, and genotype. The original CSS case studies reported cerebellar and brain stem differences in patients [9], and these areas were drastically affected in our study (Supplementary Table 2). This confirms in the mouse what has only been reported in humans. Additionally, those case studies also report a Dandy-Walker
malformation (hypoplasia of the cerebellar vermis with cystic dilation of the fourth ventricle), which is also seen in our mice.

Both ARID1B and cerebellar abnormalities have been linked to ASD [63, 64]. In fact, in humans, one of the original MRI findings was an abnormal cerebellar vermis [64], and in mice, the cerebellum has been consistently linked to autism relevant genetic mutations [12, 13, 65-67]. The neuroanatomical phenotype of the Arid1b<sup>-/-</sup> mouse is a smaller cerebellum, particularly of the white matter and vermal structures, in conjunction with a larger relative isocortex, hippocampus, and corpus callosum. This is consistent with several previously published mouse models related to autism, including Chd8<sup>-/-</sup> [11, 68, 69], 22q11.2 [70], Fmr1 [65], Itgb3 [66], En2 [12], Nrnx1a [12], and Shank3 [12]. The relationship to Chd8<sup>-/-</sup> mouse models would be expected as both models are related to chromatin modification. This suggests that a smaller cerebellum, with a larger cortex and corpus callosum, may delineate a neuroanatomical subset of ASD related to chromatin modification.

The neuroanatomical delay, interestingly, is sex dependent. Neuroanatomical differences in Arid1b<sup>-/-</sup> emerge at PND7 in males, but PND40 in females (Figure 4). These genotype based sex differences are similar to normal neuroanatomical sex differences, in that brain regions larger in male mice emerge earlier than those larger in female mice [26]. This indicates that both sex and the Arid1b<sup>-/-</sup> mutation contribute to a differential growth pattern between males and females. This may explain some discrepancies with other preclinical single sex Arid1b<sup>-/-</sup> papers. Previous authors did not look at sex differences outside of a few measures, and either used mixed groups [19], males [20], or females [18]. The sex dependent differences seen in the Arid1b<sup>-/-</sup> mice could offer some explanation for the reported female bias in CSS [71].
While there are some mixed results within previous behavioural investigations of *Arid1b*\(^{+/−}\) mice, there is a general consensus in three areas [18-20]. 1) A generalized increase in anxiety levels; corroborated in our findings. 2) A generalized decrease in sociability. We observed normal 3-chambered social approach, but found that male *Arid1b*\(^{+/−}\) mice spent less time engaging in social interactions in the male-female social dyad interaction assay [48, 49]. And finally, 3) an increase in repetitive behaviour in *Arid1b*\(^{+/−}\) mice; our tests in self-grooming and marble burying (data not shown) were unable to corroborate this phenotype.

Behavioural seizure susceptibility has not previously been investigated, but is relevant in both CSS and ASD [6]. We observed a strong seizure susceptibility phenotype in *Arid1b*\(^{+/−}\) mice (Figure 7). *Arid1b*\(^{+/−}\) mice were functionally susceptible to PTZ-induced seizures, related to GABAergic neuron plasticity, we attribute this to an excitatory-inhibition imbalance, as with other genetic models of NDD [72-74]. This hypothesis corroborates histological analysis by Jung et al. [19].

The developmental delay in the *Arid1b*\(^{+/−}\) mice, with substantial reductions in both size and basal locomotor activity, can cause interpretive issues with motor skill assessments. Control testing in open fields are commonly used in NDD. However, it cannot be overstated that the results of those observations may be driven by size or low motor skills [44, 45]: a small mouse tested in large chambers will always be at a disadvantage. We had the unique opportunity to observe the impact of an ~15% size reduction and reduced locomotive activity on the control and sophisticated behavioural outcomes. While *Arid1b*\(^{+/−}\) are small and weak throughout life, they have preserved motor coordination, gait abilities, and motor learning abilities (Figure 6).

*Arid1b*\(^{+/−}\) mice showed a deficit in novel object recognition, which aligns with cognitive deficits previously found [19]. However, when tested in motor-independent tasks, we did not observe this phenotype. *Arid1b*\(^{+/−}\) mice showed normal contextual and cued memory in a fear conditioning
paradigm and showed normal or even slightly better performance in a pairwise discrimination task. By investigating learning and memory abilities that do not rely on high levels of movement and exploration, we avoided potential confounds regarding cognitive deficits, indicating that $\text{Arid1b}^{+/c}$ mice were not impaired on motor-independent cognitive tasks. This was surprising as $\text{ARID1B}$ is a common gene mutated in patients with intellectual disability, and learning difficulties are common in children with CSS and ASD [3, 10, 75]. This may reflect limitations with mice as a model system, or the high heterogeneity observed in human population. Previous groups reported learning and memory deficits without factoring in the motor impairment despite its phenotypic presence in these reports [44, 76]. This comprehensive profile of behavioural phenotypes is essential for bridging the gap and improving translational predictability.

**Limitations**

There are a few limitations to this study. The first of which is that the imaging and behavioural testing were performed on different cohorts. The does not allow the ability to link the behavioural findings directly to specific neuroanatomy nor the reverse of linking the imaging findings to a specific behavioural trait. Additionally, this is one of several papers that have now examined the $\text{Arid1b}^{+/c}$ mouse, and some of our findings are confirmatory, but also several are contradictory to previously published works in mouse and human subjects. However, we feel that both the behaviour and imaging performed in this work are robust and reproducible, such that we have full confidence in the data presented herein.

**Conclusions**

Overall, our behavioural findings are similar to earlier reports, but we are seeing a much more subtle effect, with fewer findings in the ASD-relevant behaviours. Methodological differences,
including test order, could account for our inability to see differences as robust as the Jung et al.,
group. Their behavioural battery began with the stressful Morris water maze testing and T-maze
food restriction, while ours followed the standard order of less stressful to most [11, 77]. In
addition, our sample size was twice the sample size per sex than the Jung study. Behavioural
differences could also be attributed to background strain effects [78]. The hallmark characteristic
that appears to define the *Arid1b*+/- mice is a developmental delay, as evidenced by an overall
total brain volume loss throughout development, and reduced body weight, length, and head width
seen as early as PND8. Thus, stunted growth and failure to thrive seen in clinical populations is
robustly recapitulated in our model. The cerebellum and hippocampus are also affected
throughout development, showing both a genotype effect, a sex by genotype interaction, and an
age by genotype interaction. Interestingly, in the adult mice, this hippocampal finding is absent,
suggesting that while the developmental pattern of the hippocampus is affected, it does return to
the appropriate size by adulthood. The effect of this altered development warrants further
investigation since underlying differences in adulthood may be undetected at the mesoscopic
scale of MRI examined here. However, in learning and memory behavioural tasks, *Arid1b* mutants
did not show any deficits in hippocampal memory tasks as adults, suggesting functional effects,
if any, do not persist.
List of Abbreviations

Arid1b – AT-Rich Interactive Domain 1B
ASD – Autism Spectrum Disorder
USV – Ultrasonic Vocalizations
NDD - Neurodevelopmental Disorder
ID – Intellectual Disability
SFARI – Simons Foundation Autism Research Initiative
CSS – Coffin-Siris Syndrome
MRI – Magnetic Resonance Imaging
MEMRI – Manganese Enhanced MRI
TCP – The Centre for Phenogenomics
MICe – Mouse Imaging Centre
CMMR – Canadian Mutant Mouse Repository
WT – Wild-Type
GO – Gene Ontology
T2 – Transverse Relaxation Time
TR – Repetition Time
TE – Echo Time
PND – Post Natal Day
ANTS – Advanced Normalization Tools
NOR – Novel Object Recognition
PTZ – Pentylenetetrazol
Declarations

Ethics approval and consent to participate

All procedures were approved by the Institutional Animal Care and Use Committee either at The Centre for Phenogenomics (TCP) in Toronto, Ontario, Canada, or at the University of California Davis School of Medicine in Sacramento, CA, USA, and were conducted in accordance with the Canadian Council on Animal Care Guide to Care and Use of Experimental Animals or the National Institutes of Health Guide for the Care and Use of Laboratory Animals. Efforts were made to minimize animal suffering and the number of animals used.

Consent for publication

Not Applicable

Availability of data and Materials

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request. Additionally, The neuroimaging data will be released at https://www.braincode.ca/ after the publication of this article.

Competing Interests

The authors declare that they have no conflicts of interest.

Funding

This work was supported by the Canadian Institute for Health Research and the Ontario Brain Institute (JE, JPL) and the NIH R01NS097808 and P50 HD103526 (PI, Abbeduto) (SPP and JLS).
Author’s Contributions

JE, AK, LRQ and DF performed the MR imaging and analysis, AK and ZL maintained and bred the colonies at TCP. AW and ASN performed the RNA experiments and analysis. SPP and JLS performed the behavioural experiments and analysis. AC and LN created the mouse at TCP. SPP, JE, JLS, and JPL designed the experiments and wrote the paper.
References

1. Banerjee-Basu S, Packer A: SFARI Gene: an evolving database for the autism research community. Disease models & mechanisms 2010:133–135.

2. Nord AS, Roeb W, Dickel DE, Walsh T, Kusenda M, O'Connor KL, Malhotra D, McCarthy SE, Stray SM, Taylor SM, Sebat J, STAART Psychopharmacology Network, King B, King M-C, McClellan JM: Reduced transcript expression of genes affected by inherited and de novo CNVs in autism. Eur J Hum Genet 2011, 19:727–731.

3. Halgren C, Kjaergaard S, Bak M, Hansen C, El-Schich Z, Anderson CM, Henriksen KF, Hjalgrim H, Kirchhoff M, Bijlsma EK, Nielsen M, Hollander den NS, Ruivenkamp CAL, Isidor B, Le Caignec C, Zannolli R, Mucciolo M, Renieri A, Mari F, Anderlid B-M, Andrieux J, Dieux A, Tommerup N, Bache I: Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. Clin Genet 2012, 82:248–255.

4. Ronan JL, Wu W, Crabtree GR: From neural development to cognition: unexpected roles for chromatin. Nature Reviews Genetics 2015 17:1 2013, 14:347–359.

5. Coffin GS, Siris E: Mental retardation with absent fifth fingernail and terminal phalanx. Am J Dis Child 1970, 119:433–439.

6. Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, Stephens K, Amemiya A, Schrier Vergano S, Santen G, Wieczorek D, Wollnik B, Matsumoto N, Deardorff MA: Coffin-Siris Syndrome. 1993.

7. Tunnessen WW, McMillan JA, Levin MB: The Coffin-Siris syndrome. Am J Dis Child 1978, 132:393–395.

8. Imai T, Hattori H, Miyazaki M, Higuchi Y, Adachi S, Nakahata T: Dandy-Walker variant in Coffin-Siris syndrome. Am J Med Genet 2001, 100:152–155.

9. DeBassio WA, Kemper TL, Knoefel JE: Coffin-Siris syndrome. Neuropathologic findings. Arch Neurol 1985, 42:350–353.

10. van der Sluijs PJ, Jansen S, Vergano SA, Adachi-Fukuda M, Alanay Y, AlKindy A, Baban A, Bayat A, Beck-Wödl S, Berry K, Bijlsma EK, Bok LA, Brouwer AFJ, van der Burgt I, Campeau PM, Canham N, Chrzanowska K, Chu YWY, Chung BHY, Dahan K, De Rademaeker M, Destree A, Dudding-Byth T, Earl R, Elcioglu N, Elias ER, Fagerberg C, Gardham A, Gener B, Gerkes EH, et al.: The ARID1B spectrum in 143 patients: from nonsyndromic intellectual disability to Coffin-Siris syndrome. Genet Med 2018:1.

11. Gompers AL, Su-Feher L, Ellegood J, Copping NA, Riyadh MA, Stradleigh TW, Pride MC, Schaffler MD, Wade AA, Catta-Preta R, Zdilar I, Louis S, Kaushik G, Mannion BJ, Plajzer-Frick I, Afzal V, Visel A, Pennacchio LA, Dickel DE, Lerch JP, Crawley JN, Zarbalis KS, Silverman JL, Nord AS: Germline Chd8 haploinsufficiency alters brain development in mouse. Nat Neurosci 2017, 20:1062–1073.

12. Ellegood J, Anagnostou E, Babineau BA, Crawley JN, Lin L, Genestine M, DiCicco-Bloom E, Lai JKY, Foster JA, Peñagarikano O, Geschwind DH, Pacey LK, Hampson DR, Laliberté CL,
Mills AA, Tam E, Osborne LR, Kouser M, Espinosa-Becerra F, Xuan Z, Powell CM, Raznahan A, Robins DM, Nakai N, Nakatani J, Takumi T, van Eede MC, Kerr TM, Muller C, Blakely RD, et al.: Clustering autism: using neuroanatomical differences in 26 mouse models to gain insight into the heterogeneity. Mol Psychiatry 2015, 20:118–125.

13. Ellegood J, Lerch JP, Henkelman RM: Brain abnormalities in a Neuroligin3 R451C knockin mouse model associated with autism. Autism Res 2011, 4:368–376.

14. Zerbi V, Ielacqua GD, Markicevic M, Haberl MG, Ellisman MH, A-Bhaskaran A, Frick A, Rudin M, Wenderoth N: Dysfunctional Autism Risk Genes Cause Circuit-Specific Connectivity Deficits With Distinct Developmental Trajectories. Cereb Cortex 2018, 28:2495–2506.

15. Horder J, Petrinovic MM, Mendez MA, Bruns A, Takumi T, Spooren W, Barker GJ, Künnecke B, Murphy DG: Glutamate and GABA in autism spectrum disorder—a translational magnetic resonance spectroscopy study in man and rodent models. Transl Psychiatry 2018, 8:106.

16. Bertram J, Koschützke L, Pfannmüller JP, Esche J, van Diepen L, Kuss AW, Hartmann B, Bartsch D, Lotze M, Bohlen Und Halbach von O: Morphological and behavioral characterization of adult mice deficient for SrGAP3. Cell Tissue Res 2016, 366:1–11.

17. Ellegood J, Crawley JN: Behavioral and Neuroanatomical Phenotypes in Mouse Models of Autism. Neurotherapeutics 2015, 12:521–533.

18. Celen C, Chuang J-C, Luo X, Nijem N, Walker AK, Chen F, Zhang S, Chung AS, Nguyen LH, Nassour I, Budhipramono A, Sun X, Bok LA, McEntagart M, Gevers EF, Birnbaum SG, Eisch AJ, Powell CM, Ge W-P, Santen GW, Chahrour M, Zhu H: Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. Elife 2017, 6:2081.

19. Jung E-M, Moffat JJ, Liu J, Dravid SM, Gurumurthy CB, Kim W-Y: Arid1b haploinsufficiency disrupts cortical interneuron development and mouse behavior. Nat Neurosci 2017, 20:1694–1707.

20. Shibutani M, Horii T, Shoji H, Morita S, Kimura M, Terawaki N, Miyakawa T, Hatada I: Arid1b Haploinsufficiency Causes Abnormal Brain Gene Expression and Autism-Related Behaviors in Mice. Int J Mol Sci 2017, 18:1872.

21. Cahill LS, Laliberté CL, Ellegood J, Spring S, Gleave JA, Eede MCV, Lerch JP, Henkelman RM: Preparation of fixed mouse brains for MRI. Neuroimage 2012, 60:933–939.

22. Lerch JP, Sled JG, Henkelman RM: MRI phenotyping of genetically altered mice. Methods Mol Biol 2011, 711:349–361.

23. de Guzman AE, Wong MD, Gleave JA, Nieman BJ: Variations in post-perfusion immersion fixation and storage alter MRI measurements of mouse brain morphometry. Neuroimage 2016, 142:687–695.

24. Bock NA, Konyer NB, Henkelman RM: Multiple-mouse MRI. Magn Reson Med 2003, 49:158–167.
25. Spencer Noakes TL, Henkelman RM, Nieman BJ: **Partitioning k-space for cylindrical three-dimensional rapid acquisition with relaxation enhancement imaging in the mouse brain.** *NMR Biomed* 2017, **30**.

26. Qiu LR, Fernandes DJ, Szulc-Lerch KU, Dazai J, Nieman BJ, Turnbull DH, Foster JA, Palmet MR, Lerch JP: **Mouse MRI shows brain areas relatively larger in males emerge before those larger in females.** *Nat Commun* 2018, **9**:2615.

27. Szulc KU, Lerch JP, Nieman BJ, Bartelle BB, Friedel M, Suero-Abreu GA, Watson C, Joyner AL, Turnbull DH: **4D MEMRI atlas of neonatal FVB/N mouse brain development.** *Neuroimage* 2015, **118**:49–62.

28. Nieman BJ, Szulc KU, Turnbull DH: **Three-dimensional, in vivo MRI with self-gating and image coregistration in the mouse.** *Magn Reson Med* 2009, **61**:1148–1157.

29. Collins DL, Neelin P, Peters TM, Evans AC: **Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space.** *J Comput Assist Tomogr* 1994, **18**:192–205.

30. Avants BB, Epstein CL, Grossman M, Gee JC: **Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain.** *Med Image Anal* 2008, **12**:26–41.

31. Avants BB, Tustison NJ, Song G, Cook PA, Klein A, Gee JC: **A reproducible evaluation of ANTs similarity metric performance in brain image registration.** *Neuroimage* 2011, **54**:2033–2044.

32. Lerch JP, Carroll JB, Spring S, Bertram LN, Schwab C, Hayden MR, Henkelman RM: **Automated deformation analysis in the YAC128 Huntington disease mouse model.** *Neuroimage* 2008, **39**:32–39.

33. Nieman BJ, Flenniken AM, Adamson SL, Henkelman RM, Sled JG: **Anatomical phenotyping in the brain and skull of a mutant mouse by magnetic resonance imaging and computed tomography.** *Physiol Genomics* 2006, **24**:154–162.

34. Dorr AE, Lerch JP, Spring S, Kabani N, Henkelman RM: **High resolution three-dimensional brain atlas using an average magnetic resonance image of 40 adult C57Bl/6J mice.** *Neuroimage* 2008, **42**:60–69.

35. Steadman PE, Ellegood J, Szulc KU, Turnbull DH, Joyner AL, Henkelman RM, Lerch JP: **Genetic effects on cerebellar structure across mouse models of autism using a magnetic resonance imaging atlas.** *Autism Res* 2014, **7**:124–137.

36. Ullmann JFP, Watson C, Janke AL, Kurniawan ND, Reutens DC: **A segmentation protocol and MRI atlas of the C57BL/6J mouse neocortex.** *Neuroimage* 2013, **78**:196–203.

37. Richards K, Watson C, Buckley RF, Kurniawan ND, Yang Z, Keller MD, Beare R, Bartlett PF, Egan GF, Galloway GJ, Paxinos G, Petrou S, Reutens DC: **Segmentation of the mouse hippocampal formation in magnetic resonance images.** *Neuroimage* 2011, **58**:732–740.
38. Genovese CR, Lazar NA, Nichols T: Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *Neuroimage* 2002, 15:870–878.

39. Satterthwaite FE: Synthesis of variance. *Psychometrika* 1941, 6:309–316.

40. Fox WM: Reflex-ontogeny and behavioural development of the mouse. *Anim Behav* 1965, 13:234–241.

41. Adhikari A, Copping NA, Onaga B, Pride MC, Coulson RL, Yang M, Yasui DH, LaSalle JM, Silverman JL: Cognitive deficits in the Snord116 deletion mouse model for Prader-Willi syndrome. *Neurobiol Learn Mem* 2018.

42. Flannery BM, Silverman JL, Bruun DA, Puhger KR, McCoy MR, Hammock BD, Crawley JN, Lein PJ: Behavioral assessment of NIH Swiss mice acutely intoxicated with tetramethylenedisulfotetramine. *Neurotoxicol Teratol* 2015, 47:36–45.

43. Copping NA, Christian SGB, Ritter DJ, Islam MS, Buscher N, Zolkowska D, Pride MC, Berg EL, LaSalle JM, Ellegood J, Lerch JP, Reiter LT, Silverman JL, Dindot SV: Neuronal overexpression of Ube3a isoform 2 causes behavioral impairments and neuroanatomical pathology relevant to 15q11.2-q13.3 duplication syndrome. *Hum Mol Genet* 2017, 26:3995–4010.

44. Sukoff Rizzo SJ, Silverman JL: Methodological Considerations for Optimizing and Validating Behavioral Assays. *Curr Protoc Mouse Biol* 2016, 6:364–379.

45. Silverman JL, Ellegood J: Behavioral and neuroanatomical approaches in models of neurodevelopmental disorders: opportunities for translation. *Curr Opin Neurol* 2018, 31:126–133.

46. Sashindranath M, Daglas M, Medcalf RL: Evaluation of gait impairment in mice subjected to craniotomy and traumatic brain injury. *Behavioural Brain Research* 2015, 286:33–38.

47. Gulinello M, Mitchell HA, Chang Q, Timothy O'Brien W, Zhou Z, Abel T, Wang L, Corbin JG, Veeraragavan S, Samaco RC, Andrews NA, Fagiolini M, Cole TB, Burbacher TM, Crawley JN: Rigor and reproducibility in rodent behavioral research. *Neurobiol Learn Mem* 2018.

48. Silverman JL, Pride MC, Hayes JE, Puhger KR, Butler-Struben HM, Baker S, Crawley JN: GABAB Receptor Agonist R-Baclofen Reverses Social Deficits and Reduces Repetitive Behavior in Two Mouse Models of Autism. *Neuropsychopharmacology* 2015, 40:2228–2239.

49. Dhamne SC, Silverman JL, Super CE, Lammers SHT, Hameed MQ, Modi ME, Copping NA, Pride MC, Smith DG, Rotenberg A, Crawley JN, Sahin M: Replicable in vivo physiological and behavioral phenotypes of the Shank3B null mutant mouse model of autism. *Mol Autism* 2017, 8:26.

50. Yang M, Silverman JL, Crawley JN: Automated three-chambered social approach task for mice. *Curr Protoc Neurosci* 2011, Chapter 8:Unit 8.26.
51. Bales KL, Solomon M, Jacob S, Crawley JN, Silverman JL, Larke RH, Sahagun E, Puhger KR, Pride MC, Mendoza SP: Long-term exposure to intranasal oxytocin in a mouse autism model. Transl Psychiatry 2014, 4:e480–e480.

52. Copping NA, Berg EL, Foley GM, Schaffler MD, Onaga BL, Buscher N, Silverman JL, Yang M: Touchscreen learning deficits and normal social approach behavior in the Shank3B model of Phelan-McDermid Syndrome and autism. Neuroscience 2017, 345:155–165.

53. Scattoni ML, Ricceri L, Crawley JN: Unusual repertoire of vocalizations in adult BTBR T+tf/J mice during three types of social encounters. Genes, Brain and Behavior 2011, 10:44–56.

54. Brigman JL, Powell EM, Mittleman G, Young JW: Examining the genetic and neural components of cognitive flexibility using mice. Physiol Behav 2012, 107:666–669.

55. Brigman JL, Daut RA, Wright T, Gunduz-Cinar O, Davis MI, Jiang Z, Saksida LM, Jinde S, Pease M, Bussey TJ, Lovinger DM, Nakazawa K, Holmes A: GluN2B in corticostriatal circuits governs choice learning and choice shifting. Nat Neurosci 2013, 16:1101–1110.

56. Horner AE, Heath CJ, Hvoslef-Eide M, Kent BA, Kim CH, Nilsson SRO, Alsiö J, Oomen CA, Holmes A, Saksida LM, Bussey TJ: The touchscreen operant platform for testing learning and memory in rats and mice. Nat Protoc 2013, 8:1961–1984.

57. Dhir A: Pentylenetetrazol (PTZ) Kindling Model of Epilepsy. Volume 18. Hoboken, NJ, USA: John Wiley & Sons, Inc; 2001.

58. Branchi I, Santucci D, Alleva E: Analysis of ultrasonic vocalizations emitted by infant rodents. Curr Protoc Toxicol 2006, Chapter 13:Unit13.12.

59. Branchi I, Santucci D, Alleva E: Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development. Behavioural Brain Research 2001, 125:49–56.

60. Branchi I, Santucci D, Vitale A, Alleva E: Ultrasonic vocalizations by infant laboratory mice: a preliminary spectrographic characterization under different conditions. Dev Psychobiol 1998, 33:249–256.

61. Rocha L, Briones M, Ackermann RF, Anton B, Maidment NT, Evans CJ, Engel J: Pentylenetetrazol-induced kindling: early involvement of excitatory and inhibitory systems. Epilepsy Res 1996, 26:105–113.

62. Macdonald RL, Barker JL: Pentylenetetrazol and penicillin are selective antagonists of GABA-mediated post-synaptic inhibition in cultured mammalian neurones. Nature 1977, 267:720–721.

63. Saitoh O, Courchesne E: Magnetic resonance imaging study of the brain in autism. Psychiatry Clin Neurosci 1998, 52 Suppl:S219–22.

64. Courchesne E, Yeung-Courchesne R, Press GA, Hesselink JR, Jernigan TL: Hypoplasia of cerebellar vermal lobules VI and VII in autism. N Engl J Med 1988, 318:1349–1354.
65. Ellegood J, Pacey LK, Hampson DR, Lerch JP, Henkelman RM: **Anatomical phenotyping in a mouse model of fragile X syndrome with magnetic resonance imaging.** Neuroimage 2010, 53:1023–1029.

66. Ellegood J, Henkelman RM, Lerch JP: **Neuroanatomical Assessment of the Integrin β3 Mouse Model Related to Autism and the Serotonin System Using High Resolution MRI.** Front Psychiatry 2012, 3:37.

67. Stoodley CJ, D'Mello AM, Ellegood J, Jakkamsetti V, Liu P, Nebel MB, Gibson JM, Kelly E, Meng F, Cano CA, Pascual JM, Mostofsky SH, Lerch JP, Tsai PT: **Altered cerebellar connectivity in autism and cerebellar-mediated rescue of autism-related behaviors in mice.** Nat Neurosci 2017, 20:1744–1751.

68. Suetterlin P, Hurley S, Mohan C, Riegman KLH, Pagani M, Caruso A, Ellegood J, Galbusera A, Crespo-Enriquez I, Michetti C, Yee Y, Ellingford R, Brock O, Delogu A, Francis-West P, Lerch JP, Scattoni ML, Gozzi A, Fernandes C, Basson MA: **Altered Neocortical Gene Expression, Brain Overgrowth and Functional Over-Connectivity in Chd8 Haploinsufficient Mice.** Cereb Cortex 2018, 28:2192–2206.

69. Jung H, Park H, Choi Y, Kang H, Lee E, Kweon H, Roh JD, Ellegood J, Choi W, Kang J, Rhim I, Choi S-Y, Bae M, Kim S-G, Lee J, Chung C, Yoo T, Park H, Kim Y, Ha S, Um SM, Mo S, Kwon Y, Mah W, Bae YC, Kim H, Lerch JP, Paik S-B, Kim E: **Sexually dimorphic behavior, neuronal activity, and gene expression in Chd8-mutant mice.** Nat Neurosci 2018, 21:1218–1228.

70. Ellegood J, Markx S, Lerch JP, Steadman PE, Genç C, Provenzano F, Kushner SA, Henkelman RM, Karayiorgou M, Gogos JA: **Neuroanatomical phenotypes in a mouse model of the 22q11.2 microdeletion.** Mol Psychiatry 2014, 19:99–107.

71. Qazi QH, Heckman LS, Markouizos D, Verma RS: **The Coffin-Siris syndrome.** J Med Genet 1990, 27:333–336.

72. Souchet B, Guedj F, Sahún I, Duchon A, Daubigney F, Badel A, Yanagawa Y, Barallobre MJ, Dierssen M, Yu E, Herault Y, Arbones M, Janel N, Créau N, Delabar JM: **Excitation/inhibition balance and learning are modified by Dyrk1a gene dosage.** Neurobiology of Disease 2014, 69:65–75.

73. Golden CE, Buxbaum JD, De Rubeis S: **Disrupted circuits in mouse models of autism spectrum disorder and intellectual disability.** Curr Opin Neurobiol 2018, 48:106–112.

74. Antoine MW, Langberg T, Schnepel P, Feldman DE: **Increased Excitation-Inhibition Ratio Stabilizes Synapse and Circuit Excitability in Four Autism Mouse Models.** Neuron 2019, 101:648–661.e4.

75. Hoyer J, Ekici AB, Endele S, Popp B, Zweier C, Wiesener A, Wohleber E, Dufke A, Rossier E, Petsch C, Zweier M, Göhring I, Zink AM, Rappold G, Schröck E, Wieczorek D, Riess O, Engels H, Rauch A, Reis A: **Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability.** Am J Hum Genet 2012, 90:565–572.
76. Silverman JL, Nithianantharajah J, Der-Avakian A, Young JW, Sukoff Rizzo SJ: Lost in Translation: At the crossroads of face validity and translational utility of behavioural assays in animal models for the development of therapeutics. Neurosci Biobehav Rev.

77. Silverman JL, Yang M, Lord C, Crawley JN: Behavioural phenotyping assays for mouse models of autism. Nat Rev Neurosci 2010, 11:490–502.

78. Copping NA, Adhikari A, Petkova SP, Silverman JL: Genetic backgrounds have unique seizure response profiles and behavioral outcomes following convulsant administration. Epilepsy Behav 2019, 101(Pt A):106547.
**Figures**

**Figure 1**

(A) Scatterplot of differentially expressed genes in heterozygous mice. Genes with $p$-value < 0.05 are in red. (B) Bar plot indicating logRPKM *Arid1b* expression between wild-type (WT) and heterozygous (HT) adult mutants. (C) Table showing select pathways enriched in differentially expressed genes meeting an FDR cutoff of 0.2. Pathways enriched in up-regulated genes are shown in the peach color. Pathways enriched in down-regulated genes are shown in blue. Pathways shown have 500 or fewer genes annotated to their category. Ontologies are biological pathways (BP), molecular function (MF), or cellular component (CC).
Figure 2 – Structural Differences in Adult Arid1b+/− Mice - Neuroanatomical difference found throughout the brain in the ex vivo adult brain cohort. Shown in both absolute and relative volume in the full group, males, and females. Absolute volume differences are measured in mm³, whereas relative volume differences are calculated from a linear model in which the total brain volume is used as a covariate.
Figure 3 – Effects of Arid1b<sup>+/−</sup> on the Brain Throughout Development - A) Highlights the results from the linear mixed effects model. The first column represents the main effect of genotype across the study sample. The next highlights the sex by genotype interaction, followed by the age by genotype interaction, and the last column shows the three-way interaction between age, sex, and genotype. The growth rates of the B) hippocampal region, C) cerebellum, D) deep cerebellar nuclei, and E) fiber tracts are shown throughout development. The trend lines shown and the 95% confidence intervals are based on predictions from the linear mixed effects model.
Figure 4 – The effect of *Arid1b*+/− on the brain differs between sexes - A) Highlights the top 50 structures on coronal slices that were larger or smaller in the *Arid1b*+/− mouse at PND60. The delayed emergence of the differences in the female *Arid1b*+/− mutants are show for the smaller (B) and larger (C) structures. The dotted line shows when significant differences emerge for the *Arid1b*+/− mice vs. the WT. The trend lines and the 95% confidence intervals shown are based on predictions from the linear mixed effects model.
**Figure 5 – Neonatal ultrasonic vocalization emissions, developmental growth, neurological reflexes and developmental delays in Arid1b+/-** (A) Across neonatal developmental life, Arid1b+/- pups displayed abnormal ultrasonic vocalization (USVs) emissions including delay of peak number of calls and decreased total number of USVs on several days compared to Arid1b+/+. (B) When summed, Arid1b+/- pups emitted significantly fewer USVs compared to Arid1b+/+ littermate controls. (C) No genotype difference in axillary abdominal temperature was found, confirming fewer USVs were not the result of neonatal hypothermia. (D) Arid1b+/- weighed less across development beginning at postnatal 8 and continuously throughout development compared to sex-matched Arid1b+/+. (E) Across neonatal developmental life, Arid1b+/- pups were shorter by body length (F) and narrower by head width measurements. (G) While Arid1b+/- showed normal latencies to perform the righting reflex, a measure of limb coordination in early life, (H) Arid1b+/- had prominent deficits in negative geotaxis (i.e., incline reorientation) and (I) the ability to traverse out of circle by walking.
Figure 6 – Adult *Arid1b<sup>−/−</sup>* have diminished physical size, strength and motor ability. (A) Adult *Arid1<sup>−/−</sup>* weigh less compared to *Arid1b<sup>+/+</sup>* adults and throughout behavioural testing. (B) In a forelimb grip strength assay, *Arid1b<sup>−/−</sup>* showed decreased maximum forelimb force exerted, indicating reduced muscle strength. (C-D) Despite their smaller size, *Arid1b<sup>−/−</sup>* showed normal stride length (C) and stride frequency (D) in fore and hind paws using DigiGait automated gait analysis, indicating normal ambulation and motor coordination of the limbs. (E-H) In open field exploration assay, horizontal, vertical and total activity were recorded for a 30-minute period. Data are shown over time in 5-minute bins. *Arid1b<sup>−/−</sup>* were hypoactive with reduced horizontal activity (E), vertical activity (F) and total activity (G). (H) When summed over the 30-minute session, *Arid1b<sup>−/−</sup>* had less total activity compared to *Arid1b<sup>+/+</sup>* (H), indicating a robust motor deficit. For (A-B), *Arid1b<sup>+/+</sup>* N=28, *Arid1b<sup>−/−</sup>* N=25, for (C-D) *Arid1b<sup>−/−</sup>* N=21, *Arid1b<sup>+/+</sup>* N=19, for (E-H) *Arid1b<sup>−/−</sup>* N=28, *Arid1b<sup>−/−</sup>* N=26. * p<.05 vs *Arid1b<sup>+/+</sup>* by repeated measures two-way ANOVA or student’s unpaired t-test. † p<.09 vs *Arid1b<sup>+/+</sup>* by repeated measures two-way ANOVA. Error bars represent mean ±SEM.
Figure 7 – *Arid1b*+/− exhibited mild ASD-relevant social communication phenotypes in the direct reciprocal social interaction and seizure susceptibility but no repetitive behaviour.

Three-chambered social approach is a standard measure of social behaviour. It was used as a preliminary evaluation tool for sociability. While both genotypes exhibited normal social approach i.e., they spent more time in the chamber with a novel mouse or time sniffing the novel mouse compared to a novel object, the *Arid1b*+/− made fewer total entries during the task. (A) *Arid1b*+/− showed no deficits compared to *Arid1b*+/+ littermates in 3-chambered social approach in time spent in chamber with the novel mouse, (B) time spent sniffing the novel mouse compared to novel object also indicated sociability in both genotypes. (C) *Arid1b*+/− made fewer transitions...
compared to Arid1b	extsuperscript{+/+}. (D-F) Next, we moved to a more sensitive task to observe social dyad interactions. During male-female social dyad interactions, males are introduced to a WT, novel estrous female for 5 minutes in a novel, clean environment and social investigative and situational behaviours are evaluated and male-emitted ultrasonic vocalizations (USVs) are counted. (D) Adult Arid1b	extsuperscript{+/-} males spent less time nose-to-anogenital sniffing and (E) less time in a following posture/behaviour, the two main components of this interaction, indicating reduced social behaviour in this task. (F) Arid1b	extsuperscript{+/-} males also emitted fewer USVs during the dyad interactions compared to Arid1b	extsuperscript{+/+} male littermates which also suggested impaired social communication. (G) No increased or decreased self-grooming was observed. (H-I) We used a gold standard assay of anxiety, the elevated plus maze, to observe anxiety-like behaviour. (H) Arid1b	extsuperscript{+/-} spent fewer total seconds on the open arm which usually indicates anxiety-like behaviour. (I) Arid1b	extsuperscript{+/-} also showed fewer total entries between arms, which suggests this task may be confounded by the data in Figure 2, the robust motoric deficit. (J-L) With a high concordance of epilepsy in ASD, and excitatory inhibitory balance being a prominent theory, we investigated seizure susceptibility in Arid1b	extsuperscript{+/-}. Mice are injected intraperitoneally with pentylenetetrazol (PTZ) chemoconvulsant, and latency to seizure events including loss of righting, tonic-clonic seizure and death were recorded. (J) Arid1b	extsuperscript{+/-} showed seizure vulnerability and susceptibility by decreased in time to loss of righting after PTZ administration (K) faster onset to tonic-clonic extension and (L) latency to death, compared to Arid1b	extsuperscript{+/+}. For (A-C), Arid1b	extsuperscript{+/+} N=28, Arid1b	extsuperscript{+/-} N=25, for (D-F) Arid1b	extsuperscript{+/-} N=15, Arid1b	extsuperscript{+/-} N=17, in (G), Arid1b	extsuperscript{+/+} N=28, Arid1b	extsuperscript{+/-} N=25, for (G-H), Arid1b	extsuperscript{+/-} N=28, Arid1b	extsuperscript{+/-} N=27, for (J-L), Arid1b	extsuperscript{+/+} N=45, Arid1b	extsuperscript{+/-} N=40. * p<.05 vs Arid1b	extsuperscript{+/-} by Repeated Measures Two-Way ANOVA or Student’s Unpaired t-Test. Error bars represent mean ±SEM.
Figure 8 – Largely intact learning and memory in *Arid1b* 

To assess learning and memory and behavioural flexibility in a computerized automated task, that does not rely heavily on motor skills, we utilized a visual pairwise discrimination touchscreen task. Mice are food-restricted and trained to initiate and execute trials by touching their noses to the touchscreen. Mice are then trained to associate an image (a spider or an airplane, A inset) with a food reward. (A) *Arid1b* 

required fewer sessions to reach criteria of 80% accuracy for two consecutive days compared to *Arid1b* 

littermates indicated no deficits in ability to acquire visual discrimination. (B) As a group, *Arid1b* 

completed acquisition of the pairwise discrimination task faster than *Arid1b* 

, with larger proportions of mice finishing sooner (survival curve). Once animals achieved task criteria, they underwent reversal of the task where the opposite image now becomes the correct image (C, inset). (C-D) *Arid1b* 

showed no behavioural inflexibility or deficits in learning and memory in pairwise discrimination reversal compared to *Arid1b* 

by (C) no difference in number of sessions required to reach criteria (D) and no difference in proportion of mice reaching criteria of 75% accuracy for two consecutive days (survival curve, D). (E-F) In a canonical learning and memory tasks with less motor involvement, Pavlovian contextual and
auditory cue fear conditioning, Arid1b+/+ and Arid1b+- demonstrated normal associative learning and memory abilities. Learning and memory was assessed by percentage of time spent freezing after associating a context and tone cue with a 0.5 mA foot shock. (E) Arid1b+/− successfully acquired fear memory during training and showed intact contextual memory of their environment 24 hours later. There was no difference in % freezing between Arid1b+/+ and Arid1b+- (F). Arid1b+/− also showed intact cued learning of an auditory tone compared to Arid1b+/+.

Recognition memory was tested using a novel object recognition test in an open area. Mice are habituated to an open field arena and 24 hours, familiarized for 10 minutes to two identical objects equally spaced in the arena. After an hour long intertrial interval, one object is replaced with a novel one and mice are allowed to explore for 5 minutes. Time spent sniffing the objects is recorded. Mice exhibit novel object recognition when they spend more time investigating the novel object compared to the familiar object. (G) Arid1b+/+ successfully showed a preference for the novel object and spent more time sniffing the novel object whereas Arid1b+- failed to show a preference and spent equal time sniffing both objects. For (A-B), Arid1b+/+ N=24, Arid1b+- N=15, for (C-D), Arid1b+/+ N=20, Arid1b+- N=14, for (E-F), Arid1b+/+ N=27, Arid1b+- N=23, for (G), Arid1b+/+ N=27, Arid1b+- N=21. * p<.05 vs Arid1b+/+ by repeated measures Two-Way ANOVA or Student’s Unpaired t-Test. Error bars represent mean ±SEM.