THE FUNCTIONAL ROLES FOR SWI/SNF CHROMATIN REMODELING COMPLEXES IN PHYSIOLOGY AND DISEASE

APPROVED BY SUPERVISORY COMMITTEE

Hao Zhu, M.D.

Jiang Wu, Ph.D.

Eric Olson, Ph.D.

Jian Xu, Ph.D.

DEDICATION

Dedicated to my beloved family: Şaban Celen, Bahtışen Celen, Görkem Celen and Haşim Ekmel Ercan

For their love and support

THE FUNCTIONAL ROLES FOR SWI/SNF CHROMATIN REMODELING COMPLEXES IN PHYSIOLOGY AND DISEASE

by

CEMRE CELEN

DISSERTATION

Presented to the Faculty of the Medical School

The University of Texas Southwestern Medical Center

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center
Dallas, Texas

December, 2020

THE FUNCTIONAL ROLES FOR SWI/SNF CHROMATIN REMODELING COMPLEXES IN PHYSIOLOGY AND DISEASE

CEMRE CELEN
The University of Texas Southwestern Medical Center, 2020

Supervising Professor: HAO ZHU, M.D.

Sequencing studies have implicated multiple subunits of SWI/SNF complexes in human neurodevelopmental and psychiatric disorders, as well as in cancers. Particularly haploinsufficiency of *ARID1B*, a SWI/SNF chromatin-remodeling subunit, has been implicated in short stature, autism spectrum disorder, intellectual disability, and corpus callosum agenesis. In addition, *ARID1B* is the most common cause of Coffin-Siris Syndrome, a developmental delay syndrome characterized by some of the above abnormalities. However, its role in pathologies is not well characterized due to absence of in vivo models. Therefore, in the first part of this thesis, we generated *Arid1b* heterozygous mice, which showed social behavior impairment, altered vocalization, anxiety-like behavior, neuroanatomical abnormalities, and

growth impairment. In the brain, *Arid1b* haploinsufficiency resulted in changes in the expression of SWI/SNF-regulated genes implicated in neuropsychiatric disorders. A focus on reversible mechanisms identified insulin-like growth factor deficiency with inadequate compensation by Growth Hormone Releasing Hormone and Growth Hormone, underappreciated findings in *ARID1B* patients. Therapeutically, GH supplementation was able to correct growth retardation and muscle weakness. This model functionally validates the involvement of *ARID1B* in human disorders and allows mechanistic dissection of neurodevelopmental diseases linked to chromatin-remodeling.

ARID1A is a paralogous subunit that is commonly mutated in cancers and plays critical roles in liver regeneration. Chromatin remodeling mechanisms could be generally important for regeneration in other tissues. Since dynamic regulation of β-cell proliferation in pancreatic islets is poorly understood and better understanding could lead to therapeutic approaches for replenishing β-cell mass in type 1 and type 2 diabetes, in the second part of this thesis we focused on the role of ARID1A in β-cells. *Arid1a* is physiologically suppressed when β-cells proliferate during pregnancy or after pancreas resection. Whole-body *Arid1a* knockout mice were protected against streptozotocin induced diabetes. Cell-type and temporally specific genetic dissection showed that β-cell specific *Arid1a* deletion could potentiate β-cell regeneration in multiple contexts. Transcriptomic and epigenomic profiling of mutant islets revealed increased Neuregulin-ERBB-NR4A signaling. Functionally, chemical inhibition of ERBB or NR4A was able to block increased regeneration associated with *Arid1a* loss. Together, this work defined the role of ARID1A in β-cells and provided new insights into the molecular regulators of β-cell regeneration. Overall, we uncovered important roles of ARID1A and ARID1B-containing SWI/SNF complexes in physiological and disease states.

ACKNOWLEDGEMENTS

As much as journey to getting my Ph.D. was filled with excitement and joy, it would have been much more challenging to get through the bumps on the road if it wasn't for the support of my family, friends, mentor, and lab mates. I am grateful to many people for guiding me through my Ph.D. and helping me to develop the necessary skills and perseverance to become a better scientist.

I am grateful to have Dr. Hao Zhu as my mentor during my graduate training. There have been only few decisions that I have felt strongly about in my life and joining Dr. Hao Zhu's lab was definitely one of them. When I was a bright-eyed first year student rotating in his lab, those moments when he would walk over to my bench to exchange ideas and get excited over a piece of data that I have generated that day were so valuable. Every day I am amazed how he can conceptualize, communicate and turn his ideas into action and help you to identify the bigger picture; and I can only aspire and work to integrate some of these traits into my own scientific life. Not only he is a great scientist, but he is also a great human who cares about his trainees and works on creating a lab atmosphere where communication and collaboration between lab members are valued. I want to sincerely thank him for his guidance and support during my training.

I would like to express my gratitude to my thesis committee members Dr. Eric Olson, Dr. Jian Xu and Dr. Jiang Wu for their valuable feedback, scientific advice, and continued support over the years. I would like to thank multiple people for their assistance for my projects: Dr. Shari Birnbaum, Dr. Maria Chahrour and Nadine Nijem for their valuable help with behavior experiments, Dr. Shunli Shen for his help with surgical techniques and colony maintenance, Dr. Angela Walker, Dr. Woo-Ping Ge, Dr. Fei Chen, Dr. Lin Li for their technical help, and Dr. Gijs Santen for clinical data and perspective.

I had great lab mates during my time in Zhu lab and their presence in the lab made everything better. I would like to give special thanks to Dr. Jen-Chieh Chuang for taking the time to teach me at the initial stages of my PhD. I also would like to thank Dr. Shuyuan Zhang, Dr. Liem Nguyen, Dr. Ibrahim Nassour, Dr. Sam Wang, Austin Moore, Yu-Hsuan Lin, and Andrew Chung for the valuable discussions and fun times in and out of the lab.

I would like to acknowledge my primary school teacher Gülsevim Pehlivan for her immense contributions to my education. She laid such strong foundation that it really made an impact on my long-term learning and choices. I am grateful for her commitment and dedication.

During these years, I have also gained lots of new friends who have been there for me at various stages of my time in Dallas. I am thankful to Duygu Saatçioğlu and Didem Ağaç Çobanoğlu for welcoming me to Dallas. I would like to thank my friends Enrique Limon, Claire Taitte, Keira Hansen and Eliane Desjardins for their friendship, fun times and dances. I also would like to thank my lifelong friends Gülfem Demir, Ege Deniz Yıldırım and Müge Katırcı for their companionship at every stage of my life even when we all are spread around the world.

I would like to thank my amazing family Şaban Celen, Bahtışen Celen and Görkem Celen for their love and support throughout all my life. The path I have chosen to follow has not been the easiest and it led me to live away from home since high school. It meant a lot to have an encouraging and supportive family who let me pursue my dreams in life and was always there for me. Despite the distance, you are always in my mind. I would like to thank my grandparents for their care and love during my childhood. Last but not least, I want to thank my partner, Haşim Ekmel Ercan for making my life filled with joy and excitement every single day over last thirteen years. Your truly unconditional love and support made this life even better at good times and gave me hope and strength at dark times. Sharing this life with you puts a smile on my face and I cannot wait for the new adventures awaiting.

TABLE OF CONTENTS

PRIOR PUBLICATIONS

Celen C, Chuang JC, Shen S, Otto JE, Collings CK, Luo X, Lin Li, Wang Y, Wang Z, Jia Y, Sun X, Nassour I, Park J, Ghaben A, Wang T, Wang SC, Scherer PE, Kadoch C, and Zhu H. *Arid1a* loss potentiates β-cell regeneration through activation of EGF signaling. Biorxiv. 2020. doi: https://doi.org/10.1101/2020.02.10.942615.

Celen C*, Chuang JC*, 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 WP, Santen GW, Chahrour M, Zhu H. Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. *ELife*. 2017 Jul 11. 6: e25730. doi: 10.7554/eLife.25730. PMID: 28695822.

Wang Z, Chen K, Jia Y, Chuang JC, Sun X, Lin YH, Celen C, Li L, Huang F, Liu X, Castrillon DH, Wang T, and Zhu H. Dual ARID1A/ARID1B loss leads to rapid carcinogenesis and disruptive redistribution of BAF complexes. *Nature Cancer*. 2020. https://doi.org/10.1038/s43018-020-00109-0

Moore A, Wu L, Chuang JC, Sun X, Luo X, Gopal P, Li L, Celen C, Zimmer M, and Zhu H. *Arid1a* loss drives non-alcoholic steatohepatitis in mice via epigenetic dysregulation of hepatic lipogenesis and fatty acid oxidation. *Hepatology.* 2019 May; 69(5):1931-1945. doi: 10.1002/hep.30487. PMID: 30584660

Bergiers I, Andrews T, Vargel Bölükbaşı Ö, Buness A, Janosz E, Lopez-Anguita N, Ganter K, Kosim K, Celen C, Perçin GI, Collier P, Baying B, Benes V, Hemberg M, Lancrin C. Single cell transcriptomics reveals a new dynamical function of transcription factors during embryonic hematopoiesis. *Elife.* 2018 Mar 20; 7:e29312. doi: 10.7554/eLife.29312. PMID: 29555020

Yilmaz S, Alkaya Uludağ D, Kasapçopur O, Barut K, Akdemir E, Celen C, Yasuno K, Bilguvar K, Günel M, Tüysüz B. Genotype-phenotype investigation of thirty-five patients from 11 unrelated families with Camptodactyly-Arthropathy-Coxavara-Pericarditis (CACP) syndrome. *Molecular Genet Genomic Med*. 2018 Mar; 6(2): 230-248. doi: 10.1002/mgg3.364. PMID: 29397575

Sun X, Wang SC, Wei Y, Luo X, Jia Y, Li L, Gopal P, Zhu M, Nassour I, Chuang JC, Maples T, Celen C, Nguyen LH, Wu L, Fu S, Li W, Hui L, Tian F, Ji Y, Zhang S, Sorouri M, Hwang TH, Letzig L, James L, Wang Z, Yopp AC, Singal AG, Zhu H. (2017) *Arid1a* Has Context-Dependent Oncogenic and Tumor Suppressor Functions in Liver Cancer. *Cancer Cell*. 2017 Nov 13; 32(5): 574-589.e6. doi: 10.1016/j.ccell.2017.10.007.

Sun X, Chuang JC, Kanchwala M, Wu L, Celen C, Li L, Liang H, Zhang S, Maples T, Nguyen LH, Wang SC, Signer RAJ, Sorouri M, Nassour I, Liu X, Xu J, Wu M, Zhao Y, Kuo YC, Wang Z, Xing C, and Zhu H. Suppression of the SWI/SNF Component *Arid1a* Promotes Mammalian Regeneration. *Cell Stem Cell*. 2016 Apr 7;18(4):456-66. doi: 10.1016/j.stem.2016.03.001. PMID: 27044474

*These authors contributed equally to the work.

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ASD	Autism-spectrum disorder
BAF	Brg/Brahma-associated factors
βKO	Beta-cell specific knockout
CDKs	Cyclin-dependent kinases
CHD/M-2	Chromodomain helicase DNA binding
CSS	Coffin-Siris Syndrome
DEN	Diethylnitrosamine
EGFR	Epidermal growth factor receptor
ERBB	Erythroblastic oncogene B
ESC	Embryonic stem cell
GH	Growth hormone
GHRH	Growth hormone releasing hormone
GTT	Glucose tolerance test
ID	Intellectual disability
IGF1	Insulin-like growth hormone 1
IGFBP3	Insulin-like growth hormone binding protein 3
ISWI	Imitation SWI
ITT	Insulin tolerance test
KO	Knockout
MDa	Mega Dalton
NBS	Nicolaides-Baraitser syndrome
NR4A	Nuclear receptor subfamily 4
mSWI/SNF	Mammalian Switch/Sucrose Non-Fermenting

ncBAF	Non-canonical BAF
pBAF	Polybromo-associated BAF
PDX1	Pancreatic And Duodenal Homeobox 1
PPx	Partial pancreatectomy
STZ	Streptozotocin
Srf	Serum response factor
TFs	Transcription factors
USVs	Ultrasonic vocalizations

CHAPTER 1

PROLOGUE

Chromatin structure and regulation

Cells have developed multiple mechanisms to compact and package chromosomal DNA into chromatin. This packaging is essential as about two meters of DNA needs to fit into the nucleus in eukaryotic cells (*1*). Nucleosomes are the repeating structural building blocks of chromatin that were initially described by their similar appearance to beads on a string under the electron microscope (*2*, *3*). The core nucleosome compromises an octamer of core histones containing two copies of H2A, H2B, H3, H4 histones wrapped by 147 base pairs of DNA. These nucleosome cores together with linker DNA and linker histone H1 are assembled into nucleosomes, which can be packaged into higher order structures that allow further compaction

and organization (*4*). Not only does this compaction provide a solution for the space limitation that exists within the nucleus, it also allows chromatin to be dynamically regulated in a tissue and temporally specific manner. There are three main mechanisms that are used to regulate chromatin organization: modification of histones by histone-modifying proteins, DNA by DNA-modifying proteins, and chromatin by ATP-dependent chromatin remodeling complexes (*5*). ATP-dependent chromatin remodeling complexes can be subdivided into four subfamilies: SWI/SNF, ISWI, INO80, and CHD/M2 which all contain a varied Snf2-like ATPase domain that can catalyze the translocation of DNA around the histones. Despite having common ATP-dependent translocation activities, regulation of diverse outcomes are attributed to interaction with different binding factors with activator or repressor functions, in addition to DNA or histone targeting specificity due to diverse subunit compositions (*6*).

Discovery, composition, and function of SWI/SNF ATP-dependent chromatin remodelers

SWI/SNF ATP-dependent chromatin remodeling complexes use the energy derived from ATP hydrolysis to modulate the accessibility of chromatin. They can disrupt DNA nucleosome contacts resulting in eviction, sliding or reconstruction of nucleosomes through histone exchange. SWI/SNF activity is important for the regulation of fundamental processes such as transcription, replication, recombination and DNA repair (*7*). The discovery of ATP-dependent SWI/SNF chromatin remodeling complexes dates back to the 1980s. The first set of genes were identified in screens for signaling factors involved in sucrose fermentation. These mutations defined a new locus called SNF (Sucrose Non-Fermenting) (*8*, *9*). In a separate study five other genes called SWI1 to 5 (Switch) were discovered by screening for mutants that were defective in mating-type switching in yeast (*10*). The name SWI/SNF complex was coined later when it was understood that these genes were related to each other and were found in complexes. SWI/SNF chromatin remodeling complexes are evolutionary conserved from yeast

to flies to mammals (*11*). Indeed, the opposition between SWI/SNF and Polycomb complexes was first discovered in *Drosophila*, and the imbalance between SWI/SNF and Polycomb activity is thought to be one of the fundamental mechanisms behind the malignant and neurodevelopmental disorders caused by SWI/SNF mutations in humans (*12*, *13*).

Mammalian SWI/SNF complexes (also called BAF complexes) are 0.8 to 1.5 MDa complexes containing about 15 subunits encoded from about 29 genes. Each mSWI/SNF complex contains one of the catalytic ATPase subunits, BRG1 (SMARCA4), BRM (SMARCA2) or PBRM1 (BAF180). Other subunits carry structural or targeting roles. Mashtalir et al. in Dr. Kadoch's group recently published a detailed biochemical characterization of mSWI/SNF complexes in which they divided mSWI/SNF complexes into 3 groups according to their assembly pathways and defining subunits: BAF (Brg/Brahma-associated factors), PBAF (polybromo-associated BAF), ncBAF (non-canonical BAF). ARID1A (BAF250A) and ARID1B (BAF250B) are mutually exclusive defining subunits of 1 MDa BAF complexes. PBAF, is the largest complex of the three with a molecular weight of 1.41 MDa, contains four defining subunits including ARID2 (BAF200), PBRM1 (BAF180), BAF45A, and BRD7. The most recently identified ncBAF complex is the smallest complex (0.87 MDa molecular weight) and it does not contain four of the regular BAF complex subunits including ARID (BAF250), SMARCB1 (BAF47, SNF5, INI1), SMARCC2 (BAF170), and SMARCE1 (BAF57). The defining subunits of ncBAF are BRD9 and GLTSCR1/1L (*14*).

SWI/SNF mutations in neurodevelopmental disorders and cancer

In the last 10 years, human exome sequencing studies showed that SWI/SNF complex components are some of the most frequently mutated epigenetic regulators in both cancer and neurodevelopmental disorders. In 2012, Hoyer et al. reported de novo nonsense or frameshift mutations in *ARID1B* in 8 out of 887 (0.9%) unselected patients with unexplained intellectual

disability (ID). All of these patients displayed varying degrees of hypotonia and psychomotor retardation. Speech impairment was also observed in many of the patients along with other findings. Later, Santen et al. identified 3 patients with truncating *ARID1B* mutations exhibiting features of Coffin-Siris Syndrome (CSS) (*15*). Tsurusaki et al. also implicated mutations in multiple SWI/SNF complex components in CSS. 15 genes encoding SWI/SNF components were screened in 23 patients with CSS and 20 patients were shown to contain a de novo germline mutations in *SMARCB1*, *SMARCA4*, *SMARCA2*, *SMARCE1*, *ARID1A* and *ARID1B*. This paper was important because it highlighted the importance of multiple SWI/SNF complex components in CSS (*16*). The third study reported mutations in *SMARCA2* in patients with Nicolaides-Baraitser syndrome (NBS) (*17*, *18*).

Prominent phenotypic features of CSS include mental retardation, excessive hair growth, coarse facial features, sparse scalp hair, developmental delay, and hypoplastic or absent fifth fingers or toenails. Therefore, it has also been named as fifth-digit syndrome. Corpus callosum agenesis and congenital heart defects are among the variable penetrance phenotypes observed in these patients (*19*). CSS is now recognized as SWI/SNF-associated disorder since 60% of patients have mutation in one of the subunits of this complex. Among these SWI/SNF mutations, more than half of them are seen in *ARID1B*. Overall, 37% of CSS patients have mutations in *ARID1B.* In addition, 0.3% of autistic individuals are estimated to have mutations in *ARID1B* (*20*). Also, in a 2015 paper in Nature, the genomes of 1100 patients with unexplained intellectual disability were sequenced and 11 of them (1% of all cases) had heterozygous loss of *ARID1B*, making it the top candidate gene for undiagnosed developmental disorders (*21*). However, until our study there were no mouse models to test the causality and the molecular mechanisms associated with *Arid1b* loss, despite this gene being implicated in wide variety of disorders. One conceptually interesting observation made as a result of these human exome sequencing studies was that most mutations in *ARID1A* and *ARID1B* were

heterozygous nonsense or frameshift mutations that result in haploinsufficiency. These mutations also did not necessarily accumulate in the ARID DNA binding domain. On the other hand, SMARCA4 and SMARCA2 mutations were either nonsense or in-frame deletions accumulating in the ATPase or helicase domains (*22*).

Intriguingly, genome and exome sequencing studies also implicated SWI/SNF complex genes in a variety of human cancers. Cigall Kadoch's lab compiled and analyzed 44 published human genome/exome studies and found that SWI/SNF subunits were mutated in 20% of all cancer types. Among the subunits, *ARID1A* was the most commonly mutated gene. In particular, 45% endometrioid and clear-cell ovarian cancers contained *ARID1A* mutations. Both A*RID1A* and *ARID1B* were also found to be mutated in medulloblastoma, gastric, bladder, hepatocellular, colorectal, melanoma, lung and pancreatic cancers (*23*). Given the mutations in *ARID1A* and *ARID1B* both in neurodevelopmental disorders and cancer, it will be interesting to follow patients who harbor germline SWI/SNF mutations for cancer development. Our studies showed that germline haploinsufficient *Arid1b* mice did not develop spontaneous tumors and did not display increased liver tumor development compared to WT mice after diethylnitrosamine (DEN) exposure. Even though neurodevelopmental disorders and cancer appear to be unrelated, sequencing studies have made it clear that mutations in multiple chromatin remodelers including SWI/SNF components are common in both. One common theme that could connect these two is the role of SWI/SNF complexes in important physiological processes such as self-renewal, proliferation, and differentiation.

SWI/SNF complexes regulate self-renewal, proliferation, and differentiation

Subunit switching and different assemblies of SWI/SNF complexes exist not only between different organisms but also between different tissues and developmental stages. This combinatorial assembly and subunit switch of SWI/SNF complexes is important to generate

biological specificity. Cardiac progenitors are distinguished by the expression of BAF60C (*24*). Lessard and Wu et al. demonstrated that transitioning from proliferating neural stem and progenitor cells to post-mitotic committed neuronal lineages is accompanied by the replacement of BAF45a and BAF53a with BAF45b, BAF45c, and BAF53b subunits. In fact, prevention of this subunit switching impairs neuronal differentiation (*25*).

Deletion of catalytic and core SWI/SNF components BRG1, SNF5, or BAF155 results in lethality at the blastocyst stage, accompanied by impaired inner cell mass formation in mice (*26–28*). However, it was not clear whether non-catalytic subunits such as ARID1A and ARID1B — that were also shown to be not required for in vitro chromatin remodeling functions of SWI/SNF on nucleosome templates — were important for early developmental decisions in vivo. Interestingly, deletion of *Arid1a* resulted in early embryonic lethality in mice and impaired the self-renewal and pluripotency of ES cells (*29*). We also showed that *Arid1b* null mice die perinatally, indicating that ARID1A and ARID1B cannot compensate for each other's loss during development (*30*). Also, ARID1A is also essential for cell cycle arrest whereas ARID1B is not (*31*). In addition, *Arid1b* null ES cells were hard to maintain in undifferentiated state in vitro. They displayed reduced levels of pluripotency marker expression and their proliferation rate was slower than WT ES cells in culture (*32*). Another study by Dr. Moran's group showed that the choice between ARID1A and ARID1B is a determining factor for SWI/SNF's interactions with activator vs repressor E2Fs and HATs vs HDACs. Changes in these interactions can result in opposite roles in cell cycle regulation. In line with the anti-proliferative function of ARID1A, ARID1A-containing complexes seem to be associated with repressive E2Fs (E2F4 and E2F5) and HDACs. In contrast, ARID1B containing complexes associate with both activator E2F1 and HATs and also with repressive E2Fs (E2F4 and E2F5) and HDACs. Pro-proliferative functions of ARID1B-containing complexes still seems to be dominant despite being associated with opposing factors (*33*). Along the same lines, Sun et al.

from our lab showed that *Arid1a* loss in the mouse liver results in dissociation of repressive E2F4 from E2F target genes thereby increasing proliferation in regenerating liver (*34*). My study shows that *Arid1a* loss also promotes partial pancreatectomy (PPx) induced β-cell regeneration and protects against streptozotocin (STZ) induced diabetes, supporting a proliferation suppressor role that could be manipulated to induce the proliferation of β-cells (*35*).

Overview of β-cell regeneration approaches

At various stages of either type 1 and type 2 diabetes, there are absolute or relative deficiencies of functional beta cells. Therefore, understanding how β-cells are replenished in vivo may give us clues for developing therapeutic approaches to expand β-cell mass. Dor et al. in Dr. Melton's group used lineage tracing approaches in mouse models and demonstrated that proliferation of pre-existing β-cells is the main source new β-cells during homeostasis and after PPx in adult mice (*36*). Another lineage tracing study in mice demonstrated that extreme loss of β-cells (>99%) by diphtheria toxin results in conversion of α-cells into β-cells (*37*). Therefore, self-replication of β-cells and direct conversion of other islet cell types into β-cells are both possible in mice depending on the degree and type of injury.

There are multiple approaches to generate β-cells: 1) differentiation of ESCs or iPSCs into β-cells, 2) reprogramming of non-β-cell types into β-cells and 3) stimulating endogenous β-cell regeneration. Protocols for in vitro differentiation of hESCs and iPSCs towards pancreatic hormone expressing endocrine cells and later glucose-responsive islet clusters are continuously being improved over the years (*38–41*). Direct reprogramming of acinar cells into β-cells by adenoviral delivery of NGN3, PDX1, and MAFA have been accomplished in adult mice (*42*) and were successful in alleviating diabetes (*43*). In addition, proof of principle

experiments showed that α-cells, liver, pancreatic ductal, intestinal crypt, and stomach cells can be reprogrammed into insulin-producing cells (*44–48*).

Molecular regulation of endogenous β-cell regeneration

β-cell proliferation seems to be the main mechanism of replenishing β-cells in animal models during postnatal life, although these events appear to be infrequent (*36*, *49*). It has also been shown that the proliferation potential of β-cells decline with age in animal models and adult β-cells have a very slow turnover rate in mice (1 in 1400 mature β-cells per day) (*50*). However, injury and stress such as pregnancy and high fat diet feeding can increase this basal rate of proliferation (*51*, *52*). Therefore, understanding the underlying mechanisms related to age-dependent reduction in β-cell proliferation and the regulators of compensatory β-cell regeneration is essential before being able to conceive of therapeutic strategies.

Krishnamurthy et al. from Dr. Seung Kim's laboratory showed that the increased expression of cyclin-dependent kinase inhibitor $p16^{INK4a}$ and tumor suppressor $p19^{Arf}$ limits β-cell proliferation in aging mice (*53*). Their expression is limited by the H3K27me mark established by EZH2, the catalytic subunit of PRC2 complex. Declining levels of EZH2 in aging mice results in increased levels of $p16^{INK4a}$ and $p19^{Arf}$ and KO of *Ezh2* in β-cells was enough to cause β-cell hypoplasia and diabetes (*54*). Therefore, PRC2 appears to be regulating β-cell proliferation in the opposite direction as SWI/SNF, since we have shown that reduction in ARID1A activates β-cell proliferation. In addition, β-cell specific loss of PRC2 also correlates with β-cell dedifferentiation and features of diabetes in both mice and humans (*55*). In addition to Polycomb, PDGF signaling and p38/MAPK pathways have also been shown to control age-dependent proliferation in β-cells through the regulation of cell cycle inhibitors (*56*, *57*). Moreover, multiple growth factor signaling pathways including EGF are implicated in the regulation of β-cell regeneration. Downregulation of the EGFR pathway in pancreatic

islets results in diabetes due to impaired postnatal growth of β-cells (*58*). Moreover, specific deletion of *Egfr* in adults impaired proliferation before and after PPx through suppression of Cyclin D1 (*59*). However, until now there were no studies connecting SWI/SNF chromatin remodeling and β-cell regeneration.

SWI/SNF chromatin remodeling in pancreas

SWI/SNF chromatin remodelers had not been investigated in great detail in β-cells. Studies using in vitro mouse β-cell lines showed that PDX1, which is an important transcription factor regulating glucose-dependent insulin expression, interacts with BRG1-containing SWI/SNF complexes. Expression of PDX1 target genes including *Ins*, *MafA,* and *Glut2* was reduced upon *Brg1* knockdown, indicating that the interaction between PDX1 and BRG1-containing SWI/SNF complexes is important for the regulation of β-cell specific genes (*60*). Embryonic pancreas specific deletion of *Brg1* reduced multipotent progenitors and reduced all pancreatic cell lineages. In adulthood, only the dual deletion of *Brg1* and *Brm* resulted in impaired β-cell function but single gene deletions did not have an impact (*61*). Although this study linked reduced β-cell function with the inactivation of SWI/SNF subunits, it should be noted that this approach likely abolished BAF complex assembly due to loss of both ATPase subunits. Therefore, further investigations were needed to determine the role of other SWI/SNF subunits in β-cell regeneration and function. Given that *Arid1a* loss increases liver regeneration (*34*) and liver and pancreas arise from the same lineage during embryonic development, whether or not β-cell regeneration could be modulated by *Arid1a* loss required further investigation.

Dissertation Objectives

The main objective of this thesis was to understand the functional roles of SWI/SNF chromatin remodeling complexes in tissue regeneration and disease states. To this end, we generated multiple loss of function mouse models of *Arid1b* and *Arid1a*, genes encoding so called paralogous and mutually exclusive DNA-binding subunits of SWI/SNF chromatin remodeling complexes. We hypothesized that the balance between ARID1A and ARID1B containing complexes might affect physiological and disease states, since the amount of each appears to be regulated in context-dependent manner. Interestingly, haploinsufficient germline mutations in *ARID1B* were strongly implicated in neurodevelopmental disorders including Coffin-Siris Syndrome, autism and intellectual disability in humans. Our first study showed that an *Arid1b* haploinsufficient mouse model exhibited neuropsychiatric abnormalities reminiscent of autism, as well as the neuroanatomical and growth retardation phenotypes seen in CSS. This was the first mouse model shown to recapitulate multiple aspects of *ARID1B*-ophaties and provided mechanistic insights into the role of *ARID1B* in pathology. Another objective of this study was to identify reversible mechanisms that could be immediately translated into clinic. Our study using this model led us to discover IGF-1 and growth hormone deficiencies with a source in the pituitary gland which, at the time, were underappreciated clinical manifestations of human *ARID1B*-ophaties.

In our second study, we explored the question of why the regenerative capacity of pancreatic β-cells decline with age and whether this decline can be reversed. We started with the observation that SWI/SNF components were highly expressed in adult β-cells and particularly the expression of *Arid1a* was reduced under conditions that demand β-cell expansion, such as pregnancy and 50 % PPx. These observations led us to examine whether or not the inhibition of SWI/SNF activity could unlock the proliferative potential of β-cells and provide an advantage in the setting of diabetes. In this study, we showed that *Arid1a* deletion

increased endogenous β-cell regeneration through activation of ERBB and NR4A signaling pathways and protected these mice against STZ-induced diabetes.

CHAPTER 2

METHODOLOGY

METHODOLOGY REGARDING CHAPTER 3

Mice

All animal procedures were based on animal care guidelines approved by the Institutional Animal Care and Use Committee at University of Texas Southwestern Medical Center (UTSW). Constitutive and conditional *Arid1b* knockout mice were generated by the UTSW Transgenic Core using CRISPR/Cas9 genome editing. Guide RNAs were designed to target sequences before and after Exon 5 of *Arid1b*, creating a frame-shift mutation to induce nonsense-mediated decay. Guide RNAs, S. pyogenes *Cas9* mRNA, and oligo donors containing LoxP sequences were injected into single celled zygotes. C57BL/6J mice were used to generate these mice. To generate WT and $Arid1b^{+/-}$ study mice, C57BL/6J WT females were crossed to $Arid1b^{+/-}$ males. $Arid1b^{+/-}$ mice were tail genotyped using the primers

CTTGGTCTTACCCATTTGCACAGT (forward) and GATGGAGGATCCTTACTACAGGGGGATT (reverse). Amplicon size for WT allele is 710 bp and deletion band is 310 bp. *Arid1b* floxed mice were tail genotyped using the primers 5'-CTT GGT CTT ACC CAT TTG CAC AGT-3' (forward) and 5'-AGT GCC TAG GAA GGC AGA GTT TGA GAG-3' (reverse). Amplicon size for WT allele is 475 bp, and for floxed allele 554 bp.

Pituitary and hypothalamus dissection

The whole pituitary and the mediobasal hypothalamus (MBH), which includes both the arcuate nucleus (ARC) and the ventromedial (VMH) hypothalamus, were dissected and subjected to RT-qPCR analysis.

RNA extraction and qRT-PCR

Total RNA was isolated using TRIzol reagent (Invitrogen, catalog no. 15596018). cDNA synthesis was performed with 1 μg of total RNA using the iScript RT Supermix (BioRad, catalog no. 1708840). SYBR Green based quantitative real-time PCR was performed. Gene expression levels were measured using the ΔΔCt method. Mouse *Arid1b* primers were: forward,5'-GTTGGCTCTCCTGTGGGAAGCAA-3';
reverse,5'- GTGACTGGCTCAAGGCAGGAT-3'.

Western blot assay

Tissues were lysed in TPER Tissue Protein Extraction Reagent and homogenized in FastPrep tissue homogenizer. Western blots were performed in standard fashion. Primary antibodies were prepared in 5% BSA in PBS-T. The following primary antibody was used: Arid1b

Antibody (KMN1) X (cat #: sc-32762 X, RRID:AB_2060367). Following secondary antibody was used: anti-mouse IgG, HRP-linked antibody (Cell Signaling, #7076, RRID:AB_330924).

Immunohistochemistry (IHC)

Perinatal mice were decapitated after anesthesia. Heads were fixed in 4% PFA overnight before brains were extracted. Adult mice were anesthetized and underwent intracardial perfusion with ice-cold 0.1 M phosphate-buffered saline (PBS) followed by 4% PFA for fixation. Extracted brains were immersed for 24 hours in 4% PFA in 0.1M PBS at 4°C for post-fixation, followed by least 3 days of immersion in 30% sucrose in 0.1 M PBS with 0.01% sodium azide for cryoprotection. Brains were cut into 40 µm- (adult) or 20 µm-thick (perinatal) sections with a cryostat (model CM3050S; Leica). Brain sections were permeabilized with 0.25% Triton X-100 in 1×PBS and then blocked for 2 h with 5% BSA/3% normal goat serum (NGS) in 0.25% Triton X-100 in 1×PBS. Primary antibodies for TBR1 (1:150 dilution, rabbit, polyclonal, Abcam Cat# ab31940 RRID:AB_2200219) were applied to sections overnight at 4°C. To count cell number, brain sections were incubated with Hoechst 33342 (1 µg/ml; Cell Signaling Technology Cat# 4082S RRID:AB_10626776) alone or together with secondary antibodies (Alexa 488) for 2 h at room temperature.

For the analysis of proliferation, adult mice received one BrdU injection for five consecutive days, and three days following the last injection, brains were fixed and harvested. Slide-mounted IHC for BrdU-, Ki67-, and doublecortin- immunoreactive (+) cells in dentate gyrus was performed as described previously (*62*, *63*). Briefly, every ninth section of the hippocampus was slide-mounted onto charged slides and left for two hours to dry. Antigen retrieval was performed using 0.01 M citric acid (pH 6.0) at 100°C for 15 minutes, followed by washing in PBS at room temperature (RT). Hydrogen peroxide (0.3% H_2O_2) incubation was performed for 30 min to inhibit endogenous peroxidase activity. For BrdU IHC, permeabilization with 0.1% Trypsin in 0.1 M TRIS and 0.1% $CaCl_2$ and denaturation with 2N

HCl in 1x PBS were performed in order to allow antibody access to nuclear DNA. One hour of blocking non-specific binding was performed by incubation in 3% donkey serum, 0.3% Triton-X in PBS. Following these steps, slides were incubated with rat-α-BrdU (1:400; Bio-Rad / AbD Serotec Cat# OBT0030 RRID:AB_609568), rabbit-α-Ki67 antibody (1:500; Thermo Fisher Scientific Cat# RM-9106-S0 RRID:AB_2341197) in 3% serum and 0.3% Tween-20 overnight. Primary antibody incubation was followed by 1 x PBS rinses and a 1 hr incubation with a biotin-tagged secondary antibody targeting the respective primary antibody. Following rinses with 1x PBS, incubation with an avidin-biotin complex occurred for 90 min. Incubation with diaminobenzidine for 5-10 min was used to visualize immunoreactive cells. The counterstain Fast Red was used for nuclear visualization (~3 min incubation). Lastly, slides were placed through an ethanol dehydration series and coverslipped with DPX.

Stereological cell counts

BrdU+ and Ki67+ cells were quantified using an Olympus BX-51 microscope at 40X by an observer blind to experimental groups as previously described (*63*). Immunopositive cells were quantified in every 9th coronal section in the subgranular zone of the granular cell layer in the dentate gyrus, spanning the entire anterior-posterior axis of the hippocampus (-0.82 mm to -4.78 mm from Bregma). Manual stereological counting was performed under bright field, and total cell counts were multiplied by 9 to account for the whole hippocampus. Doublecortin+ cells were quantified using the Optical Fractionator method, and dentate gyrus and corpus callosum volume were assessed using Cavalieri analysis. Stereoinvestigator was used to perform both of these techniques. Investigators were blinded to genotypes during sectioning, counting, and analysis.

RNA-Seq and ChIP-Seq

RNA from four WT and four *Arid1b*$^{+/-}$ hippocampi from 78-82 day old females were purified with a QIAGEN miRNeasy Mini Kit. NuGEN libraries were made with these RNAs. These indexed libraries were multiplexed in a single flow cell lane and received 75 base single-end sequencing on a NextSeq 500 using the High Output Kit v2 (75 cycles) at the CRI Sequencing Facility. Raw sequencing reads were trimmed to remove adaptor and low quality sequences (Phred score<20) using trim_galore package (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Trimmed reads were aligned to mouse reference genome GRCm38/mm10 with HiSAT2 (*64*). After duplicates removal by SAMtools (*65*) and Picard (http://broadinstitute.github.io/picard.), read counts were generated for the annotated genes based GENCODE V20 (*66*) using featureCounts (*67*). Differential gene analysis was performed use edgeR (*68*), using FDR < 0.05 as cutoff. Enriched pathways were analyzed through the use of QIAGEN's Ingenuity® Pathway Analysis (IPA®, QIAGEN RedwoodCity, www.qiagen.com/ingenuity). Heatmaps to visualize the data were generated by using GENE-E (www.broadinstitute.org/cancer/software/GENE-E). RNA-Seq data is deposited to GEO database and can be accessed through GEO accession number (GSE92238). Brg1 ChIP-seq data from the forebrain was downloaded from GEO with the accession number GSM912547 in GSE37151. Files were remapped to the mm10 genome build by CrossMap. Brg1 target genes were predicted by using BETA-minus program on the Cistrome Analysis Pipeline, an integrative platform for transcriptional regulation studies. Heatmap and Metaplot were generated using deeptools, a flexible platform for exploring deep-sequencing data.

Metabolic cage studies

Metabolic cage studies were performed in a temperature-controlled room containing 36 TSE metabolic cages (The TSE Labmaster System of Germany) maintained by UTSW Metabolic Core. Three days prior to study, mice were introduced to metabolic cages and after 3-day acclimation, cages were connected to TSE system and parameters were recorded for a total of 5-days. Investigators were blinded to mouse genotypes.

Grip strength test

Muscle strength was measured by a grip strength test performed by the Neuro-Models Core Facility at UT Southwestern Medical Center in a blinded fashion. Test was conducted using a wire mesh grid connected to a horizontally-aligned force meter (San Diego Instruments, San Diego, CA). The grid was secured at a 45 degree angle, and the top rung of the grid was used for all testing. Mice were held at the base of the tail and supported ventrally while being moved into position to grasp the wire grid. Once the rug was successfully grasped, mice were gently pulled in a horizontal plane until the animal's grip was released from the grid. Peak force (in gram-force units, gf) was captured by the force meter and recorded for later analysis. Forelimb and hindlimb tests were conducted separately, with each being measured 5 times over a 2-3 minute period. Investigators were blinded to mouse genotypes and the identity of treatment groups. *Arid1b*$^{+/-}$ mice with apparent hydrocephaly were excluded. Experiment was performed once.

Ultrasonic vocalization recordings

Ultrasonic vocalizations (USVs) were recorded from both male and female pups isolated from their mothers at P4, during the daylight period of the light/dark cycle. Dams and their litters were acclimated for 30 min in the test room. Each pup was removed from the cage containing its mother and littermates and placed in a clean plastic container in a

wooden sound-attenuating recording chamber. Each pup was first acclimated in the recording chamber for 30 s then recorded for 10 min. Recordings were acquired using an UltraSoundGate CM16/CMPA condenser microphone (Avisoft Bioacoustics) positioned at a fixed height of 8 cm above the pups, and were amplified and digitized (sampled at 16 bits, 250 kHz) using UltraSoundGate 416H 1.1 hardware and Avisoft-RECORDER software (Avisoft Bioacoustics). The data were transferred to Avisoft-SASLab Pro (version 5.2) to analyze spectrograms of vocalizations with settings of 0% overlapping FlatTop windows, 100% frame size, and 256 points fast Fourier transform (FFT) length. The following measures were recorded for each group: number of USV calls, mean duration of USV calls, and mean peak frequency. Recordings were performed with the experimenter blinded to mouse genotypes.

Juvenile social interaction test

The adult male test mouse was placed into a fresh home cage and habituated in the test room with red light for 15 min before testing. A 3 week old male juvenile mouse was placed into the opposite end of the cage that the test mouse was already in. Active interactions between the mice were scored manually with 2 minutes of total test time. Only interactions when the test mouse is interacting with the juvenile, but not other way around were scored. Non-strict male littermates were used. *Arid1b*$^{+/-}$ mice with obvious hydrocephaly were excluded. The experiment was performed twice.

Grooming test

The test was performed between 10:00 am and 2:00 pm. Adult female mouse was singly placed into a new standard cage, without nestlets, food, or water, acclimated for 10 minutes, then videotaped for another 10 minutes. The amount of time spent grooming was recorded

continuously to calculate the total time spent grooming. Grooming is considered self-grooming of any part of the body (including the face, head, ears, full-body). Data is plotted as percent of total time spent grooming.

Marble burying test

The test mouse was acclimated for 30 minutes in the testing room. One standard housing cage for each test mouse was filled with clean bedding material. 15 clean marbles were arranged on top of the bedding in each cage, forming 5 even rows and 3 columns. Mice were placed individually into the prepared cages and kept undisturbed for 30 minutes. After the testing period, they were returned to their original cages. A still image of the test cage was taken to record the number and pattern of buried marbles. A marble was considered buried if more than 2/3 of its depth is covered. Results were calculated and plotted as the percentage of marbles buried per genotype.

Locomotor activity

This test was performed by UTSW Rodent Core Facilty. Experimenters were blinded to mouse genotypes. The experiment was repeated two times and combined data is included. Mice were placed individually into a clean, plastic mouse cage (18 cm x 28 cm) with minimal bedding. Each cage was placed into a dark Plexiglas box. Movement was monitored by 5 photobeams in one dimension (Photobeam Activity System, San Diego Instruments, San Diego, CA) for 2 hours, with the number of beam breaks recorded every 5 min. The movement is characterized in three ways: repetitive beam breaks of a single beam is classified as stereotypy, consecutive beam breaks of two or more beams is classified as ambulatory movements, and total beam breaks during each 5 min interval. Number of total beam breaks during 5 min interval was reported. *Arid1b*$^{+/-}$ mice with apparent hydrocephaly were excluded.

Open field activity test

The test was performed by UTSW Rodent Core Facilty. Experimenters were blinded to mouse genotypes. Experiment was repeated two times and combined data is included. Mice were placed in the periphery of a novel open field environment (44 cm x 44 cm, walls 30 cm high) in a dimly lit room and allowed to explore for 5 min. The animals were monitored from above by a video camera connected to a computer running video tracking software (Ethovision 3.0, Noldus, Leesburg, Virginia) to determine the time, distance moved and number of entries into three areas: the periphery (5 cm from the walls), the center (14 cm x 14cm) and non-periphery (area excluding periphery). The open field arenas were wiped and allowed to dry between mice. Locomotor activity test was performed prior to open field activity test in these cohorts. *Arid1b*$^{+/-}$ mice with apparent hydrocephaly were excluded.

Elevated plus maze

Test was performed by UTSW Rodent Core Facilty. Experimenters were blinded to mouse genotypes. Experiment was repeated two times and combined data is included. Mice were placed in the center of a black Plexiglas elevated plus maze (each arm 30 cm long and 5 cm wide with two opposite arms closed by 25 cm high walls) elevated 31 cm in a dimly lit room and allowed to explore for 5 min. The animals were monitored from above by a video camera connected to a computer running video tracking software (Ethovision 3.0, Noldus, Leesburg, Virginia) to determine time spent in the open and closed arms, time spent in the middle, and the number of entries into the open and closed arm. The apparatus was wiped and allowed to dry between mice. Locomotor activity and open field activity test were performed prior to elevated plus maze in these cohorts. *Arid1b*$^{+/-}$ mice with apparent hydrocephaly were excluded.

Dark-Light activity

Test was performed by UTSW Rodent Core Facilty. Experimenters were blinded to mouse genotypes. Experiment was repeated two times and combined data is included. Mice were placed into a black Plexiglas chamber (25 cm x 26 cm) and allowed to explore for 2 min. After the habituation period, a small door was opened allowing them to access the light side of the apparatus (25 cm x 26 cm lit to approximately 1700 lux) for 10 min. The animals were monitored by 7 photobeams in the dark compartment and 8 photobeams on the light side connected to a computer which recorded the time spent in each compartment, latency to enter the light side and the number of entrances to each compartment (Med-PC IV, Med Associates, St. Albans, VT). The dark-light apparatus was wiped and allowed to dry between mice. Locomotor activity, open field activity, and elevated plus maze tests were performed prior to dark-light activity test in these cohorts. *Arid1b*$^{+/-}$ mice with apparent hydrocephaly were excluded.

Morris water maze

Test was performed by UTSW Rodent Core Facilty. Experimenters were blinded to mouse genotypes. A circular pool is filled with room temperature water to a depth of approximately 12 inches. A platform (10 cm diameter) is placed in one quadrant of the pool with the top of the platform about 2 cm below the water level. White non-toxic paint is added to enhance the contrast with the animal and to hide the location of the platform. Each day the mice are placed in the pool and allowed to swim for 1 min to find the platform. The swim path and time until locating the platform is recorded via a videocamera and computer running videotracking software (Ethovision, Noldus). If the mouse does not find the platform within a minute, they are gently guided or placed on the platform for 10 sec, then removed from the pool and return to their home cage. Each animal is placed in the pool for a total of 4 times each day for 13 days.

Twenty-four hours after the last training day, a probe test is conducted in which the platform is removed from the pool and each mouse is allowed to swim for 1 minute to determine whether the animal has learned the location of the platform. The time each animal spends in the quadrant which had contained the platform on training days and the number of times that the animal crosses the location which had contained the platform are indicators of how well the animal has learned the spatial location of the platform. To control for visual problems, the mice are given 4-6 trials after the probe test using the same pool and platform, however a large black block is placed on top of the platform to clearly mark the location. The location of the platform is moved on each trial.

Fear Conditioning

Test was performed by UTSW Rodent Core Facilty. Experimenters were blinded to mouse genotypes. Fear conditioning was measured in boxes equipped with a metal grid floor connected to a scrambled shock generator (Med Associates Inc., St. Albans). For training, mice were individually placed in the chamber. After 2 min, the mice received 3 tone-shock pairings (30 sec white noise, 80 dB tone co-terminated with a 2 sec, 0.5 mA footshock, 1 min intertrial interval). The following day, memory of the context was measured by placing the mice into the same chambers and freezing was measured automatically by the Med Associates software. Forty-eight hours after training, memory for the white noise cue was measured by placing the mice in a box with altered floors and walls, different lighting, and a vanilla smell. Freezing was measured for 3 min, then the noise cue was turned on for an additional 3 min and freezing was measured.

Footshock Sensitivity

Test was performed by the UTSW Rodent Core Facility. Experimenters were blinded to mouse genotypes. The mice were placed individually into boxes equipped with a metal grid floor connected to a scrambled shock generator (Med Associates Inc., St. Albans). After approximately 1 min, the mice received a series of footshocks (2 sec each) with increasing intensity. The initial shock intensity was 0.05 mA and the amplitude was increased by 0.05 mA for each consecutive footshock with 15 sec intershock interval. The first shock intensity that each animal displayed each behaviour (flinch, jump, and vocalization) was recorded. Once the animal displayed all three behaviours, it was removed from the chamber.

GH stimulation test and ELISA experiments

Plasma IGF-1 concentration was determined using mouse/rat IGF-1 Quantikine ELISA Kit (R&D Biosystems, Cat #: MG100) without fasting. Mice were fasted for 36 hours before GH stimulation testing. Fifteen minutes after anesthesia with pentobarbital (50 mg/kg given once i.p.), 0.14 g/kg GHRH (Phoenix Pharmaceuticals, Cat #031-02) was injected i.p. Blood was sampled retro-orbitally using a capillary tube before, 5, and 15 min after injection. Plasma GH concentration was measured using a Rat/Mouse GH ELISA KIT (Millipore, Cat #: EZRMGH-45K).

Recombinant human IGF1 (rhIGF1) therapy

See Figure 5C. Mice at the age of P10 were ranked from highest to lowest body weight and even numbered mice were placed into the treatment group and odd numbered mice were placed into the vehicle group. *Arid1b*$^{+/-}$ mice with apparent hydrocephaly were excluded. Recombinant human IGF1 was purchased from Peprotech (Catalog #100-11, SCR_015509). It

was prepared according to datasheet and injected intraperitoneally. Vehicle or 0.5 mg/kg rhIGF1 dissolved in vehicle were injected starting from P11 daily.

Recombinant mouse growth hormone (rmGH) therapy

Mice at the age of P11 were ranked from highest to lowest body weight and even numbered mice were placed into the treatment group and odd numbered mice were placed into the vehicle group. *Arid1b*$^{+/-}$ mice with apparent hydrocephaly were excluded. Recombinant mouse growth hormone was obtained from National Hormone and Peptide Program (NHPP) (SCR_015508). It was prepared according to the NHPP datasheet. Injection was performed subcutaneously. 30 ug GH/mouse/day was injected between P11 to P14, 50 ug GH/mouse/day was injected between P14 to P21, 70 ug GH/mouse/day was injected between P21 to P50. Grip strength test was performed at P60 after 10 days without treatment.

Statistical analyses

Unless specified otherwise in the figure legends, statistical analyses were performed using unpaired, two-tailed, Student's t-test or ANOVA. The data bars and error bars indicate mean ± standard error mean (SEM). *, p value ≤ 0.05; **, p value ≤ 0.01; *** p value ≤ 0.001; ****, p value ≤ 0.0001; ns, not significant. No statistical methods were used to predetermine sample sizes; however, sample sizes were estimated based on similar experiments reported in the relevant literature in the field (*69*, *70*).

METHODOLOGY REGARDING CHAPTER 4

Mice

All experiments on mice were approved by and handled in accordance with the guidelines of the Institutional Animal Care and Use Committee at UTSW. All experiments were performed in in age and gender controlled fashion. Male mice were used for STZ experiments. The transgenic mouse lines used are as follows and described before: *Arid1a* floxed mice (JAX stock #027717) (*29*), *Ptf1a-Cre* (*71*), *MIP-rtTA* (*72*), *TRE-Cre* (JAX 006234) (*73*). *TRE-HER3* mouse line was generated by Jiyoung Park and Alexandra Ghaben in Dr. Scherer's group and a paper with the detailed characterization of this mouse model is in preparation.

Cell Lines.

The H2.35 (ATCC® CRL-1995™) and BTC6 cells were obtained from ATCC (ATCC® CRL-11506™) and cultured according to manufacturer's protocol. MIN6 cell line was obtained from Dr. Melanie Cobb's lab and cultured in DMEM with 15% Heat Inactivated Fetal Bovine Serum, 1% L-Glutamine, 1% Pen/Strep , 0.0005% beta mercaptoethanol.

Mouse Islet Culture.

Islets from adult mice were isolated and recovered overnight in culture medium (RPMI1640 with 10% heat-inactivated FBS, L-glutamine and Pen/Strep) in the incubator and used for experiments.

Mouse pancreatic islet isolation and dispersion of islets.

Islet isolation was done as described previously (*74*) with minor modifications. Briefly, Liberase TL (Roche, 05401020001) solution was prepared by dissolving 5 mg lyophilized Liberase TL powder in 1 ml sterile water so that the concentration is 5 mg/ml corresponding to 26 Wunsch units. Prior to use, this 1 ml Liberase solution was added to 21.6 ml serum free RPMI to obtain the working solution. Adult mice were perfused with 3 ml Liberase TL working solution through the common bile duct cannulation and inflated pancreas was put into 2 ml Liberase TL solution sitting on ice in 50 ml falcon tube and incubated at 37 °C water bath for 10-16 minutes by shaking and reaction was stopped with the addition of RPMI containing serum. Islets were dissociated from the exocrine tissues by shaking vigorously several times followed by Histopaque 1077 (Sigma-Aldrich, 10771) density centrifugation (900g, 20 minutes, acceleration:2, deceleration: 0). Purified islets were collected from the interface and washed. Intact islets were hand-picked under the dissection microscope.

Islet mitogen experiments

Islets from male CD1 mice were isolated and recovered overnight in culture medium (RPMI1640 with 10% heat-inactivated FBS, L-glutamine and Pen/Strep) in the incubator prior to various treatments. After recovering, islets were treated with either vehicle (culture medium with 11.1 mM glucose), high glucose (culture medium with 25 mM glucose), or with the addition of IGF2 (200nM), GLP-1 (100nM), IL1β (1 ng/ml) for 6 hours and RNA was harvested for qPCR analysis.

Kinase inhibitor screen

SelleckChem customized library (Z49127) was a collection of 304 kinases inhibitors. See the supplementary table for a full list of inhibitors with additional details about their targets. Isogenic WT and Arid1a KO H2.35 transformed mouse liver hepatocyte cells were used in this kinase screen and treated with kinase inhibitors. To determine if EGFR family inhibitors is among the treatment conditions that induced the highest decrease in cell viability, we performed a hypergeometric test to examine the enrichment of treatments involving these inhibitors. Briefly, we ranked all treatment conditions in descending order according to observed cell viability loss, then, we counted the number of EGFR family inhibitor treatments among the top 10% conditions. The possibility of observing at least N EGFR family inhibitor treatments by chance is calculated and reported as the p-value of EGFR enrichment.

Generation of CAS9 single cell clones

Mouse *Arid1a* gRNA (GCTGCTGCTGATACGAAGGTTGG) was cloned into LentiGuide-puro plasmid (Addgene #53963). LentiCas9-Blast plasmid was purchased from Addgene (Addgene #53962). Active lentivirus was prepared in 293T cells in 10-cm dish. The day after seeding cells, each dish was transfected with pVSV-G, pLenti-gag pol, LentiCas9-puro-sG*Arid1a* or LentiCas9-Blast plasmid by Lipofectamine 3000 (Life Technologies # L3000015). Virus containing medium was collected at 60h after transfection. For creating the Cas9 stable expressing cell line, cells were infected with Cas9 lentivirus, followed by blasticidin selection at 2ug/ml for 4 days. Then Cas9 expressing cells were infected with Arid1a gRNA lentivirus. Three days after the infection, we selected cells for 3 days with puromycin at 2ug/ml. Then, 50 to 200 cells were plated in 15cm dishes for single clone selection. Single clones were picked when they grew big enough and verified the genotype and Arid1a expression by PCR and Western blot.

In vivo drugs

Dox water (1g/L) was used to induce conditional deletion. Tamoxifen was dissolved (Sigma-Aldrich, T5648) in corn oil at a concentration of 20 mg/ml by sonication. 500uL of 20mg/mL Tamoxifen by oral gavage for two consecutive days. 20 mg/kg canertinib (LC Laboratories NC0704940) was administered daily through oral gavage for one month in STZ experiments. Either 20 mg/kg canertinib or 20 mg/kg NR4A1 antagonist C-DIM8-DIM-C-pPhOH (Axon Medchem # Axon 2827) was administered daily through gavage starting a day before PPx until sacrificing mice day 7 post-PPx.

Streptozotocin (STZ) injury

Mice were fasted for 4-6 hours prior to STZ injection. STZ (Sigma-Aldrich S0130) was dissolved in sodium-citrate (Sigma-Aldrich S4641) solution to a final concentration of 10 mg/ml freshly 10-15 min before the injection. Sodium-citrate solution was prepared by dissolving 1.47 gram of sodium-citrate powder in 50 ml ddH2O, and adjusting the pH to 4.5. Prepared STZ solution was injected via intraperitoneal injection. Different dosage was used for different strain backgrounds as indicated in the figure legends since response to STZ is strain dependent.

Glucose tolerance test (GTT)

After 16 hours fast, blood glucose was measured using a glucometer from tail tip blood before and multiple times within 2 hours of intraperitoneal injection of 2g/kg D-(+)-Glucose (Sigma Aldrich # G7528).

Insulin tolerance test (ITT).

After 6h fast, blood glucose was measured before intraperitoneal injection of insulin (0.75 mU/g body wt) and then 30, 60, 90 and 120 min after injection.

Partial pancreatectomy (PPx)

PPx was performed as described (*75*) except that mice were not fasted overnight.

RNA Extraction and RT-qPCR

Total RNA from isolated mouse islets were extracted using TRIzol reagent (Invitrogen) or The RNeasy Plus Micro Kit (cat. no. 74034). For RT-qPCR of mRNAs, cDNA synthesis was performed with 1 mg of total RNA using miScript II Reverse Transcription Kit (QIAGEN). See Supplemental Information for primers used in these experiments. Expression was measured using the ΔΔCt method.

Western blot

Isolated islets were lysed in RIPA buffer (Sigma R0278) supplemented with protease and phosphatase inhibitors. Protein concentration was determined by Pierce™ BCA Protein Assay Kit (Thermo Fisher #23225). Western blots were performed in the standard fashion. The following antibodies were used: ARID1A (Sigma-Aldrich Cat# HPA005456, RRID:AB_1078205), ARID1A (Santa Cruz Biotechnology Cat# sc-32761,

RRID:AB_673396), p-NR4A1 (Cell Signaling Technology Cat# 5095, RRID:AB_10695108), phospho-EGFR (Tyr1068) (Cell Signaling Technology Cat# 3777, RRID:AB_2096270), phospho-EGFR (Y1173) (Cell Signaling Technology Cat# 4407, RRID:AB_331795), phospho-AKT (Ser473) (Cell Signaling Technology Cat# 4060, RRID:AB_2315049), phospho-p44/42 (p-ERK) (Cell Signaling Technology Cat# 9101, RRID:AB_331646), p44/42 (ERK) (Cell Signaling Technology Cat# 9107, RRID:AB_10695739), c-FOS (Santa Cruz Biotechnology Cat# sc-166940, RRID:AB_10609634), β-Actin (Cell Signaling Technology Cat# 4970, RRID:AB_2223172), Vinculin (Cell Signaling Technology Cat# 13901, RRID:AB_2728768), anti-rabbit IgG, HRP-linked antibody (Cell Signaling Technology Cat# 7074, RRID:AB_2099233) and anti-mouse IgG, HRP-linked antibody (Cell Signaling Technology Cat# 7076, RRID:AB_330924).

Histology, immunohistochemistry, and immunofluorescence

Tissue samples were fixed in 4% paraformaldehyde (PFA) and embedded in paraffin. H&E staining was performed by the UTSW Histology Core Facility. Primary antibodies used: anti-rabbit Arid1a (Sigma-Aldrich Cat# HPA005456, RRID:AB_1078205, used for IHC), anti-rabbit Glucagon (Cell Signaling Technology Cat# 2760, RRID:AB_659831, used for IHC), anti-rabbit Insulin (Abcam Cat# ab108326, RRID:AB_10861152, used for IHC), anti-rabbit Ki-67 (Abcam Cat# ab15580, RRID:AB_443209, used for IF), anti-guinea pig Insulin (Abcam Cat# ab7842, RRID:AB_306130, used for IF). Secondary antibodies used: Alexa Fluor 488 goat anti-rabbit IgG (Thermo Fisher; 11008; RRID:AB_10563748) goat polyclonal antibody to anti-Guinea pig IgG-Alexa 568 (Abcam, ab175714). For IHC, detection was performed with the Elite ABC Kit and DAB Substrate (Vector Laboratories) followed by hematoxylin counterstaining (Sigma). VECTASHIELD® Antifade Mounting Media with DAPI

counterstain (Vector Labs, H-1200) was used. For islet number and area calculation, H&E sections were imaged using The NanoZoomer 2.0-HT whole slide imaging. Number of islets were counted in each slide. Whole H&E section area and area of each individual islet in that H&E section were measured using NDP.view software. To determine the proliferation, slices of pancreas were costained with insulin, Ki-67 and DAPI and all the islets present in the section were imaged using the same parameters. Channels were merged and images were analyzed in Image J.

RNA-sequencing

RNA was extracted from islets isolated from 2 *Arid1a*$^{Fl/Fl}$ *and 2 Ubiquitin-CreER; Arid1a*$^{Fl/Fl}$ mice. RNA-seq libraries were prepared with the Ovation RNA-Seq Systems 1-16 (Nugen) and indexed libraries were multiplexed in a single flow cell and underwent 75 base pair single-end sequencing on an Illumina NextSeq500 using the High Output kit v2 (75 cycles) at the UTSW Children's Research Institute Sequencing Facility.

Quantification and statistical analysis

RNA-Seq Analysis. RNA-Seq analysis was performed as described before (*30*). Briefly, adaptors and low quality sequences (Phred<20) were removed by trimming raw sequencing reads using galore package (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). The sequence reads were aligned to the GRCm38/mm10 with HiSAT2 (*64*). After duplicates removal by SAMtools (*65*), read counts were generated for the annotated genes based on GENCODE V20, using featureCounts (*66*, *67*). Differential gene expression analysis was

performed on edgeR, using FDR < 0.05 as cutoff (*68*, *76*). Heatmaps to visualize the data were generated by using GENE-E (*68*).

ATAC-Seq

ATAC-Seq libraries were constructed using islets that were freshly isolated from adult WT and Arid1a βKO mice under steady state conditions or 5 days post-PPx. Islet cells were lysed and 4 x10^4 nuclei was counted using a cell counter and used for the transposition reaction using TD and TDE1. ATAC-Seq protocol published by Buenrostro et al. was followed with the exception of a size selection step using Ampure XP beads before the final elution step (*77*). Briefly, transposed DNA fragments were purified using a MinElute Kit (Qiagen) and PCR-amplified using PCR primer1 and a barcoded PCR primer 2 (Nextera DNA Sample Preparation Kits - Index Kit). Thermal cycle conditions: 72 °C for 5 min, 98°C for 30 s, followed by 5 cycles of 98°C for 10 s, 63°C for 30 s and 72°C for 60 s. qPCR was performed using 5 ul of amplified DNA to determine the additional number of cycles needed for each sample. Remaining 45 ul was amplified again using determined number of cycles. Thermal cycle conditions: 1 cycles of 98°C for 30 s, N cycles of 98°C for 10 s, 63°C for 30 s, 72°C for 60 s where N is determined using qPCR. Amplified libraries were purified using MinElute Kit and library quality was assessed using the TapeStation system (Agilent). Libraries were sequenced by 75 bp single end reads using the NextSeq500 system at the Children's Research Institute at UTSW Sequencing Facility

ATAC-seq Data Processing and Analysis.

Sequences were trimmed by Cutadapt (*78*) to remove Nextera transposase adaptor. The trimmed sequences were aligned with BWA to mouse mm10. Peaks were called using MACS14 (*79*) with default parameters. Differentially bound peaks were called using DiffBind (*80*, *81*). Peaks were annotated by CHIPseeker (*82*).The peak heatmaps and average profile of peak were visualized by CHIPseeker as well.

GSEA

GSEA was used to determine significantly enriched gene sets. To perform GSEA analysis using RNA-seq data from WT and *Arid1a* KO islets, raw read counts from each sample were converted to cpm value (count per million) using the cpm function within edgeR R package. GSEA analysis was performed with a pre-ranked gene list by log fold change. GSEA was then performed against hallmark gene sets using default parameters (http://software.broadinstitute.org/gsea/index.jsp).

ChIP-seq Data Processing and Analysis.

Sequencing of ChIP-seq samples was carried out using the Illumina NextSeq technology, and reads were demultiplexed with the bcl2fastq software tool. Read quality was evaluated by FASTQC, and read alignment to the hg19 genome was executed with Bowtie2 v2.29 in the -k 1 reporting mode (*83*). Narrow peaks were detected using MACS2 v2.1.1 with a q-value cutoff of 0.001 and input as controls (*79*). Output BAM files were transformed into BigWig track files using the "callpeak" function of MACS2 v2.1.1 with the "-B --SPMR" option followed by the use of the BEDTools (*84*) "sort" function and the UCSC utility

"bedgraphToBigWig". BigWig track files were then input in IGV v2.4.3 for visualization. Heat maps were generated using ngsplot v2.61, which was also used to perform K-means clustering (*85*). Cis-regulatory function analysis was carried out using the GREAT online software suite (*86*).

Statistical analysis

Variation is indicated using standard error presented as mean ± SEM. Two-tailed Student's *t*-tests (two-sample equal variance) were used to test the significance of differences between the two groups. Statistical significance is displayed as * ($P < 0.05$), ** ($P < 0.01$), *** ($P < 0.001$), ****($P < 0.0001$). Statistical analyses were performed using GraphPad Prism unless otherwise indicated. Mice from multiple litters were used in the experiments. In STZ follow up experiments, mice were occasionally excluded from the analysis due to death.

CHAPTER 3

ARID1B HAPLOINSUFFICIENT MICE REVEAL NEUROPSYCHIATRIC PHENOTYPES AND REVERSIBLE CAUSES OF GROWTH IMPAIRMENT

INTRODUCTION

It is becoming clear that SWI/SNF chromatin-remodeling complexes have a major impact on human diseases, from cancer to neuropsychiatric disorders to body size regulation (*7*, *22*). SWI/SNF chromatin-remodeling complexes use the energy of ATP to remodel nucleosome density and position to control epigenetic states, lineage differentiation, and cellular growth during development and cancer (*7*, *87*). *ARID1B* is a 236 kDa protein that contains an AT-rich DNA interactive domain ("ARID" domain) and facilitates proper genomic

targeting of ARID1B containing SWI/SNF complexes. *ARID1B* is the most commonly mutated gene in Coffin-Siris Syndrome (CSS), a monogenic syndrome characterized by growth retardation, facial dysmorphism, and intellectual disability (ID) (*15*, *16*, *88*). In addition, *ARID1B* is among the most frequently mutated genes in autism spectrum disorders (ASD) and non-syndromic ID (*20*, *21*). In these diseases, *ARID1B* mutations are scattered across the gene without clear accumulation in particular domains (*22*, *89*). Since these are often predicted to be nonsense or frameshift mutations, heterozygous *ARID1B* loss-of-function is hypothesized to be the causative genetic mechanism. Up to this point, if and how *ARID1B* mutations translate into various human phenotypes is unknown, and animal models have not yet been used to model or devise novel treatments for these "*ARID1B*-opathies".

Here, we employ genetically engineered mouse models to elucidate the phenotypic impact of *Arid1b* mutations. We developed an *Arid1b* haploinsufficient mouse that exhibits neuropsychiatric abnormalities reminiscent of ASD, as well as the developmental and growth retardation phenotypes seen in CSS. Although not previously considered a cardinal feature of this syndrome, a meta-analysis of 60 patients by Santen et al. shows that on average, stature is considerably shortened in CSS patients by about two standard deviations (*89*). After showing the clinical relevance of our mouse model, we focused on potentially reversible etiologies of behavioral and growth phenotypes. We observed GHRH-GH-IGF1 axis deficiencies in *Arid1b* heterozygous mice and also found evidence for this in humans. GH supplementation in mice rescued growth retardation and muscle weakness, which are also salient features of human *ARID1B*-opathies. Though successful in *Mecp2* mutant mice that model Rett syndrome (*90*, *91*), intervening on the GH-IGF1 axis was not able to reverse neuropsychiatric defects associated with *Arid1b*. Our findings not only functionally validate *ARID1B's* involvement in human disease, they suggest underappreciated clinical manifestations of human *ARID1B* mutations that can be approached from a treatment-perspective.

RESULTS

Characterization of Arid1b$^{+/-}$ and Arid1b$^{-/-}$ mice

Using Cas9 germline gene-editing, we generated whole-body knockout and conditional floxed mice (Figure 1A). *Arid1b* is prominently expressed in the cortex, cerebellum, and hippocampus (*92*, *93*) (Figure 2A and ref. 15). In heterozygous mice, *Arid1b* mRNA transcripts were reduced in liver, whole brain, pituitary gland, dentate gyrus, and hypothalamus (Figure 1B). Protein levels showed a similar pattern in whole brain extracts, and homozygous P0 pups showed an absence of ARID1B protein (Figure 1C). Homozygous mice were born but died perinatally (Figure 1D). To model the genetics of haploinsufficient human ARID1B-opathies, we generated whole-body heterozygous (*Arid1b$^{+/}$)* mice [birth ratios from *Arid1b$^{+/-}$* x *Arid1b$^{+/+}$* crosses: 389/661 (58.9%) WT, 272/661 (41.1%) *Arid1b$^{+/-}$*], which survived into adulthood and appeared healthy but were small for age (Figure 1E). There were no abnormalities in electrolytes, liver function tests, or blood counts (Figure 2B-D). 16/272 (6.6%) of *Arid1b$^{+/-}$* mice had hydrocephalus, the displacement of brain parenchyma by accumulated cerebrospinal fluid, a condition that frequently accompanies Dandy-Walker malformations seen in CSS patients (*94*) (Figure 2E).

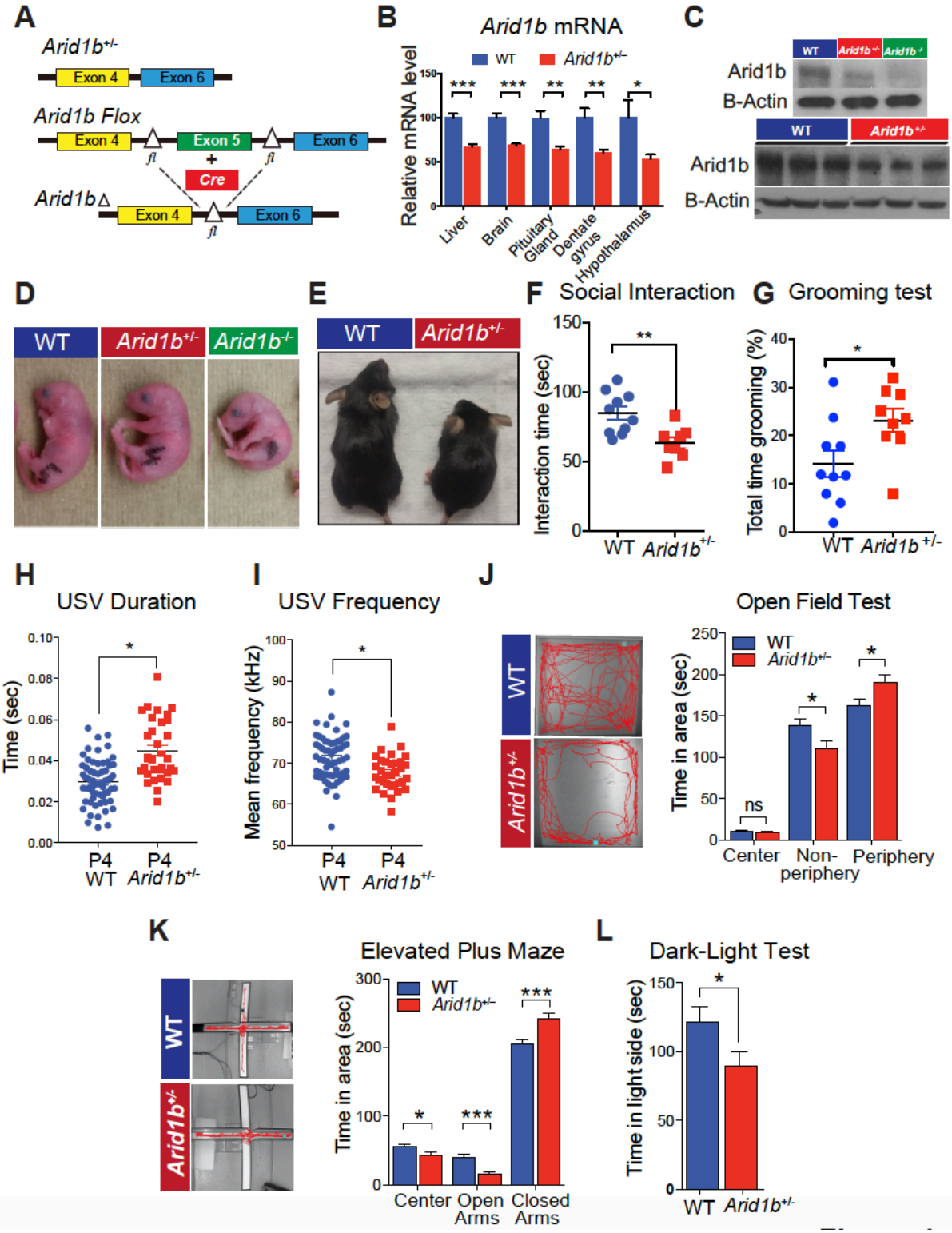

A
Arid1b+/-
Exon 4
Exon 6
Arid1b Flox
Exon 4
Exon 5
Exon 6
fl
fl
Cre
Arid1bΔ
Exon 4
Exon 6
fl
B
Arid1b mRNA
WT
Arid1b+/-
Relative mRNA level
150
100
50
0
***
***
**
**
*
Liver
Brain
Pituitary Gland
Dentate gyrus
Hypothalamus
C
WT
Arid1b+/-
Arid1b-/-
Arid1b
B-Actin
WT
Arid1b+/-
Arid1b
B-Actin
D
WT
Arid1b+/-
Arid1b-/-
E
WT
Arid1b+/-
F Social Interaction
150
100
50
0
**
Interaction time (sec)
WT
Arid1b+/-
G Grooming test
40
30
20
10
0
*
Total time grooming (%)
WT
Arid1b+/-
H
USV Duration
0.10
0.08
0.06
0.04
0.02
0.00
*
Time (sec)
P4 WT
P4 Arid1b+/-
I
USV Frequency
100
90
80
70
60
50
*
Mean frequency (kHz)
P4 WT
P4 Arid1b+/-
J
WT
Arid1b+/-
Open Field Test
250
200
150
100
50
0
WT
Arid1b+/-
*
*
*
ns
Time in area (sec)
Center
Non-periphery
Periphery
K
WT
Arid1b+/-
Elevated Plus Maze
300
200
100
0
WT
Arid1b+/-
*
***
***
Time in area (sec)
Center
Open Arms
Closed Arms
L
Dark-Light Test
150
100
50
0
*
Time in light side (sec)
WT
Arid1b+/-

Figure 1. Arid1b+/- mice exhibit physical manifestations of developmental delay, autistic-like features, and abnormal behavioral phenotypes.

(A) Schematic of *Arid1b* whole body heterozygous mice in which exon 5 is deleted (hereafter referred as *Arid1b*$^{+/-}$) and *Arid1b* floxed mice. B) Relative *Arid1b* mRNA levels in selected organs and brain regions as assessed by qPCR. (C) Relative Arid1b levels in p0 mouse limb (top panel) and whole brain extracts at p45 as assessed by western blot analysis. (D) Appearance of WT and *Arid1b*$^{+/-}$ mice at postnatal day 0. (E) Appearance of WT and *Arid1b*$^{+/-}$ littermates at 1 month of age. (F) Juvenile social interaction testing for 10 WT and 9 *Arid1b*$^{+/-}$ male mice. (G) Grooming test for 10 WT and 9 *Arid1b*$^{+/-}$ female mice. (H) The ultrasonic vocalization (USV) test measuring the duration of vocal communication in 63 WT and 33 *Arid1b*$^{+/-}$ male and female mice during separation of pups from dams at postnatal day 4. (I) The ultrasonic vocalization (USV) test measuring the frequency of vocal communication in 63 WT and 33 *Arid1b*$^{+/-}$ male and female mice during separation of pups from dams at postnatal day 4. (J) Representative traces of WT and *Arid1b*$^{+/-}$ mice in the open field and time spent in the indicated areas for 20 WT and 20 *Arid1b*$^{+/-}$ 8 week old male mice. (K) Representative traces of WT and *Arid1b*$^{+/-}$ mice in the elevated plus maze and time spent in the indicated areas for for 20 WT and 20 *Arid1b*$^{+/-}$ 8 week old male mice. (L) Dark-light box testing for 20 WT and 20 *Arid1b*$^{+/-}$ 8 week old male mice. Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes, *, p value ≤ 0.05; **, p value ≤ 0.01; *** p value ≤ 0.001; **** p value ≤ 0.0001; ns, not significant. Student's t-test (two-tailed distribution, two-sample unequal variance) was used to calculate p-values unless otherwise indicated in the corresponding figure legend.

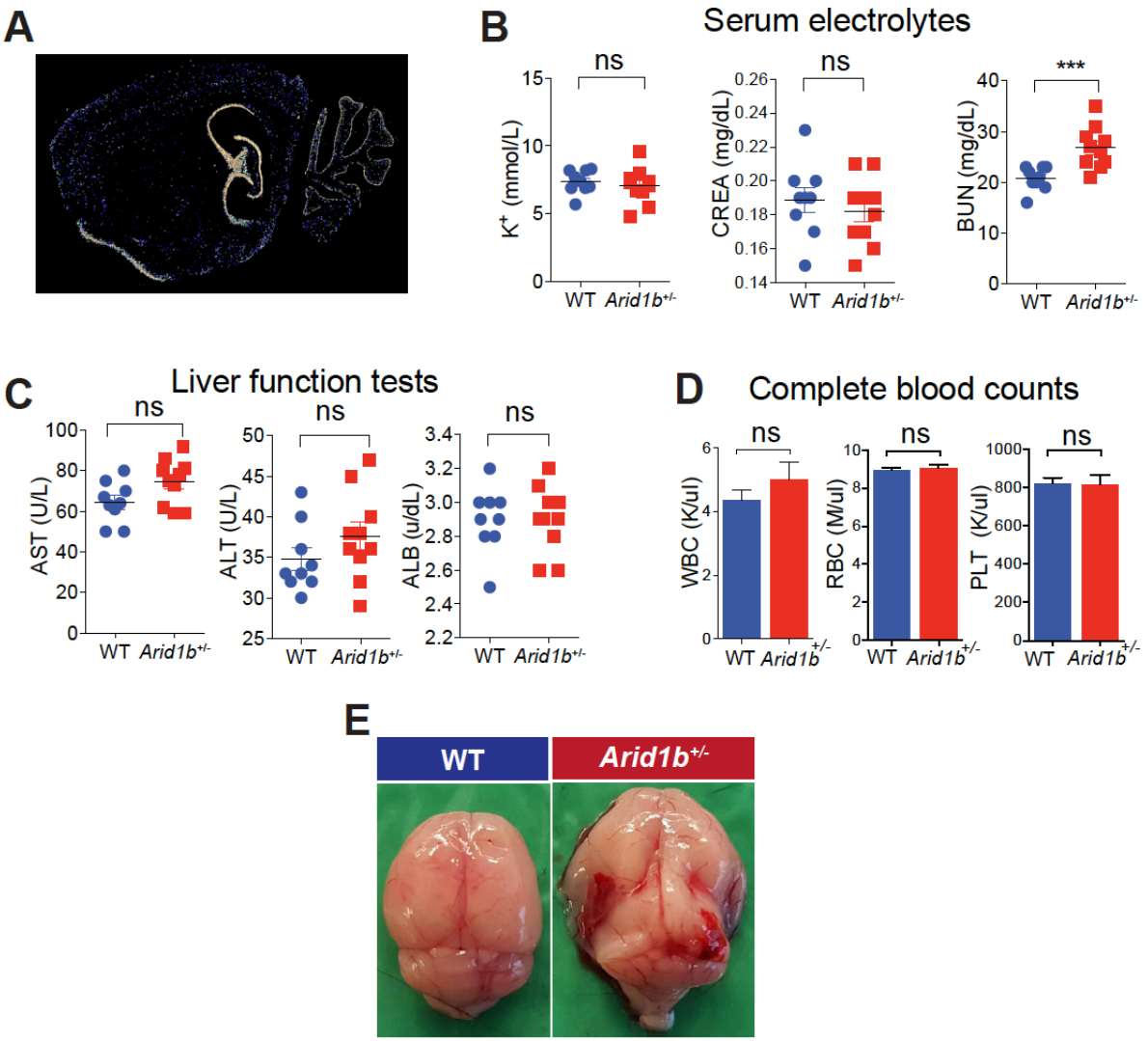


Figure 2. Additional characterization of *Arid1b*$^{+/-}$ mice.

(A) *Arid1b* in situ hybridization showing a sagittal brain section from an 8-week old male mouse. Hippocampus and cerebellum showed high levels of signal. Image taken from the Allen Brain Atlas. (B) Serum electrolytes including potassium (K^+), creatinine (CREA), and blood urea nitrogen (BUN) measured in 100 day-old females (n = 9 WT and 9 *Arid1b*$^{+/-}$). (C) Liver function tests including Aspartate transaminase (AST), Alanine amino transferase (ALT), and Albumin (ALB) in 100 day old females (n = 9 WT and 10 *Arid1b*$^{+/-}$). (D) Blood cell counts including white blood cells (WBC), red blood cells (RBC), and platelets (PLT) measured in

100 day old females (n = 9 WT and 10 *Arid1b*$^{+/-}$). (E) 0/389 WT and 16/272 (6.6%) *Arid1b*$^{+/-}$ mice had grossly appreciable hydrocephalus. Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes: *, p value ≤ 0.05; **, p value ≤ 0.01; *** p value ≤ 0.001; **** p value ≤ 0.0001; ns, not significant. Student's t-test (two-tailed distribution, two-sample unequal variance) was used to calculate p-values unless otherwise indicated in the corresponding figure legend.

Arid1b*$^{+/-}$ *mice developed abnormal social, vocal, and behavioral phenotypes

Given the associations between *Arid1b* mutations and ASD, we examined behaviors related to this disorder. To examine social interactions, we quantified the time spent interacting with a juvenile target mouse. Compared to WT littermate controls, *Arid1b*$^{+/-}$ mice spent significantly less time interacting with unfamiliar juvenile mice (Figure 1F), suggesting impaired social behavior. To enrich the connections between *Arid1b*$^{+/-}$ mice and ASD-like phenotypes, we also performed grooming and marble burying tests that examined repetitive behaviors (*95*). Consistent with other ASD mouse models, *Arid1b*$^{+/-}$ mice exhibited increased self-grooming (Figure 1G) and potentially as a consequence, buried less marbles (Figure 3A). A similar pattern of repetitive behaviors was seen with *Synapsin* knockout mice, another mouse model of ASD (*96*).

Another feature of ASD is abnormal communication and language. Several mouse models of ASD and language disorders show alterations in one or more vocalization parameters, including the number, duration, frequency, amplitude, and other characteristics of ultrasonic vocalizations (USVs) (*69*, *97*). Furthermore, ASD patients who have retained speech tend to exhibit abnormalities in voice quality and pitch (*98*, *99*). USVs emitted by *Arid1b*$^{+/-}$ mice are longer in duration, and have abnormal pitch (Figure 1H, I and Figure 3B).

Interestingly, *Arid1b*$^{+/-}$ mice emitted the same total number of USVs compared to WT mice (Figure 3C), suggesting altered modulation rather than absent vocalizations.

Anxiety-like behavior, a comorbidity of ASD, was examined using three separate tasks in male mice. In the open field test, *Arid1b*$^{+/-}$ mice spent significantly more time in the periphery while avoiding the anxiety-provoking center (Figure 1J). In the elevated plus maze, *Arid1b*$^{+/-}$ mice spent more time in the anxiety-relieving, walled arms of the maze (Figure 1K). In the dark-light box test, *Arid1b*$^{+/-}$ mice avoided exploring the brightly lit chamber (Figure 1L). WT and mutant mice traveled equal distances both initially and over a 2 hour time period, making locomotor differences less likely a confounder in simple environments (Figure 3D). These tests consistently demonstrated higher levels of anxiety-like behavior in *Arid1b*$^{+/-}$ mice compared to their WT littermates.

Given the associations between *Arid1b* haploinsufficiency and intellectual disability, we assessed cognitive functions in *Arid1b*$^{+/-}$ mice. The Morris water maze test, a contextual fear-conditioning test, and a cued fear-conditioning test each did not reveal defects in memory and learning (Figure 3E-G). The genotypes were equally able to sense the electric shock applied during fear conditioning (Figure 3H). Overall, these tests showed that *Arid1b*$^{+/-}$ mice displayed abnormal social, vocal, and behavioral phenotypes, but did not clearly have cognitive or memory deficiencies.

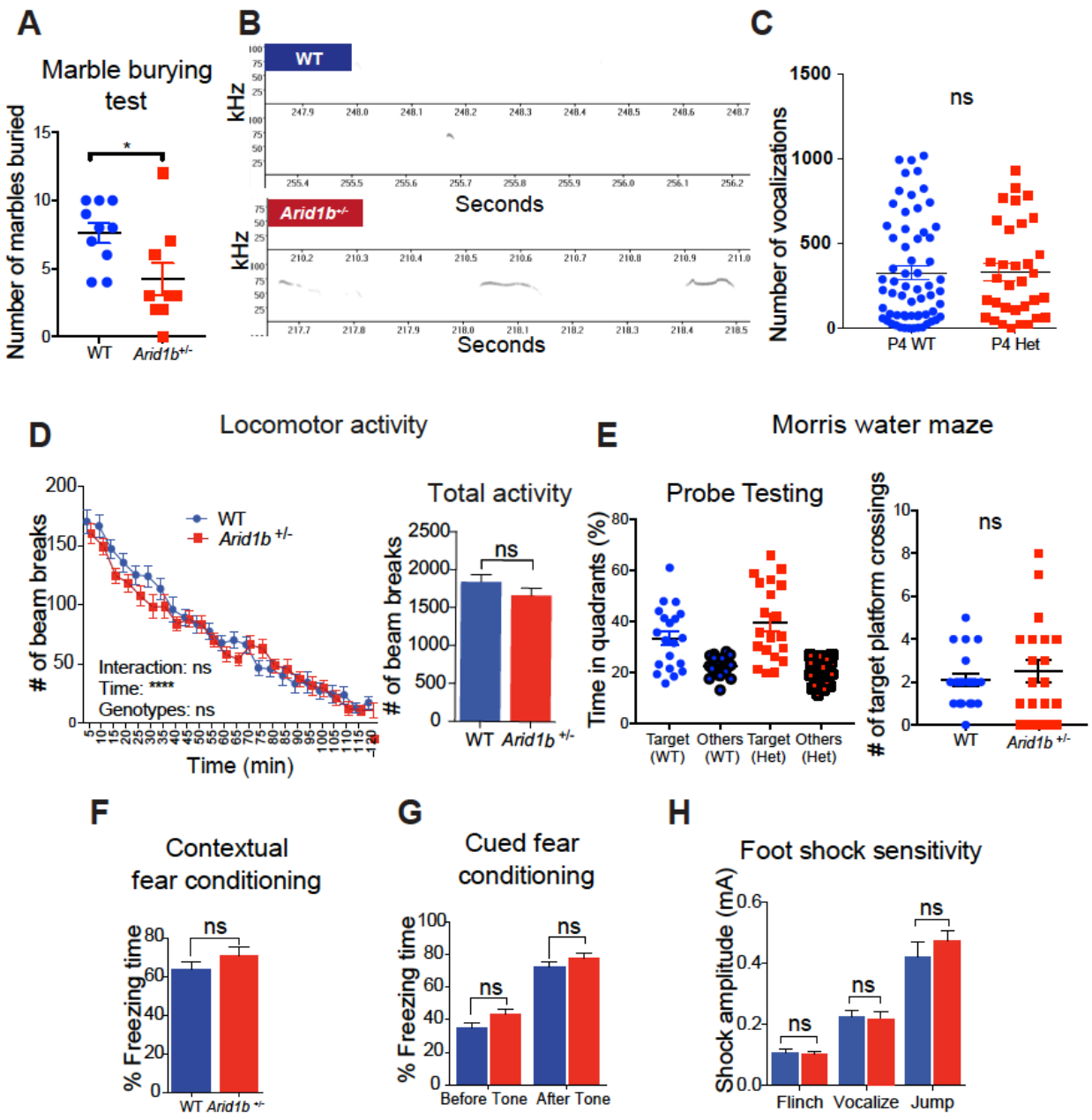


Figure 3. Additional neurobehavioral testing on *Arid1b*$^{+/-}$ mice.

(A) Marble burying test for 10 WT and 9 *Arid1b*$^{+/-}$ female mice. (B) Representative spectrograms of WT and *Arid1b*$^{+/-}$ vocalizations. (C) Number of vocalization calls of 63 WT and 33 *Arid1b*$^{+/-}$ male and female mice. (D) Locomotor activity quantified by number of beam breaks in a familiar home-cage environment within periodic time intervals over the course of

two hours. Two-way repeated measures ANOVA was used to calculate the p value (left). Total number of beam breaks during two hours, analyzed with Student's t-test (right). 8 week old male mice (n = 20 WT and 20 *Arid1b*$^{+/-}$) were used. (E) Time spent in the target quadrant vs. other quadrants in the probe testing of the water maze test at day 14. Target quadrant contains a hidden platform. 20 WT and 20 *Arid1b*$^{+/-}$ mice were examined (left). Number of target quadrant crossings in the probe testing of the water maze test at day 14. 20 WT and 20 *Arid1b*$^{+/-}$ mice were examined (right). (F) Percent freezing time was quantified in the contextual fear conditioning test (n = 10 WT and 11 *Arid1b*$^{+/-}$ 8 week old male mice). (G) Percent freezing time was quantified in the cued fear conditioning test (n = 10 WT and 11 *Arid1b*$^{+/-}$ 8 week old male mice). (H) Foot shock sensitivity testing (n = 10 WT and 11 *Arid1b*$^{+/-}$ 8 week old male mice). Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes: *, p value ≤ 0.05; **, p value ≤ 0.01; *** p value ≤ 0.001; **** p value ≤ 0.0001; ns, not significant. Student's t-test (two-tailed distribution, two-sample unequal variance) was used to calculate p-values unless otherwise indicated in the corresponding figure legend.

Arid1b haploinsufficiency resulted in neuroanatomical and gene expression abnormalities

In an effort to understand how behavioral abnormalities arose, we examined *Arid1b*$^{+/-}$ brains for other neurodevelopmental abnormalities. Because some patients with *ARID1B* mutations exhibit corpus callosum hypoplasia or agenesis (*94*) , we examined brains of *Arid1b*$^{+/-}$ mice and identified a significant reduction in corpus callosum size (Figure 4A). Consistent with studies showing that small hippocampus, dentate gyrus, and cortex size are associated with anxiety and depressive disorders in mice and human (*100–103*), *Arid1b*$^{+/-}$ mice have smaller dentate gyri (Figure 4B) and both *Arid1b*$^{+/-}$ and *Arid1b*$^{-/-}$ pups had reduced cortical

thickness with reduced TBR1 marked neuronal cellularity (Figure 5A-D). Less proliferating cells were also seen in the subgranular zone of the dentate gyrus (Figure 4C, D, F, G), especially in posterior regions (Figure 4E, H). Thus, reduced corpus callosum size, dentate gyrus size, cortex thickness and proliferation are neuroanatomical and cellular correlates of the behavioral phenotypes seen in *Arid1b* mutants.

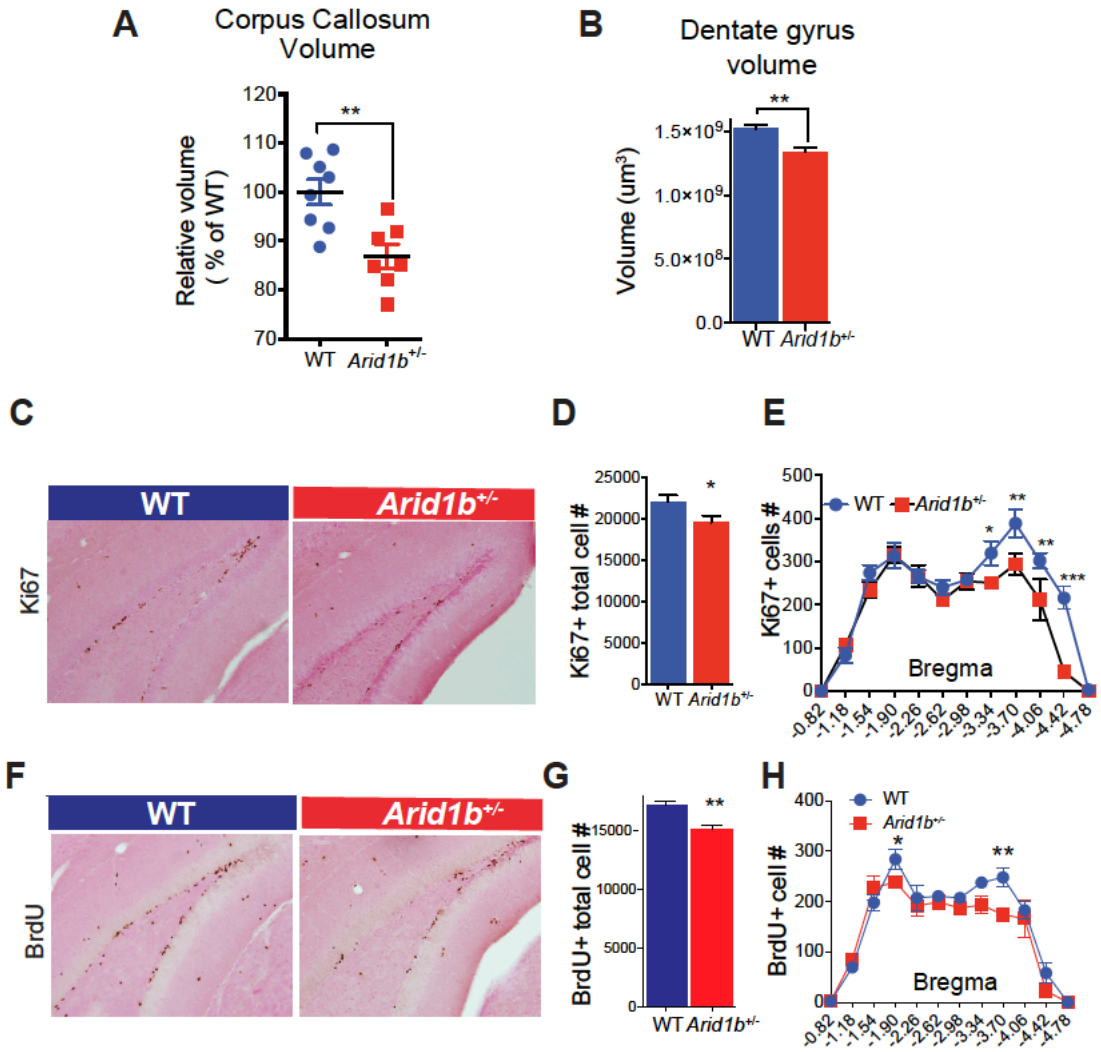


Figure 4. *Arid1b* haploinsufficiency results in neuroanatomical abnormalities implicated in neuropsychiatric diseases.

(A) Relative corpus callosum volume quantified through Cavalieri analysis (n = 8 WT and 7 *Arid1b*$^{+/-}$ brains from 50 day old females) (B) Dentate gyrus volume quantified through Cavalieri analysis (n = 7 WT and 7 *Arid1b*$^{+/-}$ brains from 50 day old females) (C) Representative Ki67+ immunostaining. (D) Quantitation of Ki67+ total cell number (8 WT and 7 *Arid1b*$^{+/-}$ brains from 50 day old females). (E) Bregma analysis was used to determine cell proliferation (Ki67) as a function of location in the subgranular zone of the dentate gyrus. Two-way ANOVA with uncorrected Fischer's Least Significant Difference (LSD) was used to calculate the statistics. (F) Representative BrdU immunostaining. WT and *Arid1b*$^{+/-}$ mice received one injection per day of the thymidine analog, bromodeoxyuridine (BrdU), for five days and brains were harvested three days following the last injection (6 WT and 4 *Arid1b*$^{+/-}$ brains from 50 days old females). (G) Quantification of BrdU + total cell number. (H) Bregma analysis was used to determine cell proliferation (BrdU) as a function of location in the subgranular zone of the dentate gyrus (n = 6 WT and 4 *Arid1b*$^{+/-}$). Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes, *, p value ≤ 0.05; **, p value ≤ 0.01; *** p value ≤ 0.001; **** p value ≤ 0.0001; ns, not significant. Student's *t*-test (two-tailed distribution, two-sample unequal variance) was used to calculate p-values unless otherwise indicated in the corresponding figure legend.

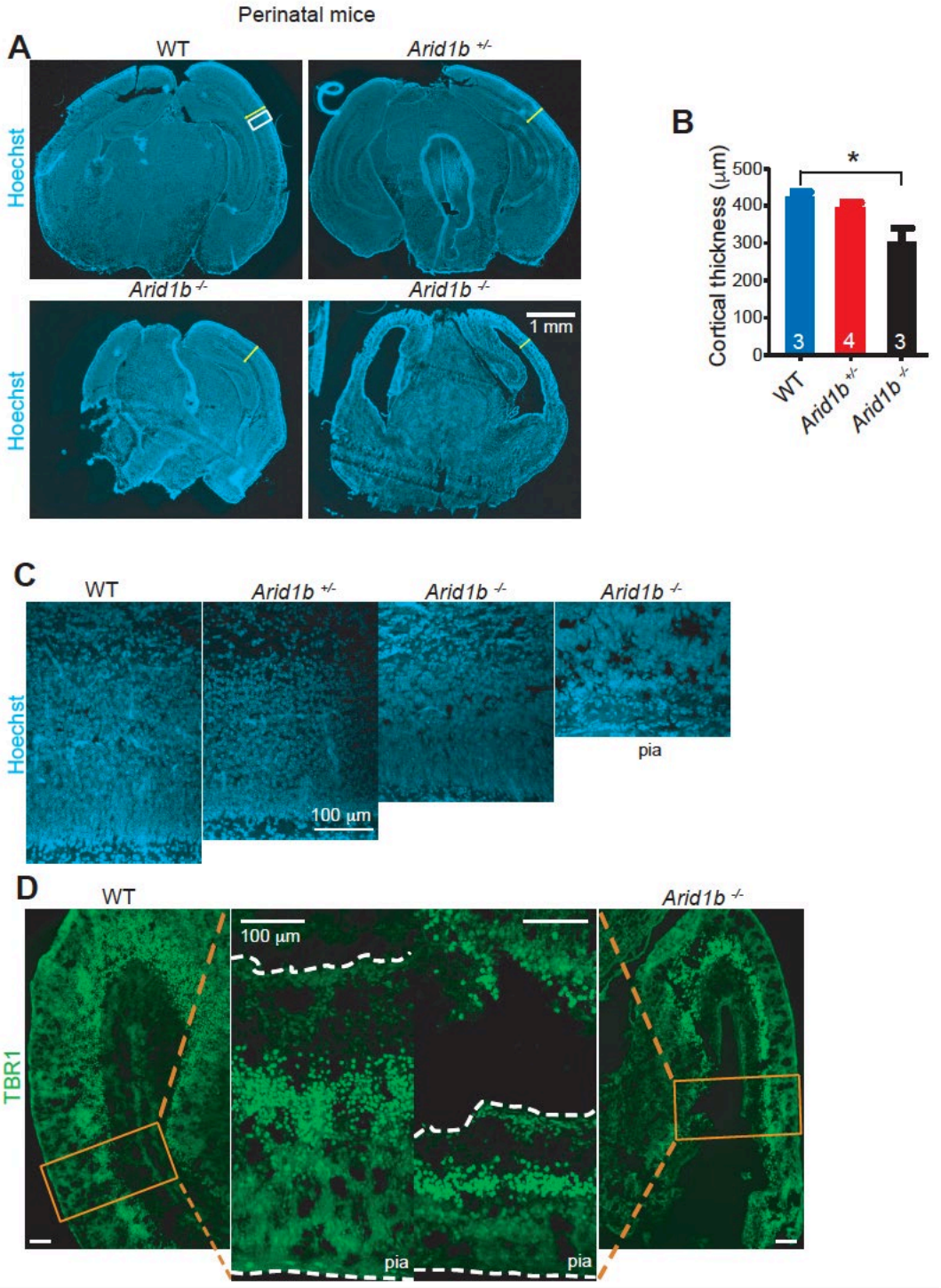

Perinatal mice
A
WT
Arid1b +/-
Arid1b -/-
Arid1b -/-
Hoechst
Hoechst
1 mm
B
Cortical thickness (μm)
500
400
300
200
100
0
*
3
4
3
WT
Arid1b +/-
Arid1b -/-
C
WT
Arid1b +/-
Arid1b -/-
Arid1b -/-
Hoechst
100 μm
pia
D
WT
Arid1b -/-
TBR1
100 μm
pia
pia

Figure 5. *Arid1b*$^{+/-}$ and *Arid1b*$^{-/-}$ brains have defects in cortical development

(A) Hoechst staining of WT, *Arid1b*$^{+/-}$, and *Arid1b*$^{-/-}$ brain sections showing cortical thickness differences outlined by the region in the white box. (B) Quantification of cortical thickness in WT, *Arid1b*$^{+/-}$, and *Arid1b*$^{-/-}$ brain sections. (C) Close up view of cortical regions in the white box from (A) above, showing reduced cortical thickness and cellularity. (D) TBR1 staining in WT and *Arid1b*$^{-/-}$ cortex showing neuron numbers.

RNA-seq was performed to examine the impact of *Arid1b* haploinsufficiency on transcriptional output in the hippocampus. Differential gene expression analysis showed 56 significantly down- and 79 upregulated mRNAs (edgeR FDR < 0.05; Figure 6A). As expected, *Arid1b* was one of the most downregulated genes. Globally, differentially regulated genes were associated with nervous system development as well as psychological, behavioral, and developmental disorders (Figure 6B). *Arid1b*$^{+/-}$ tissues also showed specific alterations in Ephrin, nNOS, axonal guidance and glutamate receptor signaling pathways (Figure 6C). 14 of 140 (10%) differentially regulated genes were among the highest ranking candidate autism risk genes identified in the SFARI gene database (*104*) (Figure 6D). To determine if some of these genes are directly regulated by SWI/SNF, we analyzed the ChIP-Seq targets of Smarca4 (Brg1), a core SWI/SNF complex subunit (*105*). 91 of 140 (65%) differentially regulated genes showed direct binding by Brg1 (Figure 6E), with positional enrichment at transcriptional start sites (TSSs) (Figure 6F, G). Arid1b-mediated SWI/SNF transcriptional activities appeared to directly regulate numerous neuropsychiatric related genes, including ones implicated in ASD (Figure 6H, I). Our data also show that haploinsufficiency is sufficient to cause broad gene

expression disruption, but future studies will be required to determine the exact downstream genes that account for the neuropsychiatric phenotypes.

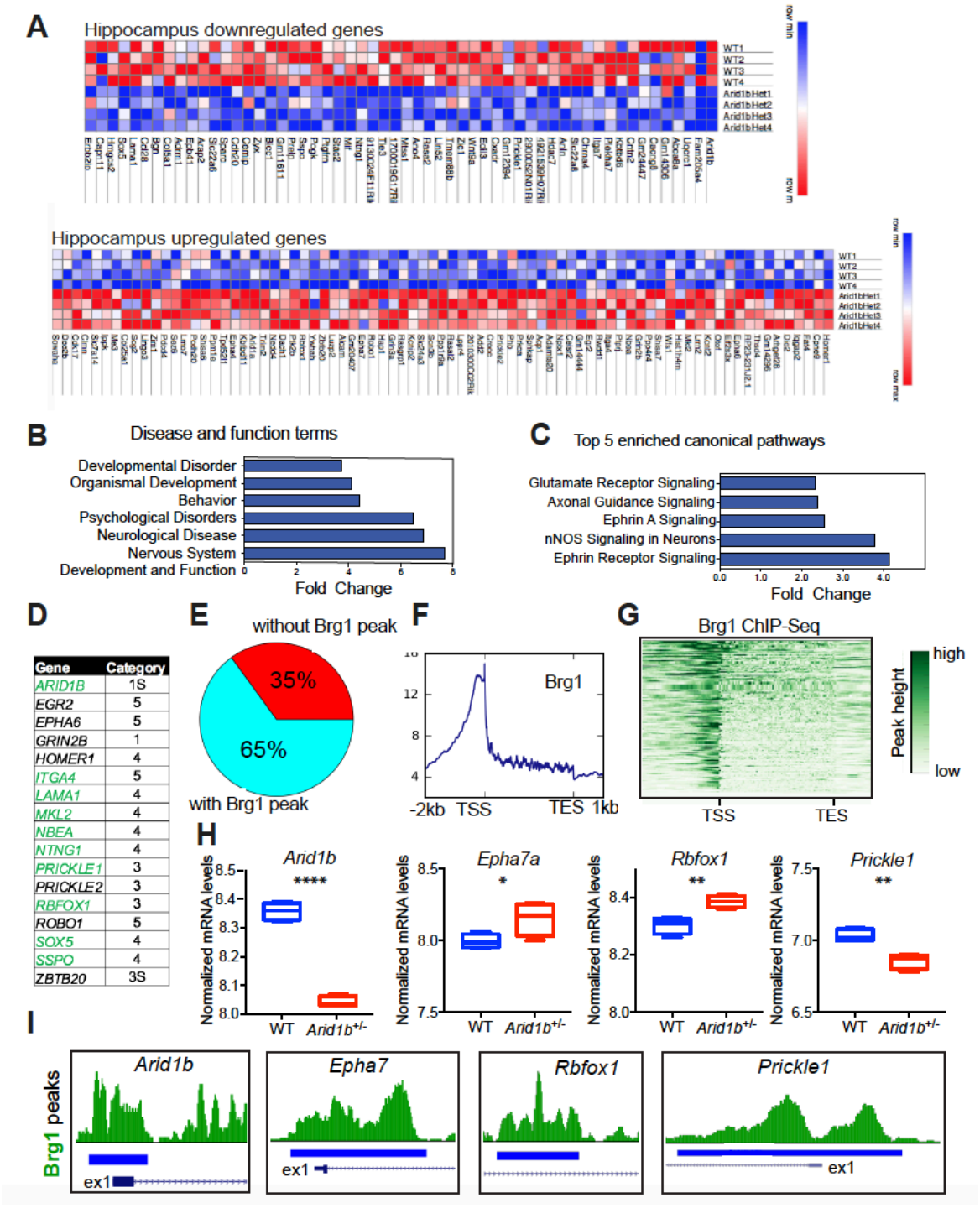


Figure 6. *Arid1b* haploinsufficiency results in changes in the expression of SWI/SNF regulated genes implicated in neuropsychiatric diseases.

(A) All significantly up- and downregulated genes in the *Arid1b*$^{+/-}$ hippocampus are ranked according to p-value (least to most significant from left to right). (B) Most enriched diseases and biological functions in hippocampus. (C) Most differentially regulated genetic pathways in the hippocampus. (D) 14 of 140 (10%) differentially regulated genes were among the highest ranking autism risk genes identified in the SFARI database. Category S: syndromic, Category 1: high confidence, Category 2: strong candidate, Category 3: suggestive evidence, Category 4: minimal evidence, Category 5: Hypothesized (*104*). (E) Pie chart showing that 91 of 140 (65%) differentially regulated genes in hippocampus are direct targets of Brg1, a core SWI/SNF complex subunit. Brg1 target genes were identified using ChIP-Seq in mouse e11.5 forebrain (*105*). (F) Metaplot showing enrichment of Brg1 at the TSSs of genes regulated by Arid1b. (G) Heatmap showing Brg1 promoter binding in these genes. (H) Differential mRNA expression of representative genes involved in neurodevelopment and ASD (Data from: SFARI database, updated September, 2016) (*104*). (I) Brg1 peaks showing direct binding of SWI/SNF at the promoters of ASD-related genes. Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes, *, p value ≤ 0.05; **, p value ≤ 0.01; *** p value ≤ 0.001; **** p value ≤ 0.0001; ns, not significant. Student's t-test (two-tailed distribution, two-sample unequal variance) was used to calculate p-values unless otherwise indicated in the corresponding figure legend.

Arid1b$^{+/-}$ *mice exhibited GHRH-GH-IGF1 axis defects*

Having established that *Arid1b* haploinsufficient mice recapitulate salient aspects of human *ARID1B*-opathies, we were particularly interested in identifying reversible pathological mechanisms and therapeutic opportunities. Since we identified neuroanatomical and neural expression aberrations in *Arid1b*$^{+/-}$ mice, we also asked if any non-neuropsychiatric syndromic features are potentially related to neurodevelopmental abnormalities. As mentioned previously, *Arid1b*$^{+/-}$ mice developed reduced nose-to-rump length and weight (Figure 7A, B). *Arid1b*$^{+/-}$ mice had disproportionally small kidneys and hearts, but no other gross organ defects (Figure 8A). Profiling using metabolic cages showed that *Arid1b*$^{+/-}$ mice had equivalent food intake and water consumption (Figure 8B, C), suggesting that size differences were unlikely due to food intake or energetic differences.

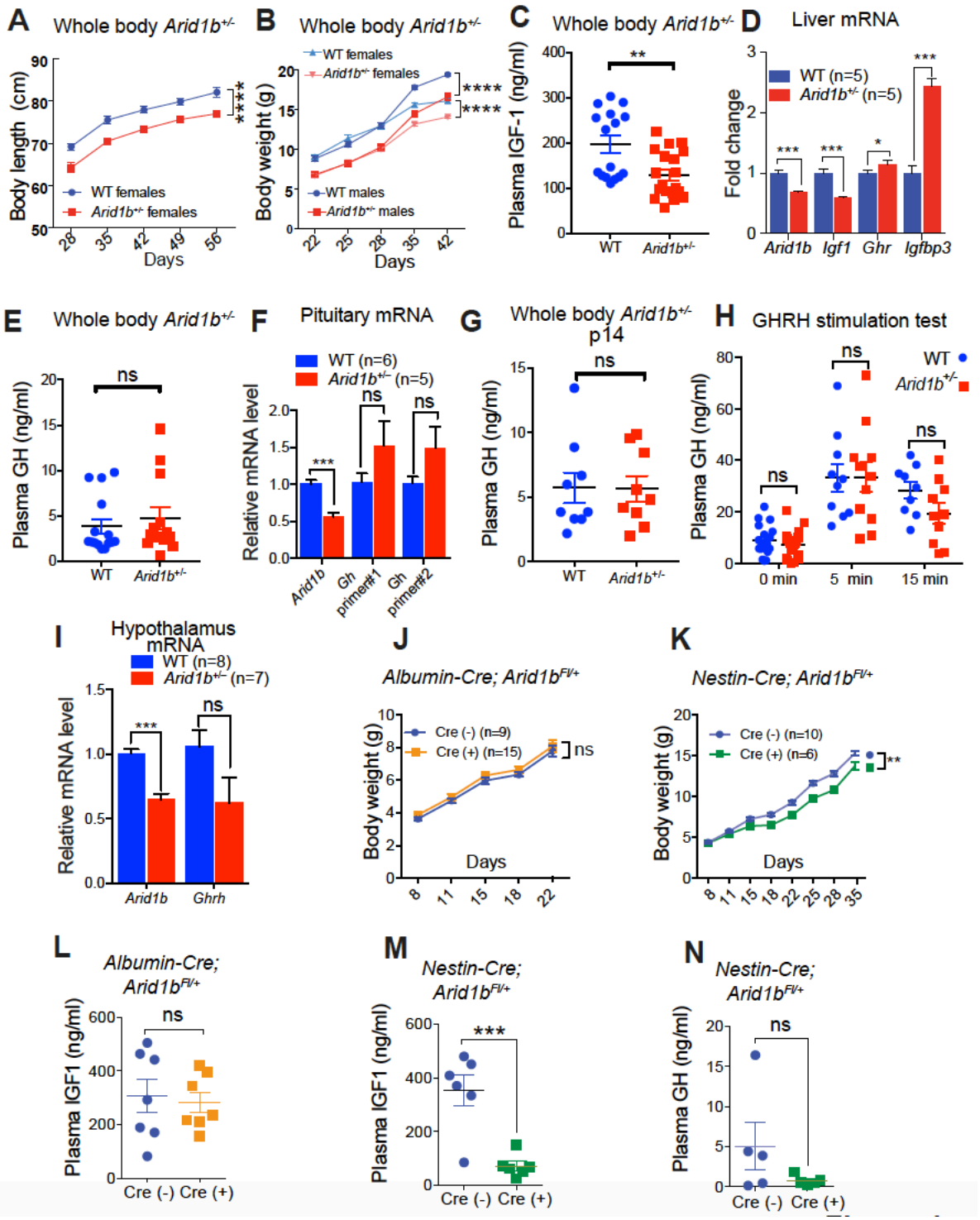

A
Whole body Arid1b+/-
Body length (cm)
WT females
Arid1b+/- females
Days
B
Whole body Arid1b+/-
Body weight (g)
WT females
Arid1b+/- females
WT males
Arid1b+/- males
Days
C
Whole body Arid1b+/-
Plasma IGF-1 (ng/ml)
WT
Arid1b+/-
D
Liver mRNA
Fold change
WT (n=5)
Arid1b+/- (n=5)
Arid1b
Igf1
Ghr
Igfbp3
E
Whole body Arid1b+/-
Plasma GH (ng/ml)
ns
WT
Arid1b+/-
F
Pituitary mRNA
Relative mRNA level
WT (n=6)
Arid1b+/- (n=5)
Arid1b
Gh primer#1
Gh primer#2
G
Whole body Arid1b+/- p14
Plasma GH (ng/ml)
H
GHRH stimulation test
Plasma GH (ng/ml)
WT
Arid1b+/-
0 min
5 min
15 min
I
Hypothalamus mRNA
Relative mRNA level
WT (n=8)
Arid1b+/- (n=7)
Arid1b
Ghrh
J
Albumin-Cre; Arid1bFl/+
Body weight (g)
Cre (-) (n=9)
Cre (+) (n=15)
Days
K
Nestin-Cre; Arid1bFl/+
Body weight (g)
Cre (-) (n=10)
Cre (+) (n=6)
Days
L
Albumin-Cre; Arid1bFl/+
Plasma IGF1 (ng/ml)
Cre (-)
Cre (+)
M
Nestin-Cre; Arid1bFl/+
Plasma IGF1 (ng/ml)
Cre (-)
Cre (+)
N
Nestin-Cre; Arid1bFl/+
Plasma GH (ng/ml)
Cre (-)
Cre (+)

Figure 7. Growth retardation in *Arid1b*$^{+/-}$ mice is due to GH-IGF1 axis deficiency with a neuronal source.

(A) Body length (nose-to-rump) curve of females (n = 9 WT and 9 *Arid1b*$^{+/-}$). Repeated ANOVA with Bonferroni's post hoc analysis used as the statistical test. (B) Body weight growth curve for males (n = 14 WT and 14 *Arid1b*$^{+/-}$) and females (n = 20 WT and 20 *Arid1b*$^{+/-}$). Repeated ANOVA with Bonferroni's post-hoc analysis used as the statistical test. (C) Plasma IGF1 as measured by ELISA (n = 16 WT and 19 *Arid1b*$^{+/-}$ 28-41 day old male and female mice). (D) *Igf1* mRNA in WT and *Arid1b*$^{+/-}$ livers as measured by qPCR (n = 5 WT and 5 *Arid1b*$^{+/-}$ livers from 45-day old female mice). (E) Plasma GH as measured by ELISA (n = 15 WT and 14 *Arid1b*$^{+/-}$ 28-41 day old male and female mice). (F) *Gh* mRNA in WT and *Arid1b*$^{+/-}$ pituitary as measured by qPCR (n = 6 WT and 5 *Arid1b*$^{+/-}$ pituitary from 33-44 day old female mice). (G) Plasma GH (n = 9 WT and 9 *Arid1b*$^{+/-}$ 2 week old male mice). (H) Plasma GH before and after stimulation by human GHRH (n = 19 WT and n = 20 *Arid1b*$^{+/-}$ mice at baseline, n = 11 WT and n = 10 *Arid1b*$^{+/-}$ mice 5 and 15 min after GHRH administration) (I) *Ghrh* mRNA in WT and *Arid1b*$^{+/-}$ mediobasal hypothalamus as measured by qPCR (n = 8 WT and 7 *Arid1b*$^{+/-}$ samples from 33-44 day old female mice). (J) Body weight curve for female *Arid1b*$^{Fl/+}$ (n = 9) and *Albumin-Cre*; *Arid1b*$^{Fl/+}$ (n = 15) mice. (K) Body weight curve for female *Arid1b*$^{Fl/+}$ (n = 10) and *Nestin-Cre*; *Arid1b*$^{Fl/+}$ (n = 6) mice. (L) Plasma IGF1 levels for 40-45 day old female *Arid1b*$^{Fl/+}$ (n = 7) and *Albumin-Cre*; *Arid1b*$^{Fl/+}$ (n = 7) mice. (M) Plasma IGF1 levels for 30-45 days old female *Arid1b*$^{Fl/+}$ (n = 6) and *Nestin-Cre*; *Arid1b*$^{Fl/+}$ (n = 6) mice. (N) Plasma GH levels for female *Arid1b*$^{Fl/+}$ (n = 5) and *Nestin-Cre; Arid1b*$^{Fl/+}$ (n = 5) mice. Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes, *, p value ≤ 0.05; **, p value

≤ 0.01; *** *p* value ≤ 0.001; **** *p* value ≤ 0.0001; ns, not significant. Student's *t*-test (two-tailed distribution, two-sample unequal variance) was used to calculate p-values unless otherwise indicated in the corresponding figure legend.

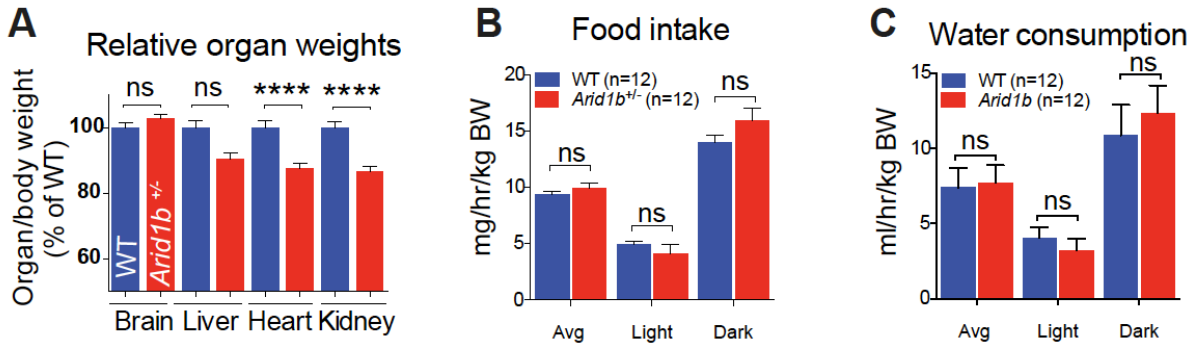


Figure 8. Growth and metabolic analysis of *Arid1b*$^{+/-}$ mice.

(A) *Arid1b*$^{+/-}$ organ/body weight ratio presented as % of WT organ/body weight ratio. (B) Food intake and (C) Water consumption was quantified in a metabolic cage over 4 days. The absolute data was normalized to body weight (n = 12 WT and 12 *Arid1b*$^{+/-}$ 8 week old male mice). Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes: *, *p* value ≤ 0.05; **, *p* value ≤ 0.01; *** *p* value ≤ 0.001; **** *p* value ≤ 0.0001; ns, not significant. Student's *t*-test (two-tailed distribution, two-sample unequal variance) was used to calculate p-values unless otherwise indicated in the corresponding figure legend.

A critical factor that regulates both body size and brain development is insulin-like growth factor (IGF1). We found a significant reduction in plasma IGF1 levels in *Arid1b*$^{+/-}$ mice (Figure 7C), and confirmed a reduction in *Igf1* mRNA in the liver, which is a major source of IGF1 (Figure 7D). To discern if there was a hypothalamic, pituitary, peripheral, or combinatorial problem that led to IGF1 deficiency, we performed a series of endocrinologic tests. In the same cohorts of mice with IGF1 deficiency, fasting GH was not significantly different in *Arid1b*$^{+/-}$ mice (Figure 7E). In addition, there were no significant *Gh* mRNA expression differences between WT and *Arid1b*$^{+/-}$ pituitary glands despite differences in *Arid1b* mRNA levels (Figure 7F). We also confirmed that GH was not altered in younger 2-week old *Arid1b*$^{+/-}$ mice, an age where GH levels are more critical for growth (Figure 7G). The combination of low IGF1 and normal GH levels pointed to a peripheral defect without appropriate GH compensation from the pituitary gland.

Because GH was not elevated as would be expected if there was only a peripheral IGF1 producing defect, we asked if the *Arid1b*$^{+/-}$ pituitary was capable of making and secreting sufficient amounts of GH in the context of GH stimulation testing. In multiple cohorts of *Arid1b*$^{+/-}$ mice, GH levels were never significantly different at baseline and also increased normally at multiple time points after stimulation with Growth Hormone Releasing Hormone (GHRH) (Figure 7H). This indicated a normal ability for the pituitary to respond to exogenous GHRH. Next, we attempted to determine if the hypothalamus was not producing enough *Ghrh*. We dissected the mediobasal hypothalamus containing *Ghrh* expressing neurons to examine *Ghrh* expression. We found that *Ghrh* mRNA levels were not significantly different (Figure 7I), indicating a lack of appropriate GHRH response to IGF1 deficiency, suggesting a central defect that contributed to growth impairment.

In addition, we sought genetic evidence for a partial central (hypothalamic or pituitary) root cause of IGF1 deficiency by generating organ-specific *Arid1b* mutant models. We used

Nestin-Cre in the brain, *Albumin-Cre* in the liver, and *Ckmm-Cre* in the skeletal muscles to spatially control *Arid1b* haploinsufficiency. Neither *Albumin-Cre* (Figure 7J) nor *Ckmm-Cre; Arid1b*$^{Fl/+}$ mice (data not shown) showed growth or morphological defects, while *Nestin-Cre; Arid1b*$^{Fl/+}$ mice recapitulated the growth impairments seen in whole body *Arid1b*$^{+/-}$ mice (Figure 7K), suggesting at least a neuronal contribution to growth impairment. While *Albumin-Cre; Arid1b*$^{Fl/+}$ mice showed no IGF1 differences, *Nestin-Cre; Arid1b*$^{Fl/+}$ mice had reduced plasma IGF1 levels (Figure 7L, M). Moreover, *Nestin-Cre; Arid1b*$^{Fl/+}$ mice showed an inappropriate lack of GH increase in the face of this IGF1 deficiency (Figure 7N). Since liver specific *Arid1b*$^{+/-}$ mice did not replicate the whole body *Arid1b*$^{+/-}$ mice, it is possible that a combination of central and multi-organ peripheral defects in the GHRH-GH-IGF1 axis were required to fully recapitulate the growth impairment of whole body *Arid1b*$^{+/-}$ mice.

GH therapy reversed growth retardation and muscle weakness

Given plasma IGF1 deficiency in *Arid1b*$^{+/-}$ cohorts, we first tested if IGF1 replacement could rescue physical aspects of developmental delay and abnormal behavioral phenotypes. Neither body size (Figure 9A) nor elevated plus maze abnormalities (Figure 9B) were rescued after treating WT and *Arid1b*$^{+/-}$ cohorts with recombinant human IGF1 (rhIGF1). This was not surprising because it is known that exogenous IGF1 is unstable and often does not efficiently reach target tissues responsible for growth (*106*).

The fact that GH was not elevated in the context of low IGF1 suggested to us that there was not adequate GH production or compensation. Thus, we asked if GH supplementation could rescue some of the physical aspects of developmental delay. WT and *Arid1b*$^{+/-}$ cohorts were treated with recombinant mouse GH (rmGH) (Figure 9C). After 40 days of treatment, *Arid1b* heterozygous mice gained significantly more body weight and nose-to-rump length than did WT mice (Figure 9D, E), demonstrating that exogenous GH supplementation was

sufficient to rescue growth retardation in *Arid1b*$^{+/-}$ mice. Given this selective efficacy for mutant mice, we asked if GH could potentially improve muscle weakness often associated with CSS. We found that at baseline, *Arid1b*$^{+/-}$ mice also had muscle weakness identified through grip strength testing. Replacement with GH was able to selectively increase muscle strength in mutant mice (Figure 9F, G). Despite improvements in physical manifestations, GH replacement was not able to reverse behavioral phenotypes such as anxiety, as measured in the elevated plus maze (Figure 9H). This suggested that correcting the GHRH-GH-IGF1 axis was not sufficient to rescue neuropsychiatric manifestations, but was able to reverse growth retardation mediated by *Arid1b* deficiency.

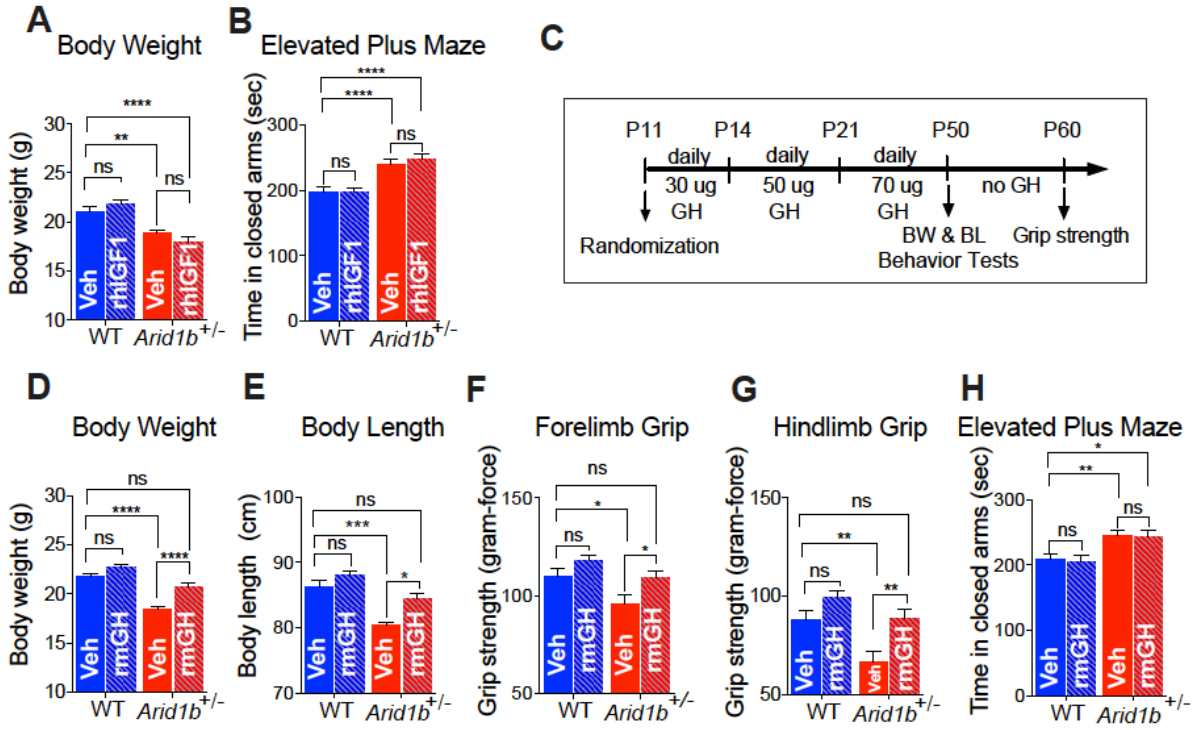
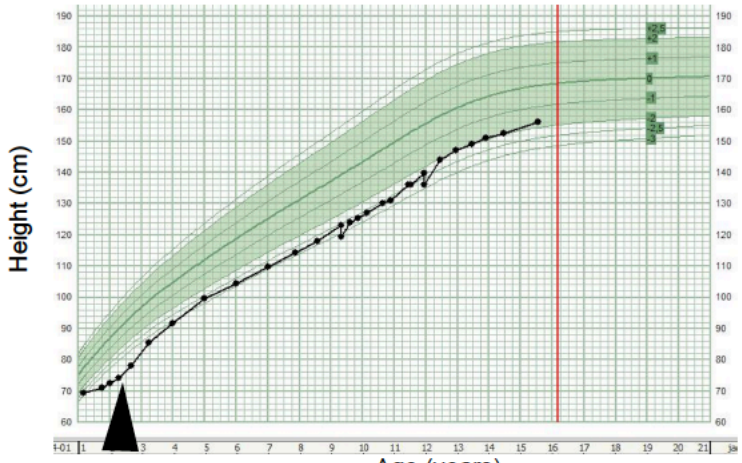


Figure 9. GH therapy reverses growth retardation and muscle weakness.

(A) Body weights at p50 (WT + vehicle (n = 12), WT + rhIGF1 (n = 12), *Arid1b*$^{+/-}$ + vehicle (n = 13) and *Arid1b*$^{+/-}$ + rhIGF1 (n = 13)). 0.5 mg/kg rhIGF1 was administrated daily starting from p11. (B) Time spent in the closed arms of elevated plus maze at p50 (WT + vehicle (n = 27), WT + rhIGF1 (n = 29), *Arid1b*$^{+/-}$ + vehicle (n = 21) and *Arid1b*$^{+/-}$ + rhIGF1 (n = 22)). 0.5 mg/kg rhIGF1 was administrated daily starting from postnatal day 11. (C) Schema showing the duration and dose of daily recombinant GH treatment. (D) Body weights at p50 (WT + saline (n = 20), WT + rmGH (n = 16), *Arid1b*$^{+/-}$ + saline (n = 16) and *Arid1b*$^{+/-}$ + rmGH (n = 16)). (E) Nose-to-rump lengths at p50 (WT + saline (n = 7), WT + rmGH (n = 7), *Arid1b*$^{+/-}$ + saline (n = 6) and *Arid1b*$^{+/-}$ + rmGH (n = 6)). (F) Forelimb grip strength at p50 (WT + saline (n = 9), WT + rmGH (n = 9), *Arid1b*$^{+/-}$ + saline (n = 9) and *Arid1b*$^{+/-}$ + rmGH (n = 9)). (G) Hindlimb grip strength at p50 (WT + saline (n = 9), WT + rmGH (n = 9), *Arid1b*$^{+/-}$ + saline (n = 9) and *Arid1b*$^{+/-}$ + rmGH (n = 9)). (H) Time spent in the closed arms of elevated plus maze at p50 (WT + vehicle (n = 19), WT + rmGH (n = 16), *Arid1b*$^{+/-}$ + vehicle (n = 19) and *Arid1b*$^{+/-}$ + rmGH (n = 16)). Values represent mean ± SEM. Asterisks indicate significant differences between indicated littermate genotypes: *, p value ≤ 0.05; **, p value ≤ 0.01; *** p value ≤ 0.001; **** p value ≤ 0.0001; ns, not significant. Two-way ANOVA was used to calculate the p value.

In an analysis of 60 *ARID1B* CSS patients, height was shown to be significantly reduced (*107*). In addition, some non-syndromic patients with missense mutations in *ARID1B* exhibited growth deficiency due to partial GH deficiency (*108*). We also obtained detailed clinical information from additional CSS patients, two with *ARID1B* mutations (from the www.arid1bgene.com database) and one with a mutation in *SMARCA4*, which encodes another SWI/SNF component. All three of these cases had deficiencies in the GH-IGF1 axis and clear beneficial responses to GH replacement therapy (growth curves for the *ARID1B* patients are shown in Figure 10A, B). These data from humans and mice suggest deficiencies at various parts of the GHRH-GH-IGF axis, leading to growth impairment responsive to GH supplementation.

A


Height (cm)
Age (years)
Beginning of GH administration

B

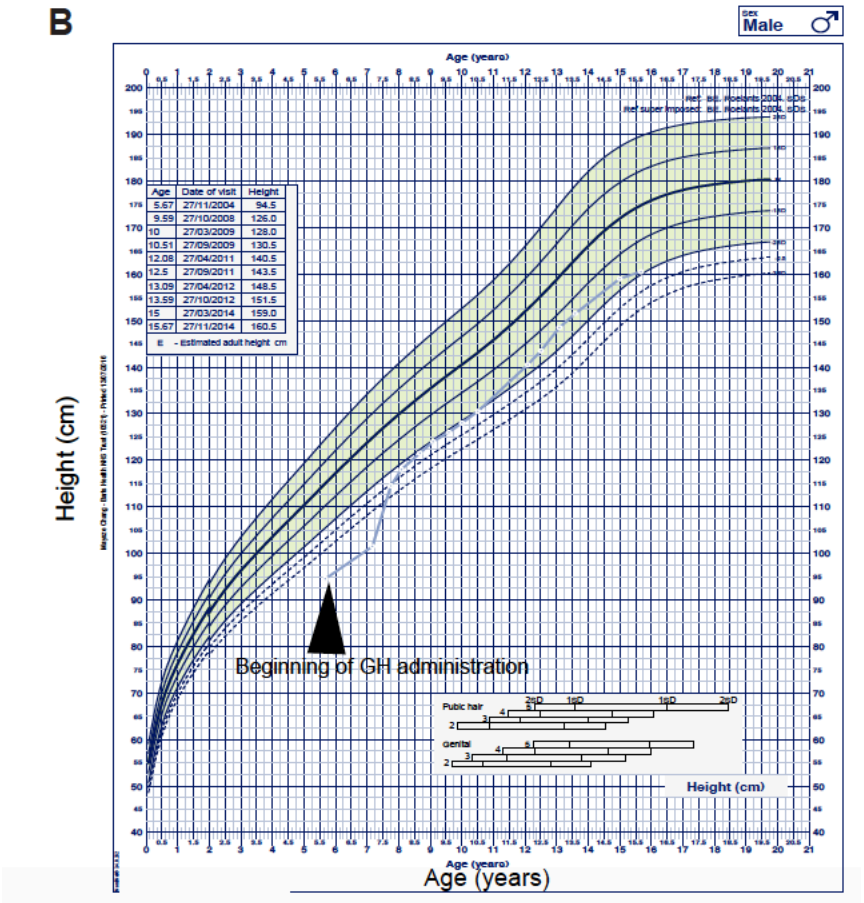

Sex
Male
Age (years)
Ref: Bill. Hoekstra 2004, sDs
Ref super imposed: Bill. Hoekstra 2004, sDs
Age | Date of visit | Height
5.67 | 27/11/2004 | 94.5
9.59 | 27/10/2008 | 126.0
10 | 27/03/2009 | 128.0
10.51 | 27/09/2009 | 130.5
12.08 | 27/04/2011 | 140.5
12.5 | 27/09/2011 | 143.5
13.09 | 27/04/2012 | 148.5
13.59 | 27/10/2012 | 151.5
15 | 27/03/2014 | 159.0
15.67 | 27/11/2014 | 160.5
tE - estimated adult height cm
Beginning of GH administration
Pubic hair
Genital
Height (cm)
Height (cm)
Age (years)

Figure 10. Two *ARID1B* mutant CSS patients are GH deficient and responsive to GH replacement.

(A) Growth curve of a CSS patient with an *ARID1B* mutation. GH treatments started at 2 years of age. Green shaded area represents the normal range. (B) Growth curve of a CSS patient with a 6q21.5 deletion, which includes the *ARID1B* locus. GH treatments started at 5.5 years of age.

Given the concern that *Arid1b* might be a tumor suppressor gene (*23*, *109*), and that GH replacement might synergistically activate malignant outgrowth, we asked if *Arid1b* haploinsufficiency with or without GH treatment influences cancer risk. Whole body *Arid1b*$^{+/-}$ mice did not develop cancer in any tissue in the first 12 months of life, even after GH treatment (data not shown). To test if they might be more susceptible to cancer after mutagenic exposure, we gave a single dose of DEN at two weeks of age and found no differences in liver cancer formation at 12 months of age (Figure 11A-E). Our data do not support the concept that heterozygous *Arid1b* mutations alone confer a significantly increased risk for cancer.

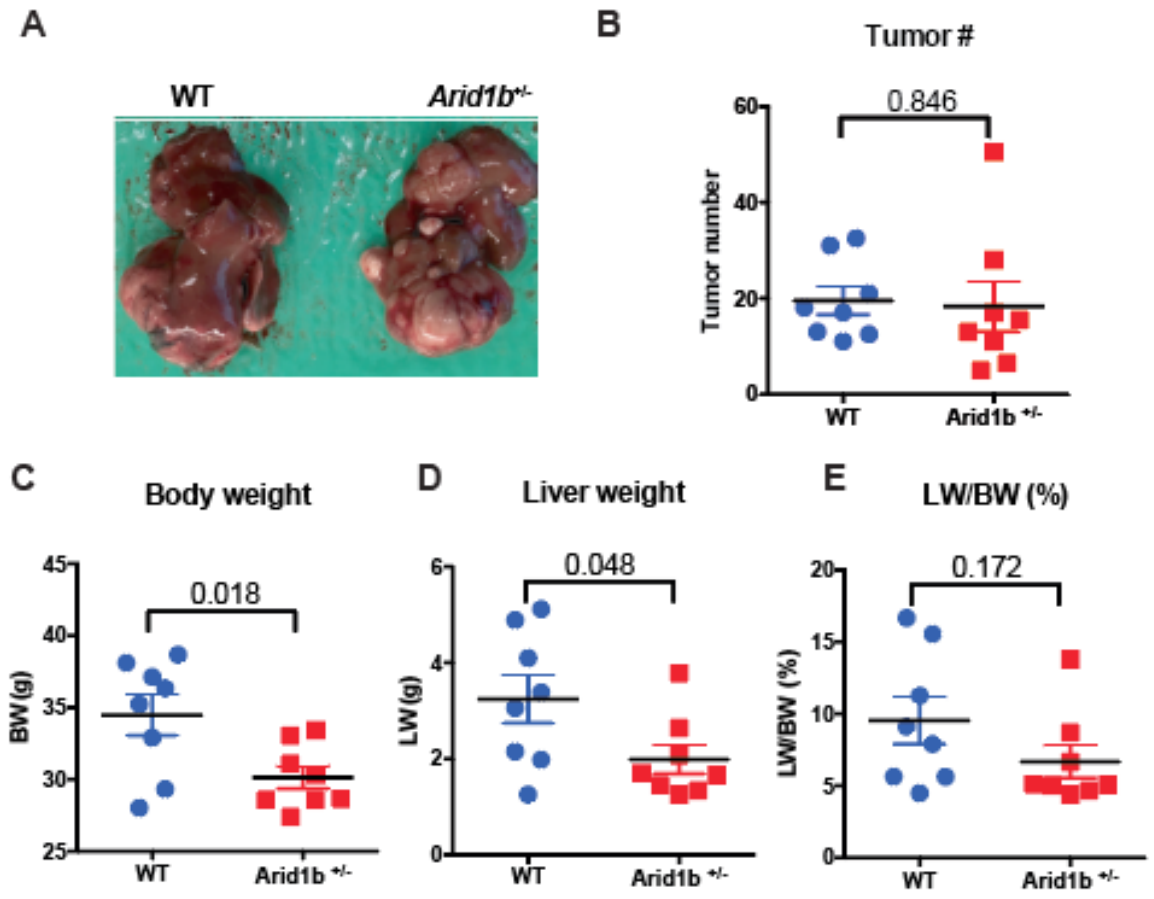


Figure 11. Whole body *Arid1b*$^{+/-}$ mice did not develop increased liver tumor burden 12 months after DEN injection.

(A) Representative images of WT and *Arid1b*$^{+/-}$ livers 12 months after DEN injection. One dose of 25 mg/kg DEN was injected at postnatal day 15. (B) Total liver surface tumor numbers in (WT + DEN (n = 8), *Arid1b*$^{+/-}$ + DEN (n = 8)) mice. (C) Body weights in (WT + DEN (n = 8), *Arid1b*$^{+/-}$ + DEN (n = 8)) mice. (D) Liver weights in (WT + DEN (n = 8), *Arid1b*$^{+/-}$ + DEN (n = 8) mice. (E) Liver weight to body weight percentage (LW/BW (%)) in (WT + DEN (n = 8), *Arid1b*$^{+/-}$ + DEN (n = 8) mice.

CONCLUDING REMARKS AND FUTURE DIRECTIONS

In summary, we have developed the first mouse model of *Arid1b* haploinsufficiency, one of the most common genetic lesions found in ASD, ID, and CSS. Several aspects of our model recapitulated the features of the overlapping disorders associated with *ARID1B* (Table 1). Our model is faithful to the heterozygosity seen in most *ARID1B*-opathies. It is interesting that another model of ASD involving a chromatin remodeling gene called *CHD8* also requires haploinsufficiency (*70*) and suggests that dose will play a critical mechanistic role in phenotypes resulting from mutations in epigenetic regulators. Our study also raises interesting questions about how genotype relates to phenotype in diseases involving *ARID1B*. Some children with *ARID1B* mutations have a subset but not all features of CSS, and others have different disorders such as ASD or isolated corpus callosum agenesis. For example, Yu *et al.* reported cases of idiopathic short stature without cognitive defects that were attributed to *de novo ARID1B* missense mutations, suggesting that there are either differences between mutations or important genetic interactions that play a key role in defining the phenotypic expression of these mutations (*108*). It is also possible that differences between the function of *Arid1b* in mouse and human brain development account for intellectual and cognitive discrepancies. Our mouse model affords the ability to interrogate these types of questions in different strain backgrounds and with genetic interactors. We are hopeful that this will advance the understanding of *ARID1B* and SWI/SNF in human diseases.

We also uncovered a role for the GHRH-GH-IGF1 axis in *ARID1B*-opathies. Previously, height was shown to be reduced in *ARID1B* patients (*107*), but there are only anecdotal findings of GH deficiencies (Figure 10A, B) (*108*). After the clinical identification of *ARID1B* or SWI/SNF mutations, interventions for short stature are usually not investigated. Thus, it is likely that GHRH-GH-IGF1 deficiency is under-diagnosed and rarely treated in this

patient population. Conversely, *ARID1B* mutations are not suspected in patients with non-syndromic short stature, and could represent a more common causative mechanism than previously suspected. The findings here should motivate deeper interrogation of the GHRH-GH-IGF1 axis and potentially GH supplementation in syndromic patients with CSS or non-syndromic patients with *ARID1B* mutations and short stature.

Our study does not pinpoint the exact source of the peripheral IGF1 deficiency since liver-specific Cre lines did not recapitulate the IGF1 deficiency seen in the whole body *Arid1b*$^{+/-}$ mice. Given the *Nestin-Cre* results, it is possible that reduced IGF1 production in the brain led to reduced plasma IGF1 and the inability to compensate with GHRH and GH exacerbated growth retardation. Another possibility that subtle peripheral defects in the liver and muscle will only manifest when combined with defects in other organs such as the brain. Future studies with tissue-specific conditional experiments could help to resolve these questions. Overall, our study provides a preclinical model for mechanistic and therapeutic dissection of *ARID1B* related diseases, and offers a translatable avenue to alleviate growth related aspects of developmental delay.

Human ARID1B features	**Diagnosis**	**References**	**Mouse**
Intellectual/ cognitive disability	CSS, ID	(*110*), (*15*), (*111*)	No
Growth retardation	CSS	(*16*)	Yes
Coarse facial features	CSS	(*112*)	Unknown
Muscle hypotonia	CSS	(*110*)	Yes
Hydrocephalus	CSS	(*113*)	Yes
Agenesis or hypoplasia of corpus callosum	CSS	(*111*), (*15*)	Yes
Brachydactyly, hypoplastic nail/finger	CSS	(*110*), (*15*), (*114*)	Unknown
Abnormal vocalization, speech impairment	ASD, CSS	(*15*)	Yes
Anxiety	ASD, CSS	(*20*), (*21*)	Yes
Social interaction deficits	ASD	(*20*), (*21*)	Yes
Repetitive behaviors	ASD	(*20*), (*21*)	Yes

Table 1. Major clinical features associated with *ARID1B* mutations and phenotypes seen in *Arid1b*$^{+/-}$ mice.

Abbreviations: CSS: Coffin-siris syndrome, ID: Intellectual disability, ASD: Autism spectrum disorder.

Acknowledgments

This chapter was published as Celen et al. 2017 (*30*).

We thank Dr. Eric Olson for providing *Ckmm-Cre* mice. We thank the Children's Research Institute Sequencing Core for sequencing and UTSW Bioinformatics Core for the analysis. We thank Dr. Erik Plautz and Laura Ingle in the UTSW Neuro-Models Facility for performing the grip strength test and UTSW Rodent Behavior Core Facility for assisting some of the behavioral tests.

Data and materials availability

RNA-Seq data is deposited to GEO database and can be accessed through GEO accession number (GSE92238).

CHAPTER 4

ARID1A LOSS POTENTIATES PANCREATIC β-CELL REGENERATION THROUGH ACTIVATION OF EGF SIGNALING

INTRODUCTION

Understanding how to recover endogenous β-cells in disease might ultimately depend on understanding how β-cell abundance is controlled in the first place. The molecular circuitry that dictates total β-cell mass during development and dynamic physiologic situations such as pregnancy is poorly understood. It is also unclear if the same mechanisms regulate β-cell expansion after injuries and during disease processes. One unexploited therapeutic strategy is to expand endogenous β-cells in type 1 and 2 diabetes, settings in which normal regenerative capacity cannot fully compensate for β-cell loss. Given the importance of dedifferentiation and proliferation in profoundly regenerative organisms such as zebrafish and planaria (*115*), it stands to reason that facilitating chromatin state changes that occur physiologically could also

increase self-renewal capacity in mammalian tissues. We hypothesized that the mammalian SWI/SNF (mSWI/SNF) ATP-dependent chromatin remodeling complex, known to promote terminal differentiation (*34*, *116–120*) and also known to limit regeneration after liver injuries (*34*), might operate as a general repressor of regeneration in tissues other than the liver. Given key similarities between β-cell and hepatocyte self-renewal, we reasoned that mSWI/SNF might also suppress regenerative capacity in β-cells, a centrally important cell type in metabolic disease.

mSWI/SNF are large 10-15 component complexes containing a core ATPase (BRG1 or BRM, also known as SMARCA4 and SMARCA2), plus non-catalytic subunits that influence targeting and complex activities (*121*) such as the mutually exclusive ARID1A (BAF250A) and ARID1B (BAF250B) subunits, which together define the canonical BAF family of mSWI/SNF complexes. The loss of ARID1A alters the assembly of the ATPase module (*14*), influencing canonical BAF complex targeting and DNA accessibility, resulting in global changes in gene regulation (*122–127*). How the molecular consequences of mSWI/SNF perturbation relate to physiologic phenotypes and disease outcomes is an active area of exploration.

Intriguingly, ARID1A and other mSWI/SNF components are downregulated during islet expansion, findings shared with other regenerative tissues such as the liver (*34*). We demonstrated that this downregulation is functionally important by using spatially and temporally specific conditional knockout mouse models subjected to multiple types of β-cell injuries. These findings support the idea that physiological events that occur during regeneration can be further amplified to accelerate tissue healing. Molecular dissection of the events occurring downstream of ARID1A loss showed that NRG-ERBB-NR4A signaling activities were increased. Interestingly, the ERBB protein family (EGFR/ERBB1, ERBB2, ERBB3, and ERBB4) of transmembrane receptor tyrosine kinases (RTKs) have been

implicated in diabetes (*58*, *59*, *128*). Remarkably, SNPs in the human *ERBB3* locus are among the strongest signals in T1D GWAS but how these SNPs affect ERBB signalling and β-cell biology were previously unknown (*129–131*). Here, we show that ARID1A-containing SWI/SNF complexes have a major and unexpected role in β-cell regeneration, operating in large part through pathways that were previously implicated by GWAS studies in diabetogenesis. This expands our emerging understanding of mSWI/SNF as a central regulator of tissue regeneration, and underscores the need to understand chromatin remodeling mechanisms that may represent therapeutic targets in regeneration.

RESULTS

Arid1a expression is suppressed during physiologic β-cell expansion

ARID1A-containing SWI/SNF chromatin remodeling complexes (or canonical BAF complexes) drive terminal differentiation and block proliferation by regulating chromatin accessibility at loci targeted by lineage specific transcription factors (*61*, *132*, *133*). In the absence of ARID1A, increased liver regeneration and accelerated wound healing were observed (*34*). Although the primary mechanism for the homeostatic maintenance of adult β-cells is self-duplication, majority of adult pancreatic β-cells are still mostly in post-mitotic state (*36*, *134*). We wondered if the inhibition of SWI/SNF activity could increase their proliferative potential. Genes that encode components of the SWI/SNF complex are expressed at high levels in the mouse β-cell line MIN6 and primary pancreatic islets compared to the mouse α-cell line ATC1 and primary hepatocytes (Figure 12A). In particular, *Arid1a* is highly expressed in quiescent β-cells within the islet (Figure 12B). Next, we examined *Arid1a* expression in islets

under conditions that demand β-cell expansion, such as pregnancy and 50% partial pancreatectomy (PPx) (*135*, *136*). *Arid1a* mRNA was lower in maternal islets during pregnancy-associated β-cell expansion, and protein levels declined until the time of birth (Figure 12C,D). Similarly, ARID1A protein levels were reduced during regeneration induced by PPx (Figure 12E).

We asked if signals known to elicit β-cell hyperplasia can influence the levels of *Arid1a* mRNA. We isolated wild-type (WT) mouse islets and treated with glucagon-like peptide-1 (GLP-1), high glucose, insulin-like growth factor-2 (IGF2), and interleukin-1β (IL-1β), all factors known to drive β-cell proliferation (*137–141*). Each factor suppressed *Arid1a* mRNA levels in islets, providing a potential molecular rationale for how the suppression of *Arid1a* during β-cell expansion is linked to upstream signals (Figure 1F). These results indicated dynamic regulation of ARID1A and SWI/SNF complex components during β-cell expansion and regeneration, suggesting functional roles during these processes.

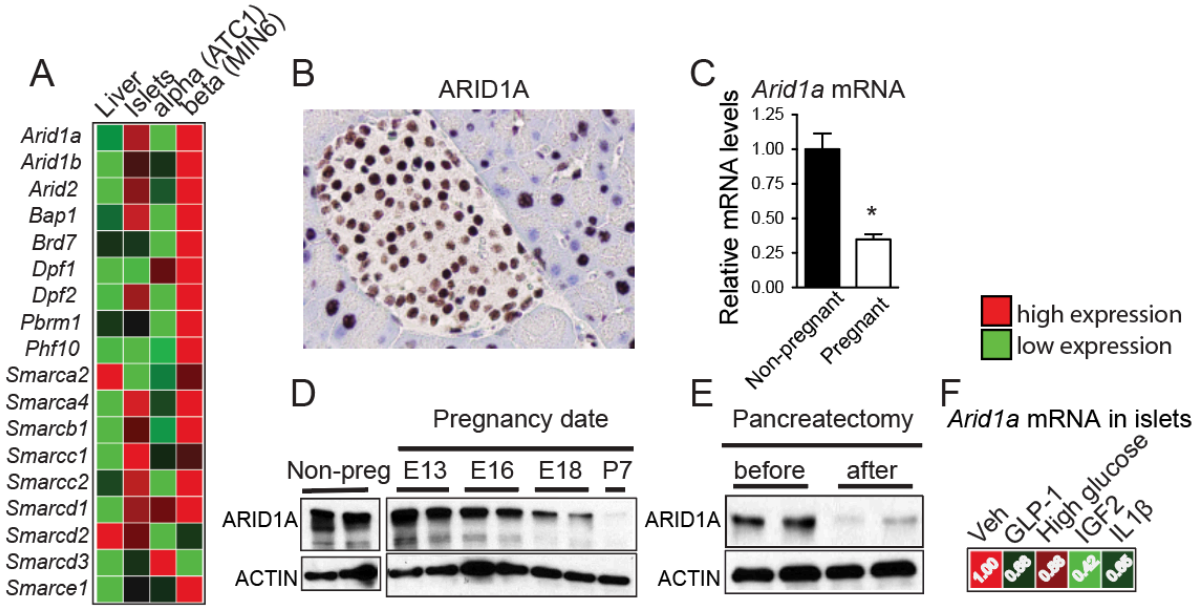


Figure 12. *Arid1a* expression is suppressed during physiologic β-cell expansion.

(A) mRNA expression heat map of SWI/SNF components in mouse liver, mouse islet, alpha cell line (ATC1) and beta cell line (MIN6) using qPCR. (B) ARID1A immunostaining in the adult mouse pancreas. An islet is shown. (C) qPCR showing *Arid1a* mRNA levels in islets isolated from non-pregnant and pregnant females at gestational day 16 (n = 3 mice per group). (D) Western blot showing ARID1A levels in the islets of non-pregnant females and at different time points during gestation. (E) Western blot showing ARID1A levels in the islets before and 5 days after 50% pancreatectomy (PPX). (F) qPCR showing the reduction in *Arid1a* levels in isolated islets after 6 hours of treatment in culture with the following stimuli: GLP-1 (100 nM), high glucose (25 mM), IGF2 (200nM) and IL1β (1 ng/ml).

Whole-body Arid1a deletion protected against type I diabetes

To determine if ARID1A restrains β-cell mass in adults, we inducibly deleted *Arid1a* in a whole body "*Arid1a* KO" mouse strain. Global deletion was induced with tamoxifen in adult *Ubiquitin-CreER; Arid1a*$^{Fl/Fl}$ (*Ubc-CreER; Arid1a*$^{Fl/Fl}$) mice (Figure 13A). Almost complete absence of the flox band and protein levels were observed (Figure 13B-D). Whole body *Arid1a* KO mice did not show overt signs of disease and appeared to be healthy. At baseline, whole body glucose tolerance was unchanged, indicating that ARID1A is not required for the regulation of glucose homeostasis or uptake under non-injury, basal conditions (Figure 13E). To determine if *Arid1a* KO mice were more able to cope with β-cell destruction in a type 1 diabetes (T1D) model, we subjected mice to streptozotocin (STZ), a chemical that ablates insulin-producing β-cells (Figure 13A). After STZ, *Arid1a* KO mice were almost completely protected against the development of diabetes as measured by fed state glucose levels (Figure 13F) and glucose tolerance testing (GTT) (Figure 13G). This left the possibility of either insulin secretion or insulin sensitivity changes, particularly because *Arid1a* was deleted in all tissues. Mice subjected to STZ did not show changes in insulin sensitivity based on insulin tolerance testing (ITT), suggesting no substantial metabolic impact of *Arid1a* loss on peripheral tissues, despite efficient deletion in those tissues (Figure 13H). However, *Arid1a* KO mice produced higher levels of insulin after glucose administration (Figure 13I), suggesting a β-cell mediated mechanism of protection against or response to STZ-induced T1D.

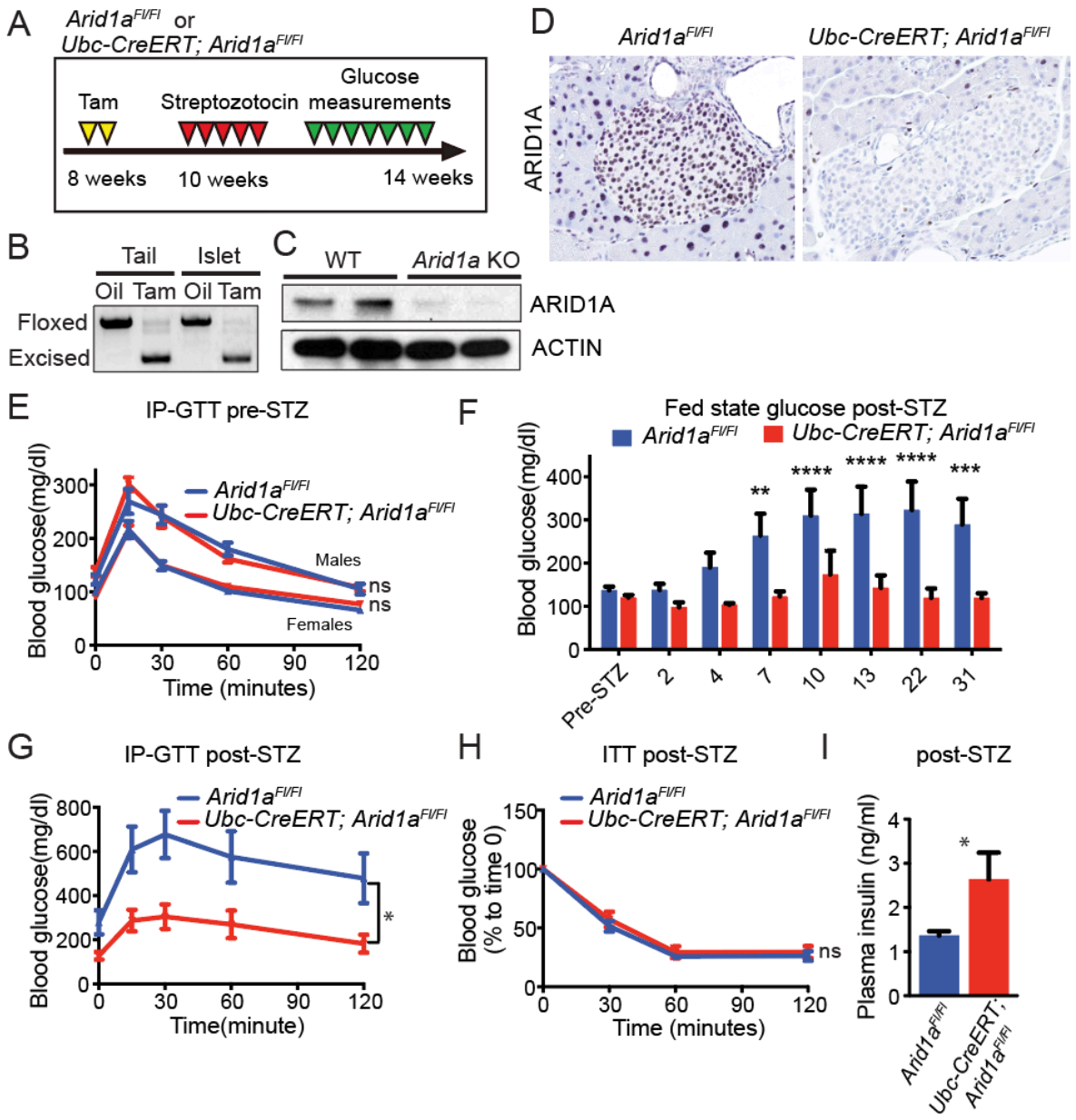


Figure 13. Whole body *Arid1a* deletion protects against STZ-induced T1D.

(A) Schema for STZ-induced diabetes studies in *Arid1a*$^{Fl/Fl}$ control and *Ubc-CreER; Arid1a*$^{Fl/Fl}$ whole body KO mice. Two weeks after administration of 20 mg/kg tamoxifen oral gavage x 2 days, diabetes was induced by injecting 50 mg/kg STZ IP for 5 consecutive days. (B) Genotyping of tail and islets to assess the recombination at the *Arid1a*$^{Fl/Fl}$ locus. (C) Western blot assessing the reduction in ARID1A protein levels. (D) Immunostaining of ARID1A in the adult pancreas of *Ubc-CreER; Arid1a*$^{Fl/Fl}$ mice. (E) IP glucose tolerance test at baseline

(Females: n=5 for *Arid1a*$^{Fl/Fl}$ and n=4 for *Ubc-CreER; Arid1a*$^{Fl/Fl}$, Males: n=3 for *Arid1a*$^{Fl/Fl}$ and n=8 for *Ubc-CreER; Arid1a*$^{Fl/Fl}$). (F) After STZ, fed state blood glucose was measured following STZ (n=11 for *Arid1a*$^{Fl/Fl}$ and n=9 for *Ubc-CreER; Arid1a*$^{Fl/Fl}$). (G) IP glucose tolerance test performed 2 weeks post-STZ. (H) Insulin tolerance test performed 2 weeks post-STZ. (I) Plasma insulin levels measured by ELISA 2 weeks post-STZ.

Arid1a deficiency leads to a β-cell-mediated anti-diabetic phenotype

Given that the whole body KO mice have lost *Arid1a* in multiple cell types, it was possible that non-β-cell autonomous mechanisms might have been at play. In WT mice, ARID1A is expressed in most acinar, all duct, and all islet cells. To rule out potential paracrine or endocrine effects, we used *Ptf1a-Cre; Arid1a*$^{Fl/Fl}$ mice to induce deletion in non-β-cells in the pancreas. *Ptf1a* is a transcription factor that is expressed in all cell types in the pancreas starting around E9.5 (*71*). In *Ptf1a-Cre; Arid1a*$^{Fl/Fl}$ mice, ARID1A was lost in all acinar and most duct cells but ARID1A protein was retained in the islets, as shown recently (*142*, *143*). GTT did not reveal any differences in glucose clearance efficiency between WT, *Ptf1a-Cre; Arid1a*$^{+/Fl}$, or *Ptf1a-Cre; Arid1a*$^{Fl/Fl}$ mice (Figure 14A). Following STZ treatment (Figure 14B), there were also no differences in blood glucose measurements over the course of 1 month (Figure 14C), ruling out a potential contribution from acinar or duct cells to the phenotypes observed in whole-body KO mice.

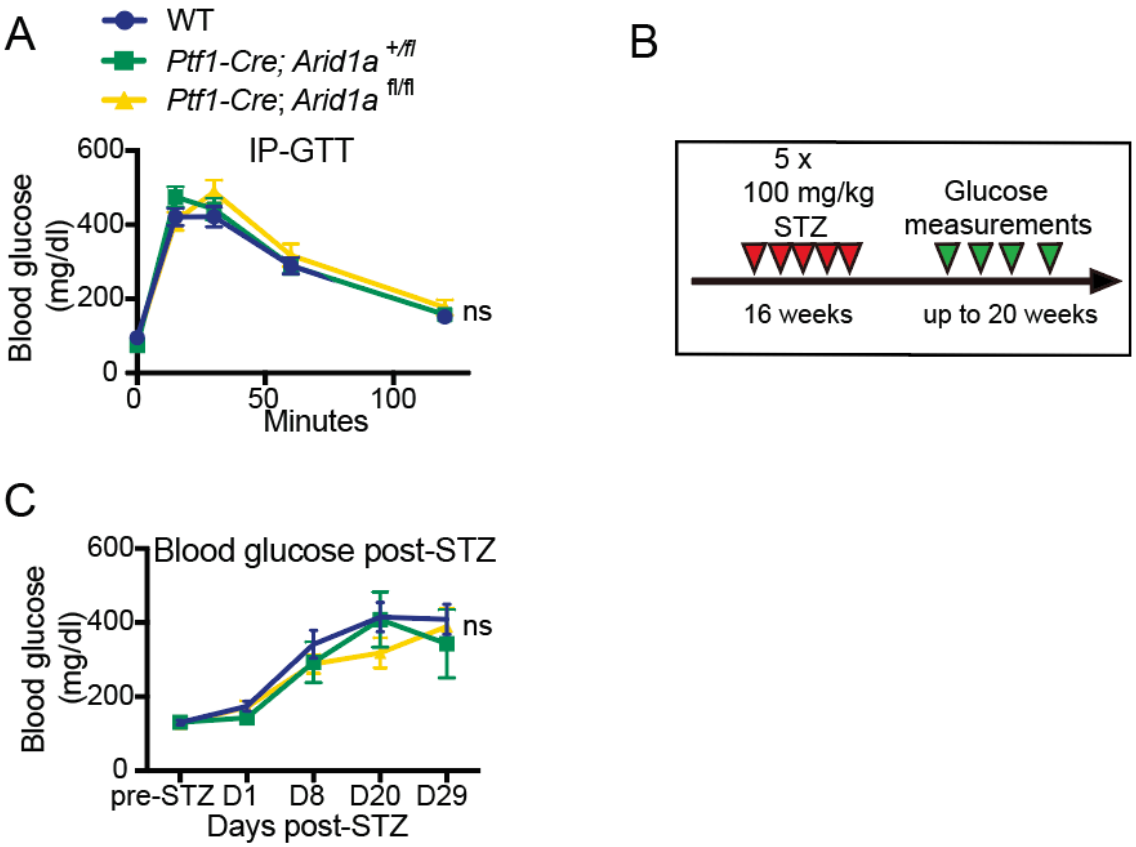


Figure 14. Acinar and ductal knockout of *Arid1a* does not phenocopy the whole body knockout.

(A) Intraperitoneal glucose tolerance test for *Arid1a*$^{Fl/Fl}$, *Ptf1a-Cre; Arid1a*$^{+/Fl}$, *and Ptf1a-Cre; Arid1a*$^{Fl/Fl}$ mice after overnight fasting. (B) Experimental timeline for STZ-induced diabetes. (C) Fed state blood glucose measurements taken post-STZ.

Next, we generated a temporally regulated, β-cell specific model of *Arid1a* loss. This model allowed us to induce β-cell specific deletion of *Arid1a* after β-cell injuries, thus giving us the ability to answer questions specifically about β-cell regeneration, rather than cell survival in the presence of injuries. We employed a *MIP (Mouse Insulin Promoter)-rtTA; TRE (Tetracycline Responsive Element promoter)-Cre* transgenic system to allow such genetic control (Figure 15A) (*72*). *MIP-rtTA; TRE-Cre; Arid1a*$^{Fl/Fl}$ *(Arid1a* βKO) mice and littermate controls without *Cre* and/or *rtTA* were used to model cell-type and temporally specific *Arid1a* deletion. Notably, some *Arid1a* floxed band was still present in islets, potentially due to α, δ and γ cell retention of the *Arid1a* floxed allele (Figure 15B). Despite the absence of ARID1A nuclear staining (Figure 15C), islet cell morphology was normal. Two weeks post-deletion, the ability to clear glucose after GTT challenge was unaltered (Figure 15D), suggesting that there was no change in β-cell function at baseline. Next, both groups of mice were given five doses of STZ (50mg/kg per day x 5 days), then fed doxycycline (1mg/mL) water to conditionally induce *Arid1a* deletion in the β-cells of the experimental group. β-cell specific deletion of *Arid1a* after STZ administration prevented the rise of blood glucose to pathological levels (Figure 15E), and reduced the profound weight loss associated with T1D (Figure 15F). In the setting of STZ, *Arid1a* βKO showed higher insulin levels than control mice. Not surprisingly, insulin levels were still lower in the STZ treated *Arid1a* βKO group than uninjured WT mice since STZ was still able to ablate a subset of β-cells (Figure 15G). These data showed that *Arid1a* deficiency in β-cells was sufficient to preserve insulin production in response to STZ-induced T1D.

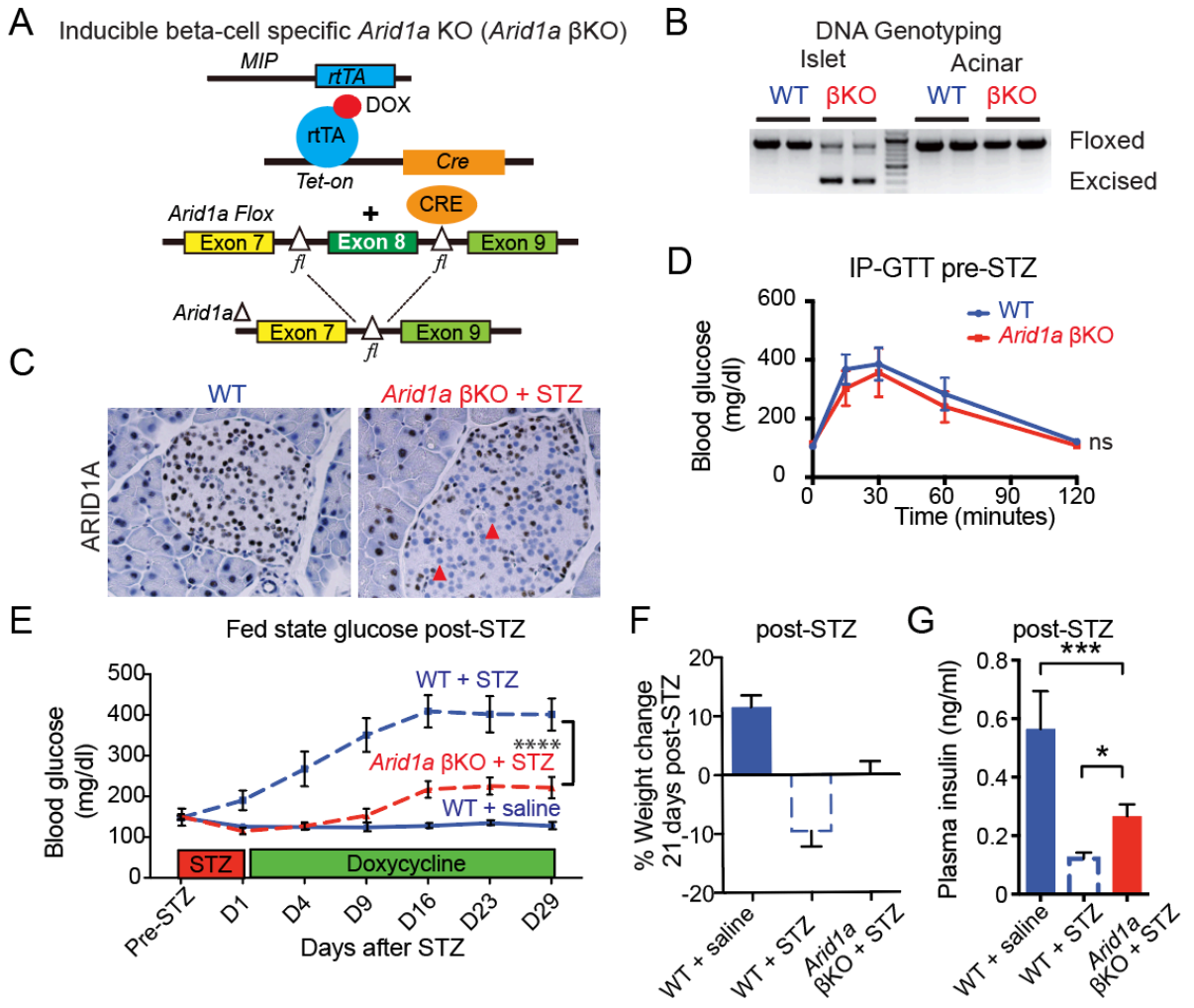


Figure 15. *Arid1a* deficiency leads to a β-cell-mediated anti-diabetic phenotype.

(A) Dox-inducible KO of *Arid1a* in mouse β-cells (*Arid1a* βKO). rtTA is expressed under the control of the mouse insulin promoter (*MIP*). In the presence of doxycycline (Dox), rtTA activates the transcription of the *TRE-Cre* transgene. CRE in turn converts the floxed *Arid1a* alleles to knockout alleles. Dox (1mg/mL) was added to the drinking water to conditionally induce *Arid1a* deletion. (B) Genotyping of islets and acinar cells to assess the recombination at the $Arid1a^{Fl/Fl}$ locus. Partial excision is expected in islets since cell types other than β cells are present. (C) Immunostaining of ARID1A in WT and *Arid1a* βKO pancreata. (D) IP glucose tolerance test 2 weeks post-dox. (E) Fed state blood glucose after STZ. 5 doses of 50 mg/kg STZ were injected before dox-induced β-cell deletion of *Arid1a*. After the last dose of STZ,

dox (1mg/mL) was provided in drinking water (n=4 mice for WT + saline, n=12 mice for WT + STZ, n=10 mice for *Arid1a* βKO + STZ). (F) % body weight change 21 days post-STZ. (G) Plasma insulin levels measured by ELISA 30 days post-STZ.

Arid1a loss results in increased β-cell survival and proliferation

We examined islets before and after STZ injury. Insulin staining in WT and *Arid1a* βKO islets at baseline showed no differences. Thirty days after STZ, *Arid1a* βKO mice had more β-cells than did WT mice (Figure 16A). In addition, WT mice showed considerable α-cell expansion as marked by glucagon staining (Figure 16B). Because resistance to diabetes could have been mediated by either a failure to lose β-cells or an increased ability to regenerate β-cells, we examined proliferation in the insulin expressing β-cell compartment. *Arid1a* deficient β-cells had greater numbers of Ki-67 positive cells. In the WT setting, a majority of the proliferating cells were positioned in the islet periphery, which is the non-insulin expressing compartment (Figure 16C,D). We then wondered whether the observed increase in proliferation after STZ had any effect on islet number or area. Although we did not detect any increase in the number of islets per unit of pancreas area (Figure 16E), there was a trend towards increased individual islet area ($p=0.06$) (Figure 16F) in βKO mice 3 weeks after STZ.

Given the remote possibility that *Arid1a* deficient β-cells could be protected from T1D due to a lack of STZ induced destruction rather than increased regeneration, we also performed PPx, a surgical assay that does not rely on the ability of cells to metabolize chemical toxins such as STZ. PPx leads to β-cell proliferation that compensates for overall islet loss. Six days after resection, the islets from regenerated pancreata of βKO mice showed a significant increase

in β-cell proliferation as measured by Ki-67 positive cell number within the insulin expressing compartment (Figure 16G). In summary, *Arid1a* loss in β-cells did not constitutively enforce islet overgrowth in the absence of injury but increased proliferation after chemical and surgical injuries.

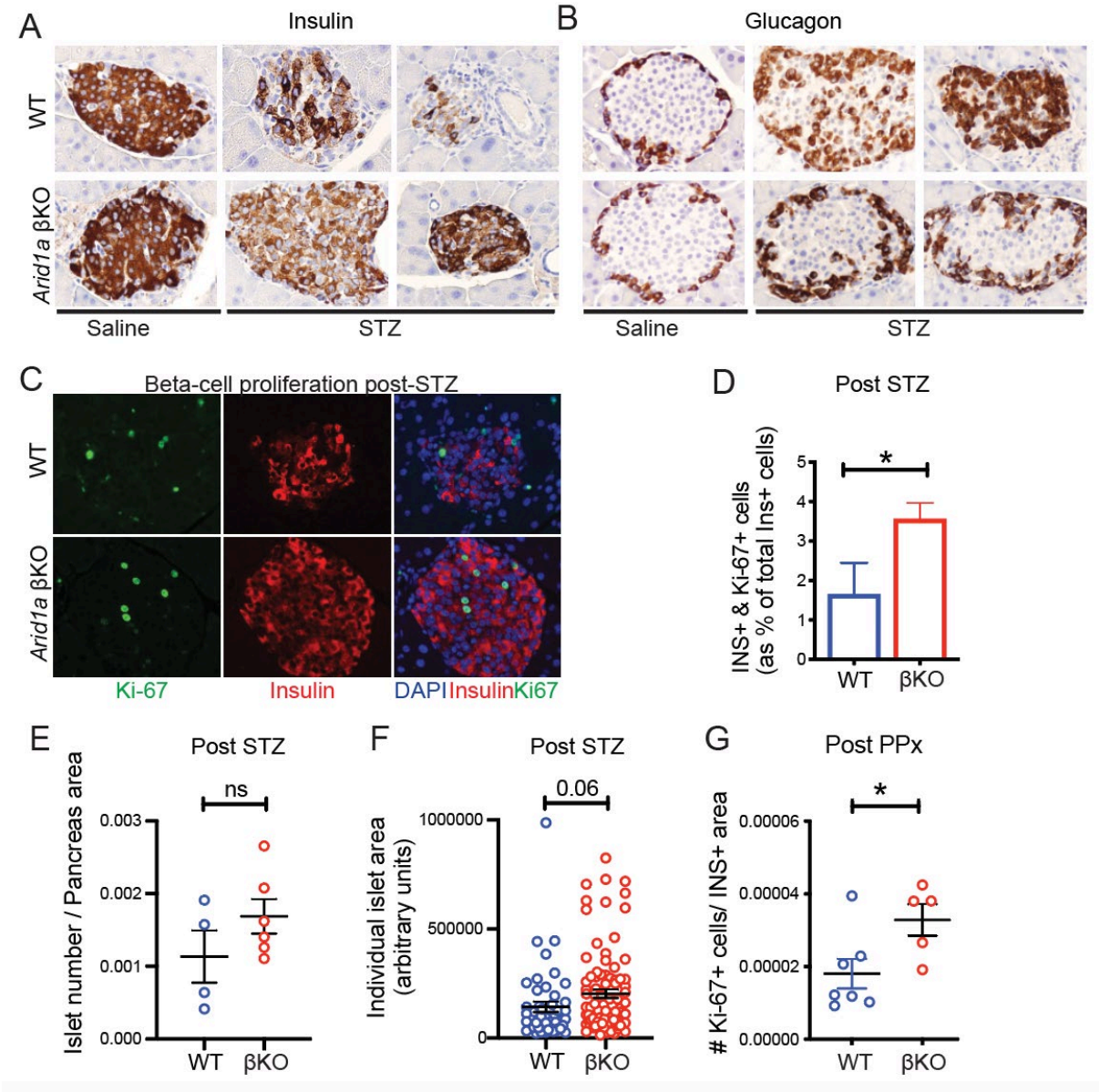


Figure 16. *Arid1a* loss results in increased β-cell survival and proliferation.

(A) Insulin staining of WT and *Arid1a* βKO pancreata treated with saline or STZ. Samples were collected 1 month post-STZ. (B) Glucagon staining of WT and *Arid1a* βKO pancreata treated with saline or STZ. Samples were collected 1 month post-STZ. (C) Representative Ki-67 staining (green) of WT and *Arid1a* βKO pancreata post-STZ. Insulin is used as a β-cell marker (red). Samples were collected 1 month post-STZ. (D) Quantification of the percentage of insulin+ & Ki-67+ cells to all insulin+ cells (total of 9 islets from 1 WT and total of 25 islets from 3 *Arid1a* βKO mice). (E) The number of islets per pancreas section area in WT and *Arid1a* βKO mice 3 weeks post-STZ. Each dot represents the total number of islets per section area in each mouse (n = 4 WT mice for WT and n = 6 *Arid1a* βKO mice). (F) Individual islet area (n = 4 WT mice and 6 *Arid1a* βKO mice, at least 10 islets were measured for each mouse). (G) Ki-67+ cell number/insulin+ area in islets from regenerated pancreata 6 days after PPx (n = 7 WT and 5 *Arid1a* βKO mice, values from ~10 islets per mice were calculated and averaged to get a single data point in the graph).

Phenotypes associated with Arid1a loss are dependent on EGF/Neuregulin hyperactivation

To probe the transcriptional programs associated with *Arid1a* deletion, we performed RNA-seq on islets from control and KO mice, before and after pancreatectomy. Of 2796 differentially-expressed genes, 1586 were up and 1210 were downregulated (Figure 17A). Gene Set Enrichment Analysis (GSEA) performed on differentially expressed genes showed that genes involved in the neuregulin (NRG) and epidermal growth factor (EGF) pathways

were collectively overproduced in KO islets (Figure 17B & Figure 17A right inset). EGF and NRG are the ligands that activate ERBB family of RTKs. Binding of the EGF family growth factors leads to dimerization and phosphorylation of RTKs which leads to the activation of ERK1/2 and phosphatidylinositol 3-kinase (PI3K) signaling pathways (*144*).

To functionally interrogate the transcriptomic findings, we performed shRNA knockdown of *Arid1a* in BTC6 and MIN6 immortalized β-cell lines to determine if *Arid1a* loss might cause increased dependency on NRG/EGF signaling. *Arid1a* knockdown in BTC6 and MIN6 caused increased phospho-EGFR (p-EGFR) after EGF ligand exposure, indicating potentiation of the pathway (Figure 17C). In addition, EGFR/ERBB inhibition with the small molecule inhibitors erlotinib and canertinib were able to ablate p-EGFR in the presence of EGF (Figure 17D). Similar proliferative phenotypes were confirmed with *Arid1a* shRNA in MIN6 cells and these small molecule inhibitors selectively abrogated the hyperproliferation associated with *Arid1a* knockdown (Figure 17E).

To corroborate the shRNA results using another *in vitro* model, we generated *Arid1a* deficient MIN6 β-cells using CRISPR. Multiple single cell MIN6 clones for each genotype were expanded and confirmed for *Arid1a* loss (Figure 17F). Similar to β-cells from mouse islets, *Arid1a* KO MIN6 clones grew more rapidly than Gal4 targeted control clones (Figure 17G). Genes identified in the RNA-seq experiment were also upregulated in KO vs. control islets (*Nr4a1, Nr4a2, Srf, JunB*; see Figure 5H). Collectively, these data show that *Arid1a* loss in β-cells causes a preferential dependency on ERBB signalling.

To determine if differentially expressed genes from the RNA-Seq analysis also showed changes in epigenetic marks, we performed H3K27 acetylation (H3K27ac) ChIP-seq. Sites marked by H3K27ac were clustered into groups that increased or decreased in H3K27ac in response to *Arid1a* deletion (Figure 17I). Interestingly, many sites increased in H3K27ac

abundance in the *Arid1a* heterozygous and homozygous clones (clusters 3 and 4). The clusters containing sites that changed in H3K27ac occupancy upon *Arid1a* loss were largely composed of sites at distal regions (Figure 17J), which is consistent with the disruption of canonical BAF complexes (*124*, *125*, *133*). Sites that increased in H3K27ac upon *Arid1a* loss included *Nrg4* and *Egfr*, consistent with transcriptional hyperactivation in the *Arid1a*-deficient setting (Figure 17K). GO biological processes associated with increased acetylation of H3K27 in the *Arid1a* knockout populations in clusters 3 and 4 contain terms involving glucocorticoid biosynthesis and hormone metabolism, and the associated mouse phenotypes included abnormal insulin secretion and pancreas secretion (Figure 18A and B). These observations are consistent with differential regulation of pathways affecting glucose metabolism and insulin secretion in KO mouse β-cells.

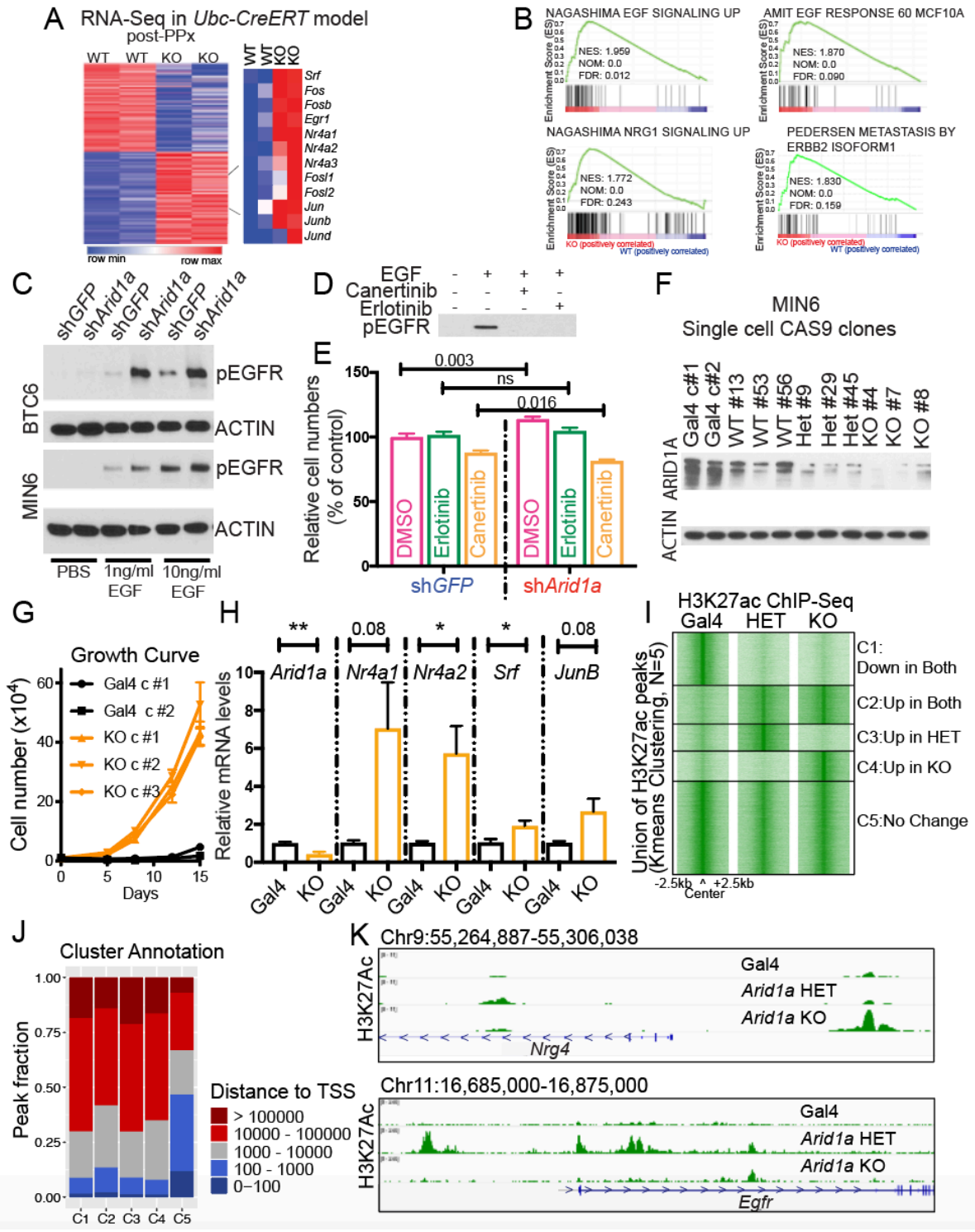
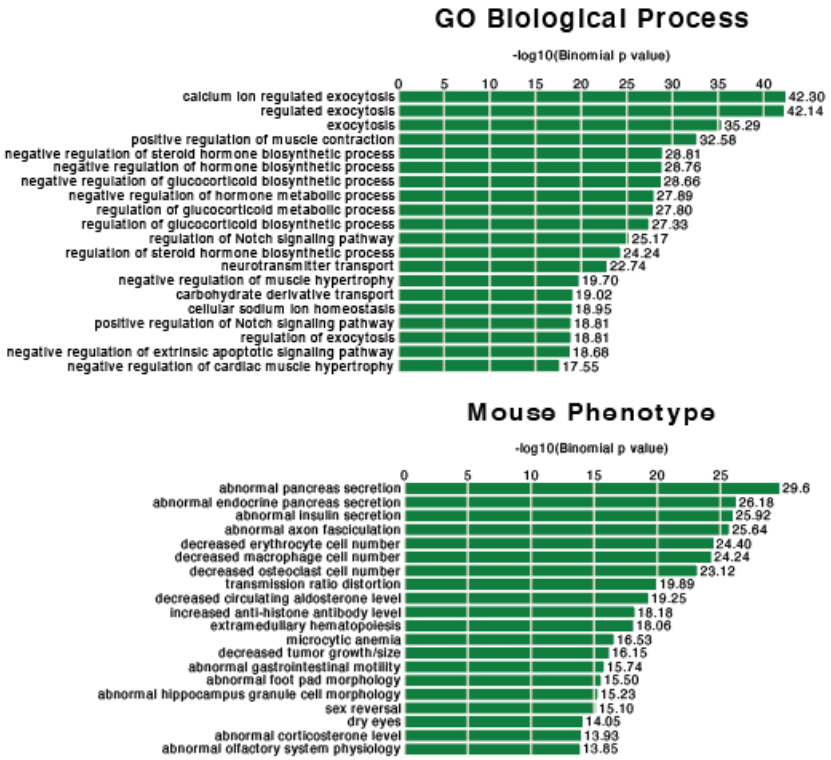

A
RNA-Seq in Ubc-CreERT model
post-PPx
WT WT KO KO
Srf
Fos
Fosb
Egr1
Nr4a1
Nr4a2
Nr4a3
Fosl1
Fosl2
Jun
Junb
Jund
row min
row max
B
NAGASHIMA EGF SIGNALING UP
NES: 1.959
NOM: 0.0
FDR: 0.012
AMIT EGF RESPONSE 60 MCF10A
NES: 1.870
NOM: 0.0
FDR: 0.090
NAGASHIMA NRG1 SIGNALING UP
NES: 1.772
NOM: 0.0
FDR: 0.243
PEDERSEN METASTASIS BY ERBB2 ISOFORM1
NES: 1.830
NOM: 0.0
FDR: 0.159
C
shGFP shArid1a shGFP shArid1a shGFP shArid1a
BTC6
pEGFR
ACTIN
MIN6
pEGFR
ACTIN
PBS
1ng/ml EGF
10ng/ml EGF
D
EGF
Canertinib
Erlotinib
pEGFR
E
Relative cell numbers (% of control)
0.003
ns
0.016
DMSO
Erlotinib
Canertinib
shGFP
shArid1a
F
MIN6
Single cell CAS9 clones
Gal4 c#1
Gal4 c#2
WT #13
WT #53
WT #56
Het #9
Het #29
Het #45
KO #4
KO #7
KO #8
ARID1A
ACTIN
G
Growth Curve
Cell number (x10^4)
Gal4 c #1
Gal4 c #2
KO c #1
KO c #2
KO c #3
Days
H
Relative mRNA levels
Arid1a
Nr4a1
Nr4a2
Srf
JunB
Gal4
KO
I
H3K27ac ChIP-Seq
Gal4 HET KO
Union of H3K27ac peaks (Kmeans Clustering, N=5)
C1: Down in Both
C2:Up in Both
C3:Up in HET
C4:Up in KO
C5:No Change
-2.5kb ^ +2.5kb
Center
J
Cluster Annotation
Peak fraction
C1 C2 C3 C4 C5
Distance to TSS
> 100000
10000 - 100000
1000 - 10000
100 - 1000
0–100
K
Chr9:55,264,887-55,306,038
H3K27Ac
Gal4
Arid1a HET
Arid1a KO
Nrg4
Chr11:16,685,000-16,875,000
H3K27Ac
Gal4
Arid1a HET
Arid1a KO
Egfr

Figure 17. *Arid1a* deficient β-cells have a dependence on increased EGF/Neuregulin signaling.

(A) Heatmap of differentially expressed genes in WT and KO islets from *Arid1a*$^{Fl/Fl}$ and *Ubc-CreER; Arid1a*$^{Fl/Fl}$ mice, isolated after PPx and detected by RNA-Seq (left). Heatmap showing a subset of overexpressed genes in KO islets (right; n = 2 and 2 mice). (B) GSEA shows that EGF and NRG1 response genes are upregulated in KO islets. The nominal enrichment score (NES), nominal p-value, and false discovery rate (FDR) q-value are shown within each GSEA plot. (C) p-EGFR western blots in MIN6 and BTC6 treated with different doses of EGF. (D) p-EGFR western blots in WT MIN6 cells treated with EGF in the presence or absence of erlotinib and canertinib. (E) Relative cell numbers for control *shGFP* and *shArid1a* MIN6 cells in the presence of canertinib and erlotinib. Shown as % of control cells (*shGFP* group treated with DMSO). (F) Western blot showing ARID1A levels in MIN6 clones. MIN6 cells were transduced with lenti-*CAS9*-blasticidin and either non-targeting lenti-sgRNA (*Gal4*)-puromycin or lenti-sgRNA (*Arid1a*)-puromycin. (G) MIN6 clone growth over 15 days measured by cell counting. (H) mRNA expression of selected genes in *Gal4* control and *Arid1a* KO MIN6 clones as measured by qPCR. (I) K-mean clustering of H3K27Ac ChIP-Seq peaks in non-targeting *Gal4*, *Arid1a* heterozygous and *Arid1a* KO MIN6 clones. (J) Annotation of clusters defined by H3K27Ac ChIP-Seq sites by distance to TSS. (K) Sample tracks for H3K27Ac ChIP-Seq in control, *Arid1a* heterozygous and *Arid1a* KO MIN6 single cell clones at the *Nrg4* and *Egfr* loci.

Cis-Regulatory Function Analysis by GREAT on Cluster 3

A



B

Cis-Regulatory Function Analysis by GREAT on Cluster 4

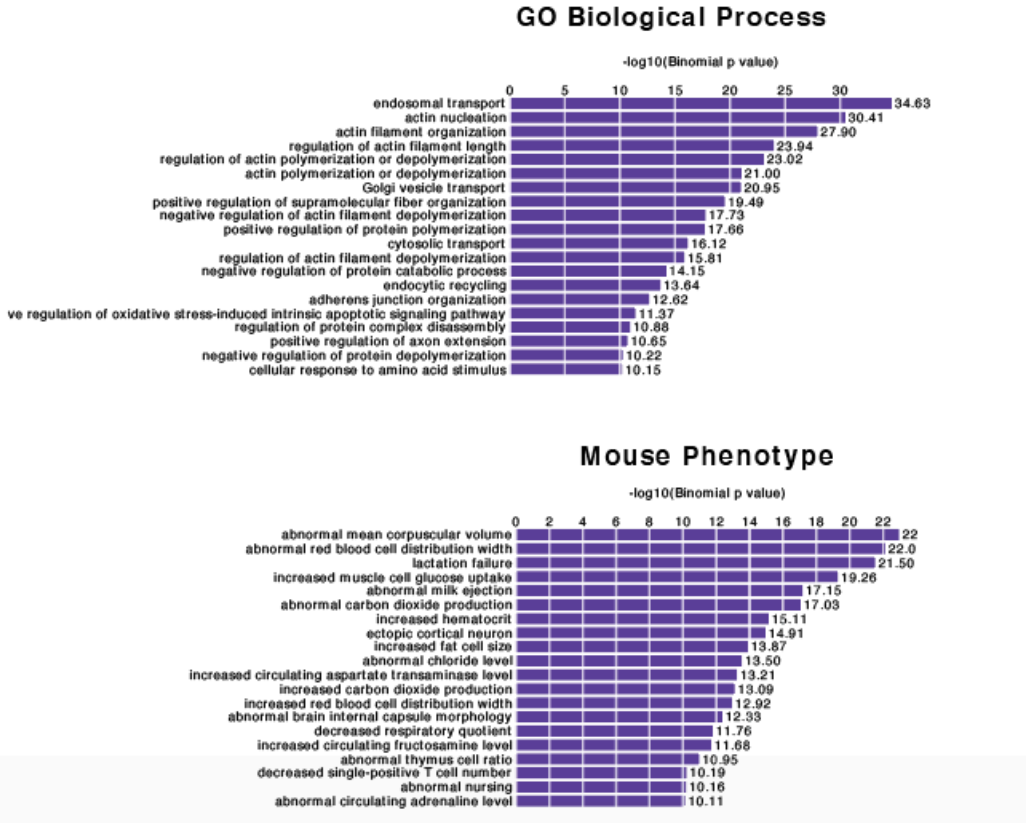
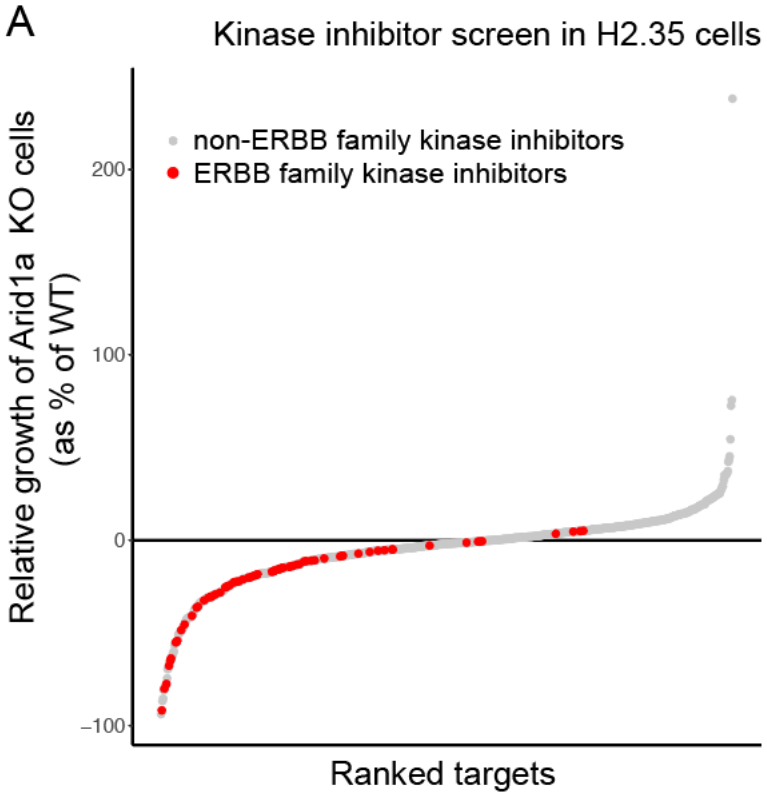


Figure 18. Cis-regulatory function analysis by GREAT for H3K27Ac peaks.

(A) Cis-regulatory function analysis by GREAT on sites within Cluster 3 (increased H3K27Ac in *Arid1a* HET condition) given as GO processes and mouse phenotypes. (B) Cis-regulatory function analysis by GREAT on sites within Cluster 4 (increased H3K27Ac in *Arid1a* KO condition) given as GO processes and mouse phenotypes.

To challenge the idea that *Arid1a* deletion causes more proliferation through the preferential promotion of ERBB signaling as opposed to a more generalized activation of multiple mitogenic signal pathways, we performed a small molecule screen to identify pathways that *Arid1a* KO cells are more dependent on for growth and survival. We screened 300 kinase inhibitors on previously generated immortalized H2.35 cells isogenic for *Arid1a* deletion (Figure 19A). Interestingly, ERBB family inhibitors were significantly enriched among treatments that induced the most prominent reductions in cell viability (P-value = 5.8e-6). There were 3 ERBB family inhibitors (Afatinib, WZ8040, canertinib) among the top 10 kinase inhibitors associated with the highest reduction of growth/viability in *Arid1a* KO cells (Figure 19B). Together with the islet transcriptome data, these observations established a strong functional connection between ARID1A and ERBB.

A



B

Top 10 inhibitors causing synthetic lethality in *Arid1a* KO cells

Inhibitor name	Target kinase	Relative growth
PHA-665752	c-MET	-94 %
Afatinib (BIBW2992)	EGFR, ERBB2	-92 %
Ponatinib (AP24534)	BCR-ABL	-87 %
R935788	SYK	-85 %
R406	SYK	-81 %
WZ8040	EGFR	-80 %
LDN193189	ALK1, ALK2, ALK3, ALK6	-79 %
R788 (Fostamatinib)	SYK	-78 %
CI-1033 (Canertinib)	EGFR, ERBB2, ERBB4	-77 %
YM201636	PIP5KIII	-75 %

Figure 19. Kinase inhibitor screen in H2.35 cells shows that *Arid1a* deficient cells are more sensitive to cell growth inhibition induced by ERBB inhibitors than WT cells.

(A) Relative cell growth (KO vs WT). About 300 kinase inhibitors were used in this screen. Each gray dot corresponds to relative growth of *Arid1a* KO vs. WT cells in the presence of a non-ERBB kinase inhibitor, whereas each red dot corresponds to an ERBB family kinase inhibitor. (B) Table showing the top 10 kinase inhibitors that caused the highest reduction of growth/viability selectively in *Arid1a* KO cells. 3 out of 10 top kinase inhibitors targeted the ERBB family of RTKs.

Pan-ERBB and NR4A1 inhibition ablated the Arid1a KO phenotype in vivo

We sought to determine whether pro-proliferative effects of *Arid1a* loss were mediated through ERBB activation in vivo as well as in vitro. We first tested the in vivo relevance of these results by examining whole body *Ubc-CreER; Arid1a* WT and KO mice described earlier. After STZ mediated islet ablation, mice were given daily doses of canertinib, an intervention that abolished the anti-diabetic effects of *Arid1a* loss (Figure 20A, compare to Figure 13F which did not include canertinib). We reasoned that ERBB inhibition in the STZ model could have been influenced by cell death in addition to proliferation, so we also performed the PPx assay to assess β-cell proliferation after surgical injury. Here, β-cell specific *Arid1a* deletion was induced one week before PPx and drug treatments were started one day before and continued until 6 days after PPx, which is the apex for β-cell proliferation (*145*). The increase in Ki-67/insulin double positive β-cells became more pronounced in the

regenerating pancreas (Figure 20B, top row & Figure 20C). Again, canertinib abolished the pro-proliferative effect seen in KO islets (Figure 20B, middle row & Figure 20C). These results showed that ERBB signaling in part mediates the β-cell regenerating effects of *Arid1a* deficiency in vivo.

Because ARID1A likely influences a large network of genes that regulate differentiation and regenerative capacity, we next asked if other genes downstream of ERBB signaling exert the effects of *Arid1a* loss. To further examine important functional networks resulting *Arid1a* KO phenotype, a protein-protein interaction network of overexpressed genes from the RNA-seq data was created using the STRING database (*146*). This interaction network showed that the SRF/JUN/FOS/EGR1 transcription factors and the NR4A family of nuclear receptor transcription factors were tightly connected and clustered together (Figure 20D). Indeed, these genes also came up under the GSEA datasets that were classified as "EGF/NRG1 signaling up" (Figure 17A) suggesting that they are downstream of ERBB signaling. The overexpression of these transcription factors were previously established as one of the hallmarks of immature, proliferative β-cells (*147*). We validated higher protein levels of c-FOS and p-NR4A1 in KO islets (Figure 20E). We then asked if ARID1A relays its effects in part through NR4A1 in vivo by treating mice that had undergone PPx with an NR4A1 antagonist (C-DIM8), in a similar fashion as the canertinib experiments. C-DIM8 also abolished the increased proliferation normally seen in *Arid1a* KO β-cells (Figure 20B, bottom row & Figure 20C), suggesting that increased NR4A1 activity is additionally required for the phenotypic effects of *Arid1a* loss.

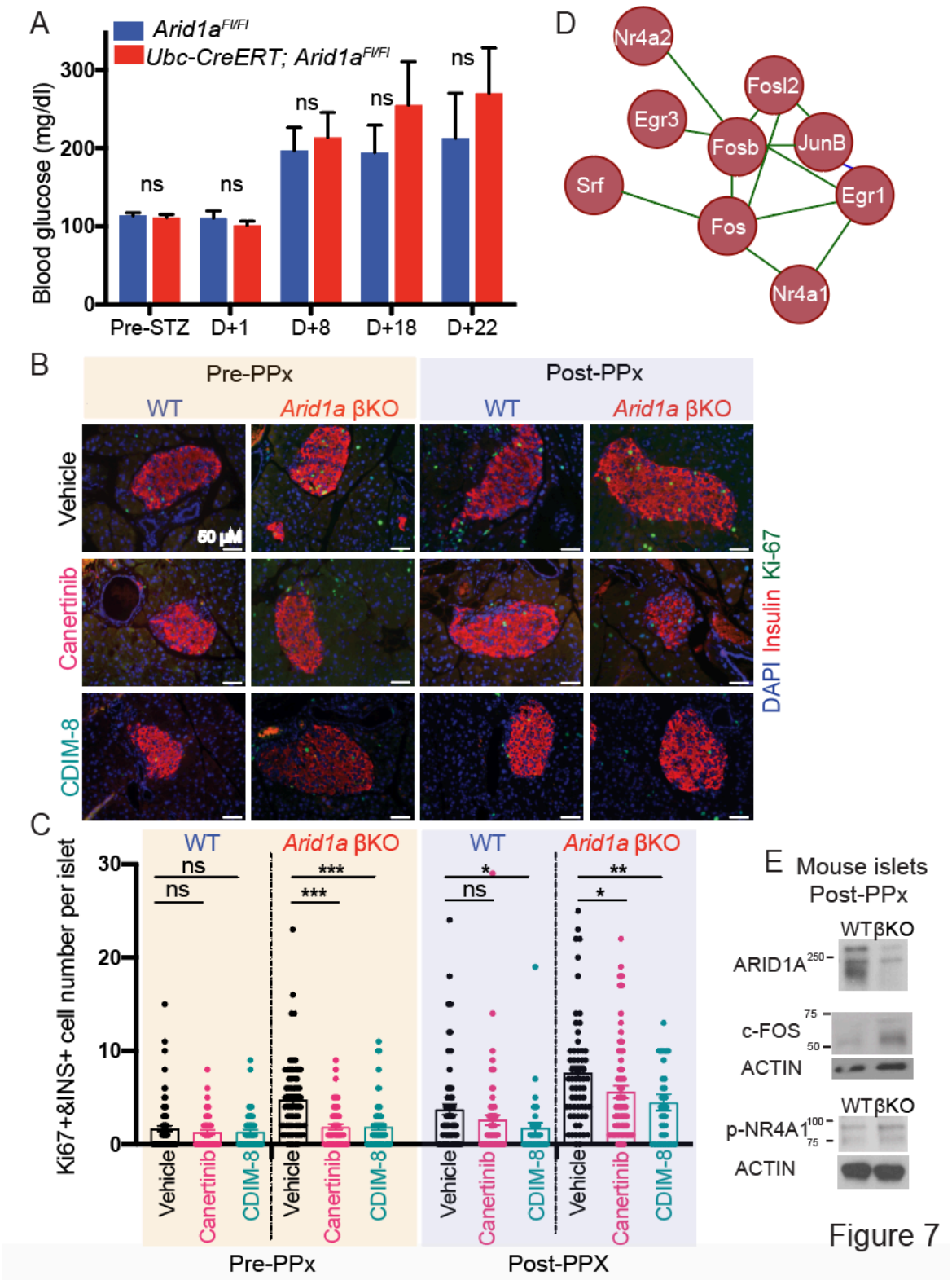

A
Arid1a^Fl/Fl
Ubc-CreERT; Arid1a^Fl/Fl
Blood glucose (mg/dl)
300
200
100
0
ns
ns
ns
ns
ns
Pre-STZ
D+1
D+8
D+18
D+22
D
Nr4a2
Fosl2
Egr3
Fosb
JunB
Srf
Egr1
Fos
Nr4a1
B
Pre-PPx
Post-PPx
WT
Arid1a βKO
WT
Arid1a βKO
Vehicle
Canertinib
CDIM-8
50 μM
DAPI Insulin Ki-67
C
Ki67+&INS+ cell number per islet
30
20
10
0
WT
Arid1a βKO
WT
Arid1a βKO
ns
ns
***
***
*
ns
**
*
Vehicle
Canertinib
CDIM-8
Pre-PPx
Post-PPX
E
Mouse islets
Post-PPx
WTβKO
ARID1A 250
c-FOS 75 50
ACTIN
WTβKO
p-NR4A1 100 75
ACTIN

Figure 7

Figure 20. Pan-ERBB and NR4A1 inhibition abrogated the *Arid1a* KO phenotype in vivo.

(A) Fed state blood glucose measurements following STZ in WT and whole body *Arid1a* KO mice. 20mg/kg canertinib was administered daily through oral gavage. Compare to Fig. 2F, which did not include canertinib. (B) Immunofluorescence for DAPI (blue), insulin (red), and Ki-67 (green) in WT and *Arid1a* βKO islets before or after PPx in the presence of vehicle, canertinib, or CDIM-8 treatment. (C) Quantification of Ki-67 and insulin double positive cell number per islet. Each dot represents an islet. Between 37-75 islets per genotype were counted from 5 to 8 mice per genotype. (D) The STRING database predicts an important protein-protein interaction network from differentially overexpressed genes found in RNA-seq data. (E) Western blot analysis of c-FOS and p-NR4A1 in mouse islets isolated 6 day post-PPx.

CONCLUDING REMARKS & FUTURE DIRECTIONS

Preserving or replenishing endogenous pancreatic β-cells is a major challenge in diabetes. To ultimately meet these goals, there is a need to understand the factors limiting β-cell proliferation. Although engineering strategies such as inducible pluripotent stem cell to β-cell conversion or the transdifferentiation of other cell types to β-cells are emerging (*148*), targeting endogenous regeneration of β-cells has potential advantages since this occurs in development, pregnancy, and obesity to meet increased physiological demands for insulin production. Here, we have identified *Arid1a* as a factor whose suppression is required for gestational β-cell expansion and is sufficient for β-cell replenishment after multiple pancreatic injuries. Previous

studies have shown that BRG1 and BRM-containing SWI/SNF complexes have context-dependent roles in modulating PDX1 activity in β-cell development and function (*60*, *61*). Wei *et. al.* reported that Vitamin D ligand causes the Vitamin D receptor to interact with an active BRD7/pBAF rather than an inactive BRD9/ncBAF complex, which reduces β-cell failure and slows diabetes progression (*149*). Our study is the first to highlight *Arid1a* and SWI/SNF as an important epigenetic node that regulates the regenerative capacity of β-cells.

By functionally validating EGF/NRG1 signaling as a critical effector of ARID1A suppression, we connected two previously unrelated, but prominent regeneration networks. EGF signaling is known to be important for liver, heart, and pancreas regeneration (*34*, *59*, *150*, *151*).

Mechanistically, our results implicate a network of EGF/NRG1 signaling genes. These include *Nr4a* nuclear receptor genes and a set of nutrient-responsive immediate early genes (*JunB*, *Fos*, *Egr1*) and their upstream activator *Srf* (*152*). Interestingly, EGR binding sites represented the most dynamic chromatin regions during whole body regeneration of *Hofstenia miamia*, an acoel worm that is capable of whole body regeneration. In these worms, EGR is essential for regeneration through its activities as a pioneer factor that directly regulated wound-induced genes including the EGFR ligands *nrg-1* and *nrg-2* (*153*). In addition, Zeng et al. recently used single cell transcriptomics to categorize β-cells from less to most mature. They showed that *Srf*/*JunB*/*Fos*/*Egr1* were among the most downregulated genes during postnatal β-cell maturation and comprise a signature of less mature, proliferative β-cells. Importantly, overexpression of upstream activator *Srf* in islets did not impair insulin secretion (*147*). Our data suggest that ARID1A containing complexes may negatively regulate this larger network of pro-regeneration genes. Importantly, β-cells without ARID1A are poised for proliferation while still able to maintain insulin production.

Our study highlights targets that could be investigated for their therapeutic potential in future studies. In contrast to the majority of nuclear receptors, the NR4A family of nuclear receptors do not have an identified physiological ligand, therefore their activity is mainly regulated at the level of protein expression and post-translational modifications (*154*). However, there is a agonist, Cytosporone-B, that can selectively increase the transcriptional activity of *Nr4a*1 (*155*). Whether this agonist could be utilized to improve the regenerative potential of β-cells is an area of interest. A more direct way to tackle this problem would be to use specific inhibitors targeting ARID1A or BAF complexes. There is a major interest in the field to develop small molecules to manipulate activities of SWI/SNF complex, and if successful, it will be interesting to determine if such inhibitors might be effective in diabetes.

BIBLIOGRAPHY

1. S. Khorasanizadeh, The nucleosome: from genomic organization to genomic regulation. *Cell*. **116**, 259–272 (2004).
2. R. D. Kornberg, Structure of chromatin. *Annu. Rev. Biochem.* **46**, 931–954 (1977).
3. C. L. Woodcock, J. P. Safer, J. E. Stanchfield, Structural repeating units in chromatin. I. Evidence for their general occurrence. *Exp. Cell Res.* **97**, 101–110 (1976).
4. K. Luger, A. W. Mäder, R. K. Richmond, D. F. Sargent, T. J. Richmond, Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*. **389**, 251–260 (1997).
5. A. M. Valencia, C. Kadoch, Chromatin regulatory mechanisms and therapeutic opportunities in cancer. *Nat. Cell Biol.* **21**, 152–161 (2019).
6. C. R. Clapier, J. Iwasa, B. R. Cairns, C. L. Peterson, Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat. Rev. Mol. Cell Biol.* **18**, 407–422 (2017).
7. C. Kadoch, G. R. Crabtree, Mammalian SWI/SNF chromatin remodeling complexes and cancer: Mechanistic insights gained from human genomics. *Sci. Adv.* **1**, e1500447 (2015).
8. M. Carlson, B. C. Osmond, D. Botstein, Mutants of yeast defective in sucrose utilization. *Genetics*. **98**, 25–40 (1981).
9. L. Neigeborn, M. Carlson, Genes affecting the regulation of SUC2 gene expression by glucose repression in Saccharomyces cerevisiae. *Genetics*. **108**, 845–858 (1984).
10. M. Stern, R. Jensen, I. Herskowitz, Five SWI genes are required for expression of the HO gene in yeast. *J. Mol. Biol.* **178**, 853–868 (1984).
11. A. K. Dingwall *et al.*, The Drosophila snr1 and brm proteins are related to yeast SWI/SNF proteins and are components of a large protein complex. *Mol. Biol. Cell*. **6**,

777–791 (1995).

12. J. W. Tamkun *et al.*, brahma: a regulator of Drosophila homeotic genes structurally related to the yeast transcriptional activator SNF2/SWI2. *Cell*. **68**, 561–572 (1992).

13. C. Kadoch *et al.*, Dynamics of BAF-Polycomb complex opposition on heterochromatin in normal and oncogenic states. *Nat. Genet.* **49**, 213–222 (2017).

14. N. Mashtalir *et al.*, Modular organization and assembly of SWI/SNF family chromatin remodeling complexes. *Cell*. **175**, 1272–1288.e20 (2018).

15. G. W. E. Santen *et al.*, Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. *Nat. Genet.* **44**, 379–380 (2012).

16. Y. Tsurusaki *et al.*, Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. *Nat. Genet.* **44**, 376–378 (2012).

17. F. Gao *et al.*, Heterozygous mutations in SMARCA2 reprogram the enhancer landscape by global retargeting of SMARCA4. *Mol. Cell*. **75**, 891–904.e7 (2019).

18. D. Wolff *et al.*, In-Frame Deletion and Missense Mutations of the C-Terminal Helicase Domain of SMARCA2 in Three Patients with Nicolaides-Baraitser Syndrome. *Mol. Syndromol.* **2**, 237–244 (2012).

19. S. S. Vergano, M. A. Deardorff, Clinical features, diagnostic criteria, and management of Coffin-Siris syndrome. *Am. J. Med. Genet. C Semin. Med. Genet.* **166C**, 252–256 (2014).

20. B. J. O'Roak *et al.*, Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science*. **338**, 1619–1622 (2012).

21. Deciphering Developmental Disorders Study, Large-scale discovery of novel genetic causes of developmental disorders. *Nature*. **519**, 223–228 (2015).

22. J. L. Ronan, W. Wu, G. R. Crabtree, From neural development to cognition: unexpected roles for chromatin. *Nat. Rev. Genet.* **14**, 347–359 (2013).

23. C. Kadoch *et al.*, Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. *Nat. Genet.* **45**, 592–601 (2013).

24. H. Lickert *et al.*, Baf60c is essential for function of BAF chromatin remodelling complexes in heart development. *Nature*. **432**, 107–112 (2004).

25. J. Lessard *et al.*, An essential switch in subunit composition of a chromatin remodeling complex during neural development. *Neuron*. **55**, 201–215 (2007).

26. S. Bultman *et al.*, A Brg1 null mutation in the mouse reveals functional differences among mammalian SWI/SNF complexes. *Mol. Cell*. **6**, 1287–1295 (2000).

27. A. Klochendler-Yeivin *et al.*, The murine SNF5/INI1 chromatin remodeling factor is essential for embryonic development and tumor suppression. *EMBO Rep.* **1**, 500–506 (2000).

28. J. K. Kim *et al.*, Srg3, a mouse homolog of yeast SWI3, is essential for early embryogenesis and involved in brain development. *Mol. Cell. Biol.* **21**, 7787–7795 (2001).

29. X. Gao *et al.*, ES cell pluripotency and germ-layer formation require the SWI/SNF chromatin remodeling component BAF250a. *Proc. Natl. Acad. Sci. USA*. **105**, 6656–6661 (2008).

30. C. Celen *et al.*, Arid1b haploinsufficient mice reveal neuropsychiatric phenotypes and reversible causes of growth impairment. *Elife*. **6** (2017), doi:10.7554/eLife.25730.

31. N. G. Nagl *et al.*, The p270 (ARID1A/SMARCF1) subunit of mammalian SWI/SNF-related complexes is essential for normal cell cycle arrest. *Cancer Res.* **65**, 9236–9244 (2005).

32. Z. Yan *et al.*, BAF250B-associated SWI/SNF chromatin-remodeling complex is required to maintain undifferentiated mouse embryonic stem cells. *Stem Cells*. **26**,

1155–1165 (2008).

33. N. G. Nagl, X. Wang, A. Patsialou, M. Van Scoy, E. Moran, Distinct mammalian SWI/SNF chromatin remodeling complexes with opposing roles in cell-cycle control. *EMBO J.* **26**, 752–763 (2007).

34. X. Sun *et al.*, Suppression of the SWI/SNF component arid1a promotes mammalian regeneration. *Cell Stem Cell*. **18**, 456–466 (2016).

35. C. Celen *et al.*, Arid1a loss potentiates pancreatic β-cell regeneration through activation of EGF signaling. *BioRxiv* (2020), doi:10.1101/2020.02.10.942615.

36. Y. Dor, J. Brown, O. I. Martinez, D. A. Melton, Adult pancreatic beta-cells are formed by self-duplication rather than stem-cell differentiation. *Nature*. **429**, 41–46 (2004).

37. F. Thorel *et al.*, Conversion of adult pancreatic alpha-cells to beta-cells after extreme beta-cell loss. *Nature*. **464**, 1149–1154 (2010).

38. K. A. D'Amour *et al.*, Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. *Nat. Biotechnol.* **24**, 1392–1401 (2006).

39. E. Kroon *et al.*, Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat. Biotechnol.* **26**, 443–452 (2008).

40. J. R. Millman *et al.*, Generation of stem cell-derived β-cells from patients with type 1 diabetes. *Nat. Commun.* **7**, 11463 (2016).

41. A. Rezania *et al.*, Reversal of diabetes with insulin-producing cells derived in vitro from human pluripotent stem cells. *Nat. Biotechnol.* **32**, 1121–1133 (2014).

42. Q. Zhou, J. Brown, A. Kanarek, J. Rajagopal, D. A. Melton, In vivo reprogramming of adult pancreatic exocrine cells to beta-cells. *Nature*. **455**, 627–632 (2008).

43. W. Li *et al.*, Long-term persistence and development of induced pancreatic beta cells generated by lineage conversion of acinar cells. *Nat. Biotechnol.* **32**, 1223–1230

(2014).

44. X. Xiao *et al.*, Endogenous Reprogramming of Alpha Cells into Beta Cells, Induced by Viral Gene Therapy, Reverses Autoimmune Diabetes. *Cell Stem Cell*. **22**, 78–90.e4 (2018).

45. N. Cerdá-Esteban *et al.*, Stepwise reprogramming of liver cells to a pancreas progenitor state by the transcriptional regulator Tgif2. *Nat. Commun.* **8**, 14127 (2017).

46. J. Lee *et al.*, Expansion and conversion of human pancreatic ductal cells into insulin-secreting endocrine cells. *Elife*. **2**, e00940 (2013).

47. Y.-J. Chen *et al.*, De novo formation of insulin-producing "neo-β cell islets" from intestinal crypts. *Cell Rep.* **6**, 1046–1058 (2014).

48. C. Ariyachet *et al.*, Reprogrammed stomach tissue as a renewable source of functional β cells for blood glucose regulation. *Cell Stem Cell*. **18**, 410–421 (2016).

49. S. Georgia, A. Bhushan, Beta cell replication is the primary mechanism for maintaining postnatal beta cell mass. *J. Clin. Invest.* **114**, 963–968 (2004).

50. M. Teta, S. Y. Long, L. M. Wartschow, M. M. Rankin, J. A. Kushner, Very slow turnover of beta-cells in aged adult mice. *Diabetes*. **54**, 2557–2567 (2005).

51. S. Rieck, K. H. Kaestner, Expansion of beta-cell mass in response to pregnancy. *Trends Endocrinol. Metab.* **21**, 151–158 (2010).

52. R. E. Mosser *et al.*, High-fat diet-induced β-cell proliferation occurs prior to insulin resistance in C57Bl/6J male mice. *Am. J. Physiol. Endocrinol. Metab.* **308**, E573-82 (2015).

53. J. Krishnamurthy *et al.*, p16INK4a induces an age-dependent decline in islet regenerative potential. *Nature*. **443**, 453–457 (2006).

54. H. Chen *et al.*, Polycomb protein Ezh2 regulates pancreatic beta-cell Ink4a/Arf expression and regeneration in diabetes mellitus. *Genes Dev.* **23**, 975–985 (2009).

55. T. T.-H. Lu *et al.*, The Polycomb-Dependent Epigenome Controls β Cell Dysfunction, Dedifferentiation, and Diabetes. *Cell Metab.* **27**, 1294–1308.e7 (2018).

56. H. Chen *et al.*, PDGF signalling controls age-dependent proliferation in pancreatic β-cells. *Nature*. **478**, 349–355 (2011).

57. E. S. M. Wong *et al.*, p38MAPK controls expression of multiple cell cycle inhibitors and islet proliferation with advancing age. *Dev. Cell*. **17**, 142–149 (2009).

58. P. J. Miettinen *et al.*, Downregulation of EGF receptor signaling in pancreatic islets causes diabetes due to impaired postnatal beta-cell growth. *Diabetes*. **55**, 3299–3308 (2006).

59. Z. Song *et al.*, Epidermal growth factor receptor signaling regulates β cell proliferation in adult mice. *J. Biol. Chem.* **291**, 22630–22637 (2016).

60. B. McKenna, M. Guo, A. Reynolds, M. Hara, R. Stein, Dynamic recruitment of functionally distinct Swi/Snf chromatin remodeling complexes modulates Pdx1 activity in islet β cells. *Cell Rep.* **10**, 2032–2042 (2015).

61. J. M. Spaeth *et al.*, The Pdx1-Bound Swi/Snf Chromatin Remodeling Complex Regulates Pancreatic Progenitor Cell Proliferation and Mature Islet β-Cell Function. *Diabetes*. **68**, 1806–1818 (2019).

62. N. A. DeCarolis *et al.*, 56Fe Particle Exposure Results in a Long-Lasting Increase in a Cellular Index of Genomic Instability and Transiently Suppresses Adult Hippocampal Neurogenesis in Vivo. *Life Sci Space Res (Amst)*. **2**, 70–79 (2014).

63. A. K. Walker *et al.*, The P7C3 class of neuroprotective compounds exerts antidepressant efficacy in mice by increasing hippocampal neurogenesis. *Mol. Psychiatry*. **20**, 500–508 (2015).

64. D. Kim, B. Langmead, S. L. Salzberg, HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods*. **12**, 357–360 (2015).

65. H. Li *et al.*, The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. **25**, 2078–2079 (2009).

66. J. Harrow *et al.*, GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res.* **22**, 1760–1774 (2012).

67. Y. Liao, G. K. Smyth, W. Shi, featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*. **30**, 923–930 (2014).

68. M. D. Robinson, D. J. McCarthy, G. K. Smyth, edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. **26**, 139–140 (2010).

69. D. J. Araujo *et al.*, FoxP1 orchestration of ASD-relevant signaling pathways in the striatum. *Genes Dev.* **29**, 2081–2096 (2015).

70. Y. Katayama *et al.*, CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature*. **537**, 675–679 (2016).

71. Y. Kawaguchi *et al.*, The role of the transcriptional regulator Ptf1a in converting intestinal to pancreatic progenitors. *Nat. Genet.* **32**, 128–134 (2002).

72. C. M. Kusminski *et al.*, MitoNEET-Parkin Effects in Pancreatic α- and β-Cells, Cellular Survival, and Intrainsular Cross Talk. *Diabetes*. **65**, 1534–1555 (2016).

73. A.-K. T. Perl, S. E. Wert, A. Nagy, C. G. Lobe, J. A. Whitsett, Early restriction of peripheral and proximal cell lineages during formation of the lung. *Proc. Natl. Acad. Sci. USA*. **99**, 10482–10487 (2002).

74. E. J. Zmuda, C. A. Powell, T. Hai, A method for murine islet isolation and subcapsular kidney transplantation. *J. Vis. Exp.* (2011), doi:10.3791/2096.

75. F. Martín *et al.*, Mechanisms of glucose hypersensitivity in beta-cells from normoglycemic, partially pancreatectomized mice. *Diabetes*. **48**, 1954–1961 (1999).

76. D. J. McCarthy, Y. Chen, G. K. Smyth, Differential expression analysis of multifactor

RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* **40**, 4288–4297 (2012).

77. J. D. Buenrostro, B. Wu, H. Y. Chang, W. J. Greenleaf, *Curr. Protoc. Mol. Biol.*, in press, doi:10.1002/0471142727.mb2129s109.

78. M. Martin, Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j.* **17**, 10 (2011).

79. Y. Zhang *et al.*, Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **9**, R137 (2008).

80. R. Stark, G. Brown, DiffBind: differential binding analysis of ChIP-Seq peak data. *R package version* (2011).

81. C. S. Ross-Innes *et al.*, Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. *Nature*. **481**, 389–393 (2012).

82. G. Yu, L.-G. Wang, Q.-Y. He, ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. *Bioinformatics*. **31**, 2382–2383 (2015).

83. B. Langmead, S. L. Salzberg, Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. **9**, 357–359 (2012).

84. A. R. Quinlan, I. M. Hall, BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. **26**, 841–842 (2010).

85. L. Shen, N. Shao, X. Liu, E. Nestler, ngs.plot: Quick mining and visualization of next-generation sequencing data by integrating genomic databases. *BMC Genomics*. **15**, 284 (2014).

86. E. P. Argiras, C. R. Blakeley, M. S. Dunnill, S. Otremski, M. K. Sykes, High PEEP decreases hyaline membrane formation in surfactant deficient lungs. *Br. J. Anaesth.* **59**, 1278–1285 (1987).

87. L. Ho, G. R. Crabtree, Chromatin remodelling during development. *Nature*. **463**, 474–

484 (2010).

88. D. Wieczorek *et al.*, A comprehensive molecular study on Coffin-Siris and Nicolaides-Baraitser syndromes identifies a broad molecular and clinical spectrum converging on altered chromatin remodeling. *Hum. Mol. Genet.* **22**, 5121–5135 (2013).

89. G. W. E. Santen *et al.*, Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. *Hum. Mutat.* **34**, 1519–1528 (2013).

90. D. Tropea *et al.*, Partial reversal of Rett Syndrome-like symptoms in MeCP2 mutant mice. *Proc. Natl. Acad. Sci. USA*. **106**, 2029–2034 (2009).

91. J. Castro *et al.*, Functional recovery with recombinant human IGF1 treatment in a mouse model of Rett Syndrome. *Proc. Natl. Acad. Sci. USA*. **111**, 9941–9946 (2014).

92. E. S. Lein *et al.*, Genome-wide atlas of gene expression in the adult mouse brain. *Nature*. **445**, 168–176 (2007).

93. M. Ka, D. A. Chopra, S. M. Dravid, W.-Y. Kim, Essential roles for ARID1B in dendritic arborization and spine morphology of developing pyramidal neurons. *J. Neurosci.* **36**, 2723–2742 (2016).

94. S. Schrier Vergano *et al.*, in *GeneReviews(®)*, R. A. Pagon et al., Eds. (University of Washington, Seattle, Seattle (WA), 1993).

95. J. L. Silverman, M. Yang, C. Lord, J. N. Crawley, Behavioural phenotyping assays for mouse models of autism. *Nat. Rev. Neurosci.* **11**, 490–502 (2010).

96. B. Greco *et al.*, Autism-related behavioral abnormalities in synapsin knockout mice. *Behav. Brain Res.* **251**, 65–74 (2013).

97. G. Konopka, T. F. Roberts, Animal models of speech and vocal communication deficits associated with psychiatric disorders. *Biol. Psychiatry*. **79**, 53–61 (2016).

98. L. Kanner, Autistic disturbances of affective contact. *Nervous child* (1943).

99. Y. S. Bonneh, Y. Levanon, O. Dean-Pardo, L. Lossos, Y. Adini, Abnormal speech spectrum and increased pitch variability in young autistic children. *Front. Hum. Neurosci.* **4**, 237 (2011).

100. A. Persson *et al.*, Decreased hippocampal volume and increased anxiety in a transgenic mouse model expressing the human CYP2C19 gene. *Mol. Psychiatry*. **19**, 733–741 (2014).

101. S. Travis *et al.*, Dentate gyrus volume and memory performance in major depressive disorder. *J. Affect. Disord.* **172**, 159–164 (2015).

102. M. Boldrini *et al.*, Hippocampal granule neuron number and dentate gyrus volume in antidepressant-treated and untreated major depression. *Neuropsychopharmacology*. **38**, 1068–1077 (2013).

103. L. Schmaal *et al.*, Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA Major Depressive Disorder Working Group. *Mol. Psychiatry*. **22**, 900–909 (2017).

104. S. N. Basu, R. Kollu, S. Banerjee-Basu, AutDB: a gene reference resource for autism research. *Nucleic Acids Res.* **37**, D832-6 (2009).

105. C. Attanasio *et al.*, Tissue-specific SMARCA4 binding at active and repressed regulatory elements during embryogenesis. *Genome Res.* **24**, 920–929 (2014).

106. H. Kletzl *et al.*, First-in-man study with a novel PEGylated recombinant human insulin-like growth factor-I. *Growth Horm IGF Res*. **33**, 9–16 (2017).

107. G. W. E. Santen, J. Clayton-Smith, ARID1B-CSS consortium, The ARID1B phenotype: what we have learned so far. *Am. J. Med. Genet. C Semin. Med. Genet.* **166C**, 276–289 (2014).

108. Y. Yu *et al.*, De novo mutations in ARID1B associated with both syndromic and non-syndromic short stature. *BMC Genomics*. **16**, 701 (2015).

109. A. Fujimoto *et al.*, Whole-genome sequencing of liver cancers identifies etiological influences on mutation patterns and recurrent mutations in chromatin regulators. *Nat. Genet.* **44**, 760–764 (2012).

110. J. Hoyer *et al.*, Haploinsufficiency of ARID1B, a member of the SWI/SNF-a chromatin-remodeling complex, is a frequent cause of intellectual disability. *Am. J. Hum. Genet.* **90**, 565–572 (2012).

111. C. Halgren *et al.*, Corpus callosum abnormalities, intellectual disability, speech impairment, and autism in patients with haploinsufficiency of ARID1B. *Clin. Genet.* **82**, 248–255 (2012).

112. J. C. H. Sim, S. M. White, P. J. Lockhart, ARID1B-mediated disorders: Mutations and possible mechanisms. *Intractable Rare Dis. Res.* **4**, 17–23 (2015).

113. T. Imai *et al.*, Dandy-Walker variant in Coffin-Siris syndrome. *Am. J. Med. Genet.* **100**, 152–155 (2001).

114. A. Brautbar, J. Ragsdale, M. Shinawi, Is this the Coffin-Siris syndrome or the BOD syndrome? *Am. J. Med. Genet. A*. **149A**, 559–562 (2009).

115. C. Jopling, S. Boue, J. C. Izpisua Belmonte, Dedifferentiation, transdifferentiation and reprogramming: three routes to regeneration. *Nat. Rev. Mol. Cell Biol.* **12**, 79–89 (2011).

116. Y. Yu *et al.*, Olig2 targets chromatin remodelers to enhancers to initiate oligodendrocyte differentiation. *Cell*. **152**, 248–261 (2013).

117. Z. Zhang *et al.*, Autism-Associated Chromatin Regulator Brg1/SmarcA4 Is Required for Synapse Development and Myocyte Enhancer Factor 2-Mediated Synapse Remodeling. *Mol. Cell. Biol.* **36**, 70–83 (2016).

118. L. Han *et al.*, Chromatin remodeling mediated by ARID1A is indispensable for normal hematopoiesis in mice. *Leukemia*. **33**, 2291–2305 (2019).

119. S. K. Hota *et al.*, Dynamic BAF chromatin remodeling complex subunit inclusion promotes temporally distinct gene expression programs in cardiogenesis. *Development*. **146** (2019), doi:10.1242/dev.174086.

120. W. Zhang *et al.*, The BAF and PRC2 complex subunits dpf2 and eed antagonistically converge on tbx3 to control ESC differentiation. *Cell Stem Cell*. **24**, 138–152.e8 (2019).

121. J. I. Wu, J. Lessard, G. R. Crabtree, Understanding the words of chromatin regulation. *Cell*. **136**, 200–206 (2009).

122. R. L. Chandler *et al.*, ARID1a-DNA interactions are required for promoter occupancy by SWI/SNF. *Mol. Cell. Biol.* **33**, 265–280 (2013).

123. X. Sun *et al.*, Arid1a Has Context-Dependent Oncogenic and Tumor Suppressor Functions in Liver Cancer. *Cancer Cell*. **33**, 151–152 (2018).

124. R. Mathur *et al.*, ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. *Nat. Genet.* **49**, 296–302 (2017).

125. T. W. R. Kelso *et al.*, Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. *Elife*. **6** (2017), doi:10.7554/eLife.30506.

126. R. T. Nakayama *et al.*, SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. *Nat. Genet.* **49**, 1613–1623 (2017).

127. J. Pan *et al.*, The ATPase module of mammalian SWI/SNF family complexes mediates subcomplex identity and catalytic activity-independent genomic targeting. *Nat. Genet.* **51**, 618–626 (2019).

128. Y. S. Oh, S. Shin, Y.-J. Lee, E. H. Kim, H.-S. Jun, Betacellulin-induced beta cell proliferation and regeneration is mediated by activation of ErbB-1 and ErbB-2 receptors. *PLoS One*. **6**, e23894 (2011).

129. A. G. Nikitin *et al.*, [Association of the polymorphisms of the ERBB3 and SH2B3 genes with type 1 diabetes]. *Mol. Biol. (Mosk)*. **44**, 257–262 (2010).

130. J. P. Bradfield *et al.*, A genome-wide meta-analysis of six type 1 diabetes cohorts identifies multiple associated loci. *PLoS Genet.* **7**, e1002293 (2011).

131. C. Sun *et al.*, ERBB3-rs2292239 as primary type 1 diabetes association locus among non-HLA genes in Chinese. *Meta Gene*. **9**, 120–123 (2016).

132. B. H. Alver *et al.*, The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. *Nat. Commun.* **8**, 14648 (2017).

133. T. Vierbuchen *et al.*, AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. *Mol. Cell*. **68**, 1067–1082.e12 (2017).

134. J. J. Meier *et al.*, Beta-cell replication is the primary mechanism subserving the postnatal expansion of beta-cell mass in humans. *Diabetes*. **57**, 1584–1594 (2008).

135. S. Rieck, K. H. Kaestner, Expansion of beta-cell mass in response to pregnancy. *Trends Endocrinol. Metab.* **21**, 151–158 (2010).

136. A. Ackermann Misfeldt, R. H. Costa, M. Gannon, Beta-cell proliferation, but not neogenesis, following 60% partial pancreatectomy is impaired in the absence of FoxM1. *Diabetes*. **57**, 3069–3077 (2008).

137. H. Alismail, S. Jin, Microenvironmental stimuli for proliferation of functional islet β-cells. *Cell Biosci.* **4**, 12 (2014).

138. D. D. De León *et al.*, Role of endogenous glucagon-like peptide-1 in islet regeneration after partial pancreatectomy. *Diabetes*. **52**, 365–371 (2003).

139. D. F. Gaddy, M. J. Riedel, S. Pejawar-Gaddy, T. J. Kieffer, P. D. Robbins, In vivo expression of HGF/NK1 and GLP-1 From dsAAV vectors enhances pancreatic ß-cell proliferation and improves pathology in the db/db mouse model of diabetes. *Diabetes*. **59**, 3108–3116 (2010).

140. C. Hajmrle *et al.*, Interleukin-1 signaling contributes to acute islet compensation. *JCI Insight*. **1**, e86055 (2016).

141. S. Porat *et al.*, Control of pancreatic β cell regeneration by glucose metabolism. *Cell Metab.* **13**, 440–449 (2011).

142. S. C. Wang *et al.*, SWI/SNF component ARID1A restrains pancreatic neoplasia formation. *Gut*. **68**, 1259–1270 (2019).

143. W. Wang *et al.*, ARID1A, a SWI/SNF subunit, is critical to acinar cell homeostasis and regeneration and is a barrier to transformation and epithelial-mesenchymal transition in the pancreas. *Gut*. **68**, 1245–1258 (2019).

144. Y. Yarden, M. X. Sliwkowski, Untangling the ErbB signalling network. *Nat. Rev. Mol. Cell Biol.* **2**, 127–137 (2001).

145. M. Peshavaria *et al.*, Regulation of pancreatic beta-cell regeneration in the normoglycemic 60% partial-pancreatectomy mouse. *Diabetes*. **55**, 3289–3298 (2006).

146. D. Szklarczyk *et al.*, STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* **43**, D447-52 (2015).

147. C. Zeng *et al.*, Pseudotemporal ordering of single cells reveals metabolic control of postnatal β cell proliferation. *Cell Metab.* **25**, 1160–1175.e11 (2017).

148. C. Aguayo-Mazzucato, S. Bonner-Weir, Pancreatic β cell regeneration as a possible therapy for diabetes. *Cell Metab.* **27**, 57–67 (2018).

149. Z. Wei *et al.*, Vitamin D switches BAF complexes to protect β cells. *Cell*. **173**, 1135–1149.e15 (2018).

150. M. Gemberling, R. Karra, A. L. Dickson, K. D. Poss, Nrg1 is an injury-induced cardiomyocyte mitogen for the endogenous heart regeneration program in zebrafish. *Elife*. **4** (2015), doi:10.7554/eLife.05871.

151. G. D'Uva *et al.*, ERBB2 triggers mammalian heart regeneration by promoting

cardiomyocyte dedifferentiation and proliferation. *Nat. Cell Biol.* **17**, 627–638 (2015).

152. M. Mina *et al.*, Promoter-level expression clustering identifies time development of transcriptional regulatory cascades initiated by ErbB receptors in breast cancer cells. *Sci. Rep.* **5**, 11999 (2015).

153. A. R. Gehrke *et al.*, Acoel genome reveals the regulatory landscape of whole-body regeneration. *Science*. **363** (2019), doi:10.1126/science.aau6173.

154. S. Safe *et al.*, Nuclear receptor 4A (NR4A) family - orphans no more. *J. Steroid Biochem. Mol. Biol.* **157**, 48–60 (2016).

155. Y. Zhan *et al.*, Cytosporone B is an agonist for nuclear orphan receptor Nur77. *Nat. Chem. Biol.* **4**, 548–556 (2008).