UC San Diego

UC San Diego Previously Published Works

Title

Heterozygous Mutations in SMARCA2 Reprogram the Enhancer Landscape by Global Retargeting of SMARCA4

Permalink

https://escholarship.org/uc/item/8326p3gx

Journal

Molecular Cell, 75(5)

ISSN

1097-2765

Authors

Gao, Fangjian
Elliott, Nicholas J
Ho, Josephine
et al.

Publication Date

2019-09-01

DOI

10.1016/j.molcel.2019.06.024

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2020 September 05.

Published in final edited form as:

Mol Cell. 2019 September 05; 75(5): 891–904.e7. doi:10.1016/j.molcel.2019.06.024.

Heterozygous mutations in SMARCA2 reprogram the enhancer landscape by global retargeting of SMARCA4

Fangjian Gao[1], **Nicholas J. Elliott**[1], **Josephine Ho**[1], **Alexzander Sharp**[2], **Maxim N. Shokhirev**[3], **Diana C. Hargreaves**[1,4,*]

[1]Molecular and Cellular Biology Laboratory, Salk Institute for Biological Studies, La Jolla, California 92037 USA

[2]University of California San Diego Biological Sciences Graduate Program, La Jolla, California 92037 USA

[3]Razavi Newman Integrative Genomics and Bioinformatics Core, Salk Institute for Biological Studies, La Jolla California 92037 USA

[4]Lead Contact

SUMMARY

Mammalian SWI/SNF complexes are multi-subunit chromatin remodeling complexes associated with an ATPase, either SMARCA4 or SMARCA2. Heterozygous mutations in the SMARCA2 ATPase cause Nicolaides-Baraitser Syndrome (NCBRS), an intellectual disability syndrome associated with delayed speech onset. We engineered human embryonic stem cells (hESCs) to carry NCBRS-associated heterozygous SMARCA2 K755R or R1159Q mutations. While SMARCA2 mutant hESCs were phenotypically normal, differentiation to neural progenitors cells (NPCs) was severely impaired. We find that SMARCA2 mutations cause enhancer reorganization with loss of SOX3-dependent neural enhancers and prominent emergence of astrocyte-specific *de novo* enhancers. Changes in chromatin accessibility at enhancers were associated with an increase in SMARCA2 binding and retargeting of SMARCA4. We show that AP-1 family member FRA2 is aberrantly overexpressed in SMARCA2 mutant NPCs, where it functions as a pioneer factor at *de novo* enhancers. Together, our results demonstrate SMARCA2 mutations cause impaired differentiation through enhancer reprogramming via inappropriate targeting of SMARCA4.

Graphical Abstract:

*To whom correspondence and material requests should be addressed: dhargreaves@salk.edu.

AUTHOR CONTRIBUTIONS

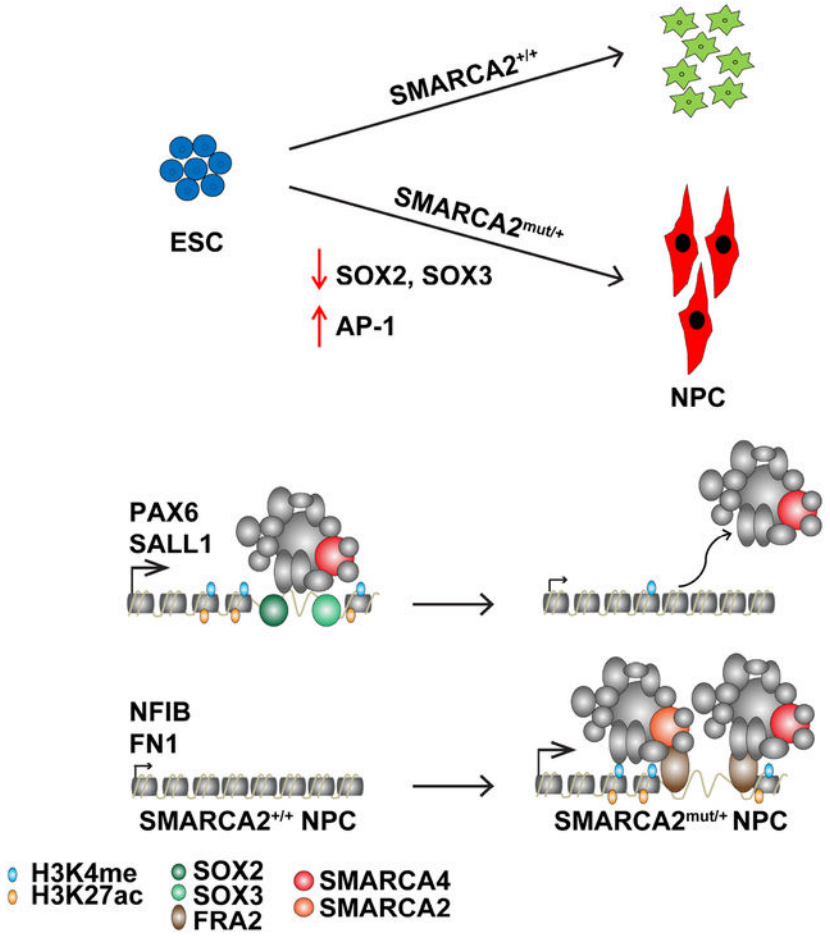
Conceptualization: F.G. and D.C.H.; Methodology: F.G. and D.C.H.; Investigation: F.G. and N.E., J.H.; Formal Analysis: F.G., N.E., D.C.H.; Validation: A.S.; Writing-Original Draft: F.G. and D.C.H.; Writing-Review and Editing: F.G. and D.C.H.; Software: M.N.S.; Project Administration: D.C.H.; Funding Acquisition: F.G. and D.C.H.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



eTOC Blurb(An eTOC blurb should also be included that is no longer than 50 words describing the context and significance of the findings for the broader journal readership. When writing this paragraph, please target it to non-specialists by highlighting the major conceptual point of the paper in plain language, without extensive experimental detail. The blurb must be written in the third person and refer to “First Author et al.”**):**

Mutations in subunits of the SWI/SNF complex cause developmental disorders associated with intellectual disability and speech delay. However, the mechanistic basis of disease is poorly understood. Gao et al. show that disease-associated SMARCA2 mutations lead to impaired neural differentiation due to changes in transcription factor expression that drive enhancer reprogramming.

INTRODUCTION

Mammalian SWI/SNF chromatin remodeling complexes play pivotal roles in regulating embryonic development and lineage determination. SWI/SNF complexes are composed of a ATPase, either SMARCA4 (BRG1) or SMARCA2 (BRM), and up to 15 associated subunits encoded from 29 genes (Hargreaves and Crabtree, 2011; Kadoch et al., 2013). Recent next-generation sequencing efforts have uncovered *de novo* mutations in subunits of the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SWI/SNF complex in individuals with Coffin-Siris syndrome (CSS) (Miyake et al., 2016; Santen et al., 2012; Santen et al., 2013; Tsurusaki et al., 2012; Zarate et al., 2016) and Nicolaides-Barraitser syndrome (NCBRS) (Tang et al., 2017; Van Houdt et al., 2012; Wolff et al., 2012), as well as in individuals with undefined developmental disorders (Deciphering Developmental Disorders, 2015; Wright et al., 2018) and autism (De Rubeis et al., 2014; Stessman et al., 2017; Yuen et al., 2017). These syndromes are characterized by intellectual disability and delayed or absent speech development, as well as various facial and skeletal abnormalities (Kosho and Okamoto, 2014; Pretegiani et al., 2016). Heterozygous missense mutations in the SMARCA2 ATPase domain are thought to be the underlying cause of NCBRS, and are now used as the genetic identifier for this syndrome (Sousa and Hennekam, 2014). However, it is not known how mutations in SWI/SNF subunit genes, and specifically missense mutations in SMARCA2, affect human neural differentiation.

During embryonic development, the SWI/SNF complex is combinatorially assembled, giving rise to cell type- and developmental stage-specific functions (Ho and Crabtree, 2010). During differentiation from embryonic stem cells (ESCs) to neural progenitor cells (NPCs), SMARCA4 expression gradually declines, while SMARCA2 is upregulated, and continues to be highly expressed in neurons (Lessard et al., 2007; Staahl et al., 2013; Wu et al., 2007). SMARCA4 is required for proliferation of neural stem cells (Lessard et al., 2007; Wu et al., 2007). In addition, conditional deletion of SMARCA4 in adult neural stem cells leads to abnormal neural differentiation, specifically through failure to regulate PAX6-dependent gene expression (Ninkovic et al., 2013). In contrast, heterozygous or complete deletion of SMARCA2 has no effect on the developing mouse, other than the fact that these mice are slightly bigger (Reyes et al., 1998). Thus, the phenotypes resulting from heterozygous SMARCA2 missense mutations in NCBRS individuals are not consistent with SMARCA2 loss-of-function.

At the molecular level, it has been suggested that ATPase mutants interfere with transcriptional activation mediated by SWI/SNF complexes. Early studies demonstrated that mutations in the SMARCA4 ATPase, namely K785R, result in reduced ATPase and chromatin remodeling activity without affecting protein stability or SWI/SNF complex assembly (Richmond and Peterson, 1996; Sif et al., 2001). Transformation of SMARCA4 K785R into swi2Δ or $SWI2^{+}$ yeast abolished SWI2-dependent transcription from the yeast HO locus (Khavari et al., 1993). Similarly, SMARCA4 K785R expression abolished SWI/SNF-dependent expression of GRE and EF-1a luciferase reporter constructs transfected into yeast or human cells, respectively (Khavari et al., 1993; Richmond and Peterson, 1996). These data collectively suggest a model in which mutations in the ATPase domain of the SMARCA4 homolog SMARCA2 could result in expression of a mutant protein that dominantly interferes with the ability of the wild-type ATPase to remodel chromatin and activate transcription. However, SMARCA2 mutants have not been characterized either in *in vitro* chromatin remodeling assays or in the context of *in vivo* differentiation.

To understand the molecular basis for SMARCA2 missense mutations, we developed a model system utilizing human ESC differentiation to NPCs of genetically engineered SMARCA2 mutant ESCs. We find that mutations in SMARCA2 disrupt neural differentiation, leading to loss of NPC self-renewal and loss of NPC marker expression. Loss

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of neural differentiation is accompanied by changes in gene expression, including loss of neural-specific factors *SOX2*, *SOX3*, and downstream targets *PAX6* and *SOX11*. Epigenomic analysis using Assay for Transposase Accessible Chromatin with high-throughput sequencing (ATAC-seq) demonstrated that SMARCA2 mutant NPCs have significant losses in chromatin accessibility at cell type-specific enhancers, primarily at SOX3 binding motifs, which is accompanied by loss of activating histone modifications and SOX3 binding. Surprisingly, loss of chromatin accessibility at these sites is correlated with loss of SMARCA4 binding and no change in already low levels of SMARCA2 binding. On the other hand, gains in co-bound SMARCA4 and SMARCA2 are correlated with *de novo* activation of distal enhancers associated with the astrocyte lineage. Mechanistically, we find that retargeting of SMARCA4 and SMARCA2 to *de novo* enhancers is driven by pioneering activity of AP-1 factors, including FRA2, which are overexpressed in SMARCA2 mutant NPCs. Our studies provide important insight into epigenetic control of neural differentiation, and challenge existing models on the molecular activity of mutant ATPases. This, in turn, has implications for intellectual disability disorders and cancer where ATPase missense mutations are prevalent in SMARCA2 and its homolog SMARCA4.

RESULTS

NCBRS-associated SMARCA2 mutations cause impaired neural progenitor differentiation

To determine how likely individual SMARCA2 missense mutations are to be pathogenic, we performed conservation analysis of the human SMARCA2 ATPase domain using ConSurf (Ashkenazy et al., 2010; Berezin et al., 2004), which scores individual amino acids from 1 (least) to 9 (most) conserved (Figure 1A). We mapped the mutations identified in NCBRS patients as well as single nucleotide variants (SNVs) from the normal population (Lek et al., 2016) and found that disease-associated missense mutations are significantly more enriched in highly conserved regions than normal SNVs, suggesting they are likely playing a causative role in producing the developmental defects observed in NCBRS individuals (Figure 1A, 1B). By mapping the NCBRS-associated mutations to the corresponding residue of the yeast SWI2 (Liu et al., 2017), we further found that these mutations all occur at residues that contact DNA (Figure 1C). These data suggest that NCBRS-associated mutations may affect SMARCA2 function in a similar way, potentially by impacting DNA binding or DNA-dependent ATPase activity.

To study how NCBRS-associated mutations affect neural development, we used CRISPR/Cas9 editing to introduce heterozygous K755R or R1159Q SMARCA2 mutations into the H9 hESC line, the former because it is known to abolish ATPase activity in yeast SWI2 (Richmond and Peterson, 1996; Sif et al., 2001) and the latter because it is the most common mutation in NCBRS. We confirmed heterozygosity of genomic DNA and cDNA of expressed transcripts (Figure S1A). We did not observe any morphological changes or differences in expression of the pluripotency markers OCT4, SOX2, or SWI/SNF subunits in $SMARCA2^{K755R/+}$ and $SMARCA2^{R1159Q/+}$ hESCs compared to WT hESCs (Figure S1B–D), indicating that SMARCA2 mutations do not grossly affect the hESC regulatory network. To study how SMARCA2 mutations affect neural differentiation, we differentiated hESCs to NPCs using an embryoid body (EB) formation protocol followed by neural rosette isolation

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Zhang et al., 2001). NPCs derived from *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ hESCs were larger than WT NPCs with less condensed nuclei and proliferated more slowly than WT NPCs with more cells in the G1 phase of the cell cycle (Figure 1D–F). We performed immunofluorescence for neural progenitor markers NESTIN, PAX6 and SOX2. While NESTIN levels were similar between WT and SMARCA2 mutant NPCs, significantly fewer *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs stained positive for PAX6 or SOX2 demonstrating that SMARCA2 mutant NPCs fail to express key transcription factors associated with NPC identity (Figure 1G–I). Consistent with our cell cycle analysis, we found that fewer SMARCA2 mutant NPCs stained positive for the proliferation marker Ki67 (Figure 1H, I). SMARCA2 mutant NPCs failed to differentiate into neurons, despite robust neuronal differentiation of WT NPCs, as indicated by Neuron-specific Class III β-tubulin (Tuj1) and microtubule-associated protein 2 (MAP2) staining and outgrowth of dendritic processes (Figure 1J, K). These data suggest SMARCA2 mutations affect NPC self-renewal and impair proper neural differentiation.

Expression of neural progenitor genes is dysregulated in SMARCA2 mutant NPCs

To gain insight into how SMARCA2 mutations affect neural gene expression on a global level, we performed RNA-seq on WT, *SMARCA2*$^{K755R/+}$, and *SMARCA2*$^{R1159Q/+}$ NPCs. Several hundred genes were differentially up and down regulated in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 2A). *SMARCA2*$^{R1159Q/+}$ NPCs exhibited a high degree of overlap with *SMARCA2*$^{K755R/+}$ NPCs, with 1244 and 526 genes up- and down-regulated in both mutants, respectively (Figure 2B, 2C). Among these differentially expressed genes (DEGs), several key regulators of neural identity were significantly downregulated in both SMARCA2 mutant NPC lines, including *SOX2*, *SOX3*, *SOX11*, *PAX6*, and *RFX4* (Figure 2C). We observed upregulation of extracellular matrix proteins such as *FN1* and *COL1A1*, the NFI family transcription factors *NFIB, NFIA*, and *NFIX* associated with the astrogliogenesis, as well as AP-1 transcription factors *FOSL1*, *FOSL2*, *JUNB*, and *JUND* in SMARCA2 mutant NPCs (Figure 2C). Gene Ontology (GO) analysis showed significant enrichment of genes involved in cell cycle, central nervous system development, and neurogenesis among downregulated DEGs, consistent with the cellular and mitotic phenotypes we observed (Figure 2D, 1D–K). Upregulated DEGs were enriched for extracellular matrix, extracellular space and structure, and tissue development terms (Figure 2D). These data demonstrate that NCBRS-associated mutations in SMARCA2 cause transcriptional changes that result in impaired neural differentiation due to downregulation of neural transcription factors and neurogenesis genes.

We also measured the expression of SWI/SNF subunits by RNA-seq and immunoblotting in SMARCA2 mutant NPCs. We found that several SWI/SNF subunits were reduced at the mRNA level, including SMARCA4, ARID1A, and SMARCD1, while SMARCA2 and SMARCD2/3 were upregulated (Figure 2E). SMARCD1 and ARID1A protein were also significantly reduced, but the protein expression of many other SWI/SNF subunits was not significantly affected (Figure 2F, G). SMARCA2, however, was robustly upregulated both at the mRNA and protein level (Figure 2E–G). Further, we found that SMARCA2 could be efficiently immunoprecipitated in SMARCA2 mutant NPCs, where it associates with known SWI/SNF complex subunits, but not SMARCA4, which is mutually-exclusive of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SMARCA2 in SWI/SNF complex assembly (Figure 2H). To compare the biochemical properties of the WT and SMARCA2 mutant proteins directly, we lentivirally expressed AM tagged-SMARCA2 WT, K755R, or R1159Q in NPCs and found that there was no difference in the ability of WT or mutant SMARCA2 to associate with SWI/SNF subunits (Figure S2A). These data suggest that missense mutations in SMARCA2 do not affect the ability of the mutant protein to be expressed or to assemble with SWI/SNF subunits, consistent with previous reports (Richmond and Peterson, 1996). However, changes in SWI/SNF subunit expression, including upregulation of SMARCA2, could contribute SMARCA2 mutant phenotypes. Thus, defects in neural differentiation in SMARCA2 mutant NPCs arise from transcriptional changes associated with the expression and incorporation of SMARCA2 mutant protein into SWI/SNF complexes.

SMARCA2 mutations lead to changes in chromatin accessibility correlated with SMARCA4 binding

A previous study showed that SMARCA4 K785R, which is analogous to SMARCA2 K755R, has reduced *in vitro* chromatin remodeling activity on nucleosomal templates (Sif et al., 2001). We thus hypothesized that chromatin remodeling activity would be compromised in cells expressing SMARCA2 ATPase mutations. We performed ATAC-seq in WT and SMARCA2 mutant NPCs and found that tens of thousands of sites were differentially regulated, with surprisingly roughly equal number of up- and down-regulated sites in both $SMARCA2^{K755R/+}$ and $SMARCA2^{R1159Q/+}$ NPCs (Figure 3A, B). We observed no significant changes between $SMARCA2^{K755R/+}$ and $SMARCA2^{R1159Q/+}$ NPCs, suggesting that these mutations have very similar effects on chromatin remodeling (Figure 3A). Thus, SMARCA2 mutations do not cause blanket inhibition of chromatin remodeling, but rather result in changes in accessibility reflecting widespread losses and gains.

To gain insight into how genomic occupancy of SMARCA2 and SMARCA4 correspond to alterations in chromatin accessibility, we performed ChIP-seq for SMARCA4 and SMARCA2 in WT, $SMARCA2^{K755R/+}$ and $SMARCA2^{R1159Q/+}$ NPCs. Surprisingly, we recovered <100 peaks for SMARCA2 in WT NPCs (Figure 3C, D). In contrast, we recovered 54,920 and 34,362 SMARCA2 peaks in $SMARCA2^{K755R/+}$ and $SMARCA2^{R1159Q/+}$ NPCs, respectively, suggesting a global gain in SMARCA2 binding to chromatin (Figure 3C). To test whether this gain in binding is the result of increased nuclear localization or DNA binding affinity of mutant SMARCA2, we performed cellular fractionation and salt extraction on WT and SMARCA2 mutant NPCs. We found that the cellular localization (Figure S3A, B) and DNA binding affinity (Figure S3C, D) of SWI/SNF subunits, including SMARCA2, was similar in WT and SMARCA2 mutant NPCs. Additionally, we found no difference in the ability of AM-SMARCA2 WT, K755R, or R1159Q to be extracted from DNA (Figure S2B, C). Thus, SMARCA2 mutant proteins do not have dramatically altered cellular distribution or DNA binding affinity that would account for the gain in SMARCA2 binding in SMARCA2 mutant NPCs.

In contrast to SMARCA2, we observed robust SMARCA4 binding at 49,280 sites in WT NPCs (Figure 3C). SMARCA4 binding was lost at 31,194/28,580 sites and gained at 70,582/73,359 sites in $SMARCA2^{K755R/+}$ and $SMARCA2^{R1159Q/+}$ NPCs relative to WT

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

NPCs, respectively (Figure 3C). Strikingly, loss of accessibility was correlated with a loss in SMARCA4 binding in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs with little to no change in already low levels of SMARCA2 binding (Figure 3D, E). On the other hand, gain of accessibility was correlated with gains in both SMARCA4 and SMARCA2 binding (Figure 3D, F). SMARCA4 and SMARCA2 binding were highly correlated at sites that gained accessibility, overlapping at 66.2% and 64.4% of the SMARCA2 binding sites in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 3G). Thus, SMARCA2 and SMARCA4 exhibit increased, coordinated, and functionally active binding to chromatin, which results in thousands of sites becoming more accessible. To compare the binding of WT and mutant SMARCA2, we performed AM-ChIP on *SMARCA2*$^{R1159Q/+}$ NPCs expressing AM-SMARCA2 WT or R1159Q and found that the binding of WT and mutant SMARCA2 is largely overlapping, with both localizing to sites that gain accessibility (Figure S2D, E). Thus, SMARCA2 retargeting in SMARCA2 mutant NPCs is not due to inherent differences in DNA binding affinity or catalytic activity allowing preferential ChIP fixation of mutant SMARCA2, as was observed in the case of ISWI (Gelbart et al., 2005). Finally, we annotated the affected sites and found that the vast majority of differentially accessible sites occur in intergenic or intronic regions (Figure 3H). These data demonstrate the surprising finding that changes in chromatin accessibility in SMARCA2 mutant NPCs are consistent with global retargeting of SMARCA4, and not loss-of-function or dominant-negative activity of mutant SMARCA2. In fact, while SMARCA2 DNA binding is dramatically increased in SMARCA2 mutant NPCs, this binding occurs specifically at sites where active chromatin remodeling is observed in the form of increased accessibility.

Neural-specific enhancers are lost in SMARCA2 mutant NPCs

We overlaid sites that either increased or decreased in accessibility in SMARCA2 mutant NPCs with ChIP-seq of histone modifications from WT NPCs (Figure 4A). Sites that lose chromatin accessibility are enriched for histone 3 lysine 4 monomethylation (H3K4me) and histone 3 lysine 27 acetylation (H3K27ac) found at poised, active, and super enhancers (Rada-Iglesias, 2018; Whyte et al., 2013), and to a lesser extent histone 3 lysine 4 trimethylation (H3K4me3) commonly found at promoters, but not at sites marked with histone 3 lysine 27 trimethylation (H3K27me3). However, there was little to no enrichment of any modification at sites that gained accessibility (Figure 4A). We then performed ChIP-seq for H3K4me and H3K27ac in SMARCA2 mutant NPCs. We found that neural-specific enhancers that lose accessibility also lose H3K4me and H3K27ac in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs, suggesting that these sites lose active histone modifications associated with enhancer activity (Figure 4B). Thus, sites that lose accessibility reflect a loss of open chromatin at neural regulatory elements, while gained sites occur at sites that lack poised and active histone marks.

Motif analysis on decreased accessible sites revealed an enrichment of motifs for SOX3, RFX1, HOXA2, MEIS1, and BRN1 (Figure 4C, D). Indeed, SOX3 motifs are present in nearly 50% of sites that lose accessibility in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 4E). To identify *bone fide* SOX3 binding sites in NPCs, we performed SOX3 ChIP-seq, as well as SOX2 ChIP-seq, as SOX2 was shown to prime SOX3 binding in mouse NPCs (Bergsland et al., 2011). We found that SOX2 and SOX3 are co-bound with each other

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and with SMARCA4 in NPCs (Figure 4F). Additionally, we found that SMARCA4 is bound to the vast majority of NPC-specific super enhancers, as well as super enhancers defined in brain tissue from the cingulate gyrus (Khan and Zhang, 2016), the majority of which are also bound by SOX2 and SOX3 (Figure 4G). To determine if SOX2 and SOX3 recruit SMARCA4, we knocked down SOX2 and SOX3 in NPCs by siRNA-mediated transfection (Figure S4A). Indeed, SMARCA4 binding was reduced at SOX3 binding sites in siSOX2/3 knockdown NPCs compared to NPCs transfected with a scrambled control (Figure 4H). Together, these data suggest that SOX2 and SOX3 are key transcription factors that function to recruit SMARCA4 to neural-specific regulatory elements in NPCs.

We then focused on the epigenetic changes at SOX3 ChIP-seq binding sites in SMARCA2 mutant NPCs. As a result of lower mRNA expression, SOX2 and SOX3 protein expression was reduced in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 4I), leading to a reduction in SOX2 and SOX3 binding in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs at SOX3 binding sites (Figure 4J, K). Consistent with a role for SOX2/3 in recruiting SMARCA4, we found that SMARCA4 binding and chromatin accessibility were dramatically reduced at SOX3 binding sites in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 4L, N), while SMARCA2 binding was low in both WT and SMARCA2 mutant NPCs at SOX3 binding sites (Figure 4M). These data suggest that the loss of SOX2 and SOX3 expression in SMARCA2 mutant NPCs leads to a failure to recruit SMARCA4-containing SWI/SNF complexes and maintain chromatin accessibility at SOX3 binding sites enriched in neural-specific enhancers. We then overlapped DEGs in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs with genes annotated to SOX2 or SOX3 binding sites and found that 43% of all SMARCA2 mutant DEGs are bound by SOX2, SOX3, or both (Figure 4O). Further, we found that over 25% of down-regulated genes in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs were also reduced by siSOX2/3 knockdown in mouse NPCs (Bergsland et al., 2011), including key neural genes *PAX6*, *POU3F2*, *SOX13*, and *NKX6.1*, while 19% of up-regulated genes including *FN1*, *COL1A1*, *JUNB*, *NFIB, and NFIX* were increased upon SOX2/3 knockdown (Figure 4P). As examples, we focused on *PAX6* and *SALL1*. These genes have SOX2 and SOX3 binding at intronic enhancers that are marked by accessibility and active histone modifications in WT NPCs, which are significantly reduced in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs, along with SOX2, SOX3, and SMARCA4 binding (Figure S4B). These results demonstrate that SOX2/3 are key NPC transcription factors that regulate neural genes in part by recruiting SMARCA4 to neural-specific enhancers. In SMARCA2 mutant NPCs, the reduced expression of SOX2/3 and other transcription factors results in the loss of SMARCA4 binding, chromatin accessibility, and active histone modifications at neural-specific enhancers, resulting in dysregulation of neural gene transcription.

SMARCA2 mutations activate an ectopic astrocyte gene program

We then examined the sites that gain accessibility in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs, 94% of which lack chromatin accessibility, H3K4me, and H3K27ac in WT NPCs (Figure 5A). We refer to these as *de novo* enhancers (DEs) to distinguish them from sites that have basal accessibility in WT NPCs (Figure 5A). Moreover, the vast majority of increased accessible sites (9,309/10,015) lack H3K4me or

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

H3K27ac in hESCs, suggesting they arise completely *de novo* and not from poised enhancers that are primed for activation in hESCs. To characterize these sites further, we overlapped sites that lost or gained accessibility in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs with super enhancers from 59 different tissues reported in the dbSUPER database (Khan and Zhang, 2016). Consistent with our findings, 30–40% of super enhancers identified in brain tissues including the hippocampus, inferior temporal lobe and mid frontal lobe lose chromatin accessibility in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 5B). Strikingly though, more than 60% of the super enhancers identified in astrocytes and myoblasts gain chromatin accessibility (Figure 5B). In addition, DEs have increased H3K4me and H3K27ac in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs, indicating that SMARCA2 mutations lead to the establishment of active enhancers (Figure 5C). Indeed, of the astrocyte super enhancers that exhibited increased accessibility in SMARCA2 mutant NPCs, nearly a third also resulted in increased gene expression of the closest annotated astrocyte-specific gene (Figure 5D). In parallel, we found that 20% of upregulated genes in SMARCA2 mutant NPCs are specifically enriched in mouse and human astrocytes (Cahoy et al., 2008; Zhang et al., 2016), about half of which also exhibited increased accessibility (Figure 5E). This list reflects an enrichment of early astrocyte precursor genes and includes several astrocyte markers (*CD44, F3, CLU,* and *FSTL1*), the NFI family of transcription factors (*NFIA*, *NFIB*, *NFIX*, *NFIC*) that can facilitate astrocyte differentiation *in vitro* and *in vivo* (Caiazzo et al., 2015; Deneen et al., 2006; Matuzelski et al., 2017), and several astrocyte growth and survival genes (*EGFR*, *NOTCH2*, and *JUNB*) (Scholze et al., 2014). These data strongly suggest that SMARCA2 mutant NPCs activate an ectopic astrocyte gene program through aberrant activation of *de novo* astrocyte enhancers.

Pioneering activity of AP-1 retargets SWI/SNF complexes to *de novo* enhancers

We found that the FRA1 (bZIP) motif specific to AP-1 factors was specifically enriched at sites that gain accessibility in SMARCA2 mutant NPCs (Figure 6A, B). AP-1 factors frequently cooperate with lineage-determining factors in cellular differentiation (Heinz et al., 2010; Vierbuchen et al., 2017). Accordingly, we found an enrichment of RUNX, TEAD, and NFI motifs +/− 50bp from the center of the AP-1 motif at sites that gained accessibility in SMARCA2 mutant NPCs. However, these motifs occurred in less than 7% of AP-1 containing increased accessible sites (Figure 6C), indicating that AP-1 is the primary driver in the establishment of newly accessible sites through a previously unappreciated pioneer activity. AP-1 members *FOSL1* (FRA1), *FOSL2* (FRA2), and *JUNB* were significantly upregulated in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 6D), as was FRA2 protein (Figure 6E). Further, we found that FRA2 ChIP binding was highly enriched at sites that gain accessibility in SMARCA2 mutant NPCs (Figure 6F). We and others have shown that accessibility at AP-1 binding sites is SWI/SNF complex-dependent (Kelso et al., 2017; Mathur et al., 2017; Nakayama et al., 2017), mediated in part through an interaction between SWI/SNF complexes and AP-1 factors (Vierbuchen et al., 2017). Indeed, we found that FRA2 associated with SMARCA4, as well as WT and mutant SMARCA2 (Figure 6G). Further, we found that both SMARCA2 and SMARCA4 binding were dramatically increased with concomitant gain in accessibility at DE sites that gain FRA2 binding in *SMARCA2*$^{K755R/+}$ and *SMARCA2*$^{R1159Q/+}$ NPCs (Figure 6H–J). These data suggest that upregulation of AP-1 factors and specifically FRA2 leads to retargeting of SMARCA4- and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SMARCA2-containing SWI/SNF complexes, driving *de novo* activation of enhancers. To test this hypothesis, we knocked down FRA1, FRA2, and JUNB in *SMARCA2*$^{R1159Q/+}$ NPCs (Figure S5A) and found that knockdown of AP-1 factors reduced the gain in accessibility (Figure 6K), as well as the recruitment of SMARCA4 and SMARCA2 (Figure 6L, M) at FRA2 binding sites in *SMARCA2*$^{R1159Q/+}$ NPCs. Thus, the recruitment of SWI/SNF complexes and gain of chromatin accessibility in SMARCA2 mutant NPCs to *de novo* enhancers is dependent on AP-1. Finally, we found that 97% of upregulated genes with corresponding increases in accessibility (n=595) are bound by FRA2 in SMARCA2 mutant NPCs (Figure 6N), e.g. *NFIB* and *FN1* (Figure S5B). Together, these data demonstrate that SMARCA2 mutations result in retargeting of SMARCA4 and SMARCA2 by pioneer factor AP-1, leading to functional remodeling of astrocyte enhancers and upregulation of astrocyte genes.

DISCUSSION

Here we report the discovery that NCBRS-associated mutations in SMARCA2 result in defects in neurogenesis stemming from a loss of SOX3 expression and SOX3-dependent recruitment of SMARCA4 to neural enhancers. This, in turn, leads to decreased expression of SOX3 target genes, including *PAX6, SALL1,* and *SOX11*. These results are consistent with several reports describing *SOX11* mutations in individuals with intellectual disability and SWI/SNF-related phenotypes (Hempel et al., 2016; Tsurusaki et al., 2014). However, it should be noted that individuals with *SMARCA2* mutations have distinct and more severe phenotypes than those with *SOX11* mutations (Bogershausen and Wollnik, 2018), suggesting that SMARCA2 mutant phenotypes likely stem from alterations in the expression of many genes. For example, we find that AP-1 and NFI factors are upregulated along with many genes associated with astrogliogenesis. AP-1 has not previously been implicated in astrogliogenesis, although it is known that MAPK signaling, which impinges on AP-1 factors, is required for astrogliogenesis in vitro (Rajan and McKay, 1998) and in the developing brain (Li et al., 2012). Interestingly, AP-1 genes and NFI family genes are also directly bound by SOX2/3 and upregulated in SOX2/3 knockdown NPCs (Bergsland et al., 2011), indicating that gliogenesis may be actively inhibited by SOX2/3 repression of these genes. Collectively, these data suggest that the morphological and phenotypic effects of SMARCA2 mutations are due to dysregulation of SOX3-dependent gene programs, leading to loss of neurogenic capacity in favor of gliogenic potential.

Perhaps the most intriguing finding from our work is the association of SMARCA4 and SMARCA2 binding in SMARCA2 mutant NPCs with changes in chromatin accessibility. Contrary to findings on the activity of mutant SWI2 (Richmond and Peterson, 1996) and mutant hBRG1 (Khavari et al., 1993), we did not observe binding of SMARCA2 in association with decreased accessibility or gene repression of SMARCA2 target genes. Instead, we observed widespread redistribution of SMARCA4, including decreased binding at sites that lose accessibility and increased binding at sites that gain accessibility. At sites that gained SMARCA4 binding and chromatin accessibility, there was a concomitant gain of both WT and mutant SMARCA2 binding due in large part to SMARCA2 upregulation, suggesting that the binding of mutant SMARCA2 has no effect on remodeling activity of SMARCA4/A2 at these sites. These data are inconsistent with models supporting the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

inhibition of wild-type activity by mutant ATPases through direct interference (Khavari et al., 1993; Richmond and Peterson, 1996) or sequestration of interaction partners as suggested for mutant SMARCA4 (Hodges et al., 2018). We find in contrast that the binding of SMARCA4 and SMARCA2 is functionally active in reprogramming the enhancer landscape, and that a key feature of SMARCA2 mutant cells is the aberrant gain in activity of the SWI/SNF complex.

Our studies implicate a previously unappreciated pioneering activity for AP-1 in the redistribution of SMARCA4 and SMARCA2. Notably, we and others show that AP-1 family members interact with SWI/SNF complexes (Vierbuchen et al., 2017). In addition, our work and others have shown that chromatin accessibility at lineage-specific enhancers highly enriched in AP-1 motifs and AP-1 binding is SWI/SNF- and AP-1-dependent (Alver et al., 2017; Kelso et al., 2017; Mathur et al., 2017; Vierbuchen et al., 2017). Here we show that AP-1 factors are sufficient to drive SWI/SNF complex redistribution and functional remodeling to *de novo* sites. These data suggest that AP-1 has pioneer activity (Zaret and Carroll, 2011), or the ability to bind its cognate site on nucleosomes, independent of cooperation with other lineage-determining factors. It is not yet clear how AP-1 accesses these sites specifically, which may be related to the location of the binding site relative to nucleosome phasing, the underlying DNA sequence, or DNA methylation status. Our results have implications for our understanding of the role of AP-1 in lineage determination and suggest that future studies are necessary to revisit the potential pioneering contribution of AP-1 and SWI/SNF complexes to the action of lineage-determining factors in defining cell type-specific enhancer landscapes.

STAR METHODS

CONTACT FOR REAGENT AND RESOURCE SHARING

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Diana C. Hargreaves (dhargreaves@salk.edu)

EXPERIMENTAL MODEL AND SUBJECT DETAILS

WT H9 female human embryonic stem cells (hESCs) were obtained from Stem Cell Core Facility of the Salk Institute. All cell lines used in this study were derived from H9 hESCs. WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ H9 human embryonic stem cells (hESCs) were cultured under standard feeder-free conditions on basement membrane extract (Trevigen) using an in-house produced medium similar to mTeSR1 as published(Ludwig et al., 2006). hESCs were maintained on Matri-gel-coated (BD biosciences, #354230) plates at 37°C, 5% CO2 with daily media changes and passaged when 80–90% confluent. WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ neural progenitor cells (NPCs) were culture in DMEM/F12, GlutaMAX (Thermo Fisher Sci #10565018) supplemented with 1× N2 (Thermo Fisher Sci #17502048), 1× B27 without Vitamin A (Thermo Fisher Sci # 12587010), 20ng/ml FGF2 (Joint protein central). NPCs were maintained on poly-l-ornithine and laminin-coated plates initially and amplified on Matri-gel coated plates at 37°C, 5% CO2 with media changes every other day.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

METHOD DETAILS

Genome Editing—SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ hESCs were generated using CRISPR(Ran et al., 2013) genome editing. Briefly, sgRNAs targeting the intronic region before exons corresponding to K755R or R1159Q were cloned into the px330 (Addgene 42230) vector. The repair template including right and left arms was synthesized by Biomatik to inlcude XhoI and EcoRI sites. The loxP-PGK-NEO-loxP was PCR amplified from pL452 (gift from Dr. Pei Wang, University of Texas Health Science Center at San Antonio), digested with XhoI and EcoRI, and inserted into the sites in the repair template. Finally the repair template was cloned into pL253 (gift from Dr. Pei Wang, University of Texas Health Science Center at San Antonio). H9 ESCs were nucleofected with 5ug of pX330-K755R or pX330-R1159 and 5 ug of linearized repair template using Human Stem Cell Nucleofector kit 1 (Lonza Bioscience, # VPH-5012) and Amaxa Nucleofector (Lonza Bioscience) with program B-016. Cells were plated on matrigel at low seeding density. Colonies were picked, expanded, and confirmed by genotyping. Cells containing corresponding mutations were nucleofected again with pSALK-CRE to remove the PGK-NEO cassette. Both genomic DNA and cDNA was sequenced using Sanger sequencing to confirm heterozygosity. All hESC clones were karyotyped by IGM Genomics Center (University of California at San Diego) and confirmed to be normal karyotype.

Neural progenitor cell generation—NPCs were differentiated from hESCs use a modified protocol based on previous published Embryoid Body (EB)-derived methods. (Boyer et al., 2012; Zhang et al., 2001) Briefly, hESCs were passaged onto γ-irradiated mouse embryonic fibroblasts (MEFs). After large colonies formed, cells were dissociated with collagenase type IV (1mg/ml in DMEM/F12, Thermo Fisher Sci #17104019), re-suspended in Neural Induction Medium (NIM, DMEM/F12, GlutaMAX supplemented with 1× N2 and 1× B27 without Vitamin A), and transferred to an ultra-low-attachment plate. EBs were grown in NIM, changing media every 2–3 days for a total of 7 days. EBs were then transferred to poly-l-ornithine and laminin-coated (P/L) plates in NIM supplemented with 1 μg/ml laminin (Thermo Fisher Sci #23017015) for attachment. EBs were grown for neural rosette formation with media changes every other day. After 7 days, neural rosettes with long projections were picked manually to a new P/L coated plate in Neural Progenitor Medium (NPM, DMEM/F12, GlutaMAX supplemented with 1× N2, 1× B27 without Vitamin A, 20ng/ml FGF2 and 1 μg/ml laminin). The next day, cultures were aspirated to remove cells without projections and fed with NPM every other day thereafter. When confluent, clumps were dissociated with Accutase (Stemcell technologies #07920), passaged to a new P/L coated plate.

Neuronal differentiation—NPCs were plated on poly-l-ornithine and laminin-coated (P/L) 4 well chamber slides (Millipore #PEZGS0416) at a density of 2×10^4/well in NPM. The next day, cells were fed with Neural Differentiation Medium (NDM, DMEM/F12, GlutaMAX supplemented with 1× N2, 1× B27 (Thermo Fisher Sci #17504044), 20ng/ml FGF2, 1μg/ml laminin, 20 ng/ml BDNF (Peprotech #450–02), 20 ng/ml GDNF (Peprotech, #450–10), 200 nM ascorbic acid (Sigma, #A0278) and 1 mM cyclic AMP (Sigma, #D0627)). Cells were differentiated at 37°C, 5% CO2 with media changes every 4 days for 6 weeks.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Immunofluorescence—Cells were fixed with 4% paraformaldehyde (Alfa Aesar #43368–9M) in phosphate-buffered saline (PBS) for 20 minutes at room temperature and washed 3 times with PBS for 10 minutes. Cells were incubated with blocking buffer (PBS with 5% donkey serum (Jackson immunoresearch #017-000-121)) containing 0.1% Triton X-100 for 15 minutes followed by blocking buffer for 30 minutes at room temperature and then incubated overnight with primary antibodies at 4 °C followed by 3 washes with PBS. C ells were blocked again with blocking buffer at room temperature for 30 minutes and incubated with secondary antibodies for 1 hour at room temperature followed by 1 wash with PBS. Cells were stained with DAPI (2ug/ml, Sigma-Aldrich, #D9542) for 10 minutes at room temperature followed by 3 washes with PBS. Slides were mounted with Shandon™ Immu-Mount™ (Thermo Fisher Sci, # 9990402). Images were taken with a Zeiss Microscope (Carl Zeiss Microscopy).

Lentivirus preparation and infection—cDNA of SMARCA2 WT, K755R or R1159Q was fused with AM tag at C-terminus and cloned into an EF1α-driven lentiviral expression vector (N103). HEK293T cells were transfected with the lentiviral constructs and packaging plasmids pMD2.G (Addgene, #12259) and psPAX2 (Addgene, #12260) using polyethylenimine (PEI)-mediated transfection. The media containing the virus was collected 48 and 72 hours post-transfection, filtered and centrifuged at 70,952 xg for two hours at 4C. The viral pellet was resuspended in 1× PBS and stored in the −80°C until use. hNPCs were seeded at $5.6\times10^4/cm^2$ 12 hours before infection. NPM containing concentrated virus was added to cells for infection. Virus containing NPM was replaced by NPM containing puromycin (1ug/ml, Thermo Fisher Sci, # A1113803) 24 hours later for a total selection of 48 hours. Infected cells were amplified for at least 1 passage before collecting for downstream experiments.

RNAi—hNPCs were seeded at $5.6\times10^4/cm^2$ 12 hours before transfection. For SOX2 and SOX3 double knockdown, SOX2 siRNA (Dharmacon, #L-011778–00) and SOX3 siRNA (Dharmacon, #L-012143–00) or equivalent non-targeting RNA (Dharmacon, #D-001810) was mixed in Opt-MEM (Thermo Fisher Sci #31985–070) and Lipofectamine RNAiMAX (Thermo Fisher Sci #13778030) as suggested by reagent menu. For AP-1 triple knockdown, FOSL1 siRNA (Dharmacon, #L-004341–00), FOSL2 siRNA (Dharmacon, #L-004110–00) and JUNB siRNA (Dharmacon, #L-003269–00) or equivalent non-targeting RNA (Dharmacon, #D-001810) was mixed in Opt-MEM (Thermo Fisher Sci #31985–070) and Lipofectamine RNAiMAX (Thermo Fisher Sci #13778030) as suggested by reagent menu. Seventy-two hours post transfection, cells were collected for ChIP-seq and ATAC-seq.

Conservation analysis—SNF2 domain conservation was analyzed with Consurf 2016(Ashkenazy et al., 2016) (http://consurf.tau.ac.il/2016/) using human SMARCA2 ATPase domain amino acids sequence as input. Default parameters were used.

Structure visualization—The Yeast Snf2 domain structure was visualized using PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.).

Proliferation assay—NPCs were plated on a 24 well plates with a density of 4×10^4/ well. The next day, cells were fed with fresh medium to remove floating dead cells and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

imaged every 3 hours with a 10× objective in an IncuCyte (Essen BioScience). At least 9 images were taken from each well and every sample had duplicate wells for each experiment. Pictures were processed with IncuCyte ZOOM 2016B to measure well confluency based on phase contrast. Cell confluency of each well was exported for further analysis.

Cell cycle analysis—NPCs cell cycle were measured with propidium iodide (PI) staining using flow cytometry analysis. Briefly, NPCs were collected and washed once with PBS. Cells were fixed with cold 70% ethanol for 30 minutes at 4°C followed with 2 w ash in PBS. Cells were re-suspended with 50 μl of RNAase (100 μg/ml), followed by 200 μl PI (50 μg/ml, Thermo Fisher Sci, P3566) and analyzed by flow cytometry. LSRII (BD biosciences) was used for the flow cytometry analysis and data was processed with Flowjo software (Flowjo).

Immunobloting (IB) assay—Protein samples were run on 4–12% Bis-Tris gels (Life Technologies). After primary antibody incubation, blots were probed with 1:20,000 v/v dilution of fluorescently-labeled secondary antibodies (Life Technologies #A21058, #SA535571) in 2% BSA in PBST for an hour at room temperature. Fluorescent images were developed using the Odyssey Imager (LI-COR Biosciences). In Figure 1H, figure was generated from composite images taken from a single gel with no deletions for the purpose of lane alignment.

Immunoprecipitation—Nuclear lysates were collected from NPCs following a revised Dignam protocol(Andrews and Faller, 1991). Cells were re-suspended in Buffer A (10 mM Hepes pH 7.9, 1.5 mM MgCl2, 10 mM KCl) supplemented with 1 mM DTT, 1 mM PMSF, 1 μM pepstatin, 10 μM leupeptin and 10 μM chymostatin, and cells were lysed by homogenization using a 21-gauge needle with six to eight strokes. If lysis was incomplete, cells were treated with 0.05–0.1% NP-40 for ten minutes on ice prior to nuclei collection. Nuclei were spun down at 6000 × g for 2 minutes then re-suspended in Buffer C (20 mM Hepes pH 7.9, 20% glycerol, 420 mM NaCl, 1.5 mM MgCl2, 0.2 mM EDTA) supplemented with 1 mM DTT, 1 mM PMSF, 1 μM pepstatin, 10 μM leupeptin and 10 μM chymostatin. After 45 minutes of end-to-end rotation in the cold room, sample was clarified at 21,000 × g for ten minutes. Supernatant was collected and flash frozen in liquid nitrogen, if necessary. Prior to the IP, nuclear lysates were diluted with two times of original volume of 20 mM HEPES pH 7.9, 0.3% NP-40 to dilute the NaCl concentration to 140mM. 200–300 μg of nuclear lysates was used per IP with antibodies overnight at 4°C. Precipitated proteins were bound to 50:50 Protein A, Protein G Dynabeads (Thermo Fisher Sci #10002D, 10004D) for one to two hours and washed with Wash Buffer (50 mM Tris pH 8, 150 mM NaCl, 1 mM EDTA, 10% glycerol, 0.1% NP-40). Proteins were eluted in SDS-PAGE loading solution with boiling and analyzed by IB.

RNA-Seq—RNA from 2×10^6 WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs was extracted and purified with the Zymo Research Quick-RNA miniprep kit according to manufacturer's instructions. RNA-seq libraries were prepared using Illumina TruSeq

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Stranded mRNA kit following manufacturer's instructions with 5 μg of input RNA. All RNA-seq analysis was performed on 2 biological replicates.

ATAC-Seq—50,000 NPCs were collected and washed with cold PBS. Cells were re-suspended in resuspension buffer (10mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2) and then lysed in lysis buffer (10mM Tris-HCl, pH 7.4, 10 mM NaCl, 3 mM MgCl2, 0.1% NP-40). Nuclei were collected and incubated in transposition mix containing Tn5 transposase (Illumina, #15027865) for 30 minutes at 37°C. DNA was then purified and ligated with adapters and amplified by PCR using NEBNext® High-Fidelity 2× PCR Master Mix (New England BioLabs, #M0541L). DNA library was size selected using AMPure XP Beads (Beckman Coulter, A63882) to remove adapters and large DNA fragments and send for paired-end sequencing on NextSeq 500. All ATAC-seq analysis was performed on 2 biological replicates.

ChIP-seq—For ChIP-Seq of histone modifications, 1.5×10^6 NPCs were collected and cross-linked in 1% formaldehyde. For ChIP-Seq of SMARCA2, SMARCA4, FRA2, SOX2, SOX3, and AM, 6×10^6 NPCs were first collected and cross-linked in 2mM disuccinimidyl glutarate (DSG) then in 1% formaldehyde. After quenching the excess formaldehyde with 125 mM glycine, the fixed cells were washed, pelleted and flash-frozen. Upon thawing, the cells were re-suspended in lysis solution (50 mM HEPES-KOH pH 8, 140 mM NaCl, 1 mM EDTA, 10% glycerol, 0.5% NP40, 0.25% Triton X-100 and incubated on ice for ten minutes. The isolated nuclei were washed with wash solution (10 mM Tris-HCl pH 8, 1 mM EDTA, 0.5 mM EGTA, 200 mM NaCl) and shearing buffer (0.1% SDS, 1 mM EDTA, 10 mM Tris-HCl pH 8) then sheared in a Covaris E220 sonicator for 10–20 minutes to generate DNA fragments between ~200–1000bp. After clarification of insoluble material by centrifugation, the chromatin was immunoprecipitated overnight at 4°C with antibodies against H3K4me, H3K4me3, H3K27me3, H3K27ac, SMARCA2, SMARCA4, FRA2, SOX2, SOX3 and AM bound to Protein A+G Dynabeads (Invitrogen) in ChIP buffer (50 mM HEPES-KOH pH 7.5, 300 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% DOC, 0.1% SDS). Antibody-bound DNA was washed and treated with Proteinase K (50 μg/ml, Thermo Fisher Sci, #17916) and RNase A (0.2 mg/ml, Thermo Fisher Sci, #EN0531) and cross-linking was reversed by incubation at 55°C for two and a ha lf hours. Purified ChIP DNA was used for library generation (NuGen Ovation Ultralow Library System V2, NuGEN Technologies, #0344NB) according to manufacturer's instructions for subsequent sequencing. ChIP-seq analysis of SMARCA2, SMARCA4, and histone modifications in WT and mutant NPCs was performed on 2 biological replicates. ChIP-seq of SOX2, SOX3, and FRA2 in WT and mutant NPCs, SMARCA2 and SMARCA4 after SOX2/3 RNAi, and AM in $SMARCA2^{R1159Q/+}$ NPCs was performed once.

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA-Seq analysis—Single-end 50bp reads were aligned to hg38 using STAR alignment tool (V2.5)(Dobin et al., 2013) (STAR: ultrafast universal RNA-seq aligner. Bioinformatics). RNA expression was quantified as raw integer counts using analyzeRepeats.pl in HOMER(Heinz et al., 2010) using the following parameters: -strand both -count exons -condenseGenes-noadj. To identify differentially expressed genes, we performed

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

getDiffExpression.pl in HOMER, which uses the DESeq2(Love et al., 2014) R package to calculate the biological variation within replicates. Cutoffs were set at log2 FC = 0.585 and FDR at 0.05 (Benjamin-Hochberg). Repeats were analyzed using analyzeRepeats.pl to quantify gene expression by Fragments Per Kilobase of Million mapped reads (FPKM) across the biological replicates.

ATAC-seq analysis—Paired-end 42bp reads were aligned to hg38 using STAR alignment tool (V2.5) (STAR: ultrafast universal RNA-seq aligner. Bioinformatics). ATAC-seq peaks were called using the findPeaks program within HOMER using parameters for DNAse-seq (-style dnase). Peaks were called when enriched >4 fold over local tag counts. Differential peaks were found using DESeq2 R package by merging peaks from WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ samples and called using getDifferentialPeaksReplicates.pl with fold change ≥ 2 or ≤ −2, FDR < 0.05. Differential peaks were annotated by mapping to the nearest TSS using the annotatePeaks.pl command.

ChIP-seq analysis—Single-end 50 bp reads were aligned to hg38 using STAR alignment tool (V2.5) (STAR: ultrafast universal RNA-seq aligner. Bioinformatics). ChIP-Seq peaks were called using findPeaks within HOMER using default parameters for histone (-style histone) or transcription factor (-style factor). Peaks were called when enriched > two-fold over input and > four-fold over local tag counts, with FDR 0.001 (Benjamin-Hochberg). For histone ChIP, peaks within a 1000 bp range were stitched together to form regions. ChIP-Seq peaks or regions were annotated by mapping to the nearest TSS using the annotatePeaks.pl command. Differential ChIP peaks were found by merging peaks from control and experiment groups and called using getDifferentialPeaks with fold change ≥ 1.5 or ≤ −1.5, Poisson p value < 0.0001. Significance of peak overlap was determined by calculating the number of peaks co-occurring across the entire genome using the HOMER mergePeaks program. For hESC histone modification enrichment analysis, we used publicly available H9 hESC ChIP-seq data histone modifications (Rada-Iglesias et al., 2011) (see Data Availability). Enhancers were called by identifying all H3K4me-positive regions that are at least 1 kb away from the nearest TSS or H3K4me3 mark. These were sub-divided as active (H3K27ac-positive) or poised (H3K27ac-negative) (Creyghton et al., 2010). We then differentiated the H3K4me-positive, H3K27ac-positive regions into active and super enhancers by ranking the regions by H3K27ac ChIP-seq tag density and using the tangent of the curve to call super enhancers (Whyte et al., 2013).

Gene Ontology analysis—Gene ontology analysis was performed on the GSEA website (GSEA homepage (Subramanian et al., 2005), [www.gsea-msigdb.org/], 2004–2017). FDR values were calculated according to Benjamin-Hochberg multiple testing.

Motif analysis—Sequences within 200 bp of peak centers were compared to known motifs in the HOMER database using the findMotifsGenome.pl command with the following fragment size: -size 200. Random GC content-matched genomic regions were used as background (default). Enriched motifs are statistically significant motifs in input over background by a p value of less than 0.05. P values were calculated using cumulative binomial distribution.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Super enhancer data analysis—Super enhancer peaks were downloaded from dbSUPER database (Khan and Zhang, 2016). Direct overlap between our differential accessibility peaks and super enhancers were calculated by HOMER mergePeaks function with -matrix option. Percentage overlapping of each super enhancer is calculated and rank color sorted in excel.

Statistical tests

Conservation Analysis Quantification: Figure 1B, Statistically significant differences was calculated by two-tailed t tests which were performed to calculate P values using Excel function ttest.

Proliferation assay Quantification: Figure 1E, Cell confluency was measured by InCucyte for over 45 hours and normalized to 0h time point.

Two-tailed t tests were performed for the last time point to calculate P values using Excel function ttest.

Immunofluorescence Quantification: Figure 1I, Positive staining cells were normalized to DAPI stained nuclei. Statistically significant differences were calculated by Two-tailed t tests on Graphpad Prism Version 7.

Immunoblotting Quantification: Figure 2G, ARID1A n=12, ARID1B n=8, SMARCA4 n=12, SMARCA2 n=12, SMARCD1 n=12, SMARCE1 n=10, ACTL6A n=10, SMARCB1 n=9. Quantification was done by Image Studio™ (V2.1, LI-COR) and statistically significant differences were calculated by Two-tailed t tests on Graphpad Prism Version 7.

Correlation Analysis: Figure 3G, correlation of SMARCA2 and SMARCA4 log2 fold change values between datasets: Goodness of fit (R2) was analyzed using linear regression on Graphpad Prism Version 7.

Observed over Expected Enrichment: Figure 4A, the natural log of observed/expected was calculated using hypergeometric distribution.

DATA AND SOFTWARE AVAILABILITY

Data availability and data sets used—RNA-seq, ATAC-seq and ChIP-seq data that support the findings of this study have been deposited in the Gene Expression Omnibus under the accession code GSE122631. We also used publicly available sequencing data, which were processed using HOMER v4.10 (Christopher Benner, HOMER, http://homer.ucsd.edu/homer/index.html). H9 hESC histone modification data sets were downloaded from GEO (GSE24447) and mapped to genome hg38. Peaks were called as described in ChIP-seq data analysis session. Raw Immunofluorescent and immunoblotting data has been deposited to Mendeley: http://dx.doi.org/10.17632/hrpkrwt9sc.1

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ACKNOWLEDGEMENTS

We are grateful to Salk Core Facilities supported by NIH-NCI CCSG: P30 014195 and the Helmsley Trust. We thank Ken Diffenderfer, Aimee Pankonin, James Pierpoint, Nasun Hah, Thu Nguyen, Uri Manor, Tong Zhang, Carolyn O'Connor, and Conor Fitzpatrick for technical support, Drs. Carole Marchetto, Fred Gage and Nicola Allen (Salk) for helpful discussions, Dr. Jovylyn Gatchalian (Salk) for critical reading of the manuscript and to the members of the Hargreaves Lab for their valuable feedback. DCH is supported by the NIH R35 GM128943–01, NIH R00 CA184043–03, the V Foundation for Cancer Research V2016–006, and the Helmsley Trust. FG is supported by Salk Pioneer Fund Postdoctoral Scholar Award.

REFERENCES

Alver BH, Kim KH, Lu P, Wang X, Manchester HE, Wang W, Haswell JR, Park PJ, and Roberts CW (2017). The SWI/SNF chromatin remodelling complex is required for maintenance of lineage specific enhancers. Nat Commun 8, 14648. [PubMed: 28262751]

Andrews NC, and Faller DV (1991). A rapid micropreparation technique for extraction of DNA-binding proteins from limiting numbers of mammalian cells. Nucleic Acids Res 19, 2499. [PubMed: 2041787]

Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, and Ben-Tal N (2016). ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 44, W344–350. [PubMed: 27166375]

Ashkenazy H, Erez E, Martz E, Pupko T, and Ben-Tal N (2010). ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. Nucleic Acids Res 38, W529–533. [PubMed: 20478830]

Berezin C, Glaser F, Rosenberg J, Paz I, Pupko T, Fariselli P, Casadio R, and Ben-Tal N (2004). ConSeq: the identification of functionally and structurally important residues in protein sequences. Bioinformatics 20, 1322–1324. [PubMed: 14871869]

Bergsland M, Ramskold D, Zaouter C, Klum S, Sandberg R, and Muhr J (2011). Sequentially acting Sox transcription factors in neural lineage development. Genes Dev 25, 2453–2464. [PubMed: 22085726]

Bogershausen N, and Wollnik B (2018). Mutational Landscapes and Phenotypic Spectrum of SWI/SNF-Related Intellectual Disability Disorders. Front Mol Neurosci 11, 252. [PubMed: 30123105]

Boyer LF, Campbell B, Larkin S, Mu Y, and Gage FH (2012). Dopaminergic differentiation of human pluripotent cells. Curr Protoc Stem Cell Biol , *Chapter 1* Unit1H 6.

Cahoy JD, Emery B, Kaushal A, Foo LC, Zamanian JL, Christopherson KS, Xing Y, Lubischer JL, Krieg PA, Krupenko SA, et al. (2008). A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. J Neurosci 28, 264–278. [PubMed: 18171944]

Caiazzo M, Giannelli S, Valente P, Lignani G, Carissimo A, Sessa A, Colasante G, Bartolomeo R, Massimino L, Ferroni S, et al. (2015). Direct conversion of fibroblasts into functional astrocytes by defined transcription factors. Stem Cell Reports 4, 25–36. [PubMed: 25556566]

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, et al. (2010). Histone H3K27ac separates active from poised enhancers and predicts developmental state. Proc Natl Acad Sci U S A 107, 21931–21936. [PubMed: 21106759]

De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, Kou Y, Liu L, Fromer M, Walker S, et al. (2014). Synaptic, transcriptional and chromatin genes disrupted in autism. Nature 515, 209–215. [PubMed: 25363760]

Deciphering Developmental Disorders S (2015). Large-scale discovery of novel genetic causes of developmental disorders. Nature 519, 223–228. [PubMed: 25533962]

Deneen B, Ho R, Lukaszewicz A, Hochstim CJ, Gronostajski RM, and Anderson DJ (2006). The transcription factor NFIA controls the onset of gliogenesis in the developing spinal cord. Neuron 52, 953–968. [PubMed: 17178400]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, and Gingeras TR (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15–21. [PubMed: 23104886]

Gelbart ME, Bachman N, Delrow J, Boeke JD, and Tsukiyama T (2005). Genome-wide identification of Isw2 chromatin-remodeling targets by localization of a catalytically inactive mutant. Genes Dev 19, 942–954. [PubMed: 15833917]

Hargreaves DC, and Crabtree GR (2011). ATP-dependent chromatin remodeling: genetics, genomics and mechanisms. Cell Res 21, 396–420. [PubMed: 21358755]

Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, and Glass CK (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38, 576–589. [PubMed: 20513432]

Hempel A, Pagnamenta AT, Blyth M, Mansour S, McConnell V, Kou I, Ikegawa S, Tsurusaki Y, Matsumoto N, Lo-Castro A, et al. (2016). Deletions and de novo mutations of SOX11 are associated with a neurodevelopmental disorder with features of Coffin-Siris syndrome. J Med Genet 53, 152–162. [PubMed: 26543203]

Ho L, and Crabtree GR (2010). Chromatin remodelling during development. Nature 463, 474–484. [PubMed: 20110991]

Hodges HC, Stanton BZ, Cermakova K, Chang CY, Miller EL, Kirkland JG, Ku WL, Veverka V, Zhao K, and Crabtree GR (2018). Dominant-negative SMARCA4 mutants alter the accessibility landscape of tissue-unrestricted enhancers. Nat Struct Mol Biol 25, 61–72. [PubMed: 29323272]

Kadoch C, Hargreaves DC, Hodges C, Elias L, Ho L, Ranish J, and Crabtree GR (2013). Proteomic and bioinformatic analysis of mammalian SWI/SNF complexes identifies extensive roles in human malignancy. Nat Genet 45, 592–601. [PubMed: 23644491]

Kelso TWR, Porter DK, Amaral ML, Shokhirev MN, Benner C, and Hargreaves DC (2017). Chromatin accessibility underlies synthetic lethality of SWI/SNF subunits in ARID1A-mutant cancers. Elife 6.

Khan A, and Zhang X (2016). dbSUPER: a database of super-enhancers in mouse and human genome. Nucleic Acids Res 44, D164–171. [PubMed: 26438538]

Khavari PA, Peterson CL, Tamkun JW, Mendel DB, and Crabtree GR (1993). BRG1 contains a conserved domain of the SWI2/SNF2 family necessary for normal mitotic growth and transcription. Nature 366, 170–174. [PubMed: 8232556]

Kosho T, and Okamoto N (2014). Genotype-phenotype correlation of Coffin-Siris syndrome caused by mutations in SMARCB1, SMARCA4, SMARCE1, and ARID1A. Am J Med Genet C Semin Med Genet 166C, 262–275. [PubMed: 25168959]

Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB, et al. (2016). Analysis of protein-coding genetic variation in 60,706 humans. Nature 536, 285–291. [PubMed: 27535533]

Lessard J, Wu JI, Ranish JA, Wan M, Winslow MM, Staahl BT, Wu H, Aebersold R, Graef IA, and Crabtree GR (2007). An essential switch in subunit composition of a chromatin remodeling complex during neural development. Neuron 55, 201–215. [PubMed: 17640523]

Li X, Newbern JM, Wu Y, Morgan-Smith M, Zhong J, Charron J, and Snider WD (2012). MEK Is a Key Regulator of Gliogenesis in the Developing Brain. Neuron 75, 1035–1050. [PubMed: 22998872]

Liu X, Li M, Xia X, Li X, and Chen Z (2017). Mechanism of chromatin remodelling revealed by the Snf2-nucleosome structure. Nature 544, 440–445. [PubMed: 28424519]

Love MI, Huber W, and Anders S (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15, 550. [PubMed: 25516281]

Ludwig TE, Bergendahl V, Levenstein ME, Yu J, Probasco MD, and Thomson JA (2006). Feeder-independent culture of human embryonic stem cells. Nat Methods 3, 637–646. [PubMed: 16862139]

Mathur R, Alver BH, San Roman AK, Wilson BG, Wang X, Agoston AT, Park PJ, Shivdasani RA, and Roberts CW (2017). ARID1A loss impairs enhancer-mediated gene regulation and drives colon cancer in mice. Nat Genet 49, 296–302. [PubMed: 27941798]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Matuzelski E, Bunt J, Harkins D, Lim JWC, Gronostajski RM, Richards LJ, Harris L, and Piper M (2017). Transcriptional regulation of Nfix by NFIB drives astrocytic maturation within the developing spinal cord. Dev Biol 432, 286–297. [PubMed: 29106906]

Miyake N, Abdel-Salam G, Yamagata T, Eid MM, Osaka H, Okamoto N, Mohamed AM, Ikeda T, Afifi HH, Piard J, et al. (2016). Clinical features of SMARCA2 duplication overlap with Coffin-Siris syndrome. Am J Med Genet A 170, 2662–2670. [PubMed: 27264538]

Nakayama RT, Pulice JL, Valencia AM, McBride MJ, McKenzie ZM, Gillespie MA, Ku WL, Teng M, Cui K, Williams RT, et al. (2017). SMARCB1 is required for widespread BAF complex-mediated activation of enhancers and bivalent promoters. Nat Genet 49, 1613–1623. [PubMed: 28945250]

Ninkovic J, Steiner-Mezzadri A, Jawerka M, Akinci U, Masserdotti G, Petricca S, Fischer J, von Holst A, Beckers J, Lie CD, et al. (2013). The BAF complex interacts with Pax6 in adult neural progenitors to establish a neurogenic cross-regulatory transcriptional network. Cell Stem Cell 13, 403–418. [PubMed: 23933087]

Pretegiani E, Mari F, Renieri A, Penco S, and Dotti MT (2016). Nicolaides-Baraitser syndrome: defining a phenotype. J Neurol 263, 1659–1660. [PubMed: 27286846]

Rada-Iglesias A (2018). Pioneering of Enhancer Landscapes during Pluripotent State Transitions. Cell Stem Cell 23, 149–151. [PubMed: 30075122]

Rada-Iglesias A, Bajpai R, Swigut T, Brugmann SA, Flynn RA, and Wysocka J (2011). A unique chromatin signature uncovers early developmental enhancers in humans. Nature 470, 279–283. [PubMed: 21160473]

Rajan P, and McKay RD (1998). Multiple routes to astrocytic differentiation in the CNS. J Neurosci 18, 3620–3629. [PubMed: 9570793]

Ran FA, Hsu PD, Wright J, Agarwala V, Scott DA, and Zhang F (2013). Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8, 2281–2308. [PubMed: 24157548]

Reyes JC, Barra J, Muchardt C, Camus A, Babinet C, and Yaniv M (1998). Altered control of cellular proliferation in the absence of mammalian brahma (SNF2alpha). Embo J 17, 6979–6991. [PubMed: 9843504]

Richmond E, and Peterson CL (1996). Functional analysis of the DNA-stimulated ATPase domain of yeast SWI2/SNF2. Nucleic Acids Res 24, 3685–3692. [PubMed: 8871545]

Santen GW, Aten E, Sun Y, Almomani R, Gilissen C, Nielsen M, Kant SG, Snoeck IN, Peeters EA, Hilhorst-Hofstee Y, et al. (2012). Mutations in SWI/SNF chromatin remodeling complex gene ARID1B cause Coffin-Siris syndrome. Nat Genet 44, 379–380. [PubMed: 22426309]

Santen GW, Aten E, Vulto-van Silfhout AT, Pottinger C, van Bon BW, van Minderhout IJ, Snowdowne R, van der Lans CA, Boogaard M, Linssen MM, et al. (2013). Coffin-Siris syndrome and the BAF complex: genotype-phenotype study in 63 patients. Hum Mutat 34, 1519–1528. [PubMed: 23929686]

Scholze AR, Foo LC, Mulinyawe S, and Barres BA (2014). BMP signaling in astrocytes downregulates EGFR to modulate survival and maturation. PLoS One 9, e110668. [PubMed: 25330173]

Sif S, Saurin AJ, Imbalzano AN, and Kingston RE (2001). Purification and characterization of mSin3A-containing Brg1 and hBrm chromatin remodeling complexes. Genes Dev 15, 603–618. [PubMed: 11238380]

Sousa SB, and Hennekam RC (2014). Phenotype and genotype in Nicolaides-Baraitser syndrome. Am J Med Genet C Semin Med Genet 166C, 302–314. [PubMed: 25169058]

Staahl BT, Tang J, Wu W, Sun A, Gitler AD, Yoo AS, and Crabtree GR (2013). Kinetic analysis of npBAF to nBAF switching reveals exchange of SS18 with CREST and integration with neural developmental pathways. J Neurosci 33, 10348–10361. [PubMed: 23785148]

Stessman HA, Xiong B, Coe BP, Wang T, Hoekzema K, Fenckova M, Kvarnung M, Gerdts J, Trinh S, Cosemans N, et al. (2017). Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. Nat Genet 49, 515–526. [PubMed: 28191889]

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al. (2005). Gene set enrichment analysis: a knowledge-based

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102, 15545–15550. [PubMed: 16199517]

Tang S, Hughes E, Lascelles K, Simpson MA, and Pal DK (2017). New SMARCA2 mutation in a patient with Nicolaides-Baraitser syndrome and myoclonic astatic epilepsy. Am J Med Genet A 173, 195–199. [PubMed: 27665729]

Tsurusaki Y, Koshimizu E, Ohashi H, Phadke S, Kou I, Shiina M, Suzuki T, Okamoto N, Imamura S, Yamashita M, et al. (2014). De novo SOX11 mutations cause Coffin-Siris syndrome. Nat Commun 5, 4011. [PubMed: 24886874]

Tsurusaki Y, Okamoto N, Ohashi H, Kosho T, Imai Y, Hibi-Ko Y, Kaname T, Naritomi K, Kawame H, Wakui K, et al. (2012). Mutations affecting components of the SWI/SNF complex cause Coffin-Siris syndrome. Nat Genet 44, 376–378. [PubMed: 22426308]

Van Houdt JK, Nowakowska BA, Sousa SB, van Schaik BD, Seuntjens E, Avonce N, Sifrim A, Abdul-Rahman OA, van den Boogaard MJ, Bottani A, et al. (2012). Heterozygous missense mutations in SMARCA2 cause Nicolaides-Baraitser syndrome. Nat Genet 44, 445–449, S441. [PubMed: 22366787]

Vierbuchen T, Ling E, Cowley CJ, Couch CH, Wang X, Harmin DA, Roberts CWM, and Greenberg ME (2017). AP-1 Transcription Factors and the BAF Complex Mediate Signal-Dependent Enhancer Selection. Mol Cell 68, 1067–1082 e1012. [PubMed: 29272704]

Whyte WA, Orlando DA, Hnisz D, Abraham BJ, Lin CY, Kagey MH, Rahl PB, Lee TI, and Young RA (2013). Master transcription factors and mediator establish super-enhancers at key cell identity genes. Cell 153, 307–319. [PubMed: 23582322]

Wolff D, Endele S, Azzarello-Burri S, Hoyer J, Zweier M, Schanze I, Schmitt B, Rauch A, Reis A, and Zweier C (2012). 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. [PubMed: 22822383]

Wright CF, McRae JF, Clayton S, Gallone G, Aitken S, FitzGerald TW, Jones P, Prigmore E, Rajan D, Lord J, et al. (2018). Making new genetic diagnoses with old data: iterative reanalysis and reporting from genome-wide data in 1,133 families with developmental disorders. Genet Med 20, 1216–1223. [PubMed: 29323667]

Wu JI, Lessard J, Olave IA, Qiu Z, Ghosh A, Graef IA, and Crabtree GR (2007). Regulation of dendritic development by neuron-specific chromatin remodeling complexes. Neuron 56, 94–108. [PubMed: 17920018]

Yuen RC, Merico D, Bookman M, J LH, Thiruvahindrapuram B, Patel RV, Whitney J, Deflaux N, Bingham J, Wang Z, et al. (2017). Whole genome sequencing resource identifies 18 new candidate genes for autism spectrum disorder. Nat Neurosci 20, 602–611. [PubMed: 28263302]

Zarate YA, Bhoj E, Kaylor J, Li D, Tsurusaki Y, Miyake N, Matsumoto N, Phadke S, Escobar L, Irani A, et al. (2016). SMARCE1, a rare cause of Coffin-Siris Syndrome: Clinical description of three additional cases. Am J Med Genet A 170, 1967–1973. [PubMed: 27264197]

Zaret KS, and Carroll JS (2011). Pioneer transcription factors: establishing competence for gene expression. Genes Dev 25, 2227–2241. [PubMed: 22056668]

Zhang SC, Wernig M, Duncan ID, Brustle O, and Thomson JA (2001). In vitro differentiation of transplantable neural precursors from human embryonic stem cells. Nat Biotechnol 19, 1129–1133. [PubMed: 11731781]

Zhang Y, Sloan SA, Clarke LE, Caneda C, Plaza CA, Blumenthal PD, Vogel H, Steinberg GK, Edwards MS, Li G, et al. (2016). Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. Neuron 89, 37–53. [PubMed: 26687838]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights:

- Heterozygous missense mutations in SMARCA2 impair neural differentiation
- SMARCA2 mutations result in enhancer reprogramming due to SMARCA4 retargeting
- SOX3-dependent neural enhancers are lost in SMARCA2 mutant neural progenitor cells
- AP-1 factors drive *de novo* enhancer formation in SMARCA2 mutant cells

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

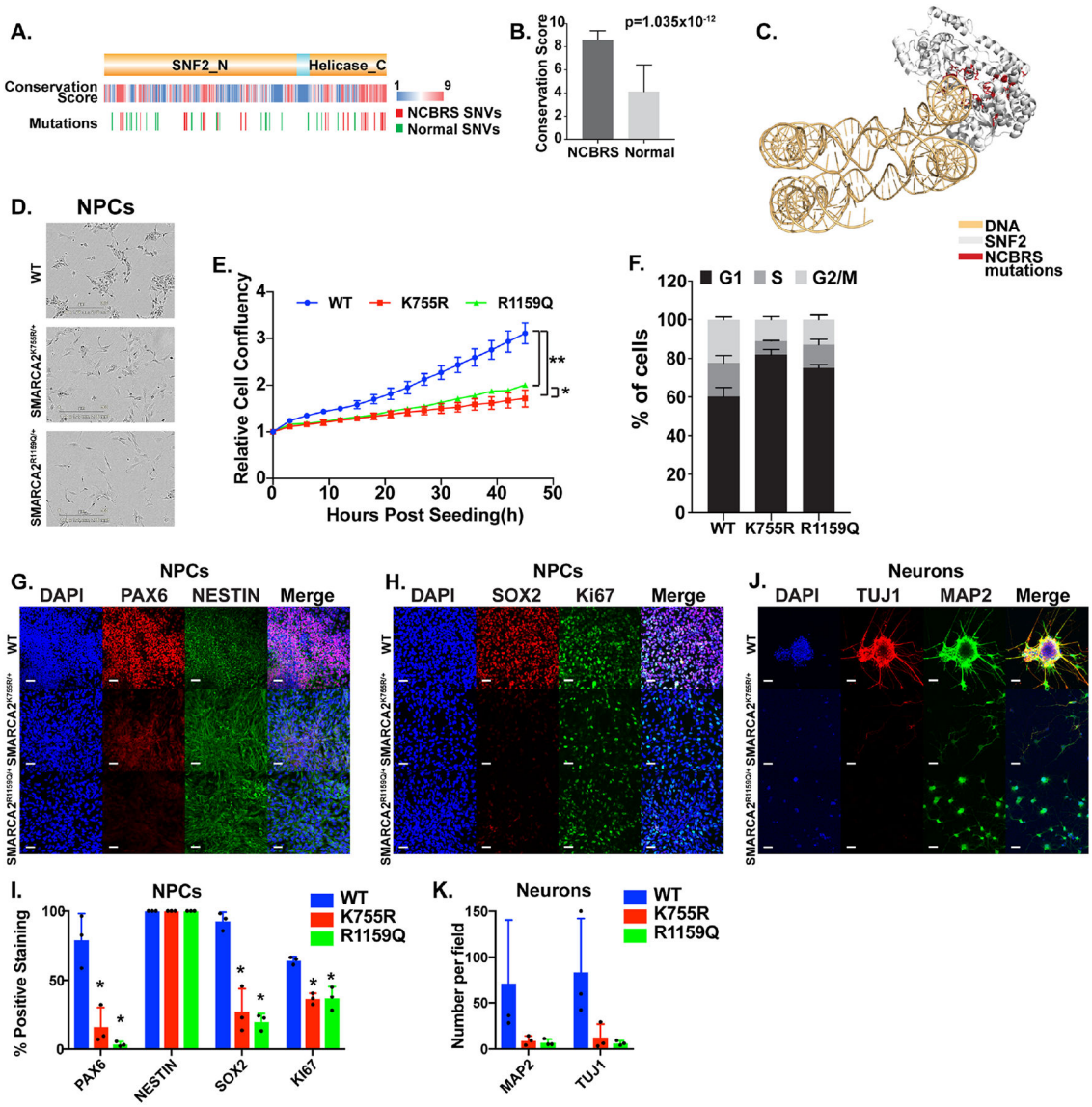


Figure 1. SMARCA2 mutations impair neural progenitor differentiation.

A. Schematic diagram of SMARCA2 ATPase domain with conservation score. B. Conservation analysis comparing NCBRS SNVs with normal SNVs. Error bars = mean ± SD, n=28 (NCBRS), 24 (Normal) C. Structure of yeast SNF2 domain and DNA with NCBRS mutations labeled in red. D. Light microscopy of SMARCA2$^{+/+}$, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. E. Cell confluency of WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. **: p<0.01, * p<0.05 Error bars = mean ± SD, n=4. F. Cell cycle analysis of WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs as measured by propidium iodide staining. Error Bar = mean ± SD, n=3. G and H. Immunofluorescence of WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs stained with PAX6 and NESTIN (left), or SOX2 and Ki67 (right), along with DAPI. Scale bar=50μm. I. Quantitation of the percent of positively stained cells in G and H, Error Bar = mean ± SD, n=3. J. Immunofluorescence of WT, SMARCA2$^{K755R/+}$ and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SMARCA2$^{R1159Q/+}$ neurons stained with TUJ1, MAP2, and DAPI. Scale bar=50μm. K. Quantitation of the number of positively stained cells per field in J, Error Bar = mean ± SD, n=3. See also Figure S1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

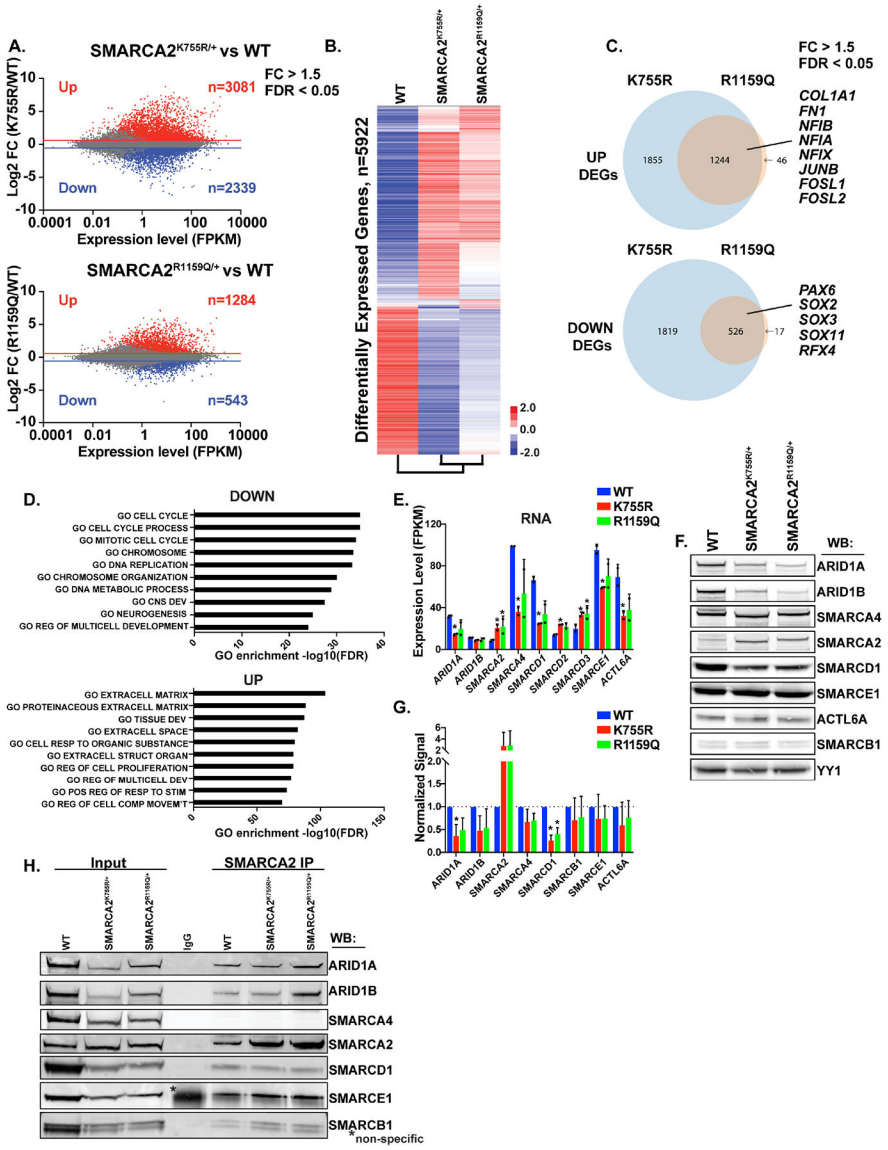


Figure 2. SMARCA2 mutant NPCs have altered gene expression.

A. Scatterplot of average log2 Fragments Per Kilobase of transcripts per Million mapped reads (FPKM) mRNA expression level against log2 fold change (FC) in expression of SMARCA2$^{K755R/+}$/WT (top) or SMARCA2$^{R1159Q/+}$/WT NPCs (bottom). B. Hierarchical clustering of log2 FPKM of differentially expressed genes (DEGs) in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. C. Overlap of upregulated and downregulated DEGs in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. D. Significance of gene ontology (GO) analysis of genes that are downregulated (top) or upregulated (bottom) in both SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. E. FPKM of SWI/SNF subunits in WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. Error bars = mean ± SD, n=2. F. Immunoblotting of protein level of SWI/SNF subunits in WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. G. Quantitation of western blots as represented in F, Error bars = mean ± SD, sample size see method. H. SMARCA2 immunoprecipitation in WT,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs and immunoblotting for the indicated SWI/SNF subunits. See also Figure S2.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

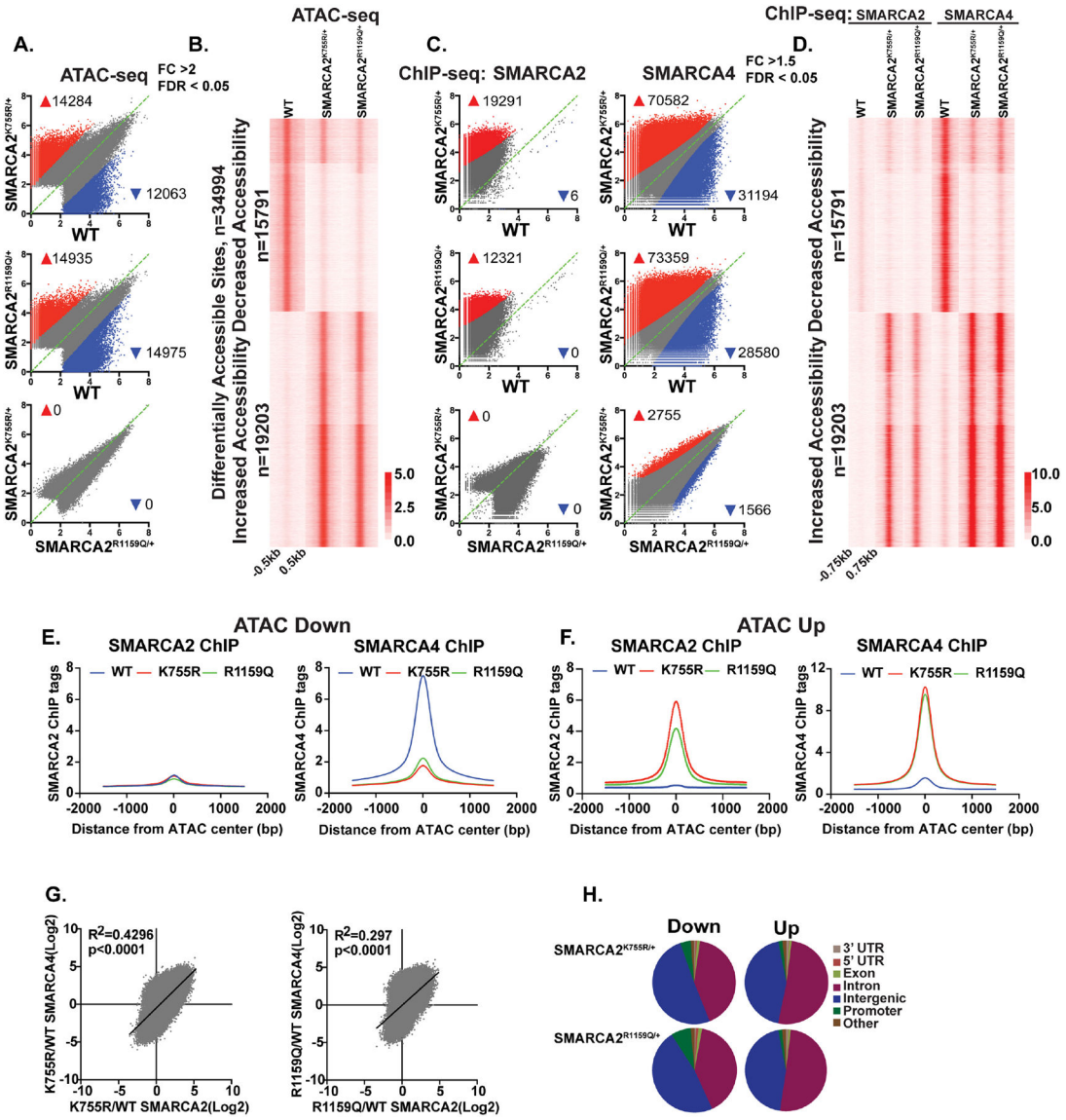


Figure 3. Changes in chromatin accessibility in SMARCA2 mutant NPCs are correlated with the binding of SMARCA4.

A. Scatterplot of average Log2 ATAC-seq tag density of WT against SMARCA2$^{K755R/+}$ (top), WT against SMARCA2$^{R1159Q/+}$ (middle) and SMARCA2$^{K755R/+}$ against SMARCA2$^{R1159Q/+}$ (bottom). B. Heat map of log2 ATAC tag density at differentially accessible sites in WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. C. Scatterplot of average Log2 ChIP-seq tag density of WT against SMARCA2$^{K755R/+}$ (top), WT against SMARCA2$^{R1159Q/+}$ (middle) and SMARCA2$^{K755R/+}$ against SMARCA2$^{R1159Q/+}$ (bottom). D. Heat map of SMARCA2 and SMARCA4 ChIP-seq tags at sites with increased or decreased accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs, clustered as in B. E and F. Histogram of Normalized ChIP-seq signal for SMARCA2 or SMARCA4 at sites that decreased (E, n=15,791) or increased (F, n=19,203) accessibility in SMARCA2$^{K755R/+}$ or SMARCA2$^{R1159Q/+}$ NPCs. G. Scatterplot of log2 fold change (FC) of SMARCA2 against SMARCA4 ChIP density in SMARCA2$^{K755R/+}$ (Left) or SMARCA2$^{R1159Q/+}$ (Right) versus

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

WT. H. Pie chart of composition of sites with decreased (down) or increased (up) accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs compared to WT NPCs. See also Figure S2 and S3.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

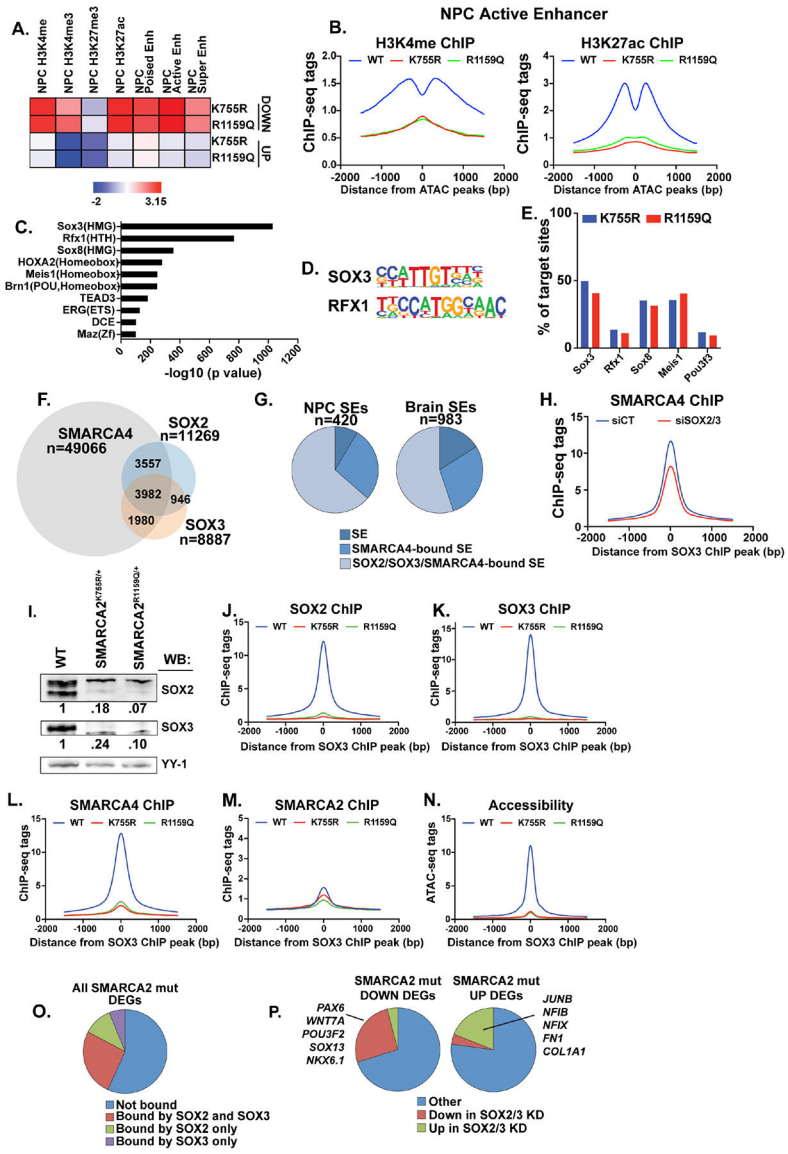


Figure 4. SOX3-dependent neural enhancers are lost in SMARCA2 mutant NPCs.

A. Enrichment of sites that increased or decreased accessibility in both SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs for histone modifications and enhancer groups in WT NPCs. B. Histogram of normalized ChIP-seq tags for H3K4me and H3K27ac at the NPC active enhancers (n=4,977) that lose accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. C. Significance of de novo motif enrichment on sites that lose accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. D. The SOX3 and RFX1 motifs identified in C. E. Proportion of decreased accessible sites in SMARCA2$^{K755R/+}$ or SMARCA2$^{R1159Q/+}$ NPCs containing indicated motif. F. Overlap of SOX2, SOX3, and SMARCA4 ChIP-seq binding sites. G. Proportion of NPC super enhancers (SEs) or Brain (cingulate gyrus) super enhancers that are bound by SMARCA4, or SMARCA4, SOX2, and SOX3. H. Histogram of normalized ChIP-seq tags for SMARCA4 in WT NPCs transfected with siRNAs against control or SOX2 and SOX3 at SOX3 binding sites that lose chromatin

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

accessibility (n=3,190). I. Immunoblot of SOX2/3 in WT, SMARCA2$^{K755R/+}$, SMARCA2$^{R1159Q/+}$ NPCs. J-N. Histogram of normalized ChIP-seq tags for SOX2 (J), SOX3 (K), SMARCA4 (L), SMARCA2 (M) or ATAC-seq tags (N) in WT, SMARCA2$^{K755R/+}$ or SMARCA2$^{R1159Q/+}$ NPCs at SOX3 binding sites that lose accessibility (n=3,190). O. Fraction of all Differentially Expressed Genes (DEGs) from SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs bound by SOX2, SOX3, or both. P. Fraction all DEGs from SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs down- or up-regulated in SOX2/SOX3 knockdown NPCs. See also Figure S4.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

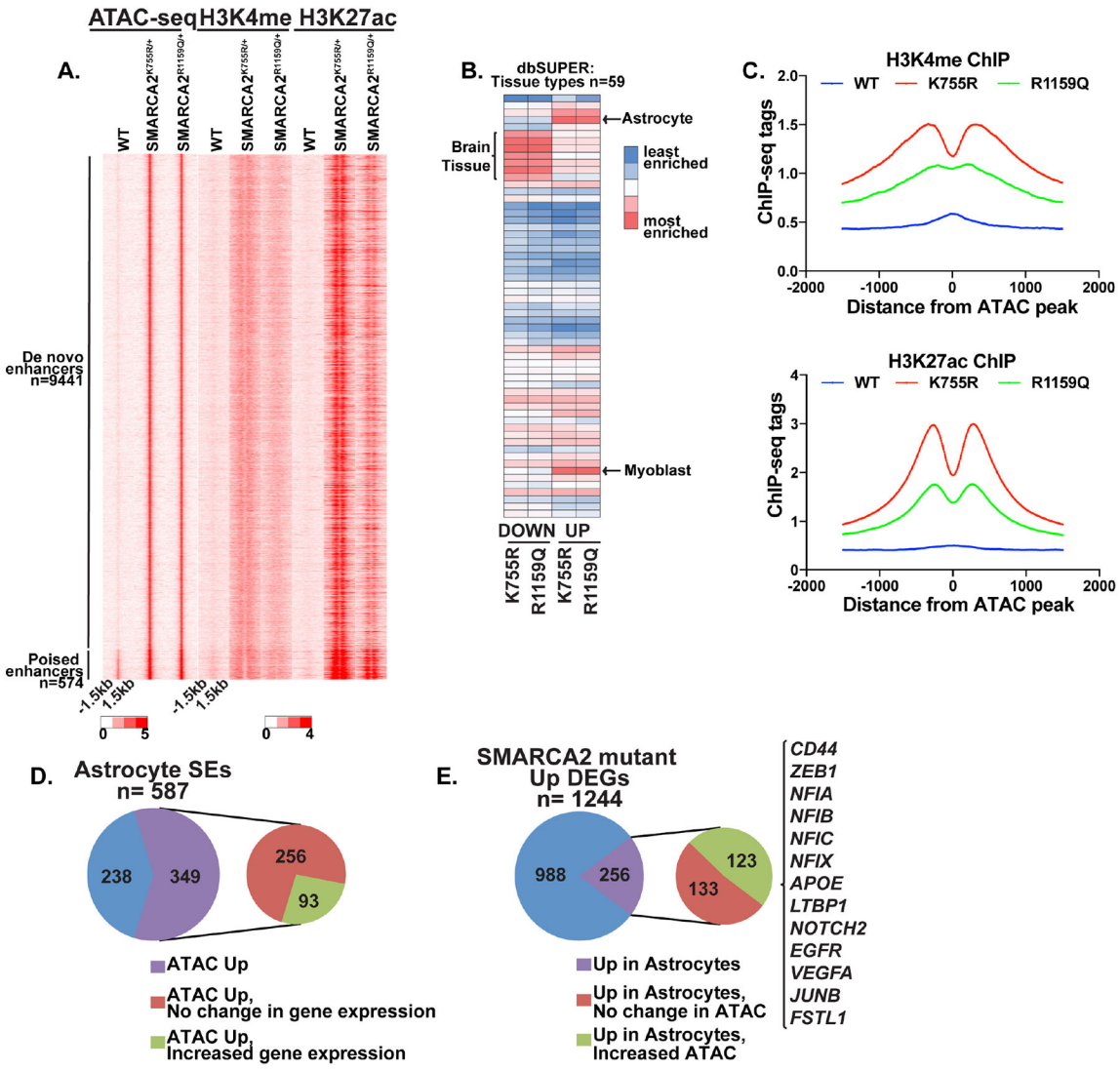


Figure 5. SMARCA2 mutations lead to de novo activation of astrocyte-specific enhancers.

A. Heat map of log2 ATAC tag density or ChIP-seq tags at increased accessible sites in WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs defined as having no accessibility (De novo) or a peak of accessibility (Poised) in NPCs. B. Rank of percent overlap of sites that increased or decreased accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs with super enhancers from 59 tissues in the dbSUPER database. C. Histogram of normalized ChIP-seq tags for H3K4me (top) and H3K27ac (bottom) at the DE sites that gain accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs (n=9,441). D. Proportion of astrocyte Super Enhancer (SE)-associated genes with increased accessibility and gene expression in SMARCA2 mutant NPCs. E. Proportion of SMARCA2 mutant upregulated genes that are selectively expressed in astrocytes and have increased accessibility in SMARCA2 mutant NPCs.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

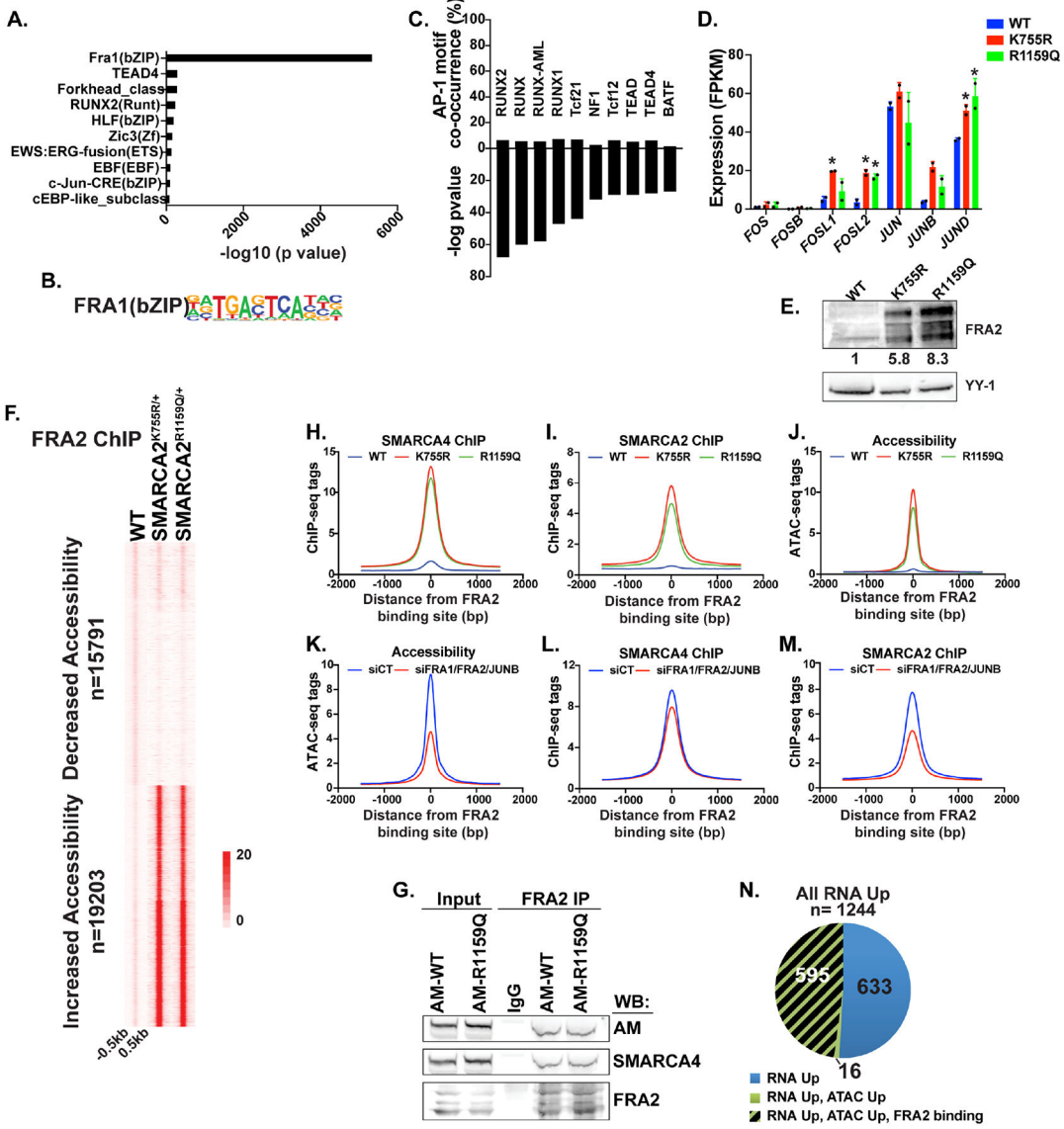


Figure 6. Pioneering activity of AP-1 retargets SWI/SNF complexes to de novo enhancers.

A. Enrichment of de novo motifs on sites that gain accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. B. De novo FRA1 (bZIP) motif identified in A. C. (Bottom) Enrichment (-log p value) of motifs +/− 50bp of AP-1 sites with increased accessibility in SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. (Top) Percent of AP-1 sites with increased accessibility that have the indicated motif. D. FPKM of RNA-seq of AP-1 family members in WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. Error bars = mean ±SD,n=2. *FDR < 0.05. E. Immunoblot and quantitation of FRA2 in WT, SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs. F. Heat map of FRA2 ChIP-seq tags at differentially accessible sites clustered as in Figure 3B. G. FRA2 co-immunoprecipitation of SMARCA4, AM-SMARCA2, and AM-SMARCA2 R1159Q in WT NPCs. H-J. Histogram of normalized ChIP-seq tags for SMARCA4 (H), SMARCA2 (I), and ATAC-seq tags (J) at DE sites that gain FRA2 binding in both SMARCA2$^{K755R/+}$ and SMARCA2$^{R1159Q/+}$ NPCs

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(n=7,312). K-M. Histogram of normalized ATAC-seq tags (K), or ChIP-seq tags for SMARCA4 (L) and SMARCA2 (M) in SMARCA2$^{K755R/+}$ NPCs transfected with siRNAs against control or FRA1/FRA2/JUNB at DE sites that gain FRA2 binding in SMARCA2$^{K755R/+}$ NPCs (n=10,900). N. Proportion of genes with increased expression, accessibility, and FRA2 binding in SMARCA2 mutant NPCs. See also Figure S5.