Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Mol Cell. Author manuscript; available in PMC 2022 November 18.

Published in final edited form as:

Mol Cell. 2021 November 18; 81(22): 4663–4676.e8. doi:10.1016/j.molcel.2021.09.020.

NRF1 Association with AUTS2-Polycomb Mediates Specific Gene Activation in the Brain

Sanxiong Liu[1,2], **Kimberly A. Aldinger**[3], **Chi Vicky Cheng**[3], **Takae Kiyama**[4,5], **Mitali Dave**[3], **Hanna K. McNamara**[3], **Wukui Zhao**[1], **James M. Stafford**[1,2], **Nicolas Descostes**[1,2], **Pedro Lee**[1,2], **Stefano G. Caraffi**[6], **Ivan Ivanovski**[6,7], **Edoardo Errichiello**[8,19], **Christiane Zweier**[9,20], **Orsetta Zuffardi**[8], **Michael Schneider**[10], **Antigone S. Papavasiliou**[11], **M. Scott Perry**[12], **Jennifer Humberson**[13], **Megan T. Cho**[14], **Astrid Weber**[15], **Andrew Swale**[15,16], **Tudor C. Badea**[5,17], **Chai-An Mao**[4,5], **Livia Garavelli**[6], **William B. Dobyns**[3,18], **Danny Reinberg**[1,2,21,*]

[1]Department of Biochemistry and Molecular Pharmacology, New York University Langone School of Medicine, New York, NY 10016, USA

[2]Howard Hughes Medical Institute, Chevy Chase, MD 20815, USA

[3]Center for Integrative Brain Research, Seattle Children's Research Institute, Seattle, WA, USA

[4]Ruiz Department of Ophthalmology and Visual Science, McGovern Medical School at The University of Texas Health Science Center at Houston (UTHealth), Houston, TX 77030, USA

[5]National Eye Institute, NIH, Bethesda, MD 20892, USA

[6]Struttura Semplice Dipartimentale di Genetica Medica, Azienda USL-IRCCS di Reggio Emilia, Reggio Emilia, Italy

[7]Institute of Medical Genetics, University of Zürich, Zürich, Switzerland

[8]Dipartimento di Medicina Molecolare, Università di Pavia, Pavia, Italy

[9]Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU), Schwabachanlage 10, 91054 Erlangen, Germany

[10]Carle Physicians Group, Section of Neurology, St. Christopher's Hospital for Children, Urbana, IL, USA

[11]Neurology, IASO Children's Hospital, Athens, Greece

[*]Correspondence: danny.reinberg@nyumc.org, https://doi.org/10.1016/j.molcel.2021.09.020.

AUTHOR CONTRIBUTIONS

S.L. and D.R. conceived the project, designed the experiments, and wrote the paper. S.L. performed the vast majority of the experiments and all bioinformatics analyses. K.A.A., C.V.C., M.D., and H.K.M. performed the clinical and genetic analyses under the supervision of W.B.D. W.Z. helped with the luciferase assay. J.M.S., N.D., and P.L. helped with the ChIP-seq and IP-MS in mouse brain. S.G.C., I.I., E.E., C.Z., O.Z., M.S., A.S.P., M.S.P., J.H., M.T.C., A.W., A.S., and L.G. collected the clinical data. T.K. performed the Nrf1 mouse studies under the supervision of T.C.B. and C.-A.M. D.R. supervised the study.

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 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

D.R. is a cofounder of Constellation and Fulcrum Pharmaceuticals. Other authors declare no competing interests.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

[12]Comprehensive Epilepsy Program, Jane and John Justin Neuroscience Center, Cook Children's Medical Center, Fort Worth, TX 76104, USA

[13]Division of Genetics, Department of Pediatrics, University of Virginia Children's Hospital, Charlottesville, VA, USA

[14]GeneDx, Inc., Gaithersburg, MD, USA

[15]Liverpool Women's Hospital, Liverpool, UK

[16]Manchester Centre for Genomic Medicine, Manchester, UK

[17]Research and Development Institute, Transilvania University of Brasov, School of Medicine, Brasov, Romania

[18]Department of Pediatrics (Genetic Medicine), University of Washington, Seattle, WA, USA

[19]IRCCS Mondino Foundation, Pavia, Italy

[20]Department of Human Genetics, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland

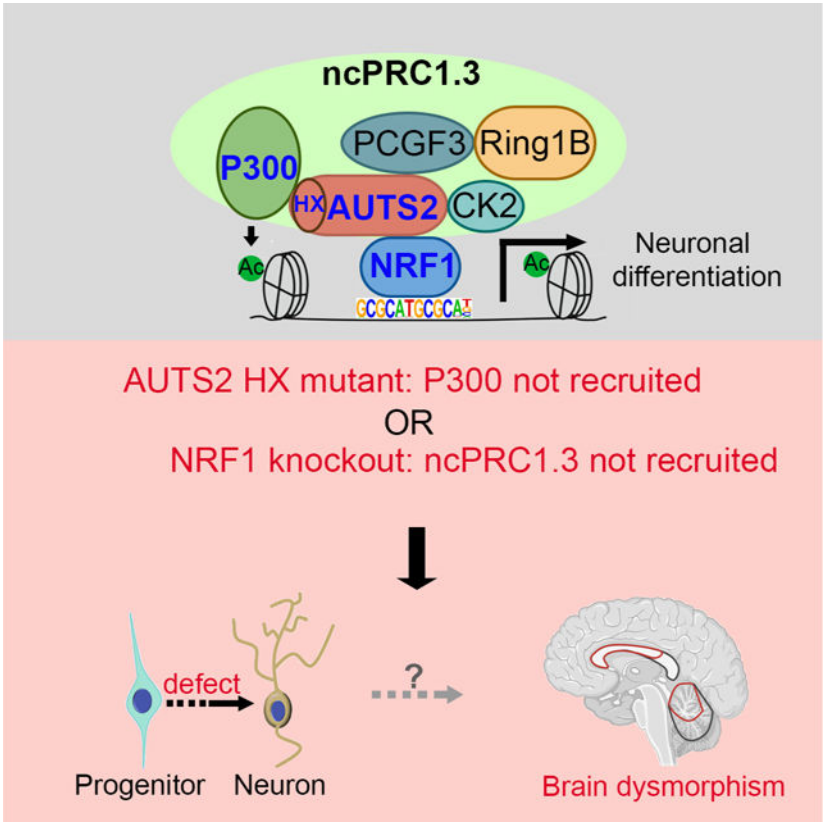
[21]Lead contact

SUMMARY

The heterogeneous family of complexes comprising Polycomb Repressive Complex 1 (PRC1) is instrumental to establishing facultative heterochromatin that is repressive to transcription. Yet, two PRC1 species, ncPRC1.3 and ncPRC1.5, are known to comprise novel components, AUTS2, P300, and CK2 that convert this repressive function to that of transcription activation. Here, we report that patients harboring mutations in the HX repeat domain of AUTS2 exhibit defects in AUTS2 and P300 interaction as well as a developmental disorder reflective of Rubinstein-Taybi syndrome, which is mainly associated with a heterozygous pathogenic variant in *CREBBP/EP300*. Moreover, the absence of AUTS2 or mutation in its HX repeat domain gives rise to a mis-regulation of a subset of developmental genes and curtails motor neuron differentiation of mouse embryonic stem cells. Notably, the transcription factor, Nuclear Respiratory Factor 1 (NRF1) exhibits a novel and integral role in this neurodevelopmental process, being required for ncPRC1.3 recruitment to chromatin.

Graphical Abstract

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



eTOC blurb

Liu et al. report that patient-derived mutations in the HX repeat of AUTS2 disrupt its interaction with P300, thereby thwarting AUTS2-ncPRC1.3-mediated active transcription. Such mutations reflect those in P300 proper with respect to RSTS. Moreover, NRF1-mediated AUTS2 recruitment is paramount to the activation of AUTS2-ncPRC1.3 targets and brain development.

INTRODUCTION

The establishment, maintenance, inheritance, and regulated dissolution of facultative heterochromatin (Trojer and Reinberg, 2007) are key to attaining the distinctive gene expression profiles that arise during differentiation in multi-cellular organisms. In contrast to chromatin domains engaged in transcription, our previous findings demonstrate that it is the repertoire of repressive chromatin domains that are conveyed to daughter cells upon DNA replication, pointing to their role in preserving a given cell identity during cell division (Escobar et al., 2019, 2021). The formation of facultative heterochromatin during development and its integrity in adulthood depend on the activities of the Polycomb group (PcG) of proteins (Bonasio et al., 2010; Margueron and Reinberg, 2011; Di Croce and Helin, 2013).

Two multi-subunit complexes, Polycomb Repressive Complex-1 and -2 (PRC1 and PRC2, respectively) act in concert to establish facultative heterochromatin (Schuettengruber et al., 2017). PRC2 comprises EZH1/-2, the sole activity that catalyzes mono-, di, and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

tri-methylation of histone H3 at lysine 27 (H3K27me1, -me2, -me3, respectively), with chromatin domains comprising H3K27me2/me3 providing the platform for chromatin compaction (Lau et al., 2017; Oksuz et al., 2018; Yu et al., 2019; Kim and Kingston, 2020). PRC1 complexes comprise other subsets of the PcG protein family. We and others previously characterized at least six heterogeneous PRC1 subcomplexes (PRC1.1-PRC1.6), each of which comprise one of the six Polycomb Group Ring Finger (PCGF1–6) components and all of which comprise RING1A and/or RING1B (Gao et al., 2012; Tavares et al., 2012; Hauri et al., 2016). This array of complexes was broadly classified into two major PRC1 subcomplexes, canonical and non-canonical PRC1, which network with PRC2 in distinct manners.

Canonical PRC1 (cPRC1) comprise one of several CBX proteins that bind to chromatin containing nucleosomes decorated with H3K27me3, resulting in chromatin compaction and thus, transcription repression (Min et al., 2003; Francis et al., 2004; Gao et al., 2012; Lau et al., 2017; Kim and Kingston, 2020). On the other hand, the RYBP or YAF2 components of noncanonical PRC1 (ncPRC1) stimulate RING1-mediated catalysis of mono-ubiquitination of lysine 119 of histone H2A (H2AK119ub1) (Rose et al., 2016), which correlates with transcription inhibition (Stock et al., 2007; Zhou et al., 2008). ncPRC1 coordinates with PRC2 through the PRC2 accessory protein, Jarid2, which interacts with H2AK119ub1 (Kasinath et al., 2021), and also stimulates PRC2 activity (Li et al., 2010; Pasini et al., 2010). While the joint recruitment of PRC2 with either cPRC1 or ncPRC1 is distinct, both versions of PRC1 are found in proximity at select genome-wide regions (Gao et al., 2012; Scelfo et al., 2019). Most importantly, this PRC2/PRC1 network is paramount to fostering the appropriate profiles of facultative heterochromatin evident during development and in adulthood (Margueron and Reinberg, 2011; Aloia et al., 2013; Schuettengruber et al., 2017).

Our previous characterization of PRC1 complexes revealed that a subset of ncPRC1 that comprise either PCGF3 (ncPRC1.3) or PCGF5 (ncPRC1.5) unexpectedly contain three non-PcG proteins: AUTS2, casein kinase 2 (CK2), and P300 (Gao et al., 2012, 2014). Remarkably, these factors act in concert to not only subvert the repressive function of ncPRC1.3 or ncPRC1.5, but to convert the respective complex into a transcriptional activator (Gao et al., 2014). Specifically, CK2 phosphorylates the RING1A/B subunit of PRC1, thereby inhibiting its catalysis of H2AK119ub1 and thus, thwarting repression. AUTS2 interacts with PCGF3/-5 and importantly, recruits P300/CBP, a known transcriptional co-activator that conveys histone acetyltransferase activity (Bannister and Kouzarides, 1996; Ogryzko et al., 1996). These findings point to AUTS2 having a profound impact on gene expression in the context of defined aspects of development. Indeed, AUTS2-ncPRC1.3 is important during development of the central nervous system (CNS), as well as in the post-developmental stage (see below). In contrast, AUTS2-ncPRC1.5 appears to function during the establishment and maintenance of other cell lineages (our unpublished data).

The gene encoding AUTS2 was designated as such based on the identification of its translocation in a pair of monozygotic twins that were diagnosed with autism (Sultana et al., 2002); yet its role in Autism Spectrum Disorders (ASD) is still putative. Nonetheless, the role of AUTS2 in neurodevelopment has been more widely established through the identification of variants in *AUTS2* that are associated with various neurological diseases,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

including intellectual disability (ID), epilepsy, bipolar disorder and others (Hattori et al., 2009; Mefford et al., 2010; Elia et al., 2010; Oksenberg and Ahituv, 2013; Beunders et al., 2013; Kapoor et al., 2013; Hori and Hoshino, 2017). Notably, a thorough and informative report analyzing AUTS2 ChIP-Seq in E16.5 mouse forebrain reveals AUTS2 occupancy at gene promoters and enhancers whose function appear to be important during neurodevelopment (Oksenberg et al., 2014), and points to its substantive role in activating genes required for appropriate CNS development and function.

Here, we report heterozygous *de novo* variants of *AUTS2* in patients who exhibit a severe phenotype overlapping that of Rubinstein-Taybi syndrome (RSTS), a neurodevelopmental disorder characterized by distinctive facial features, short stature, and intellectual disability (Stevens, 1993; Ajmone et al., 2018). Given that AUTS2 interacts with P300, converting ncPRC1.3 into a transcriptional activator, it is notable that RSTS is mainly associated with heterozygous pathogenic variants in *EP300* or *CREBBP* (CREB Binding Protein/CBP). Strikingly, we found that the AUTS2 variants are defective in P300 interaction, underscoring the biological relevancy of P300 incorporation into AUTS2-ncPRC1.3 with respect to appropriate neurodevelopment and brain function in human.

We further determined the means by which AUTS2-ncPRC1.3 accesses specific chromatin sites in the brain to facilitate transcription and identified the transcription factor, Nuclear Respiratory Factor 1 (NRF1), as being instrumental to this process. Previous studies implicated NRF1 in mitochondrial biogenesis (Scarpulla, 2011) and retinal development (Hsiao et al., 2013; Kiyama et al., 2018). We found that NRF1 mediates AUTS2-ncPRC1.3 recruitment to a subset of neurodevelopmental genes during differentiation of mouse embryonic stem cells to motor neurons, as well as in the mouse brain during early development. These findings expose a novel and key role for NRF1 in facilitating appropriate AUTS2-ncPRC1.3-mediated activation of genes involved in neurodevelopment as a consequence of AUTS2-mediated recruitment of P300.

RESULTS

ncPRC1.3 occupies active genes during early development in mouse brain

The *Auts2* gene encodes 2 major AUTS2 isoforms in the mouse brain (Hori et al., 2014). The long isoform (AUTS2_L, 1–1261 aa, Figure 1B) arises from the full-length (FL) mouse Auts2 transcript (Auts2-1, Figure 1A), which includes 19 exons. The short isoform (AUTS2_S, 458-1261 aa, Figure 1B) (Hori et al., 2014), arises from an Auts2 transcript (Auts2–2, Figure 1A), which initiates from a transcriptional start site near exon 7 and comprises a translational start site in the middle of exon 8. Both AUTS2 isoforms contain a PY motif and an HX repeat that are located, in the case of AUTS2-L, within its N-terminus between two proline rich regions (PR1 and PR2) (Figure 1B). The PY motif is a potential WW domain-binding region in various activating transcription factors (Yagi et al., 1999; Lin et al., 2019). The HX repeat (aa 525–542) comprises alternating HQ (x6) or HT (x3) residues, mutations of which in two other genes (*ATN1* and *RERE*), are related to neurodevelopmental disorders (Jordan et al., 2018; Palmer et al., 2019, respectively).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

To investigate the role of AUTS2 in brain development, specifically in the context of the ncPRC1.3/1.5 complex, we examined its expression and that of the core PRC1 components in mouse brain throughout early development. Western blot analysis of whole mouse brain lysates using an antibody against the AUTS2 C-terminus [1160–1259 amino acids of human AUTS2 protein (Gao et al., 2014)], detected both the FL (~170 kDa) and shorter (~95 kDa) isoforms of AUTS2, with the latter being predominant (Figure 1C). The expression of both isoforms gradually decreased throughout early development (Figure 1C). In accordance, the expression of RING1B and PCGF3 subsided dramatically from postnatal day 5 (P5) (Figure 1C), as did that of PCGF5 (Figure 1C, also see below).

As AUTS2 is incorporated into both ncPRC1.3 comprising PCGF3 and ncPRC1.5 comprising PCGF5 in 293 T-REx cells (Gao et al., 2012), we examined the expression of PCGF3 and PCGF5 in whole brain lysates at postnatal day 1 (P1). RNA-Seq data revealed that *Pcgf3*, but not *Pcgf5*, was predominantly expressed in the mouse brain (Figure 1D). Following co-immunoprecipitation (co-IP) experiments using AUTS2 antibody and whole brain lysate, mass spectrometry (MS) revealed considerably more peptides from PCGF3 than from PCGF5 (Figure 1E). Two other ncPRC1.3 components, RING1A/B and CK2, were also observed (Figure 1E), as reported previously (Gao et al., 2014). Importantly, brain-specific conditional knockout of *Pcgf3* (*Pcgf3*$^{loxp/loxp}$:*Nes*^{Cre}) caused lethality (data not shown), suggesting a critical role for AUTS2-ncPRC1.3 during early brain development.

We next characterized the genomic localization of AUTS2, P300, and the RING1B, RYBP, and PCGF3 components of ncPRC1.3 using ChIP followed by deep sequencing (ChIP-seq) in whole brain lysate at postnatal day 1. Consistent with our previous ChIP-seq data in mouse brain (Gao et al., 2014), AUTS2 associated with ncPRC1.3 components, including P300 in the promoter of active genes that were devoid of histone post-translational modifications (hPTMs) associated with transcription repression, and instead exhibited strong signals for those associated with active transcription, *e.g.* H3K27ac and H3K4me3 (Figure 1F). Genome-wide analysis (Figure 1G), corroborated this finding and together provided strong evidence that AUTS2-ncPRC1.3 is involved in active transcription in the mouse brain. GO analysis revealed that terms related to RNA processing and neuronal development were enriched in genes within AUTS2-bound regions (Figure S1).

Patients with mutations in the *AUTS2* HX repeat share features with Rubinstein-Taybi syndrome

We previously reported that a truncated form of AUTS2 (404–913 aa) is sufficient to mediate transcription activation through its recruitment of P300 (Gao et al., 2014). Yet, the AUTS2 residues key to its interaction with P300 and the physiological relevance of this interaction during brain development remained largely unexplored. In these regards and as part of an ongoing effort to identify genetic variants associated with developmental brain disorders (Aldinger et al., 2019), we discovered a novel *de novo* missense variant in *AUTS2*. This missense variant was detected in a boy with multiple congenital anomalies and a proposed diagnosis of Rubinstein-Taybi syndrome (RSTS). His phenotype was more severe than the syndrome previously reported among individuals with heterozygous *AUTS2* deletions (Beunders et al., 2013). To investigate the clinical phenotype associated with

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

AUTS2 mutations, we identified 6 additional individuals with *de novo* intragenic variants that were clustered in exon 9 of *AUTS2* (NM_015570.2) (Figures 2A, S2 and Table S1). All 7 individuals displayed dysmorphic features and feeding difficulties in infancy, and most had moderate to severe intellectual disability and hypotonia (Table S1). Importantly, 5 of these individuals had severe phenotypes (Table S1), and harbored mutations within the HX repeat domain (aa 525–542), which comprises alternating HQ (x6) or HT (x3) residues. Our original proband (LR05–007) had a missense mutation, p.Thr534Pro, while the remaining four individuals (LR15–003, LR18–404, LR19–314, LR19–506) had an identical recurrent small deletion within exon 9: p.His535_Thr542del (Figure 2A).

All 5 patients with mutations in the HX repeat of AUTS2 had a dysmorphic facial appearance dominated by features seen in RSTS, although less severe than classic RSTS (Table S1). RSTS is a complex multiple congenital anomaly syndrome characterized by short stature, distinctive facial features, and varying degrees of intellectual disability (RUBINSTEIN and TAYBI, 1963; Wiley et al., 2003). A clinical diagnosis of RSTS was suggested for 2 of the 5 individuals prior to genetic testing (Figure 2B and Table S1). The two other patients with mutations outside of the HX repeat domain did not exhibit a phenotype overlapping that of RSTS (Figure 2B and Table S1), but instead displayed other neurological defects. Moreover, neuroimaging studies from 4 of the 5 patients with mutations in the HX repeat showed hypoplasia of the corpus callosum (n=3), cerebellar hypoplasia and small posterior fossa (n=4), and Chiari malformation type 1 (n=1) (Figure 2C and Table S1). Such brain malformations are also reported in individuals with classic RSTS (Cantani and Gagliesi, 1998; Ajmone et al., 2018). Given that RSTS in most individuals is associated with mutations in *CREBBP* or *EP300*, or a microdeletion of 16p13.3 that includes *CREBBP* (Stevens, 1993), we hypothesized that the HX repeat in AUTS2 coordinates with CBP/P300 in regulating proper gene expression in the brain.

Integrity of both the HX repeat in AUTS2 and ncPRC1.3/1.5 are required for P300 recruitment and transcription activation

That AUTS2 interacts with P300 in the context of ncPRC1.3 and that mutations in either the AUTS2 HX repeat or CBP/P300 are associated with RSTS, pointed to AUTS2-HX mutants being defective in P300 interaction. To examine this possibility, we initially expressed Flag-tagged AUTS2, either wild-type (WT) or patient-derived mutant forms including two variants within the HX repeat (T534P and 535–542aa del) and one outside this region (P517L within the PY motif), in 293 T-REx cells (Figures 3A and 2A). Strikingly, co-IP experiments using Flag-WT or Flag-mutant AUTS2 revealed that both mutations within the HX repeat (T534P and 535–542aa del), but not P517L, disrupted interaction with P300 (Figure 3B). This finding is in accordance with the clinical diagnosis of RSTS for patients harboring mutations within the HX repeat, but not the P517L mutation (Table S1). Of note, WT and all AUTS2 variants interacted stably with RING1B (Figure 3B), suggesting that AUTS2 regions outside the PY motif and HX repeat could mediate AUTS2 incorporation into the ncPRC1. Importantly, reciprocal co-IP experiments against endogenous P300 confirmed that AUTS2 mutated in its HX repeat disrupted P300 interaction (Figure 3C). Similar results were observed using a more biologically relevant system: cells undergoing differentiation into motor neurons (see below).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Since P300 is required for AUTS2-mediated transcriptional activation (Gao et al., 2014), we tested AUTS2 HX mutants for the ability to activate transcription. GAL4–AUTS2, either WT or mutant in the PY motif or the HX repeat, or GAL4 alone were inducibly expressed in 293 T-REx cells containing an integrated luciferase reporter with a UAS element comprising five consecutive GAL4 DNA binding sites (Figure 3D). Indeed, the GAL4-AUTS2 mutants in the HX repeat exhibited a severe defect in doxycycline-mediated induction of luciferase activity, compared to WT (Figure 3E). Although the mutation in the PY motif did not appear to affect AUTS2-P300 interaction (Figures 3B and 3C), it did attenuate AUTS2-mediated transcriptional activation (Figures 3D and 3E). As the PY motif associates with WW domain-containing proteins to facilitate transcriptional activation (Yagi et al., 1999; Lin et al., 2019), the integrity of both the PY motif and the HX repeat may be required for AUTS2-induced gene activation.

To corroborate the critical role of the HX repeat domain in mediating AUTS2-P300 interaction, we sought a more relevant system: *in vitro* neuronal differentiation (and see below). We genetically modified mESC to harbor one of two patient-derived mutations in the HX repeat of the endogenous *Auts2* locus (T534P or 535–542 aa deletion, Figures S3A and S3B, respectively), which could then be differentiated towards motor neurons (see Figure 5A for more details). Expression levels of AUTS2 were unaffected by either mutation (Figure S3C). Consistent with our observations in 293T cells, both mutations of the HX repeat domain of AUTS2 disrupted its interaction with P300, but not its incorporation into the ncPRC1 complex (Figures 3F and 3G).

Since AUTS2 associates with ncPRC1.3/1.5 core components and with P300 (Gao et al., 2012, 2014), we wondered if PRC1 core components are required for efficient P300 recruitment and transcription activation. Indeed, shRNA-mediated silencing of PCGF3 and PCGF5 (Figure S3D) led to a dramatic loss in both GAL4–AUTS2-mediated activation of the luciferase reporter (Figure 3H) and recruitment of P300 to the promoter of the reporter (Figure 3I). Altogether, these results establish that the AUTS2 HX repeat domain and ncPRC1.3/1.5 core components engage in efficient P300 recruitment for transcription activation.

Our results thus far underscore that the context in which P300 functions, *i.e.* ncPRC1.3, is of prime importance and likely relevant in the case of patients who express AUTS2 HX repeat mutants reported above and exhibit RSTS, a syndrome previously associated with CBP/P300 mutations (Stevens, 1993). As shown above, these AUTS2 mutants thwart P300 interaction, essentially rendering ncPRC1.3 target genes defective in P300-mediated activation. To understand the role of ncPRC1.3 in fostering appropriate P300 activity during neurodevelopment, we next sought the means by which ncPRC1.3 accesses its target genes and the importance of this process for appropriate neurodevelopment.

AUTS2 and NRF1 co-localize within chromatin and interact in the mouse brain

To identify the factor(s) involved in the key process by which AUTS2 accesses chromatin, we first determined the motifs of transcription factors (TFs) enriched in AUTS2-bound sites in the mouse brain and identified an overrepresentation for that of Nuclear Respiratory Factor 1 (NRF1) (Figure 4A). NRF1 is a TF involved in mitochondrial biogenesis

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(Scarpulla, 2011), which binds to GC-rich DNA elements in the promoters of many mitochondrial biogenesis-related genes (Evans and Scarpulla, 1990; Gleyzer et al., 2005). As well, NRF1 is associated with the regulation of neurite outgrowth (Chang et al., 2005; Tong et al., 2013), and exhibits essential roles in retinal development (Hsiao et al., 2013; Kiyama et al., 2018); yet its function and regulation in the CNS is largely unknown. To validate our computational prediction, we performed ChIP-seq for NRF1 using 2 different antibodies and lysates from whole mouse brain and ascertained that the majority of AUTS2 peaks (1545 of 2005 total peaks) were associated with NRF1 binding (Figures 4B and 4C).

Based on this high overlap between chromatin-bound AUTS2 and NRF1, we next tested the possibility that NRF1 might physically interact with AUTS2. Reciprocal co-IP assays using nuclear extracts from whole mouse brain revealed that indeed, endogenous NRF1 and AUTS2 were physically associated (Figure 4D). Notably, the core component of AUTS2-ncPRC1.3, PCGF3, also co-immunoprecipitated with NRF1, indicating that AUTS2 interaction with NRF1 occurred within the context of the ncPRC1.3 complex in mouse brain (Figure 4D). Moreover, the interaction between AUTS2 and NRF1 was independent of the AUTS2 HX repeat domain (Figure S3E). Interestingly, the expression of NRF1 recapitulated the pattern of AUTS2-ncPRC1.3 and CBP/P300 expression during early brain development (Figures 4E and 1C). These data strongly suggest that NRF1 contributes to the recruitment of AUTS2 and its associated ncPRC1 complex, although it is also clear that both have independent targets, likely due to their additional functions and partnership with other factors.

AUTS2 and NRF1 colocalize with ncPRC1.3 at actively transcribed loci in motor neurons

We next gauged the profile of chromatin binding and transcription regulation inherent to AUTS2-ncPRC1.3 and NRF1 in the context of a system by which differentiated motor neurons (MN) are attained (Wichterle et al., 2002; Mazzoni et al., 2013; Narendra et al., 2015). Under these conditions, the expression of AUTS2 and of PCGF3 were significantly up-regulated in MN, while that of NRF1 decreased by half at both the protein (Figure 5A), and RNA levels (Figure S4A). In contrast, the overall levels of PRC1 complex as reflected by that of RING1B were down-regulated (Figures 5A and S4A), consistent with its essential role in maintaining mESC identity (Endoh et al., 2008). To complement our previous ChIP-seq for RNA polymerase II (PolII) in both mESC and MN (Narendra et al., 2015; LeRoy et al., 2019), we performed similar ChIP-seq for AUTS2 and NRF1. The majority of regions that gained AUTS2 binding upon MN differentiation, also accumulated NRF1 binding (Figure 5B). Importantly, these regions became actively transcribed during differentiation as evidenced by increased PolII binding (Figure 5B).

To examine whether AUTS2 cooperates with ncPRC1.3 for active transcription in MN, we analyzed the genome-wide distribution of a set of hPTMs (H2AK119ub1, H3K27me3, H3K27ac and H3K4me3) and the core ncPRC1 subunits, RING1B and RYBP. Consistent with studies in other systems (Kloet et al., 2016; Cohen et al., 2018; Loubiere et al., 2020), k-means clustering revealed three discrete classes of RING1B-bound regions in MN (Figure 5C). In cluster 1, we observed strong and broad ChIP-seq signals for RING1B, H2AK119ub1, and H3K27me3, and the absence of signals for H3K27ac, H3K4me3, PolII

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and the ncPRC1 component, RYBP/YAF2 (RYBP antibodies do not distinguish RYBP and YAF2), suggesting that these regions are co-repressed by PRC2 and cPRC1. The second cluster exhibited lower levels of H2AK119ub1 and H3K27me3, and increased levels of H3K27ac and H3K4me3 mostly abundant at the peak center, suggesting that these RING1B-bound regions featured bivalency (Bernstein et al., 2006; Voigt et al., 2013). The third cluster exhibited both H3K27ac- and H3K4me3-marked active regions enriched for developmental GO terms (Figure S4B). Notably, both AUTS2 and NRF1 binding were specifically enriched in cluster 3, which also exhibited elevated levels of the ncPRC1 components, RYBP/YAF2 (Figure 5C). Accordingly, genes flanking cluster 3 regions were expressed at levels significantly higher than those in clusters 1 and 2 (Figure 5D). These data strongly suggest that NRF1 associates with AUTS2-ncPRC1 to facilitate transcription in MN.

NRF1 directs AUTS2-ncPRC1 chromatin binding

To ascertain whether AUTS2 binding to chromatin is dependent on NRF1 or vice versa, we first performed ChIP-seq for the presence of AUTS2 or NRF1 as a function of either NRF1 or AUTS2 depletion, respectively. Exon 9 of the *Auts2* gene was targeted by CRISPR-Cas9 to remove both the long and short forms of the protein (Figures S4C–E, Table S2), while exon 4 of the *Nrf1* gene was targeted by CRISPR-Cas9 to remove NRF1 protein (Figures S4F–H, Table S2) in mESC from which MN were then derived. As a consequence of NRF1 depletion, AUTS2-binding at most of its targets were decreased in MN (Figures 5E and S5A). However, NRF1 ChIP-seq signals remained largely unaltered upon AUTS2 depletion (Figures 5F and S5B), demonstrating that while AUTS2 binding is NRF1-dependent, NRF1 binding to chromatin is AUTS2-independent.

We next probed how NRF1-directed AUTS2 binding might modulate ncPRC1.3-associated active transcription. Under NRF1-depleted conditions, we performed ChIP-seq for RING1B and a set of hPTMs (H2AK119ub1, H3K27me3, and H3K27ac), followed by k-means clustering analysis. Remarkably, Ring1B binding in the NRF1-KO MN was lost at the majority of cluster 3 regions (1030 of 1171 peaks) found in WT (Figure 5G). Moreover, we did not recover the cluster of RING1B-bound active regions (labeled by H3K27ac, but not by H3K27me3 or H2AK119ub1), relative to the control (Figures 5H and 5C), strongly suggesting that the absence of NRF1 was detrimental to ncPRC1.3-mediated active transcription. As evident in Figure 5C, the H3K27ac signal at cluster 3 regions was dramatically reduced in NRF1-KO MN, while the H3K27me3 signal was increased (Figure 5I). Collectively, these data demonstrate a pivotal role for NRF1 in facilitating ncPRC1.3-associated active transcription by directing AUTS2 binding to chromatin in MN.

We attempted to delete NRF1 in the mouse brain from the embryonic stage in order to validate the NRF1-mediated AUTS2 recruitment we observed in cell culture, but such embryos did not survive, consistent with a previous report that *Nrf1*-null mouse embryos die between embryonic day 3.5 (E3.5) and E6.5 (Huo and Scarpulla, 2001). Instead, we chose the $Tbr1^{CreERT2}$ line to strategically delete *Nrf1* in the adult mouse brain to examine its role in the CNS (Figures S6A–C, see Methods for details). Importantly, we noticed several histological anomalies in these mutant mice ($Tbr1^{CreERT2/+}$: $Nrf1^{fx/fx}$: $Pou4f1^{CKO/+}$),

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

compared to control mice (*Tbr1*$^{CreERT2/+}$*: Nrf1*$^{fx/+}$*: Pou4f1*$^{CKO/+}$), including a reduction in the size of the hippocampus and the width of the corpus callosum, as well as an enlarged lateral ventricle, indicating neuronal loss in both the cortex and hippocampus (Figures S6D and S6E). As well, a significant loss in retinal ganglion cells (RGCs) was observed in retinas collected from these mutant mice (Figures S6F and S6G), consistent with its previously elucidated function in retinal development (Kiyama et al., 2018).

The HX repeat in AUTS2 and NRF1-directed binding are required for PNP to MN differentiation

That both AUTS2 and NRF1 might coordinately regulate the process of neuronal differentiation has not been previously recognized. Thus, we next characterized gene expression profiles during the transition from multipotent, posterior neural progenitors (PNPs) to terminally differentiated MNs as a function of the presence of AUTS2 or NRF1. We performed single-cell RNA sequencing (scRNA-seq) on MN differentiated from mESC under conditions of *Auts2* knockout (*Auts2*-KO), mutations in the AUTS2 HX repeat (*Auts2*-HX*, 535–542 aa del) or *Nrf1* knockout (*Nrf1*-KO) (see Methods), compared to WT. We obtained 816 high-quality (cells with >3000 detected genes) single-cell transcriptomes from all samples (WT, 228 cells; *Auts2*-KO, 221 cells; *Auts2*-HX*, 184 cells; *Nrf1*-KO, 183 cells) for in-depth analyses (Figures 6A and S7A). To identify major cell types, we performed unsupervised clustering on a graph-based representation of the cellular transcriptomes. 5 major clusters were visualized in a uniform manifold approximation and projection (UMAP) embedding (Butler et al., 2018; Becht et al., 2019), as represented by color-coded dashed-line circles (Figure 6A). Clusters were annotated according to known markers and previously established lineage information (Wichterle et al., 2002; Briggs et al., 2017); *i.e.*, posterior neural progenitor (PNP) expressing *Sox3*, posterior and ventral neural progenitor (PVNP) expressing *Hoxd4*, newborn motor neuron (NMN) expressing *Neurog2*, and motor neuron (MN) expressing *Mnx1* and *Chat* (Figures S7B and S7C). Importantly, cells from all samples were clustered by cell type identity rather than sample identity (Figure 6A), indicating little or no batch effect. Notably, *Auts2* and *Nrf1* were highly expressed in all cell types (Figure S7C), suggesting their involvement in all of the different stages of differentiation.

To pursue the potential functional requirement of the intact HX repeat in AUTS2 and of NRF1 for proper MN differentiation, we first compared the percentage of MN (including NMN and MN) and PNP under WT, *Auts2*-KO, Auts2-HX* and *Nrf1*-KO conditions. The percentage of PNP was retained at a much higher level in the case of either mutant (52% in WT, compared to 67% in *Auts2*-KO, 60% in *Auts2*-HX* and 68% in *Nrf1*-KO, Figure 6B). Moreover, we observed a slight decrease in the percentage of MN in *Auts2*-KO and *Auts2*-HX* (32% in WT, 26% in both *Auts2*-KO and *Auts2*-HX*) and a more severe defect in *Nrf1*-KO (32% to 16%) (Figure 6B). The lower percentage of MN under these mutant conditions was not due to apoptosis (Figure S7D), but instead, pointed to a defect in PNP differentiation into MN. To gain more insight into the molecular mechanism by which AUTS2 and NRF1 contribute to the transition from PNP to MN, we further analyzed the differentially expressed genes (DEGs) specifically in the MN population from either WT or *Auts2*-KO (see Methods). Among the top 500 DEGs, 458 genes were down-regulated

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

in *Auts2*-KO (Figure 6C). Strikingly, 405 of these 458 genes were also down-regulated in *Auts2*-HX mutant MN, in accordance with the critical role of the HX repeat domain in mediating AUTS2-P300 interaction and transcriptional activation (Figures 6C, and 3B–G). Furthermore, about half of these 458 genes (205 of 458) were also down-regulated in *Nrf1*-KO MN compared to WT MN (C1, C2 and C3 labeled on the right, Figure 6C). To confirm whether NRF1-directed AUTS2 binding is required for the transcriptional activation of AUTS2-ncPRC1.3 associated active genes (cluster 3 region, Figures 5C and 5H), we compared the DEGs in *Nrf1*-KO MN with the genes located in cluster 3 and cluster 1 regions from the RING1B ChIP-Seq. Consistent with a KO of *Nrf1* leading to a loss in AUTS2-ncPRC1 binding to active genes (Figures 5C, 5G and 5H), genes located in cluster 3 regions were enriched in the down-regulated category in *Nrf1*-KO MN (Figure S8). In contrast, very few of the genes located in the cluster 1 region (co-repressed by cPRC1 and PRC2) were affected in the *Nrf1*-KO MN (Figure S8). These results strongly suggest that AUTS2 and NRF1 function coordinately in regulating transcriptional activation (Figures 5C and 5H).

We next asked whether the genes that were down-regulated in *Auts2*-KO MN reflected those associated with the transition from PNP to MN. Indeed, 186 of the 458 genes were normally up-regulated during PNP to MN differentiation in WT, but were defective in activation in the *Auts2*-KO (C1, C2 and C3 labeled on the left, Figure 6C). For example, up-regulation of *Asic2*, a member of the sodium channel superfamily that regulates synaptic function (Zha et al., 2009) and of *Pnpla6*, a phospholipase that functions in neurite outgrowth (Guerreiro et al., 2015) were significantly attenuated under conditions of AUTS2 depletion. Finally, genes that were down-regulated in *Auts2*-KO MN were enriched for GO terms related to neuronal differentiation and function (Figure 6D).

As shown above, we demonstrated that both the AUTS2-HX repeat domain and NRF1 are required for AUTS2-mediated transcription activation in the context of ncPRC1.3. Comparison of the expression levels of AUTS2-NRF1 co-targeted genes demonstrated that the vast majority of NRF1/AUTS2 co-targets (674 of 837 genes) were down-regulated in AUTS2 HX mutant MN compared to WT (Figure 6E). Collectively, these data demonstrate that AUTS2 and NRF1 function coordinately to foster the appropriate differentiation of PNP to MN by directly binding to and activating a subset of the relevant genes.

Discussion

Taken together, our findings support a model in which NRF1 directs AUTS2-ncPRC1.3 binding to a subset of neuronal differentiation-related genes that are thereby subjected to activation by P300 through its interaction with the AUTS2 HX repeat domain (Figure 7). Given the existence of ncPRC1.3 in which AUTS2 conveys transcription activation, along with the reported association of AUTS2 haploinsufficiency in AUTS2-syndrome (Beunders et al., 2013), and possibly in ASD (Sultana et al., 2002), AUTS2 had appeared key for regulating appropriate neurodevelopment. Here, we identified the critical role of its HX repeat domain by interrogating mutations found in individuals exhibiting a distinct and severe neurodevelopmental syndrome that overlaps with RSTS. While RSTS is largely associated with pathogenic variants in the *EP300*/*CREBBP* genes, the patient-associated

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

AUTS2-HX mutants reported here essentially reflect a defect in P300 function. Our *in vitro* and cell-based studies demonstrated that RSTS-associated AUTS2 mutations in the HX repeat domain, disrupt AUTS2-P300 interaction and attenuate AUTS2-mediated active transcription. Of note, a recent study reports a patient with a syndromic neurodevelopmental disorder harboring a different mutation (532–541 aa deletion) in the AUTS2 HX repeat domain (Martinez-Delgado et al., 2020), further pointing to its critical role in normal brain functioning. Moreover, mutations within exon 9 outside the HX repeat such as the PY motif, as well as a mutation at residue 495, result in individuals that display severe behavioral phenotypes such as epilepsy, in lieu of RSTS (Table S1); further stressing the role of AUTS2 in normal brain function. Interestingly, a recent report shows that AUTS2 controls neuronal differentiation in a PRC1-independent manner through BMP inhibition, pointing to the multifaceted actions of AUTS2 in neurodevelopment (Geng et al., 2021).

Our study directs attention to NRF1 in facilitating chromatin access by AUTS2, key to its role at the appropriate target genes. We found that most AUTS2 binding events require NRF1, while most NRF1 binding is AUTS2-independent. A previous report also noted that the motif of NRF1 is significantly enriched in AUTS2-bound regions (Oksenberg et al., 2014), although the enrichment is less dramatic than observed here. On the other hand, a recent study identified TF USF1/2 as being key to PCGF3 chromatin binding in mESC (Scelfo et al., 2019); yet the motif corresponding to USF1/2 was not recovered in our study, suggesting that cell type/tissue specific mechanisms might dictate ncPRC1.3 recruitment to chromatin.

Our previous mechanistic findings resolved the means by which the function of ncPRC1.3/1.5 comprising AUTS2, P300, CK2, PCGF3 or PCGF5, RING1A and/or -1B, and RYBP or YAF2 is converted from a typical PRC1 that facilitates transcription repression to that of a transcriptional activator (Gao et al., 2014). This conversion involves AUTS2-mediated recruitment of P300 and CK2-mediated phosphorylation of serine 168 of the integral PRC1 subunit, RING1, which thwarts its catalysis of H2AK119ub1 (Gao et al., 2014). Given that RYBP/YAF2 within other ncPRC1 complexes stimulate such RING1A/RING1B-mediated ubiquitination, their presence within ncPRC1.3 may indicate additional function(s) for these components. Of note, distinct ncPRC1.3/1.5 complexes comprise another non-PcG protein, FBRSL1, which shares a high degree of sequence similarity with AUTS2 (Gao et al., 2012). Interestingly, AUTS2 and FBRSL1 bind competitively to the respective PCGF subunit of ncPRC1.3/1.5 (Gao et al., 2014). Consistent with our data in supporting the importance of ncPRC1.3 in neurodevelopment, mutations in FBRSL1 are also associated with a neurodevelopmental syndrome (Ufartes et al., 2020). However, it has yet to be determined whether FBRSL1-ncPRC1.3 acts in a repressive or activating manner.

As shown here, AUTS2 interacts with P300 through its HX repeat domain and NRF1 associates with AUTS2, yet intriguingly, AUTS2 and P300 convey efficient transcription activation only in the context of ncPRC1.3/1.5. We speculate that ncPRC1.3 provides a conformational context within which P300 recruitment is facilitated or stabilized and/or its activity is optimized. Perhaps, this module exposes an interacting surface(s) for NRF1, as well. The characterization of NRF1 performed by others demonstrate its ability to dimerize and its phosphorylation at several serine residues within its amino-terminus (Gugneja and

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Scarpulla, 1997). These phosphorylation events do not regulate NRF1 dimerization, but instead mutation of these sites compromise NRF1 DNA binding activity (Gugneja and Scarpulla, 1997). The study also indicates that CK2 could stimulate the DNA binding activity of NRF1 *in vitro*. As CK2 is an integral component of ncPRC1.3 that inhibits its repressive activity by phosphorylating its RING1A/B component (Gao et al., 2014), CK2 might also promote ncPRC1.3-mediated transcription activation by enhancing NRF1 binding activity *in vivo*, resulting in the optimal activation of ncPRC1.3-AUTS2 target genes in the brain.

Evidence involving NRF1 have highlighted its importance in mitochondrial biogenesis (Scarpulla, 2011) and retinal development (Hsiao et al., 2013; Kiyama et al., 2018). Intriguingly, the pathways fostering mitochondrial integrity might be critical to those regulating distinct developmental pathways. It is important to emphasize that the reduced volume of corpus callosum observed in *Nrf1* mutant ($Tbr1^{CreERT2/+}$; $Nrf1^{fx/fx}$; $Pou4f1^{CKO/+}$) mice (Figure S6), is a key syndromic feature observed in RSTS patients (Cantani and Gagliesi, 1998), as well as in patients harboring mutations in the AUTS2 HX repeat domain reported here, strongly supporting that NRF1 and AUTS2 function coordinately in regulating brain development. The corpus callosum (CC) connects the cerebral hemispheres and is the largest fiber tract in the brain (Edwards et al., 2014). During development, defects in neurogenesis, telencephalic midline patterning, neuronal migration and specification, axon guidance, and post-guidance development can interrupt CC formation (Reyes et al., 2020).

As shown here, ablation of NRF1 or AUTS2 led to defective progenitor to MN differentiation *in vitro*. Deletion of NRF1 in the mouse brain and mutations in the AUTS2 HX repeat domain in human show CC malformation, a common feature of RSTS patients (Figure 7). The precise mechanisms by which NRF1 and AUTS2-ncPRC1.3 coordinately regulate gene expression, neuronal differentiation and thus, mouse and human brain development and function in postnatal individuals require further investigation. Alternate approaches are needed as deletion of either NRF1 or AUTS2 (both long and short isoforms) lead to early embryonic lethality. The generation of mouse models carrying mutations in the AUTS2 HX repeat domain might expedite future studies should they recapitulate the RSTS phenotype or other neurological diseases.

Limitations of the Study

The data presented here use homozygous mutations/knockout cellular models due to the lack of appropriate materials from patients that exhibit heterozygous mutations in AUTS2. This limitation of our study awaits a future determination as to whether the function of the AUTS2 HX repeat domain is essential for brain development *in vivo*. We generated homozygous knock-in mutations in the HX repeat domain of AUTS2 in mESC, which were then differentiated to MN to study how the mutations found in patients might affect neurodevelopment. Yet, the patients reported in this study harbor heterozygous mutations. Thus, future functional studies based on human iPSCs or mouse model(s) will provide further conceptual advances towards understanding the etiology of RSTS. Another limitation is that the identification of NRF1-mediated AUTS2-ncPRC1.3 recruitment arises from

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

ChIP-Seq analyses of AUTS2 and NRF1 in whole brain extract having assorted cell type specificities. While NRF1 expression is high in the hippocampus and low in other brain regions *(e.g.*, cortex), the levels of AUTS2 and PCGF3 are abundant in these NRF1-deficient regions. Thus, our studies are limited to the brain regions having NRF1 expression. Factors other than NRF1, *e.g.* USF1/2, likely contribute to AUTS2-ncPRC1.3 recruitment in brain regions deficient in NRF1 expression.

STAR*METHODS

Resource availability

Lead contact—Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Danny Reinberg (danny.reinberg@nyumc.org).

Materials Availability—All unique/stable reagents generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

Data and Code Availability

- The accession numbers for the raw data FASTQ files and processed files for all sequencing data are deposited in NCBI GEO are GEO: GSE161808. Original gel imaging data can be accessed from Mendeley: DOI: doi:10.17632/69vsfxr2n6.1.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Experimental Model and Subject Details

Animals—We did not observe any sex/gender influence on results derived from mice or patients. All mice were housed with a 12 hour light-dark cycle. Mixed cohorts of female and male mice were used for all experiments to minimize gender effects. Mice used for ChIP-Seq and biochemical assays were from various developmental stages as indicated in Figure legend. Mice used for Tamoxifen injection to delete NRF1 were 2 to 3 months old. All animal procedures followed the US Public Health Service Policy on Humane Care and Use of Laboratory Animals and were approved by the Animal Welfare Committee at New York University and the University of Texas McGovern Medical School at Houston.

Cell lines and culture condition—All ESC lines (E14 and derivatives) were grown in DMEM supplemented with 15% FBS, L-glutamine, penicillin/streptomycin, sodium pyruvate, non-essential amino acids, 0.1 mM β-mercaptoethanol, LIF, and 2i inhibitors, which include 1 μM MEK1/2 inhibitor (PD0325901) and 3 μM GSK3 inhibitor (CHIR99021) on 0.1% gelatin coated plates.

HEK293 T-Rex and HEK293T 5XGal4TK-Luc cells were cultured in standard DMEM supplemented with 10% FBS, 100 U/mL penicillin-streptomycin.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

WT and mutant NFH-AUTS2 inducible cell lines were obtained by transfecting each pINTO-NFH plasmid into 293 T-Rex cells, and WT and mutant Gal4-AUTS2 inducible cell lines were obtained by transfecting each pINTO-Gal4 plasmid into HEK293T 5XGal4TK-Luc cells. Transfected cells were seeded at limiting dilutions, and isolated clones were screened by western blot.

Clinical Cohort—The initial proband (LR05–007) was identified through trio-based exome sequencing in a cohort of 100 individuals with cerebellar malformations (Aldinger et al., 2019). We recruited 6 additional individuals with *de novo* pathogenic or likely pathogenic variants in *AUTS2* by sharing data through GeneMatcher (Sobreira et al., 2015) or through collaboration with colleagues. The 7 individuals in this cohort (2 females, 5 males) were 1 to 15.5 years of age at the time of their most recent evaluation. We obtained clinical data for all patients, including features tabulated in a reported AUTS2 clinical severity score (Beunders et al., 2013, 2015, 2016). Written informed consent was obtained from all participants through protocols approved by Institutional Review Boards at the local institution or at Seattle Children's Hospital.

METHOD DETAILS

Immunoprecipitation and Proteomics—Cell pellets were prepared from cell culture plates or mouse brain. Nuclei were extracted using HMSD buffer (20 mM HEPES, pH 7.5 at 4°C, 5 mM $MgCl_2$, 250 mM sucrose, 25 mM NaCl, 1 mM DTT) supplemented with protease inhibitors (0.2 mM PMSF, 1 μg/mL Pepstatin A, 1 μg/mL Leupeptin, and 1 μg/mL Aprotinin) and phosphatase inhibitors (10 mM NaF and 1 mM Na_3VO_4), and incubated on ice for 5 min. Lysates were pelleted at 3,000 rpm for 5 min at 4°C and nuclei pellets were washed one more time with HMSD buffer. The resulting nuclei pellets were resuspended in BC420 high salt buffer (20 mM Tris-HCl at pH 7.9, 1.5 mM $MgCl_2$, 0.42 M NaCl, 0.5 mM DTT, 0.2 mM EDTA) supplemented with protease inhibitors and phosphatase inhibitors for lysing at 4°C for 1 hr with occasional pipetting. Lysates were then pelleted at 20,000 × g for 15 min at 4°C. Finally, supernatants were collected and subjected to dialysis in Buffer D (20 mM HEPES, 150 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, and 5% glycerol) overnight at 4°C. Prior to any subsequent applications, nuclear extracts were centrifuged again at 20,000 g for 10 min at 4°C to remove any precipitate. Supernatants were collected, and protein concentrations were quantified via bicinchonic acid (BCA) assay. For immunoprecipitation, 1–2 mg of nuclear extract was incubated with 1~3 μg of antibody. After incubation at 4°C for 2 hr, 30 ul of protein G beads were added for incubation at 4°C overnight. Beads were washed three times with Buffer D (20 mM HEPES, 150 mM NaCl, 1.5 mM $MgCl_2$, 0.2 mM EDTA, and 5% glycerol), and eluted with 0.2 M glycine (pH 2.6) or 1× SDS loading buffer. Proteins from immunoprecipitation were separated by SDS–PAGE, using 4%–12% NuPAGE Novex Bis–Tris gels and then stained with Coomassie Blue. Bands were excised from gels and digested with trypsin, followed by standard LC-MS/MS procedure.

Whole cell extract and western blotting—Cells were harvested and lysed with CHAPS-Urea buffer (50 mM Tris-HCl, pH 7.9, 8M Urea, and 1% CHAPS) containing protease inhibitors and phosphatase inhibitors as mentioned above. The cell suspension was briefly sonicated (40% amplitude, 5 strokes) and centrifuged at 20,000 × g at 4°C

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

for 20 min. The supernatant was collected and protein concentrations were quantified via bicinchonic acid (BCA) assay. Proteins were separated using a 6%–12% SDS PAGE gel, and transferred onto a PVDF membrane. Membranes were blocked with 5% milk in PBST at RT for 1 hr and incubated with primary antibody overnight at 4°C. Membranes were washed 3 times with PBST and then incubated with HRP-conjugated secondary antibodies for 1 hr at RT, followed by exposure.

shRNA-mediated genome editing—Oligos designed for knockdown (see Table S2) were annealed and cloned into PLKO.1-shRNA vector using Age1 and EcoR1 restriction sites. The resulting plasmids were transfected into 293TRex cells using Neon Transfection System with parameter: 1500V, 1 pulse of 30ms. The cells were subject to puromycin selection at 1 ug/ml for 4 days. The efficiency of knockdown was confirmed by western blot.

CRISPR-mediated genome editing—To generate stable *Auts2* and *Nrf1* KO cell lines, sgRNAs were designed using CRISPR design tool in https://benchling.com. sgRNAs in Table S1 were cloned in pSpCas9(BB)-2A-GFP (PX458, a gift from Feng Zhang, Addgene plasmid #48138) and transfected into mESCs, using Lipofectamine 2000 (Life Technologies). GFP-positive cells were sorted 48 hr after transfection and 20,000 cells were plated on a 15 cm dish. Single mESC was allowed to grow to a colony for ~5 days and then was picked, trypsinized in Accutase for 5 min, and split into two individual wells of two 96-well plates for genotyping and culture, respectively. Genomic DNA was extracted using lysis buffer (50 mM Tris-HCl, pH 8, 2 mM NaCl, 10 mM EDTA, 0.1% SDS) supplemented with protease K, and genotyping PCRs were performed using primers (Table S2) surrounding the target site. The resulting PCR products were sent for sequencing to determine the presence of a deletion or a mutation event. Clones were further confirmed by western blot.

For endogenously knock-in the mutations in AUTS2 HX repeat in mESC, a ssODN donor used for homology directed repair and CRISPR/Cas9 plasmid (px458) with designed sgRNA (see Table S2 also) were co-transfected into mESCs, using Lipofectamine 2000. The following FACS, colony picking and characterization are the same as generating KO lines described above.

Motor Neuron Differentiation—E14 mouse embryonic stem cells (mESCs) were cultured in standard medium supplemented with LIF, and 2i conditions as described above. For motor neuron (MN) differentiation, the previously described protocol was applied (Narendra et al., 2015). Briefly, about 4 million mESCs were plated in a 500 cm^2 square dish and differentiated into embryoid bodies in AK medium (250 ml advanced DMEM/F12, 250 ml neurobasal medium, 75 ml knockout serum, L-glutamine, penicillin/streptomycin, 0.1 mM β-mercaptoethanol) for 2 days. Embryoid bodies were then diluted by 1:4 and further patterning was induced by freshly adding 1 μM all-transretinoic acid (RA) and 0.5 μM smoothened agonist (SAG) for an additional 4 days. Fresh medium was added after 2 days to support motor neuron survival.

Luciferase reporter assay—HEK293T 5XGal4 TK-Luc cells stably transfected with pINTO-GAL4 vector control or with inserts of interest were treated with 100 ng/ml doxycycline. Cells were lysed by adding 250 ul of ice-cold lysis buffer (0.2% Triton X-100,

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

100 mM potassium phosphate, pH 7.8, and 1 mM DTT) and shaking for 10 min at 4°C. The cell lysate was centrifuged at 20,000g for 10 min and the protein concentration of the resulting supernatant was determined by bradford Assay. 30 ug of the supernatant was assayed for luciferase activity using luciferase assay substrate (Promega).

X-gal staining—Mouse brain was fixed by perfusion with 10% neutral buffered formalin. Extracted brain was embedded in OCT compound then sectioned into 50 um thickness. Sections were dried at RT for 3 hr and then washed with wash buffer (0.1 M sodium phosphate containing 2 mM $MgCl_2$, 0.01% deoxycholate, and 0.02% Nonidet P-40, pH 7.3). LacZ color reaction was performed in wash buffer containing 5 mM potassium ferrocyanide, 5 mM potassium ferricyanide, and 1 mg/ml X-gal at 37°C overnight. Color reaction was terminated by incubation in 10% formalin for 10 min. Post-fixed sections were washed, dehydrated, and mounted with Cytoseal 60 (Thermo Fisher Scientific). Images were collected with a Canon EOS 10 digital camera (Melville, NY) mounted on an Olympus IX71 microscope.

Immuno-histochemical analysis—Mouse brain was fixed by perfusion with 10% neutral buffered formalin. Extracted brain was embedded in OCT compound, and then sectioned into 100 um thickness. Sections were incubated with anti-GFP (1:1000, Invitrogen) antibody. Alexa-488 conjugated secondary antibody was used in 1:800 dilution (Jackson Immuno-reserach).

Alkaline phosphatase (AP) staining— $Tbr1^{CreERT2/+}:Nrf1^{fx/+}:Pou4f1^{CKO/+}$ and $Tbr1^{CreERT2/+}:Nrf1^{fx/fx}:Pou4f1^{CKO/+}$ mice (2 to 3 months old) were intraperitoneally injected with tamoxifen (5 consecutive daily injections of 100 ug/g body weight), and then used for AP staining 3 months after tamoxifen injection. Brains were fixed by perfusion with 10% neutral buffered formalin. Fixed brain was embedded in OCT compound and sectioned into 100 um thickness. Retinal flat-mounts were fixed in 10% neutral buffered formalin for 10 min at RT. Brain sections and retinas were incubated in heated PBS for 30 min in a 65°C water bath to inactivate endogenous AP activity. AP color reaction was performed in 0.1 M Tris, pH 9.5, 0.1 M NaCl, 50 mM $MgCl_2$, 0.34 g/ml nitroblue tetrazolium and 0.175 g/ml 5-bromo-4-chrolo-3-indolyl-phosphate for overnight at RT. Stained tissues were washed three times in PBS, post-fixed with 10% neutral buffered formalin, dehydrated with a series of ethanol, then cleared with 2:1 benzyl benzoate/benzyl alcohol. Tiled images were collected using a Zeiss Axio Imager 2 microscope (Carl Zeiss).

ChIP-seq library preparation—For cross-linking, ESCs were fixed in 1% formaldehyde for 10 min at RT directly on plates and quenched with 125 mM glycine for 5 min at RT. For cross-linking of MN, ESC-derived motor neuron cultured for 6 days were dissociated with 0.05% trypsin, neutralized, fixed in 1% formaldehyde for 15 min at RT and then quenched with 125 mM glycine for 5 min at RT. For cross-linking of mouse brain, mouse whole brains were quickly dissected at postnatal day one and homogenized with a glass douncing homogenizer using first a loose, then a tight pestle. The cell homogenate was fixed with a final concentration of 1% formaldehyde for 10 min at RT and the reaction was quenched with 0.125 M glycine for 5 min at RT.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Cell pellets were washed twice in PBS and nuclei were isolated using buffers in the following order: LB1 (50 mM HEPES, pH 7.5 at 4°C, 140 mM NaCl, 1 mM EDTA, 10% Glycerol, 0.5% NP40, 0.25% Triton X; 10 min at 4°C), LB2 (10 mM Tris, pH 8 at 4°C, 200 mM NaCl, 1 mM EDTA, 0.5 mM EGTA; 10 min at 4°C), and LB3 (10 mM Tris, pH 7.5 at 4°C, 1 mM EDTA, 0.5 mM EGTA, and 0.5% N-Lauroylsarcosine sodium salt). Chromatin was fragmented to an average size of 250 bp in LB3 buffer using a Diagenode Bioruptor. 200 μg sonicated chromatin, 4 ug antibody and 20 ul Dynabeads were used in each ChIP reaction supplemented with 0.5x volumn of incubation buffer (3% Triton X, 0.3% Na Deoxycholate, 15 mM EDTA). 1 μg of Drosophila chromatin and 0.2 μg of anti-Drosophila H2A.X antibody were added in each ChIP reaction as spike-in references, except 3 μg of Drosophila chromatin was used for H3K27ac and H2AK119ub ChIP. After 5 consecutive washes with RIPA buffer (50 mM HEPES, pH 7.5 at 4°C, 0.7% Na Deoxycholate, 1 mM EDTA, 1% NP40, 500 mM LiCl) and one wash with TE+50 mM NaCl, the beads-bound DNA was eluted in freshly prepared elution buffer (50 mM Tris, pH 8, 10 mM EDTA, 1% SDS) at 65°C for 20 min. Eluted DNA was de-crosslinked at 65°C overnight, followed by protease K and RNase A treatment.

For Library preparation, IP'ed DNA (~1–30 ng) was end-repaired using End-It Repair Kit, tailed with deoxyadenine using Klenow exo-, and ligated to custom adapters with T4 Rapid DNA Ligase (Enzymatics). Fragments of 200–600 bp were size-selected using Agencourt AMPure XP beads (0.5X and 0.3X), and subjected to PCR amplification using Q5 DNA polymerase. Libraries were size-selected using Agencourt AMPure XP beads (0.75X), quantified by Qubit™ dsDNA HS Assay Kit and quality checked by High Sensitivity D1000 ScreenTape. Libraries were sequenced as 50 bp single-end reads on the Illumina HiSeq 4000 platform.

RNA-seq library preparation—Total RNA was isolated with Tripure isolation reagent and gDNA was removed by RNeasy Plus mini kit. PolyA+ RNA was isolated from 5 ug total RNA using Dynabeads Oligo$(dT)_{25}$, fragmented with Mg^{2+} contained in the 1st strand buffer at 94°C for 15 min, and reverse transcribed using Superscript III and random hexamers to synthesize the first strand cDNA. Single strand cDNA was precipitated and dTTP was removed by phenol/chloroform/isoamyl alcohol extraction and ethanol precipitation. Second strand cDNA was synthesized with dUTP to generate strand asymmetry using DNA Pol I and E. coli ligase, and then purified by MinElute PCR Purification Kit. Double-stranded DNA was end-repaired, A-tailed, and ligated to custom barcode adapters as described above. RNA-seq libraries were sequenced as 50 bp paired-end reads on the Illumina HiSeq 4000 platform or NovaSeq 6000 platform.

Single cell RNA-seq library preparation—For single cell RNA-seq library preparation, we chose the newly developed Smart-seq3 technique (Hagemann-Jensen, 2020; Hagemann-Jensen et al., 2020), an improved version of Smart-seq2 with a 5'-unique molecular identifier RNA counting strategy and a much higher sensitivity that detects thousands more transcripts per cell. The library was generated according to the published protocol (Hagemann-Jensen, 2020; Hagemann-Jensen et al., 2020), with the following modifications. ESC-derived motor neuron cultured for 6 days were dissociated with 0.05%

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

trypsin and stained with 0.2 uM Calcein AM and 8 uM Ethidium homodimer-1 at RT for 15 min. Single viable cells were sorted using Fluorescence Activated Cell Sorting to single wells of 96-well fully-skirted Eppendorf PCR plates in 3 uL lysis buffer (5% PEG8000, 0.1% Triton X-100, 0.5 unit/ul RNAse Inhibitor, 0.5 uM OligodT_{30}VN, 0.5 mM dNTP in nuclease free water). The plates were immediately covered, spun at 2000 rpm for 1 min at 4°C, and stored at −80°C until further analysis. In each plate, well A1 was left empty and 100 cells were sorted to well H1 for quality control and they were excluded for downstream analysis.

Plates were incubated at 72°C for 10 min for lysing the cells and denaturing the RNA. Next, 1 μl of reverse transcription mix (25 mM Tris-HCl, pH 8, 30 mM NaCl, 1 mM GTP, 2.5 mM $MgCl_2$, 8 mM DTT, 0.5 unit/ul RNAse Inhibitor, 2 μM Smart-seq3 TSO, 2 unit/ul Maxima H-minus reverse transcriptase enzyme) was added to each well for reverse transcription and template switching at 42°C for 90 min followed by 10 cycles at 50°C for 2 min and at 42°C for 2 min and the reaction was inactivated at 85°C for 5 min. cDNA pre-amplification was performed by adding 6 μl of PCR mix (1×Kapa HiFi HotStart buffer, 0.02 unit/ul KAPA HiFi DNA polymerase, 0.5 mM $MgCl_2$, 0.3 mM dNTPs, 0.5 μM Smartseq3 forward PCR primer and 0.1 μM Smartseq3 reverse PCR primer) with the following protocol: 3 min at 98°C for initial denaturation, 18–24 cycles of 20 s at 98°C, 30 s at 65°C and 6 min at 72°C, followed by a final elongation for 5 min at 72°C. PCR cycles were determined for each cell type by prior experiments and 19 cycles were used for motor neuron to obtain ~10 ng purified cDNA.

Pre-amplified cDNA was purified by 0.6x volume of AMpure XP beads, eluted in 14 ul nuclease free water, quantified using the Quant-iT PicoGreen dsDNA Assay Kit and size distributions were checked on a high-sensitivity DNA chip. cDNA was then diluted to 200 pg/μl, and 1 ul was used for tagmentation by mixing with 1 μl of tagmentation reaction mixture (10 mM TAPS, pH 8.5, 5 mM $MgCl_2$, 8% PEG8000, 0.08 ul Tn5 mix (illumina)). The reaction was performed at 55°C for 7 min, followed by heat inactivation at 80°C for 5 min. Library amplification of the tagmented samples was performed using custom-designed index primers and by adding 5 μl of PCR mix (1× Phusion High-Fidelity buffer, 0.01 unit/ul Phusion High-Fidelity DNA Polymerase, 0.2 mM dNTPs, 0.1 μM forward indexed primer and 0.1 μM reverse indexed primer). Amplification was performed as follows: 3 min 72°C; 30 s at 95°C; 12 cycles of (10 s at 95°C; 30 s at 55°C; 30 s at 72°C); and 5 min at 72°C. Samples from one 96-well plate were pooled, and then purified with 0.6x volume of Ampure XP beads. Libraries were quantified by Qubit™ dsDNA HS Assay Kit and quality checked on a high-sensitivity DNA chip. The size averaged at ~1 Kb should be expected. scRNA-seq libraries were sequenced as 50 bp paired-end reads on the Illumina NovaSeq 6000 platform and importantly, ~20% phiX spike-in was added to resolve the low complexity issue in the first 20 bps. All oligos used are listed in Table S3.

QUANTIFICATION AND STATISTICAL ANALYSIS

RNA-seq data analysis—Reads were aligned to the mouse reference genome mm10 using STAR with parameters: --outFilterMismatchNoverLmax 0.2 --outFilterMultimapNmax 1 --outSAMstrandField intronMotif --outSAMmapqUnique

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

60 --twopassMode Basic --outSJfilterReads Unique --outFilterIntronMotifs RemoveNoncanonical. Gene counts were calculated using featureCounts with parameters: -p -s 2 -t exon, and RefSeq mm10 annotation downloaded from GENCODE. The output gene count tables were used as input into DeSeq2 for normalization and differential expression analysis. For comparing the expression level of different genes within a sample, TPM (transcripts per kilobase million) is calculated as: TPM = (CDS read count * mean read length * 10^6) / (CDS length * total transcript count)

ChIP-seq data analysis—Reads were aligned to the mouse reference genome mm10 and dm6 for spike-in samples, using Bowtie2 with default parameters. Reads of quality score less than 30 were removed using samtools and PCR duplicates were removed using picard. Regions in mm10 genome blacklist was removed using bedtools and bigwig files were generated using deeptools and parameters: --binSize 50 --normalizeUsing RPKM --ignoreDuplicates --ignoreForNormalization chrX --extendReads 250 for visualization in IGV. Peaks were called using MACS2 with parameters: -g mm --keep-dup 1 --nomodel --extsize 300. Genomic peak annotation was performed with the R package ChIPseeker considering the region ± 3 kb around the TSS as the promoter. Peak overlapping analysis was performed using the Python package Intervene and visualized using the Python package Matplotlib. Motif discovery was performed using narrowPeak files generated by MACS2 and findMotifsGenome.pl function from HOMER with default parameters: mm10 -size given -mask –preparse.

For visualization of ChIP-seq, uniquely aligned reads mapping to the mouse genome were normalized using dm6 spike-in as described previously(Orlando et al., 2014). Heatmaps were performed using the functions computeMatrix followed by plotHeatmap and plotProfile from deepTools. All ChIP-Seq files presented in Figures 5C and 5H were used as input to computeMatrix and plotHeatmap for K-means clustering. Violin plots for ChIP-Seq signal were prepared by multiBigwigSummary in BED-file mode from deepTools using as bed file the regions corresponding to Ring1B cluster 3 regions. The resulting tab-delimited raw-count file was used as input to ggplot2 for Violin plot.

scRNA-seq data analysis—Raw non-demultiplexed fastq files were processed using zUMIs with STAR to generate expression profiles for both the 5′ ends containing UMIs as well as combined full-length and UMI data. The parameters: find pattern: ATTGCGCAATG, UMI (12–19), cDNA (23–50) were specified for Read 1. The UMI count table containing both intron and exon reads was used by Seurat for downstream analysis. Cells with more than 3000 genes detected and less than 5% of reads mapping to mitochondrial genome were retained. Data were normalized (scTransform) and used for principal component analysis dimensionality reduction, followed by louvain clustering and UMAP dimensionality reduction. Major cell types were readily identifiable by common marker genes: posterior neural progenitor (PNP) expressing Sox3, posterior and ventral neural progenitor (PVNP) expressing Hoxd4, newborn motor neuron (NMN) expressing Neurog2, and motor neuron (MN) expressing Mnx1 and Chat. The percentage of each cell type from each sample was used as input to ggplot2 for stack area plot.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Differential gene expression analysis between defined cell clusters was performed using R package presto. The top 500 DEGs were ordered by p value. Averaged expression among the defined cell clusters was scaled by row and used as input to R package ComplexHeatmap for visualization. Regions of AUTS2-PRC1 and NRF1 co-localization were identified using peak files generated by MACS2 and bedtools intersect function with default parameters. Co-targeted genes with its promoter located in the overlapped regions were identified by R package ChIPseeker with default parameters. The normalized expression of those genes in MN population was calculated using R package presto and used as input to ggplot2 for scatter plot.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

We thank Drs. Lynne Vales, Esteban Mazzoni, Anne Schaefer for critical reading of the manuscript as well as past and current Reinberg laboratory members for critical comments and discussions. We also thank the New York University Langone Medical Center (NYULMC) Genome Technology Center for help with sequencing, the NYULMC Cytometry and Cell Sorting Core for help with FACS, and the NYULMC Animal Facility for help with mice housing. This study utilized computing resources at the High-Performance Computing Facility of the Center for Health Informatics and Bioinformatics at the NYULMC. C.-A.M. and T.K. are supported by grants from the National Institutes of Health-National Eye Institute (EY024376, P30EY028102). The work in Reinberg lab is supported by grants to D.R. from the NIH (R01NS100897, R01CA199652) and the Simon Foundation Grant #240344 to D.R.) and the Howard Hughes Medical Institute.

REFERENCES

Ajmone PF, Avignone S, Gervasini C, Giacobbe A, Monti F, Costantino A, Esposito S, Marchisio P, Triulzi F, and Milani D (2018). Rubinstein-Taybi syndrome: New neuroradiological and neuropsychiatric insights from a multidisciplinary approach. Am J Med Genet B Neuropsychiatr Genet 177, 406–415. [PubMed: 29637745]

Aldinger KA, Timms AE, Thomson Z, Mirzaa GM, Bennett JT, Rosenberg AB, Roco CM, Hirano M, Abidi F, Haldipur P, et al. (2019). Redefining the Etiologic Landscape of Cerebellar Malformations. The American Journal of Human Genetics 105, 606–615. [PubMed: 31474318]

Aloia L, Stefano BD, and Croce LD (2013). Polycomb complexes in stem cells and embryonic development. Development 140, 2525–2534. [PubMed: 23715546]

Badea TC, Cahill H, Ecker J, Hattar S, and Nathans J (2009). Distinct Roles of Transcription Factors Brn3a and Brn3b in Controlling the Development, Morphology, and Function of Retinal Ganglion Cells. Neuron 61, 852–864. [PubMed: 19323995]

Bannister AJ, and Kouzarides T (1996). The CBP co-activator is a histone acetyltransferase. Nature 384, 641–643. [PubMed: 8967953]

Becht E, McInnes L, Healy J, Dutertre C-A, Kwok IWH, Ng LG, Ginhoux F, and Newell EW (2019). Dimensionality reduction for visualizing single-cell data using UMAP. Nature Biotechnology 37, 38–44.

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, et al. (2006). A Bivalent Chromatin Structure Marks Key Developmental Genes in Embryonic Stem Cells. Cell 125, 315–326. [PubMed: 16630819]

Beunders G, Voorhoeve E, Golzio C, Pardo LM, Rosenfeld JA, Talkowski ME, Simonic I, Lionel AC, Vergult S, Pyatt RE, et al. (2013). Exonic Deletions in AUTS2 Cause a Syndromic Form of Intellectual Disability and Suggest a Critical Role for the C Terminus. The American Journal of Human Genetics 92, 210–220. [PubMed: 23332918]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Beunders G, de Munnik SA, Van der Aa N, Ceulemans B, Voorhoeve E, Groffen AJ, Nillesen WM, Meijers-Heijboer EJ, Frank Kooy R, Yntema HG, et al. (2015). Two male adults with pathogenic AUTS2 variants, including a two-base pair deletion, further delineate the AUTS2 syndrome. European Journal of Human Genetics 23, 803–807. [PubMed: 25205402]

Beunders G, Kamp J. van de, Vasudevan P, Morton J, Smets K, Kleefstra T, Munnik S.A. de, Schuurs-Hoeijmakers J, Ceulemans B, Zollino M, et al. (2016). A detailed clinical analysis of 13 patients with AUTS2 syndrome further delineates the phenotypic spectrum and underscores the behavioural phenotype. Journal of Medical Genetics 53, 523–532. [PubMed: 27075013]

Bonasio R, Tu S, and Reinberg D (2010). Molecular Signals of Epigenetic States. Science 330, 612–616. [PubMed: 21030644]

Briggs JA, Li VC, Lee S, Woolf CJ, Klein A, and Kirschner MW (2017). Mouse embryonic stem cells can differentiate via multiple paths to the same state. ELife 6, e26945. [PubMed: 28990928]

Butler A, Hoffman P, Smibert P, Papalexi E, and Satija R (2018). Integrating single-cell transcriptomic data across different conditions, technologies, and species. Nature Biotechnology 36, 411–420.

Cantani A, and Gagliesi D (1998). Rubinstein-Taybi syndrome. Review of 732 cases and analysis of the typical traits. Eur Rev Med Pharmacol Sci 2, 81–87. [PubMed: 10229563]

Chang W-T, Chen H, Chiou R-J, Chen C-Y, and Huang A-M (2005). A novel function of transcription factor α-Pal/NRF-1: Increasing neurite outgrowth. Biochemical and Biophysical Research Communications 334, 199–206. [PubMed: 15992771]

Cohen I, Zhao D, Bar C, Valdes VJ, Dauber-Decker KL, Nguyen MB, Nakayama M, Rendl M, Bickmore WA, Koseki H, et al. (2018). PRC1 Fine-tunes Gene Repression and Activation to Safeguard Skin Development and Stem Cell Specification. Cell Stem Cell 22, 726–739.e7. [PubMed: 29727681]

Di Croce L, and Helin K (2013). Transcriptional regulation by Polycomb group proteins. Nat Struct Mol Biol 20, 1147–1155. [PubMed: 24096405]

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]

Edwards TJ, Sherr EH, Barkovich AJ, and Richards LJ (2014). Clinical, genetic and imaging findings identify new causes for corpus callosum development syndromes. Brain 137, 1579–1613. [PubMed: 24477430]

Elia J, Gai X, Xie HM, Perin JC, Geiger E, Glessner JT, D'arcy M, deBerardinis R, Frackelton E, Kim C, et al. (2010). Rare structural variants found in attention-deficit hyperactivity disorder are preferentially associated with neurodevelopmental genes. Molecular Psychiatry 15, 637–646. [PubMed: 19546859]

Endoh M, Endo TA, Endoh T, Fujimura Y, Ohara O, Toyoda T, Otte AP, Okano M, Brockdorff N, Vidal M, et al. (2008). Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development 135, 1513–1524. [PubMed: 18339675]

Escobar TM, Oksuz O, Saldaña-Meyer R, Descostes N, Bonasio R, and Reinberg D (2019). Active and Repressed Chromatin Domains Exhibit Distinct Nucleosome Segregation during DNA Replication. Cell 179, 953–963.e11. [PubMed: 31675501]

Escobar TM, Loyola A, and Reinberg D (2021). Parental nucleosome segregation and the inheritance of cellular identity. Nature Reviews Genetics 1–14.

Evans MJ, and Scarpulla RC (1990). NRF-1: a trans-activator of nuclear-encoded respiratory genes in animal cells. Genes Dev 4, 1023–1034. [PubMed: 2166701]

Francis NJ, Kingston RE, and Woodcock CL (2004). Chromatin Compaction by a Polycomb Group Protein Complex. Science 306, 1574–1577. [PubMed: 15567868]

Gao Z, Zhang J, Bonasio R, Strino F, Sawai A, Parisi F, Kluger Y, and Reinberg D (2012). PCGF Homologs, CBX Proteins, and RYBP Define Functionally Distinct PRC1 Family Complexes. Molecular Cell 45, 344–356. [PubMed: 22325352]

Gao Z, Lee P, Stafford JM, von Schimmelmann M, Schaefer A, and Reinberg D (2014). An AUTS2–Polycomb complex activates gene expression in the CNS. Nature 516, 349–354. [PubMed: 25519132]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Geng Z, Wang Q, Miao W, Wolf T, Chavez J, Giddings E, Hobbs R, DeGraff DJ, Wang Y, Stafford J, et al. (2021). AUTS2 controls neuronal lineage choice through a novel PRC1-independent complex and BMP inhibition. BioRxiv 2021.06.29.450402.

Gleyzer N, Vercauteren K, and Scarpulla RC (2005). Control of Mitochondrial Transcription Specificity Factors (TFB1M and TFB2M) by Nuclear Respiratory Factors (NRF-1 and NRF-2) and PGC-1 Family Coactivators. Mol Cell Biol 25, 1354–1366. [PubMed: 15684387]

Gu Z, Eils R, and Schlesner M (2016). Complex heatmaps reveal patterns and correlations in multidimensional genomic data. Bioinformatics 32, 2847–2849. [PubMed: 27207943]

Guerreiro R, Brás J, and Hardy J (2015). SnapShot: Genetics of ALS and FTD. Cell 160, 798–798.e1. [PubMed: 25679767]

Gugneja S, and Scarpulla RC (1997). Serine Phosphorylation within a Concise Amino-terminal Domain in Nuclear Respiratory Factor 1 Enhances DNA Binding*. Journal of Biological Chemistry 272, 18732–18739.

Hagemann-Jensen M (2020). Smart-seq3 Protocol.

Hagemann-Jensen M, Ziegenhain C, Chen P, Ramsköld D, Hendriks G-J, Larsson AJM, Faridani OR, and Sandberg R (2020). Single-cell RNA counting at allele and isoform resolution using Smart-seq3. Nature Biotechnology 38, 708–714.

Hattori E, Toyota T, Ishitsuka Y, Iwayama Y, Yamada K, Ujike H, Morita Y, Kodama M, Nakata K, Minabe Y, et al. (2009). Preliminary genome-wide association study of bipolar disorder in the Japanese population. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics 150B, 1110–1117.

Hauri S, Comoglio F, Seimiya M, Gerstung M, Glatter T, Hansen K, Aebersold R, Paro R, Gstaiger M, and Beisel C (2016). A High-Density Map for Navigating the Human Polycomb Complexome. Cell Reports 17, 583–595. [PubMed: 27705803]

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. Molecular Cell 38, 576–589. [PubMed: 20513432]

Hori K, and Hoshino M (2017). Neuronal Migration and AUTS2 Syndrome. Brain Sciences 7, 54.

Hori K, Nagai T, Shan W, Sakamoto A, Taya S, Hashimoto R, Hayashi T, Abe M, Yamazaki M, Nakao K, et al. (2014). Cytoskeletal Regulation by AUTS2 in Neuronal Migration and Neuritogenesis. Cell Reports 9, 2166–2179. [PubMed: 25533347]

Hsiao H-Y, Jukam D, Johnston R, and Desplan C (2013). The neuronal transcription factor erect wing regulates specification and maintenance of Drosophila R8 photoreceptor subtypes. Developmental Biology 381, 482–490. [PubMed: 23850772]

Huo L, and Scarpulla RC (2001). Mitochondrial DNA Instability and Peri-Implantation Lethality Associated with Targeted Disruption of Nuclear Respiratory Factor 1 in Mice. Molecular and Cellular Biology 21, 644–654. [PubMed: 11134350]

Jordan VK, Fregeau B, Ge X, Giordano J, Wapner RJ, Balci TB, Carter MT, Bernat JA, Moccia AN, Srivastava A, et al. (2018). Genotype–phenotype correlations in individuals with pathogenic RERE variants. Human Mutation 39, 666–675. [PubMed: 29330883]

Kapoor M, Wang J-C, Wetherill L, Le N, Bertelsen S, Hinrichs AL, Budde J, Agrawal A, Bucholz K, Dick D, et al. (2013). A meta-analysis of two genome-wide association studies to identify novel loci for maximum number of alcoholic drinks. Hum Genet 132, 1141–1151. [PubMed: 23743675]

Kasinath V, Beck C, Sauer P, Poepsel S, Kosmatka J, Faini M, Toso D, Aebersold R, and Nogales E (2021). JARID2 and AEBP2 regulate PRC2 in the presence of H2AK119ub1 and other histone modifications. Science 371.

Kim J, and Kingston RE (2020). The CBX family of proteins in transcriptional repression and memory. J Biosci 45, 16. [PubMed: 31965994]

Kiyama T, Chen C-K, Wang SW, Pan P, Ju Z, Wang J, Takada S, Klein WH, and Mao C-A (2018). Essential roles of mitochondrial biogenesis regulator Nrf1 in retinal development and homeostasis. Molecular Neurodegeneration 13, 56. [PubMed: 30333037]

Kiyama T, Long Y, Chen C-K, Whitaker CM, Shay A, Wu H, Badea TC, Mohsenin A, Parker-Thornburg J, Klein WH, et al. (2019). Essential Roles of Tbr1 in the Formation and Maintenance

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

of the Orientation-Selective J-RGCs and a Group of OFF-Sustained RGCs in Mouse. Cell Reports 27, 900–915.e5. [PubMed: 30995485]

Kloet SL, Makowski MM, Baymaz HI, van Voorthuijsen L, Karemaker ID, Santanach A, Jansen PWTC, Di Croce L, and Vermeulen M (2016). The dynamic interactome and genomic targets of Polycomb complexes during stem cell differentiation. Nat Struct Mol Biol 23, 682–690. [PubMed: 27294783]

Korotkevich G, Sukhov V, and Sergushichev A (2019). Fast gene set enrichment analysis. BioRxiv 060012.

Korsunsky I, Nathan A, Millard N, and Raychaudhuri S (2019). Presto scales Wilcoxon and auROC analyses to millions of observations. BioRxiv 653253.

Langmead B, and Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nat Methods 9, 357–359. [PubMed: 22388286]

Lau MS, Schwartz MG, Kundu S, Savol AJ, Wang PI, Marr SK, Grau DJ, Schorderet P, Sadreyev RI, Tabin CJ, et al. (2017). Mutation of a nucleosome compaction region disrupts Polycomb-mediated axial patterning. Science 355, 1081–1084. [PubMed: 28280206]

LeRoy G, Oksuz O, Descostes N, Aoi Y, Ganai RA, Kara HO, Yu J-R, Lee C-H, Stafford J, Shilatifard A, et al. (2019). LEDGF and HDGF2 relieve the nucleosome-induced barrier to transcription in differentiated cells. Science Advances 5, eaay3068. [PubMed: 31616795]

Li G, Margueron R, Ku M, Chambon P, Bernstein BE, and Reinberg D (2010). Jarid2 and PRC2, partners in regulating gene expression. Genes Dev 24, 368–380. [PubMed: 20123894]

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, and Subgroup, 1000 Genome Project Data Processing (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078–2079. [PubMed: 19505943]

Liao Y, Smyth GK, and Shi W (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30, 923–930. [PubMed: 24227677]

Lin Z, Yang Z, Xie R, Ji Z, Guan K, and Zhang M (2019). Decoding WW domain tandem-mediated target recognitions in tissue growth and cell polarity. ELife 8, e49439. [PubMed: 31486770]

Loubiere V, Papadopoulos GL, Szabo Q, Martinez A-M, and Cavalli G (2020). Widespread activation of developmental gene expression characterized by PRC1-dependent chromatin looping. Science Advances 6, eaax4001. [PubMed: 31950077]

Margueron R, and Reinberg D (2011). The Polycomb complex PRC2 and its mark in life. Nature 469, 343–349. [PubMed: 21248841]

Martinez-Delgado B, Lopez-Martin E, Lara-Herguedas J, Monzon S, Cuesta I, Juliá M, Aquino V, Rodriguez-Martin C, Damian A, Gonzalo I, et al. (2020). De novo small deletion affecting transcription start site of short isoform of AUTS2 gene in a patient with syndromic neurodevelopmental defects. Am J Med Genet A.

Mazzoni EO, Mahony S, Peljto M, Patel T, Thornton SR, McCuine S, Reeder C, Boyer LA, Young RA, Gifford DK, et al. (2013). Saltatory remodeling of Hox chromatin in response to rostrocaudal patterning signals. Nature Neuroscience 16, 1191–1198. [PubMed: 23955559]

Mefford HC, Muhle H, Ostertag P, Spiczak S. von, Buysse K, Baker C, Franke A, Malafosse A, Genton P, Thomas P, et al. (2010). Genome-Wide Copy Number Variation in Epilepsy: Novel Susceptibility Loci in Idiopathic Generalized and Focal Epilepsies. PLOS Genetics 6, e1000962. [PubMed: 20502679]

Min J, Zhang Y, and Xu R-M (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev. 17, 1823–1828. [PubMed: 12897052]

Narendra V, Rocha PP, An D, Raviram R, Skok JA, Mazzoni EO, and Reinberg D (2015). CTCF establishes discrete functional chromatin domains at the Hox clusters during differentiation. Science 347, 1017–1021. [PubMed: 25722416]

Ogryzko VV, Schiltz RL, Russanova V, Howard BH, and Nakatani Y (1996). The Transcriptional Coactivators p300 and CBP Are Histone Acetyltransferases. Cell 87, 953–959. [PubMed: 8945521]

Oksenberg N, and Ahituv N (2013). The role of AUTS2 in neurodevelopment and human evolution. Trends in Genetics 29, 600–608. [PubMed: 24008202]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Oksenberg N, Haliburton GDE, Eckalbar WL, Oren I, Nishizaki S, Murphy K, Pollard KS, Birnbaum RY, and Ahituv N (2014). Genome-wide distribution of Auts2 binding localizes with active neurodevelopmental genes. Translational Psychiatry 4, e431–e431. [PubMed: 25180570]

Oksuz O, Narendra V, Lee C-H, Descostes N, LeRoy G, Raviram R, Blumenberg L, Karch K, Rocha PP, Garcia BA, et al. (2018). Capturing the Onset of PRC2-Mediated Repressive Domain Formation. Molecular Cell 70, 1149–1162.e5. [PubMed: 29932905]

Orlando DA, Chen MW, Brown VE, Solanki S, Choi YJ, Olson ER, Fritz CC, Bradner JE, and Guenther MG (2014). Quantitative ChIP-Seq Normalization Reveals Global Modulation of the Epigenome. Cell Reports 9, 1163–1170. [PubMed: 25437568]

Palmer EE, Hong S, Al Zahrani F, Hashem MO, Aleisa FA, Ahmed HMJ, Kandula T, Macintosh R, Minoche AE, Puttick C, et al. (2019). De Novo Variants Disrupting the HX Repeat Motif of ATN1 Cause a Recognizable Non-Progressive Neurocognitive Syndrome. The American Journal of Human Genetics 104, 542–552. [PubMed: 30827498]

Parekh S, Ziegenhain C, Vieth B, Enard W, and Hellmann I (2018). zUMIs - A fast and flexible pipeline to process RNA sequencing data with UMIs. Gigascience 7.

Pasini D, Cloos PAC, Walfridsson J, Olsson L, Bukowski J-P, Johansen JV, Bak M, Tommerup N, Rappsilber J, and Helin K (2010). JARID2 regulates binding of the Polycomb repressive complex 2 to target genes in ES cells. Nature 464, 306–310. [PubMed: 20075857]

Quinlan AR, and Hall IM (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841–842. [PubMed: 20110278]

Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, and Manke T (2016). deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44, W160–W165. [PubMed: 27079975]

Reyes NSDL, Bragg-Gonzalo L, and Nieto M (2020). Development and plasticity of the corpus callosum. Development 147.

RUBINSTEIN JH, and TAYBI H (1963). Broad thumbs and toes and facial abnormalities. A possible mental retardation syndrome. Am J Dis Child 105, 588–608. [PubMed: 13983033]

Scarpulla RC (2011). Metabolic control of mitochondrial biogenesis through the PGC-1 family regulatory network. Biochim Biophys Acta 1813, 1269–1278. [PubMed: 20933024]

Scelfo A, Fernández-Pérez D, Tamburri S, Zanotti M, Lavarone E, Soldi M, Bonaldi T, Ferrari KJ, and Pasini D (2019). Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent- and RING1A/B-Independent-Specific Activities. Molecular Cell 74, 1037–1052.e7. [PubMed: 31029542]

Schuettengruber B, Bourbon H-M, Di Croce L, and Cavalli G (2017). Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 171, 34–57. [PubMed: 28938122]

Sobreira N, Schiettecatte F, Valle D, and Hamosh A (2015). GeneMatcher: A Matching Tool for Connecting Investigators with an Interest in the Same Gene. Human Mutation 36, 928–930. [PubMed: 26220891]

Stevens CA (1993). Rubinstein-Taybi Syndrome. In GeneReviews®, Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJ, Stephens K, and Amemiya A, eds. (Seattle (WA): University of Washington, Seattle), p.

Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, and Pombo A (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol 9, 1428–1435. [PubMed: 18037880]

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 approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences 102, 15545–15550.

Sultana R, Yu C-E, Yu J, Munson J, Chen D, Hua W, Estes A, Cortes F, de la Barra F, Yu D, et al. (2002). Identification of a Novel Gene on Chromosome 7q11.2 Interrupted by a Translocation Breakpoint in a Pair of Autistic Twins. Genomics 80, 129–134. [PubMed: 12160723]

Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, Bezstarosti K, Taylor S, Ura H, Koide H, et al. (2012). RYBP-PRC1 Complexes Mediate H2A Ubiquitylation at Polycomb Target Sites Independently of PRC2 and H3K27me3. Cell 148, 664–678. [PubMed: 22325148]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Tong C-W, Wang J-L, Jiang M-S, Hsu C-H, Chang W-T, and Huang A-M (2013). Novel genes that mediate nuclear respiratory factor 1-regualted neurite outgrowth in neuroblastoma IMR-32 cells. Gene 515, 62–70. [PubMed: 23219993]

Trojer P, and Reinberg D (2007). Facultative Heterochromatin: Is There a Distinctive Molecular Signature? Molecular Cell 28, 1–13. [PubMed: 17936700]

Ufartes R, Berger H, Till K, Salinas G, Sturm M, Altmüller J, Nürnberg P, Thiele H, Funke R, Apeshiotis N, et al. (2020). De novo mutations in FBRSL1 cause a novel recognizable malformation and intellectual disability syndrome. Hum Genet 139, 1363–1379. [PubMed: 32424618]

Voigt P, Tee W-W, and Reinberg D (2013). A double take on bivalent promoters. Genes Dev. 27, 1318–1338. [PubMed: 23788621]

Wichterle H, Lieberam I, Porter JA, and Jessell TM (2002). Directed Differentiation of Embryonic Stem Cells into Motor Neurons. Cell 110, 385–397. [PubMed: 12176325]

Wiley S, Swayne S, Rubinstein JH, Lanphear NE, and Stevens CA (2003). Rubinstein-Taybi syndrome medical guidelines. American Journal of Medical Genetics Part A 119A, 101–110. [PubMed: 12749047]

Yagi R, Chen L-F, Shigesada K, Murakami Y, and Ito Y (1999). A WW domain-containing Yes-associated protein (YAP) is a novel transcriptional co-activator. The EMBO Journal 18, 2551–2562. [PubMed: 10228168]

Yu G, Wang L-G, and He Q-Y (2015). ChIPseeker: an R/Bioconductor package for ChIP peak annotation, comparison and visualization. Bioinformatics 31, 2382–2383. [PubMed: 25765347]

Yu J-R, Lee C-H, Oksuz O, Stafford JM, and Reinberg D (2019). PRC2 is high maintenance. Genes Dev. 33, 903–935. [PubMed: 31123062]

Zha X, Costa V, Harding AMS, Reznikov L, Benson CJ, and Welsh MJ (2009). ASIC2 Subunits Target Acid-Sensing Ion Channels to the Synapse via an Association with PSD-95. J. Neurosci 29, 8438–8446. [PubMed: 19571134]

Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, et al. (2008). Model-based Analysis of ChIP-Seq (MACS). Genome Biology 9, R137. [PubMed: 18798982]

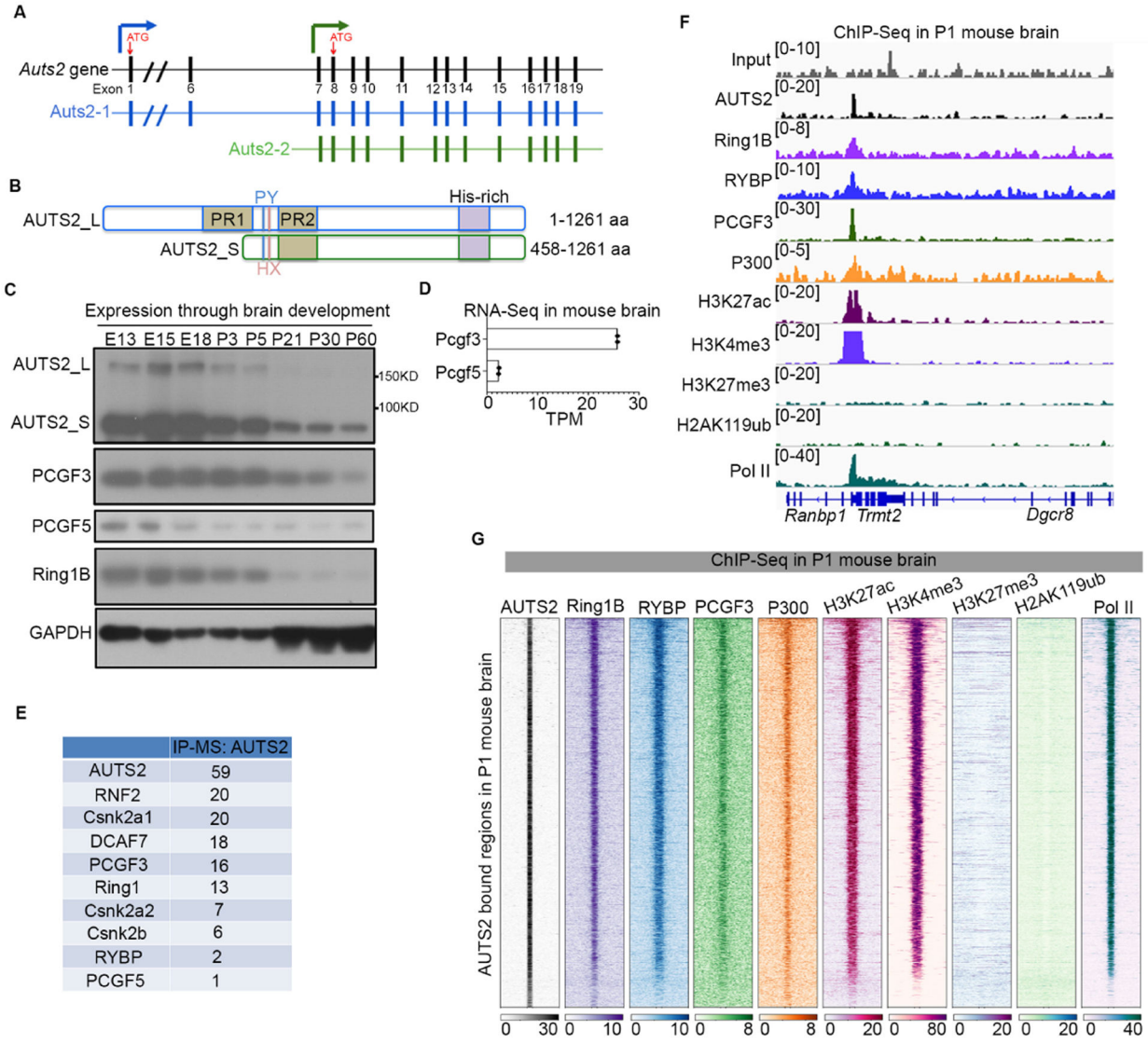
Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK, and Rosenfeld MG (2008). Histone H2A Monoubiquitination Represses Transcription by Inhibiting RNA Polymerase II Transcriptional Elongation. Molecular Cell 29, 69–80. [PubMed: 18206970]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Highlights

- The AUTS2 HX repeat recruits P300 for ncPRC1.3-mediated transcription activation
- Mutations in the HX repeat phenocopy *CREBBP/EP300* mutations in RSTS
- NRF1 recruits ncPRC1.3 that activates a subset of genes fostering neuronal development
- Recruitment of P300 and ncPRC1.3 are required for neuronal progenitor differentiation

Author Manuscript Author Manuscript Author Manuscript Author Manuscript



E

	IP-MS: AUTS2
AUTS2	59
RNF2	20
Csnk2a1	20
DCAF7	18
PCGF3	16
Ring1	13
Csnk2a2	7
Csnk2b	6
RYBP	2
PCGF5	1

Figure 1. AUTS2-ncPRC1.3 targets active genes in mouse brain

(A) Schematic showing the mouse *Auts2* gene structure and its two major transcripts in the mouse brain. Red arrows indicate the translational start codons used for each AUTS2 isoform.

(B) Schematic showing the domains of the long and short isoforms of mouse AUTS2 protein. PR, proline-rich region; PY, PPPY motif; HX, HX repeat motif comprising alternating HQ (x6) or HT (x3) residues; His-rich, eight histidine repeats.

(C) Expression of AUTS2 and core ncPRC1.3/1.5 components (PCGF3 and PCGF5, respectively, and RING1B) in the mouse brain. Immunoblotting was performed with whole brain extracts at various developmental stages, as indicated.

(D) Bar graphs showing the value of transcripts per kilobase million (TPM) for Pcgf3 and Pcgf5 revealed by RNA-Seq from whole brain lysate at postnatal day 1 (P1).

(E) Proteomic mass spectrometry results of immunoprecipitation (IP) using AUTS2 antibody in whole brain lysate at P1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(F) IGV browser views showing ChIP-seq for input, AUTS2, RING1B, RYBP, PCGF3, P300, H3K27ac, H3K4me3, H3K27me3, H2AK119ub1, and RNA Polymerase II (PolII) at the representative loci. ChIP-seq was performed in whole brain lysate at P1.
(G) Heatmap showing AUTS2, RING1B, RYBP, PCGF3, P300, H3K27ac, H3K4me3, H3K27me3, H2AK119ub1, and PolII ChIP-seq signals centered on AUTS2 bound regions (±5 kb). ChIP-seq was performed in whole brain lysate at P1.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

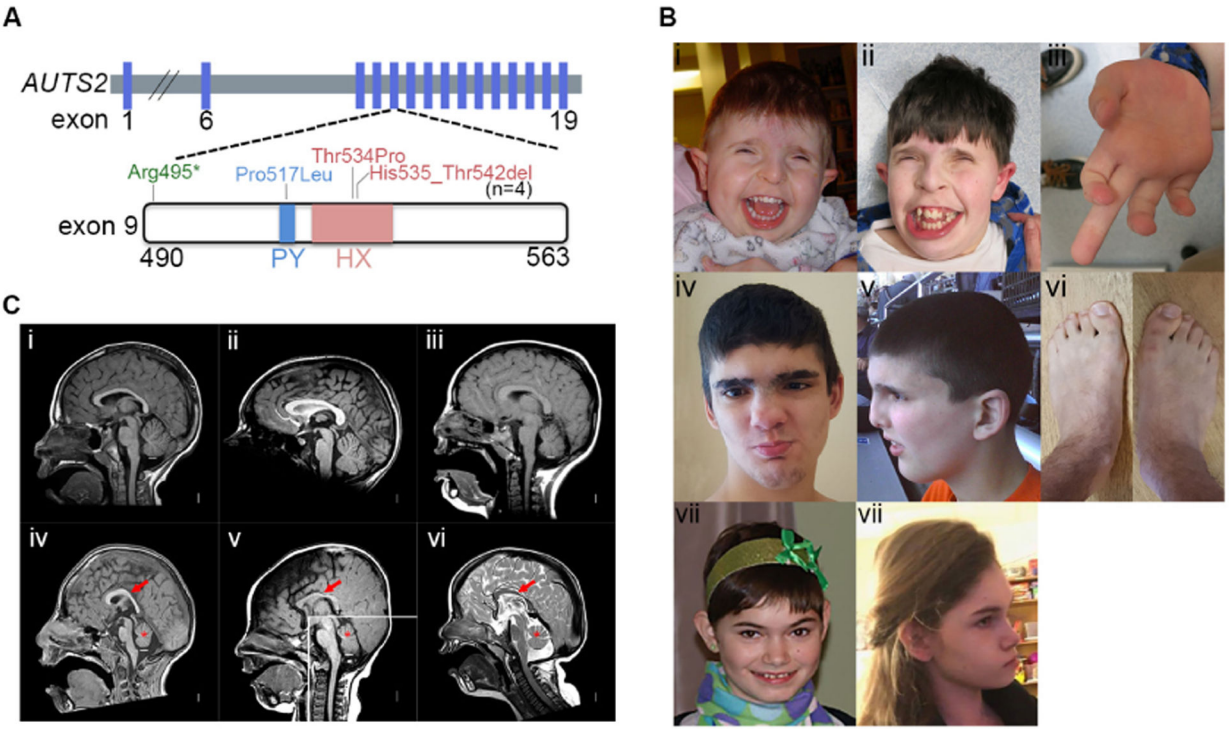


Figure 2. Patients with mutations in the AUTS2 HX repeat have features overlapping those of Rubinstein-Taybi syndrome.

(A) Schematic illustrating mutations in the *AUTS2* gene from individual patients as identified through trio-based exome sequencing. Mutations resulting in similar clinical features are labeled with same color. PY, PPPY motif; HX, HX repeat motif comprising alternating HQ (x6) or HT (x3) residues.

(B) Clinical features of individuals with mutations in AUTS2 PY motif and HX repeat domain. Frontal photos of our original proband (LR05–007 with p.Thr534Pro mutation) at ~1 (**i**) and 13 (**ii**) years show reflexive eye closure with smiling, as well as a low hanging columella, features often seen in RSTS (**i-ii**). Photo of his right hand shows severe symphalangism of the right third finger (**iii**). Frontal (**iv**) and profile (**v**) photos of a boy (LR15–003 with p.His535_Thr542 del) at 17 years show thick horizontal eyebrows with synophrys, prominent (high) nasal bridge, broad nose with mildly low columella, and posteriorly-rotated left ear. Photos of his feet show mildly broad halluces (**vi**). Frontal (**vii**) and profile (**viii**) photos of a girl (LR15–004 with p.Pro517Leu) show a normal facial appearance.

(C) Magnetic resonance images from 5 individuals with missense variants in AUTS2 exon 9 and a normal control. The three midline sagittal images in the top row from a normal control (**i**), a girl with a missense variant (p. Pro517Leu) (**ii**), and a boy with the recurrent INDEL (p.His535_Thr542 del) (**iii**) respectively, all show normal midline structures. The three midline sagittal images in the bottom row come from a boy with a missense mutation (p.Thr534Pro) (**iv**), and two boys with the recurrent INDEL (p.His535_Thr542 del) (**v-vi**). All three show thin and dysplastic corpus callosum and small cerebellar vermis (arrows and asterisks, respectively, in **iv-vi**). The horizontal white or black lines mark the level of the obex, the usual lower extent of the vermis.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

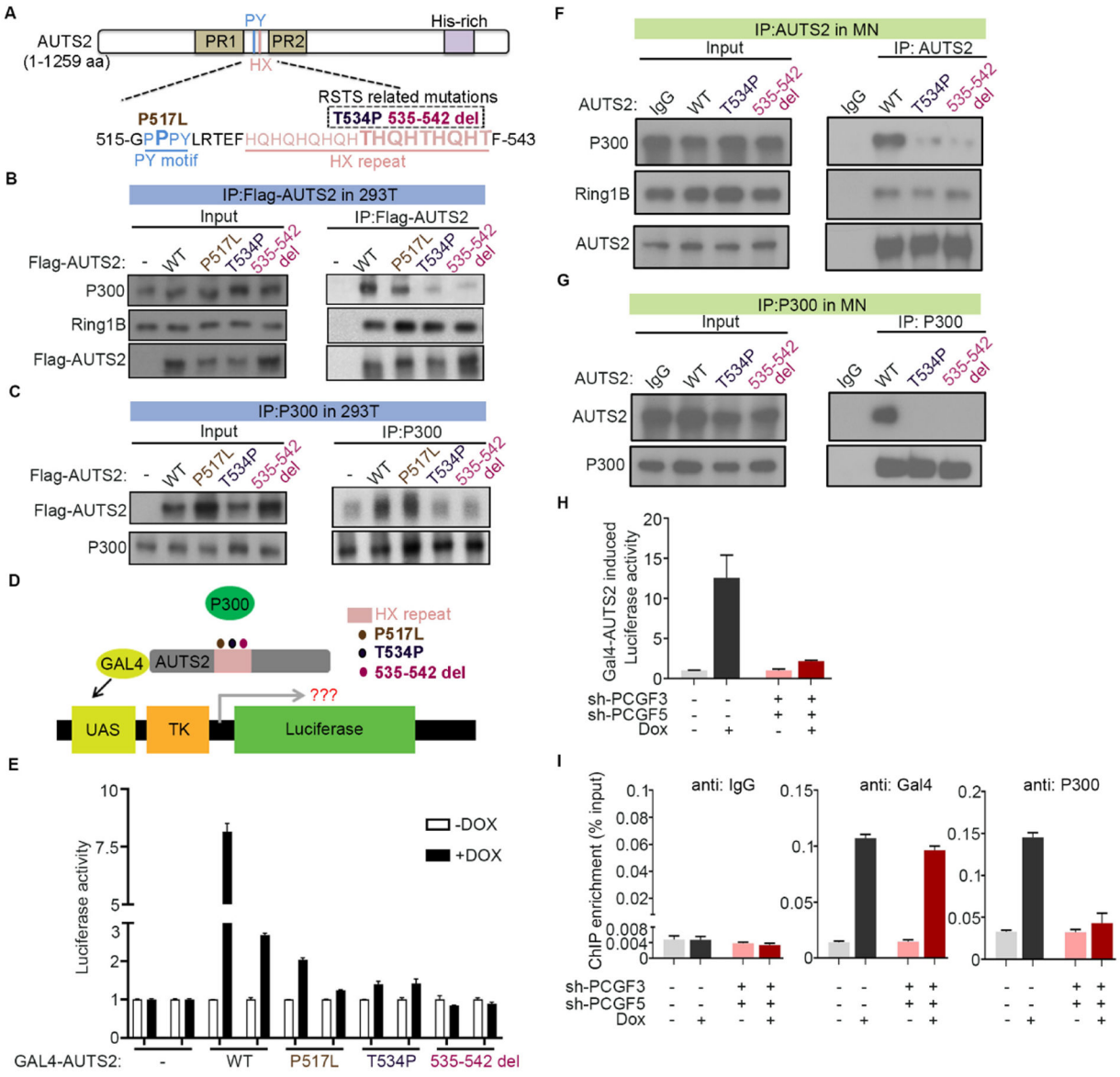


Figure 3. The HX repeat domain and ncPRC1.3/1.5 core components are required for efficient P300 recruitment and transcriptional activation.

(A) Schematic showing human AUTS2 variants constructed and expressed in 293 T-REx cells.

(B-C) Western blots show co-IP results from nuclear extract of 293 T-REx cells expressing Flag–AUTS2, either WT or mutant versions as indicated, using Flag antibody (B) or reciprocal IP using P300 antibody (C).

(D) Schematic of the reporter construct for the luciferase assay in the context of GAL4-AUTS2, either WT or mutant versions as indicated.

(E) Luciferase activity in cells expressing GAL4-AUTS2, either WT or mutant versions, before and after doxycycline treatment.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(F-G) Western blots show co-IP results from nuclear extract of MN differentiated from WT and *Auts2* HX mutant (T534P and 535–542 aa deletion, respectively) mESC as indicated, using AUTS2 antibody (F) or reciprocal IP using P300 antibody (G).
(H) Luciferase activity in cells expressing GAL4-AUTS2, either in the presence or absence of PCGF3 and PCGF5, before and after doxycycline treatment.
(I) ChIP-qPCR at the UAS element using the antibodies as indicated in cells expressing GAL4-AUTS2, either in the presence or absence of PCGF3 and PCGF5, before and after doxycycline treatment.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

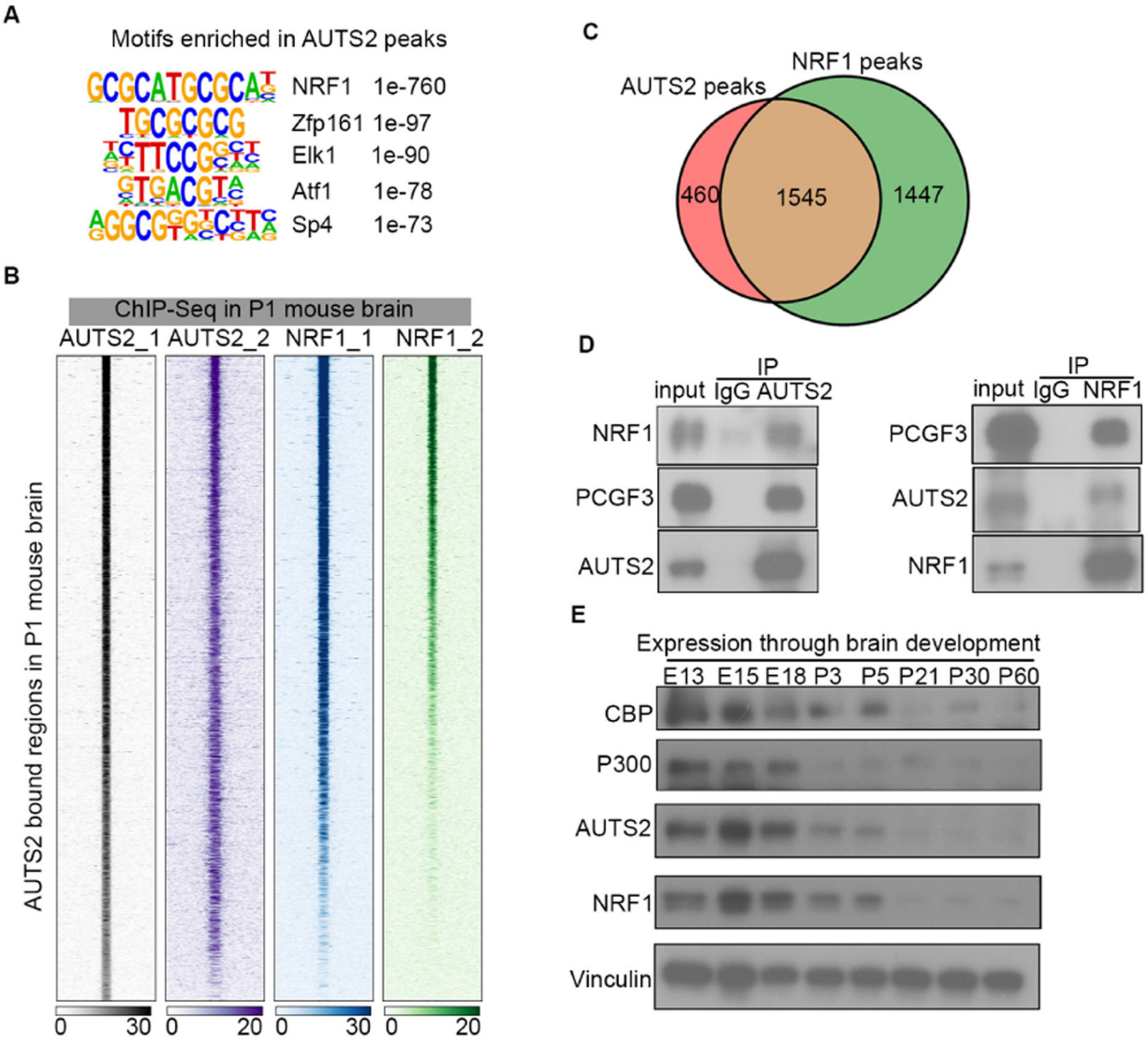


Figure 4. Transcription factor NRF1 colocalizes with AUTS2 on chromatin and physically interacts with AUTS2 in mouse brain

(A) Top 5 enriched motifs identified in AUTS2-bound regions in mouse brain using HOMER.

(B) Heatmap showing AUTS2 and NRF1 ChIP-seq signals centered on AUTS2-bound regions (±5 kb) with two replicates. ChIP-seq was performed in whole brain lysate at P1.

(C) Venn diagram showing the extent of overlap for AUTS2- and NRF1-bound regions revealed by ChIP-seq from (B).

(D) Reciprocal co-IP and western blot analyses demonstrating AUTS2 and NRF1 interaction in whole brain lysate at P1.

(E) Expression of CBP, P300, AUTS2, and NRF1 in mouse brain. Immunoblotting was performed with whole brain extracts at various developmental stages, as indicated.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

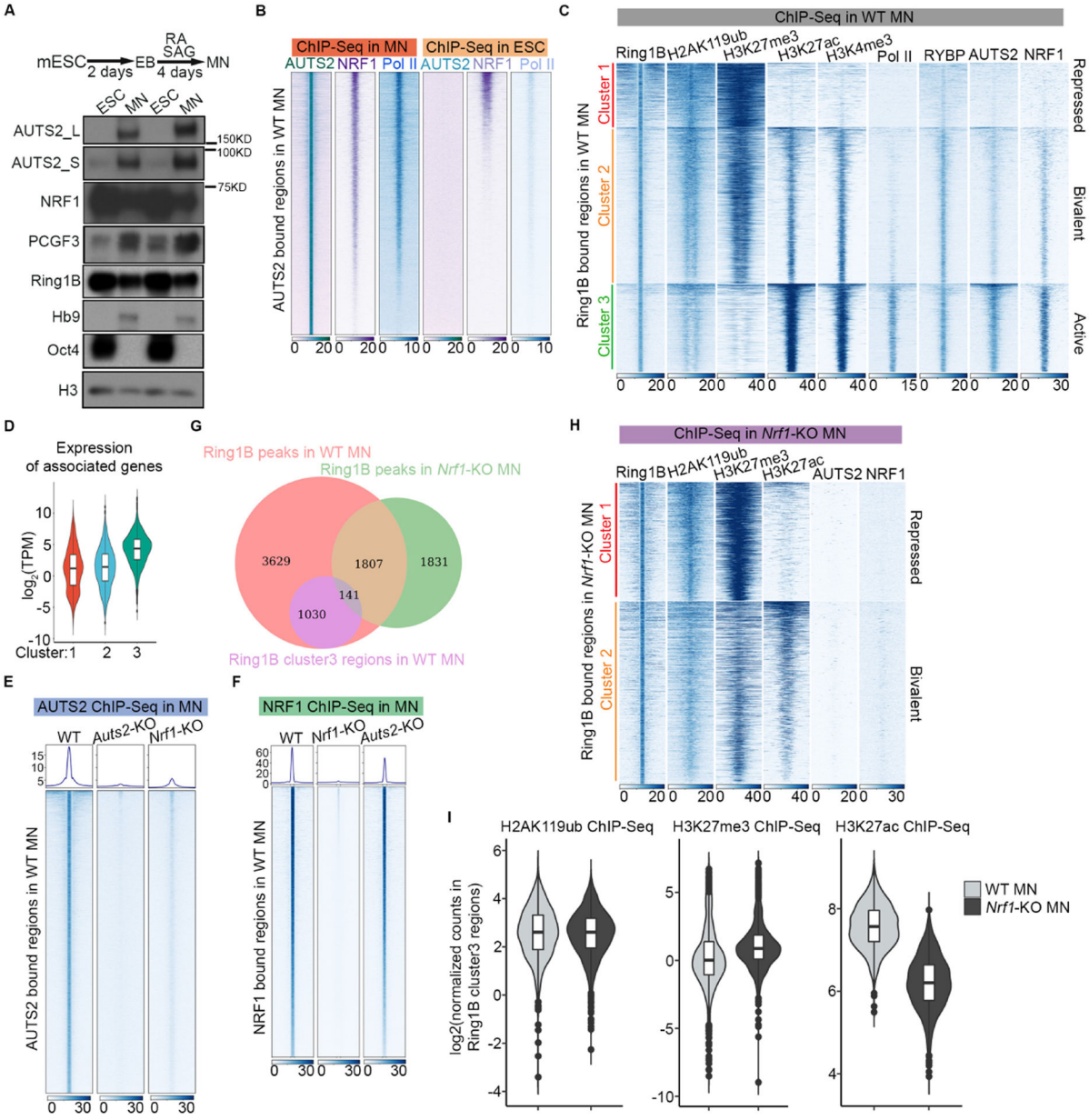


Figure 5. NRF1 is crucial for AUTS2-ncPRC1.3 associated active transcription in MN

(A) The schematic at top depicts the protocol for differentiation of mouse embryonic stem cells (mESC) to motor neurons (MN) using retinoic acid (RA) and smoothened agonist (SAG). EB, embryoid bodies. Below is a western blot showing the expression of AUTS2, NRF1, PCGF3, and RING1B in mESC and MN.

(B) Heatmap showing AUTS2, NRF1 and RNA Pol II ChIP-seq signals centered on AUTS2-bound regions identified in WT MN (±5 kb).

(C) k-means clustering of RING1B, H2AK119ub1, H3K27me3, H3K27ac, H3K4me3, RNA Pol II, RYBP, AUTS2, and NRF1 ChIP-seq signals from WT MN centered on RING1B-bound regions identified in WT MN (±5 kb).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(D) Violin plot of the $\log_2$(TPM) of genes assigned to each cluster [as indicated in (C)], quantified from RNA-seq in WT MN.

(E) Average density profiles (top) and heatmap (bottom) showing AUTS2 ChIP-seq signals from WT, *Auts2*-KO, and *Nrf1*-KO MN centered on AUTS2-bound regions identified in WT MN (±5 kb).

(F) Average density profiles (top) and heatmap (bottom) showing NRF1 ChIP-seq signals from WT, *Nrf1*-KO, and *Auts2*-KO MN centered on NRF1-bound regions identified in WT MN (±5 kb).

(G) Venn diagram showing the extent of overlap for RING1B-bound regions in WT MN, RING1B cluster 3 regions [as indicated in (C)], and RING1B-bound regions in *Nrf1*-KO MN revealed by ChIP-seq.

(H) k-means clustering of RING1B, H2AK119ub1, H3K27me3, H3K27ac, AUTS2, and NRF1 ChIP-seq signals from *Nrf1*-KO MN centered on RING1B-bound regions identified in *Nrf1*-KO MN (±5 kb).

(I) Violin plot of the $\log_2$(normalized counts), quantified from H2AK119ub, H3K27me3, and H3K27ac ChIP-Seq signal at RING1B cluster 3 regions [as indicated in (C)] in WT and *Nrf1*-KO MN.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

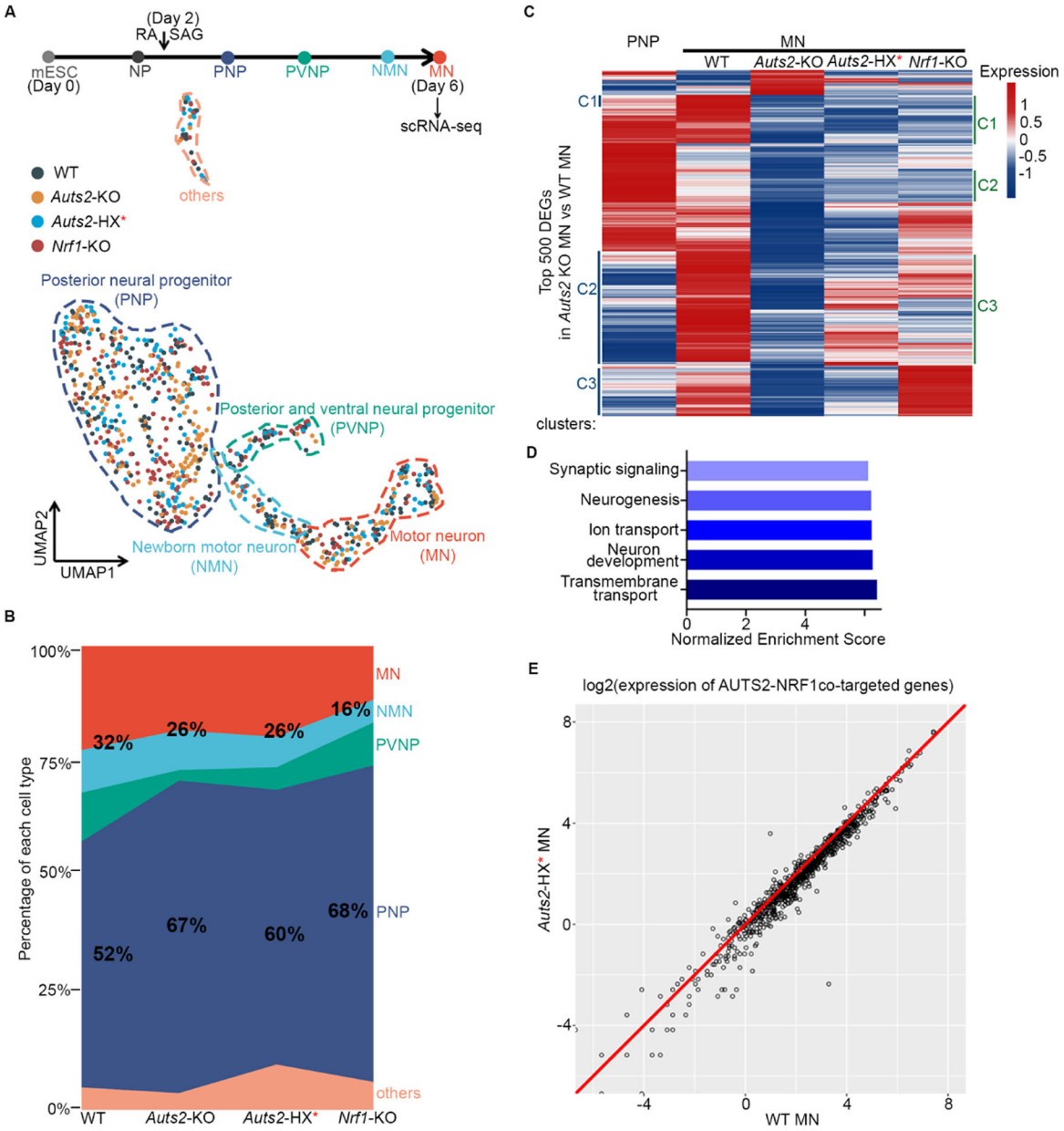


Figure 6. Defect in PNP to MN differentiation under *Auts2*-KO, mutation in *Auts2* HX repeat or *Nrf1*-KO as revealed by scRNA-Seq

(A) The schematic at top depicts cell lineage transitions from mESC to MN. MN (day 6) differentiated from WT, *Auts2* knockout, *Auts2* HX mutant (535–542 aa deletion) and *Nrf1* knockout ESC were harvested for scRNA-Seq. Below is the dimensionality reduction (UMAP) of 816 cells from WT, *Auts2*-KO, *Auts2* HX mutant, and *Nrf1*-KO samples, sequenced with the Smart-seq3 technique and colored by sample identity (WT, 228 cells; *Auts2*-KO, 221 cells; *Auts2*-HX*, 184 cells; *Nrf1*-KO, 183 cells). Five cell classes revealed by unsupervised clustering of cellular transcriptomics are represented by color-coded circles of dashed-lines and annotated based on marker gene expression. Colors of dashed-lines match those for the cell types shown at top.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

(B) Proportional stacked area graph showing the abundances of each cell type in MN differentiated from WT, *Auts2*-KO, *Auts2* HX mutant (535–542 aa deletion), and *Nrf1*-KO ESC. The percentage of MN (MN+NMN) and PNP in each sample are labelled.
(C) Top 500 DEGs were identified by comparing *Auts2*-KO MN versus WT MN. Expression of these DEGs across WT PNP, WT MN, *Auts2*-KO MN, *Auts2* HX mutant MN, and *Nrf1*-KO MN is shown by heatmap. Color scale represents the averaged and scaled expression values from each cell population. Labelling of clusters (C1, C2, C3) on left is based on gene expression differences in WT PNP, WT MN, and *Auts2*-KO MN. Labelling of clusters (C1, C2, C3) on right is based on gene expression differences in WT MN, *Auts2*-KO MN, and *Nrf1*-KO MN.
(D) Bar plot summarizing results of GO analysis for genes downregulated in *Auts2* KO MN versus WT MN.
(E) Scatter plot of $\log_2$(normalized expression of AUTS2-NRF1 co-targeted genes) in WT and *Auts2* HX mutant MN from scRNA-Seq.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

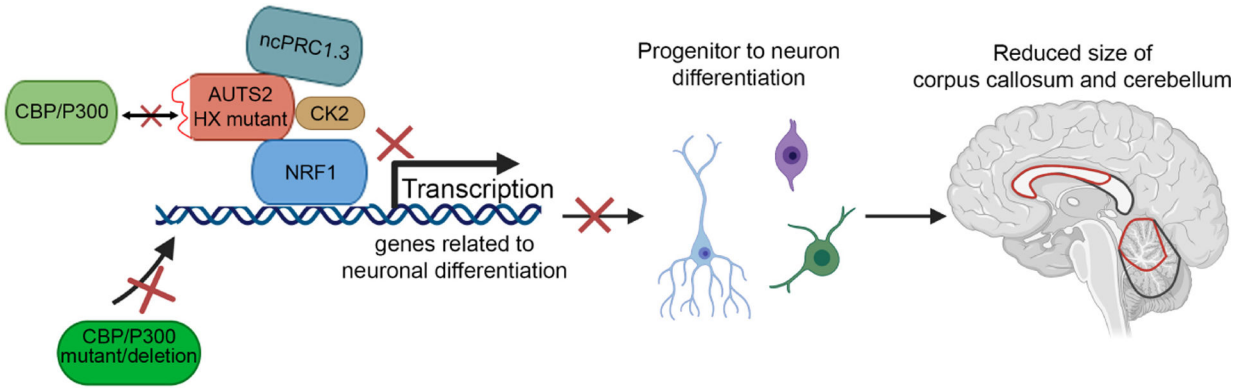


Figure 7. Model: AUTS2-ncPRC1.3 activates its targeted genes for brain development through NRF1-mediated recruitment and HX repeat-mediated P300 interaction.

Mutations or deletions of CBP or P300 impede neuronal differentiation during early brain development and result in the malformation of several brain regions, including the corpus collosum and cerebellum in RSTS patients. As shown here, mutations in the AUTS2 HX repeat domain impair AUTS2 interaction with P300 within the context of ncPRC1.3, effectively disabling transcription activation. We speculate such disruptions to appropriate gene activation lead to a defect in brain development that overlaps with RSTS.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-AUTS2	(Gao et al., 2014)	N/A
Rabbit anti-PCGF3	Abcam	Cat# ab201510
Rabbit anti-PCGF5	Abcam	Cat# ab201511
Rabbit anti-Ring1B	Bethyl Laboratories	Cat# A302–869A, RRID:AB_10632773
Rabbit anti-RYBP	Sigma	Cat# PRS2227, RRID:AB_1847589
Rabbit anti-NRF1	Abcam	Cat# ab34682, RRID:AB_2236220
Mouse anti-NRF1	Abcam	Cat# ab55744, RRID:AB_2154534
Mouse anti-P300	Active motif	Cat# 61401, RRID:AB_2716754
Rabbit anti-CBP	Cell Signaling Technology	Cat# 7389, RRID:AB_2616020
Goat anti-Oct4	Santa Cruz Biotechnology	Cat# sc-8628, RRID:AB_653551
Mouse anti-HB9	Developmental Studies Hybridoma Bank	Cat# 81.5C10, RRID:AB_2145209
Rabbit anti-H3	Abcam	Cat# ab1791, RRID:AB_302613
Rabbit anti-H2AK119ub1	Cell Signaling Technology	Cat# 8240; RRID: AB_10891618
Rabbit anti-H3K27me3	Cell Signaling Technology	Cat# 9733; RRID: AB_2616029
Rabbit anti-H3K27ac	Abcam	Cat# ab4729; RRID: AB_2118291
Rabbit anti-H3K4me3	Abcam	Cat# ab8580; RRID: AB_306649
Rabbit anti-Flag	Proteintech	Cat# 20543–1-AP; RRID: AB_11232216
Rabbit anti-Gal4	MilliporeSigma	Cat# 06–262; RRID: AB_310083
Rabbit anti- Cleaved Caspase-3	Cell Signaling Technology	Cat# 9664; RRID: AB 2070042
Rabbit anti-Vinculin	Cell Signaling Technology	Cat# 13901; RRID: AB_2728768
Rabbit anti-GAPDH	Cell Signaling Technology	Cat# 5174; RRID: AB_10622025
Rabbit anti-Drosophila-specific H2Av	Active motif	Cat# 39715, RRID:AB_2793318
Bacterial and Virus Strains		
Rosetta (DE3) Competent Cells	Novagen	Cat# 70954
One-Shot Stbl3 chemically competent cells	Invitrogen	Cat# C7373–03
DH10Bac	Thermo Fisher Scientific	Cat# 10361012
Chemicals, Peptides, and Recombinant Proteins		
Lipofectamine 2000	Thermo Fisher Scientific	Cat# 11668027
Retinoic Acid	Sigma	Cat# R2625
Smoothened Agonist, SAG	EMD/ Calbiochem	Cat# 566660
Anti-FLAG M2 affinity gel	Sigma	Cat# A2220
FLAG® Peptide	Sigma	Cat# F3290
Leukemia inhibitory factor	Reinberg lab	N/A
CHIR99021	Tocris	Cat# 4423

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
PD0325901	Sigma	Cat# PZ0162
Dynabeads protein G beads	Thermo Fisher Scientific	Cat# 10003D
SYBR-Gold Nucleic Acid Gel Stain	Thermo Fisher Scientific	Cat# S11494
Agencourt AMPure XP Beads	Beckman Coulter	Cat# A63882
RNase Inhibitor	NEB	Cat# M0314L
Poly Ethylene Glycol (PEG) 8000	Sigma	Cat# 89510–250G-F
GTP	Thermo Fisher Scientific	Cat# R1461
Calcein_AM	Thermo Fisher Scientific	Cat# C1430
Ethidium homodimer-1	Sigma	Cat# 46043–1MG-F
Tamoxifen	Sigma	Cat #T5648
Critical Commercial Assays		
Agilent High Sensitivity DNA Kit	Agilent Technologies	Cat# 5067–4626
Nextera XT DNA Library Preparation Kit	illumina	Cat# FC-131–1024
Maxima H Minus Reverse Transcriptase	Thermo Fisher	Cat# EP0751
KAPA HiFi Hotstart PCR kit	Roche	Cat# KK2502
Phusion High-Fidelity DNA Polymerase	NEB	Cat# M0530L
Quant-iT™ PicoGreen™ dsDNA Assay Kit	Thermo Fisher	Cat# P7589
RNEasy plus Mini Kit	QIAGEN	Cat# 74134
MinElute PCR Purification Kit	QIAGEN	Cat# 28006
BCA Protein Assay Kit	Thermo Scientific	Cat# 23225
Luciferase Assay System	Promega	Cat# E1500
End-It™ DNA End-Repair Kit	Lucigen	Cat# ER81050
Klenow Fragment (3’→5’ exo-)	NEB	Cat# M0212L
T4 DNA Ligase (Rapid)	QIAGEN	Cat# 50-305-905
Q5® High-Fidelity DNA Polymerase	NEB	Cat# M0491L
Deposited Data		
Mouse brain AUTS2, NRF1, RYBP, PCGF3 ChIP-Seq	This study	GEO: GSE161808
E14_ESC AUTS2, NRF1 ChIP-Seq	This study	GEO: GSE161808
Motor neuron AUTS2, NRF1, Ring1B, RYBP, H2AK119ub, H3K27me3, and H3K27ac ChIP-Seq	This study	GEO: GSE161808
Mouse brain, E14_ESC and Motor neuron RNA-Seq	This study	GEO: GSE161808
Motor neuron scRNA-Seq	This study	GEO: GSE161808
Mouse brain input, AUTS2, P300, H2AK119ub, H3K27me3, H3K27ac, H3K4me3, Ring1B and Pol II ChIP-Seq	(Gao et al., 2014)	GEO: GSE60411
E14_ESC input, Pol II ChIP-Seq	(LeRoy et al., 2019)	GEO: GSE117155
Motor neuron Pol II, H3K4me3 ChIP-Seq	(Narendra et al., 2015)	GEO: GSE60240
Experimental Models: Cell Lines		

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Human: HEK293 T-REx	Reinberg lab	N/A
Human: HEK293T 5XGal4TK-Luc	Reinberg lab	N/A
Human: HEK293T 5XGal4TK-Luc GAL4-AUTS2	Reinberg lab	N/A
Human: HEK293T 5XGal4TK-Luc GAL4-AUTS2 P517L	Reinberg lab	N/A
Human: HEK293T 5XGal4TK-Luc GAL4-AUTS2 T534P	Reinberg lab	N/A
Human: HEK293T 5XGal4TK-Luc GAL4-AUTS2 535-542 aa deletion	Reinberg lab	N/A
Human: HEK293T 5XGal4TK-Luc GAL4-AUTS2 sh-PCGF3/5	Reinberg lab	N/A
Mouse: ES cell line E14	Reinberg lab	N/A
Mouse: ES cell line E14 *Auts2* −/−	This study	N/A
Mouse: ES cell line E14 *Nrf1* −/−	This study	N/A
Mouse: ES cell line E14 *Auts2* T534P	This study	N/A
Mouse: ES cell line E14 *Auts2* 535–542 aa deletion	This study	N/A
Experimental Models: Organisms/Strains		
Mouse: C57BL/6J	Jackson Laboratory	Strain 000664
Mouse: *Nrf1*-CKO	(Kiyama et al., 2018)	N/A
Mouse: *Nrf1*-LacZ	(Kiyama et al., 2018)	N/A
Mouse: *Tbr1*-TauGFP-IRESCreERT2	(Kiyama et al., 2019)	N/A
Mouse: *Pou4f1*-CKOAP	(Badea et al., 2009)	Strain 010558
Oligonucleotides		
SgRNA and ssODN for *Auts2, Nrf1* knockout, *Auts2* HX mutant knock-in; shRNA targeting human PCGF3/5	This study	See Supplementary Table 2
Oligonucleotides, Barcode sequence for scRNA-Seq	This study	See Supplementary Table 3
Recombinant DNA		
pSpCas9(BB)-2A-GFP	Addgene	Cat # 48138
pSpCas9(BB)-2A-GFP-Auts2-KO-sgRNA	This study	N/A
pSpCas9(BB)-2A-GFP-Auts2-T534P-sgRNA	This study	N/A
pSpCas9(BB)-2A-GFP-Auts2-535-542-del-sgRNA	This study	N/A
pSpCas9(BB)-2A-GFP-Nrf1-KO-sgRNA	This study	N/A
PLKO.1 sh-hPCGF3	This study	N/A
PLKO.1 sh-hPCGF5	This study	N/A
pINTO-NFH(empty vector)	Reinberg lab	N/A
pINTO-NFH-AUTS2	This study	N/A
pINTO-NFH-AUTS2-P517L	This study	N/A
pINTO-NFH-AUTS2-T534P	This study	N/A
pINTO-NFH-AUTS2-535-542 del	This study	N/A
pINTO-Gal4(empty vector)	Reinberg lab	N/A

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pINTO-Gal4-AUTS2	This study	N/A
pINTO-Gal4-AUTS2- P517L	This study	N/A
pINTO-Gal4-AUTS2-T534P	This study	N/A
pINTO-Gal4-AUTS2-535-542 del	This study	N/A
Software and Algorithms		
Bowtie2 v2.4.1	(Langmead and Salzberg, 2012)	https://github.com/BenLangmead/bowtie2
STAR v2.6.1d	(Dobin et al., 2013)	https://github.com/alexdobin/STAR
MAC2 v2.1.1	(Zhang et al., 2008)	https://github.com/macs3-project/MACS
GSEA	(Subramanian et al., 2005)	http://software.broadinstitute.org/gsea/index.isp
fgsea	(Korotkevich et al., 2019)	https://github.com/ctlab/fgsea
BEDTools v2.27.1	(Quinlan and Hall, 2010)	https://github.com/arq5x/bedtools
Picard v2.18.11	Broad Institute	https://github.com/broadinstitute/picard
HOMER v4.10	(Heinz et al., 2010)	http://homer.ucsd.edu/
deepTools v3.2.1	(Ramírez et al., 2016)	https://github.com/deeptools/deepTools
Samtools v1.9	(Li et al., 2009)	https://github.com/samtools/
ChIPseeker v1.8.6	(Yu et al., 2015)	https://github.com/YuLab-SMU/ChIPseeker
Integrative Genomics Viewer (IGV)	Broad Institute	http://www.broadinstitute.org/igv/
Prism	GraphPad Software	https://www.graphpad.com/scientific-software/prism/
Benchling	Benchling.com	https://www.benchling.com/
SRA Toolkit	NCBI	https://github.com/ncbi/sra-tools
RStudio	RStudio	https://rstudio.com/
R V4.0.0	R-project	https://www.r-project.org/
featureCounts v1.6.3	(Liao et al., 2014)	https://github.com/torkian/subread-1.6.1
ComplexHeatmap	(Gu et al., 2016)	https://github.com/iokergoo/ComplexHeatmap
Seurat v3.1.4	(Butler et al., 2018)	https://github.com/satijalab/seurat
zUMIs v2.9.4	(Parekh et al., 2018)	https://github.com/sdparekh/zUMIs
Presto	(Korsunsky et al., 2019)	https://github.com/immunogenomics/presto