Parallel Functional Interrogations of Autism Risk Genes Reveal Drivers of Transcriptomic Convergence and the Female Protective Effect

Rebecca E. Andersen[1,2,3,4,*], Maya Talukdar[1,2,3,5,*], Tyler Sakamoto[1,6], David Exposito-Alonso[1,2,4,7,8], Janet H.T. Song[1,2,4,7,8], Xuyu Qian[1,2,4,7,8], Seungil Lee[1,6], Nila Murugan[1,9], Ryan N. Delgado[10], Sijing Zhao[1,11], Gwenyth Eichfeld[1,12], Julia Harms[1,13], David C. Page[14,15,16], and Christopher A. Walsh[1,2,3,4,7,8]

Affiliations:
[1]Division of Genetics and Genomics and Manton Center for Orphan Diseases, Boston Children's Hospital, Boston, MA, USA; [2]Department of Pediatrics, Harvard Medical School, Boston, MA, USA; [3]Broad Institute of MIT and Harvard, Cambridge, MA, USA; [4]Allen Discovery Center for Human Brain Evolution, Boston, MA, USA; [5]Harvard-MIT MD/PhD Program, Program in Biomedical Informatics, Boston, MA, USA; [6]Harvard College, Cambridge, MA, USA; [7]Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA, USA; [8]Department of Neurology, Harvard Medical School, Boston, MA, USA; [9]University of Massachusetts Boston, Boston, MA, USA; [10]Department of Genetics, Blavatnik Institute, Howard Hughes Medical Institute, Harvard Medical School, Boston, MA, USA; [11]Harvard BBS PhD Program, Boston, MA, USA; [12]Colgate University, Hamilton, NY, USA; [13]University of California Berkeley, Berkeley, CA, USA; [14]Whitehead Institute, Cambridge, MA 02142, USA; [15]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [16]Howard Hughes Medical Institute, Whitehead Institute, Cambridge, MA 02142, USA.

*These authors contributed equally.

Corresponding author and lead contact:

Christopher A. Walsh, christopher.walsh@childrens.harvard.edu

SUMMARY

The surprising number and functional diversity of genes implicated in autism spectrum disorder (ASD) has made it challenging to identify core pathophysiological mechanisms or envision interventions with broad therapeutic potential. Here, parallel CRISPR-Cas13-based knockdown of 28 ASD genes and neighboring long non-coding RNAs reveals striking convergence on shared transcriptomic effects and neurodevelopmental phenotypes in human neural progenitor cells and cerebral organoids. Perturbations of single ASD genes caused the widespread dysregulation of other ASD genes, and *de novo* reconstruction of gene regulatory networks uncovered the prominent autism risk gene *CHD8* as a critical driver of this transcriptomic convergence. The transcriptional activator *ZFX*, which escapes X chromosome inactivation in females, was also identified as a key regulator of ASD genes, revealing genetic underpinnings of the female protective effect. Thus, this study provides a crucial framework for uncovering how variants in diverse genes can cause convergent pathophysiological effects that ultimately result in a shared diagnosis.

INTRODUCTION

Autism spectrum disorder (ASD) is a common neurodevelopmental condition characterized by deficits in social communication as well as restricted and/or repetitive behaviors. Although ASD is clinically heterogeneous, these central features can be used diagnostically to distinguish ASD from other neurodevelopmental disorders (DSM-5, 2013), indicating that certain biological foundations of ASD are shared across individuals. Moreover, ASD has a significant genetic component (Bai et al. 2019), and several large-scale studies have identified genes implicated in ASD (ASD genes) (Grove et al. 2019; Satterstrom et al. 2020; Trost et al. 2022). These studies have enabled the compilation of ASD genes into large databases such as the Simons Foundation Autism Research Initiative (SFARI) Gene database (Abrahams et al. 2013), which has surprisingly revealed that hundreds of genes with diverse functions are implicated in ASD. Understanding how disruptions in such seemingly disparate genes lead to the core pathophysiology of ASD remains a major challenge (Liao et al. 2023).

Furthermore, while previous studies have primarily focused on identifying protein-coding genes (PCGs) implicated in ASD, there is emerging evidence that non-coding sequences can also play important roles in brain development. For instance, genetic disruptions in multiple long non-coding RNAs (lncRNAs) have now been implicated in ASD (Noor et al. 2010; Luo et al. 2018; Ang et al. 2019; Andersen et al. 2024), suggesting that mutations in lncRNAs can cause ASD. Moreover, transcriptomic analyses of ASD brain tissue have found that lncRNAs are frequently dysregulated in ASD (Ziats and Rennert 2013; Parikshak et al. 2016), including lncRNAs that are in close proximity to ASD-associated PCGs. As certain classes of lncRNAs have been found to modulate the expression and/or function of their nearby genes (Luo et al. 2016; Wang et al. 2020), this raises the possibility that lncRNAs neighboring ASD PCGs could play particularly critical roles in ASD-relevant molecular pathways even if they are not directly disrupted.

Additionally, ASD has a significantly higher prevalence in males, with approximately 3-4 males diagnosed for every female (Loomes et al. 2017; Maenner et al. 2023). Although ASD genes have

been identified on the X chromosome (Mendes et al. 2024), X-linked causes of ASD alone cannot fully explain the male-biased prevalence of ASD (Ropers and Hamel 2005; Werling 2016; Martin et al. 2021). Moreover, females with ASD tend to have a greater genetic load of autosomal ASD variants than males (Jacquemont et al. 2014; Pinto et al. 2014; Zhang et al. 2020; Napolitano et al. 2022; Wigdor et al. 2022; Antaki et al. 2022; Warrier et al. 2022), suggesting that females require more autosomal genetic risk to manifest ASD. However, the genetic mechanisms of this "female protective effect" (Werling 2016; Dougherty et al. 2022) have yet to be uncovered.

Here, functional interrogation of ASD genes and neighboring lncRNAs in human neural progenitor cells (NPCs) and cerebral organoids reveals remarkably widespread transcriptomic convergence, which we define as statistically significant overlap of the same individual differentially expressed genes (DEGs) across different conditions. Strikingly, most perturbations resulted in DEGs that were enriched for other ASD genes. *De novo* reconstruction of a gene regulatory network (GRN) enabled the identification of central regulators, including the well-known ASD gene *CHD8* as well as novel candidates such as *REST*, that drive transcriptomic convergence in ASD. Furthermore, the X-linked transcriptional activator *ZFX*, which escapes X chromosome inactivation (XCI) in females, emerged as a key regulator of ASD genes, demonstrating that expression from the "inactive" X chromosome is an important biological foundation of the female protective effect. Disruption of several ASD genes and neighboring lncRNAs also led to phenotypic convergence, affecting the critical processes of proliferation in NPC cultures and neurogenesis in cerebral organoids. Thus, this work reveals how key GRNs become broadly dysregulated upon disruption of a single ASD gene, uncovering molecular mechanisms of transcriptomic convergence in ASD. Moreover, these results highlight central regulators such as *CHD8* and *REST*, which are promising targets for future therapeutic interventions that could provide broad benefits across diverse genetic causes of ASD.

RESULTS

Consensus-ASD Reveals Shared and Unique Features Across Classes of Genetic Variation

The genetic underpinnings of ASD have primarily been investigated in the context of genes that are recurrently disrupted by rare *de novo* copy number variants (CNVs) or single nucleotide variants (SNVs) in ASD. However, ASD has a complex genetic architecture with contributions from multiple distinct genetic variant classes (Iakoucheva et al. 2019; Dias and Walsh 2020; Willsey et al. 2022). Thus, we first systematically aggregated variants implicated in ASD from 32 different studies across seven different genetic variant classes: 1) germline SNVs and indels implicated in idiopathic ASD, 2) syndromic variants with ASD as a recurrent feature, 3) somatic SNVs and indels, 4) germline CNVs, 5) somatic CNVs, 6) inherited recessive variants, and 7) common variants (GWAS SNPs). This dataset, which we refer to as *Consensus-ASD* (**Fig. 1A**), enabled us to identify shared and unique features across variant classes (**Fig. S1-S2** and **Table S1**), highlighting differences in genetic constraint (Lek et al. 2016) and developmental expression patterns.

Depletion of ASD Genes and Adjacent lncRNAs Causes Similar Effects on Gene Expression

Functional interrogation of ASD genes and their adjacent lncRNAs revealed strongly overlapping effects on downstream targets. Knockdown (KD) studies were performed for a total of 38 genes, including 19 SFARI genes and 17 neighboring lncRNAs, as closely adjacent lncRNAs have often been found to regulate the expression and/or function of their neighbors (Villegas and Zaphiropoulos 2015; Luo et al. 2016; Wang et al. 2020). This included the lncRNA *NR2F1-AS1*, which has itself been implicated in ASD (Ang et al. 2019). *NR2F1-AS1* is adjacent to the SFARI gene *NR2F1*, which was also included in our study. Additionally, we included the lncRNA *SOX2-OT*, a lncRNA recently implicated in ASD (Andersen et al. 2024), along with its overlapped protein-coding gene (PCG) *SOX2*, a transcription factor known to play important roles in neural development (Graham et al. 2003; Lee et al. 2014).

For each of these genes (**Table S2**), CRISPR-Cas13 (Konermann et al. 2018; Wei et al. 2023) and an array of three guide RNAs (gRNAs) targeting the gene (**Table S3**) were used for KD in human XY (male) NPC cultures (**Fig. 1B-C**). RNA-seq was performed after 24 hours to determine the acute effects of these perturbations. Most (28/38, 73.7%) of the target genes exhibited statistically significant KD relative to non-targeting control (NTC) samples (**Fig. 1D** and **Table S4**), including 19/20 (95%) of the PCGs and 9/18 (50%) of the lncRNAs. Subsequent analyses were limited to the 28 conditions that demonstrated significant KD.

While only 1 lncRNA (1/19, 5.26%) was significantly differentially expressed upon KD of its neighboring PCG, the majority of lncRNA KDs (5/9, 55.6%) led to significant differential expression of their neighboring PCG (**Fig. 1E** and **S3**). Of these, most (4/5, 80%) of the lncRNA KDs resulted in decreased expression of the PCG neighbor, consistent with previous reports that closely adjacent lncRNAs often positively regulate the expression of their nearby neighbors (Villegas and Zaphiropoulos 2015; Luo et al. 2016; Wang et al. 2020). Transcriptome-wide analysis, limited to the 25 samples with at least 50 differentially expressed genes (DEGs), revealed a strong and significant correlation in fold change of DEGs between each neighboring lncRNA-PCG pair (8/8, $p < 2.2e\text{-}16$ in each comparison; testing beta $\neq 0$) (**Fig. S4**). Linear modeling determined that lncRNA-KD fold changes were smaller than PCG-KD fold changes on a per-gene basis (8/8, $p < 2.2e\text{-}16$ in each comparison; testing beta < 0) (**Fig. S4**), indicating that perturbation of these lncRNAs and their neighboring PCGs affect shared downstream genes but that the magnitude of the effect tends to be weaker upon KD of the lncRNA. This is consistent with lncRNAs often functioning to fine-tune the expression of critical genes and signaling pathways (Zhao et al. 2021). Thus, lncRNAs adjacent to ASD genes, and particularly the lncRNAs highlighted here, warrant further consideration as potential contributors to ASD.

ASD-Related Perturbations Result in Widespread Dysregulation of Other ASD Genes

Transcriptome-wide analysis revealed that the ASD-relevant perturbations caused remarkably broad disruption of other ASD genes. Most sets of DEGs from the individual KDs exhibited enrichment for genes in *Consensus-ASD* (21/25, 84%) (**Fig. 2A**) as well as SFARI genes (19/25, 76%) (**Fig. S5A**). This enrichment was uniformly weaker when considering genes implicated in intellectual disability (SysID) (Kochinke et al. 2016) (**Fig. S5A**), suggesting that perturbation of ASD genes particularly affects other ASD genes as opposed to neurodevelopmental genes more broadly. Thus, perturbation of known ASD genes and neighboring lncRNAs leads to widespread disruption of other ASD genes.

Surprisingly, we found a striking pattern of transcriptomic convergence on the same individual DEGs across the different perturbations. This was particularly evident when focusing on high confidence SFARI (HC-SFARI) genes, the SFARI genes in the highest evidentiary tier for association with ASD (**Fig. 2B**). Not only were the same specific genes shared across multiple perturbations, but they were also largely affected in the same direction by the different KDs. Moreover, HC-SFARI genes tended to shift in a concordant direction as a result of the different perturbations even when they did not rise to the level of a statistically significant DEG within a particular KD sample (**Fig. 2C** and **S5B**). While previous functional studies of ASD genes have identified certain overlapping changes in cell type composition and behavior, evidence of transcriptomic convergence is quite limited (Paulsen et al. 2022; Weinschutz Mendes et al. 2023; Li et al. 2023). Indeed, based on functional studies in cerebral organoids it has been proposed that different ASD mutations largely impact distinct downstream genes (Paulsen et al. 2022). Thus, the strong convergence on shared ASD DEGs that we observed across numerous ASD perturbations was highly unexpected. Nevertheless, common patterns of DEGs have been identified in brain tissue from ASD cases compared to neurotypical controls (Voineagu et al. 2011; Parikshak et al. 2016; Ramaswami et al. 2020; Gandal et al. 2022), consistent with our findings.

Further analysis of transcriptomic convergence enabled the discovery of candidate ASD risk genes. We identified genes with following features: 1) significantly downregulated by at least three

of the ASD KDs; 2) at least 80% concordant in the direction of significant differential expression across all ASD KDs; and 3) intolerant to loss of function variants, as determined by LOEUF scores of less than 0.6 ("constrained"). This analysis resulted in a set of 584 genes (**Table S5**), of which 138 (23.6%) were SFARI genes. This strong enrichment (p = 6.47e-08) of SFARI genes within the candidate ASD risk genes demonstrates that these criteria can effectively prioritize genes that indeed contribute to ASD.

Key Transcription Factors Including *REST* Target Recurrently Differentially Expressed Genes

Transcription factors (TFs) that mediate transcriptomic convergence in ASD were uncovered through MAGIC analysis (Roopra 2020), which identified TFs whose target genes were significantly differentially expressed in at least 20/25 (80%) of the perturbations (recurrent MAGIC TFs) (**Fig. 2D** and **S5C**). Many of these significant TFs are involved in pathways or complexes previously linked to ASD, such as SWI/SNF (Wenderski et al. 2020; Valencia et al. 2023), WNT (Kalkman 2012), and CTCF (Price et al. 2023), or have themselves been implicated in ASD. Of the recurrent MAGIC TFs that have not been directly implicated in ASD, several exhibited high genetic constraint and co-expression with known ASD genes, features that have been used to nominate candidate ASD risk genes (Liao et al. 2023).

Of particular interest, RE1 silencing transcription factor (REST) was a significant MAGIC TF in 24/25 (96%) of our perturbations. While REST has not previously been directly associated with ASD, it is known to carry out critical neurodevelopmental roles, functioning as a master negative regulator of neurogenesis (Schoenherr and Anderson 1995; Chong et al. 1995). Re-analysis of REST binding sites from embryonic stem cells (ESCs) (Rockowitz and Zheng 2015) revealed strong enrichment of SFARI and *Consensus-ASD* genes within REST target genes from both human (adj. p = 2.7e-47 and 1.0e-43, respectively) and mouse (adj. p = 4.2e-40 and 4.2e-40, respectively) (**Fig. 2E**). Genes implicated in intellectual disability (SysID) were also enriched (adj. p = 7.6e-13 and 2.3e-10, respectively); however, this enrichment was notably

weaker than that for ASD genes, suggesting that REST is especially important in regulating ASD-relevant gene expression.

Gene Regulatory Network Analysis Identifies *CHD8* as a Top Driver of Convergence in ASD

De novo reconstruction of an NPC gene regulatory network (GRN) uncovered critical regulatory relationships underlying transcriptomic convergence. The bioinformatic tool ARACNe (Margolin et al. 2006; Lachmann et al. 2016) was used to build GRNs linking transcription factors (Lambert et al. 2018) and epigenetic regulators (Boukas et al. 2019) to their downstream targets (regulons). To confirm the validity of this approach, we first constructed a GRN that excluded data from the seven samples in which the KD target was a potential regulator, which enabled evaluation of the predictive ability of the ARACNe-derived GRN through comparison with the held-out RNA-seq data. For all 7/7 (100%) of the KDs, strong and significant cross-enrichment was found between the downregulated DEGs and the 1st and 2nd degree ARACNe regulons (i.e., direct targets as well as their direct targets) ($p < 2.2e-16$, Fisher's combined probability test) (**Fig. 3A** and **S6**). Thus, ARACNe is able to effectively capture true regulatory relationships. We then built an ARACNe GRN from our full set of RNA-seq data. In total, this identified 101,918 putative regulatory relationships between 1,098 known transcription factors or epigenetic regulators and 9,464 target genes, including many highly interconnected relationships between known ASD genes (**Fig. 3B**).

Genes implicated in ASD are central regulators of the NPC GRN, as demonstrated by analysis using the PageRank algorithm (Page et al. 1999) (**Fig. 3C**). Genes in *Consensus-ASD* were strongly enriched for high PageRank scores (adj. p = 1.3e-9), indicating their centrality to the GRN. This enrichment was also significant for SFARI genes (adj. p = 6.7e-4), while SysID genes demonstrated much weaker enrichment (adj. p = 0.03). Analysis of Hallmark Molecular Signature Databases pathways (Subramanian et al. 2005) further identified strong enrichment for genes associated with cell division and proliferation such as G2M checkpoint (adj. p = 1.8e-

3) and mitotic spindle (adj. p = 1.1e-4) within the top PageRank genes, consistent with previous findings of ASD-related effects on NPC proliferation (Connacher et al. 2022) and tubulin biology (Lasser et al. 2023; Sun et al. 2024; Kostyanovskaya et al. 2025).

Further GRN analysis identified key regulators driving transcriptomic convergence in ASD. Regulators whose activity was consistently altered by the KD of ASD genes were identified using the bioinformatic tool VIPER (Alvarez et al. 2016). Of these, we focused on top drivers of convergence in ASD by identifying regulators whose targets were enriched for genes in *Consensus-ASD* and for genes that are highly central to the network (i.e. gene with high PageRank scores), which identified 78 key regulators (**Fig. S7**). Over half of these top regulators are *Consensus-ASD* genes themselves (41/78, 52.6%, adj. p = 1.5e-19), and over one quarter are HC-SFARI genes (21/78, 26.9%, adj. p = 2.e3-20). These top regulators exhibited remarkably strong cross-regulation (p = 0), as determined by mean normalized degree (**Fig. S7**). Analysis of the STRING database of protein-protein interactions (PPIs) (Szklarczyk et al. 2023) further confirmed that these top regulators are highly interconnected (p = 2.75e-11, STRING PPI Enrichment Score) (**Fig. S8**). Taken together, this suggests that these top regulators not only modulate many ASD genes but also cross-regulate each other, driving transcriptomic convergence in ASD.

Iterative walktrap clustering identified 5 gene modules within the 78 top regulators that were especially strongly interconnected (**Fig. 3D** and **S9**). Of particular interest, the largest of these (**Fig. S9C**) demonstrates a high degree of cross-interaction throughout the module and contains several notable ASD genes including *CHD8*, one of the strongest known risk genes for ASD (Haddad Derafshi et al. 2022). The inferred *CHD8* regulons from the GRN were strongly enriched (1^{st} degree regulon adj. p = 8.7e-8, 1^{st} and 2^{nd} degree regulon adj. p = 9.0e-24) for genes that were previously identified as direct CHD8 targets through ChIP-seq from NPCs (Sugathan et al. 2014) (**Fig. 3E**). The CHD8 ChIP-seq targets were also highly enriched for genes in *Consensus-ASD* (adj. p = 4.64e-20), SFARI genes (adj. p = 1.13e-7), and HC-SFARI

genes (adj. p = 1.13e-7), further supporting the role of *CHD8* as a key regulator of ASD gene expression.

Furthermore, the *CHD8*-containing module included three regulators (*SETD2*, *SETD5*, and *KMT2E*) that we had directly perturbed and which we would predict to share many downstream targets with *CHD8*. Indeed, the DEGs from these KDs also exhibited enrichment for CHD8 ChIP-seq targets (adj. p = 8.81e-6, 1.60e-7, and 1.36e-4, respectively), further demonstrating the convergent regulation of shared targets by critical ASD genes. Interestingly, *CHD8* is not differentially expressed upon *KMT2E* KD; similarly, *KMT2E* was not identified as a direct target of CHD8 by ChIP-seq (Sugathan et al. 2014). Consistent with this, ARACNe analysis did not identify a direct interaction between *CHD8* and *KMT2E*. This suggests that even ASD genes that do not directly regulate each other may jointly regulate a shared set of downstream genes, leading to convergent changes upon their disruption. Moreover, this analysis highlights *CHD8*, a high confidence ASD risk gene, as a key regulator driving transcriptomic convergence in ASD.

The X-Linked Transcription Factor *ZFX* is a Key Contributor to the Female Protective Effect

Sex-differential patterns of gene expression could contribute to the female protective effect (FPE) in ASD. Prior studies have found that autosomal genes that are disrupted or downregulated in ASD tend to be more highly expressed in the brains of neurotypical females compared to males, suggesting that female-biased expression of ASD-relevant genes may contribute to the FPE (Werling et al. 2016; Velmeshev et al. 2023). We hypothesized that X-linked transcriptional regulators that escape XCI could drive these female-biased expression patterns. Thus, we first sought to identify constrained (LOEUF < 0.6) X-linked regulators of the candidate ASD risk genes (**Table S5**) defined earlier in this study.

MAGIC analysis revealed 155 constrained regulators whose targets were significantly enriched for candidate ASD risk genes (**Table S6**), including 39 regulators that are themselves

SFARI genes. Eight of the regulators were X-linked, and their targets were also found to be significantly enriched within the full set of DEGs from many of the individual ASD gene KDs (**Fig. 4A**). To further identify transcriptional regulators of candidate ASD risk genes, we analyzed our NPC GRN, which identified 229 constrained regulators whose regulons were significantly enriched for candidate ASD risk genes (**Table S7**), including 74 regulators that were SFARI genes and ten regulators that were X-linked (**Fig. 4B** and **S10A**). Four of these X-linked regulators could also be analyzed by MAGIC: all of them (4/4, 100%) were positively identified as significant MAGIC regulators, further reinforcing the strongly concordant results between these fundamentally different analytical approaches. Together, these analyses identified 14 constrained X-linked regulators of candidate ASD risk genes, half of which (7/14, 50%) have themselves been implicated in ASD and are considered SFARI genes. This highlights several X-linked genes that could contribute to the greater prevalence of ASD in males through their regulation of candidate ASD risk genes (**Fig. 4C**).

We then sought to uncover regulators that could contribute to the FPE. Of the 14 regulators identified above, only *ZFX* has been shown to escape XCI in females (San Roman et al. 2023) and exhibit female-biased expression in the developing cerebral cortex (Kissel et al. 2024). Indeed, *ZFX* is consistently found to be one of the most highly female-biased genes across a variety of tissue and cell types (Oliva et al. 2020; Blanton et al. 2024). Moreover, *ZFX* is a transcriptional activator that has been shown to mediate widespread upregulation of many autosomal genes (Raznahan et al. 2018; San Roman et al. 2024), supporting its potential contribution to the FPE. Interestingly, pathogenic variants in *ZFX* have recently been directly implicated in ASD (Shepherdson et al. 2024). In our perturbation studies, KD of several different ASD genes led to downregulation of *ZFX* (**Fig. 4D**), suggesting that decreased *ZFX* activity can also occur in ASD without requiring direct mutations in *ZFX*. As females exhibit higher expression of *ZFX*, they are better buffered against this change, and thus *ZFX* could contribute to sex-differential outcomes even when it is not directly mutated.

We further investigated the role of *ZFX* in ASD through re-analysis of *ZFX* targets that were previously identified based on ChIP-seq and CRISPR KD experiments in two separate cell lines (San Roman et al. 2024). Our analysis demonstrated that these *ZFX* targets are enriched for ASD-relevant gene sets including SFARI genes and genes that exhibit female-biased expression in the brain (**Fig. 4E**). The enrichment of ASD-relevant gene sets was further validated through re-analysis of additional ChIP-seq data (Rhie et al. 2018) (**Fig. S10B**). Furthermore, we analyzed our NPC GRN and found that the *ZFX* 1st and 2nd degree regulon includes many (12/78, 15.4%) of the top drivers of transcriptomic convergence in ASD that we had previously identified (**Fig. 4F**). Thus, *ZFX* is an XCI-escaping female-biased regulator of ASD risk genes that may provide a crucial foundation for the FPE by driving sex-differential expression of genes that are disrupted in ASD.

ASD Genes and Adjacent lncRNAs Regulate NPC Proliferation

We next focused on uncovering potential phenotypic convergence and found that several ASD gene perturbations caused similar effects on NPC proliferation. Analysis using the heritable CytoTrack dye (**Fig. 5A**) revealed that 8/16 (50%) of the SFARI gene KDs led to a statistically significant decrease in proliferation (**Fig. 5B** and **Table S8**). This was also the case for the lncRNA *SOX2-OT*, which has previously been implicated in ASD (Andersen et al. 2024), as well as its overlapped PCG *SOX2*, a TF known to regulate NPC proliferation (Graham et al. 2003; Lee et al. 2014). In contrast, 1/16 (6.25%) SFARI KDs and 3/8 (37.5%) lncRNA KDs led to increased NPC proliferation (**Fig. 5B**), while the remaining 11 perturbations yielded insignificant and/or inconsistent effects (**Fig. S11**). The RNA-seq data from these perturbations did not cluster based on proliferation phenotype (**Fig. S12**), suggesting that phenotypic effects for processes such as proliferation may be difficult to predict from RNA-seq data alone, and underscoring the importance of experimental analysis of such phenotypes.

Disruption of ASD Genes and Adjacent lncRNAs Alters Cerebral Organoid Development

Inducible KD in cerebral organoids further revealed shared neurodevelopmental phenotypes in a complex model system. We created the FLEx-based Inducible CRISPR Knockdown (FLICK) construct by engineering several features into a backbone incorporating the FLEx design (Schnütgen et al. 2003). Ultimately, the FLICK construct enables control over the timing of KD through doxycycline-inducible genetic recombination that results in the constitutive expression of Cas13 and an EGFP reporter (**Fig. 5C**), and this was confirmed to provide strong KD (**Fig. S13A**). The FLICK constructs also contained identifying (ID) sequences that can be captured during scRNA-seq and used to determine the KD target in a given cell, enabling the analysis of mosaic cerebral organoids. For long-term studies, we targeted ASD genes and adjacent lncRNAs whose KDs had not been found to impair NPC proliferation (**Fig. 5B** and **Fig. S11**) or survival (not shown). FLICK constructs targeting these various genes were separately transposed into iPSCs, and subsequently the cells were pooled to generate mosaic organoids (d0). KD was induced by one day of treatment with doxycycline on d14-15, when the organoids predominantly consist of PAX6+ NPCs (**Fig. S13B**). The organoids were grown until they reached one or two months of development (**Fig. 5D**), at which point analysis through scRNA-seq enabled the identification of multiple cell types including radial glial progenitors (RGs), intermediate progenitor cells (IPCs), excitatory neurons (ExNs), and inhibitory neurons (INs) (**Fig. 5E-F** and **Fig. S13C-D**).

These long-term organoid KDs revealed shared changes in cell type composition across multiple perturbations. Cells with FLICK constructs targeting the SFARI genes *BAZ2B*, *CLASP1*, and *EHMT1* as well as the lncRNA *ST7-AS1* appeared relatively depleted from the organoids at both d30 and d60 (**Fig. 6A**), despite our previous finding that these KDs did not impair NPC proliferation or survival (**Fig. 5B** and **Fig. S11**). Furthermore, these KDs also altered cell type proportions, leading to relatively more ExNs and fewer RGs (**Fig. 6B**). Together, this suggests that disruption of these genes may result in increased ExN differentiation at the expense of RG

self-renewal, leading to an overall reduction in cell number in the developing organoids. Additionally, several other KDs appeared to result in an increased proportion of ExNs by the 2-month timepoint, including the SFARI genes *PPP3CA* and *WDFY3* as well as the lncRNAs *NR2F1-AS1* and *WDFY3-AS2*. Thus, the disruption of multiple different ASD genes and neighboring lncRNAs results in this convergent neurodevelopmental phenotype, consistent with previous observations (Paulsen et al. 2022; Jourdon et al. 2023).

Furthermore, depletion of ASD genes and adjacent lncRNAs resulted in prominent transcriptomic convergence (**Fig. 6C** and **S14A**). Effects on gene expression were analyzed for each major cell type, which revealed highly overlapping DEGs in RGs with different KDs, further supporting our previous findings from NPC cultures. Indeed, for many of the KD targets, the DEGs that resulted from KD in organoid RGs strongly overlapped with the DEGs that we had previously identified from NPC cultures (**Fig. S14B**), suggesting that these effects on gene expression are not specific to a particular model system. Furthermore, strongly overlapping DEGs were also observed within additional cell types including ExNs (**Fig. 6C**) and IPCs (**Fig. S14A**), demonstrating that the convergence of ASD genes on shared downstream effects is not unique to neural progenitors. However, some of the specific genes that were convergently disrupted across perturbations differed between cell types, even when only considering genes expressed in all of the cell types (**Fig. 6D**), suggesting that cell-type-specific features impact the transcriptomic consequences of perturbing ASD genes. Additionally, down-sampling analysis revealed that the detection of overlapping DEGs heavily depended upon the number of cells sequenced for each perturbation (**Fig. S14C**). This demonstrates that strong transcriptomic convergence cannot be detected unless sufficiently large numbers of cells are analyzed.

Several ASD gene and adjacent lncRNA perturbations resulted in DEGs that were enriched for other ASD-relevant gene sets, including SFARI genes (**Fig. S14D**). This enrichment was observed across multiple cell types at both the d30 and d60 timepoints, further demonstrating that disruption of a single target gene can lead to broad dysregulation of many ASD genes.

Furthermore, shared DEGs implicated key transcriptional regulators as potential drivers of transcriptomic convergence. Analysis using the bioinformatics tool MAGIC (Roopra 2020) identified regulators whose putative target genes were disrupted by several of the perturbations (**Fig. 6E** and **Fig. S15**). Of the top 25 recurrent regulators identified from the RGs of 1-month organoids (**Fig. S15**), over half (14/25, 56%) overlapped with the top recurrent MAGIC TFs that we had previously identified from NPCs (**Fig. 2D**). Moreover, this analysis revealed recurrent regulators in ExNs from 1-month and 2-month organoids (**Fig. S15**), 21 of which were shared by ExNs from both timepoints. This set contained several SFARI genes (5/21, 23.8%), including CHD family members *CHD1* and *CHD7* (*CHD8* is not analyzed by MAGIC due to lack of ENCODE data at the time of MAGIC's creation). It also contained multiple genes that had been identified as top candidate regulators through previous GRN analysis (5/21, 23.8%) including *REST*. Interestingly, ten regulators including *REST* were also shared with the 1-month RGs, suggesting that these regulators are important drivers of convergence in multiple cellular contexts.

Transcriptomic Convergence is Further Revealed by Acute Perturbations in Cerebral Organoids

Acute disruption of ASD genes and adjacent lncRNAs in 2-month organoids revealed even greater convergence on shared DEGs. We allowed an additional set of mosaic FLICK organoids to grow for two months prior to inducing KD by treatment with doxycycline at d61-d62 and performing scRNA-seq two days later, allowing us to determine the acute effects of disrupting these genes within organoids that had developed normally. This approach resulted in similar numbers of cells with each of the different KDs (**Fig. 7A** and **S16A**), enabling deeper analysis of the KDs that had been depleted from the organoids with long-term KD (**Fig. 6A**). These organoids contained a mixture of cell types including RGs, IPCs, and ExNs (**Fig. 7B** and **Fig. S15B-C**); however, we focused on ExNs, as they were by far the most abundant cell type at this timepoint. Interestingly, when limiting our analysis to genes that are expressed in both the

organoid ExNs and our earlier NPC cultures, many of the KDs resulted in ExN DEGs that were enriched for the NPC DEGs that we had previously identified (**Fig. 7C**). While there are certainly cell-type-specific effects of the different KDs, this demonstrates that there are also shared changes in gene expression across cell types that can be partially modeled in experimentally tractable NPC cultures.

Furthermore, for perturbations that had been depleted from organoids over long-term KD, acute KD enabled us to obtain much greater cell numbers, revealing that these KDs similarly resulted in transcriptomic convergence (**Fig. 7D** and **S16D-E**). The DEGs exhibited strong enrichment of ASD-relevant genes, including SFARI genes, across several of the perturbations (**Fig. S16F**). Analysis of transcriptional regulators through MAGIC (Roopra 2020) identified 31 transcription factors and co-factors whose targets were enriched within the DEGs from at least half of the perturbations (**Fig. 7E**). These included 8/10 (80%) of the top regulators that had previously been identified in all three analyses from the long-term organoid KDs (1-month RGs, 1-month ExNs, and 2-month ExNs). Moreover, seven of these regulators, including *REST*, had also been identified as top recurrent MAGIC TFs from the previous NPC studies (**Fig. 2D**), further suggesting that they are key drivers of transcriptomic convergence in multiple cellular contexts and across acute and long-term time scales. Together, these studies reveal ASD genes and neighboring lncRNAs whose disruptions result in convergent effects on ASD gene expression as well as shared neurodevelopmental phenotypes in NPCs and cerebral organoids.

DISCUSSION

Here, we present several functional studies of ASD genes and neighboring lncRNAs in human NPCs and cerebral organoids, which demonstrated strong transcriptomic convergence including widespread dysregulation of ASD genes. GRN analysis revealed regulatory relationships that highlighted key chromatin modifiers and transcription factors, including *CHD8* and *REST*, as critical drivers of transcriptomic convergence. These regulators were highly

interconnected, with many demonstrating mutual regulation of one another, revealing how disruption of a single regulator can lead to changes that propagate throughout the network. Furthermore, the X-linked transcription factor *ZFX* was identified as a key contributor to the female protective effect through its female-biased expression and transcriptional activation of ASD-associated genes, which could buffer the effects of disrupting these genes in females. More broadly, this highlights a role for the “inactive” X chromosome in potentially mediating sex differences in ASD. Several ASD gene or neighboring lncRNA perturbations also altered NPC proliferation, demonstrating that these disruptions can result in shared phenotypic effects. Moreover, several KDs in cerebral organoids led to altered cell type proportions, resulting in a relative increase in ExNs. Thus, this work revealed critical insights into how diverse genes, including lncRNAs, can converge on core transcriptomic changes and neurodevelopmental phenotypes, ultimately resulting in ASD.

To uncover transcriptomic convergence, we initially focused on functional studies of ASD genes and neighboring lncRNAs in highly homogeneous human NPC cultures. This enabled bulk RNA-seq of each perturbation, providing the deep transcriptomic data that was essential for robust GRN reconstruction. Building an understanding of the complex relationships within GRNs is critical for identifying key “rate-limiting” factors that affect the GRN as a whole and ultimately lead to a particular biological outcome (Chakravarti and Turner 2016). For instance, studies of Hirschsprung disease have identified a crucial GRN centered on the rate-limiting gene *RET* that can be impacted by direct mutations in *RET* itself or in upstream regulators of *RET* (Chatterjee et al. 2016). Here, we have similarly identified critical genes such as *CHD8* and *REST* whose effects can propagate throughout the GRN, driving transcriptomic convergence in ASD.

While some of the specific regulatory relationships uncovered here may be unique to NPCs, our subsequent studies in cerebral organoids implicated several of these regulators in ExNs as well. Furthermore, the organoid analyses demonstrated that ExNs also exhibit convergence of ASD genes on shared downstream targets. Thus, future studies to reconstruct GRNs from

457 additional cell types including ExNs will be highly valuable for understanding how regulatory
458 relationships are reorganized in different cellular contexts. While bioinformatic pipelines for
459 reconstructing GRNs from scRNA-seq data are actively being refined (Vlahos et al. 2021;
460 Obradovic et al. 2021), these approaches are currently hampered by limited RNA capture
461 inherent to the scRNA-seq techniques. As these continue to improve, it will be feasible to
462 reconstruct robust GRNs from heterogeneous models, including organoids, as well as patient
463 tissue samples, which will further refine our understanding of gene regulation in complex
464 biological systems.

465 This study also identified several lncRNAs that affect the expression of their ASD gene
466 neighbor or other ASD genes more broadly. Intriguingly, the transcriptomic changes resulting
467 from the lncRNA KDs in NPCs were highly correlated with those from the KD of their ASD gene
468 neighbors, though the lncRNAs generally had milder effects, consistent with previous findings
469 that lncRNAs tend to be more subtle modulators of gene expression (Gao et al. 2020). This was
470 also the case for *NR2F1-AS1* (Ang et al. 2019) and *SOX2-OT* (Andersen et al. 2024), lncRNAs
471 that have been directly implicated in ASD. Thus, these lncRNAs may play neurodevelopmental
472 roles through their effects on ASD genes, providing context to previous findings that many
473 lncRNAs are dysregulated in ASD (Parikshak et al. 2016). Importantly, lncRNAs have often
474 been excluded or overlooked in genetic studies (Mattick et al. 2023), and these findings
475 highlight the value of especially careful consideration of lncRNAs that closely neighbor known
476 critical genes.

477 Taken together, these results have broad implications for our understanding of
478 neurodevelopment and disorders such as ASD. In particular, this work identified several
479 regulators that drive transcriptomic convergence, including the prominent ASD gene *CHD8* as
480 well as candidates including *REST*. While *REST* has not been directly implicated in ASD,
481 individuals with ASD exhibit decreased expression of *REST* targets (Katayama et al. 2016), and
482 treatment with a *REST* inhibitor can ameliorate deficits in social interaction in a valproic acid

mouse model of ASD (Kawase et al. 2019). Our findings further support that *REST* is a critical regulator of ASD-associated gene expression and suggest a general paradigm for identifying regulators of transcriptomic changes in ASD that may ultimately suggest new therapeutic targets. More broadly, while this work focused on ASD, the approaches used here are widely applicable to different cell types and conditions. Thus, this study serves as a framework for uncovering critical regulators across a wide diversity of contexts, from development through disease.

ACKNOWLEDGEMENTS

We thank members of the Walsh Lab for helpful discussions. We are also grateful to Dr. August Yue Huang for technical feedback regarding statistical analyses, and to Dr. Adrianna San Roman and Dr. Marla Tharp of the Page Lab for their insights regarding the role of *ZFX*. This work was facilitated by the BCH Flow Cytometry Research Core, the BCH IDDRC Molecular Genetics Core Facility, the BCH Cellular Imaging Core, and Lingsheng Dong of the HMS Research Computing Core. This project was supported by Autism Speaks Postdoctoral Fellowship 13008 and NIGMS T32 GM007748 (R.E.A.); T32 GM007753 and T32 GM144273 (M.T.); the Harvard Program for Research in Science and Engineering and the Harvard College Research Program (T.S.); a Howard Hughes Medical Institute Fellowship from the Helen Hay Whitney Foundation and Grant K99MH136290 from NIMH/NIH (J.H.T.S); K99 NS135123 from the NINDS/NIH (X.Q.); the Howard Hughes Medical Institute and Grant K99EY034603 from NEI/NIH (R.N.D); the Harvard-Amgen Scholars Program (G.E.); and grants from the Simons Foundation and Autism BrainNet (953759), NIMH (U01MH106883), the Hock E. Tan and K. Lisa Yang Center for Autism Research at Harvard University, and the Allen Discovery Center program, a Paul G. Allen Frontiers Group advised program of the Paul G. Allen Family Foundation (C.A.W). C.A.W. is an investigator of the Howard Hughes Medical Institute.

REFERENCES

Abrahams BS, Arking DE, Campbell DB, et al (2013) SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). Mol Autism 4:36. https://doi.org/10.1186/2040-2392-4-36

Alvarez MJ, Shen Y, Giorgi FM, et al (2016) Functional characterization of somatic mutations in cancer using network-based inference of protein activity. Nat Genet 48:838–47. https://doi.org/10.1038/ng.3593

Andersen RE, Alkuraya IF, Ajeesh A, et al (2024) Chromosomal structural rearrangements implicate long non-coding RNAs in rare germline disorders. Hum Genet 143:921–938. https://doi.org/10.1007/s00439-024-02693-y

Ang CE, Ma Q, Wapinski OL, et al (2019) The novel lncRNA lnc-NR2F1 is pro-neurogenic and mutated in human neurodevelopmental disorders. Elife 8:. https://doi.org/10.7554/eLife.41770

Antaki D, Guevara J, Maihofer AX, et al (2022) A phenotypic spectrum of autism is attributable to the combined effects of rare variants, polygenic risk and sex. Nat Genet 54:1284–1292. https://doi.org/10.1038/s41588-022-01064-5

Bai D, Yip BHK, Windham GC, et al (2019) Association of Genetic and Environmental Factors With Autism in a 5-Country Cohort. JAMA Psychiatry 76:1035. https://doi.org/10.1001/jamapsychiatry.2019.1411

Blanton L V, San Roman AK, Wood G, et al (2024) Stable and robust Xi and Y transcriptomes drive cell-type-specific autosomal and Xa responses in vivo and in vitro in four human cell types. Cell genomics 4:100628. https://doi.org/10.1016/j.xgen.2024.100628

Boukas L, Havrilla JM, Hickey PF, et al (2019) Coexpression patterns define epigenetic regulators associated with neurological dysfunction. Genome Res 29:532–542. https://doi.org/10.1101/gr.239442.118

Chakravarti A, Turner TN (2016) Revealing rate-limiting steps in complex disease biology: The crucial importance of studying rare, extreme-phenotype families. Bioessays 38:578–86. https://doi.org/10.1002/bies.201500203

Chatterjee S, Kapoor A, Akiyama JA, et al (2016) Enhancer Variants Synergistically Drive Dysfunction of a Gene Regulatory Network In Hirschsprung Disease. Cell 167:355-368.e10. https://doi.org/10.1016/j.cell.2016.09.005

Chong JA, Tapia-Ramírez J, Kim S, et al (1995) REST: a mammalian silencer protein that restricts sodium channel gene expression to neurons. Cell 80:949–57. https://doi.org/10.1016/0092-8674(95)90298-8

Connacher R, Williams M, Prem S, et al (2022) Autism NPCs from both idiopathic and CNV 16p11.2 deletion patients exhibit dysregulation of proliferation and mitogenic responses. Stem Cell Reports 17:1380–1394. https://doi.org/10.1016/j.stemcr.2022.04.019

Dias CM, Walsh CA (2020) Recent Advances in Understanding the Genetic Architecture of Autism. Annu Rev Genomics Hum Genet 21:289–304. https://doi.org/10.1146/annurev-genom-121219-082309

Dougherty JD, Marrus N, Maloney SE, et al (2022) Can the "female protective effect" liability threshold model explain sex differences in autism spectrum disorder? Neuron 110:3243–3262. https://doi.org/10.1016/j.neuron.2022.06.020

Gandal MJ, Haney JR, Wamsley B, et al (2022) Broad transcriptomic dysregulation occurs across the cerebral cortex in ASD. Nature 611:532–539. https://doi.org/10.1038/s41586-022-05377-7

Gao F, Cai Y, Kapranov P, Xu D (2020) Reverse-genetics studies of lncRNAs—what we have learnt and paths forward. Genome Biol 21:93. https://doi.org/10.1186/s13059-020-01994-5

Graham V, Khudyakov J, Ellis P, Pevny L (2003) SOX2 functions to maintain neural progenitor identity. Neuron 39:749–65. https://doi.org/10.1016/s0896-6273(03)00497-5

Grove J, Ripke S, Als TD, et al (2019) Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 51:431–444. https://doi.org/10.1038/s41588-019-0344-8

Haddad Derafshi B, Danko T, Chanda S, et al (2022) The autism risk factor CHD8 is a chromatin activator in human neurons and functionally dependent on the ERK-MAPK pathway effector ELK1. Sci Rep 12:22425. https://doi.org/10.1038/s41598-022-23614-x

Iakoucheva LM, Muotri AR, Sebat J (2019) Getting to the Cores of Autism. Cell 178:1287–1298. https://doi.org/10.1016/j.cell.2019.07.037

Jacquemont S, Coe BP, Hersch M, et al (2014) A higher mutational burden in females supports a "female protective model" in neurodevelopmental disorders. Am J Hum Genet 94:415–25. https://doi.org/10.1016/j.ajhg.2014.02.001

Jourdon A, Wu F, Mariani J, et al (2023) Modeling idiopathic autism in forebrain organoids reveals an imbalance of excitatory cortical neuron subtypes during early neurogenesis. Nat Neurosci 26:1505–1515. https://doi.org/10.1038/s41593-023-01399-0

Kalkman H (2012) A review of the evidence for the canonical Wnt pathway in autism spectrum disorders. Mol Autism 3:10. https://doi.org/10.1186/2040-2392-3-10

Katayama Y, Nishiyama M, Shoji H, et al (2016) CHD8 haploinsufficiency results in autistic-like phenotypes in mice. Nature 537:675–679. https://doi.org/10.1038/nature19357

Kawase H, Ago Y, Naito M, et al (2019) mS-11, a mimetic of the mSin3-binding helix in NRSF, ameliorates social interaction deficits in a prenatal valproic acid-induced autism mouse model. Pharmacol Biochem Behav 176:1–5. https://doi.org/10.1016/j.pbb.2018.11.003

Kissel LT, Pochareddy S, An J-Y, et al (2024) Sex-Differential Gene Expression in Developing Human Cortex and Its Intersection With Autism Risk Pathways. Biological psychiatry global open science 4:100321. https://doi.org/10.1016/j.bpsgos.2024.100321

Kochinke K, Zweier C, Nijhof B, et al (2016) Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. Am J Hum Genet 98:149–64. https://doi.org/10.1016/j.ajhg.2015.11.024

Konermann S, Lotfy P, Brideau NJ, et al (2018) Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. Cell 173:665-676.e14. https://doi.org/10.1016/j.cell.2018.02.033

Kostyanovskaya E, Lasser MC, Wang B, et al (2025) Convergence of autism proteins at the cilium. bioRxiv. https://doi.org/10.1101/2024.12.05.626924

Lachmann A, Giorgi FM, Lopez G, Califano A (2016) ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics 32:2233–2235. https://doi.org/10.1093/bioinformatics/btw216

Lambert SA, Jolma A, Campitelli LF, et al (2018) The Human Transcription Factors. Cell 172:650–665. https://doi.org/10.1016/J.CELL.2018.01.029

Lasser M, Sun N, Xu Y, et al (2023) Pleiotropy of autism-associated chromatin regulators. Development 150:dev201515. https://doi.org/10.1242/dev.201515

Lee KE, Seo J, Shin J, et al (2014) Positive feedback loop between Sox2 and Sox6 inhibits neuronal differentiation in the developing central nervous system. Proc Natl Acad Sci U S A 111:2794–9. https://doi.org/10.1073/pnas.1308758111

Lek M, Karczewski KJ, Minikel E V., et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536:285–291. https://doi.org/10.1038/nature19057

Li C, Fleck JS, Martins-Costa C, et al (2023) Single-cell brain organoid screening identifies developmental defects in autism. Nature 621:373–380. https://doi.org/10.1038/s41586-023-06473-y

Liao C, Moyses-Oliveira M, De Esch CEF, et al (2023) Convergent coexpression of autism-associated genes suggests some novel risk genes may not be detectable in large-scale genetic studies. Cell Genomics 3:100277. https://doi.org/10.1016/j.xgen.2023.100277

606 Loomes R, Hull L, Mandy WPL (2017) What Is the Male-to-Female Ratio in Autism Spectrum
607 Disorder? A Systematic Review and Meta-Analysis. J Am Acad Child Adolesc Psychiatry
608 56:466–474. https://doi.org/10.1016/j.jaac.2017.03.013
609 Luo S, Lu JY, Liu L, et al (2016) Divergent lncRNAs Regulate Gene Expression and Lineage
610 Differentiation in Pluripotent Cells. Cell Stem Cell 18:637–652.
611 https://doi.org/10.1016/j.stem.2016.01.024
612 Luo T, Liu P, Wang X-Y, et al (2018) Effect of the autism-associated lncRNA Shank2-AS on
613 architecture and growth of neurons. J Cell Biochem. https://doi.org/10.1002/jcb.27471
614 Maenner MJ, Warren Z, Williams AR, et al (2023) Prevalence and Characteristics of Autism
615 Spectrum Disorder Among Children Aged 8 Years - Autism and Developmental Disabilities
616 Monitoring Network, 11 Sites, United States, 2020. MMWR Surveill Summ 72:1–14.
617 https://doi.org/10.15585/mmwr.ss7202a1
618 Margolin AA, Nemenman I, Basso K, et al (2006) ARACNE: an algorithm for the reconstruction
619 of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics 7 Suppl
620 1:S7. https://doi.org/10.1186/1471-2105-7-S1-S7
621 Martin HC, Gardner EJ, Samocha KE, et al (2021) The contribution of X-linked coding variation
622 to severe developmental disorders. Nat Commun 12:627. https://doi.org/10.1038/s41467-
623 020-20852-3
624 Mattick JS, Amaral PP, Carninci P, et al (2023) Long non-coding RNAs: definitions, functions,
625 challenges and recommendations. Nat Rev Mol Cell Biol 24:430–447.
626 https://doi.org/10.1038/s41580-022-00566-8
627 Mendes M, Chen DZ, Engchuan W, et al (2024) Chromosome X-wide common variant
628 association study in autism spectrum disorder. Am J Hum Genet 0:
629 https://doi.org/10.1016/j.ajhg.2024.11.008
630 Napolitano A, Schiavi S, La Rosa P, et al (2022) Sex Differences in Autism Spectrum Disorder:
631 Diagnostic, Neurobiological, and Behavioral Features. Front Psychiatry 13:889636.
632 https://doi.org/10.3389/fpsyt.2022.889636
633 Noor A, Whibley A, Marshall CR, et al (2010) Disruption at the PTCHD1 Locus on Xp22.11 in
634 Autism spectrum disorder and intellectual disability. Sci Transl Med 2:49ra68.
635 https://doi.org/10.1126/scitranslmed.3001267
636 Obradovic A, Chowdhury N, Haake SM, et al (2021) Single-cell protein activity analysis
637 identifies recurrence-associated renal tumor macrophages. Cell 184:2988-3005.e16.
638 https://doi.org/10.1016/J.CELL.2021.04.038
639 Oliva M, Muñoz-Aguirre M, Kim-Hellmuth S, et al (2020) The impact of sex on gene expression
640 across human tissues. Science 369:. https://doi.org/10.1126/science.aba3066
641 Page L, Brin S, Motwani R, Winograd T (1999) The PageRank Citation Ranking : Bringing Order
642 to the Web. The Web Conference
643 Parikshak NN, Swarup V, Belgard TG, et al (2016) Genome-wide changes in lncRNA, splicing,
644 and regional gene expression patterns in autism. Nature 540:423–427.
645 https://doi.org/10.1038/nature20612
646 Paulsen B, Velasco S, Kedaigle AJ, et al (2022) Autism genes converge on asynchronous
647 development of shared neuron classes. Nature 602:268–273.
648 https://doi.org/10.1038/s41586-021-04358-6
649 Pinto D, Delaby E, Merico D, et al (2014) Convergence of genes and cellular pathways
650 dysregulated in autism spectrum disorders. Am J Hum Genet 94:677–94.
651 https://doi.org/10.1016/j.ajhg.2014.03.018
652 Price E, Fedida LM, Pugacheva EM, et al (2023) An updated catalog of CTCF variants
653 associated with neurodevelopmental disorder phenotypes. Front Mol Neurosci 16:1185796.
654 https://doi.org/10.3389/fnmol.2023.1185796

Ramaswami G, Won H, Gandal MJ, et al (2020) Integrative genomics identifies a convergent molecular subtype that links epigenomic with transcriptomic differences in autism. Nat Commun 11:4873. https://doi.org/10.1038/s41467-020-18526-1

Raznahan A, Parikshak NN, Chandran V, et al (2018) Sex-chromosome dosage effects on gene expression in humans. Proc Natl Acad Sci U S A 115:7398–7403. https://doi.org/10.1073/pnas.1802889115

Rhie SK, Yao L, Luo Z, et al (2018) ZFX acts as a transcriptional activator in multiple types of human tumors by binding downstream from transcription start sites at the majority of CpG island promoters. Genome Res 28:310–320. https://doi.org/10.1101/gr.228809.117

Rockowitz S, Zheng D (2015) Significant expansion of the REST/NRSF cistrome in human versus mouse embryonic stem cells: potential implications for neural development. Nucleic Acids Res 43:5730–43. https://doi.org/10.1093/nar/gkv514

Roopra A (2020) MAGIC: A tool for predicting transcription factors and cofactors driving gene sets using ENCODE data. PLoS Comput Biol 16:e1007800. https://doi.org/10.1371/journal.pcbi.1007800

Ropers H-H, Hamel BCJ (2005) X-linked mental retardation. Nat Rev Genet 6:46–57. https://doi.org/10.1038/nrg1501

San Roman AK, Godfrey AK, Skaletsky H, et al (2023) The human inactive X chromosome modulates expression of the active X chromosome. Cell genomics 3:100259. https://doi.org/10.1016/j.xgen.2023.100259

San Roman AK, Skaletsky H, Godfrey AK, et al (2024) The human Y and inactive X chromosomes similarly modulate autosomal gene expression. Cell genomics 4:100462. https://doi.org/10.1016/j.xgen.2023.100462

Satterstrom FK, Kosmicki JA, Wang J, et al (2020) Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism Article Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. Cell 1–17. https://doi.org/10.1016/j.cell.2019.12.036

Schnütgen F, Doerflinger N, Calléja C, et al (2003) A directional strategy for monitoring Cre-mediated recombination at the cellular level in the mouse. Nat Biotechnol 21:562–5. https://doi.org/10.1038/nbt811

Schoenherr CJ, Anderson DJ (1995) The neuron-restrictive silencer factor (NRSF): a coordinate repressor of multiple neuron-specific genes. Science 267:1360–3. https://doi.org/10.1126/science.7871435

Shepherdson JL, Hutchison K, Don DW, et al (2024) Variants in ZFX are associated with an X-linked neurodevelopmental disorder with recurrent facial gestalt. Am J Hum Genet 111:487–508. https://doi.org/10.1016/j.ajhg.2024.01.007

Subramanian A, Tamayo P, Mootha VK, 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. https://doi.org/10.1073/pnas.0506580102

Sugathan A, Biagioli M, Golzio C, et al (2014) CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. Proc Natl Acad Sci U S A 111:E4468-77. https://doi.org/10.1073/pnas.1405266111

Sun N, Teyssier N, Wang B, et al (2024) Autism genes converge on microtubule biology and RNA-binding proteins during excitatory neurogenesis. bioRxiv. https://doi.org/10.1101/2023.12.22.573108

Szklarczyk D, Kirsch R, Koutrouli M, et al (2023) The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res 51:D638–D646. https://doi.org/10.1093/nar/gkac1000

Trost B, Thiruvahindrapuram B, Chan AJS, et al (2022) Genomic architecture of autism from comprehensive whole-genome sequence annotation. Cell 185:4409-4427.e18. https://doi.org/10.1016/j.cell.2022.10.009

Valencia AM, Sankar A, van der Sluijs PJ, et al (2023) Landscape of mSWI/SNF chromatin remodeling complex perturbations in neurodevelopmental disorders. Nat Genet 55:1400–1412. https://doi.org/10.1038/s41588-023-01451-6

Velmeshev D, Perez Y, Yan Z, et al (2023) Single-cell analysis of prenatal and postnatal human cortical development. Science (1979) 382:. https://doi.org/10.1126/science.adf0834

Villegas VE, Zaphiropoulos PG (2015) Neighboring gene regulation by antisense long non-coding RNAs. Int J Mol Sci 16:3251–66. https://doi.org/10.3390/ijms16023251

Vlahos L, Obradovic A, Worley J, et al (2021) Systematic, Protein Activity-based Characterization of Single Cell State. bioRxiv 2021.05.20.445002

Voineagu I, Wang X, Johnston P, et al (2011) Transcriptomic analysis of autistic brain reveals convergent molecular pathology. Nature 474:380–4. https://doi.org/10.1038/nature10110

Wang Y, Chen S, Li W, et al (2020) Associating divergent lncRNAs with target genes by integrating genome sequence, gene expression and chromatin accessibility data. NAR Genom Bioinform 2:lqaa019. https://doi.org/10.1093/nargab/lqaa019

Warrier V, Zhang X, Reed P, et al (2022) Genetic correlates of phenotypic heterogeneity in autism. Nat Genet 54:1293–1304. https://doi.org/10.1038/s41588-022-01072-5

Wei J, Lotfy P, Faizi K, et al (2023) Deep learning and CRISPR-Cas13d ortholog discovery for optimized RNA targeting. Cell Syst 14:1087-1102.e13. https://doi.org/10.1016/j.cels.2023.11.006

Weinschutz Mendes H, Neelakantan U, Liu Y, et al (2023) High-throughput functional analysis of autism genes in zebrafish identifies convergence in dopaminergic and neuroimmune pathways. Cell Rep 42:112243. https://doi.org/10.1016/j.celrep.2023.112243

Wenderski W, Wang L, Krokhotin A, et al (2020) Loss of the neural-specific BAF subunit ACTL6B relieves repression of early response genes and causes recessive autism. Proc Natl Acad Sci U S A 117:10055–10066. https://doi.org/10.1073/pnas.1908238117

Werling DM (2016) The role of sex-differential biology in risk for autism spectrum disorder. Biol Sex Differ 7:58. https://doi.org/10.1186/s13293-016-0112-8

Werling DM, Parikshak NN, Geschwind DH (2016) Gene expression in human brain implicates sexually dimorphic pathways in autism spectrum disorders. Nat Commun 7:10717. https://doi.org/10.1038/ncomms10717

Wigdor EM, Weiner DJ, Grove J, et al (2022) The female protective effect against autism spectrum disorder. Cell genomics 2:100134. https://doi.org/10.1016/j.xgen.2022.100134

Willsey HR, Willsey AJ, Wang B, State MW (2022) Genomics, convergent neuroscience and progress in understanding autism spectrum disorder. Nat Rev Neurosci 23:323–341. https://doi.org/10.1038/s41583-022-00576-7

Zhang Y, Li N, Li C, et al (2020) Genetic evidence of gender difference in autism spectrum disorder supports the female-protective effect. Transl Psychiatry 10:4. https://doi.org/10.1038/s41398-020-0699-8

Zhao S, Zhang X, Chen S, Zhang S (2021) Long noncoding RNAs: fine-tuners hidden in the cancer signaling network. Cell Death Discov 7:283. https://doi.org/10.1038/s41420-021-00678-8

Ziats MN, Rennert OM (2013) Aberrant expression of long noncoding RNAs in autistic brain. Journal of Molecular Neuroscience 49:589–593. https://doi.org/10.1007/s12031-012-9880-8

(2013) Diagnostic and statistical manual of mental disorders: DSM-5™, 5th ed. American Psychiatric Publishing, Inc., Arlington, VA, US

FIGURE LEGENDS

Figure 1

A) Schematic of the major classes of genetic variants within *Consensus-ASD*. **B)** Experimental design schematic. **C)** Immunocytochemistry of iPSC-derived NPC cultures. **D)** Relative expression of the targeted gene within each KD sample compared to non-targeting control (NTC) samples. Only targets with statistically significant KD are shown (see **Table S4**). Error bars: 95% confidence interval. KD target genes that have been directly implicated in ASD are labeled in bold text. **E)** Proportion of lncRNAs significantly differentially expressed upon KD of their neighboring PCG, and proportion of PCGs significantly differentially expressed upon KD of their neighboring lncRNA.

Figure 2

A) Bar plot depicting enrichment of *Consensus-ASD* genes within the DEGs of each KD condition. **B)** Circos plot depicting HC-SFARI DEGs shared between different KD conditions. **C)** Heatmap showing the relative expression of HC-SFARI genes that are DEGs in at least 5 KD conditions (see also **Fig. S5B**). Statistically significant DEGs are indicated with *. **D)** MAGIC analysis identifying transcriptional regulators whose targets are enriched within the DEGs from each KD (see also **Fig. S5C**). **E)** Bar plot depicting enrichment of REST targets within SFARI, *Consensus-ASD*, or SysID genes in human and mouse embryonic stem cells. REST target data from Rockowitz et al., 2015. Throughout the figure, KD target genes that have been directly implicated in ASD are labeled in bold text.

Figure 3

A) Volcano plot depicting differential expression upon KD of SETD5. Statistically significant DEGs with decreased expression are shown in blue, while DEGs with increased expression are shown in red. Genes within the 1st and 2nd degree SETD5 regulon, derived from an ARACNe

GRN constructed with SETD5 samples held out, are highlighted in pink. Genes within the 1st degree regulon are depicted with dark borders. Statistically significant enrichment of 1st and 2nd degree regulon genes within the downregulated DEGs is indicated with the adjusted p-value. **B)** Depiction of a subset of the NPC GRN constructed using ARACNE. All regulators are shown along with first degree target genes that are in *Consensus-ASD* or the HC-SFARI gene set. **C)** Gene set enrichment analysis (GSEA) demonstrating enrichment of genes ordered by PageRank score within several gene sets. **D)** Depiction of modules of top regulators identified through iterative walktrap clustering. Genes of particular interest are highlighted in red. **E)** Bar plot depicting enrichment of CHD8 targets within several gene categories. Data from CHD8 ChIP-seq in NPCs from Sugathan et al., 2014.

Figure 4

A) MAGIC analysis identifying X-linked transcriptional regulators whose targets are enriched within the DEGs from the NPC KDs. **B)** Binary heatmap of GRN-identified, X-linked regulators whose regulons are enriched for candidate ASD risk genes. Only target genes that are members of at least five of these aforementioned regulons are depicted. **C)** Network visualization of GRN-identified, X-linked regulators of candidate autism risk genes and their target genes. Edges directed to the top 20% most frequently targeted genes are depicted in black, and edges connecting two regulators are further bolded. All other edges are depicted in gray. Nodes representing SFARI genes are outlined in black. **D)** Heatmap showing the relative expression of *ZFX* upon KD of the indicated target genes. Statistically significant differential expression is indicated with *. **E)** Enrichment of gene sets relevant to autism and the female protective effect within *ZFX* targets identified from ChIP-seq and CRISPR KD experiments in San Roman et al., 2024. **F)** Depiction of the regulatory relationships between *ZFX* and target genes that are top drivers of transcriptomic convergence in ASD (previously identified in **Fig.**

3D). Throughout the figure, KD target genes that have been directly implicated in ASD are labeled in bold text.

Figure 5

A) Schematic of experimental design to analyze proliferation phenotypes using CytoTrack. **B)** Relative CytoTrack intensity over time for each KD condition that exhibited statistically significant proliferation phenotypes that were consistent over time. Error bars indicate standard error of the mean. Statistical significance was determined using 2-way ANOVA. See **Table S8** for full statistical results. KD target genes that have been directly implicated in ASD are labeled in bold text. **C)** Design of FLICK constructs for doxycycline-inducible Cas13 KD. **D)** Immunohistochemistry of FLICK organoid at day 60. EGFP = FLICK construct marker; PAX6 = NPC marker; TBR2 = intermediate progenitor cell marker; CTIP2 = excitatory neuron marker. **E)** UMAP of integrated organoid scRNA-seq data collected at d30 and d60. **F)** Expression of canonical brain marker genes used for identifying cell types.

Figure 6

A) Longitudinal analysis of the proportion of cells for each KD target in mosaic organoids in which the FLICK constructs were induced by treatment with doxycycline on d14-d15. KD targets that appear depleted over time are highlighted with arrows. **B)** Cell type composition per KD target. KD targets with fewer progenitors and an increased relative proportion of excitatory neurons are highlighted with arrows. **C)** Heatmap of cross-enrichment of DEGs for radial glia and excitatory neurons. **D)** Network visualization of DEGs amongst different KDs. The thickness of edges represents the number of shared DEGs. Nodes plotted based on clustering with the Fruchterman-Reingold algorithm (*igraph*). Nodes that are closer together have more shared DEGs on average. **E)** MAGIC analysis of DEGs from excitatory neurons (ExNs) in one-month

organoids (see also **Fig. S15**). Throughout the figure, KD target genes that have been directly implicated in ASD are labeled in bold text.

Figure 7

A) Proportion of cells with each KD target from scRNA-seq of mosaic organoids with acute KD. **B)** UMAP of acute KD mosaic organoids from scRNA-seq. **C)** Bar plot of cross-enrichment of DEGs from excitatory neurons (ExNs) from acute KD mosaic organoids versus DEGs from previous NPC cultures upon KD of the same target gene. **D)** Heatmap of cross-enrichment between different KD samples in ExNs. **E)** MAGIC analysis of ExN DEGs. Throughout the figure, KD target genes that have been directly implicated in ASD are labeled in bold text.

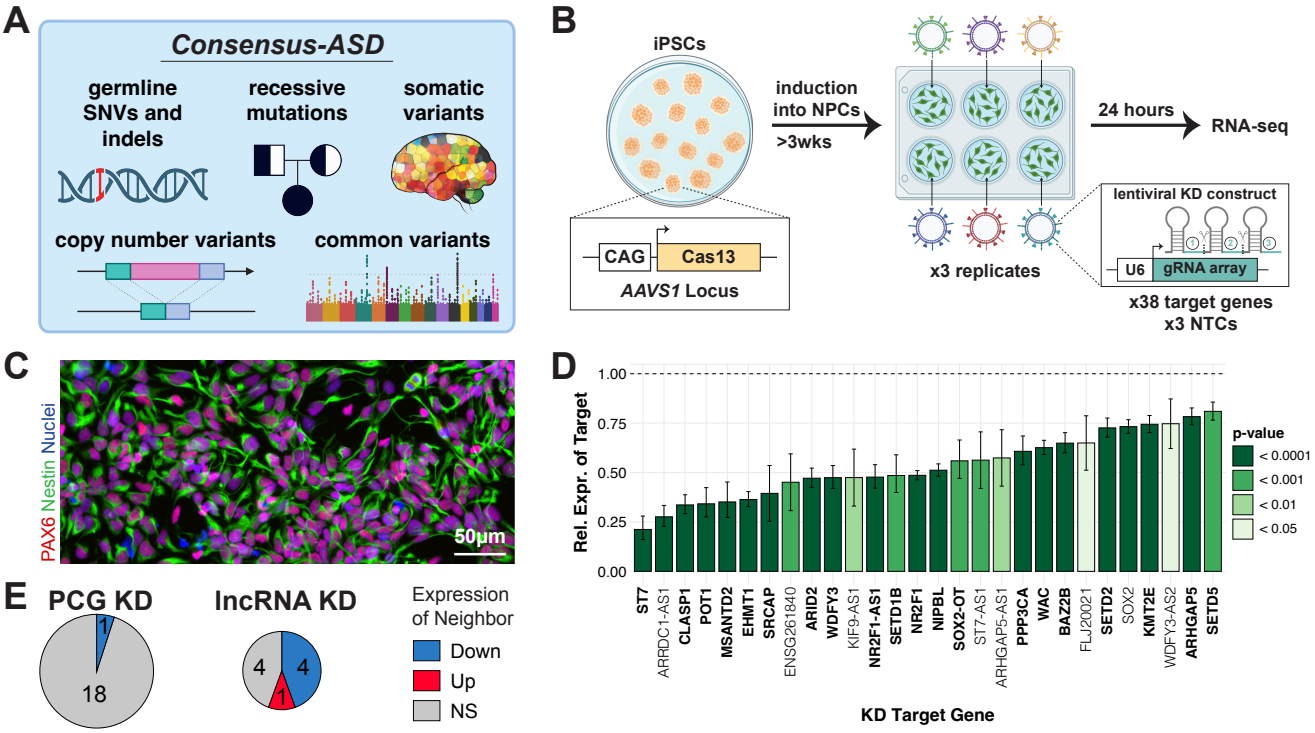

Figure 1

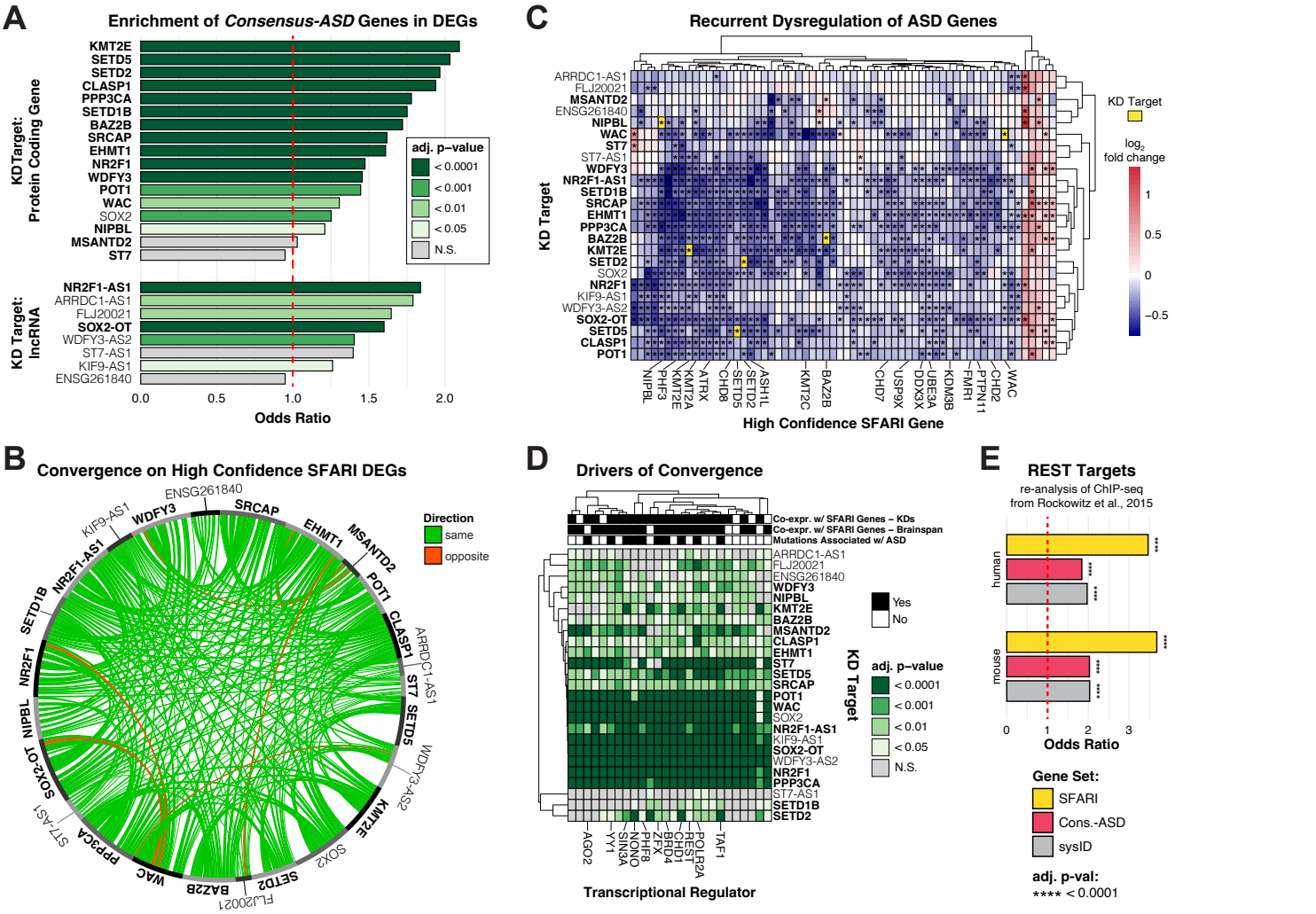

Figure 2

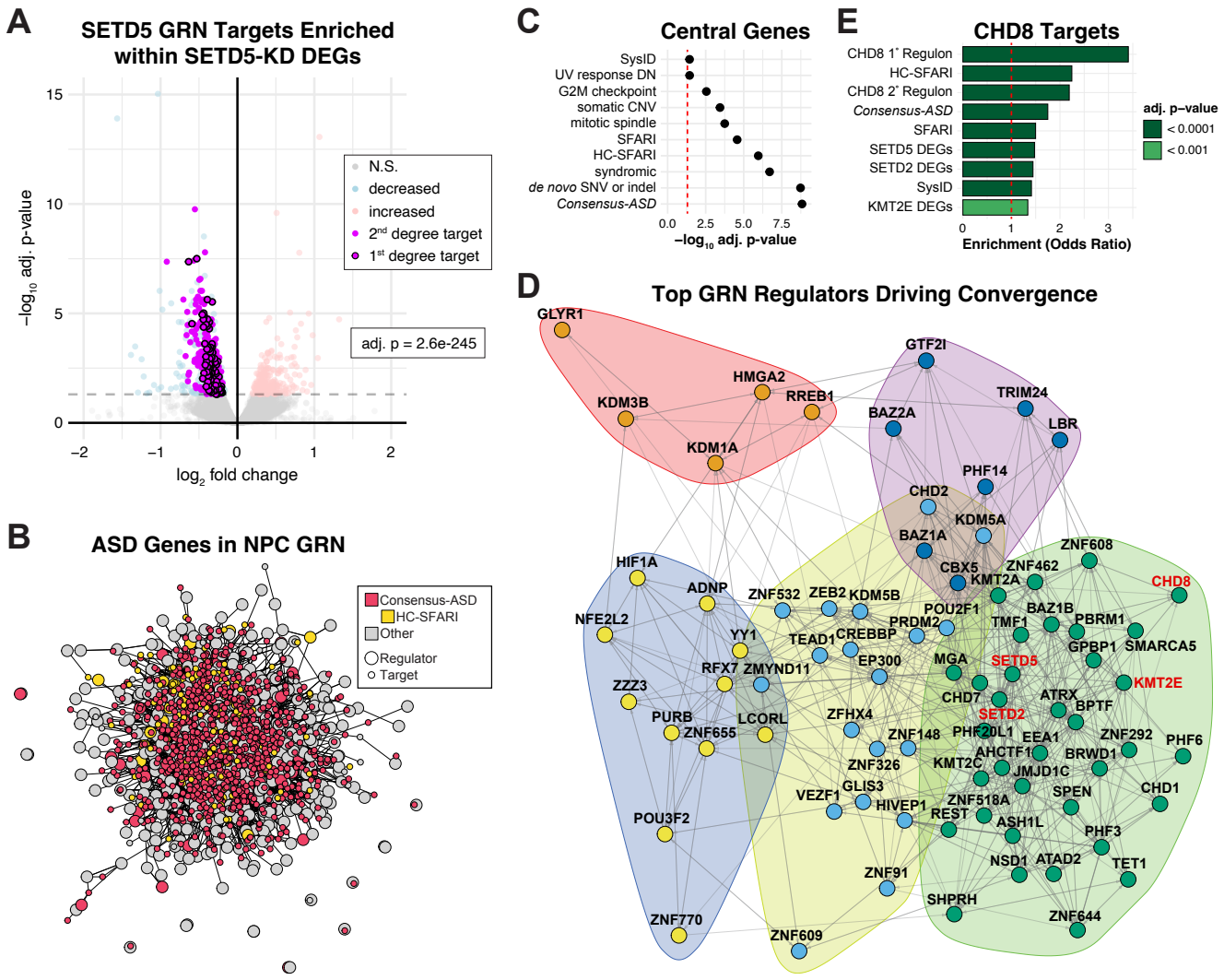

Figure 3

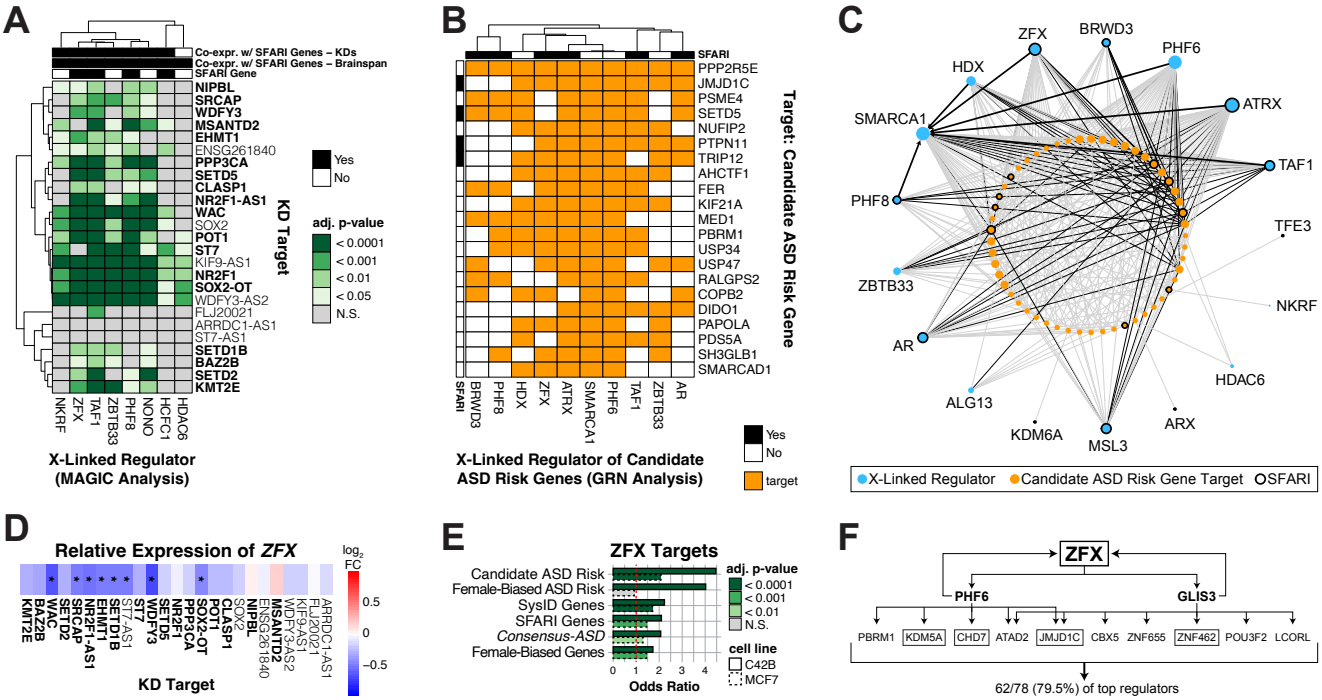

Figure 4

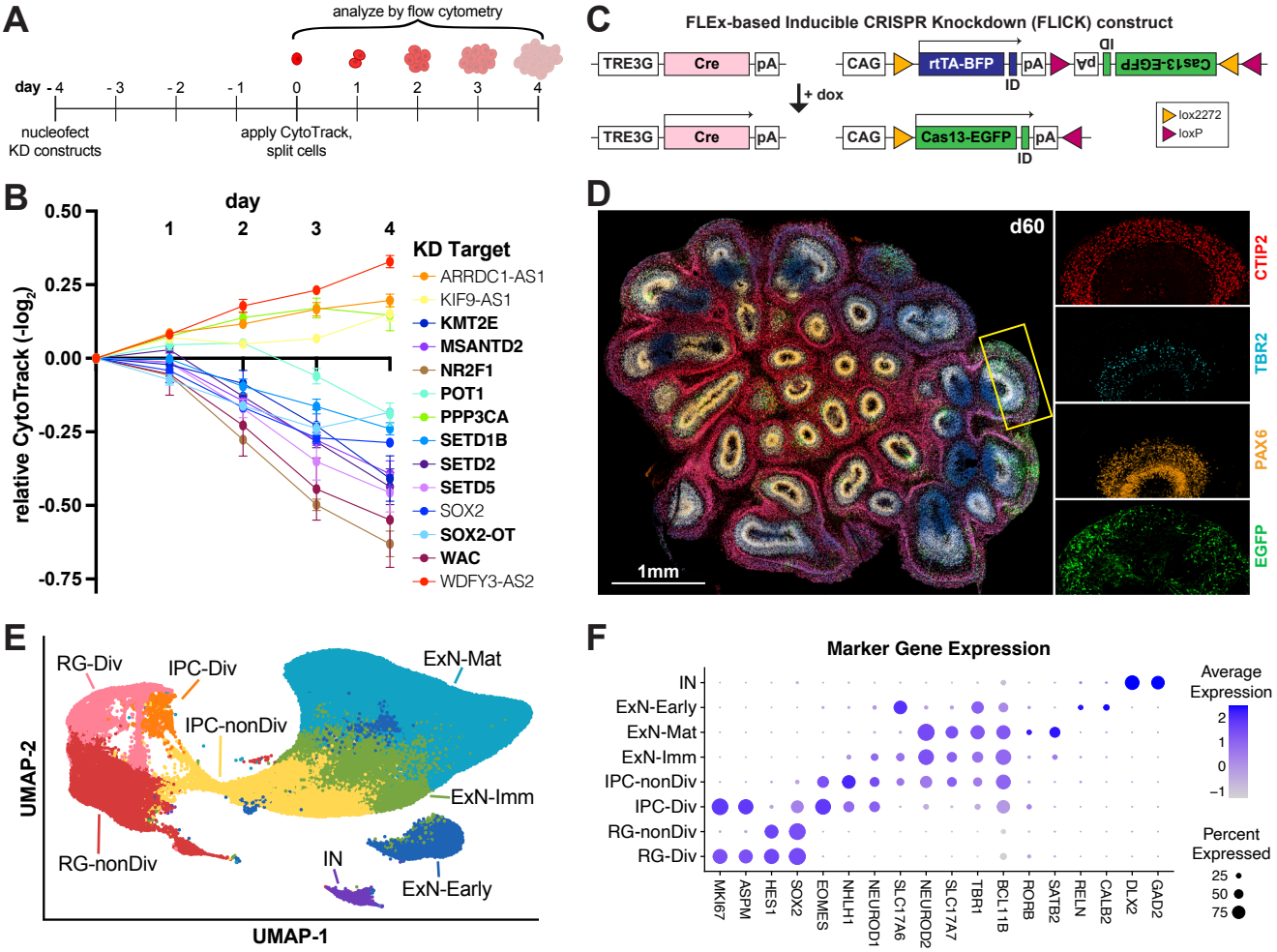

Figure 5

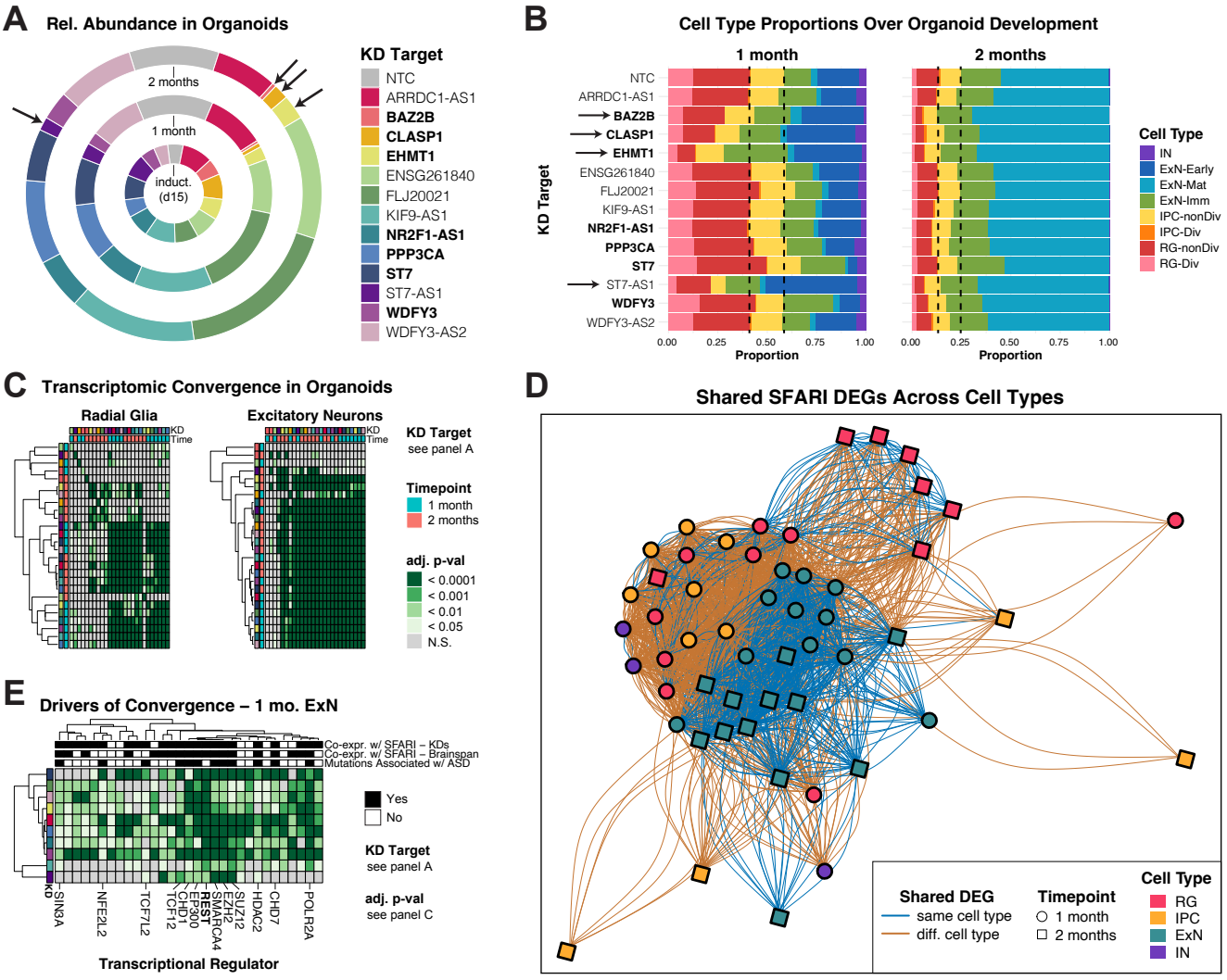

Figure 6

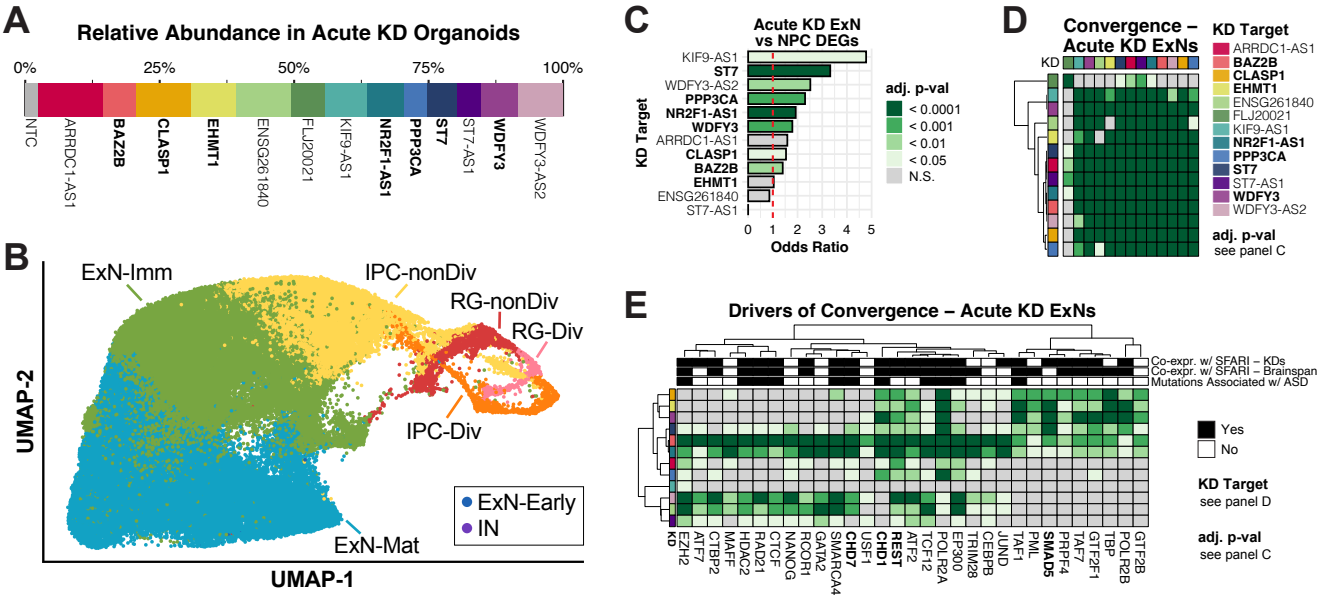

Figure 7