ARX regulates interneuron subtype differentiation and migration

Ginam Cho[1,*], Youngshin Lim[1], Shyam K Akula[3, 4, 5], Abigail K Myers[2], Connie Chen[1], Katherine A Rafael[1], Christopher A Walsh[3, 4, 5], and Jeffrey A Golden[1,*]

[1]Department of Pathology and Laboratory Medicine, Cedars-Sinai Medical Center, Los Angeles, CA 90048; [2]Hamilton College, Clinton, NY 13323; [3]Division of Genetics and Genomics, Manton Center for Orphan Disease Research and [4]Howard Hughes Medical Institute, Boston Children's Hospital, Boston, MA 02115; [5]Department of Pediatrics and Department of Neurology; Harvard-Massachusetts Institute of Technology MD/PhD Program, Program in Neuroscience, Harvard Medical School, Boston, MA 02115

* Correspondences: Ginam Cho (Ginam.Cho@cshs.org) and Jeffrey A Golden (Jeffrey.Golden@cshs.org)

Abstract

Mutations in the aristaless-related homeobox (*ARX*) gene are associated with a spectrum of neurodevelopmental disorders with neurologic phenotypes including epilepsy, intellectual disabilities and autism spectrum disorders, with or without structural malformations. Some features of these disorders have been linked to abnormal cortical interneuron (IN) development and function. Currently, more than 50 subtypes of cortical INs have been identified thanks to the single cell transcriptomics. The existence of these many IN subtypes, with differing functions, raise the likelihood that distinct phenotypic dysfunction be associated with specific IN subtype defects. Despite the functional significance of each IN subtype, the molecular mechanisms underlying their development remain poorly understood. To further extend our understanding of IN subtype development and evaluate the implications of subtype defects, different types of *Arx* mouse models were investigated: conditional knock-out (Arx cKO) driven by four IN-specific Cre lines (Gad2Cre, Nkx2.1Cre, SstCre, Dlx5/6Cre) and a poly-alanine tract expansion mutation ($Arx^{(GCG)7}$, also known as Arx^{PA1-22}). Single cell as well as total transcriptomic analysis, along with ARX chromatin immunoprecipitation sequencing (ChIP-seq), reveal that ARX regulates transcription of cortical IN development-associated genes. Our data obtained in embryonic and perinatal stage indicate that the relatively early loss of ARX in the progenitor cells or immature INs leads to IN subtype changes and migration defects, whereas its loss in postmitotic INs did not. On the other hand, the poly-alanine tract expansion mutation results in IN migration defects without affecting IN subtype development during embryonic stages, although postnatal reduction of certain subtypes of INs were detected. Our data provide new insights into how different mutations in the same gene can result in a spectrum of clinical phenotypes.

Introduction

The mammalian cerebral cortex is composed of excitatory projection neurons (PNs) and inhibitory interneurons (INs) neurons that together organize into functional cortical circuits [1–4]. PNs, accounting for 70-80% of all neocortical neurons[5] release glutamate and mediate communication between different brain regions by connecting with local networks as well as with distant targets. On the other hand, INs, constituting the remaining 20-30% neurons, release GABA and primarily modulate the timing, shape and coordination of neural circuits by establishing local connections with PNs and other Ins[4]. In mice, these two classes of neurons are derived from distinct progenitor populations. Cortical PNs are derived from the progenitors in the ventricular zone (VZ) of the dorsal telencephalon and radially migrate out to the cortical plate, which is relatively in short distance from their birthplace. In contrast, cortical INs produced from the ganglionic eminences (medial and caudal) and preoptic area which are located in the ventral telencephalon, take a long journey to the dorsal telencephalon and populate the neocortex. In human, some cortical INs seem to be derived from the dorsal telencephalon as well[6]. Imbalance in the excitation to inhibition ratio is thought to underlie the phenotypic expression of some neurological and neuropsychiatric disorders, such as autism spectrum disorder (ASD), epilepsy, and schizophrenia[7]. Therefore, studying the molecular mechanisms of the generation and development of these neurons, is of great importance in understanding and tackling these disorders.

Classically, cortical INs have been divided into different groups based on their morphologies, chemical signatures, and electrical firing properties. Three major classes of INs include calcium binding protein parvalbumin (PV)-, the neuropeptide somatostatin (SST)-, and serotonin receptor (5HTR3a)- expressing neurons [1–4]. These major IN groups can be further broken down into subclasses based on various factors such as synaptic specificity and expression of other markers[4]. The most abundant cINs (accounting for ~40% INs) expressing PV (PV INs), for instance, can be divided into 'basket cells' targeting the soma of excitatory cells, and 'chandelier cells' innervating the axon initial segment. The second most abundant SST-expressing cINs (accounting for ~30% INs) consist of two main types: 'Martinotti cells' that inhibit the distal dendrites of PNs or target other INs, and non-Martinotti cells that are present all cortical layers and extend their axons locally[4,8]. Thirdly, 5HTR3a INs (accounting ~30% INs) constitutes a

diverse group of cell types with different laminar distribution, morphologies, functional properties, and connectivity patterns, e.g., vasointestinal peptide (VIP) expressing bipolar cells [4,8]. Furthermore, recent advances in single cell transcriptomics led to the identification of more than 50 subtypes of INs in mouse visual cortex based on transcriptomic signature and unbiased statistical classification[9]. How these diverse groups of INs are generated and differentiated into different subtypes as well as how their migration is molecularly regulated remain incompletely understood.

Transcription factors play key roles in spatial and temporal specification of cortical INs. For example, the transcription factor Nkx2-1 (expressed in the VZ of the embryonic GE) is essential for the specification of SST and PV INs in the MGE[2,10]. It is also required for the initiation of the transcription factor Lhx6 expression, which is critical for the maturation of SST and PV INs[11,12]. Subsequent downregulation of Nkx2-1 is required for the migration of MGE-derived INs toward the cortex[13]. Moreover, transcription factors c-Maf and Maf-b (enriched in MGE lineages, from late-secondary progenitors in the SVZ of the embryonic MGE to mature PV and SST INs) control the numbers of MGE-lineage cINs, suppress SST cell identity and promote PV fate[14,15]. Overall, the establishment of IN subtypes is a complex process that involves the coordinated spatial and temporal expression of multiple transcription factors.

To further understand the relationship between transcriptional networks functioning in IN development, and the functional consequences of human disorder related mutations, we have investigated ARX, a transcription factor implicated in IN development and linked to human neurodevelopmental disorders[16], using two murine Arx mutant lines: the first, a floxed allele permitting conditional gene abrogation (cKO)[17] and the second, an alanine expansion allele in the first poly-Ala tract ($Arx^{(GCG)7}$)[18]. Defects in both IN subtype generation and migration were identified in both models. In the Arx-cKO we found ARX regulates Maf and Mef2c by binding to their genomic regulatory domain and the loss of ARX results in the overproduced Sst+ IN during embryonic development. The mutant showed a migration defect in marginal zone migration associated with the reduced expression of Cxcr4 that is regulated by LMO1 which is normally repressed by ARX. In contrast, the brain of $Arx^{(GCG)7}$ exhibit reduction of INs without subtype change and a similar migration defect, though mild, seen in cKO. Together, these findings provide new insights into the molecular mechanisms underlying IN development, the

roles ARX plays in IN migration and subtype specification, and ultimately how different mutations can explain the spectrum of overlapping phenotypes found in patients.

Materials and methods

Mouse

All procedures and animal care were approved and performed in accordance with the Cedars-Sinai Medical Center Laboratory Animal Research Center guidelines (IACUC). All mice strains have been previously published: Ai14 Rosa 26-tdTomato Cre-reporter[19], Gad2-Cre[20], Nkx2.1-Cre, Sst-IRES-Cre[20], Dlx5/6-Cre-IRES-GFP[21], Arx flox[17], and $Arx^{(GCG)7}$(also referred to Arx^{PA1-22})[18]. Mice were maintained in a C57BL/6J background. For timed pregnancies, noon on the day of the vaginal plug was counted as embryonic day 0.5.

Single-cell sample

E16.5 brains were harvested and dissociated with papain solution (10xgenomics). The red-positive cells were collected by cell sorter (BD FACSymphony S6) and loaded onto a 10X genomics single-cell preparation platform following the manufacturer's instruction (10xgenomics) to target 10,000 cell recovery per sample (2 wild types and 3 mutants). Library preparation was performed using 10X Chromium Single 3' Reagent Kits Version 2 (10xgenomics). The library length and concentration were quantified using a Bioanalyzer (Agilent) according to the manufacturer's protocol. Paired-end sequencing was performed on the Illumina NovaSeq.

Microarray analysis and NGS transcriptome data analysis

cDNA was amplified from extracted RNA using the QIAGE RNAeasy kit after purification of GFP-positive interneuron from the forebrain of E14.5 of wild type and PA1-22(Ala mutant). For each microarray, 4 μg of cDNA was labeled and hybridized to an Affymetrix Mouse Whole Genome 430 2.0 array, with three independent replicates per condition. The Bioconductor software package was used to background-correct and normalize the data and calculate mean fold change (FC) and false discovery rates (FDR) for various pair-wise comparisons.

Raw base call (BCL) files were converted into FASTQ files and demultiplexed to align cells with their own barcodes using Cell Ranger V.6 (10xgenomics). Transcriptome alignment

(mm10), cell filtering, barcode counting, UMI counting as well as matrix data output for downstream clustering and differential expression analyses were also performed using Scanpy[22].

RNA scope

The cryo-tissue (E14.5) or paraffine section(P0) was prepared as below. The day before the experiment, tissues were defrosted and based at 60ºC for an hour for better attachment, followed by PFA post-fixation at 4 ºC overnight. The *in situ* hybridization was performed following ACD protocol for the multiplex fluorescent kit (version 2). The RNAscope probes were Mm-Maf (Cat. 412958) and Mm-*Mef2c* (Cat. 421011). Slides were mounted with prolong glass mounting media (Invitrogen, P36984) containing DAPI. We used multiple sections from 2 WT and 2 mutant animals for validation and quantification to double-confirm the experiment.

Immunostaining

All P0 and P30 animals were sequentially perfused with PBS containing heparin, 4% PFA, followed by 4% PFA overnight fixation. E14.5 brains were dissected and fixed 4% PFA overnight. E14.5 and P0 tissues were transferred to 30% sucrose the next day for cryoprotection. Brains were embedded with OCT, and tissues were sectioned coronally at 14 um. Sections were all mounted onto glass slides before staining or ISH procedures with RNAscope (ACDbio). All immunofluorescence staining was performed with TBS containing 0.05% Tween20. All primary and secondary antibodies were diluted in a non-animal blocking solution (Vector lab). Primary antibodies were incubated with tissue overnight, while biotinylated or florescence-conjugated secondary antibodies were incubated for 1 hr at room temperature (RT). Immunofluorescent labeling was performed with the following primary antibodies: rabbit anti-RFP (Thermofisher), rabbit-Cxcr4 (Thermofisher), and rat anti-somatostatin (Millipore). The appropriate 488, 594 or 647 Alexa-conjugated secondary antibodies (1:250) were from Thermofisher. After the staining, sections were cover slipped with Prolog glass (Thermofisher) containing DAPI (Vector labs) In case of DAB staining, the ABC kit (Vector lab) and detection system were used.

Tissue clearing and counting cells.

All P0 and P30 animals were sequentially perfused with PBS containing heparin and 4% PFA, followed by 4% PFA overnight fixation. The tissue clearing step was performed by using the SHIELD protocol by LifeCanvas Technologies[23]. Samples were cleared for 7 days with Clear+ delipidation buffer and then immunolabeled using SmartLabel. Each sample was labeled with 5µg mouse anti-NeuN (RRID:AB_2572267) and 4µg rabbit anti-parvalbumin (Swant PV27; 1:400) and rat anti-somatostatin (Millipore) followed by fluorescent secondary antibodies. Samples were incubated in EasyIndex for a refractive index of 1.52 and imaged at 3.6X using SmartSPIM microscopy. Images were tile-corrected, de-striped, and registered to the Allen Brain Atlas (https://portal.brain-mapping.org). The NeuN channel was registered to the 8-20 atlas-aligned reference samples using successive rigid, affine, and b-spline warping (SimpleElastix: https://simpleelastix.github.io). Average alignment to the atlas was generated across all intermediate reference sample alignments to serve as the final atlas alignment value per sample. Fluorescent measurements from the acquired images were projected onto the Allen Brain Atlas to quantify the total fluorescence intensity per region defined by the Allen Brain Atlas. These values were then divided by the volume of the corresponding regional volume to calculate the intensity per voxel measurements. Data was plotted using Prism (version 10).

Quantification and statistical analyses

All bar graphs were shown as mean ± SEM. All statistical analyses were performed using Graphpad Prism (version 10) and a p value of < 0.05 was considered significant. The specific n for each experiment as well as the post-hoc test and corrected p-values, can be found in the Results section, in the Figure legends or in supplementary tables.

Results

Early embryonic loss of Arx in the subventricular zone of the GE results in abnormal numbers and distribution of cortical interneurons.

To study the role of ARX in interneuron (IN) development, we took advantage of Gad2-Cre mouse line, a pan-IN specific driver line where CRE activity turns on in Gad2^{+} cells in the subventricular zone (SVZ) of the ventral telencephalon from embryonic day (E)10.5 and on [20]. We first generated Arx-flox mice[17] in Ai14 background [19], which expresses tdTomato (red fluorescent protein) only in the cells with the history of CRE expression (i.e., CRE-lineage cells), and then crossed them with Gad2Cre mice to conditionally abrogate Arx in Gad2-lineage INs and permanently label these Arx-deficient INs with tdTomato. These conditional knockout (cKO) mice (hereafter referred to as 'Arx$^{F/Y}$; Ai14; Gad2Cre or simply Gad2Cre-Arx cKO; Ai14) die shortly after birth as seen in the germline mutant mice (the reason of death not certain)[16]. When checked while alive at E14.5 or postnatal day (P) 0, no changes in overall brain size or gross morphology were detected, unlike the germline mutant mice (data not shown)[16].

Despite no changes in overall brain size, however, the number of cortical INs (cINs) were reduced in these cKO mice (Fig 1A-L). When the number of INs in different zones of the neural tube (divided into 10 equal distance bins spanning from the ventricular zone, VZ, to marginal zone, MZ) was counted in the WT mice at E14.5, we found that the number of INs migrating along the MZ is higher than those along the intermediate zone (IZ) (Fig 1G). On the contrary, in the cKO mice, tdTomato^{+} INs in the migration path along the MZ was nearly completely absent, whereas those migrating along the IZ were still detected with increased numbers (Fig 1G). In agreement with these data, a smaller number of tdTomato+ INs come out of the GE in these cKO mice compared to the WT littermate controls (Fig 1B, E). These results suggest that Gad2^{+} cINs migrating along MZ are selectively affected in Gad2Cre-Arx cKO mice. Consistent with these results in the embryonic brain, in the control mice at P0, tdTomato^{+} cortical INs (cINs) are localized densely in the MZ and populated mostly in the upper layers; in the cKO mice, tdTomato^{+} cINs are disproportionally localized in the IZ and deeper layers, while scarcely distributed in the MZ and upper layers (Fig 1H-L).

Given that Gad2^{+} INs are originated from the entire ganglionic eminences, in Gad2Cre-Arx cKO mice Arx is deficient in most INs originated from the MGE, LGE, and CGE. To study a

specific role(s) of Arx in MGE-derived INs, we took advantage of Nkx2.1-Cre line which can abrogate *Arx* in progenitor cells of the MGE when crossed to Arx flox mice. With the similar breeding scheme used above with Gad2Cre line, we generated Nkx2.1Cre-Arx cKO; Ai14 mice (also referred to as $Arx^{F/Y}$; Ai14; Nkx2.1Cre). These Arx cKO mice driven by Nkx2.1Cre showed similar IN phenotypes detected in Gad2Cre-driven cKO mice, with dramatic reduction of migrating INs along the MZ while relatively spared INs along the IZ despite reduced number compared to the WT littermate controls (Fig 1M-P). Remarkably, most cINs migrating out of the MGE were stuck around the ventro-lateral boundary of the GE and failed to migrate out to the cortex (arrows in Fig IN).

To explore if these IN defects observed in Gad2Cre- and Nkx2.1Cre-driven Arx cKO mice are specifically due to early loss of Arx in the SVZ cells, we employed Sst-IRES-Cre line which abrogates Arx in SST+ postmitotic INs (SST expression turns on in the postmitotic INs outside of the SVZ), as well as Ai14 line. Surprisingly, cIN distribution in the cortical layers of these cKO mice driven by Sst-IRES-Cre did not show any obvious differences compared to the WT littermate controls (P0) (Fig 1Q-T). Taken together, our data with three different lines of Arx cKO mice strongly suggest that the early loss of Arx in the SVZ (by Gad2Cre- and Nkx2.1Cre), but not its loss in the differentiated postmitotic INs (by Sst-IRES-Cre), affects cIN migration.

Deletion of Arx leads to interneuron differentiation defects (scRNA-seq).

To explore how changes in gene expression impact the observed phenotype and other aspects of cIN development, transcriptome profiling was conducted at the single-cell level via single-cell RNA sequencing (scRNA-seq) using wild-type (WT) and cKO samples (E16.5) genetically labeled by Gad2-Cre; Ai14 (Fig 2A). The $tdTomato^{+}$ INs of the forebrain were isolated via fluorescence-activated cell sorting (FACS) and used for scRNA-seq by the 10Xgenomics platform. After quality control, removing doublet, and normalization of sequencing results, we obtained a total of 27,585 high-quality cells across 5 samples (2 WT and 3 MT brains), including 11,224 WT cells and 16,361 Arx-deficient mutant cells. Our clustering analysis revealed one MGE, one CGE, three LGE, one septal, and one excitatory clusters, which we annotated using a predefined set of marker genes[4], in addition to two unknown clusters (Fig 2B, C). The LGE

cluster was subdivided into three additional clusters (D1 MSN Tshz1, D1 MSN Pdyn, and D2 MSN), based on the predefined set of markers[24,25].

Next, we aligned the two single-cell datasets using Scanpy, specifically the scvi-tools, to compare clusters between WT and Arx-deficient mutant, (Fig 2B). Subsequent composition analysis identified several significant subtype changes in the Arx mutant group (Fig 2D). The mutant samples showed a lower proportion of CGE and a higher proportion of MGE cells when compared to WT samples. Additionally, we noticed subtype alterations in the LGE clusters. Specifically, there was a decrease in D1 MSN Pdyn, increase in D1 MSN Tshz1, and no detectable changes in D2 MSN cell clusters (Fig 2B). These results are in accordance with the observation that Arx is present only in the first two clusters but not in the third cluster (D2 MSN) (data not shown).

To gain further insights into the differentiation process, we performed RNA velocity analysis to infer the directionality and kinetics of gene expression changes in the cells. A few differences were noted when the WT and mutant clusters were compared. For example, compared to a clear and well-defined differentiation pseudo time flow mapped in the WT MGE cluster (Fig 2E), in the mutant MGE, we observed broken trajectories in MGE clusters indicating an impaired or disrupted differentiation process. Also, while the WT exhibited distinct differentiation patterns between MGE and CGE without any connected trajectories, the mutant samples displayed a continued differentiation line without clear boundaries. Conversely, in the case of the LGE, no significant differences between WT and mutant samples were noted. Together these findings suggest that the loss of Arx in the Gad2^{+} INs impacts their proper differentiation, particularly MGE- and CGE-derived cINs and some LGE-derived INs. Of note, most cINs are derived from the MGE and CGE, but not from the LGE (LGE gives rise to subcortical INs and some projection neurons).

Differentially expressed genes play key roles in IN subtype differentiation and migration.

Our results so far with the brain imaging and cell cluster analysis (based on scRNA-seq data) strongly suggest that the IN phenotypes observed in Arx-deficient mice may be associated with changes in IN subtype development and defects in IN migration. To gain insights on molecular players involved in these defects, we first identified differentially expressed genes (DEGs) (using

scRNA-seq data) in Arx-deficient INs (of the Gad2Cre-Arx cKO; Ai14 mice), by conducting machine learning-based single cell transcriptome analysis using scvi-tools package[26]. In the MGE, we identified a total of 50 genes with differential expression between the two groups and top 20 genes are shown in Fig 3A. Among these genes, Lmo1, Lmo3, Lmo4, and Nkx2-1 were found to be upregulated in the mutant cells, while Maf, Mef2c, Cxcr4, Ackr3, and Cux2 genes were downregulated (Fig 3A). These genes are known to play roles in IN development and migration. For instance, transcription factors Lmo and Nkx2-1 are normally expressed in progenitor cells of the GE; Maf and Mef2c, transcription factors, are known to positively influence the development of parvalbumin expressing (PV+) INs; Cxcr4 is required for IN migration. The ven diagrams of differentially expressed genes identified in MGE, LGE, and CGE demonstrate that MGE and CGE have more up- or down-regulated genes compared to LGE, reflecting that Arx is primarily expressed in MGE and CGE and exerts its function in these two areas than LGE (Fig 3D). Consistent with this, the gene ontology (GO) analysis on the downstream genes revealed enrichment of the pathways associated with SST differentiation and IN migration (data not shown). In contrast, upregulated genes were enriched in the pathways related to the mis-regulated IN development (data not shown).

To validate DEGs identified from the scRNA-seq data analysis, we performed immunofluorescent labeling with a panel of primary antibodies on cKO (Gad2Cre-Arx cKO; Ai14) and WT (WT;Ai14) brains. Nkx2-1, which is a key transcription factor governing MGE interneuron development, was expressed at higher levels in the ventral telencephalon of the cKO mice compared to the WT, at E14.5 (Fig 4A, B, D, E). Surprisingly, Nkx2-1+ cells were also detected in the dorsal telencephalon of the cKO mice (Fig 4D, F); Nkx2-1 is normally only briefly expressed in the progenitor cells of the cortical IN lineage in the ventral telencephalon and becomes downregulated in migrating immature cells as part of their differentiation program (Fig 4A, C) [5,27–29]. Intriguingly, in these cKO mice, Nkx2-1 expression stayed in the Arx-deficient, tdTomato+ INs in their route to the cortex when examined at P0, while none of tdTomato+ INs in the WT cortex was Nkx2-1 positive (Fig 4G-N). These observations suggest an abnormal IN development in cKO, perhaps a delay in IN differentiation or failure in progressing to differentiation status from immature fate.

We have validated other differentially expressed genes identified in scRNA-seq analysis; the two transcription factors, Maf and Mef2c, which are known to positively influence the

development of the parvalbumin expressing (PV+) INs[15], were among the genes decreased in cKO in single cell analysis. Our immunolabeling data on Maf and Mef2c confirmed their reduced expression in the cortical INs at E14.5 (Fig 4O-T). As these cKO mice die shortly after birth but before PV expression starts, we were not able to check if PV+ INs were reduced or not in these cKO mice. Interestingly, SST+ cells were dramatically increased in the cKO mice (E14.5) as shown with immunolabeling assay (Fig 4U, V). This might have resulted from IN subtype shift from PV+ to SST+ probably due to decreased expression of Maf and Mef2c, although we do not have any evidence to support this speculation. Together, these findings provide insight into interneuron subtype alterations in the Gad2Cre-Arx cKO brains.

Lmo genes (*Lmo*1, 3, and 4) were also found in the list of upregulated genes in our single-cell analysis (Fig 3A). They encode transcription factors that are downregulated as the INs start to migrate out to the cortex[30]. As Lmo1 was consistently detected as a differentially expressed gene in multiple transcriptomic analyses with different Arx mutant mouse lines (Fig 3A, B) [17,31,32] , we set out to explore its function by overexpressing it to the GE of the WT mice. When the control construct expressing GFP only (pCIG) was electroporated directly into the GE of slice cultures from E14.5 CD1 brains, the electroporated INs migrated to the cortex normally (Fig 5C, E-F). In striking contrast, most INs electroporated with Lmo1 expression construct (pCIG-Lmo1) failed to migrate to the cortex (Fig 5D, G-H). These results provide evidence that upregulated Lmo1 in Arx-deficient INs (of Gad2Cre-Arx cKO) might be at least partially responsible for IN migration/distribution defects. Interestingly, when these slice cultures were immunolabeled with CXCR4 antibody, migrating INs electroporated with Lmo1 did not express Cxcr4, while those electroporated with GFP showed normal Cxcr4 expression (Fig 5E-H). Based on these results, we postulate that during normal development, Arx suppresses Lmo1 expression, thus lifting up Cxcr4 repression by Lmo1 and facilitating INs to migrate out to the cortex. This hypothesis is in line with our findings in Arx deficient INs, where i) Lmo1 is upregulated (Fig 3A, B) while ii) Cxcr4 is downregulated as shown in single cell analysis (Fig 3A) as well as confirmed by immunolabeling in embryonic brain of cKO mice (most of the Cxcr4 expression detected in the ventral telencephalon as well as in the cortex of the WT brain are almost completely absent in the cKO mice) (Fig 5A), and iii) IN migration to the cortex is compromised (Fig 1). Notably, IN migration defects in the cKO mice resemble the phenotypes observed in Cxcr4 knockout models[33], supporting the role of Cxcr4 in IN migration phenotype of the Arx cKO mice. Taken

together, our results implicate that Arx suppresses progenitor fate and promotes IN subtype development and migration during normal development, by repressing genes involved in proliferation and activating transcription factors involved in IN differentiation and migration, either directly or indirectly.

ARX directly regulates transcription of a subset of genes involved in IN differentiation and migration.

In order to find out whether ARX directly regulates differentially expressed genes described above (including Lmo1, Lmo4, Maf, Mef2c, and Cxcr4), we next performed ARX chromatin immunoprecipitation-sequencing (ChIP-seq) analysis using E14.5 forebrain samples (Fig 6). Binding site analyses show ARX binding sites are located near transcription start site (TSS) and enriched in promoters and exons in the genome (Fig 6A-B). We also analyzed consensus binding motif of ARX as shown in Fig 6C. Out of 1484 differentially expressed genes and 6590 identified in Arx ChIP-seq, 607 genes were overlapped (Fig 6D). Out of 14049 Dlx2 ChIP-seq genes and 8764 Arx ChIP-seq genes, xx genes were overlapped (Fig 6E). Our ChIP-seq data demonstrated that the anti-ARX antibody effectively immunoprecipitated Lmo1, Maf, and Mef2c loci indicating a direct interaction between ARX and these loci (Fig 7). However, we did not observe an interaction between ARX and the Cxcr4 locus, suggesting that ARX may not directly control Cxcr4 expression (data not shown).

Ala expansion mice ($Arx^{(GCG)7/Y}$*) have reduced number of INs and IN migration defects.*

Studying IN differentiation in Arx knockout (KO) models can be challenging, as INs do not mature fully until around P30 and Gad2Cre-Arx cKO mouse is lethal at birth[16]. Thus, we turned to another Arx mouse model, $Arx^{(GCG)7/Y}$, which contains insertional mutation with seven additional Ala residues to the normal stretch of Ala residues in the poly Ala tract 1 (PA1) of Arx gene[18]. The ARX protein contains several alanine (Ala) stretches of around 5-15 residues in its primary sequence (four poly Ala tracts in ARX are called PA1, PA2, PA3, and PA4, respectively)[34], and changes in the number of Ala residues in PA 1 or PA2, the most common mutations in patients with *ARX* mutations, have been linked to epilepsy and mental issues in

humans and mice. Arx$^{(GCG)7/Y}$ mice usually survive until weaning age, making it possible to study IN development. Previously published studies reported no significant changes in IN numbers or distribution in this mouse model[35], but we took a different approach, which became available more recently, to investigate any defects in IN development, which were unrecognized due to the weakness of the previously used methodologies.

Using P30 mice, we took advantage of tissue clearing technique for brain image analysis combined with SST and PV immunolabeling (at P30) [23]. The resulting 3D image analysis revealed that the density of SST INs in Arx$^{(GCG)7/Y}$ mice is reduced in the cerebral cortex, hippocampus, and striatum, while no significant differences were observed in the midbrain and hindbrain, compared to the WT (Fig 8 A). Among the neocortical areas, the differences of the SST IN in the motor cortex and somatosensory cortex were most severe, while the auditory cortex and visual cortex showed statistically not significant changes (Fig 8B-C). We focused on layer-specific distributions of SST INs in the somatosensory cortex. In Layer 2-6a, Arx$^{(GCG)7/Y}$ brain showed severe loss of SST INs (Fig 8B). Unlike SST, the density of PV or VIP INs did not show statistically significant reduction in most areas except auditory cortex (Fig 8D-F and data not shown).

To study IN migration phenotype in Arx$^{(GCG)7/Y}$ mice, we first examined IN migration streams in the embryonic cortex (E14.5) with DLX2 immunolabeling, which is a pan-IN specific marker and detected reduction in the number of Dlx2+ cortical INs along MZ of the cortical migration path (Fig 9A-B). We also investigated the level of Cxcr4 expression (which is a critical chemokine receptor for IN migration) in the migrating cortical INs at E14.5 and found a dramatic reduction along the migration path (Fig 9C). Notably, Cxcr4 KO mice exhibits a similar migration phenotype detected in Arx mutant mice[33], supporting the role of Cxcr4 in IN defects observed in Arx$^{(GCG)7/Y}$ mice.

We next performed RNA microarray analysis with E14.5 GE of Arx$^{(GCG)7/Y}$ and of WT mice to identify differentially regulated genes in the mutant mice. We have compared the list of upregulated and downregulated genes in Arx$^{(GCG)7/Y}$ GE cells with those obtained from the microarray analysis with the GFP-positive GE cells of the E14.5 Arx cKO by Dlx5/6CreGFP (which deletes Arx in Dlx5/6 lineage cells) (Fig 9D) [21]. As anticipated, Arx$^{(GCG)7/Y}$ GE had fewer mis-regulated genes (with fold change >1.5 and adj-P value <0.05), compared to Arx-deficient

GE (Arx cKO by Dlx5/6Cre-GFP). Among the total 45 upregulated genes in $Arx^{(GCG)7/Y}$ GE, only 15 of them (including Lmo1 and Lmo4) were overlapped with Arx cKO (total 1329 genes were upregulated in Arx cKO GE). Similarly, among 22 downregulated genes in $Arx^{(GCG)7/Y}$ GE, only Cxcr4 and Lhx8 were also downregulated in Arx cKO GE (total 845 genes downregulated in Arx cKO GE) (Fig 9D).

Discussion

Cortical INs play key roles in modulating neural circuit activity in the brain by connecting with local PNs as well as other INs. It is for this reason perhaps our brain generates tremendously diverse, both structurally and functionally, cortical INs. Our study presented here is aimed at providing better understanding of IN development and migration. Our analyses on Arx cKO as well as Arx Ala expansion mice ($Arx^{(GCG)7}$, also known as Arx^{PA1-22}) provide a comprehensive perspective on the impact of the Arx mutations on cellular gene expression and cell type composition at a single cell level as well as in the whole cortex. The findings from our study indicate potential associations between the mutant genotype and specific cellular processes and cell types, which could have significant implications for understanding the molecular mechanisms underlying the mutant phenotype.

Arx serves as a pan-interneuron molecular marker, including MGE, CGE, and LGE, with the exception of VZ. Loss of Arx in the interneuron population leads to a significant shift in cortical interneuron subtypes derived from MGE/CGE and LGE. The MGE-like cell population in the cKO brain increases, while CGE cells decrease in number. Additionally, a shift between MSN D1 and D2 is observed, indicating that Arx plays a role in balancing MGE, CGE, and LGE populations. Regarding MGE, the Arx cKO brain exhibits elevated levels of MGE cells expressing Nkx2-1, Tox, Tox3, and Lmo1, the ventral VZ markers, compared to the control. This phenotype resembles that of Sp9 KO[35], indicating a genetic pathway connection.

The Arx cKO shows altered gene expression characteristics of CGE (Prox1 and Nr2f2) and LGE (Meis2 and Etv1) in the MGE cluster. In the LGE cluster, Sox6, a gene specific to MGE, is upregulated in cKO, while the CGE cluster exhibits ectopic expression of LGE markers, including Etv1. Since these upregulated genes are transcription factors, the cKO clusters lack distinct characteristics of their own, suggesting that Arx controls the determination of interneuron region specification. Despite Arx not being expressed in the LGE D2 MSN cluster, the LGE D2 MSN cluster of cKO leads to reduction of cells in the number compared to the control. This change may be a result of a non-cell-autonomous manner, though the exact mechanism remains unclear.

The loss of Arx in the SVZ results from the use of Gad2-Cre for generating cKO, which turns on Cre expression in SVZ of GE, but not in VZ where Arx is not expressed. Consequently, the

phenotypes, such as interneuron migration defects at the marginal zone, delayed migration, and altered subtype specification, result from the loss of Arx in SVZ. Interestingly, deletion of Arx by Sst-Cre, activated after SVZ, does not show an apparent phenotype, similar to the situation in Maf/Marb cKO[15]. Our cKO analysis suggests that SVZ plays a critical role in interneuron subtype specification, deviating from the existing model that indicates a PV vs. Sst-specific predisposed progenitor zone in VZ.

Various transcription factors are associated with interneuron subtype specification. Nkx2.1, expressed exclusively in MGE, gives rise to cortical interneurons, and its absence leads to a significant reduction in PV and SST interneurons, with less impact on CR interneurons. Lhx6, another critical factor in cortical interneurons, also contributes to the reduction of PV interneurons, with less effect on SST interneurons. Additionally, Nr2f2 KO results in a reduction in SST interneurons, while Nr2f1 conditional knockout leads to fewer PV interneurons[36]. Furthermore, Sat1b KO is associated with a more severe reduction in Sst-interneurons compared to PV-interneurons[37], and Sp9 KO affects both Sst- and PV-interneurons in the cortex [35].

In $Arx^{(GCG)7}$ mutant's transcriptome analysis, we identified a few differentially expressed genes that do not share with Arx KO as a subset. However, Lmo1, expressed in VZ is up-regulated, and Cxcr4, the receptor for interneuron migration, is down-regulated. In the current study, we showed that forced overexpression of Lmo1 as mimicked in Arx mutant mice can delay interneuron migration and repress the expression of Cxcr4. We speculate that loss of Lmo1 expression outside of the VZ can derepress Cxcr4 expression to migrate out from GE. Mis-regulation of Lmo1 in ventral neurogenesis can affect migration. $Arx^{(GCG)7}$ has a subtype change phenotype which is a more severe loss of Sst-interneuron, when examined at P30. But pan-interneuron-conditionally deleted Arx has an overproduction of Sst-interneuron, less PV-interneuron (or less Maf/Mef2c-interneuron to be exact) and less CGE-derived interneuron in the number. However, due to perinatal lethality in the cKO, we do not know if overproduced SST IN will be maintained or undergo cell death during postnatal pruning period.

References

1. Hu, J. S., Vogt, D., Sandberg, M. & Rubenstein, J. L. Cortical interneuron development: a tale of time and space. *Development* **144**, 3867–3878 (2017).

2. Lim, L., Mi, D., Llorca, A. & Marín, O. Development and Functional Diversification of Cortical Interneurons. *Neuron* **100**, 294–313 (2018).

3. Fishell, G. & Kepecs, A. Interneuron Types as Attractors and Controllers. *Annu. Rev. Neurosci.* **43**, 1–30 (2019).

4. Wamsley, B. & Fishell, G. Genetic and activity-dependent mechanisms underlying interneuron diversity. *Nat. Rev. Neurosci.* **18**, 299–309 (2017).

5. DeFelipe, J. *et al.* New insights into the classification and nomenclature of cortical GABAergic interneurons. *Nat. Rev. Neurosci.* **14**, 202–216 (2013).

6. Yu, X. & Zecevic, N. Dorsal Radial Glial Cells Have the Potential to Generate Cortical Interneurons in Human But Not in Mouse Brain. *J. Neurosci.* **31**, 2413–2420 (2011).

7. Paterno, R., Casalia, M. & Baraban, S. C. Interneuron deficits in neurodevelopmental disorders: Implications for disease pathology and interneuron-based therapies. *Eur. J. Paediatr. Neurol.* **24**, 81–88 (2020).

8. Llorca, A. & Deogracias, R. Origin, Development, and Synaptogenesis of Cortical Interneurons. *Front. Neurosci.* **16**, 929469 (2022).

9. Tasic, B. *et al.* Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci* **19**, 335–346 (2016).

10. Sussel, L., Marin, O., Kimura, S. & Rubenstein, J. L. R. Loss of Nkx2.1 homeobox gene function results in a ventral to dorsal molecular respecification within the basal telencephalon: evidence for a transformation of the pallidum into the striatum. *Development* **126**, 3359–3370 (1999).

11. Liodis, P. *et al.* Lhx6 Activity Is Required for the Normal Migration and Specification of Cortical Interneuron Subtypes. *J. Neurosci.* **27**, 3078–3089 (2007).

12. Zhao, Y. *et al.* Distinct molecular pathways for development of telencephalic interneuron subtypes revealed through analysis of Lhx6 mutants. *J. Comp. Neurol.* **510**, 79–99 (2008).

13. Nóbrega-Pereira, S. *et al.* Postmitotic Nkx2-1 Controls the Migration of Telencephalic Interneurons by Direct Repression of Guidance Receptors. *Neuron* **59**, 733–745 (2008).

14. Pai, E. L.-L. *et al.* Maf and Mafb control mouse pallial interneuron fate and maturation through neuropsychiatric disease gene regulation. *Elife* **9**, e54903 (2020).

15. Pai, E. L.-L. *et al.* Mafb and c-Maf Have Prenatal Compensatory and Postnatal Antagonistic Roles in Cortical Interneuron Fate and Function. *Cell Rep.* **26**, 1157-1173.e5 (2019).

16. Kitamura, K. *et al.* Mutation of ARX causes abnormal development of forebrain and testes in mice and X-linked lissencephaly with abnormal genitalia in humans. *Nat. Genet.* **32**, 359–369 (2002).

17. Fulp, C. T. *et al.* Identification of Arx transcriptional targets in the developing basal forebrain. *Hum. Mol. Genet.* **17**, 3740–3760 (2008).

18. Kitamura, K. *et al.* Three human ARX mutations cause the lissencephaly-like and mental retardation with epilepsy-like pleiotropic phenotypes in mice. *Hum Mol Genet* **18**, 3708–3724 (2009).

19. Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat. Neurosci.* **13**, 133–140 (2010).

20. Taniguchi, H. *et al.* A Resource of Cre Driver Lines for Genetic Targeting of GABAergic Neurons in Cerebral Cortex. *Neuron* **71**, 995–1013 (2011).

21. Stenman, J., Toresson, H. & Campbell, K. Identification of two distinct progenitor populations in the lateral ganglionic eminence: implications for striatal and olfactory bulb neurogenesis. *J. Neurosci. : Off. J. Soc. Neurosci.* **23**, 167–74 (2003).

22. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol.* **19**, 15 (2018).

23. Park, Y.-G. *et al.* Protection of tissue physicochemical properties using polyfunctional crosslinkers. *Nat Biotechnol* **37**, 73–83 (2019).

24. Shi, Y. *et al.* Mouse and human share conserved transcriptional programs for interneuron development. *Science* **374**, eabj6641 (2021).

25. Zhao, Z. *et al.* Evolutionarily conservative and non-conservative regulatory networks during primate interneuron development revealed by single-cell RNA and ATAC sequencing. *Cell Res.* **32**, 425–436 (2022).

26. Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep Generative Modeling for Single-cell Transcriptomics. *Nat. methods* **15**, 1053–1058 (2018).

27. Tasic, B. *et al.* Shared and distinct transcriptomic cell types across neocortical areas. *Nature* **563**, 72–78 (2018).

28. Krienen, F. M. *et al.* Innovations present in the primate interneuron repertoire. *Nature* **586**, 262–269 (2020).

29. Hodge, R. D. *et al.* Conserved cell types with divergent features in human versus mouse cortex. *Nature* **573**, 61–68 (2019).

30. Chen, Y.-J. J. *et al.* Single-cell RNA sequencing identifies distinct mouse medial ganglionic eminence cell types. *Sci Rep-uk* **7**, 45656 (2017).

31. Drongitis, D. *et al.* Deregulation of microtubule organization and RNA metabolism in Arx models for Lissencephaly and developmental epileptic encephalopathy. *Hum. Mol. Genet.* **31**, ddac028- (2022).

32. Lee, K., Mattiske, T., Kitamura, K., Gecz, J. & Shoubridge, C. Reduced polyalanine-expanded Arx mutant protein in developing mouse subpallium alters Lmo1 transcriptional regulation. *Hum. Mol. Genet.* **23**, 1084–1094 (2014).

33. Wang, Y. *et al.* CXCR4 and CXCR7 Have Distinct Functions in Regulating Interneuron Migration. *Neuron* **69**, 61–76 (2011).

34. Friocourt, G. & Parnavelas, J. G. Mutations in ARX Result in Several Defects Involving GABAergic Neurons. *Front. Cell. Neurosci.* **4**, 4 (2010).

35. An Epilepsy-Related ARX Polyalanine Expansion Modifies Glutamatergic Neurons Excitability and Morphology Without Affecting GABAergic Neurons Development. *Cerebral Cortex* **23**, 1–11 (2012).

36. Liu, Z. *et al.* Sp9 Regulates Medial Ganglionic Eminence-Derived Cortical Interneuron Development. *Cereb. Cortex* **29**, 2653–2667 (2019).

Figure Legends

Figure 1. Interneuron (IN) generation, migration, and distribution defects in ARX cKO;Ai14 mice driven by Gad2Cre and Nkx2.1Cre but not by SstCre. **A-F; H-K,** Representative images of the immunolabeling of tdTomato on coronal sections of ARX cKO under Gad2Cre (A-C) and wild type (WT) littermate control (D-F) taken at E14. 5 (A-F) and P0 (H-K). B and C are magnified images of the boxed areas in A; E and F are magnified images of the boxed areas in E and F. **G, L,** Quantification of the percentage of the tdTomato+ cell numbers in each bin to the total tdTomato+ cell numbers in all 10 bins at E 14.5 (G) and P0 (L). **M-P,** Representative images of the immunolabeling of tdTomato (red) on coronal sections of ARX cKO under Nkx2.1Cre^{ER} (M, O) and WT littermate control (N, P) taken at E14. 5. DAPI (blue) was used for nuclear staining. Arrows in N indicate INs stuck in the GE and failed to migrate out to the cortex. O and P are magnified images of the boxed areas in M and N, respectively. **Q-T,** Representative images of the immunolabeling of tdTomato (red) on coronal sections of ARX cKO under SstCre (Q, S) and WT littermate control (R, T) taken at E14. 5. S and T are magnified images of the boxed areas in Qand R, respectively. CTX, cortex; GE, ganglionic eminence; IZ, intermediate zone; MZ, marginal zone.

Figure 2. Alterations in IN subtypes revealed by single cell transcriptome analysis of Gad2Cre-ARX cKO and WT INs. **A,** Schematic diagram of overall experimental approach. The tdTomato+ INs were collected using FACS from E16.5 forebrains (n=3 for WT and n=2 for MT). **B,** UMAP projection of INs identified 8 clusters (left panel) and the results from WT (orange) and Gad2Cre-Arx cKO (blue) were overlaid to each other (right panel). **C,** Classification of INs into several IN subtypes. Heatmap of clusters indicates expression of maker genes differentially expressed that best present interneuron subtype. **D,** Composition analysis of clusters. WT MGE cell population is increased compared to the cKO, but CGE population is decreased in the cKO. **E,** Inferring differentiation trajectories by RNA velocity analysis. The overlaid arrows indicate inferred developmental trajectory.

Figure 3. Differentially expressed genes in Gad2Cre-ARX cKO INs grouped by cell types. **A-C,** Heatmap of MGE (A), CGE (B) and LGE (C) clusters displaying differentially expressed genes, compared to controls (WT). **D,** Venn diagram of up-regulated and down-regulated genes in MGE, LGE and CGE INs. CGE, caudal ganglionic eminence; LGE, lateral ganglionic eminence; MGE, medial ganglionic eminence.

Figure 4. Abnormal IN subtype development in ARX cKO. **A-N,** Representative images of NKX2-1 immunolabeling on E14.5 (A-F) or P0 (G-N) coronal sections of the WT littermate (A-C; G-J) and Gad2Cre-ARX cKO (D-F; K-N) brains. cKO cortical INs (cINs) migrating to the cortex (F) as well as those arrived at the cortical plate (M and N) abnormally express Nkx2-1, a marker for MGE progenitors (normally not expressed in migrating or mature cINs in the cortex). Nkx2-1 positive cell is not localized on the marginal zone (F). Lateral side of cKO GE (E) shows more Nkx2-1-positive cells, compared to WT (B). B, C; E, F; H, I, J; L, M, N are magnified images of the boxed areas in A, D, G, K, respectively. **O-T,** Representative images of RNA-scope assay with probe *Maf* (P, S) and *Mef2c* (Q, T) on E14.5 coronal sections of the WT (O-Q) and Gad2Cre-Arx ckO (R-T). *Maf2* and *Mef2c* expressions are diminished in cKO. P, Q and S, T are magnified images of the boxed areas in O and R, respectively. **U-V,** Representative coronal section images of SST immunolabeling on WT (U) and Gad2Cre-Arx ckO (V). Dotted circles in U indicate increased SST expression in cKO.

Figure 5. Upregulated *Lmo1* inhibits IN migration by repressing *Cxcr4* expression, explaining IN migration defects in Arx cKO. **A-B,** Representative coronal section images of CXCR4 immunolabeling in WT(A) or Gad2Cre-Arx cKO (B) brain. **C-H,** Forced expression of pCIG (C, E, F) or pCIG-Lmo1 (D, G, H) into the GE of WT brain slice cultures via electroporation at E14.5 (analyzed 2 days later). GFP+ cells (green, electroporated cells) migrate normally to the cortex in the slice culture electroporated with control vector (pCIG) (C), but not in the one electroporated with Lmo1 expression construct (pCIG-Lmo1) (D). Migrating INs electroporated with pCIG (green cells marked with dotted circles in E) expresses Cxcr4 (red in F), but those with pCIG-Lmo1 (green cells marked with dotted circles in G) lost Cxcr4 expression (note no signal in dotted cells in H),

Figure 6. ChIP-seq analysis with anti-ARX antibody. **A,** Heatmap plot of ARX binding sites near transcription start site (TSS). **B,** ARX binding regions of the mouse genome. **C,** ARX binding site sequences in ChIPed DNA. **D,** Comparison of differentially expressed genes identified in scRNA-seq vs. the genes identified in ARX ChIP-seq. **E,** Comparison of ARX ChIP-seq with DLX2 ChIP-seq results showing the overlap between DLX2 and ARX binding regions in mouse genome.

Figure 7. Examples of gene loci identified as ARX binding sites in ARX ChIP-seq analysis. **A-D,** ChIP-seq analysis showing ARX binding regions in *Maf* (A), *Mef2c* (B), *Lmo1* (C), and *Lmo2* (D) gene loci, overlaid with DLX2 binding regions. Red circles indicate examples of ARX specific peaks in each gene loci.

Figure 8. Loss of SST and PV cortical interneurons in $Arx^{(GCG)7}$ mice. Analysis of 3D immunolabeling of SST and PV on adult brains (P30) of $Arx^{(GCG)7}$ and WT mice processed with tissue clearing technique. **A, D,** Quantification of SST IN (A) or PA IN (D) cell density (cells/mm^3) in each indicated brain area of the WT (labeled as W) and $Arx^{(GCG)7}$(labeled as A). **B, E,** Quantification of SST IN (B) or PA IN € cell density in each indicated layer of the somatosensory cortex of the Wild Type and $Arx^{(GCG)7}$(labeled as Ala Exp). **C, F,** Heat map on the coronal section illustrations showing the densities of the SST or PV immunolabeled cells of the WT and $Arx^{(GCG)7}$(labeled as Ala Exp) brains.

Figure 9. Interneuron migration defects and differentially expressed genes in $Arx^{(GCG)7}$. **A, C,** Representative images of DLX2 (A) or CXCR4 (C) immunolabeling on coronal sections of the embryonic (E14.5) WT and $Arx^{(GCG)7}$ forebrains. Right panels are magnified images of the boxed areas in the left panels. **B,** Quantification of the percentage of DLX2+ cells in each bin over the total number of DLX2+ cells in all 10 bins. **C,** Venn diagram of upregulated and downregulated genes identified from RNA microarray analysis of the ARX cKO (by Dlx5/6Cre) vs. $Arx^{(GCG)7}$ GE tissues, compared to each WT littermate set.

Figure 1

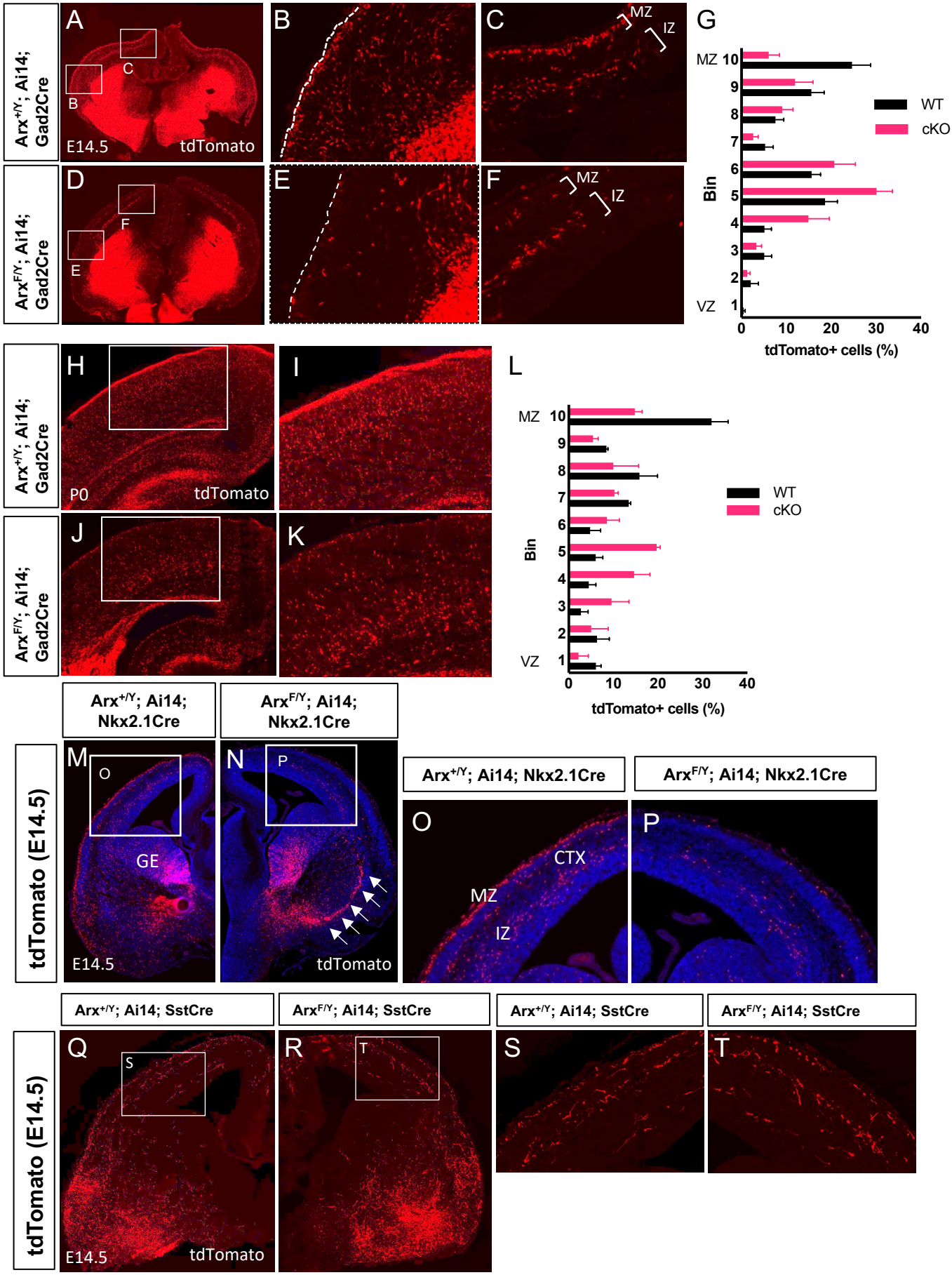

Arx+/Y; Ai14; Gad2Cre
ArxF/Y; Ai14; Gad2Cre
A
B
C
D
E
F
E14.5
tdTomato
MZ
IZ
G
MZ
VZ
Bin
WT
cKO
tdTomato+ cells (%)
H
I
J
K
P0
L
Arx+/Y; Ai14; Nkx2.1Cre
ArxF/Y; Ai14; Nkx2.1Cre
tdTomato (E14.5)
M
N
O
P
GE
CTX
Arx+/Y; Ai14; SstCre
ArxF/Y; Ai14; SstCre
Q
R
S
T

Figure 2

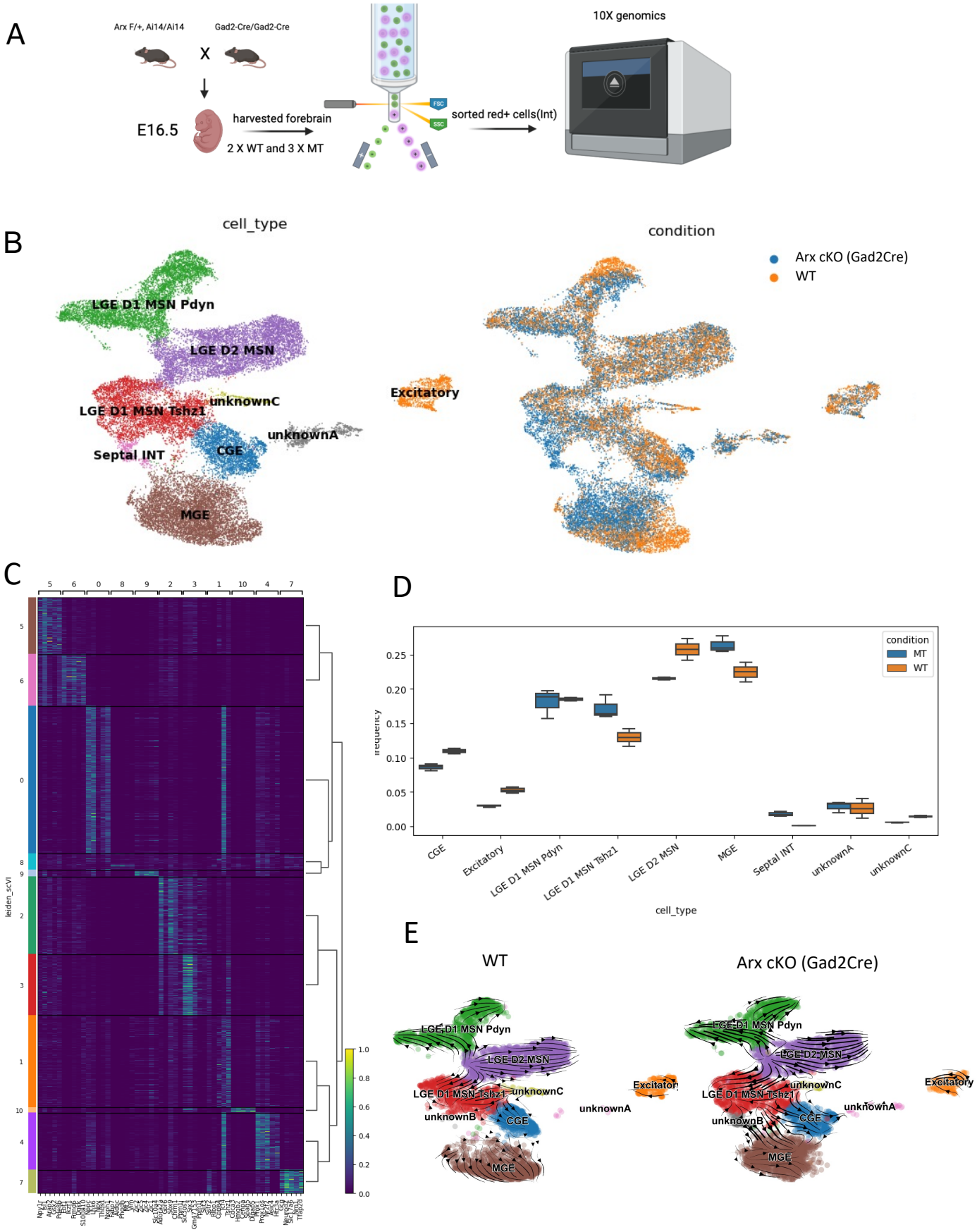
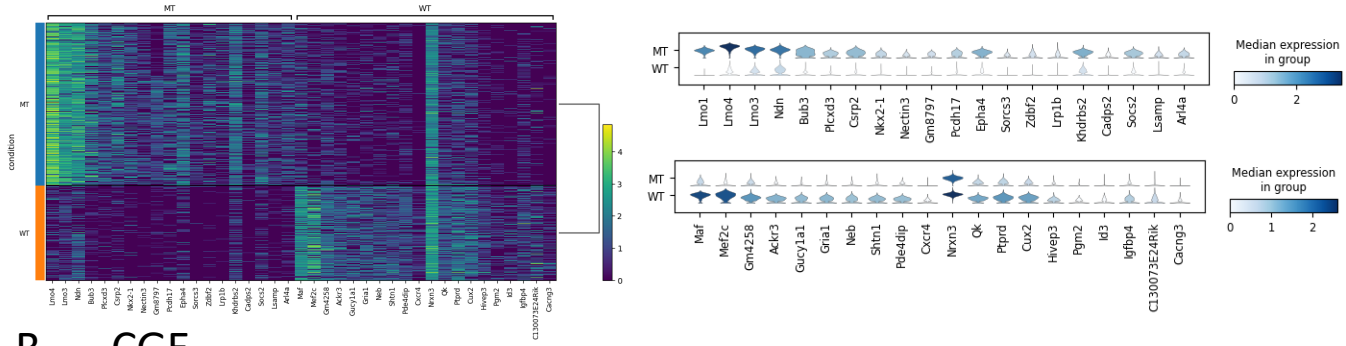
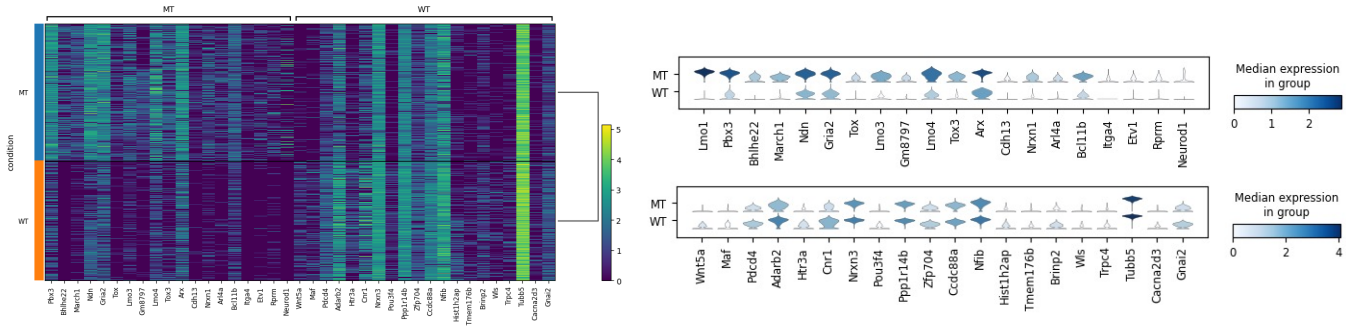

A
Arx F/+, Ai14/Ai14
Gad2-Cre/Gad2-Cre
X
E16.5
harvested forebrain
2 X WT and 3 X MT
FSC
SSC
sorted red+ cells(Int)
10X genomics
B
cell_type
LGE D1 MSN Pdyn
LGE D2 MSN
LGE D1 MSN Tshz1
unknownC
unknownA
Septal INT
CGE
MGE
Excitatory
condition
Arx cKO (Gad2Cre)
WT
C
leiden_scVI
D
condition
MT
WT
frequency
CGE
Excitatory
LGE D1 MSN Pdyn
LGE D1 MSN Tshz1
LGE D2 MSN
MGE
Septal INT
unknownA
unknownC
cell_type
E
WT
Arx cKO (Gad2Cre)
LGE D1 MSN Pdyn
LGE D2 MSN
LGE D1 MSN Tshz1
unknownC
unknownB
CGE
unknownA
MGE
Excitatory

Figure 3

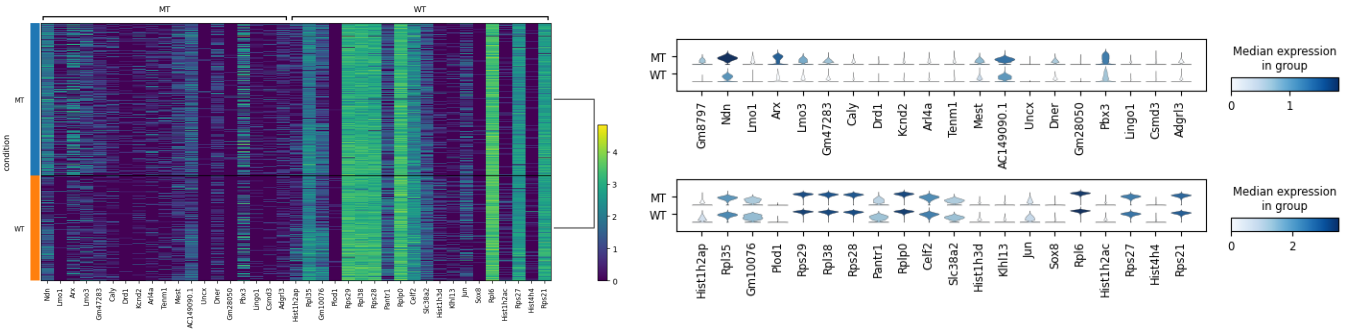
A MGE



B CGE



C LGE



D

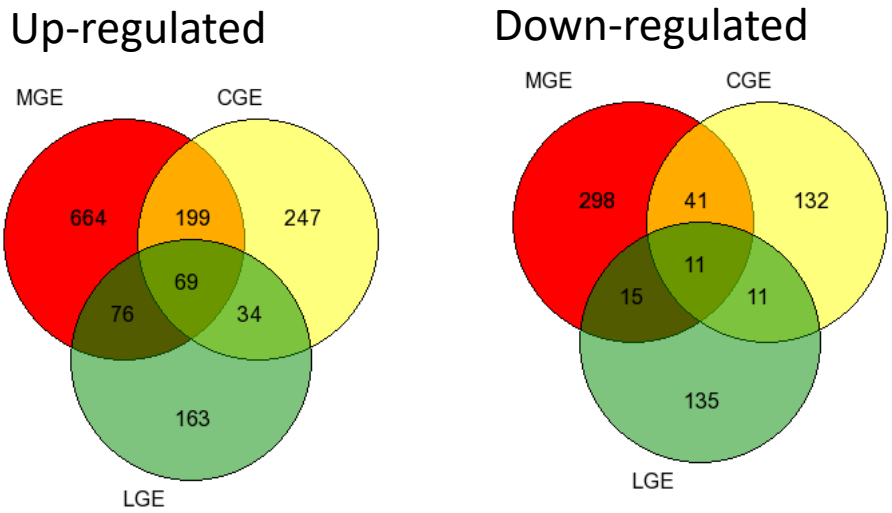


Figure 4

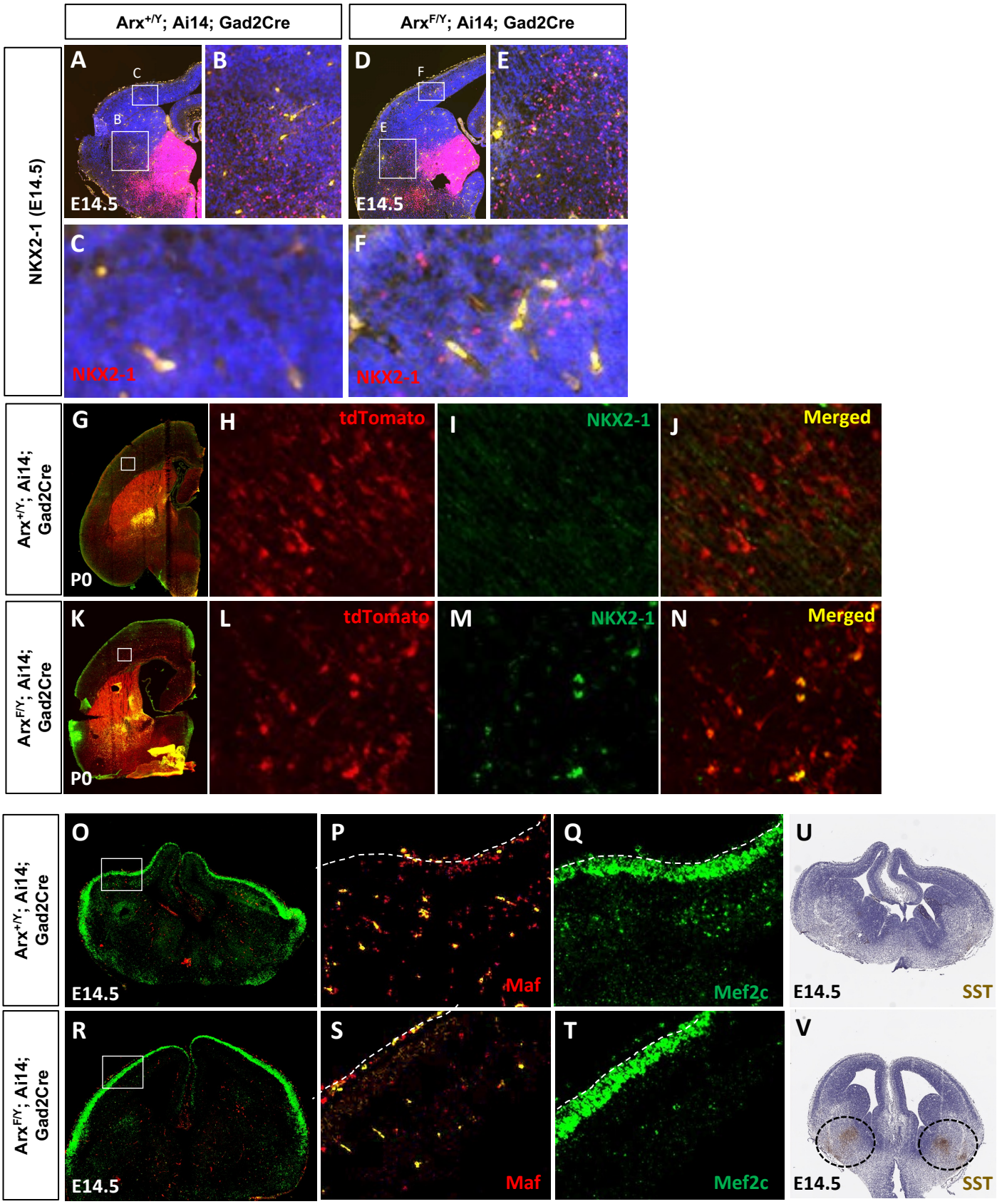

Arx+/Y; Ai14; Gad2Cre
ArxF/Y; Ai14; Gad2Cre
NKX2-1 (E14.5)
A
B
C
D
E
F
E14.5
NKX2-1
Arx+/Y; Ai14; Gad2Cre
ArxF/Y; Ai14; Gad2Cre
G
H
I
J
K
L
M
N
P0
tdTomato
NKX2-1
Merged
O
P
Q
U
R
S
T
V
E14.5
Maf
Mef2c
SST

Figure 5

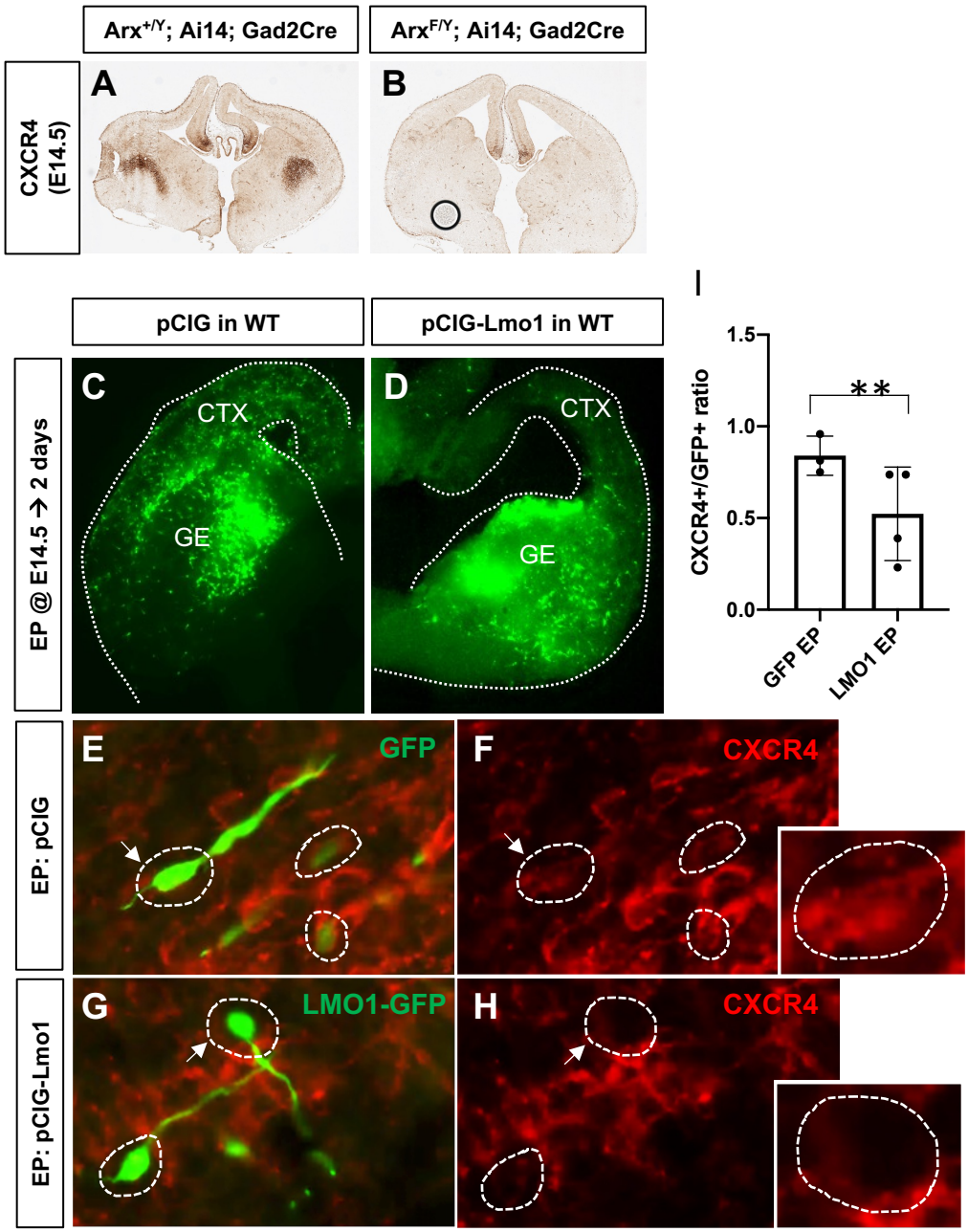

Arx+/Y; Ai14; Gad2Cre
ArxF/Y; Ai14; Gad2Cre
CXCR4 (E14.5)
A
B
pCIG in WT
pCIG-Lmo1 in WT
EP @ E14.5 → 2 days
C
D
CTX
GE
I
CXCR4+/GFP+ ratio
1.5
1.0
0.5
0.0
**
GFP EP
LMO1 EP
EP: pCIG
E
GFP
F
CXCR4
EP: pCIG-Lmo1
G
LMO1-GFP
H
CXCR4

Figure 6

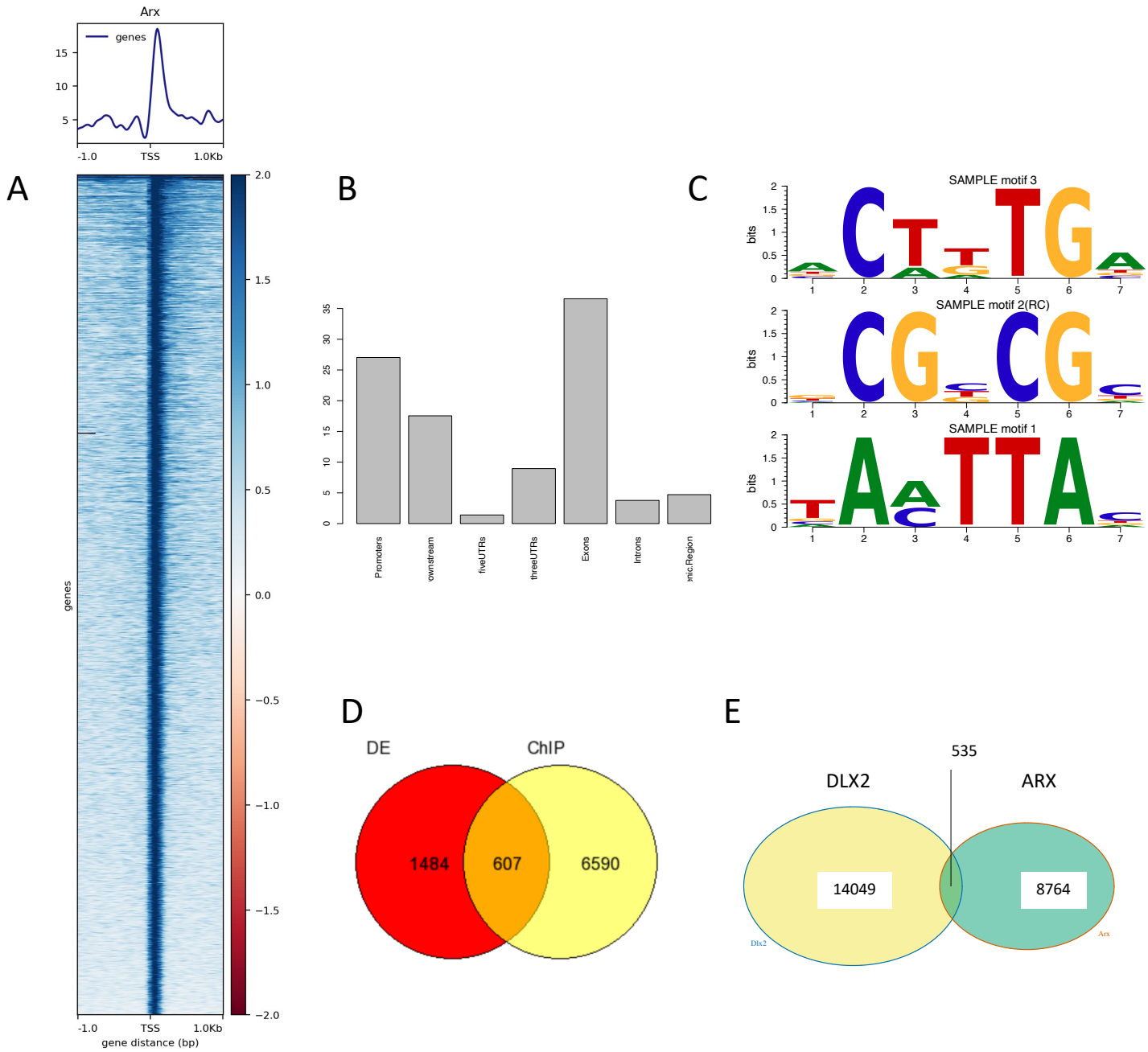

Arx
genes
15
10
5
-1.0
TSS
1.0Kb
A
genes
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
−1.5
−2.0
-1.0
TSS
1.0Kb
gene distance (bp)
B
35
30
25
20
15
10
5
0
Promoters
fiveUTRs
threeUTRs
Exons
Introns
C
SAMPLE motif 3
SAMPLE motif 2(RC)
SAMPLE motif 1
bits
D
DE
ChIP
1484
607
6590
E
535
DLX2
ARX
14049
8764
Dlx2
Arx

Figure 7

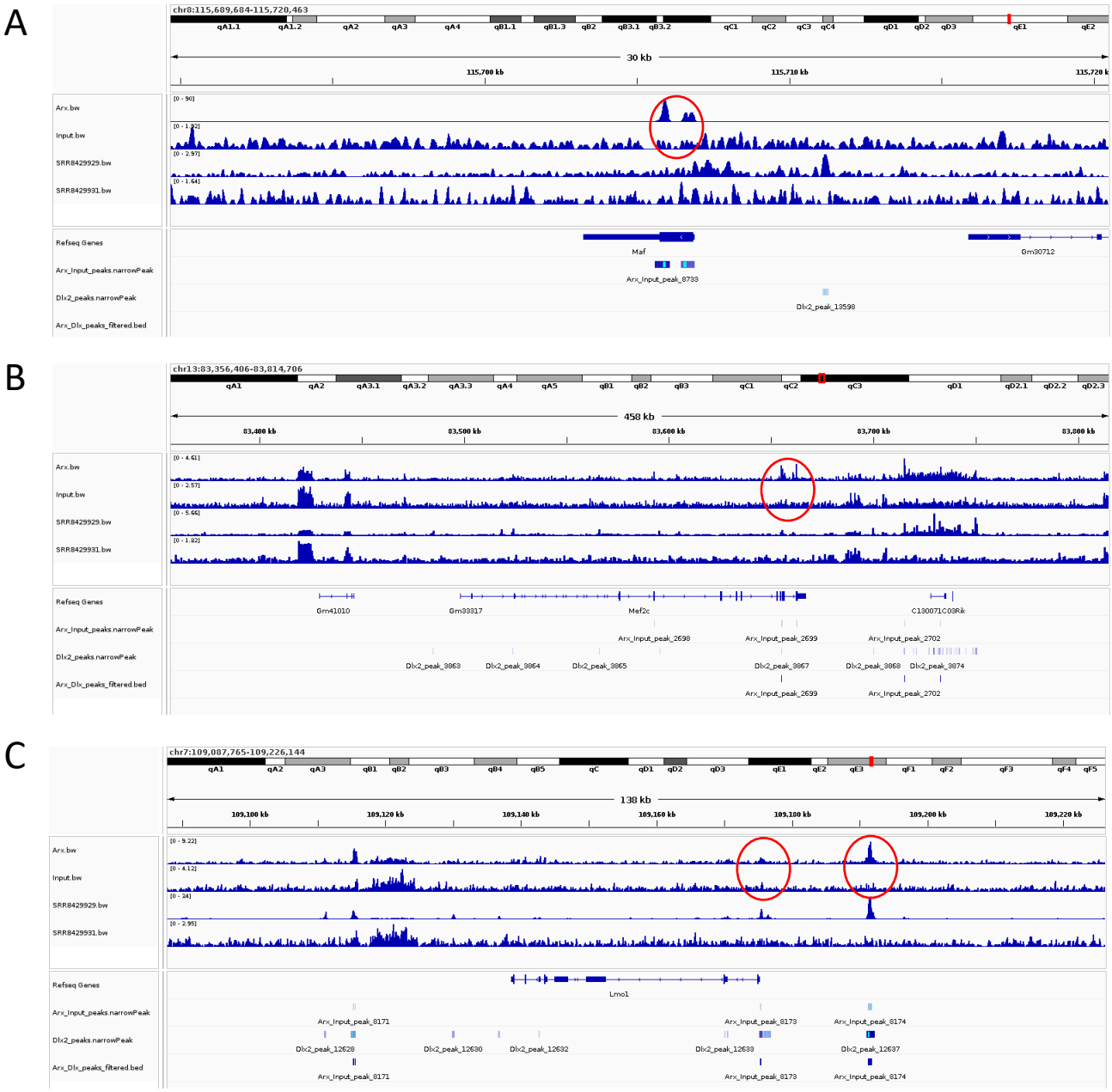

A
chr8:115,689,684-115,720,463
30 kb
115,700 kb
115,710 kb
Arx.bw
Input.bw
SRR8429929.bw
SRR8429931.bw
Refseq Genes
Arx_Input_peaks.narrowPeak
Dlx2_peaks.narrowPeak
Arx_Dlx_peaks_filtered.bed
Maf
Gm30712
Arx_Input_peak_8733
Dlx2_peak_13598
B
chr13:83,356,406-83,814,706
458 kb
83,400 kb
83,500 kb
83,600 kb
83,700 kb
83,800 kb
Gm41010
Gm38317
Mef2c
C130071C03Rik
Arx_Input_peak_2698
Arx_Input_peak_2699
Arx_Input_peak_2702
Dlx2_peak_3863
Dlx2_peak_3864
Dlx2_peak_3865
Dlx2_peak_3867
Dlx2_peak_3868
Dlx2_peak_3874
C
chr7:109,087,765-109,226,144
138 kb
109,100 kb
109,120 kb
109,140 kb
109,160 kb
109,180 kb
109,200 kb
109,220 kb
Lmo1
Arx_Input_peak_8171
Arx_Input_peak_8173
Arx_Input_peak_8174
Dlx2_peak_12628
Dlx2_peak_12630
Dlx2_peak_12632
Dlx2_peak_12633
Dlx2_peak_12637

Figure 8

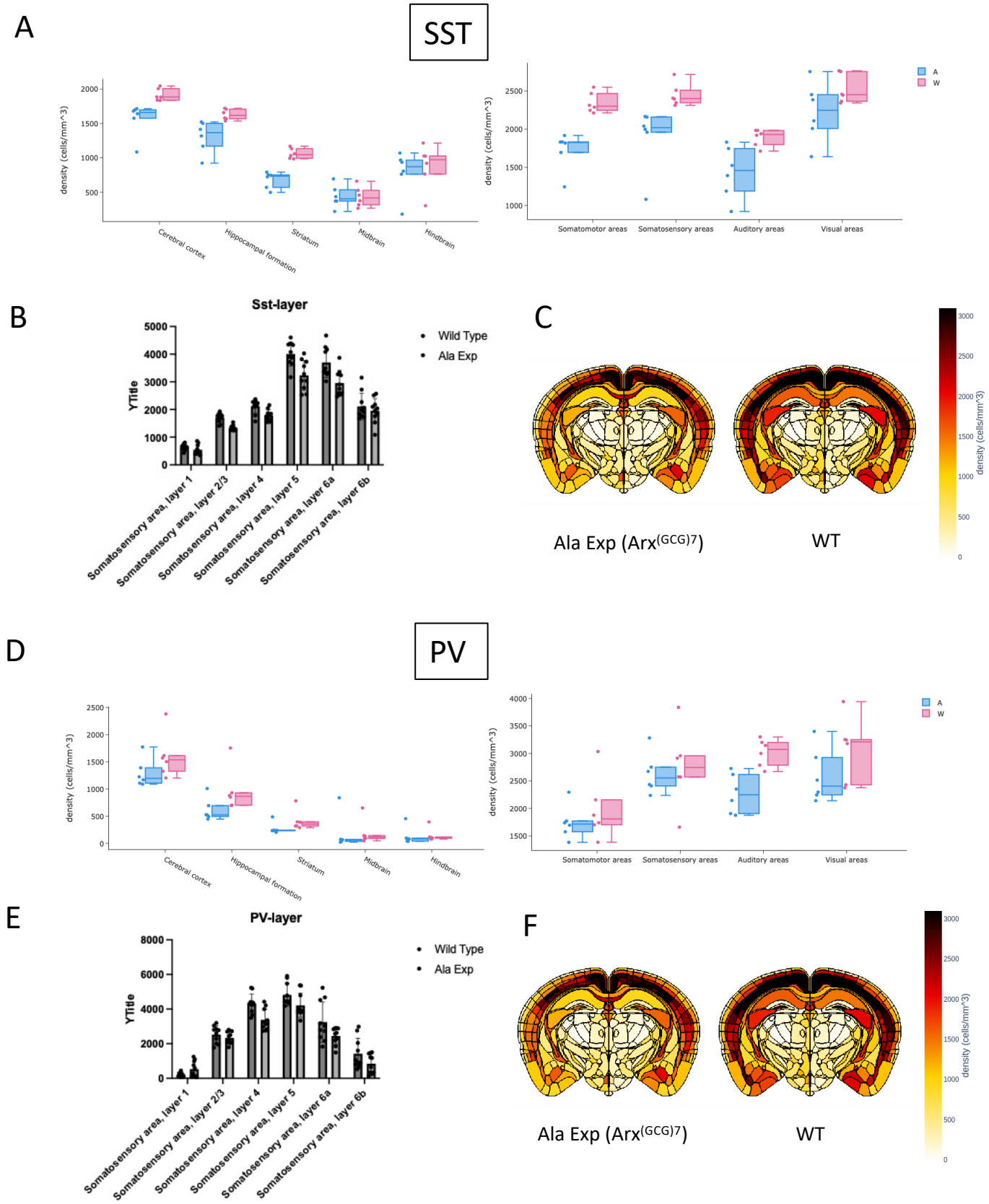
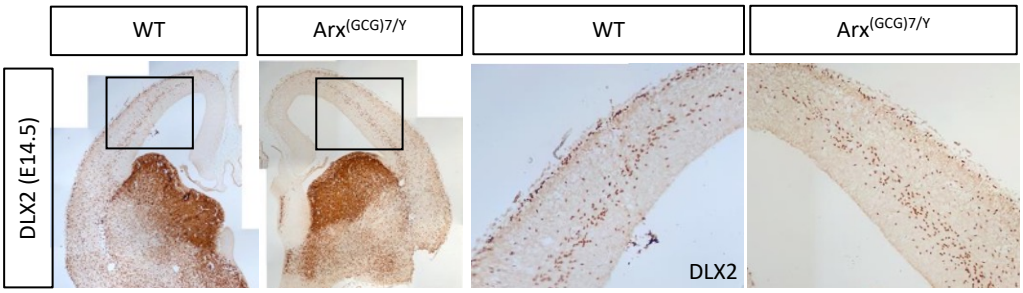
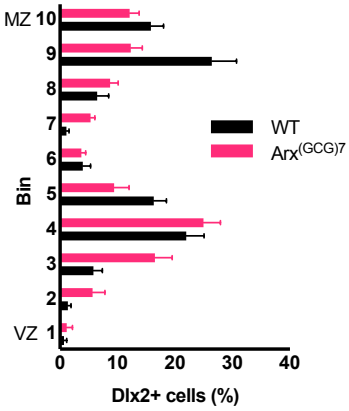

A
SST
density (cells/mm^3)
Cerebral cortex
Hippocampal formation
Striatum
Midbrain
Hindbrain
Somatomotor areas
Somatosensory areas
Auditory areas
Visual areas
A
W
B
Sst-layer
YTitle
Wild Type
Ala Exp
Somatosensory area, layer 1
Somatosensory area, layer 2/3
Somatosensory area, layer 4
Somatosensory area, layer 5
Somatosensory area, layer 6a
Somatosensory area, layer 6b
C
Ala Exp (Arx(GCG)7)
WT
density (cells/mm^3)
D
PV
E
PV-layer
F

Figure 9

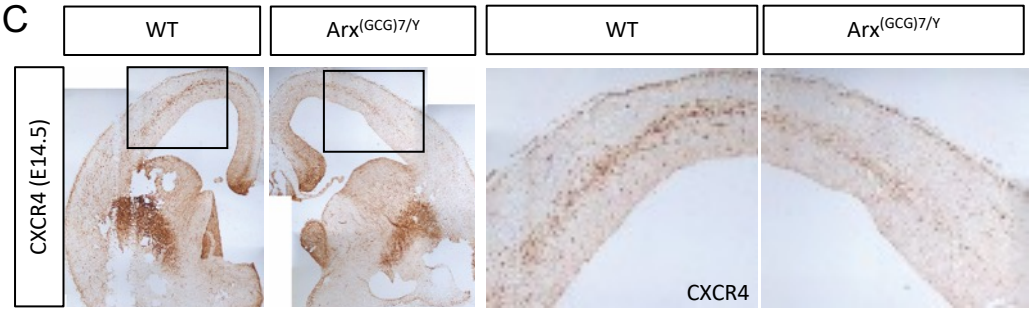
A



B



C

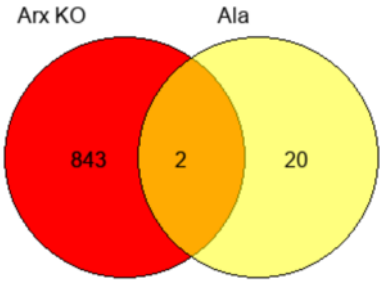


D

Upregulated in MT

Cadps2
Dlx1as?
Etv1
Kcnd2
Khdrbs2
Kitl
Lmo1
Lmo4
Mef2a
Nptx2
Srprb
Stac2
Syn2
Tox
Vldlr

Downregulated in MT



Cxcr4
Lhx8