Analysis of 14q12 microdeletions reveals novel regulatory loci for the neurodevelopmental disorder-related gene, *FOXG1*

Aishwarya Ramamurthy[1], Masha D Bandouil[1], Likhita Aluru[1], Esther Yoon[1], Nicholas Bodkin[1], Jennifer Z Cheng[1], Carina G Biar[1], Jeffrey D Calhoun[1], Gemma L Carvill[1,2] *

Affiliations:

1. Ken and Ruth Davee Department of Neurology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
2. Department of Pharmacology, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Abstract

Structural variants (SVs) are genetic variants that can vary widely in size, but are generally greater than 50bp. While there are many different types of SVs, deletions are perhaps the easiest to interpret given the loss of genetic material. As such up to 17% of neurodevelopmental disorders (NDDs) can be explained by SVs that disrupt coding regions causing gene dosage defects. Conversely, the contribution of non-coding SVs to the etiology of NDDs is unknown., From a few NDDs, and other mendelian disorders we know that non-coding SVs can perturb *cis*-regulatory elements (CREs) and disrupt downstream gene expression. In this study, we identified multiple deletions in the 14q12 region downstream of NDD-related gene *FOXG1* in individuals with epilepsy, microcephaly, intellectual disability (ID) and/or developmental delays. This gene encodes a transcription factor essential for cerebral corticogenesis, and *FOXG1* haploinsufficiency is associated with epilepsy and other NDDs. To test whether the 14q12 deletions alter *FOXG1* expression, we chose the minimum region overlapping these deletions (MRO), and modeled this deletion in HAP1s, a haploid line expressing *FOXG1.* Deletion of the MRO reduced *FOXG1* mRNA and protein levels, while altering FOXG1's native genomic interactions. We identified at least two novel putative *FOXG1* CREs in the MRO and propose multiple CREs coregulate *FOXG1*. Moreover, changes in the transcriptomic profiles seen with MRO loss overlap in part with *FOXG1* loss, indicating converging molecular pathways of *FOXG1* haploinsufficiency. These findings expand the scope of *FOXG1's* complex regulatory mechanisms, and more broadly, of regulatory SVs in NDD susceptibility. A comprehensive understanding of regulatory SVs will enable the identification of regulatory elements that are potential targets for gene-targeting therapies, including in FOXG1-related NDDs.

Introduction

Neurodevelopmental disorders (NDDs) are a heterogenous collection of developmental disorders that impair functions of the brain and the nervous system. These disorders have childhood onset and include, among others, intellectual disability (ID), autism spectrum disorder (ASD) and epilepsy [1–3]. It is now well established that disease-causing protein-coding variants in hundreds of genes account for up to 40% of NDDs [4–6]. This still leaves a large fraction of cases unexplained, rekindling interest in investigating the non-coding regulatory regions of the genome for causative variants in individuals with NDDs [7–10].

In particular, structural variants (SVs) are attractive candidates for understanding non-coding variation due to the thousands to millions of base pairs that can be disrupted and thus, in theory, easier interpretation. SVs, defined as alterations greater than 50 base pairs, include deletions, duplications, insertions, inversions, and translocations. SVs can overlap known NDD-related genes and cause these conditions through gene dosage effects. Coding SVs account for 15-17% of NDDs in general [11,12] and 10-12% of specific NDDs, such as the genetic epilepsies [13–15]. However, some SVs do not span coding regions but rather disrupt cis-regulatory elements (CREs) or higher order chromatin structures, including topologically associated domains (TADs). When SVs disrupt CREs, promoter-enhancer interactions can be perturbed, which affects gene transcription [16–18]. For example, noncoding single nucleotide variants and duplications in the *GATA2* regulatory region cause hereditary congenital facial paresis. These variants cause disease by aberrant prolonged *GATA2* expression which in turn impairs an essential neural fate switch during development [19]. SVs can also disrupt TAD boundaries, leading to an alteration of the native chromatin architecture, promoting non-canonical promoter-enhancer interactions and gene misexpression. TAD disruption by SVs has been reported in human limb malformations and congenital disorders [20,21]. Moreover, in NDDs, there is an enrichment of non-coding variants in the TAD encompassing NDD-related gene, *MEF2C*, and deletion of a proximal *MEF2C* TAD boundary resulted in reduced gene expression [22]. Collectively, these studies highlight an emerging role for non-coding SVs in the etiology of NDDs. Here we add to this burgeoning field by describing NDD-related microdeletions at the 14q12 locus flanking the gene, *FOXG1*.

FOXG1 encodes Forkhead Box G1, a key transcription factor in forebrain development that is highly dosage sensitive [23,24]. *FOXG1* expression is especially critical during corticogenesis, triencephalon specification and development, as well as maintenance of mature neurons [25–27]. Neurodevelopment is reliant not only on precise spatio-temporal *FOXG1* expression, but also on its cellular expression levels; both increased and decreased *FOXG1* expression is associated with poor neurological outcomes [24,28,29]. Heterozygous loss of *FOXG1* causes *FOXG1* syndrome – a congenital NDD characterized by epilepsy, ID, microcephaly, developmental delays, loss of motor skills and overlapping clinical features with Rett syndrome [30–32]. Several heterozygous *de novo* whole gene deletions, truncating variants and loss of function missense variants cause *FOXG1* syndrome. The pathogenic mechanism is thus haploinsufficiency, wherein one autosomal allele of *FOXG1* is functional, but insufficient to prevent disease [31,33]. In addition, *de novo* duplications of *FOXG1* result in increased gene dosage

and are associated with developmental epilepsy, ID, and severe speech impairment [34]. Interestingly, 14q12 deletions not encompassing *FOXG1* have been reported in individuals with overlapping clinical features of *FOXG1* haploinsufficiency [32,35–37].

Using DECIPHER, we identified several individuals with clinical features overlapping *FOXG1* haploinsufficiency phenotypes, without coding *FOXG1* variants, but rather downstream microdeletions that potentially disrupt CREs. We refine the regulatory landscape of *FOXG1* expression and demonstrate that deletion of a minimum region of overlap (MRO) reduced gene expression, likely contributing to haploinsufficiency in individuals with NDDs.

Results

Deletion of downstream *FOXG1* 14q12 minimum region of overlap leads to reduced *FOXG1* expression

We identified 9 microdeletions in the 14q12 region downstream of *FOXG1* described in DECIPHER. These deletions were reported in individuals (6 females and 3 males) with ID (n=8), seizures (n=3) and microcephaly (n=6). (**Figure 1 and Table 1**). Most of these deletions were *de novo* (patients 2-7); one deletion was maternally inherited (patient 9) and two were of unknown inheritance (patient 1 and 8). None of these deletions encompassed any other gene except *PRKD1.* This gene is completely tolerant to loss of function in the general population with a pLI score of 0 (gnomAD, [38]) and is not associated with NDDs [35]. Thus, we hypothesized that disruption of CREs in the region likely contributed to the phenotype via disruption of *FOXG1* expression. We selected the minimum region of overlap (MRO) between these deletions for modeling in the HAP1 cell line. This region (hg38: chr14:29,704,736-29,803,897) spans ~0.1 Mb and is ~0.93 Mb downstream of *FOXG1*. The MRO was examined for variants in the general population using gnomAD and no large deletions were observed in this locus (**Figure S1**).

We selected the HAP1 cell line due to its haploid nature, which facilitates genome editing, as well as the endogenous expression of *FOXG1* (**Figure 2**). We successfully deleted the MRO using dual flanking guides in HAP1s (14q12-MRO-Del); the *FOXG1* coding region was similarly deleted as a positive control ($FOXG1^{KO}$). We verified the genotypes of 14q12-MRO-Del (deletion of hg38: chr14:29,704,715-29,803,900) and $FOXG1^{KO}$ (deletion of hg38: chr14:28767269-28768758) using Sanger sequencing, and in the 14q12-MRO-Del confirmed the absence of major off-target SVs using genome sequencing (**Figure S2A-B).** As the MRO encompasses exon 2 as well as flanking intron of *PRKD1*, an expected loss of *PRKD1* expression was observed in 14q12-MRO-Del (**Figure S2C**).

Using qRT-PCR, we demonstrated complete loss of *FOXG1* mRNA expression in $FOXG1^{KO}$ and a 68% reduction in the 14q12-MRO-Del line (p-value 0.0006, Welch's T test) (**Figure 2A**). At the protein level, western blotting revealed similar loss of protein in $FOXG1^{KO}$ and a 45% reduction in the 14q12-MRO-Del (p-value 0.0005, Welch's T test) (**Figure 2B**). Collectively these results demonstrate that deleting the MRO reduces *FOXG1* expression, though not to the same degree as deletion of the complete coding region.

Deletion of the MRO locus alters *FOXG1's* native contacts with downstream CREs

To examine the causal mechanism of reduced *FOXG1* expression with MRO deletion, we examined chromatin interactions of the *FOXG1* locus with downstream CREs in wildtype and 14q12-MRO-Del lines using Unique Molecular Identifier (UMI)-4C. We anchored the viewpoint proximal to *FOXG1's* promoter (**Table S4**) and identified 9997 UMIs in WT and 6805 UMIs in 14q12-MRO-Del. UMI-4C contact intensity profiles were generated which highlight multiple promoter – CRE interactions across the 14q12 region. The overwhelming majority of contacts lie downstream of *FOXG1*, with very few contacts, and no peaks upstream. Moreover, in both HAP1s and neural cell lineages, peaks of UMI-4C interactions correlate with CREs defined by the enhancer histone modification, H3K27Ac, as well as VISTA enhancers and ENCODE Candidate CREs (cCREs) [39–44] (**Figure 3A**).

As expected, in the 14q12-MRO-Del, physical genomic interactions were lost between the *FOXG1* promoter and the MRO locus (**Figure 3A/B**). The 14q12-MRO-Del is expected to lead to the deletion of at least two, but possibly more enhancer elements, as defined by H3K27Ac peaks in the HAP1 cell line (**Figure 3B**). Moreover, there is an overlap of these HAP1 enhancer elements, with those present in stem cell derived neural precursor lines, as well as human brain (**Figure 3B**), suggesting that this locus is also regulatory in a neuronal context. Loss of these enhancer elements is likely what underpins reduced *FOXG1* expression in the 14q12-MRO-Del. In addition, native *FOXG1* genomic interactions were altered when the MRO was deleted, and significant differences in UMI contacts between WT and 14q12-MRO-Del were observed at several loci in the 14q12 region (**Figure S3**). Collectively, deleting the MRO locus impacted the physical genomic interaction of the *FOXG1* promoter and downstream CREs, including deletion of putative *FOXG1* enhancers at the MRO and enhanced contacts across the region; these alterations likely underpin the reduced *FOXG1* expression observed with MRO loss.

Gene expression analysis identifies shared DEGs between 14q12-MRO-Del and *FOXG1*^{KO}

Given *FOXG1's* function as a transcription factor, we evaluated the effects of *FOXG1* downregulation on overall gene expression changes using RNA-seq. We identified 1036 differentially expressed genes (DEGs) in 14q12-MRO-Del, and 517 DEGs in FOXG1^{KO}, compared to WT. Furthermore, we identified 86 DEGs that overlapped between the two datasets. The shared DEGs included genes involved in cell cycle regulation (*CCND1, BMP7)* [45–47] as well as those implicated in epilepsy (*FOSL2, EVAVL3, EPHA5*, *OLFM3, ARHGAP4)* [48–52], ID/microcephaly (*XYLT1, COLEC11, CADPS2)* [53–55] and other NDDs (*KY, ADGRL3, TCEAL9)* [56–58] **(Figure 4, Table S5).** Functional enrichment analysis of DEGs revealed a few shared molecular functions and biological pathways between 14q12-MRO-Del and FOXG1^{KO} (**Figure S4**). In essence, loss of the MRO and the subsequent reduction of *FOXG1* expression resulted in a transcriptome distinct from that of the wildtype, with a subset of DEGs shared with the *FOXG1* coding loss, suggesting shared downstream effects of loss of this transcription factor.

Discussion

FOXG1 is a dosage sensitive gene where both over- and under expression are associated with epilepsy and NDDs [23,24,59]. SVs, especially microdeletions in the 14q12 region downstream of *FOXG1* are reported in several individuals with clinical phenotypes that resemble *FOXG1* haploinsufficiency [32,35–37]. Whether the observed phenotypes are a result of the deletions was unclear, and the complex mechanisms of *FOXG1* regulation by these microdeletions are not fully understood. We identified nine 14q12 microdeletions downstream of *FOXG1* in DECIPHER in individuals with epilepsy, ID and/or microcephaly, and having no other clear causal variants. The deletions ranged from ~151.8 kb to 1.17Mb in size and were of mostly unspecified pathogenicity. We identified a shared 0.1Mb region (MRO), deleted in all patients, nearly 0.93 Mb distal to *FOXG1* that regulates *FOXG1* expression.

Deleting the MRO decreased *FOXG1* expression by more than half (68% reduction of mRNA, 45% reduction of protein) in the haploid HAP1 cell line, suggesting that the MRO is regulatory in nature. To determine what drives this reduced expression, we defined *FOXG1*'s physical genomic interaction landscape using the targeted and quantitative chromatin capture. We show that the *FOXG1* promoter interacts with the MRO, albeit with relatively low contact intensities, and these interactions are eliminated with the deletion of MRO. Along with ENCODE cCRE data, enhancer annotations from VISTA, and H3K27Ac occupancy data from HAP1 and neural lines, we identified putative CREs in the MRO that promote expression of *FOXG1*. We posit that deletion of these CREs leads to the reduction of *FOXG1* expression in the HAP1 cell line and by extension, likely contributes to reduced *FOXG1* expression in individuals with 14q12 deletions.

While *FOXG1* expression was considerably downregulated when the MRO was deleted, it was not fully lost, indicating that multiple enhancers outside the MRO regulate expression. Indeed, we observed enhanced chromatin contacts across the region in the 14q12-MRO-Del (**Figure S3**). Moreover, the MRO from this study lies within a larger regulatory region of *FOXG1* designated as the SRO (smallest region of overlap) (**Figure 3A** [35]). This region was previously defined by *de novo* 14q12 deletions downstream of *FOXG1* in three individuals with *FOXG1* syndrome related phenotypes including epilepsy, ID, developmental delays and severe postnatal microcephaly (patient 3, 5 and 7 [35]). The SRO is a complex region with multiple regulatory elements. A critical regulatory region within the SRO [37] consists of at least five Vista enhancers, hs1193, hs1201, hs1062, hs555 [35] and hs598 [60], predicted to regulate *FOXG1* (**Figure 3A**). Moreover, a human embryonic stem cell specific TAD boundary, hESCb_2291 [20,37] also resides within this critical regulatory region (**Figure 3A**). In our present study, we refined the SRO with a focus on the MRO based on the distal deletions in patient 8 and 9 (**Figure 1**). Of note, these two individuals, with deletions that encompass less of the SRO and associated regulatory elements, present only with ID. In contrast, individuals with more complete deletion of regulatory elements (patient 1-7) generally present with all features of *FOXG1* haploinsufficiency: ID, seizures and microcephaly. This, taken together with the reduction but not complete loss of *FOXG1* expression by deletion of the MRO suggests that disruption of multiple CREs is required to completely ablate *FOXG1* expression. Moreover, the number of disrupted CREs may underpin genotype-

phenotype relationships and as genome sequencing facilitates more widespread and smaller SV detection, these relationships may become more apparent.

The modulation of *FOXG1* via CREs in the SRO and across the 14q12 region that show physical interactions are likely to be important not only in ID/epilepsy, but also neurological conditions more broadly. Indeed, there are multiple GWAS loci in the region that have been associated with increased risk for ADHD and schizophrenia (**Figure 3C** [61]). Furthermore, deletion of a ~500bp region flanking schizophrenia-associated SNP rs1191551 reduced *FOXG1* expression by ~20% in HEK293T cells [62]. Further, a luciferase assay also demonstrated reduced activity of the risk rs1191551 allele, suggesting that this particular variant can modulate FOXG1 expression. Collectively, these studies highlight the need to systematically assess both disease-associated SVs and single nucleotide changes across this important regulatory region in neurological disease broadly.

Overall, our work demonstrates that there is enhancer redundancy across the 14q12 locus, where multiple enhancers exist for a single gene. This is a common phenomenon in mammalian development where key developmental genes that are dosage sensitive are tightly regulated by multiple CREs [63]. Such a redundancy in gene regulation is considered to function as a protective mechanism against deleterious genomic aberrations [64]. For instance, individual deletion of each of ten loci critical for mouse limb development resulted in normal limb phenotypes, whereas pairwise deletions of these loci resulted in noticeably different limb formations [64]. Redundant, or shadow enhancers regulate the same target gene and overlap in part or full, in the developmental space and time to drive expression patterns; they work in complex ways to buffer gene expression and fine tune gene dosage [63]. *FOXG1* is highly regulated for its spatiotemporal expression during corticogenesis. While multiple regulatory features in the critical regulatory region govern *FOXG1* expression, we extend the gene's regulatory region to the MRO and propose that shadow enhancers for *FOXG1* exist across the SRO region including the MRO. Furthermore, the effect of shadow enhancers on gene expression is not always additive; depending on the availability of transcription factors at each of the enhancers and the buffering requirement of the cell, the effects can be superadditive, subadditive, or even repressive [63,65]. This could potentially explain the differences in severity of *FOXG1* haploinsufficiency phenotypes seen in individuals with identical deletion breakpoints (patient 2, 3 from this study and Case 1 [36]).

In addition to neurodevelopmental programming, *FOXG1* is also involved in the transcriptional regulation of numerous other key regulators in the developmental pathways [24]. *FOXG1* modulates BMP signaling and the cell cycle, ensuring cells remain in proliferative states [28,47]. Since *FOXG1* levels were decreased with MRO loss, we evaluated its downstream effects on global transcription through gene expression analysis. Several genes that were differentially expressed with MRO loss were also differentially expressed with *FOXG1* loss, notably, BMP family protein *BMP7* [47], *CCND1*, a G1 phase cyclin [46] and *EGR1*, a transcription factor and member of the EGR-*FOXG1* coupled regulatory network [66]. These results indicate that the impact of MRO loss extends beyond *FOXG1* downregulation and leads to more complex

perturbations of downstream homeostatic and developmental pathways, some of which intersect with those seen with complete *FOXG1* loss.

In summary, our work refines the regulatory landscape of the *FOXG1* locus, and we identify a novel ~0.1 Mb regulatory region downstream of *FOXG1* in the 14q12 region. Within this MRO, we define putative *FOXG1* CREs and emphasize that this locus exhibits both enhancer redundancy as well as shadow enhancer activity. The complex regulatory nature of the *FOXG1* locus is reflected in its importance in disease etiology. We define the causal mechanisms of a subset of clinical features of *FOXG1* syndrome and posit that deletion of multiple CREs is necessary to cause *FOXG1* haploinsufficiency. Overall, this critical regulatory locus should be examined across both NDDs and neuropsychiatric disease for causative and contributory variants and presents unique opportunities for gene-targeting therapies via modulation of *FOXG1* expression.

Methods

Selection of candidate 14q12 microdeletions, HAP1 maintenance and cell line generation

We used DECIPHER (DatabasE of genomiC varIation and Phenotype in Humans using Ensembl Resources) [67] to identify 14q12 microdeletions downstream of the NDD-related gene, *FOXG1* in individuals with epilepsy and related NDDs. Deletions within *FOXG1's* TAD that excluded *FOXG1* and extended up to hg38: chr14:30,550,000 (approaching the distal TAD boundary limit) were considered [68]. Larger deletions spanning multiple genes were excluded as gene dosage effects from multiple encompassed genes are likely to confound patient phenotypes.

Human eHAP (fully haploid engineered HAP1) cells from Horizon Discovery BioSciences (C669) were maintained in DMEM containing 4.5 g/l glucose (Thermo Fisher Scientific) supplemented with 10% Tet-approved FBS (Atlanta Biologicals) and 1% penicillin/streptomycin (Thermo Fisher Scientific). The cells were tested for mycoplasma contamination and all experiments were performed within 3 to 10 passages.

CRISPR genome editing was used to create 14q12-MRO-Del and FOXG1^{KO} eHAP lines. Guide RNAs were designed to target each of the two breakpoints of the MRO and cloned into a Cas9-GFP vector (PX458, Addgene #48138) or a Cas9-mCherry vector (a PX458 derivative engineered in-house) (**Table S1** for guide RNA sequences). Dual guide electroporation was performed using the Neon Transfection System. 24-48h post transfection, FACS was performed for single cell clonal isolation and expansion of GFP and mCherry double positive cells. Clonal lines were genotyped by PCR, followed by Sanger sequencing, to detect the presence of the deletion (**Table S1** for primer sequences). Whole genome sequencing (WGS) was performed to further confirm the intended genotype. Moreover, we used the MetaSVs pipeline to detect the presence of any large unintended SVs, and any lines harboring such variants were excluded [69]. FOXG1^{KO} lines were similarly generated, with dual guide RNAs targeting the start and

end of *FOXG1*'s coding region (**Table S1**). Clones of 14q12-MRO-Del and FOXG1^{KO}, along with the parental/unedited line (WT Clone A) and a wildtype clone from the CRISPR screen (WT Clone B) as isogenic controls, were selected for experimentation (**Table S2**).

***FOXG1* mRNA and protein quantification**

Quantitative reverse transcription PCR was performed to assay *FOXG1* mRNA levels. RNA was extracted and purified from HAP1s in culture using TRIzol reagent (Invitrogen #15596026). qRT-PCR reactions were performed using Biorad iTaq Universal One-Step RT-qPCR Kit (1725150). 25ng RNA per reaction was used and triplicate qRT-PCR reactions were carried out for each target (**Table S1**) on a Biorad CFX384 Real-Time PCR Detection system. Raw Ct values were obtained for each sample. Ct values for *FOXG1* were transformed by first normalizing to EAR and then normalizing to WT to calculate relative fold change of *FOXG1* [70]. Two Sample T-Test (Welch's T-Test) or Wilcoxon Rank Sum Test were performed to test statistical significance between samples.

Western blotting was performed to assay FOXG1 protein levels. Whole cell lysates were prepared using RIPA Lysis and Extraction Reagent (Thermo Scientific 89900) in the presence of protease inhibitors (cOmplete™, Mini, EDTA-free Protease Inhibitor Cocktail, Sigma 11836170001). 20 µg of protein per well was separated on NuPage 3-8% Tris-Acetate precast gels (Invitrogen) and transferred onto a PVDF membrane (Millipore). Immunoblotting was performed using primary antibodies (**Table S3**) and detected with peroxidase-conjugated secondary antibodies (**Table S3**). This was followed by detection with ECL peroxidase substrate (Cytiva Amersham RPN2232). Protein band intensities on immunoblots were quantified by densitometric analysis. Two Sample T-Test (Welch's T-Test) was performed to test statistical significance between samples.

Locus specific chromatin capture

UMI-4C was performed for targeted and quantitative capture of *FOXG1*'s genomic interactions with the following specifications [22]. *FOXG1* viewpoint was defined as a 1kb window at the 5' end of the gene. Crosslinked chromatin from 5 x 10^6 cells were used per sample. UMI-4C libraries were constructed by enriching the 4C library for *FOXG1* contacts using viewpoint specific primers (**Table S1**) [71]. Two unique barcode adapters per sample were utilized to generate five UMI-4C reactions per adapter, which were pooled to make a final UMI-4C library. UMI-4C experiments were performed at least twice per clone. Libraries were subject to 75 bp paired end sequencing on the Illumina MiniSeq platform at a sequencing depth of 1-2M reads per sample.

Raw read quality was assessed using FastQC [72]. The fastq files from sequencing were analyzed using the UMI-4C R Package as described [22,73]. Since the UMI-4C library is made up of adapter ligated DNA fragments that consist of both the bait and its spatially proximal genomic region, it generates sequencing reads that map to both these loci. To quantify these spatially proximal genomic interactions, the R package

transforms the raw sequencing reads to unique genomic interactions (UMI counts) per genomic restriction fragment and the resulting data is saved as genomic tracks. From these tracks, we extracted viewpoint-specific UMI counts for a given region of interest using the package. UMI counts were normalized to the total UMI counts within the region of interest and a smoothed interaction profile was produced using default parameters in the package. In this way, we generated UMI-4C profiles for the 14q12 region of interest (hg38: chr14: 28,965,045-30,631,045) for WT and 14q12-MRO-Del samples. Differential genomic interactions between samples were evaluated using the UMI4Cats R package [74]. Significant differential contacts between WT and 14q12-MRO-Del were inferred by the Wald Test using the waldUMI4C() function with FDR adjusted *p*-value of 0.05.

Analysis of H3K27Ac data across 14q12 *FOXG1* regulatory region

H3K27Ac ChIP-seq data was assessed from published data in HAP1s (GSE167721), adult brain hippocampus (GSM1112791) and temporal lobe (GSM772995), the embryonic stem cell line H1 (GSM466732) as well as neural precursor cells (NPCs) derived from H1 (GSM753429) and ganglionic eminence neurospheres derived from H1 (GSM1127083).

Bulk RNA-Seq library prep and data analysis

RNA was extracted and purified from HAP1s in culture by the guanidinium thiocyanate-phenol-chloroform extraction method using TRIzol™ reagent (Invitrogen #15596-026). Purified RNA samples were assayed on a TapeStation 4150 (Agilent) to determine the concentration and RIN (RNA Integrity Number) values. RNA samples with RIN > 9 were sent to Novogene for mRNA library preparation and subject to 150 bp paired end sequencing on a NovaSeq platform at a sequencing depth of at least 20 M reads per sample.

Raw read quality was assessed using FastQC [72]. Reads were trimmed to remove adapter sequences using HTStream, aligned to the human reference genome GRCh38/hg38 and GENCODE v38 using STAR [75]. The read count matrix was generated using GENCODE v38 using Rsubread [76]. Differentially expressed genes (DEGs) were detected using DESeq2 using an adjusted p-value threshold of 0.05 and log2Fold change greater/less than absolute value of 1 [77]. Functional enrichment analysis was performed using gene ontology in the clusterProfiler package using a Benjamini-Hochberg corrected p-value threshold of 0.05 [78–81].

Data availability

Data for the study are available at NCBI GEO database with accession number GEO: GSEXXXXXX.

Author contributions: AR, MB, LA, EY, JZC and CB generated the experimental data. JDC assisted with experiments. AR, NB, JDC, GLC performed data analysis, and interpretation of results. GLC and AR conceived the presented idea, developed the theory, wrote this manuscript and generated figures for this manuscript. GLC supervised the project.

Acknowledgements: Our work was supported by the following funding sources: the Chicago Biomedical Consortium Catalyst Award to GLC and the American Epilepsy Society Predoctoral Fellowship to AR.

Competing interests: All authors declare no competing interests

References

[1] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders. Diagnostic and Statistical Manual of Mental Disorders 2022. https://doi.org/10.1176/APPI.BOOKS.9780890425787.

[2] Bozzi Y, Casarosa S, Caleo M. Epilepsy as a Neurodevelopmental Disorder. Front Psychiatry 2012;3:19. https://doi.org/10.3389/FPSYT.2012.00019.

[3] Mullin AP, Gokhale A, Moreno-De-Luca A, Sanyal S, Waddington JL, Faundez V. Neurodevelopmental disorders: mechanisms and boundary definitions from genomes, interactomes and proteomes. Transl Psychiatry 2013;3. https://doi.org/10.1038/TP.2013.108.

[4] McRae JF, Clayton S, Fitzgerald TW, Kaplanis J, Prigmore E, Rajan D, et al. Prevalence and architecture of de novo mutations in developmental disorders. Nature 2017 542:7642 2017;542:433–8. https://doi.org/10.1038/nature21062.

[5] van Slobbe M, van Haeringen A, Vissers LELM, Bijlsma EK, Rutten JW, Suerink M, et al. Reanalysis of whole-exome sequencing (WES) data of children with neurodevelopmental disorders in a standard patient care context. Eur J Pediatr 2024;183:345–55. https://doi.org/10.1007/S00431-023-05279-4/TABLES/4.

[6] Jansen S, Vissers LELM, de Vries BBA. The Genetics of Intellectual Disability. Brain Sciences 2023, Vol 13, Page 231 2023;13:231. https://doi.org/10.3390/BRAINSCI13020231.

[7] Mefford HC, Yendle SC, Hsu C, Cook J, Geraghty E, McMahon JM, et al. Rare copy number variants are an important cause of epileptic encephalopathies. Ann Neurol 2011;70:974. https://doi.org/10.1002/ANA.22645.

[8] Mullen SA, Carvill GL, Bellows S, Bayly MA, Berkovic SF, Dibbens LM, et al. Copy number variants are frequent in genetic generalized epilepsy with intellectual disability. 2013.

[9] Short PJ, McRae JF, Gallone G, Sifrim A, Won H, Geschwind DH, et al. De novo mutations in regulatory elements in neurodevelopmental disorders. Nature 2018 555:7698 2018;555:611–6. https://doi.org/10.1038/nature25983.

[10] Turner TN, Eichler EE. The Role of De Novo Noncoding Regulatory Mutations in Neurodevelopmental Disorders. Trends Neurosci 2019;42:115–27. https://doi.org/10.1016/J.TINS.2018.11.002/ASSET/04FBDE66-24A1-4649-AC1B-2CEABF84D314/MAIN.ASSETS/GR4.JPG.

[11] Kaminsky EB, Kaul V, Paschall J, Church DM, Bunke B, Kunig D, et al. An evidence-based approach to establish the functional and clinical significance of copy number variants in intellectual and developmental disabilities. Genet Med 2011;13:777–84. https://doi.org/10.1097/GIM.0B013E31822C79F9.

[12] D’haene E, Vergult S. Interpreting the impact of noncoding structural variation in neurodevelopmental disorders. Genetics in Medicine 2021;23:34–46. https://doi.org/10.1038/S41436-020-00974-1.

[13] Hebbar M, Mefford HC. Recent advances in epilepsy genomics and genetic testing. F1000Res 2020;9. https://doi.org/10.12688/F1000RESEARCH.21366.1.

[14] Helbig KL, Farwell Hagman KD, Shinde DN, Mroske C, Powis Z, Li S, et al. Diagnostic exome sequencing provides a molecular diagnosis for a significant proportion of patients with epilepsy. Genet Med 2016;18:898–905. https://doi.org/10.1038/GIM.2015.186.

[15] Retterer K, Juusola J, Cho MT, Vitazka P, Millan F, Gibellini F, et al. Clinical application of whole-exome sequencing across clinical indications. Genet Med 2016;18:696–704. https://doi.org/10.1038/GIM.2015.148.

[16] Cauwelier B, Dastugue N, Cools J, Poppe B, Herens C, De Paepe A, et al. Molecular cytogenetic study of 126 unselected T-ALL cases reveals high incidence of TCRβ locus rearrangements and putative new T-cell oncogenes. Leukemia 2006 20:7 2006;20:1238–44. https://doi.org/10.1038/sj.leu.2404243.

[17] Weischenfeldt J, Dubash T, Drainas AP, Mardin BR, Chen Y, Stütz AM, et al. Pan-cancer analysis of somatic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. Nature Genetics 2016 49:1 2016;49:65–74. https://doi.org/10.1038/ng.3722.

[18] Wang X, Xu J, Zhang B, Hou Y, Song F, Lyu H, et al. Genome-wide detection of enhancer-hijacking events from chromatin interaction data in rearranged genomes. Nature Methods 2021 18:6 2021;18:661–8. https://doi.org/10.1038/s41592-021-01164-w.

[19] Tenney AP, Di Gioia SA, Webb BD, Chan WM, de Boer E, Garnai SJ, et al. Noncoding variants alter GATA2 expression in rhombomere 4 motor neurons and cause dominant hereditary congenital facial paresis. Nature Genetics 2023 55:7 2023;55:1149–63. https://doi.org/10.1038/s41588-023-01424-9.

[20] Ibn-Salem J, Köhler S, Love MI, Chung HR, Huang N, Hurles ME, et al. Deletions of chromosomal regulatory boundaries are associated with congenital disease. Genome Biol 2014;15:423. https://doi.org/10.1186/S13059-014-0423-1/FIGURES/5.

[21] Lupiáñez DG, Kraft K, Heinrich V, Krawitz P, Brancati F, Klopocki E, et al. Disruptions of Topological Chromatin Domains Cause Pathogenic Rewiring of Gene-Enhancer Interactions. Cell 2015;161:1012–25. https://doi.org/10.1016/J.CELL.2015.04.004.

[22] Mohajeri K, Yadav R, D'haene E, Boone PM, Erdin S, Gao D, et al. Transcriptional and functional consequences of alterations to MEF2C and its topological organization in neuronal models. Am J Hum Genet 2022;109:2049–67. https://doi.org/10.1016/J.AJHG.2022.09.015.

[23] Li J, Chang H. W, Lai E, Parker E. J, Vogt P. K. The oncogene qin codes for a transcriptional repressor - PubMed. Cancer Res 1995;55:5540–4.

[24] Hettige NC, Ernst C. FOXG1 Dose in Brain Development. Front Pediatr 2019;7:482. https://doi.org/10.3389/FPED.2019.00482/BIBTEX.

[25] Xuan S, Baptista CA, Balas G, Tao W, Soares VC, Lai E. Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres. Neuron 1995;14:1141–52. https://doi.org/10.1016/0896-6273(95)90262-7.

[26] Dou C, Lee J, Liu B, Liu F, Massague J, Xuan S, et al. BF-1 Interferes with Transforming Growth Factor β Signaling by Associating with Smad Partners. Mol Cell Biol 2000;20:6201–11. https://doi.org/10.1128/MCB.20.17.6201-6211.2000/ASSET/74BF586E-AF6B-4CB5-BEA1-7DFD301C5DC4/ASSETS/GRAPHIC/MB1700132008.JPEG.

[27] Dastidar SG, Landrieu PMZ, D'Mello SR. FoxG1 Promotes the Survival of Postmitotic Neurons. Journal of Neuroscience 2011;31:402–13. https://doi.org/10.1523/JNEUROSCI.2897-10.2011.

[28] Hanashima C, Shen L, Li SC, Lai E. Brain Factor-1 Controls the Proliferation and Differentiation of Neocortical Progenitor Cells through Independent Mechanisms. Journal of Neuroscience 2002;22:6526–36. https://doi.org/10.1523/JNEUROSCI.22-15-06526.2002.

[29] Hettige NC, Peng H, Wu H, Zhang X, Yerko V, Zhang Y, et al. FOXG1 dose tunes cell proliferation dynamics in human forebrain progenitor cells. Stem Cell Reports 2022;17:475–88. https://doi.org/10.1016/J.STEMCR.2022.01.010.

[30] Ariani F, Hayek G, Rondinella D, Artuso R, Mencarelli MA, Spanhol-Rosseto A, et al. FOXG1 Is Responsible for the Congenital Variant of Rett Syndrome. The American Journal of Human Genetics 2008;83:89–93. https://doi.org/10.1016/J.AJHG.2008.05.015.

[31] Mitter D, Pringsheim M, Kaulisch M, Plümacher KS, Schröder S, Warthemann R, et al. FOXG1 syndrome: genotype–phenotype association in 83 patients with FOXG1 variants. Genetics in Medicine 2018 20:1 2017;20:98–108. https://doi.org/10.1038/gim.2017.75.

[32] Kortüm F, Das S, Flindt M, Morris-Rosendahl DJ, Stefanova I, Goldstein A, et al. The core FOXG1 syndrome phenotype consists of postnatal microcephaly, severe mental retardation, absent language, dyskinesia, and corpus callosum hypogenesis. J Med Genet 2011;48:396–406. https://doi.org/10.1136/JMG.2010.087528.

[33] Shoichet SA, Kunde SA, Viertel P, Schell-Apacik C, von Voss H, Tommerup N, et al. Haploinsufficiency of novel FOXG1B variants in a patient with severe mental retardation, brain malformations and microcephaly. Hum Genet 2005;117:536–44. https://doi.org/10.1007/S00439-005-1310-3/FIGURES/4.

[34] Brunetti-Pierri N, Paciorkowski AR, Ciccone R, Mina E Della, Bonaglia MC, Borgatti R, et al. Duplications of FOXG1 in 14q12 are associated with developmental epilepsy, mental retardation, and severe speech impairment. European Journal of Human Genetics 2011 19:1 2010;19:102–7. https://doi.org/10.1038/ejhg.2010.142.

[35] Allou L, Lambert L, Amsallem D, Bieth E, Edery P, Destrée A, et al. 14q12 and severe Rett-like phenotypes: new clinical insights and physical mapping of FOXG1-regulatory elements. Eur J Hum Genet 2012;20:1216–23. https://doi.org/10.1038/EJHG.2012.127.

[36] Ellaway CJ, Ho G, Bettella E, Knapman A, Collins F, Hackett A, et al. 14q12 microdeletions excluding FOXG1 give rise to a congenital variant Rett syndrome-like phenotype. European Journal of Human Genetics 2013 21:5 2012;21:522–7. https://doi.org/10.1038/ejhg.2012.208.

[37] Mehrjouy MM, Fonseca ACS, Ehmke N, Paskulin G, Novelli A, Benedicenti F, et al. Regulatory variants of FOXG1 in the context of its topological domain organisation. European Journal of Human Genetics 2017 26:2 2017;26:186–96. https://doi.org/10.1038/s41431-017-0011-4.

[38] Karczewski KJ, Francioli LC, Tiao G, Cummings BB, Alföldi J, Wang Q, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. Nature 2020 581:7809 2020;581:434–43. https://doi.org/10.1038/s41586-020-2308-7.

[39] Dunham I, Kundaje A, Aldred SF, Collins PJ, Davis CA, Doyle F, et al. An integrated encyclopedia of DNA elements in the human genome. Nature 2012;489:57–74. https://doi.org/10.1038/NATURE11247.

[40] Abascal F, Acosta R, Addleman NJ, Adrian J, Afzal V, Aken B, et al. Expanded encyclopaedias of DNA elements in the human and mouse genomes. Nature 2020 583:7818 2020;583:699–710. https://doi.org/10.1038/s41586-020-2493-4.

[41] Hitz BC, Lee J-W, Jolanki O, Kagda MS, Graham K, Sud P, et al. The ENCODE Uniform Analysis Pipelines. BioRxiv 2023. https://doi.org/10.1101/2023.04.04.535623.

[42] Luo Y, Hitz BC, Gabdank I, Hilton JA, Kagda MS, Lam B, et al. New developments on the Encyclopedia of DNA Elements (ENCODE) data portal. Nucleic Acids Res 2020;48:D882–9. https://doi.org/10.1093/NAR/GKZ1062.

[43] Visel A, Minovitsky S, Dubchak I, Pennacchio LA. VISTA Enhancer Browser--a database of tissue-specific human enhancers. Nucleic Acids Res 2007;35. https://doi.org/10.1093/NAR/GKL822.

[44] Kosicki M, Baltoumas FA, Kelman G, Boverhof J, Ong Y, Cook LE, et al. VISTA Enhancer browser: an updated database of tissue-specific developmental enhancers. Nucleic Acids Res 2025;53. https://doi.org/10.1093/NAR/GKAE940.

[45] Pedraza N, Rocandio D, Monserrat MV, Zammou B, Ortiz-Brugués A, Marfull-Oromí P, et al. Cytoplasmic expression of the cell cycle regulator cyclin D1 in radial glial progenitor cells modulates brain cortex development. BioRxiv 2024:2024.12.23.630056. https://doi.org/10.1101/2024.12.23.630056.

[46] Lottini G, Baggiani M, Chesi G, D'Orsi B, Quaranta P, Lai M, et al. Zika virus induces FOXG1 nuclear displacement and downregulation in human neural progenitors. Stem Cell Reports 2022;17:1683. https://doi.org/10.1016/J.STEMCR.2022.05.008.

[47] Hettige NC, Fleming P, Semenak A, Zhang X, Peng H, Hagel MD, et al. FOXG1 targets BMP repressors and cell cycle inhibitors in human neural progenitor cells. Hum Mol Genet 2023;32:2511. https://doi.org/10.1093/HMG/DDAD089.

[48] Cospain A, Rivera-Barahona A, Dumontet E, Gener B, Bailleul-Forestier I, Meyts I, et al. FOSL2 truncating variants in the last exon cause a neurodevelopmental disorder with scalp and enamel defects. Genetics in Medicine 2022;24:2475–86. https://doi.org/10.1016/J.GIM.2022.09.002.

[49] Wutikeli H, Xie T, Xiong W, Shen Y. ELAV/Hu RNA-binding protein family: key regulators in neurological disorders, cancer, and other diseases. RNA Biol 2025;22:1–11. https://doi.org/10.1080/15476286.2025.2471133.

[50] Liu TT, Xiao B, Li SY, Li GL, Lu XQ, Liu WP. [Changes in the expression of EphA5/ephrinA5 in the CA3 region of the hippocampus in rats with epilepsy and their role in the pathogenesis of temporal lobe epilepsy]. Zhongguo Dang Dai Er Ke Za Zhi 2017;19:1272–7. https://doi.org/10.7499/J.ISSN.1008-8830.2017.12.010.

[51] Tang S, Wang T, Zhang X, Guo Y, Xu P, Zeng J, et al. Olfactomedin-3 Enhances Seizure Activity by Interacting With AMPA Receptors in Epilepsy Models. Front Cell Dev Biol 2020;8:557009. https://doi.org/10.3389/FCELL.2020.00722/BIBTEX.

[52] Hu YY, Song W, Liu ZG, Ye XG, Zhang HW, Li X, et al. ARHGAP4 variants are associated with X-linked early-onset temporal lobe epilepsy. World J Pediatr 2024;20. https://doi.org/10.1007/S12519-024-00830-1.

[53] LaCroix AJ, Stabley D, Sahraoui R, Adam MP, Mehaffey M, Kernan K, et al. GGC Repeat Expansion and Exon 1 Methylation of XYLT1 Is a Common Pathogenic Variant in Baratela-Scott Syndrome. Am J Hum Genet 2019;104:35–44. https://doi.org/10.1016/j.ajhg.2018.11.005.

[54] Gilissen C, Hehir-Kwa JY, Thung DT, Van De Vorst M, Van Bon BWM, Willemsen MH, et al. Genome sequencing identifies major causes of severe intellectual disability. Nature 2014 511:7509 2014;511:344–7. https://doi.org/10.1038/nature13394.

[55] Bonora E, Graziano C, Minopoli F, Bacchelli E, Magini P, Diquigiovanni C, et al. Maternally inherited genetic variants of CADPS2 are present in autism spectrum disorders and intellectual disability patients. EMBO Mol Med 2014;6:795–809. https://doi.org/10.1002/EMMM.201303235.

[56] Arif B, Rasheed A, Kumar KR, Fatima A, Abbas G, Wohler E, et al. A novel homozygous KY variant causing a complex neurological disorder. Eur J Med Genet 2020;63:104031. https://doi.org/10.1016/J.EJMG.2020.104031.

[57] Menzel MHMM, Kohler A, Hudemann M, Jüngling J, Biskup S. Case Report of a Juvenile Patient with Autism Spectrum Disorder with a Novel Combination of Copy Number Variants in ADGRL3 (LPHN3) and Two Pseudogenes. Application of Clinical Genetics 2022;15:125–31. https://doi.org/10.2147/TACG.S361239.

[58] Albuainain F, Shi Y, Lor-Zade S, Hüffmeier U, Pauly M, Reis A, et al. Confirmation and expansion of the phenotype of the TCEAL1-related neurodevelopmental disorder. European Journal of Human Genetics 2024 32:3 2024;32:350–6. https://doi.org/10.1038/s41431-023-01530-6.

[59] Seltzer LE, Ma M, Ahmed S, Bertrand M, Dobyns WB, Wheless J, et al. Epilepsy and outcome in FOXG1-related disorders. Epilepsia 2014;55:1292–300. https://doi.org/10.1111/EPI.12648.

[60] Lu G, Zhang Y, Xia H, He X, Xu P, Wu L, et al. Identification of a de novo mutation of the FOXG1 gene and comprehensive analysis for molecular factors in Chinese FOXG1-

related encephalopathies. Front Mol Neurosci 2022;15:1039990. https://doi.org/10.3389/FNMOL.2022.1039990/FULL.

[61] Cerezo M, Sollis E, Ji Y, Lewis E, Abid A, Bircan KO, et al. The NHGRI-EBI GWAS Catalog: standards for reusability, sustainability and diversity. Nucleic Acids Res 2025;53:D998–1005. https://doi.org/10.1093/NAR/GKAE1070.

[62] Won H, De La Torre-Ubieta L, Stein JL, Parikshak NN, Huang J, Opland CK, et al. Chromosome conformation elucidates regulatory relationships in developing human brain. Nature 2016;538:523. https://doi.org/10.1038/NATURE19847.

[63] Kvon EZ, Waymack R, Gad M, Wunderlich Z. Enhancer redundancy in development and disease. Nature Reviews Genetics 2021 22:5 2021;22:324–36. https://doi.org/10.1038/s41576-020-00311-x.

[64] Osterwalder M, Barozzi I, Tissiéres V, Fukuda-Yuzawa Y, Mannion BJ, Afzal SY, et al. Enhancer redundancy provides phenotypic robustness in mammalian development. Nature 2018 554:7691 2018;554:239–43. https://doi.org/10.1038/nature25461.

[65] Waymack R, Fletcher A, Enciso G, Wunderlich Z. Shadow enhancers can suppress input transcription factor noise through distinct regulatory logic. Elife 2020;9:1–57. https://doi.org/10.7554/ELIFE.59351.

[66] Hou PS, Miyoshi G, Hanashima C. Sensory cortex wiring requires preselection of short- and long-range projection neurons through an Egr-Foxg1-COUP-TFI network. Nature Communications 2019 10:1 2019;10:1–18. https://doi.org/10.1038/s41467-019-11043-w.

[67] Firth H V., Richards SM, Bevan AP, Clayton S, Corpas M, Rajan D, et al. DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. The American Journal of Human Genetics 2009;84:524–33. https://doi.org/10.1016/J.AJHG.2009.03.010.

[68] Wang Y, Song F, Zhang B, Zhang L, Xu J, Kuang D, et al. The 3D Genome Browser: A web-based browser for visualizing 3D genome organization and long-range chromatin interactions. Genome Biol 2018;19:1–12. https://doi.org/10.1186/S13059-018-1519-9/FIGURES/5.

[69] Mohiyuddin M, Mu JC, Li J, Bani Asadi N, Gerstein MB, Abyzov A, et al. MetaSV: an accurate and integrative structural-variant caller for next generation sequencing. Bioinformatics 2015;31:2741–4. https://doi.org/10.1093/BIOINFORMATICS/BTV204.

[70] Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402–8. https://doi.org/10.1006/METH.2001.1262.

[71] Hamerlinck L, D'haene E, Loon N Van, Vaughan MB, Baca M del RP, Leimbacher S, et al. Non-coding structural variants identify a commonly affected regulatory region steering FOXG1 transcription in early neurodevelopment. MedRxiv 2025:2025.03.10.25323301. https://doi.org/10.1101/2025.03.10.25323301.

[72] Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ 2010. https://doi.org/https://qubeshub.org/resources/fastqc.

[73] Schwartzman O, Mukamel Z, Oded-Elkayam N, Olivares-Chauvet P, Lubling Y, Landan G, et al. UMI-4C for quantitative and targeted chromosomal contact profiling. Nat Methods 2016;13:685–91. https://doi.org/10.1038/NMETH.3922.

[74] Ramos-Rodríguez M, Subirana-Granés M, Pasquali L. UMI4Cats: an R package to analyze chromatin contact profiles obtained by UMI-4C. Bioinformatics 2021;37:4240–2. https://doi.org/10.1093/BIOINFORMATICS/BTAB392.

[75] Dobin A, Gingeras TR. Mapping RNA-seq Reads with STAR. Curr Protoc Bioinformatics 2015;51:11.14.1-11.14.19. https://doi.org/10.1002/0471250953.BI1114S51.

[76] Liao Y, Smyth GK, Shi W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. Nucleic Acids Res 2019;47:e47–e47. https://doi.org/10.1093/NAR/GKZ114.

[77] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550. https://doi.org/10.1186/s13059-014-0550-8.

[78] Yu G, Wang LG, Han Y, He QY. ClusterProfiler: An R package for comparing biological themes among gene clusters. OMICS 2012;16:284–7. https://doi.org/10.1089/OMI.2011.0118/ASSET/IMAGES/LARGE/FIGURE1.JPEG.

[79] Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. The Innovation 2021;2:100141. https://doi.org/10.1016/J.XINN.2021.100141.

[80] Yu G. Thirteen years of clusterProfiler. The Innovation 2024;5:100722. https://doi.org/10.1016/J.XINN.2024.100722.

[81] Xu S, Hu E, Cai Y, Xie Z, Luo X, Zhan L, et al. Using clusterProfiler to characterize multiomics data. Nature Protocols 2024 19:11 2024;19:3292–320. https://doi.org/10.1038/s41596-024-01020-z.

Table 1: Genetic and clinical details of 14q12 heterozygous deletion variants in individuals with epilepsy, ID and/or microcephaly considered in defining the minimum region of overlap (MRO)

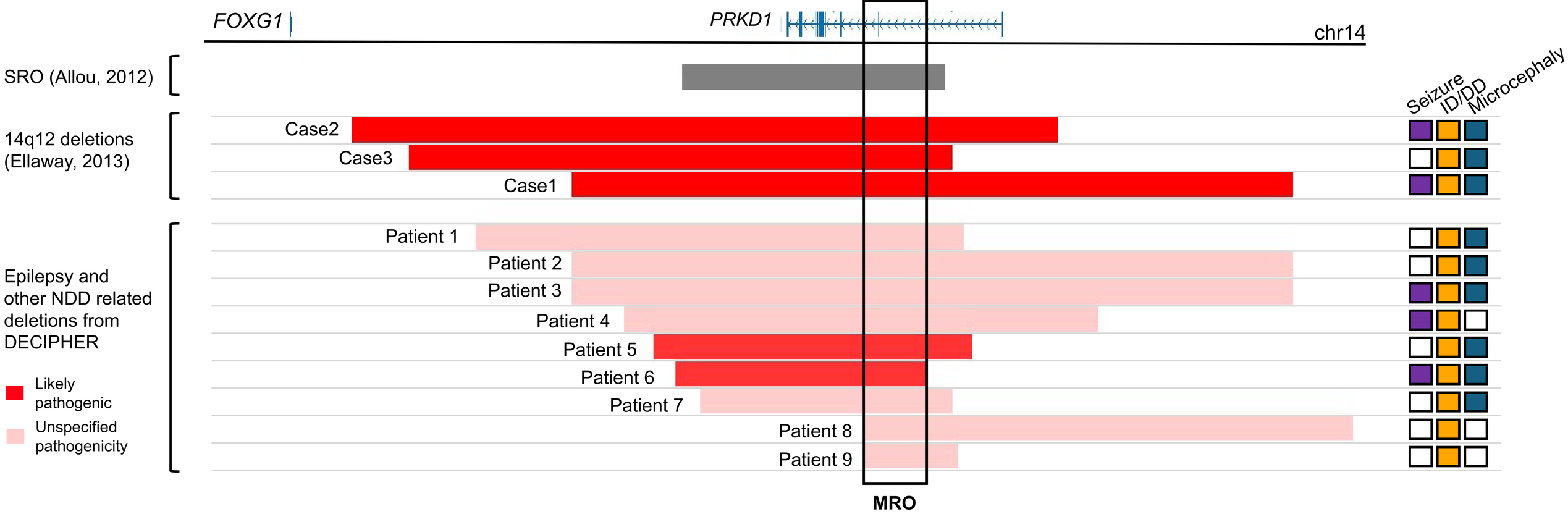
Patient	Decipher ID	Chromosomal/ Phenotypic sex	Deletion coordinates (hg38)	Inheritance	Pathogenicity assignment in DECIPHER	Seizures	Intellectual Disability (ID)	Microcephaly	Additional clinical features
DECIPHER									
1	260836	46XY	14:29068799-29865858	Unknown	Unspecified	-	Yes	Yes	
2	248405	46XY	14:29226052-30403168	*De novo*	Unspecified	-	Yes	Yes	
3	3810	46XX	14:29226052-30403168	*De novo*	Unspecified	Yes	Yes	Yes	Bruxism, hypoplasia of the corpus callosum, delayed speech and language development
4	258584	46XX	14:29312198-30083730	*De novo*	Unspecified	Yes	Yes	-	Atypical behavior, delayed speech and language development
5	358785	46XX	14:29360119-29879052	*De novo*	Likely Pathogenic	-	Yes	Yes (secondary)	Agenesis of corpus callosum
6	414554	Female	14:29395780-29803897	*De novo*	Likely Pathogenic	Yes	-	Yes	Delayed myelination, global developmental delay, simplified gyral pattern
7	252353	46XX	14:29435514-29847454	*De novo*	Unspecified	-	Yes	Yes	Hypoplasia of the corpus callosum
8	259935	46XX	14:29704706-30500956	Unknown	Unspecified	-	Yes	-	Atypical behavior
9 [37]	284715	46XY	14:29704736-29856618	Maternally inherited	Unspecified	-	Yes (mild)	-	Delayed speech and language development
Published									
Case 1 [36]	-	Female	14:29226052-30403168	*De novo*	-	Yes	-	Yes	Intellectual impairment, agenesis of corpus callosum, global developmental delay
Case 2 [36]	-	Female	14:28868223-30018792	*De novo*	-	Yes	-	Yes	Intellectual impairment, agenesis of corpus callosum, global developmental delay
Case 3 [36]	-	Male	14:28960335-29847513	*De novo*	-	-	-	Yes	Intellectual impairment, severe developmental delay

Figure 1: 14q12 microdeletions downstream of *FOXG1* in this study. Deletions present in individuals with epilepsy and other NDDs in DECIPHER are represented by horizontal bars and color-coded for pathogenicity. Additionally, three deletions from prior work that met our selection criteria are also shown (red horizontal bars) [36]. The overlap between the 14q12 deletions was used to define the minimum region of overlap (MRO), in this case by the deletion breakpoints in patients 6, 8 and 9 (black enclosure). The panel on the right indicates whether individuals with the respective 14q12 deletions presented with seizures, intellectual disability (ID) or developmental delays (DD) and microcephaly. Grey bar represents a single overlapping region (SRO) in the 14q12 region previously predicted to regulate *FOXG1* [35].

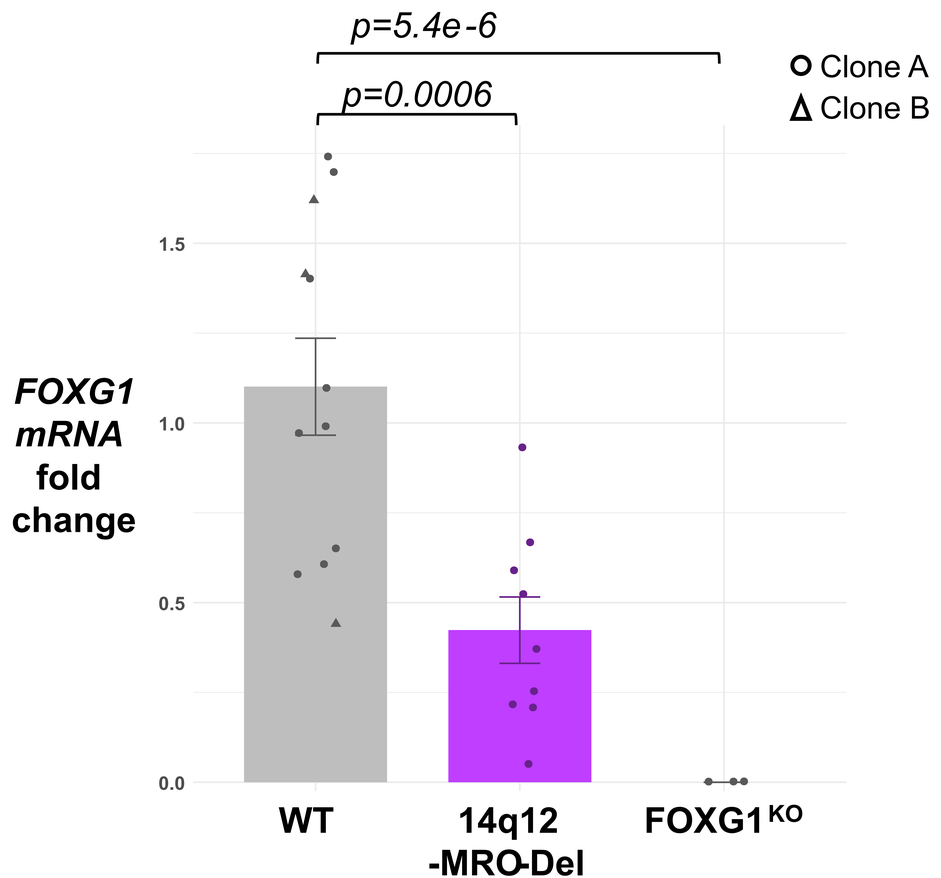
Figure 2: *FOXG1* expression is reduced with the loss of the MRO locus (A) *FOXG1* mRNA fold change relative to *EAR* assayed by qRT-PCR (n: WT Clone A = 9, WT Clone B = 3, 14q12-MRO-Del = 9, $FOXG1^{KO}$ = 3) (Two sample T-test, WT vs 14q12-MRO-Del, p= 0.0006; WT vs $FOXG1^{KO}$, p = 5.4e-6). (B) FOXG1 protein levels relative to TAF5 were assayed using Western Blotting (n: WT Clone A =15, WT Clone B = 3, 14q12-MRO-Del = 9, $FOXG1^{KO}$ = 6) (Two sample T-test, WT vs 14q12-MRO-Del, p= 0.0005; WT vs $FOXG1^{KO}$, p = 4.7e-10). Error bars represent one standard error.

Figure 3: The MRO locus harbors putative *FOXG1* CREs. (A) The MRO defined by the 14q12 patient deletions in this study is highlighted in red. *FOXG1* UMI-4C profiles represent genomic interactions of *FOXG1* in the 14q12 region (WT, grey; 14q12-MRO-Del; purple). ADHD- and schizophrenia-associated GWAS loci and CREs defined by the enhancer histone modification, H3K27Ac, as well as VISTA enhancers and ENCODE Candidate CREs (cCREs) in this region overlap peaks from the UMI-4C contact profiles. A previously described *FOXG1* regulatory region, SRO (smallest region of overlap, horizontal grey bar) in 14q12 encompasses VISTA enhancers hs1193, hs1201, hs1062, hs555 [35] and hs598 [60] (in bold). The schizophrenia-associated SNP rs1191551(pale yellow vertical highlight) [62] and the stem cell specific *FOXG1* TAD boundary, hESCb_2291, (pale grey vertical highlight) [37] also lie within the SRO. (B) A focus on the MRO region (red highlight) showcases native *FOXG1* genomic interactions in WT that are lost in 14q12-MRO-Del. Putative *FOXG1* CREs in the MRO are identified by the two H3K27Ac peaks conserved in HAP1s (ENCODE) and neural cell types (RoadMap Epigenomic Project).

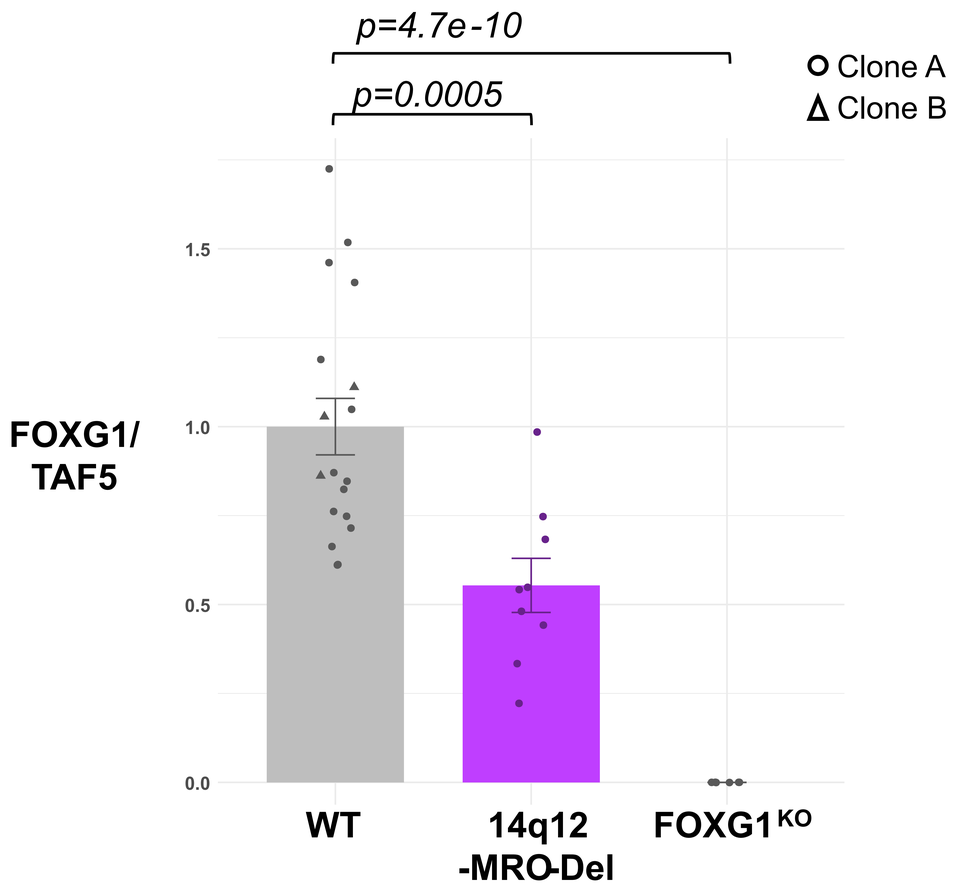
Figure 4: Differential gene expression analysis in MRO and *FOXG1* loss. Volcano plot of DEGs in (A) 14q12-MRO-Del vs WT and (B) $FOXG1^{KO}$ vs WT highlighting down regulated (blue: log2FC<1) and upregulated (orange:log2FC>1) DEGs (adjusted p-value<0.05). Genes that overlap both MRO and FOXG1 loss datasets are indicated by triangles. (C) Heatmap of gene expression changes for overlapping DEGs between 14q12-MRO-Del, and $FOXG1^{KO}$ HAP1s. (n=3 for WT Clone A, WT Clone B, 14q12-MRO-Del and $FOXG1^{KO}$). Genes implicated in *FOXG1* pathophysiology and clinical phenotypes are highlighted (panels A and B) and appear in bold in panel C.

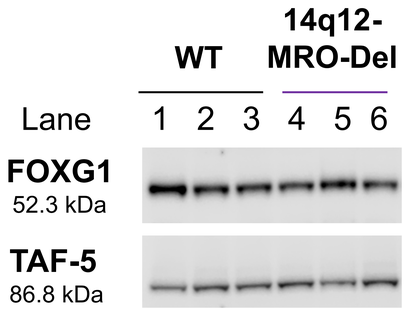

FOXG1
PRKD1
chr14
SRO (Allou, 2012)
14q12 deletions (Ellaway, 2013)
Case2
Case3
Case1
Epilepsy and other NDD related deletions from DECIPHER
Patient 1
Patient 2
Patient 3
Patient 4
Patient 5
Patient 6
Patient 7
Patient 8
Patient 9
Likely pathogenic
Unspecified pathogenicity
Seizure
ID/DD
Microcephaly
MRO

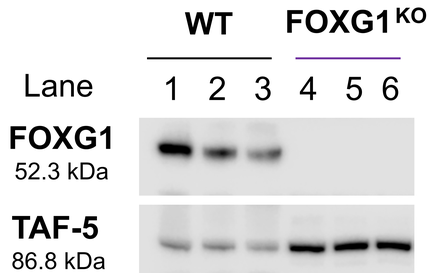
A. Relative fold change of *FOXG1* mRNA

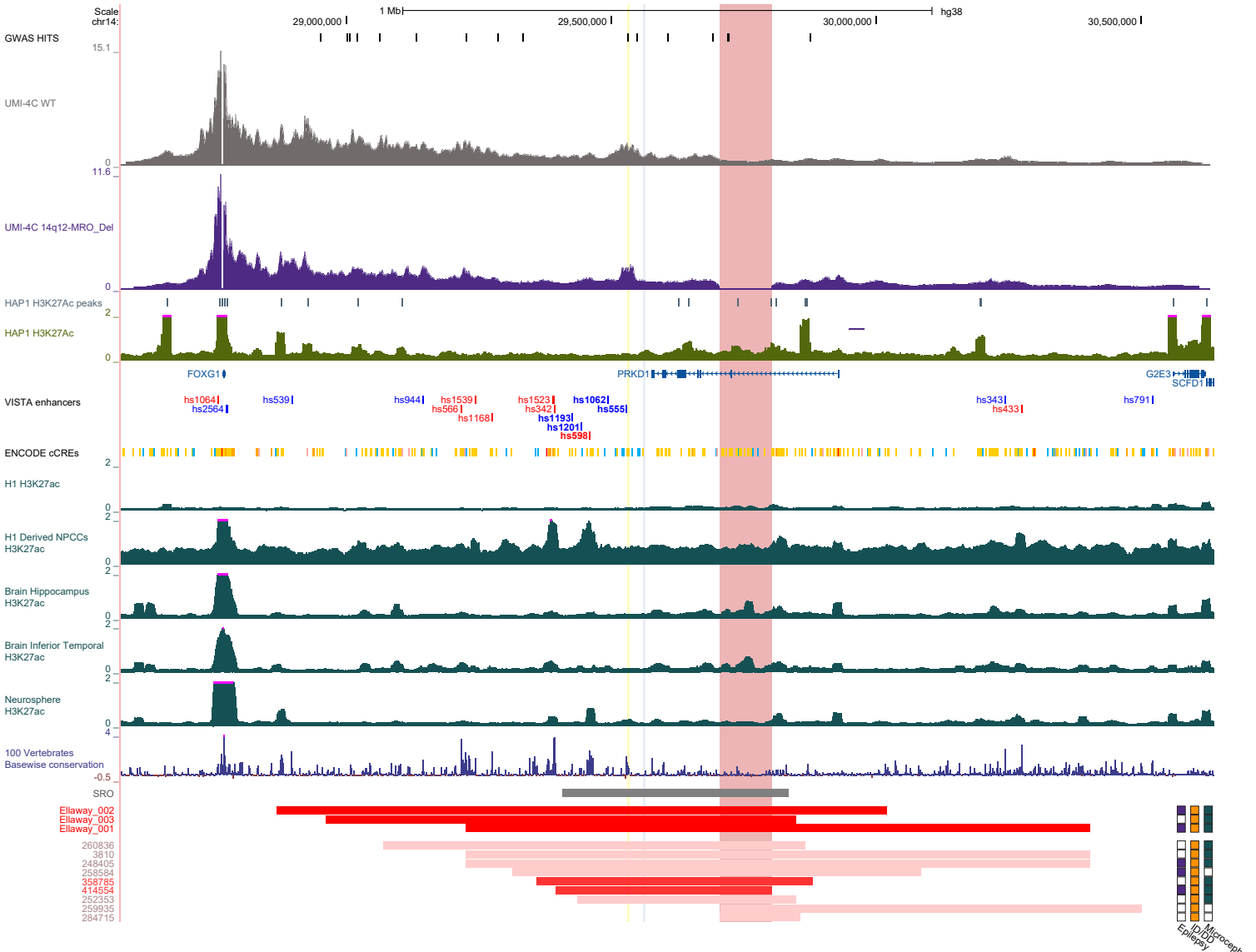


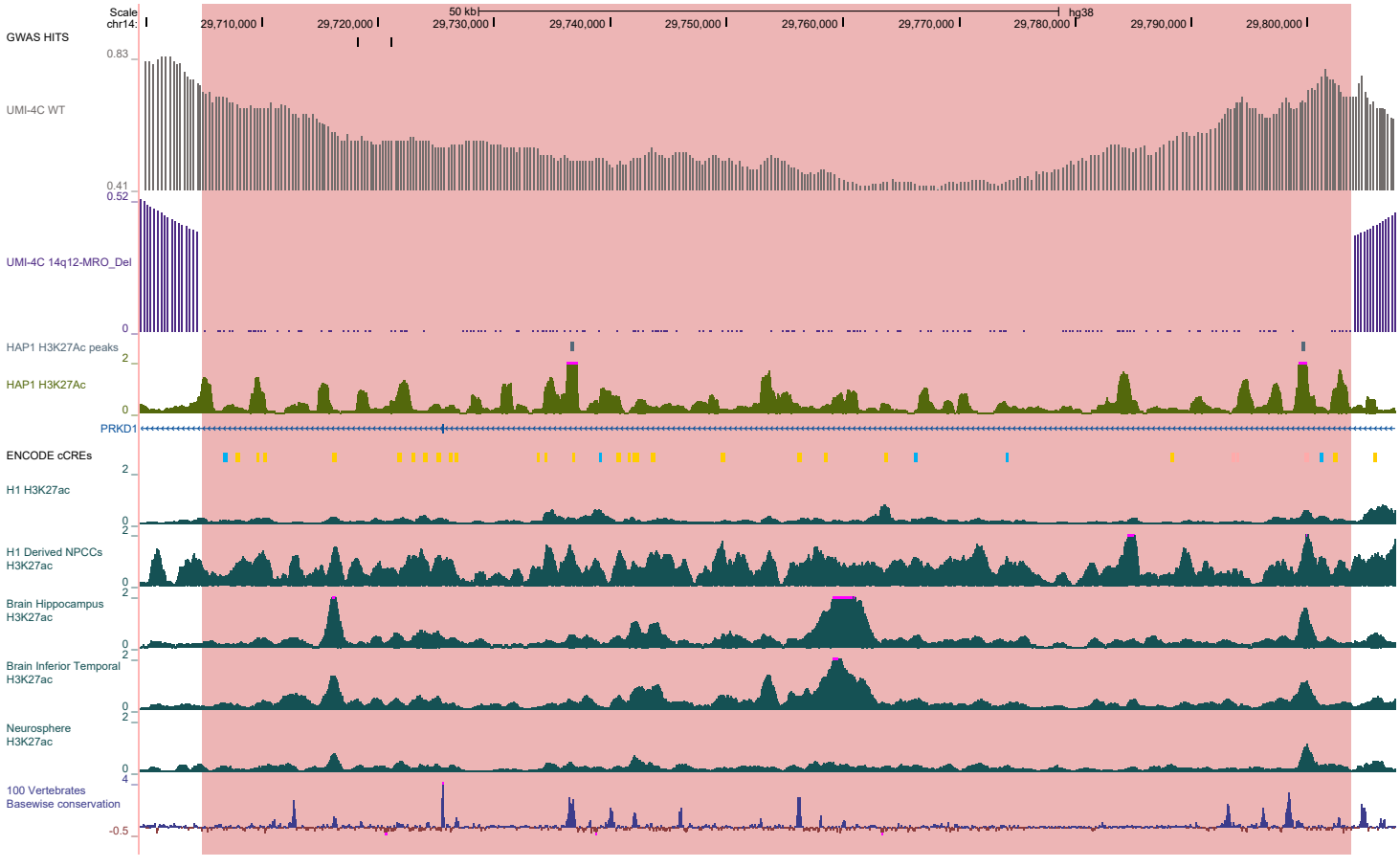
B. Relative fold change of FOXG1 protein

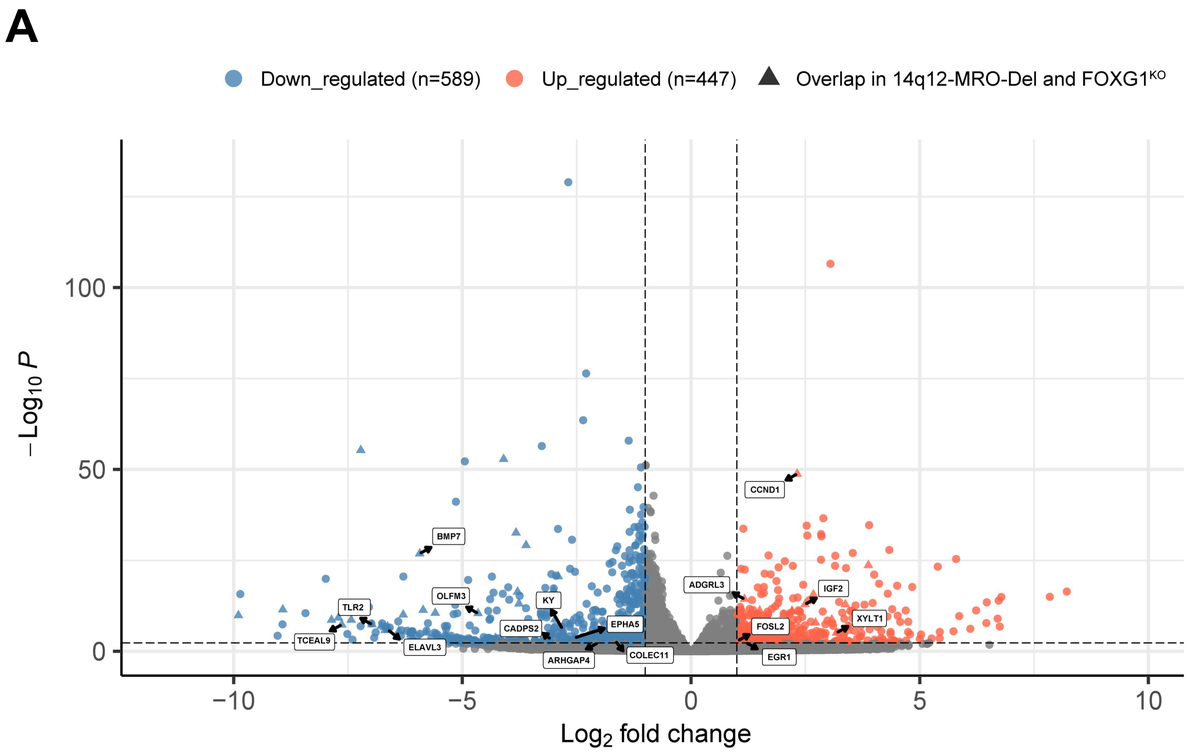


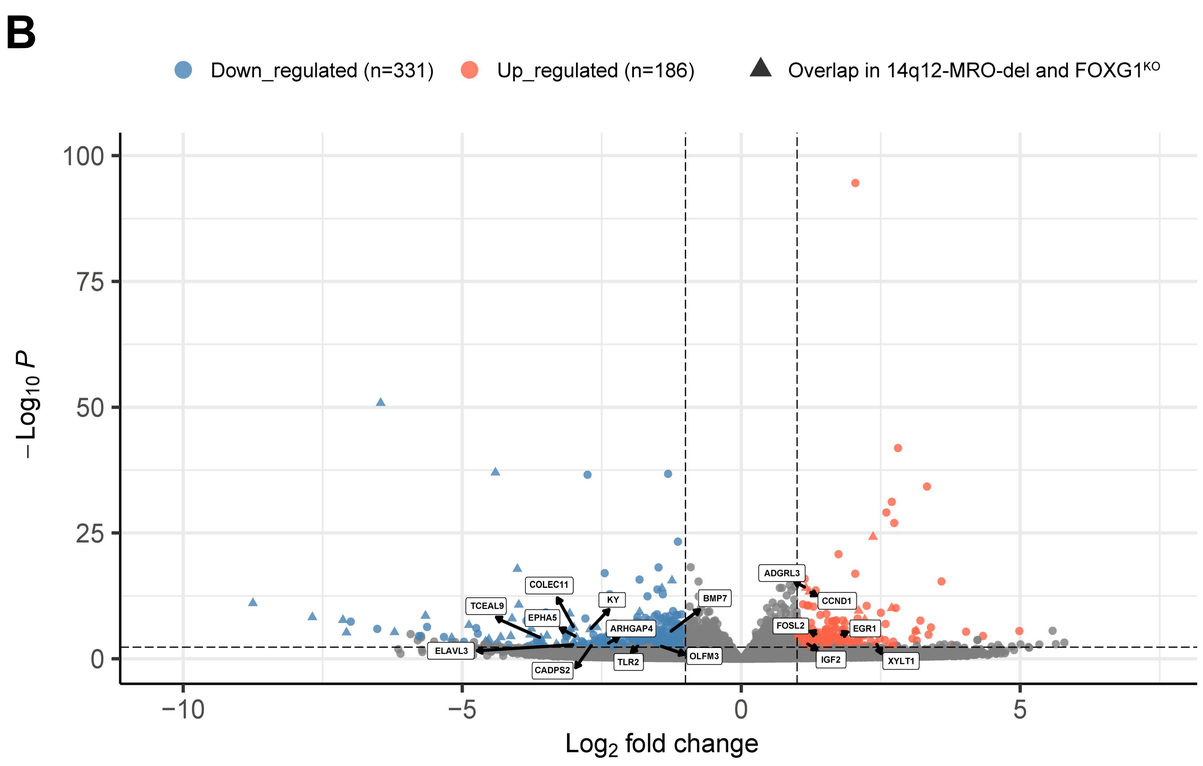


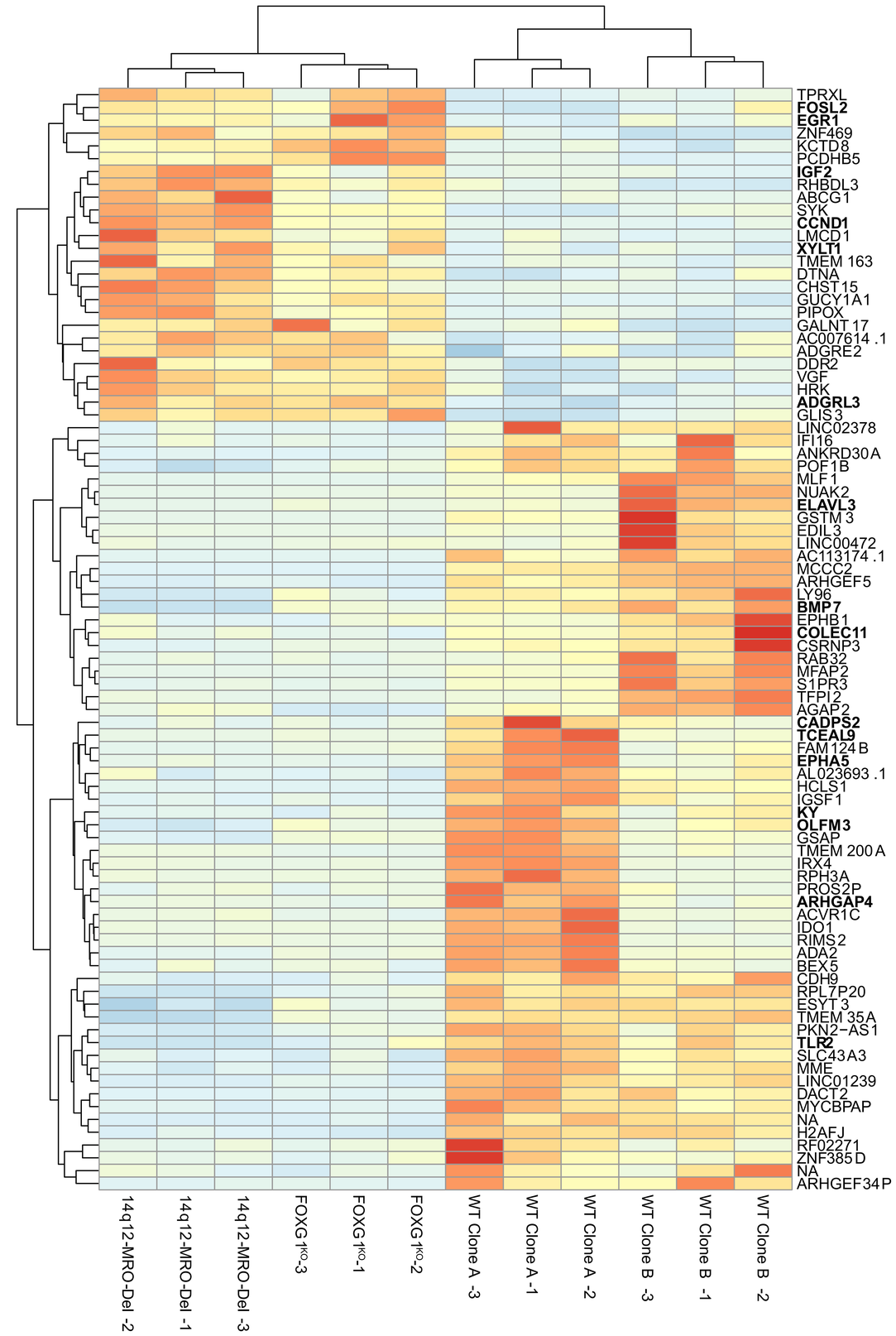



Scale
chr14:
1 Mb
hg38
29,000,000
29,500,000
30,000,000
30,500,000
GWAS HITS
UMI-4C WT
15.1
0
UMI-4C 14q12-MRO_Del
11.6
0
HAP1 H3K27Ac peaks
HAP1 H3K27Ac
2
0
FOXG1
PRKD1
G2E3
SCFD1
VISTA enhancers
hs1064
hs2564
hs539
hs944
hs1539
hs566
hs1168
hs1523
hs342
hs1062
hs555
hs1193
hs1201
hs598
hs343
hs433
hs791
ENCODE cCREs
H1 H3K27ac
H1 Derived NPCCs H3K27ac
Brain Hippocampus H3K27ac
Brain Inferior Temporal H3K27ac
Neurosphere H3K27ac
100 Vertebrates Basewise conservation
4
-0.5
SRO
Ellaway_002
Ellaway_003
Ellaway_001
260836
3810
248405
258584
358785
414554
252353
259935
284715
Epilepsy
ID/DD
Microcephaly


Scale
chr14:
50 kb
hg38
29,710,000
29,720,000
29,730,000
29,740,000
29,750,000
29,760,000
29,770,000
29,780,000
29,790,000
29,800,000
GWAS HITS
UMI-4C WT
0.83
0.41
UMI-4C 14q12-MRO_Del
0.52
0
HAP1 H3K27Ac peaks
HAP1 H3K27Ac
2
0
PRKD1
ENCODE cCREs
H1 H3K27ac
H1 Derived NPCCs
H3K27ac
Brain Hippocampus
H3K27ac
Brain Inferior Temporal
H3K27ac
Neurosphere
H3K27ac
100 Vertebrates
Basewise conservation
4
-0.5


A
Down_regulated (n=589)
Up_regulated (n=447)
Overlap in 14q12-MRO-Del and FOXG1KO
−Log10 P
Log2 fold change
CCND1
BMP7
ADGRL3
IGF2
TLR2
OLFM3
KY
CADPS2
EPHA5
FOSL2
XYLT1
TCEAL9
ELAVL3
ARHGAP4
COLEC11
EGR1


B
Down_regulated (n=331)
Up_regulated (n=186)
Overlap in 14q12-MRO-del and FOXG1KO
−Log10 P
Log2 fold change
COLEC11
TCEAL9
EPHA5
KY
BMP7
ARHGAP4
ADGRL3
CCND1
FOSL2
EGR1
ELAVL3
CADPS2
TLR2
OLFM3
IGF2
XYLT1


TPRXL
FOSL2
EGR1
ZNF469
KCTD8
PCDHB5
IGF2
RHBDL3
ABCG1
SYK
CCND1
LMCD1
XYLT1
TMEM 163
DTNA
CHST15
GUCY1A1
PIPOX
GALNT17
AC007614.1
ADGRE2
DDR2
VGF
HRK
ADGRL3
GLIS3
LINC02378
IFI16
ANKRD30A
POF1B
MLF1
NUAK2
ELAVL3
GSTM3
EDIL3
LINC00472
AC113174.1
MCCC2
ARHGEF5
LY96
BMP7
EPHB1
COLEC11
CSRNP3
RAB32
MFAP2
S1PR3
TFPI2
AGAP2
CADPS2
TCEAL9
FAM124B
EPHA5
AL023693.1
HCLS1
IGSF1
KY
OLFM3
GSAP
TMEM 200A
IRX4
RPH3A
PROS2P
ARHGAP4
ACVR1C
IDO1
RIMS2
ADA2
BEX5
CDH9
RPL7P20
ESYT3
TMEM 35A
PKN2-AS1
TLR2
SLC43A3
MME
LINC01239
DACT2
MYCBPAP
NA
H2AFJ
RF02271
ZNF385D
NA
ARHGEF34P
2
1
0
−1
−2
14q12-MRO-Del -2
14q12-MRO-Del -1
14q12-MRO-Del -3
FOXG1KO-3
FOXG1KO-1
FOXG1KO-2
WT Clone A -3
WT Clone A -1
WT Clone A -2
WT Clone B -3
WT Clone B -1
WT Clone B -2