nature portfolio

Peer Review File

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

Corresponding Author: Dr Gemma Carvill

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Ramamurthy et al. present a detailed analysis of a novel regulatory locus downstream of FOXG1 disrupting gene regulation of NDD-associated FOXG1. They identified a minimal region of overlap (MRO) in patients carrying deletions (n=9 from DECIPHER) refining a previously published smallest region of overlap (SRO; n=3 patients) identified from a published study from Ellaway et al (2013)) downstream of FOXG1 that also interrupts PRKD1. Premising that PRKD1 LoF does not lead to detrimental phenotypes given gnomad data (pLI 0) and no known NDD association, they hypothesize that the deletions impact expression of FOXG1, in which haploinsufficiency is known to cause NDD. Authors generated isogenic Hap1 lines carrying the minimal deletion and a FOXG1 deletion. This resulted in ~50% reduction of FOXG1 expression and altered chromatin looping using a 4C assay with putative CREs from published epigenomic data. Finally, authors compared transcriptomic profiles of the KO and FOXG1-del lines and observe significant overlap in differentially expressed genes, supporting shared mechanisms.

The study is well written and easy to follow, but would benefit from more discussion on study limitations. Given the missing genetic risk in NDD, identification of functional assessment of noncoding regulatory mutations are difficult but an important next step. While the authors have convincingly shown the focal PRKD1 deletion results in altered FOXG1 expression in Hap1 cells, connecting it to NDD phenotypes is still tenuous. Further, the proposed mechanism of altered cis regulation of a putative enhancer has not been fully verified. Some comments to help address these points.

Major comments:
1. Ideally FOXG1 expression from a subset of patient samples would be tested to verify the reduction in expression in the presence of downstream deletions. Perhaps this was performed previously by Allou et al (2012) when defining the SRO.

2. It is unclear if the heterozygous deletion in a diploid individual would lead to sufficient FOXG1 downregulation producing ID phenotypes observed in patients 8 & 9. 50% Hap1 downregulation would result in ~25% downregulation in a diploid state (in the absence of any compensation of the non-deleted allele). Is there evidence in other models that this level of FOXG1 knockdown would lead to cellular or developmental phenotypes?

4. While it is true that PRKD1 LoF mutations segregate in humans, gnomad shows that no homozygous variants have been reported to date. It is possible that complete LoF of PRKD1 likely produced in the 14q13-MRO-Del Hap1 line might lead to cellular phenotypes resulting in secondary FOXG1 downregulation. This might be controlled for/assayed with efficient knockdown via siRNA against PRKD1 and testing FOXG1 expression.

5. A scatterplot comparing log2 fold change of the FOXG1 KO vs the 14q13-MRO-Del line would better delineate similarities and differences btwn the two edited cell lines. This would enable an assessment of correlations as well as differences between the two that might be driven by the complete loss of FOXG1vs PRKD1. Was FOXG1 flagged as a DEG in both lines? PRKD1 in the deletion line?

6. While the 4C experiment is nicely performed in Hap1 cells, such looping is often tissue/cell-type specific, making it uncertain if this pattern would be conserved in neural cell types. Ideally the experiment would be performed in a neural cell

type to verify similar looping. At a minimum, including the general TAD structure using published brain/neural Hi-C data of the locus would enable the reader to understand the broader chromatin architecture of the locus and if looping occurs within a TAD (more likely) or crosses TAD boundaries (less likely). This would also help place the Figure 3 hESC-specific TAD boundary into context.

7. Functional validation of putative CREs identified in the deletion locus, perhaps using a reporter assay in a relevant cell line, would help support the claim that the deletion impacts gene regulation in brain/neural cells via disruption of enhancers.

Minor comments:

1. In the Discussion: Authors mention how the two patients defining the proximal breakpoint of the MRO (DECIPHER patients 8 and 9) present with less severe phenotypes. It is also important to note that the pathogenicity of the MRO-defining deletions is more uncertain (unknown parental status and inherited). In the case of the inherited variant, what is the mother's ID diagnosis?

2. Presentation of the baseline expression of FOXG1 in Hap1 cells relative to a biologically relevant cell/tissue type would better provide context of the relative expression change of the deletion allele.

3. An important limitation is that experiments were not performed in cell types (e.g., neural) that would be relevant for NDD phenotypes in patients. This should be noted in the Discussion.

4. The designation of MRO (this study) and SRO (previous Allou 2012) is confusing in its current presentation, with SRO mentioned first in the Discussion yet included in Figures 1 and 3. Consider presenting the SRO in the beginning of the Results, as it's also shown at the top of Figure 1, and then describe how the addition of nine new samples helps to further delineates the locus to the MRO.

5. In Figure 2A: why is the mean FOXG1 transcript abundance relative to WT fold change of the wild-type cell lines not 1.0?

Reviewer #2

(Remarks to the Author)
Dr Carvill and her lab assembled data on a cohort of 12 patients with microdeletions in the 14q12 cytoband downstream of FOXG1. They showed that the deletion of the minimal region of overlap (MRO) of these microdeletions in HAP1 cells resulted in decreased expression of both FOXG1 RNA and protein. Transcriptome profiling comparisons showed that similar pathways are perturbed upon deletion of FOX1G and the above-described MRO, further suggesting the presence of FOX1G regulatory elements in the latter interval.

1. While stating that "None of these deletions encompassed any other gene except PRKD1. PRKD1 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]." is formally correct, it does not stand scrutiny and the use of a better suited continuous measure of constraint. Indeed, gnomAD indicates that PRKD1 is under some constraint with observed/expected = 0.6 and pLOEUF = 0.76. The comparison of the GTEX expression patterns of PRKD1 and FOXG1, for example, appear like a more compelling argument to focus on the latter gene. Please rephrase your argumentation.
2. The comparison of the transcriptome profiles of HAP, MRO and KO cells can be more compelling. It would be interesting to stratify differentially expressed genes (DEG) in direct targets of FOXG1 and downstream effects. Are the overlapping DEG more likely to be direct targets? For FOXG1 & partners and their targets see for example biorxiv Jeon et al., "The patient-specific mouse model with Foxg1 frameshift mutation provides insights into the pathophysiology of FOXG1 syndrome". It is also somewhat worrisome that only 17% of DEG overlap (8% depending on how you look at it).
3. Conservation of the regulatory elements (e.g. Vista enhancers) should be discussed
4. HAP1 is a fantastic model but it comes with its caveat. It is neither a diploid cell nor a full organism. To follow the state of the art the authors should assess their regulatory element in the mouse.

Reviewer #3

(Remarks to the Author)
Major Comment:

The manuscript reported by Ramamurthy et al describes the identification of a minimal critical region within the PRKD1 gene resulting in a loss of FOXG1 expression in genetically modified HAP1s cells. While the manuscript is very nicely written demonstrating that 14q12 deletions are associated to FOXG1 haploinsufficiency causing FOXG1 syndrome, it still lacks the full information regarding whether these regulatory elements are indeed enhancers and where are they located within the MRO region. H3K27ac ChIP-seq data indicate the presence of two putative enhancer elements within this genomic region. However, these have not been tested using a LacZ reporter assay or a Luciferase reporter assay to demonstrate that they can drive the expression of a gene. Indeed, these elements could also be tethering elements facilitating the communication between FOXG1 regulatory enhancers and the gene itself.

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
Authors have addressed well all of my comments and concerns, with additional experiments and a more in depth Discussion. Congrats to authors on an exciting study

Reviewer #2

(Remarks to the Author)
The authors have satisfactorily addressed my queries, for example by including a functional validation of MRO-embedded enhancers and via a more systematic evaluation of FOXG1 targets

Reviewer #3

(Remarks to the Author)
The authors have addressed my comments.

RESPONSE TO REVIEWERS

The reviewer comments appear in black and our responses in blue.

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

we also provide an updated data availability statement and provide thefollowing link for the reviewers.
https://dataview.ncbi.nlm.nih.gov/object/PRJNA1348521?reviewer=a1hqe9lc5rlum2oegr8jjrtrj1

Most importantly, we added functional testing of the candidate *cis* regulatory elements (cCREs) in the MRO and strengthen links to relevance of this MRO to the neurological condition observed in patients. This change addressed one of the unifying themes amongst reviewers, while two others (cell line choice, and RNA-seq analysis) were also evident. We thus provide our response up front for these three themes and have hyperlinked to this section in the specific response to reviewers as well. Moreover, on revisiting the manuscript, there was some data that was in the discussion that was likely better presented in results, this has now been completely restructured to better support our conclusions. Figures 3 and 4 are also revised/new.

1. **Functional validation of cCREs in the MRO in HAP1 and neuronal cell lines**

In our initial submission, we demonstrated that deletion of the MRO led to decreased *FOXG1* mRNA expression and protein abundance in a haploid HAP1 cell model. However, we did not provide extensive evidence for the regulatory potential of this region. To address this, we tested the regulatory activity of two cCREs, designated by HAP1 H3K27Ac peaks, within the MRO using a luciferase-based reporter assay in both HAP1s, and a mouse hippocampal cell line (HT-22s). In this system we show that at least one of the cCREs, highly conserved across species, leads to increased expression of the reporter gene, in both HAPs as well as a neuronal context (HT-22s). In addition, we also added data from a massively parallel reporter assay (MPRA) from primary cortical cells derived from human fetal brain and induced pluripotent stem cell (iPSC)-derived cerebral organoids (Figure 3/4) (PMID: 38781390). This data showed multiple active enhancer elements across the region, including two active CREs in the MRO. These CREs were different to the CREs we functionally validated in HAPS and HT22s, suggesting that the MRO has multiple active regulatory elements, that when deleted can perturb *FOXG1* expression and are disease relevant, particularly in the context of individuals with neurodevelopmental disorders harboring 14q12 deletions.

We have also provided better context for the entire 14q12 FOXG1-regulatory region, adding data from primary and neuronal H3K27Ac enrichment, UMI-4C/Hi-C, MPRAs and vista enhancers from multiple sources. In addition, we added the complementary and supportive data described by Dr. Vergult's manuscript, on MedRxiv (2025.03.10.25323301v1), also under consideration at Nature Communications. Collectively, we are now confident that we have provided a better picture of the

complex regulatory nature of the 14q12 locus which provides improved context for how deletions in this region cause neurodevelopmental disorders, as well as other neurological conditions.

All of these experiments and changes have been made in the manuscript **results** as below. In addition, we did a major restructure of the entire **discussion** and include a new **Figure 3 and 4** and added some details to **Table 1**.

"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) (Fig. 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. PRKD1 is tolerant to loss of function in the general population with a pLI score of 0 and LOEUF of 0.76 (gnomAD, [28,29]), has very low expression in the brain compared to other tissues (GTEx), and the mouse knockout has no neurological phenotype (MGI:99879), collectively suggesting PRKD1 loss is not associated with NDDs. 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 (Supplementary Fig. 2).

We selected the HAP1 cell line due to its haploid nature, which facilitates genome editing, as well as the endogenous expression of FOXG1 (Supplementary Fig. 1 and Fig. 2). We successfully deleted the MRO using dual flanking guides in HAP1s (14q12-MRO-Del); the entire 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: chr14:28,767,269-28,768,758) using Sanger sequencing, and in the 14q12-MRO-Del confirmed the absence of major off-target SVs using genome sequencing (Supplementary Fig. 3). We confirmed the loss of PRKD1 exon 2 and flanking intronic sequence at both the genomic DNA and mRNA level in 14q12-MRO-Del (Supplementary Fig. 3-4).

Using qRT-PCR, we demonstrated complete loss of FOXG1 mRNA expression in $FOXG1^{KO}$ and a 61% reduction in the 14q12-MRO-Del line (p-value 0.0006, Welch's T test) (Fig. 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) (Fig. 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."

Discussion in manuscript, did not copy here.

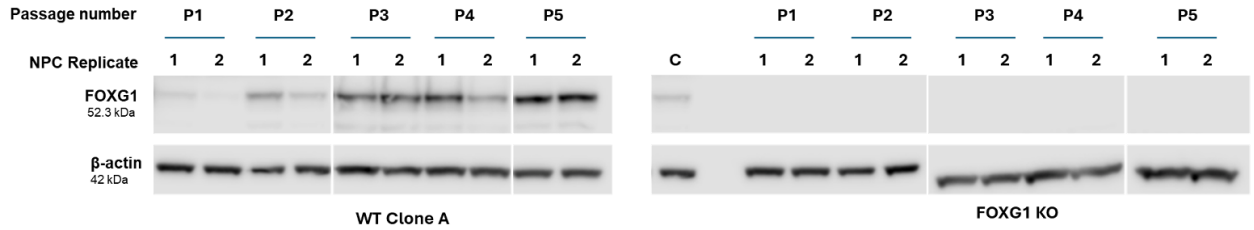
2. Modeling the MRO deletion in stem cell models.

We also show here data from our attempts to model the effects of MRO deletion in iPSC-derived neural lines. To surmise up-front, we (like others) could not reliably and reproducibly detect *FOXG1* in iPSC-derived NPCs as expression of this transcription factor is fleeting and neuronal differentiation protocols are still variable. We present the data here for the reviewers, but do not include in the manuscript given the challenges described below. Though, we go back and forth on its inclusion in the manuscript; on the one hand, it is beneficial to point out these challenges and argue for more rigor in the iPSC field, on the other, this does not particularly impact the conclusions of the paper, and is challenging to add, without disrupting the 'flow'. In the end we added a few sentences in the discussion, but we welcome input from reviewers in this regard.

We did not have access to cell lines or samples from affected individuals with 14q12 deletions encompassing the MRO, thus we utilized a healthy control iPSC line to model the MRO deletion. Using the same dual-guide sgRNA/Cas9 approach we describe in HAPs, we deleted the MRO (14q12-MRO$^{Del/WT}$) as well as FOXG1 (FOXG1^{KO}) in iPSCs. We differentiated these iPSCs to neural precursor cells (NPCs) using the well-established dual SMAD inhibition neural induction approach and evaluated FOXG1 expression.

FOXG1 promotes the self-renewal and expansion of neural progenitors in the forebrain (PMIDs 7605629 and 12151532). Thus, we anticipated transient expression in NPCs and performed a time course experiment to capture peak FOXG1 abundance. FOXG1 expression stabilized ~3 weeks (passage 3 - P3) of differentiation and was not detected in our FOXG1^{KO} lines (Figure R1A). However, expression was variable even between differentiation replicates of the same clone (1 and 2 – Figure R1A). We selected passage 5 (P5) as a time point with the most robust FOXG1 abundance and differentiated two wildtype isogenic lines and two 14q12-MRO$^{Del/WT}$ clones. However, we now detected no FOXG1 at either the protein or mRNA level even in the isogenic wildtype replicates (Figure R1B). We attempted multiple iterations of this experiment but were unable to reliably detect FOXG1 in NPCs, invalidating this approach as a reliable method of measuring FOXG1 abundance. Furthermore, we differentiated WT iPSCs to generate induced neuronal (iN) cells using NGN2 overexpression (PMID 23764284), but did not detect FOXG1, likely because the direct NGN2 differentiation process omits precursor stages of neural differentiation. These results emphasize that FOXG1 expression is transient during NPC differentiation, yet this temporal variability is not consistent across repeated differentiations (even side-by-side in the same clone) and is too variable to reliably measure FOXG1 abundance, precluding MRO deletion evaluation. In the future, a reporter-based approach and live-cell imaging will likely need to be used to robustly detect changes in FOXG1 abundance in differentiating cells. Heterogenous expression of FOXG1 across differentiation replicates in wildtype NPCs has been previously reported (PMIDs 35148845 and 37216650).

A. FOXG1 abundance abundance across NPC differentiation



B. FOXG1 abundance at P5 in WT and 14q12 MRO deletion lines

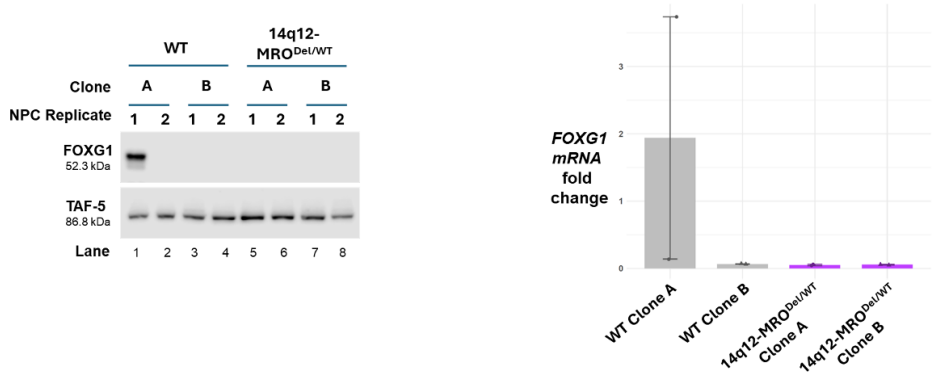


Figure R1. Time course evaluation of *FOXG1* expression during NPC differentiation. iPSCs were differentiated in duplicate (1 and 2). (A) *FOXG1* is detected as early as P1 and an increase of expression with each passage is observed in wildtype NPCs. As expected, *FOXG1* is undetected at all time points in FOXG1^{KO} NPCs. (B) iPSCs were differentiated in duplicate and NPCs at P5 were harvested for expression studies. FOXG1 protein levels relative to TAF5 were assayed using Western Blotting in NPCs (left). *FOXG1* mRNA fold change relative to *EAR* assayed by qRT-PCR in NPCs (right). NPC differentiation replicates = 2 for WT Clone A, WT Clone B, 14q12-MRO$^{Del/WT}$ Clone A and 14q12-MRO$^{Del/WT}$ Clone B.

We added the following text to the discussion:

"…While these results are encouraging, they do highlight a limitation of our study, in that we do not demonstrate the impact of MRO loss in a neuronal cell line. We selected HAP1s because of the ease of editing and robust FOXG1 abundance, and are confident that at least some of the regulatory landscape is overlapping, but we acknowledge that this is not a neuronal line. We did also generate iPSC lines heterozygous for MRO deletion and differentiated them to neural cell types. However, we found inconsistent FOXG1 expression in iPSC-derived NPCs and iPSC-derived neurons that was not reproducible, even in wildtype control lines (data not shown). High variability in the endogenous expression of this gene is also seen in previous iPSC studies [20,40]. Therefore, collectively we could not use these neuronal models as an effective experimental system to study the effects of MRO- and FOXG1 loss. In the future, reporter-based systems, and live-imaging will be necessary to accurately quantify the temporal expression of FOXG1 in iPSC-derived models, to counteract the variability inherent in these systems."

3. Additional RNA-seq analysis

To more systematically evaluate the FOXG1 target genes that are also differentially expressed genes (DEGs) in the FOXG1^{KO} and 14q12-MRO-Del lines, we retrieved ChIP-seq peaks from embryonic day 15.5 mouse cortices (PMID: 30392794) and CUT&Tag peaks from embryonic day 16.5 mouse ganglionic eminence cells (PMID: 40447563). We merged these two lists retrieving a comprehensive (i.e. present in either dataset) set of FOXG1 peaks which were annotated to the nearest gene (n=8007) and show that there are an equal proportion of direct FOXG1 targets (Supplementary Data 1). We have rewritten the **results** section, also including abundance of *FOXG1* in the text as follows:

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

Given FOXG1's role 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. FOXG1 was a DEG in both lines with a larger effect in FOXG1^{KO} (log2FC -2.753916 and adjusted p-value 8.484237e-34), as compared to the 14q12-MRO-Del line (log2FC -0.7397485 and adjusted p-value 0.002290243). Furthermore, we identified 86 DEGs that overlapped between the two datasets. The shared DEGs included genes involved in cell cycle regulation (CCND1, BMP7) [38–40] as well as those implicated in epilepsy (FOSL2, ELAVL3, EPHA5, OLFM3, ARHGAP4) [41–45], ID/microcephaly (XYLT1, COLEC11, CADPS2) [46–48], and other NDDs (KY, ADGRL3, TCEAL9) [49–51] (Fig. 5 and Supplementary Data 1). Functional enrichment analysis of DEGs revealed a few shared molecular functions and biological pathways between 14q12-MRO-Del and FOXG1^{KO} including terms relevant in a neuronal context: "regulation of nervous system development", "axon/neuron projection guidance" and "regulation of neurotransmitter levels" (Supplementary Fig. 9). We also used two existing chromatin immunoprecipitation sequencing experiments from mouse embryonic brain (E15.5 cortex [52] and E16.5 ganglionic eminence [53]) to identify 8007 genes bound by Foxg1. Of these direct Foxg1 targets, a comparable fraction were also DEGs in the FOXG1^{KO} and 14q12-MRO-Del lines (17% (86/517 DEGs) in FOXG1^{KO} and 15% (154/1036 DEGs) in 14q12-MRO-Del) (Supplementary Data 1). These results are comparable to the 33% of DEGs that are direct targets of Foxg1 in the mouse ganglionic eminence [53]. Thus collectively, even though we modeled MRO loss in a non-neuronal HAP1 cell line, we observe overlapping DEGs, including direct Foxg1 targets, in the MRO deletion and FOXG1 coding loss, suggesting shared downstream effects of reduced expression of this transcription factor."

Reviewer #1 (Remarks to the Author):

Ramamurthy et al. present a detailed analysis of a novel regulatory locus downstream of *FOXG1* disrupting gene regulation of NDD-associated *FOXG1*. They identified a minimal region of overlap (MRO) in patients carrying deletions (n=9 from DECIPHER) refining a previously published smallest region of overlap (SRO; n=3 patients) identified from a published study from Ellaway et al (2013)) downstream of *FOXG1* that also interrupts *PRKD1*. Premising that *PRKD1* LoF does not lead to detrimental phenotypes given gnomad data (pLI 0) and no known NDD association, they hypothesize that the deletions impact expression of *FOXG1*, in which haploinsufficiency is known to cause NDD. Authors generated isogenic Hap1 lines carrying the minimal deletion and a *FOXG1* deletion. This resulted in ~50% reduction of *FOXG1* expression and altered chromatin looping using a 4C assay with putative CREs from published epigenomic data. Finally, authors compared transcriptomic profiles of the KO and FOXG1-del lines and observe significant overlap in differentially expressed genes, supporting shared mechanisms.

The study is well written and easy to follow, but would benefit from more discussion on study limitations. Given the missing genetic risk in NDD, identification of functional assessment of noncoding regulatory mutations are difficult but an important next step. While the authors have convincingly shown the focal *PRKD1* deletion results in altered *FOXG1* expression in Hap1 cells, connecting it to NDD phenotypes is still tenuous. Further, the proposed mechanism of altered cis regulation of a putative enhancer has not been fully verified.

We thank the reviewers for their enthusiasm and constructive feedback. The detailed comments have helped improve our manuscript considerably, we show this in a point-by-point fashion below. In addition, to address the proposed mechanism by which the MRO may alter FOXG1 expression and links to the neurological phenotypes, including our attempts to model the MRO deletion in iPSC-derived NPCs, we describe our changes in the preamble (hyperlinks should link to top of this document)

Finally, we added the following sections to the **discussion** highlighting the limitations of our study:

"Here we demonstrate that there are multiple CREs within the MRO that may regulate FOXG1 expression, but that there are likely additional redundant enhancers across the 14q12 locus region. Testing this hypothesis, however, is challenging, and a limitation of the dual sgRNA guide genome-editing strategy used in this study to introduce the ~100 kb MRO deletion. While newer methods such as dCas9-controlled CRISPR/Cas3 can generate deletions on a megabase scale [61], our approach has too low an efficiency to introduce the >400 kb SRO. Future studies, using Cas3-like deletion approaches in cellular and animal models are key for examining the redundancy of CREs in the 14q12 region, and the absolute number of CREs that need to be deleted to abrogate FOXG1 expression. Furthermore, massively parallel reporter assays (MPRA) in human neurodevelopmental stem cell models and human tissues, such as those described [33], are likely to be essential for assessing the activity of individual CREs."

And

"While these results are encouraging, they do highlight a limitation of our study, in that we do not demonstrate the impact of MRO loss in a neuronal cell line. We selected HAP1s because of the ease of editing and robust FOXG1 abundance, and are confident that at least some of the regulatory

landscape is overlapping, but we acknowledge that this is not a neuronal line. We did also generate iPSC lines heterozygous for MRO deletion and differentiated them to neural cell types. However, we found inconsistent FOXG1 expression in iPSC-derived NPCs and iPSC-derived neurons that was not reproducible, even in wildtype control lines (data not shown). High variability in the endogenous expression of this gene is also seen in previous iPSC studies [20,40]. Therefore, collectively we could not use these neuronal models as an effective experimental system to study the effects of MRO- and FOXG1 loss. In the future, reporter-based systems, and live-imaging will be necessary to accurately quantify the temporal expression of FOXG1 in iPSC-derived models, to counteract the variability inherent in these systems."

Some comments to help address these points.

Major comments:

1. Ideally *FOXG1* expression from a subset of patient samples would be tested to verify the reduction in expression in the presence of downstream deletions. Perhaps this was performed previously by Allou et al (2012) when defining the SRO.

We agree with the reviewer that evaluating *FOXG1* expression from patient samples would be ideal to determine expression levels; however, samples from patients in this study were not available. Moreover, *FOXG1* expression is minimally detected in the most commonly accessible patient-derived biospecimens, such as blood (TPM = 0) and fibroblasts (TPM =0.14). These levels of expression preclude accurate quantification of *FOXG1* levels. To highlight this challenge, two existing studies have attempted to capture the impact of putative regulatory 14q12 deletions on *FOXG1* expression in patient-specific biospecimens. The first showed elevated *FOXG1* mRNA and protein levels in patient-specific fibroblasts and saliva (PMID: 22739344). Conversely, *FOXG1* mRNA expression was downregulated in blood samples a different group of individuals (PMID: 22968132). These contradicting outcomes, in patients with similar downstream 14q12 deletions, are likely the result of virtually absent *FOXG1* mRNA levels in the biospecimens tested and the lack of rigorous controls and replicates, which collectively render inaccurate assessment of FOXG1 abundance.

We now show the cell/tissue specific expression data in **Supplementary Figure 1** and have added to the **introduction** to highlight previous studies that evaluated *FOXG1* expression from individuals harboring 14q12 deletions downstream of *FOXG1*.

"SVs overlapping putative CREs have also been described in individuals with overlapping clinical features of FOXG1 haploinsufficiency, but without coding variants [23,25–27]. While the impact of these deletions on FOXG1 expression was examined, results have been contradictory, likely owing to the virtually absent FOXG1 expression in accessible peripheral patient biospecimens (Supplementary Fig. 1). Specifically, one group showed increased FOXG1 in blood and saliva from individuals with 14q12 deletions and predicted a single region of overlap (SRO) to be regulatory [25]. Conversely, another showed reduced FOXG1 expression in blood, despite shared regions of overlap across patient-specific deletions (Fig. 1) [26]. To address these inconsistencies, here we refined the SRO to a new minimum region of overlap (MRO) in individuals with FOXG1 haploinsufficient clinical features and downstream deletions, and modeled the MRO in the HAP1 cell line which robustly expresses FOXG1."

We also provide data for our attempts to model the MRO deletion in iPSC-derived NPCs in the introduction.

2. It is unclear if the heterozygous deletion in a diploid individual would lead to sufficient *FOXG1* downregulation producing ID phenotypes observed in patients 8 & 9. 50% Hap1 downregulation would result in ~25% downregulation in a diploid state (in the absence of any compensation of the non-deleted allele). Is there evidence in other models that this level of *FOXG1* knockdown would lead to cellular or developmental phenotypes?

This is a very fair and interesting point, we too were initially intrigued by the 50% reduction, vs a complete loss of expression, which one might perhaps expect to equate to a 50% reduction in a diploid cell. While there are multiple reasons for this observation, we propose the most likely explanation is that we selected a minimal region of overlap (MRO) to model in HAP1s which was smaller than both deletions in patients 8 and 9, and indeed all deletions in this region. Thus, while we adopted a classic genetics approach, i.e. looking for the critical region of overlap amongst individuals, this may actually be the 'wrong' approach in a cis regulatory element (CRE) study, as deletion of multiple CREs is likely necessary to abrogate expression completely. Moreover, genetic studies in other neurological disorders might offer some insights into reduced, but not haploinsufficient, levels of FOXG1 and clinical outcomes. For instance, a SNP associated with schizophrenia reduced FOXG1 levels 20%.We have refined the **discussion** section to more clearly reflect this:

*"Deletion of the MRO reduced FOXG1 expression but did not eliminate it, indicating that multiple CREs across the region need to be deleted to abrogate FOXG1 expression. Indeed, the MRO lies within a larger regulatory region of FOXG1 designated as the SRO (smallest region of overlap); 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 . The SRO is a complex region with multiple validated, highly conserved active enhancers in mouse, zebrafish and human fetal brain tissue and organoid models [30–33]. We refined the SRO to the MRO based on the distal deletions in patient 8 and 9. Of note, these two individuals present with ID only. In contrast, individuals with deletion of both the SRO and MRO (patient 1-7) generally present with more severe clinical features of FOXG1 haploinsufficiency: ID, seizures and microcephaly (**Table 1**). These varied clinical presentations, taken together with the reduction, but not complete loss of FOXG1 by MRO deletion, suggests that disruption of multiple CREs is required to completely ablate FOXG1 expression. Moreover, the number of disrupted CREs may underpin genotype-phenotype relationships; where individuals with deletion of more CREs have lower FOXG1 expression and clinical features more reminiscent of individuals with FOXG1 coding variants, while individuals with fewer CREs deleted (e.g. patients 8 and 9) have less severe clinical presentation correlated with higher FOXG1 abundance. Indeed, this modulated view of FOXG1 expression finds support in other neurological conditions. For instance, there are multiple GWAS loci in the 14q12 regulatory region that have been associated with increased risk for ADHD and schizophrenia [56]. Deletion of a ~500 bp region flanking a schizophrenia-associated SNP rs1191551 reduced FOXG1 expression by ~20% in HEK293T cells [57]. Moreover, a luciferase assay demonstrated reduced activity of the risk rs1191551 allele, suggesting that even this single nucleotide change can modulate FOXG1 expression. Collectively, these studies highlight the need to systematically assess both disease-associated SVs and single nucleotide changes across this regulatory region in neurological disease broadly, and that variable gene expression across*

neurodevelopment may have varied phenotypic consequences. As genome sequencing facilitates more widespread and smaller SV and SNV detection, these genotype-phenotype relationships may become more apparent."

4. While it is true that *PRKD1* LoF mutations segregate in humans, gnomad shows that no homozygous variants have been reported to date. It is possible that complete LoF of *PRKD1* likely produced in the 14q13-MRO-Del Hap1 line might lead to cellular phenotypes resulting in secondary *FOXG1* downregulation. This might be controlled for/assayed with efficient knockdown via siRNA against *PRKD1* and testing *FOXG1* expression.

This is indeed theoretically possible. Deletion of the MRO leads to the complete deletion of exon 2 and partial intron 1 and 2 of PRKD1 (0.99Mb region). We confirmed at the mRNA level in our RNA-seq data that this exon is skipped, generating a transcript that is out of frame and likely triggers nonsense mediated decay (NMD) (**Supplementary Figure 4**). That said, *PRKD1* had only a small change in expression in the 14q12-MRO-Del line (log2FC -1.030001) that was not significant (adjusted p-value 0.1613391), possibly due to the incomplete and dynamic nature of NMD. We did not observe alternative splicing that might rescue this out of frame transcript. However, there is an alternate start codon downstream which may be used and would lead to a smaller protein, spanning 747 of the 912 amino acids and domains as shown below (Figure 2R).

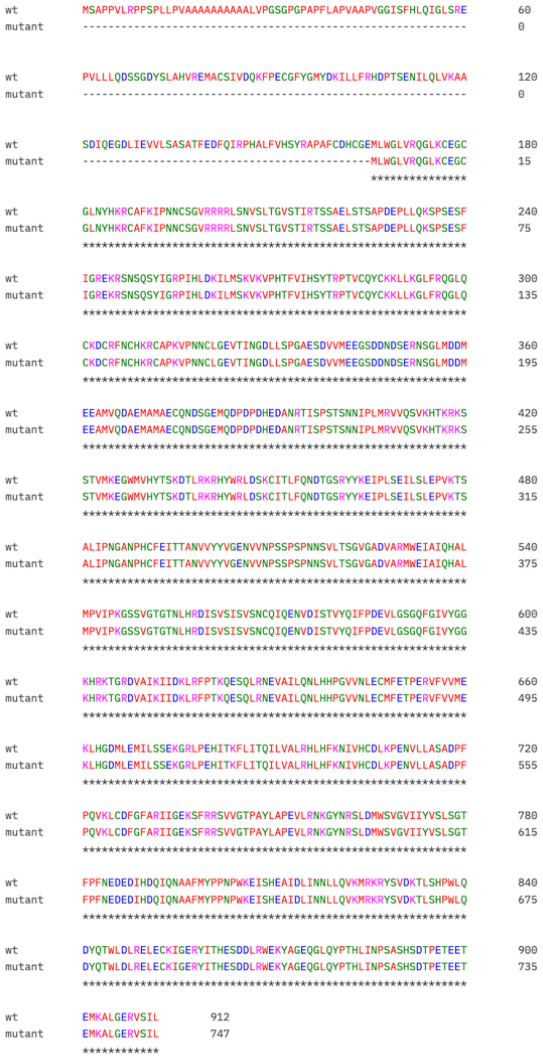


Figure R2: Protein alignment for PRKD1 for the full length and putative shorter isoform via an alternative methionine 172 amino acids downstream.

Regardless we do not know the cellular consequences of this putative complete loss of *PRKD1*. From gnomAD we know heterozygous loss is well tolerated, we would anticipate that if reduced *PRKD1* led to reduced *FOXG1*, then even in a haploinsufficient state we might expect to see neurological phenotypes, as *FOXG1* was reduced 20% in a cellular system in the context of a risk allele for schizophrenia. Thus overall, we are confident that PRKD1 reduction does not impact FOXG1 expression, though it cannot be ruled out. We updated this rationale in the **results** section as follows:

"None of these deletions encompassed any other gene except PRKD1. PRKD1 is tolerant to loss of function in the general population with a pLI score of 0 and LOEUF of 0.76 (gnomAD, [28,29]), has very low expression in the brain compared to other tissues (GTEx), and the mouse knockdown has no neurological phenotype (MGI:99879), collectively suggesting PRKD1 loss is not associated with NDDs."

5. A scatterplot comparing log2 fold change of the FOXG1 KO vs the 14q13-MRO-Del line would better delineate similarities and differences btwn the two edited cell lines. This would enable an assessment of correlations as well as differences between the two that might be driven by the complete loss of *FOXG1* vs *PRKD1*. Was *FOXG1* flagged as a DEG in both lines? *PRKD1* in the deletion line?

Expression of *PRKD1* is addressed in the previous comment. *FOXG1* was a DEG in both lines with a larger effect in the *FOXG1* knockout line, we added these details to the **results**:

"FOXG1 was a DEG in both lines with a larger effect in FOXG1^{KO} (log2FC -2.753916 and adjusted p-value 8.484237e-34) as well as the 14q12-MRO-Del line (log2FC -0.7397485 and adjusted p-value 0.002290243)."

We also show here for the reviewer the scatterplot for the FOXG1 KO vs the 14q12-MRO-Del, which detected 2115 DEGs (**Figure R3**). However, caution here is warranted as our comparisons between the two WT clones (the parent and the unedited line that went through dual sgRNAs and Cas9 transfection) and each of the edited lines (FOXG1/14q12-MRO-Del) have greater power to detect DEGs.

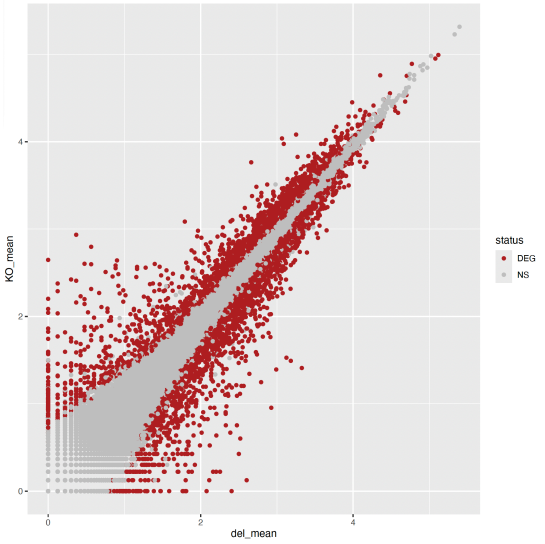


Figure R3: Scatterplot of gene expression for FOXG1 KO HAP1 line (KO_mean) against the 14q12-MRO-Del HAP1 line (del_mean). Differentially expressed gene n=2115 (DEG) (log2FC > abs 1, adjusted p-value <0.05), NS = not significant (n= 33667)

Finally, in response to another reviewer, we also provide additional RNA-seq analysis in the introduction.

6. While the 4C experiment is nicely performed in Hap1 cells, such looping is often tissue/cell-type specific, making it uncertain if this pattern would be conserved in neural cell types. Ideally the experiment would be performed in a neural cell type to verify similar looping. At a minimum, including the general TAD structure using published brain/neural Hi-C data of the locus would enable the reader to understand the broader chromatin architecture of the locus and if looping occurs within a TAD (more likely) or crosses TAD boundaries (less likely). This would also help place the Figure 3 hESC-specific TAD boundary into context.

As suggested, we have added the general TAD structure at 14q12 from HAP1s, embryonic stem cells (ESCs), iPSC-derived NPCs, and NPC-derived neurons (Supplementary Fig. 7-8). The distal TAD boundary containing *FOXG1* resides within the SRO (Figure 3), and is conserved in HAP1s, ESCs, NPCs and neurons (Supplementary Fig. 8). In individuals with *FOXG1* syndrome carrying deletions spanning the SRO, the pathogenic mechanism was proposed to be the disruption of this TAD boundary resulting in the adoption of enhancers from the neighboring TAD (PMID 25315429). However, a later study (PMID 29289958) demonstrated that this distal TAD boundary was not critical for gene regulation, and that the regulatory landscape of *FOXG1* extended into the neighboring TAD, further evidenced by cross boundary looping interactions of *FOXG1* in various cell types from other studies (Supplementary Fig. 8). Neither the distal TAD boundary, nor the TAD encompassing *FOXG1* are disrupted to enable gene expression in NPCs and neurons where *FOXG1* expression occurs, suggesting *FOXG1* regulation is independent of higher order TAD structures. We expanded our discussion of the 14q12 *FOXG1* regulatory region in the **results:**

"Deleting the MRO not only led to loss of UMI-4C contacts in this region as expected, but also altered native FOXG1 genomic interactions. Specifically, significant differences in UMI contacts between WT and 14q12-MRO-Del were observed at six 14q12 loci, including four regions with increased contacts and two that were decreased (Fig. 3, Supplementary Fig. 6). Of note, four of these enhanced contacts (3 increased, one decreased) were located 1.5 Mb downstream of FOXG1, extending beyond a TAD boundary present in HAP1s, embryonic stem cells (ESCs), neural progenitor cells (NPCs) and neurons (Fig. 3). These contacts suggest that FOXG1 regulation extends beyond this TAD boundary, substantiated by the presence of inter boundary looping interactions as well (Supplementary Fig. 7-8)."

7. Functional validation of putative CREs identified in the deletion locus, perhaps using a reporter assay in a relevant cell line, would help support the claim that the deletion impacts gene regulation in brain/neural cells via disruption of enhancers.

We agree with the reviewer and have addressed this comment in the preamble response to all reviewers: cCRE validation

Minor comments:

1. In the Discussion: Authors mention how the two patients defining the proximal breakpoint of the MRO (DECIPHER patients 8 and 9) present with less severe phenotypes. It is also important to note that the pathogenicity of the MRO-defining deletions is more uncertain (unknown parental status and inherited). In the case of the inherited variant, what is the mother's ID diagnosis?

This information is unavailable.

2. Presentation of the baseline expression of *FOXG1* in Hap1 cells relative to a biologically relevant cell/tissue type would better provide context of the relative expression change of the deletion allele.

This has been addressed in comment 1 (**Supplementary Figure 1**).

3. An important limitation is that experiments were not performed in cell types (e.g., neural) that would be relevant for NDD phenotypes in patients. This should be noted in the Discussion.

We address the comment in the preamble section and added to the limitations in the **discussion**.

"While these results are encouraging, they do highlight a limitation of our study, in that we do not demonstrate the impact of MRO loss in a neuronal cell line. We selected HAP1s because of the ease of editing and robust FOXG1 abundance, and are confident that at least some of the regulatory landscape is overlapping, but we acknowledge that this is not a neuronal line. We did also generate iPSC lines heterozygous for MRO deletion and differentiated them to neural cell types. However, we found inconsistent FOXG1 expression in iPSC-derived NPCs and iPSC-derived neurons that was not reproducible, even in wildtype control lines (data not shown). High variability in the endogenous expression of this gene is also seen in previous iPSC studies [20,40]. Therefore, collectively we could not use these neuronal models as an effective experimental system to study the effects of MRO- and FOXG1 loss. In the future, reporter-based systems, and live-imaging will be necessary to accurately quantify the temporal expression of FOXG1 in iPSC-derived models, to counteract the variability inherent in these systems."

4. The designation of MRO (this study) and SRO (previous Allou 2012) is confusing in its current presentation, with SRO mentioned first in the Discussion yet included in Figures 1 and 3. Consider presenting the SRO in the beginning of the Results, as it's also shown at the top of Figure 1, and then describe how the addition of nine new samples helps to further delineates the locus to the MRO.

We appreciate this recommendation and have incorporated changes accordingly, to introduce the SRO with a reference to Figure 1 in the introduction, prior to introducing the MRO.

5. In Figure 2A: why is the mean *FOXG1* transcript abundance relative to WT fold change of the wild-type cell lines not 1.0?

We thank the reviewer for highlighting this observation. We have edited Figure 2A (and Supplementary Figure 3C) which now reflect a mean *FOXG1* (or *PRKD1* respectively) transcript abundance relative to WT fold change of 1 in the wild-type cell lines.

Reviewer #2 (Remarks to the Author):

Dr Carvill and her lab assembled data on a cohort of 12 patients with microdeletions in the 14q12 cytoband downstream of *FOXG1*. They showed that the deletion of the minimal region of overlap (MRO) of these microdeletions in HAP1 cells resulted in decreased expression of both *FOXG1* RNA and protein. Transcriptome profiling comparisons showed that similar pathways are perturbed upon deletion of FOX1G and the above-described MRO, further suggesting the presence of *FOX1G* regulatory elements in the latter interval.

We thank the reviewer for their thorough evaluation of our manuscript and providing us with helpful feedback.

1. While stating that "None of these deletions encompassed any other gene except *PRKD1*. *PRKD1* 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]." is formally correct, it does not stand scrutiny and the use of a better suited continuous measure of constraint. Indeed, gnomAD indicates that *PRKD1* is under some constraint with observed/expected = 0.6 and pLOEUF = 0.76. The comparison of the GTEX expression patterns of *PRKD1* and *FOXG1*, for example, appear like a more compelling argument to focus on the latter gene. Please rephrase your argumentation.

We agree on all points and have modified the **results** section as follows:

"*"None of these deletions encompassed any other gene except PRKD1. PRKD1 is tolerant to loss of function in the general population with a pLI score of 0 and LOEUF of 0.76 (gnomAD, [28,29]), has very low expression in the brain compared to other tissues (GTEx), and the mouse knockdown has no neurological phenotype (MGI:99879), collectively suggesting PRKD1 loss is not associated with NDDs."*

2. The comparison of the transcriptome profiles of HAP, MRO and KO cells can be more compelling. It would be interesting to stratify differentially expressed genes (DEG) in direct targets of *FOXG1* and downstream effects. Are the overlapping DEG more likely to be direct targets? For *FOXG1* & partners and their targets see for example biorxiv Jeon et al., "The patient-specific mouse model with Foxg1 frameshift mutation provides insights into the pathophysiology of *FOXG1* syndrome". It is also somewhat worrisome that only 17% of DEG overlap (8% depending on how you look at it).

We thank the reviewer for this helpful suggestion and have added additional RNA-seq analysis preamble response to all reviewers: RNA-seq analysis

Of note the BioRxiv manuscript referred to by the reviewer, used ChIP-seq data from PMID: 30392794, this data was used in the data described above.

3. Conservation of the regulatory elements (e.g. Vista enhancers) should be discussed

We agree, we added details on conservation of regulatory elements in the **results** and conservation is also highlighted in **Figure 3 and 4**:

"The CRE landscape across this region is complex with many known and predicted regulatory regions and most have high levels of conservation (Fig. 3). This includes nine VISTA enhancers active in

neurodevelopment, of which six show forebrain-specific activity (Hs1064, Hs566, Hs1523, Hs342, Hs433, Hs344) (Supplementary Fig. 5)"

"We selected these two H3K27Ac peaks (cCRE1 and cCRE2) in the MRO to test for enhancer activity in the HAP1 cell line and the HT-22 cell line derived from mouse hippocampus. cCRE1, which has high levels of conservation, showed a statistically significant 1.6-fold and 3.4-fold increase in expression in HAP1s and HT-22s, respectively".

Also in the **discussion** section:
"Deletion of the MRO led to reduced FOXG1 expression by roughly half in the HAP1 cell line, possibly due to the deletion of cCRE1 which activated expression of a reporter gene in HAP1s. This cCRE is highly conserved across evolution and correlates with a peak of H3K27Ac in HAP1s and a distal enhancer signature in ENCODE, collectively suggesting it functions as an enhancer in HAP1s. However, the second H3K27Ac peak, cCRE2 which is poorly conserved, reduced expression of the reporter gene in HAP1s and shows correlation with H3K4me3 and DNAse hypersensitivity from ENCODE CREs usually demarcating promoter regions."

"The SRO is a complex region with multiple validated, highly conserved active enhancers in mouse, zebrafish and human fetal brain tissue and organoid models [30–33].:

4. HAP1 is a fantastic model but it comes with its caveat. It is neither a diploid cell nor a full organism. To follow the state of the art the authors should assess their regulatory element in the mouse.

We partially address this comment in the iPSC neuronal model in the preamble. While we value the role of modeling regulatory regions in mice, this is beyond scope of the present study and our lab's expertise. We also added this paragraph to the discussion:

"While these results are encouraging, they do highlight a limitation of our study, in that we do not demonstrate the impact of MRO loss in a neuronal cell line. We selected HAP1s because of the ease of editing and robust FOXG1 abundance, and are confident that at least some of the regulatory landscape is overlapping, but we acknowledge that this is not a neuronal line. We did also generate iPSC lines heterozygous for MRO deletion and differentiated them to neural cell types. However, we found inconsistent FOXG1 expression in iPSC-derived NPCs and iPSC-derived neurons that was not reproducible, even in wildtype control lines (data not shown). High variability in the endogenous expression of this gene is also seen in previous iPSC studies [20,40]. Therefore, collectively we could not use these neuronal models as an effective experimental system to study the effects of MRO- and FOXG1 loss. In the future, reporter-based systems, and live-imaging will be necessary to accurately quantify the temporal expression of FOXG1 in iPSC-derived models, to counteract the variability inherent in these systems."

Reviewer #3 (Remarks to the Author):

Major Comment:

The manuscript reported by Ramamurthy et al describes the identification of a minimal critical region within the *PRKD1* gene resulting in a loss of *FOXG1* expression in genetically modified HAP1s cells. While the manuscript is very nicely written demonstrating that 14q12 deletions are associated to *FOXG1* haploinsufficiency causing *FOXG1* syndrome, it still lacks the full information regarding whether these regulatory elements are indeed enhancers and where are they located within the MRO region. H3K27ac ChIP-seq data indicate the presence of two putative enhancer elements within this genomic region. However, these have not been tested using a LacZ reporter assay or a Luciferase reporter assay to demonstrate that they can drive the expression of a gene. Indeed, these elements could also be tethering elements facilitating the communication between *FOXG1* regulatory enhancers and the gene itself.

We thank the reviewer for their helpful feedback and suggestions. We agree with the reviewer that validation of the putative enhancers within the MRO deletion is necessary and accordingly performed the experiments described in the introduction (should hyperlink to top).