Investigating SOX10 Dosage Sensitivity in the developing Enteric Nervous System.

Background: The Enteric Nervous System (ENS) is composed of millions of neurons and glia that regulate key processes involved in gut homeostasis. Development of the ENS requires Enteric Neural Crest Cells (ENCCs) - progenitor cells of the ENS - to coordinate distinct cellular functions that include migration and commitment to ENS specific cellular lineages (*1*). The transcription factor (TF) SOX10 is a master regulator of ENCC gene expression and is implicated in human disorders of ENS formation(*2*). Changes in SOX10 levels coincides with distinct cell states during ENS development, and in a clinical setting, SOX10 haploinsufficiency results in Hirschsprung disease - a condition characterized by aganglionosis of the gut(*3*). Despite the cellular and clinical consequences of perturbations in SOX10 levels, systematic investigation mapping the effects of SOX10 dosage on human ENCC function is lacking. We aim to address this deficit by understanding how quantitative changes in SOX10 levels alter critical gene regulatory programs necessary for ENCC migration, proliferation, and ENS lineage commitment.

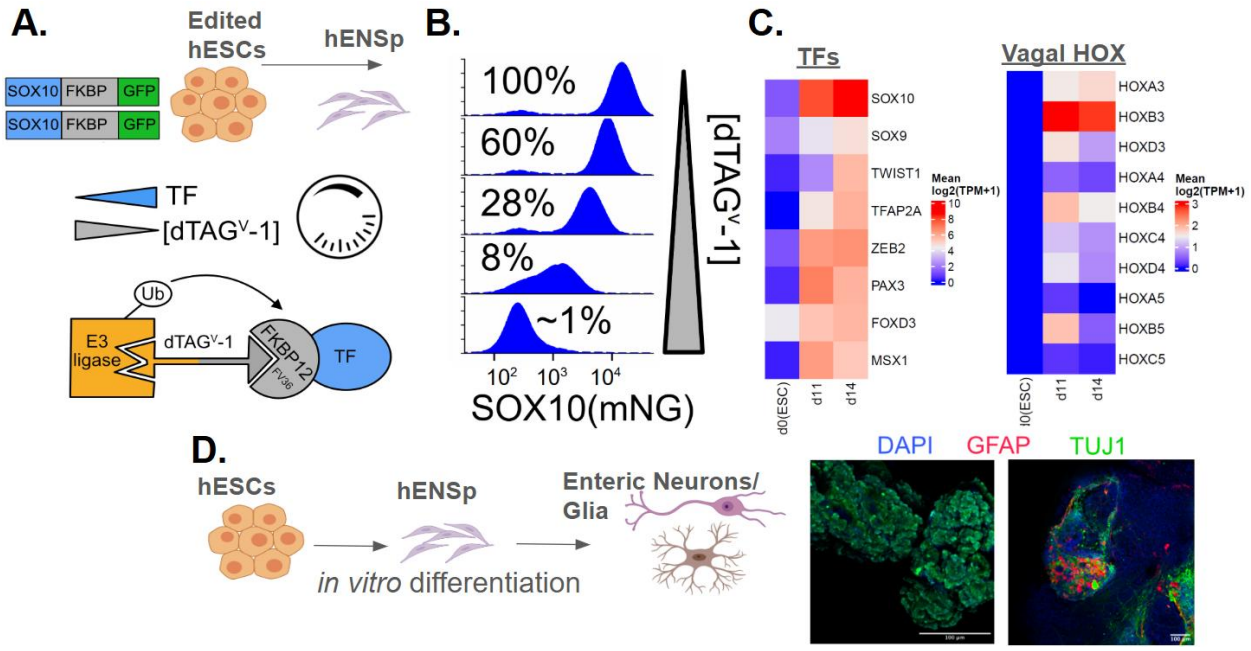


Figure 1: Chemical Genetics Approach to interrogate SOX10 dosage effects in a model of human ENS development. A,B. Endogenous tagging of the SOX10 locus in frame with the FKBP12^{F36V} degron allows for tunable decrease in SOX10 levels with dTAG^{V}-1 treatment. **C.** Robust protocol to derive human enteric nervous system progenitors (hENSp) from hESCs - demonstrated by expression of Vagal HOX genes and neural crest Transcription Factors. **D.** Prolonged culture of hENSp generates markers of enteric neurons (TUJ1) and glia (GFAP). Confocal images represent cross-section of 3-dimensional ganglioid tissue after culturing hENSp neurospheres for an extended period between 30 to 60 days in cultures.

Experimental Approach: To exercise control over endogenous SOX10 levels in an ENS developmental model, we used Crispr-Cas9 gene editing to engineer the native SOX10 locus to express a C-terminal fusion of the FKBP12^{F36V} degron (*4*) along with an mNEON green reporter. Differentiating bi-allelic SOX10 edited human embryonic stem cell (hESC) clones into human enteric nervous system progenitor cells (hENSps) demonstrates robust expression of SOX10 as measured through mNEON green expression (Figures 1A,B). Application of dTAGv1 at 5 concentrations (0nM, 0.05nM, 0.5M, 1.6nM, 500nM) generates distinct SOX10 dosages. Importantly, SOX10 edited hESCs retain the ability to differentiate into hENSps - as evidenced by expression of Vagal HOX marker genes and neural crest TFs (Figure 1C). Furthermore, use of previously published in vitro protocols (*5*) allows for a platform to specify hENSps into enteric neurons and glia - allowing us to interrogate the role of SOX10 over the course of ENS development (Figure 1D).

Results:

We hypothesized that developmental enhancers and gene expression programs required for ENS development respond distinctly to changing SOX10 dosage. To assess how chromatin accessibility and gene expression respond to quantitative perturbations in SOX10 levels, we differentiated 3 SOX10 edited hESC lines into hENSps, and treated the derived cells with dTAGv1 at previously determined concentrations for either 3h or 24h (Figure 2A). We then performed bulk-transcriptomic profiling to measure gene expression changes and performed ATAC-seq to assess SOX10 dependence in maintaining chromatin accessibility. Full depletion of SOX10 (500 nM dTAGv-1, 24 h) alters both the

transcriptional and accessibility landscape of hENSps: we identify genes important for progenitor maintenance (the receptor tyrosine kinase (*ERBB3*), glial lineage characteristics (*PLP1, S100B),* and cell polarity and migration (*NEDD9)* that require SOX10 for expression. Of the 130,080 designated accessible regions identified through ATAC-seq, more than 30,805 peaks (5% FDR) are SOX10 dependent (Figure 2B). SOX10-dependent regions showed progressively decreasing accessibility with longer depletion duration, and many mapped near genes robustly expressed in hENSps. A subset of SOX10 dependent accessible regions are associated with dormant genes that are expressed later in ENS development such as the early neuronal specifier *ASCL1* and cholinergic neuronal marker *ChAT* (Figure 2C) - indicating that priming of distinct chromatin regions by SOX10 occurs early in ENS specification.

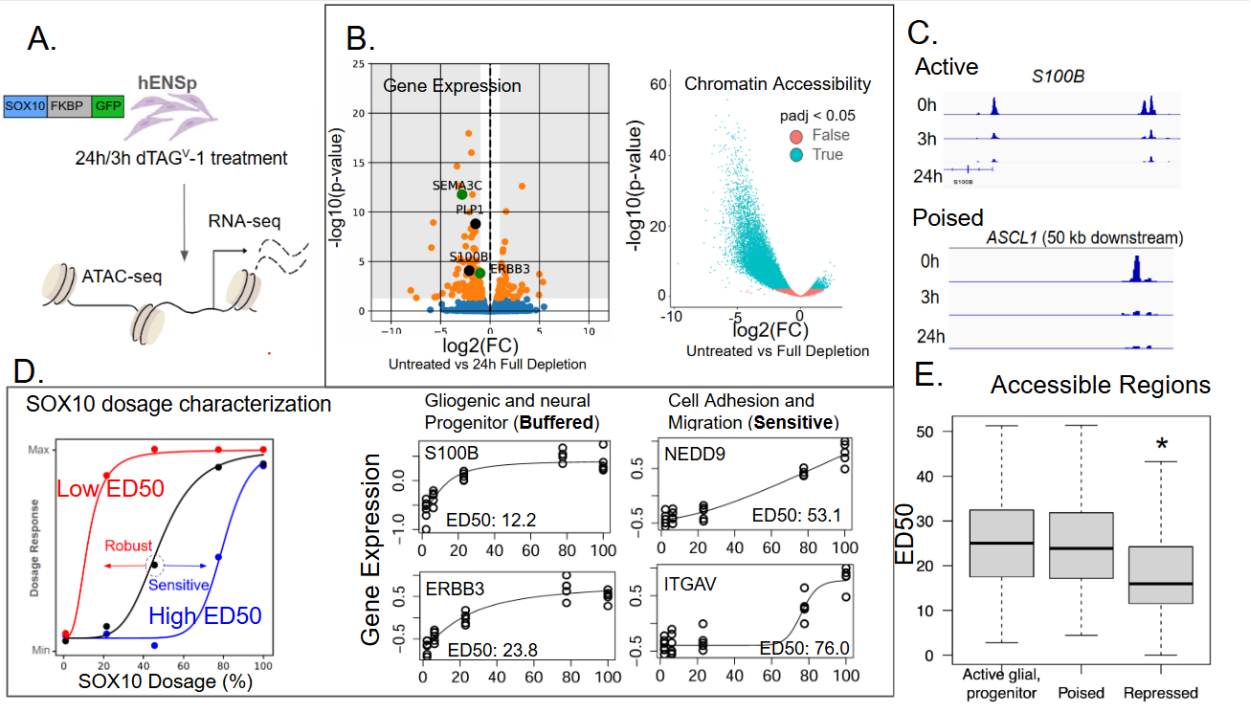


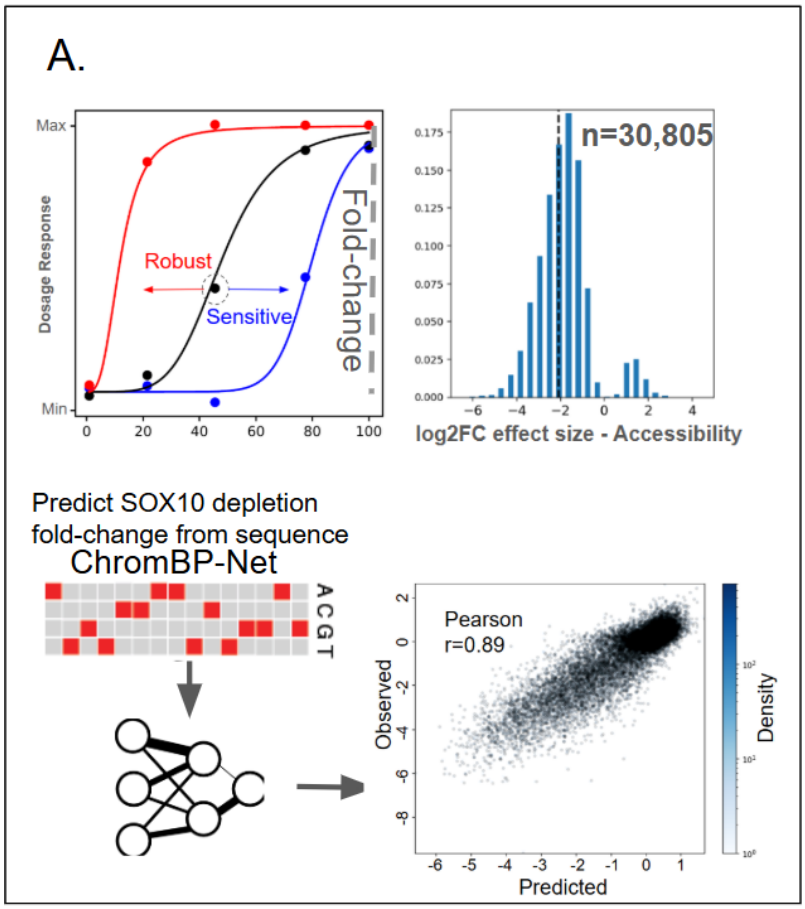
Figure 2: Molecular Characterization of SOX10 dosage effects in hENSp reveals biological processes that are sensitive to SOX10 depletion. A. Schematic of experimental procedure. **B.** Gene expression and chromatin accessibility changes after 24hr full depletion of SOX10 levels; highlighted genes drive proliferation, migration, and progenitor maintenance. **C.** Example accessible regions demonstrating SOX10 dependence. Active regions lie near genes expressed in hENSps; poised regions lie near genes expressed later during neuronal lineage commitment **D.** Hill functions showcasing curves that are robust or sensitive to SOX10 depletion. SOX10 dosage data across 5 dTAGv1 concentrations across genes driving distinct hENSp cellular functions. **E.** Distribution of fitted ED50 values across SOX10 dependent accessible genomic regions. Repressed regions are accessible loci in the vicinity of genes that increase in expression on SOX10 depletion.

Driven by our initial molecular characterization demonstrating the reliance of critical genes and their neighboring accessible loci on SOX10 activity, we fitted Hill functions to individual dosage curves for all SOX10-dependent genes and accessible regions and observed a wide range of dose responses, as quantified by fitted ED50 values (D). Genes involved in maintaining hENSp progenitor status, proliferative capacity, and early gliogenic potential tend to be highly buffered to changes in SOX10, while SOX10 dependent genes involved in cell polarity, adhesion, and motility are susceptible to small perturbations in SOX10 levels. Dosage responses of the SOX10 dependent accessibility regions confirm this observation: peak regions close to non-neuronal lineage specifying genes derepressed by loss of SOX10 activity (Repressed) demonstrate higher robustness (lower ED50 values) relative to other regions that regulate genes driving cell motility, migration, progenitor maintenance, and gliogenic programs (Active/Poised). Overall, loss of SOX10 seems to have critical effects on genes driving gut migration and colonization, while lineage specification gene programs tend to maintain expression across reduced SOX10 dosages - suggesting that SOX10 loss is concentrated in specific aspects of ENS progenitor activity.

Motivated by our observation that specific gene expression programs are more sensitive to SOX10 dosage than others, we asked what regulatory mechanisms shape these transcriptional responses. We focused on identifying sequence features that determine how large the SOX10 depletion effect (magnitude of depletion) is across SOX10 dependent chromatin accessible regions. We used the fold-change in accessibility across individual loci as this was a more robust

estimate than the fitted ED50 values. Given the wide range in effect sizes for the fold-change in accessibility across these genomic regions, we hypothesized that proximal DNA sequence features explain this variation.

Figure 3: Deep-learning model trained to predict SOX10 dependent chromatin accessibility changes uncovers TF interactions that shape hENSp gene regulation. A. Dynamic range of chromatin accessibility change assessed by ATAC-seq measurements in full SOX10 depletion condition and its distribution. A ChromBPNet model was trained to predict log2 fold-change in accessibility from DNA sequence alone on a held-out test set (Pearson r = 0.89 vs. mean-accessibility baseline] . B) Attribution of sequence sequence features from the neural network uncovers specific Transcription Factors modulating changes in chromatin accessibility upon SOX10 depletion.



We trained the ChromBP-Net model, a deep learning architecture that explicitly learns and isolates regulatory sequences driving chromatin accessibility (*5*), and fine-tuned the model to identify specific sequence features that predict the magnitude in accessibility changes on SOX10 depletion in hENSp (*6*). Applying the trained model on a held-out test set reveals strong model performance (Figure 3A), revealing that local DNA features in proximity to these SOX10 dependent accessible peaks are predictive. DeepLift-based attribution (7) identified the SOX10 palindrome and the neural crest TF motifs AP2 and NR2F as strong predictors of accessibility loss. Notably, AP2 and NR2F motifs also predicted peaks that gained accessibility upon SOX10 depletion, suggesting these TFs may regulate distinct targets through cooperative or antagonistic interplay with SOX10 at these regulatory elements. Together, these results indicate that the regulatory grammar of SOX10, NR2F, and AP2 binding across enhancers shapes the accessibility landscape of hENSps and gene programs essential for enteric neural progenitor function

Future Applications: Our approach using a degron-based strategy to modulate SOX10 levels in a human centered model of ENS development has uncovered specific gene regulatory programs that are particularly sensitive to SOX10 dosage mediated through the regulatory activity of SOX10 and other neural crest TFs such as AP2 and NR2F. The observation of chromatin accessibility poising – SOX10 dependent accessibility around genes that are expressed later in ENS development – raises the possibility of a sub-population of hENSps that are primed toward neuronal and glial lineages through the activity of these TFs. Single nucleus ATAC-seq on these hENSps will be informative in instructing if these poised-state subpopulations exits and in combination with SOX10 dosage reduction in our in vitro lineage specification platform (Figure 1D) allow us to determine if these poised cells are critically able to generate enteric neurons and glia by single-cell transcriptomic profiling across perturbations in SOX10 levels.

References

1. Rao M, Gershon MD. Enteric nervous system development: what could possibly go wrong? Nat Rev Neurosci. 2018 Sep;19(9):552-565. doi: 10.1038/s41583-018-0041-0. PMID: 30046054; PMCID: PMC6261281.
2. Betancur P, Bronner-Fraser M, Sauka-Spengler T. Genomic code for Sox10 activation reveals a key regulatory enhancer for cranial neural crest. Proc Natl Acad Sci U S A. 2010 Feb 23;107(8):3570-5. doi: 10.1073/pnas.0906596107. Epub 2010 Feb 5. PMID: 20139305; PMCID: PMC2840498.
3. Niesler B, Kuerten S, Demir IE, Schäfer KH. Disorders of the enteric nervous system - a holistic view. Nat Rev Gastroenterol Hepatol. 2021 Jun;18(6):393-410. doi: 10.1038/s41575-020-00385-2. Epub 2021 Jan 29. PMID: 33514916.
4. Nabet B, Roberts JM, Buckley DL, Paulk J, Dastjerdi S, Yang A, Leggett AL, Erb MA, Lawlor MA, Souza A, Scott TG, Vittori S, Perry JA, Qi J, Winter GE, Wong KK, Gray NS, Bradner JE. The dTAG system for immediate and target-specific protein degradation. Nat Chem Biol. 2018 May;14(5):431-441. doi: 10.1038/s41589-018-0021-8. Epub 2018 Mar 26. PMID: 29581585; PMCID: PMC6295913.
5. Barber K, Studer L, Fattahi F. Derivation of enteric neuron lineages from human pluripotent stem cells. Nat Protoc. 2019 Apr;14(4):1261-1279. doi: 10.1038/s41596-019-0141-y. Epub 2019 Mar 25. PMID: 30911172; PMCID: PMC6548871.
6. Pampari A, Shcherbina A, Kvon EZ, Kosicki M, Nair S, Kundu S, Kathiria AS, Risca VI, Kuningas K, Alasoo K, Greenleaf WJ, Pennacchio LA, Kundaje A. ChromBPNet: bias factorized, base-resolution deep learning models of chromatin accessibility reveal cis-regulatory sequence syntax, transcription factor footprints and regulatory variants. bioRxiv [Preprint]. 2025 Jan 8:2024.12.25.630221. doi: 10.1101/2024.12.25.630221. PMID: 39829783; PMCID: PMC11741299.
7. Naqvi S, Kim S, Tabatabaee S, Pampari A, Kundaje A, Pritchard JK, Wysocka J. Transfer learning reveals sequence determinants of the quantitative response to transcription factor dosage. Cell Genom. 2025 Mar 12;5(3):100780. doi: 10.1016/j.xgen.2025.100780. Epub 2025 Feb 27. PMID: 40020686; PMCID: PMC11960506.
8. Avsec Ž, Weilert M, Shrikumar A, Krueger S, Alexandari A, Dalal K, Fropf R, McAnany C, Gagneur J, Kundaje A, Zeitlinger J. Base-resolution models of transcription-factor binding reveal soft motif syntax. Nat Genet. 2021 Mar;53(3):354-366. doi: 10.1038/s41588-021-00782-6. Epub 2021 Feb 18. PMID: 33603233; PMCID: PMC8812996.