bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Title: Temporal analysis of enhancers during mouse brain development reveals dynamic regulatory function and identifies novel regulators of cerebellar development.

Authors: Miguel Ramirez[1,2], Yuliya Badayeva[1,2], Joanna Yeung[1,2], Joshua Wu[1,2], Erin Yang[1,2], FANTOM 5 Consortium[3], Brett Trost[4], Stephen W. Scherer[4], Daniel Goldowitz[1,2]*

Author Affiliations: 1. Centre for Molecular Medicine and Therapeutics, BC Children's Hospital Research Institute, Vancouver, BC, Canada V6H 3V5; 2. University of British Columbia, Vancouver, BC, Canada V6T 1Z4; 3. RIKEN, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan; 4. The Centre for Applied Genomics, The Hospital for Sick Children, Toronto, ON, Canada M5G 0A4

Corresponding Author: Daniel Goldowitz, Centre for Molecular Medicine and Therapeutics, BC Children's Hospital Research Institute, 950 W 28th Ave, Vancouver, BC, Canada V6H 3V5

Classification: Biological Sciences, Developmental Biology

Keywords: Enhancers, Brain development, Gene expression regulation, Epigenetics, Transcription factors, Mouse, Autism Spectrum Disorder

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

In this study, we identified active enhancers in the mouse cerebellum at embryonic and postnatal stages establishing the first catalog of enhancers active during embryonic cerebellum development. The majority of cerebellar enhancers have dynamic activity between embryonic and postnatal development. Cerebellar enhancers were enriched for neural transcription factor binding sites with temporally specific expression. Putative gene targets displayed spatially restricted expression patterns, indicating cell-type specific expression regulation. Functional analysis of target genes indicated that enhancers regulate processes spanning several developmental epochs such as specification, differentiation and maturation. We use these analyses to discover one novel regulator and one novel marker of cerebellar development: Bhlhe22 and Pax3, respectively. We identified an enrichment of *de novo* mutations and variants associated with autism spectrum disorder in cerebellar enhancers. Our study provides insight into the dynamics of gene expression regulation by enhancers in the developing brain and delivers a rich resource of novel gene-enhancer associations providing a basis for future in-depth studies in the cerebellum.

Introduction

Neuronal development is a complex and dynamic process that involves the coordinated generation and maturation of countless cell types. For the most numerous neuron in the brain, the cerebellar granule cell, neuronal differentiation consists of several steps beginning with the commitment of neural stem cells to become specified neural precursors, followed by multiple migratory stages to reach and mature at its final destination (Consalez, Goldowitz, Casoni, & Hawkes, 2021). Underpinning these events is the expression of gene regulatory networks that drive dynamic molecular processes required for proper brain formation (Ziats, Grosvenor, & Rennert, 2015). However, the transcriptional mechanisms that precisely regulate these gene expression programs have not been fully described.

Gene expression is typically activated when transcription factors (TFs) bind to non-coding regulatory elements and recruit the necessary components to begin transcription. Among the several classes of non-coding sequences that regulate gene expression, enhancers are the most common, with thousands predicted to coordinate transcriptional regulation during development (Heinz, Sven, Romanoski, Benner, & Glass, 2015). Enhancers are stretches of DNA

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

that bind to TFs and upregulate distal target gene expression. In the brain, enhancers help to ensure that gene expression is spatially- and temporally-specific, defining what genes will be active during distinct stages of development (Nord & West, 2020). Transcriptional regulation by enhancers has been shown to be critical for cellular identity, maturation during central nervous system (CNS) development, and activity-dependent responses in mature neurons (Frank et al., 2015; Pattabiraman et al., 2014). A detailed understanding of the enhancers that govern changes in gene expression during embryonic and early postnatal brain development remains limited. Profiling genome-wide enhancer activity at different time points and identifying their gene regulatory targets can provide insight into developmental processes regulated by enhancer elements.

Several molecular properties have been associated with enhancer activity, and the advancement of sequencing technology has facilitated their identification genome-wide in several developing brain structures (Carullo & Day, 2019). Enhancers are marked with histone post-translational modifications H3K4me1 and H3K27ac, both of which contribute to opening chromatin for TF binding (Calo & Wysocka, 2013). H3K27ac delineates active from poised elements, and has been a reliable marker for enhancer activity genome-wide (Creyghton et al., 2010). Analysis of these marks, in conjunction with transcriptomic and epigenomic datasets, has revealed that the vast majority of non-coding variants associated with neurological and psychiatric disorders are found within these regulatory elements, highlighting their importance in functional readout in the brain (Barešić, Nash, Dahoun, Howes, & Lenhard, 2020). Thus, profiling enhancer-associated histone modifications in the brain across time provides a comprehensive understanding of gene-regulatory principles, disease-associated variants, and the genetics of brain development (Nott et al., 2019).

The cerebellum has been a long-standing model to study the developmental genetics of the brain. This is, in part, due to the limited number of cell types, well-defined epochs of development for these cell types and a simple trilaminar structure in which these cells are organized, making for an enhanced resolution of events in time and space (Wang, V. Y. & Zoghbi, 2001). More recently, the study of cerebellar development has gained added interest through its documented role in the etiology of ASD (Stoodley & Limperopoulos, 2016). Previously, we developed a 12-timepoint transcriptional analysis of the developing cerebellum leading to the discovery of novel TFs critical for proper development (Zhang, P. G. Y. et al., 2018). More

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

recently, the developing cerebellum has served as an ideal setting for pioneering single-cell RNA-seq time course studies (Carter et al., 2018; Peng et al., 2019; Wizeman, Guo, Wilion, & Li, 2019). At the level of gene expression regulation, chromatin accessibility and enhancer activity have been examined previously in the postnatal cerebellum, leading to the discovery of distinct transcriptional profiles between immature and mature neurons, coordinated by non-coding *cis*-regulatory sequences (Frank et al., 2015). However, a comprehensive atlas of enhancers defining the role they play during embryonic and early postnatal cerebellar development has yet to be established. Profiling these non-coding regulatory elements and their target genes will discover novel genetic drivers of the precisely-timed and cell-specific molecular events in the developing cerebellum.

We utilize chromatin immunoprecipitation followed by sequencing (ChIP-seq) of enhancer associated histone marks H3K4me1 and H3K27ac at 3 stages of embryonic and early postnatal cerebellar development. We identify temporally specific enhancers using a differential peak analysis comparing postnatal and embryonic timepoints. Transcription factor motif enrichment and prediction of gene targets led to the elucidation of molecular processes regulated by enhancers during these stages. We use these analyses to discover two novel regulators of cerebellar development, Pax3 and Bhlhe22: a novel marker of GABAergic progenitors and a regulator of postnatal granule cell migration, respectively. Finally, we identify an enrichment of autism spectrum disorder (ASD) associated SNPs and *de novo* variants found in ASD-affected individuals in cerebellar enhancers, functionally annotating ASD-associated variation. Our study provides further insight into the dynamics of gene expression regulation by enhancers in the developing brain and delivers a rich resource to help understand the developmental and functional genetics of the developing cerebellum.

Results

Enhancer identification during cerebellar development

To identify enhancers active during embryonic and postnatal cerebellar development, we generated genome-wide H3K27ac and H3K4me1 ChIP-seq profiles from mouse cerebella dissected at embryonic day 12 (E12), postnatal day 0 (P0) and postnatal day 9 (P9) (**Figure 1A**). These developmental days represent 3 distinct stages of murine cerebellar development, each with its own developmental profile (Goldowitz & Hamre, 1998). H3K27ac and H3K4me1 signals

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

were reproducible between biological replicates as exemplified in a region on chromosome 14 (**Figure 1B**). There was a high correlation between replicates for both marks at each age (**Supplementary Figure 1A**). Therefore, we had confidence in using our H3K27ac and H3K4me1 data in downstream analyses. Robust cerebellar enhancers were identified by the presence of overlapping peaks between the two enhancer-associated histone marks at each age. This highlighted a total of 9,622 peaks; 5,859, 474, and 3,289 peaks that were in both the H3K27ac and H3K4me1 datasets at E12, P0, and P9, respectively (**Figure 1C**). Duplicate peaks between ages were removed, producing a list of **7,024** active cerebellar enhancers derived from overlapping H3K27ac and H3K4me1 signals (Supplementary Data 1).

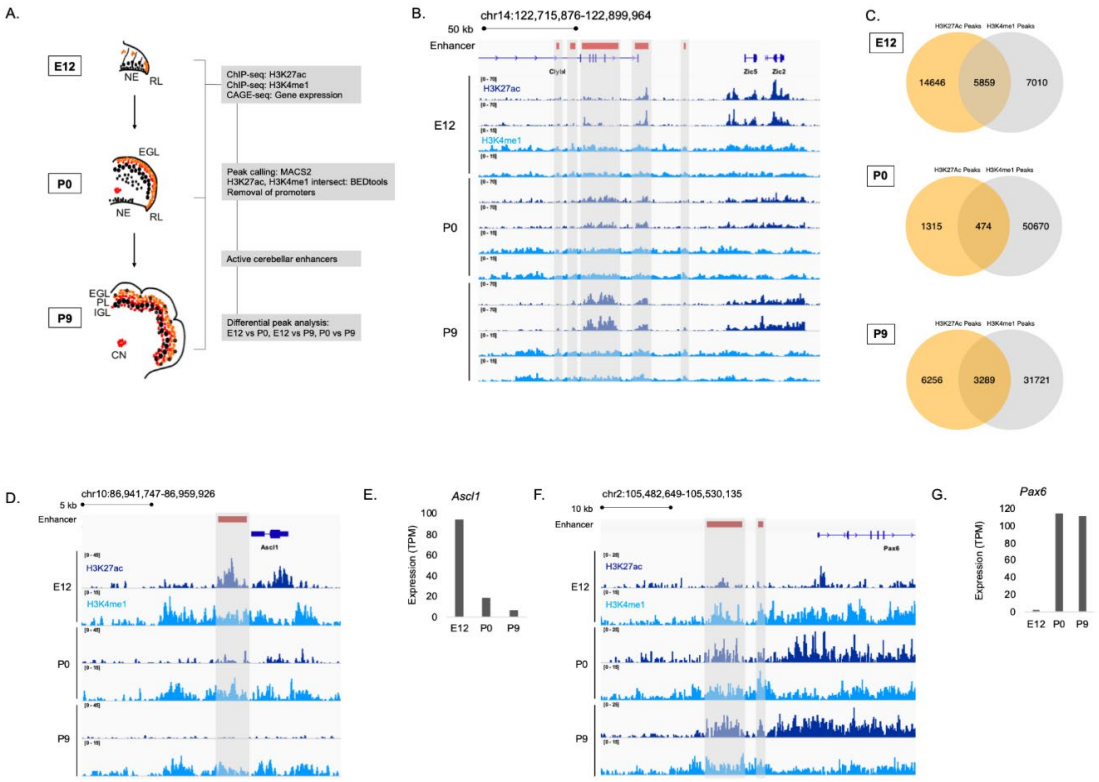

Figure 1. Enhancer identification during cerebellar development. A) An overview of the stages of cerebellar development profiled in this study. The datasets collected at these ages and the downstream analyses are shown. Labels: NE: Neuroepithelium, RL: Rhombic lip, EGL: External granular layer, PL: Purkinje layer, IGL: Inner granular layer, CN: Cerebellar nuclei **B)** A region of the mouse genome chr14:122,715,876-122,899,964 (mm9) in the Integrative Genomics Viewer (IGV) showing H3K27ac and H3K4me1 profiles across biological replicates of E12, P0, P9 cerebella. Active cerebellar enhancers are highlighted (gray box). **C)** Venn diagrams displaying overlap between H3K27ac and H3K4me1 peaks at each E12, P0 and P9. **D-E)** An example of a cerebellar enhancer identified from the E12 cerebella. Shown is normalized H3K27ac and H3K4me1 signal at the enhancer (gray box), as well as **(E)** normalized CAGE-seq expression of the nearest gene, *Ascl1*, across developmental time, at E12, P0, P9. TPM, Transcripts Per Million. **(F-G)** An example of a cerebellar enhancer identified from the P9 cerebella. Shown is normalized H3K27ac and H3K4me1 signal at the enhancer (gray box), as well as **(G)** normalized (TPM) CAGE-seq expression of the nearest gene, *Pax6*, across developmental time, at E12, P0, P9.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The relationship between enhancer activity and genes relevant to cerebellar development is shown in genomic regions flanking *Ascl1* and *Pax6,* two genes critical to cerebellar development (Kim, Battiste, Nakagawa, & Johnson, 2008; Yeung, Ha, Swanson, & Goldowitz, 2016b). We identified an enhancer active at E12 located in close proximity to *Ascl1* (**Figure 1D**). A decrease in the H3K27ac ChIP-seq signal at this enhancer corresponded to a decrease in *Ascl1* gene expression (**Figure 1E**). We identified two active enhancers at P9 located near *Pax*6 (**Figure 1F**). H3K27ac ChIP-seq signal also showed a pattern of activity similar to *Pax6* expression, increasing from embryonic to postnatal ages (**Figure 1G**). These results provide validation for the enhancers identified in our dataset in regulating genes critical to cerebellar development.

We compared our list of robust cerebellar enhancers to three previously published enhancer datasets. First, P7 H3K27ac ChIP-seq and DNase-seq profiles previously generated by Frank et al (2015) were compared to robust cerebellar enhancers. Greater than 90% of our reported cerebellar enhancers are replicated by H3K27ac and DNAse-seq peaks from this study (**Supplementary Figure 1B-C**). Second, enhancers retrieved from the enhancer database EnhancerAtlas 2.0, reporting enhancer activity in the mouse cerebellum at P0-P14 (Gao & Qian, 2020), were compared to robust cerebellar enhancers were compared to. We found that 73%, and 80% of our enhancers overlapped with the postnatal cerebellum enhancer dataset at P0, and P9, respectively (**Supplementary Figure 1D**). Third, mouse enhancers that had experimentally validated hindbrain activity at E11.5 from the VISTA Enhancer Browser (Visel, Minovitsky, Dubchak, & Pennacchio, 2007) were compared to cerebellar enhancers. We found that 56% of VISTA enhancers overlap with our cerebellar enhancer sequences at E12 (**Supplementary Figure 1E**). These confirmative findings indicate our approach was effective in capturing active cerebellar enhancers.

Enhancer dynamics during cerebellar development

The dynamics of enhancer activity over cerebellar development were examined through a differential peak analysis of H3K27ac signal. The majority, **89% (6238/7023),** of cerebellar enhancers had significant differences in peak signal (adjusted p-value ≤ 0.05) throughout cerebellar development (**Figure 2A**). At P9, **1273** cerebellar enhancers were significantly active compared to either P0 or E12 (Supplementary Data 2). At E12, **4432** active enhancers were

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

differentially active compared to either P9 or P0 (Supplementary Data 2). At P0, in contrast, only a small number of enhancers with differential signal was identified (403 and 154 showed significant changes when compared to E12 and P9, respectively). However, none of these P0 cerebellar enhancers were differentially active when compared to both E12 and P9, indicating that enhancer activity did not spike at birth. Taken together, this analysis highlights two temporally specific windows of enhancer activity at Early (embryonic) and Late (postnatal) stages (**Figure 2B).**

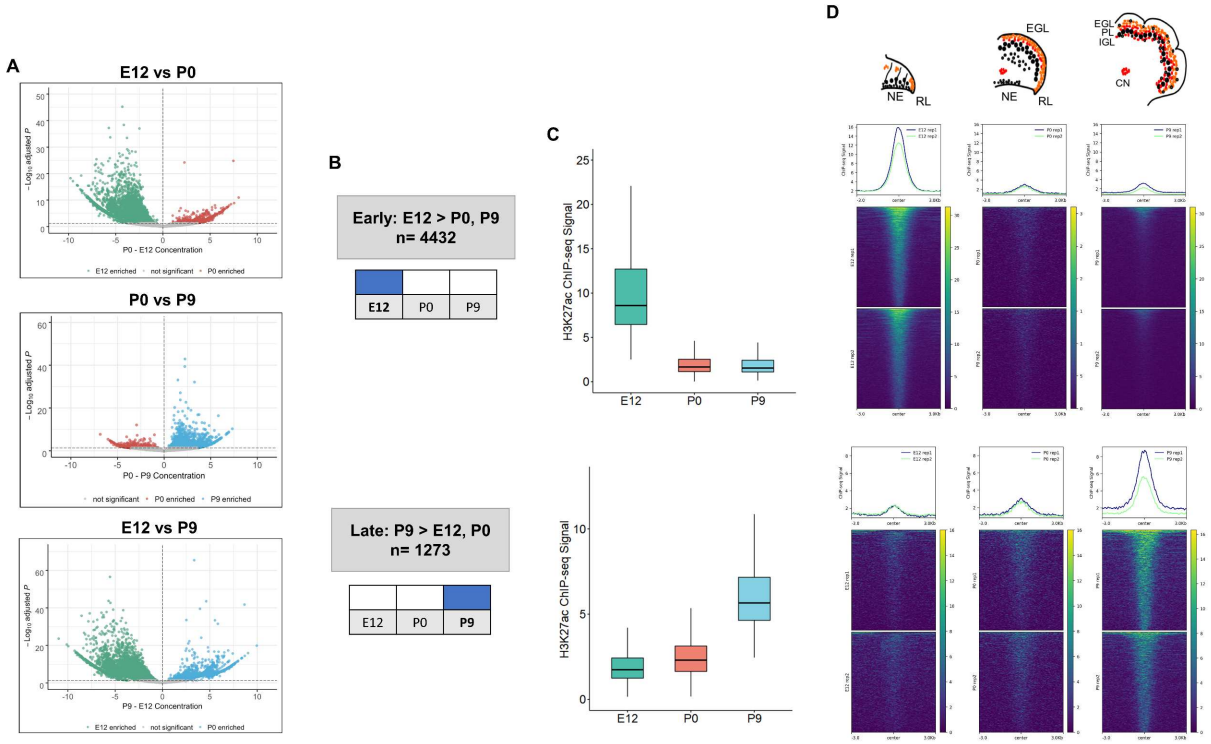

Figure 2. Enhancer activity is dynamic throughout cerebellar development. A) Volcano plots showing robust cerebellar enhancers with differential H3K27ac peak signal for three comparisons: E12 vs P9, E12 vs P0, and P0 vs P9. Differential signal strength was identified for 4433 and 4355 robust cerebellar enhancers when comparing E12 to P9 and to P0, respectively. At P9, 1275 and 403 robust cerebellar enhancers had differential signal when compared to E12 and P0, respectively. Enhancers with significant differential activity are colored at a cutoff of an adjusted p-value < 0.05. Displayed on the y-axis is the negative log10 adjusted p-value and on the x-axis is the difference in ChIP-seq signal between to the ages for a given peak. **B)** A diagram displaying how Early and Late active cerebellar enhancers were classified based on differential peak analysis results. **C)** A boxplot showing mean ChIP-seq signal (y-axis) for all Early (upper) and Late (lower) active enhancers. Error bars represent the standard error of the mean. **D)** Mean profile and heatmaps of H3K27ac signal at the midpoint of our predicted cerebellar enhancers (rows ± 3kb) in Early and Late groups at E12, P0 and P9.

Distinct patterns of enhancer activity were observed for temporally classified enhancers. For Early active enhancers, there was a loss of mean H3K27ac signal over time, with a steep decline after E12 (**Figure 2C**). Late active enhancers exhibited a gain in activity over time, with

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mean H3K27ac signal increasing steadily through development. These patterns are seen when looking at the changes in signal flanking the summits of our peaks across time (**Figure 2D**). These results indicate that the majority of cerebellar enhancers are dynamic throughout time and exhibit temporally specific activity.

Cerebellar enhancers are enriched for neural transcription factor binding sites in an age-dependent manner

We then sought to identify transcription factors whose activity is dictated by the availability of robust cerebellar enhancers, as many neural lineage-defining factors drive cell commitment in the developing brain through enhancer binding (Elsen et al., 2018; Lindtner et al., 2019). We used HOMER to search for enriched motifs (adjusted p-value < 1E-11) in Late and Early active cerebellar enhancers and then matched them to known transcription factor motifs in the JASPAR database (Heinz, S. et al., 2010). This analysis revealed a distinct set of significantly enriched motifs for Early and Late enhancers matching predicted TFs with both known and novel regulatory roles in cerebellar development (**Figure 3A**). TFs enriched in the Early active enhancers show a decrease in expression over time while TFs enriched in the Late active enhancer group show an increase in expression over time. This correspondence between enriched TF expression and enhancer activity provides validation for our findings and indicates the timing of enhancer activity may be dictated by the expression and binding of these enriched TFs.

The top three enriched TF motifs for Early active enhancers were Ascl1, Meis2 and Atoh1 (**Figure 3A**). These TFs have established roles in cerebellar development, acting as markers of GABAergic or glutamatergic cell types and regulators of differentiation (Ben-Arie et al., 1997; Kim et al., 2008; Wizeman et al., 2019). Importantly, many of the motifs enriched in the Early group matched with TFs which have received little to no attention in the cerebellum, including Sox4, Lhx2, Rfx4, Pou3f1 and Pax3 (**Figure 3A**). These TFs have been previously associated with the development of other brain areas (Frantz, Bohner, Akers, & McConnell, 1994; Porter et al., 1997; Su et al., 2016; Zhang, D. et al., 2006). In contrast, the TFs matching the motifs enriched in the Late active enhancers have a previously identified role in cerebellar development; but not necessarily involved in the same processes (**Figure 3A)**. For example, the top 3 enriched motifs matched with Neurod1, Nfia/b/x, and NF1, which have all been associated with granule cell differentiation (Miyata, Maeda, & Lee, 1999; Sanchez-Ortiz et al., 2014; Wang, W. et al., 2007).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

However, two other TFs with enriched binding sites, Pax6 and Smad4 have been found to be critical for granule cell precursor proliferation, a process preceding differentiation (Swanson & Goldowitz, 2011). These results suggest a dynamic role for the majority of our Early and Late active enhancers, driven by TFs involved in distinct stages of neuron development.

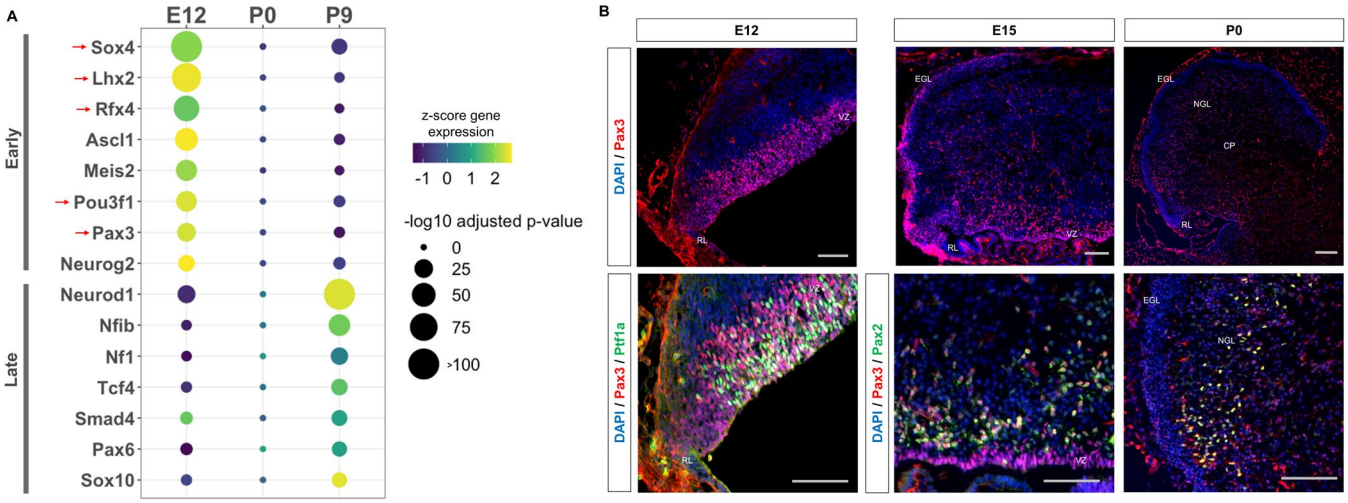

Figure 3. Neural transcription factors with known and novel function in the developing cerebellum are enriched in dynamic cerebellar enhancers. A) Dot plot displaying significantly enriched (adjusted p-value < 1E-11) motifs and the predicted matching transcription factor (TF). Displayed are the results for Early (top) and Late (bottom) active enhancers. TFs with an unknown functional role in cerebellar development are indicated with a red arrow. Size of the dots indicate the negative log10 adjusted p-value for a given motif and the color scale displays the z-score normalized expression throughout the cerebellar developmental time course. **B**) Top: Immunofluorescent staining of Pax3 in the mouse cerebellum at E12, E15 and P0. Bottom: Pax3 and Ptf1a immunofluorescent co-staining of the E12 mouse cerebellum. Immunofluorescent co-staining of Pax3 and Pax2 in the mouse cerebellum at E15 and P0. Labels: CP: Cerebellar parenchyma, EGL: External granular layer, NGL: Nascent granular layer, RL: Rhombic lip, VZ: Ventricular zone, Scalebars = 100um.

Early active enhancers are enriched for Pax3 binding sites, a novel marker for GABAergic cells

The TF motif enrichment analysis of Early enhancers led to the discovery of several TFs with novel in the context of embryonic cerebellar development; potentially involved in seminal aspects of development such as cellular specification or commitment. As a case study, we focused on Pax3, as other members of the Pax protein family have been shown to play key roles in the developing cerebellum (Leto et al., 2009; Urbánek, Fetka, Meisler, & Busslinger, 1997; Yeung, Ha, Swanson, & Goldowitz, 2016a). Indeed, we observed robust expression in the ventricular zone (VZ); a neural progenitor region for GABAergic cells in the cerebellum (Leto, Carletti, Williams,

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Magrassi, & Rossi, 2006) (**Figure 3B**). Colocalization between Pax3 and Ptf1a, the GABAergic lineage-defining molecule in the cerebellum (Hoshino et al., 2005), confirmed expression within GABAergic neural progenitors. At E15, Pax3+ cells are seen in the region just dorsal to the VZ, which consist of post-proliferative cells such as Purkinje cells and interneurons (Hoshino et al., 2005; Leto et al., 2006) (**Figure 3B**). We examined Pax3 co-labeling with markers for these cell types Foxp2 and Pax2, respectively (Fujita et al., 2008; Maricich & Herrup, 1999). While colocalization between Pax3 and Pax2 was found (**Figure 3B),** no co-staining between Pax3 and Foxp2 was observed (**Supplemental Figure 3A**). These results extend to P0, where Pax3+ cells are found in the nascent granule cell layer as well as the cerebellar parenchyma; ie co-labeling with Pax2 and not Calbindin, a Purkinje cell marker (**Figure 3E, Supplemental Figure 3B**). Thus, Pax3 is a novel marker for GABAergic progenitors and interneuron precursors in the developing cerebellum.

Co-expressed putative target genes are expressed in spatially distinct areas of the developing cerebellum.

We next investigated the molecular processes regulated by robust cerebellar enhancers through predicting their downstream targets (Osterwalder et al., 2018; Yao et al., 2015). This was done by calculating the correlation between H3K27ac signal and gene expression at E12, P0, and P9 (Zhang et al., 2018) for genes located within the same conserved topological associating domain (TAD) identified previously (Dixon et al., 2012) (See Methods). Overall, at least one positively correlated target gene was identified for **5815/7023 (70.61%)** cerebellar enhancers with an average Pearson correlation coefficient of **0.86** (Supplementary Data 3). In total, we identified **2261** target genes. Using the Mouse Genome Informatics (MGI) database, we identified **98** target genes that when knocked out result in a cerebellar phenotype; demonstrating the validity of our high throughput approach.

An unbiased *k-means* clustering was then conducted for Early and Late target genes to delineate them into the various co-expression programs coordinating molecular events during development. For this analysis, the target gene expression time course was expanded to 12 different timepoints during cerebellar development, quantified previously by CAGE-seq (Zhang et al., 2018). For Early active enhancers, 4 Clusters of co-expressed target genes were identified (**Figure 4A**). Genes in these clusters had decreased expression over time, similar to their

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

corresponding enhancer activity. However, a distinct mean expression profile was observed for each Cluster (**Figure 4B)**. Interestingly, genes with known function during cerebellar development showed distinct spatial expression patterns, observed using ISH data from the Developing Mouse Brain Atlas (Thompson et al., 2014) (**Figure 4C**).

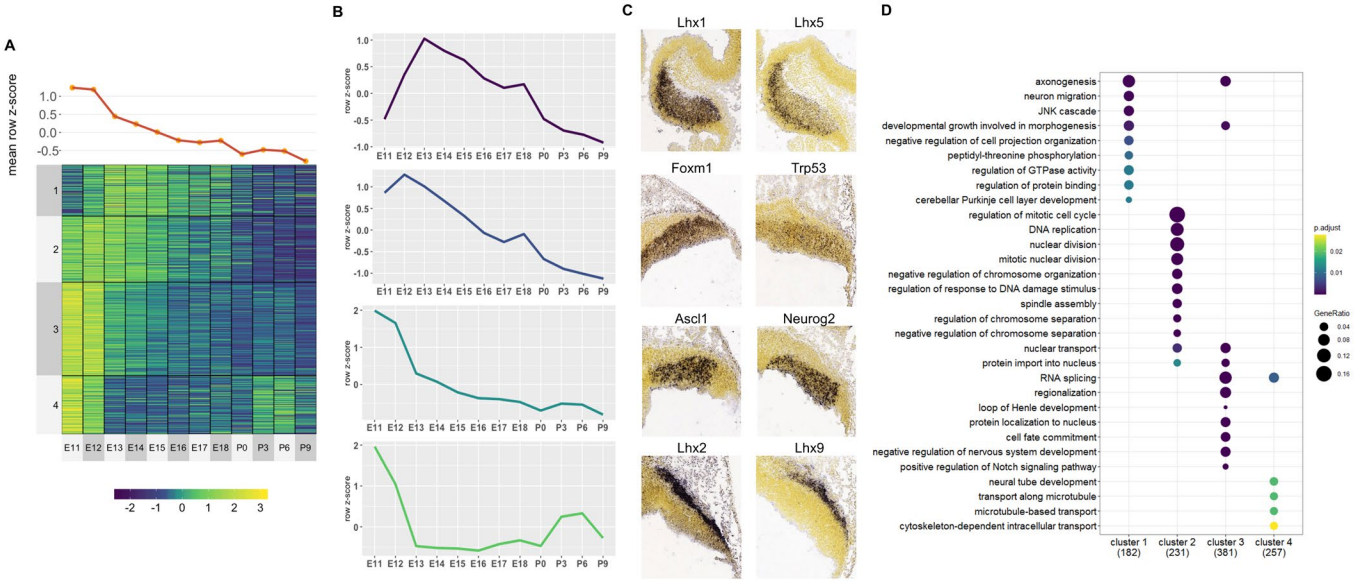

Figure 4. Co-expressed Early target genes are expressed in spatially distinct areas and have diverse roles in cerebellar development. A) Line plot and heatmap showing mean z-score expression for Early target genes throughout the cerebellar time course. **B)** Line graph representation of expression pattern throughout time for each cluster. **C)** Known cerebellar genes in each cluster and *in situ* hybridization (ISH) images showing spatial expression at peak expression ages. ISH images were taken from the Developing Mouse Atlas conducted at E13.5 for clusters 1 and 2, and E11.5 for clusters 3 and 4. **D)** Gene Ontology (GO) enrichment analysis of target genes from each cluster, displaying the top enriched GO terms. Size of the dots indicates the gene ratio for a given cluster which is equal to the number of genes within the GO category divided by the total number of genes in the cluster. Color scale indicates the adjusted p-value for each GO term.

For example, in **Cluster 3**, cerebellar genes *Ascl1* and *Neurog2* are expressed exclusively in the ventricular zone at E11.5 while **Cluster 4** contains *Lhx9* and *Meis2* which are expressed in the Nuclear Transitory Zone (neurons destined for the cerebellar nuclei). A Gene Ontology (GO) enrichment analysis revealed that each cluster is enriched for molecular processes known to be regulated by cerebellar genes within the cluster (**Figure 4D**). For example, **Cluster 1** is enriched for **axonogenesis** (GO:0007409, p-value: 3.31E-4), **neuron migration** (GO:0001764, p-value: 3.3E-4) and **Purkinje layer development** (GO:0021691, p-value: 0.01) and also contains *Lhx1* and *Lhx5* which are expressed in migrating Purkinje cells in cerebellar parenchyma and has previously been associated with the regulation of Purkinje cell differentiation during embryonic

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cerebellar development (Zhao et al., 2007). Together, these findings support the notion that Early active enhancers regulate their targets in a spatially-specific manner, regulating distinct processes in their respective cell types.

For the Late active enhancers, 4 Clusters of co-expressed target genes were identified (**Figure 5A**). We observed relatively distinct expression patterns in each of the 4 Clusters with a gradual rise in mean expression throughout time **(Figure 5B)**. Genes with known function during cerebellar development also show distinct spatial expression patterns, identified using the Developmental Mouse Atlas (**Figure 5C**). For example, **Cluster 1** and **3** contained known cerebellar genes critical for granule cell development, such as *Neurod1* and *Zic1*, while **Cluster 2** and **4** contained cerebellar genes important for in Purkinje cell development, such as *Atxn1* and *Hcn1* (**Figure 5C**) (Aruga & Millen, 2018; Ebner et al., 2013; Miyata et al., 1999; Rinaldi et al., 2013).

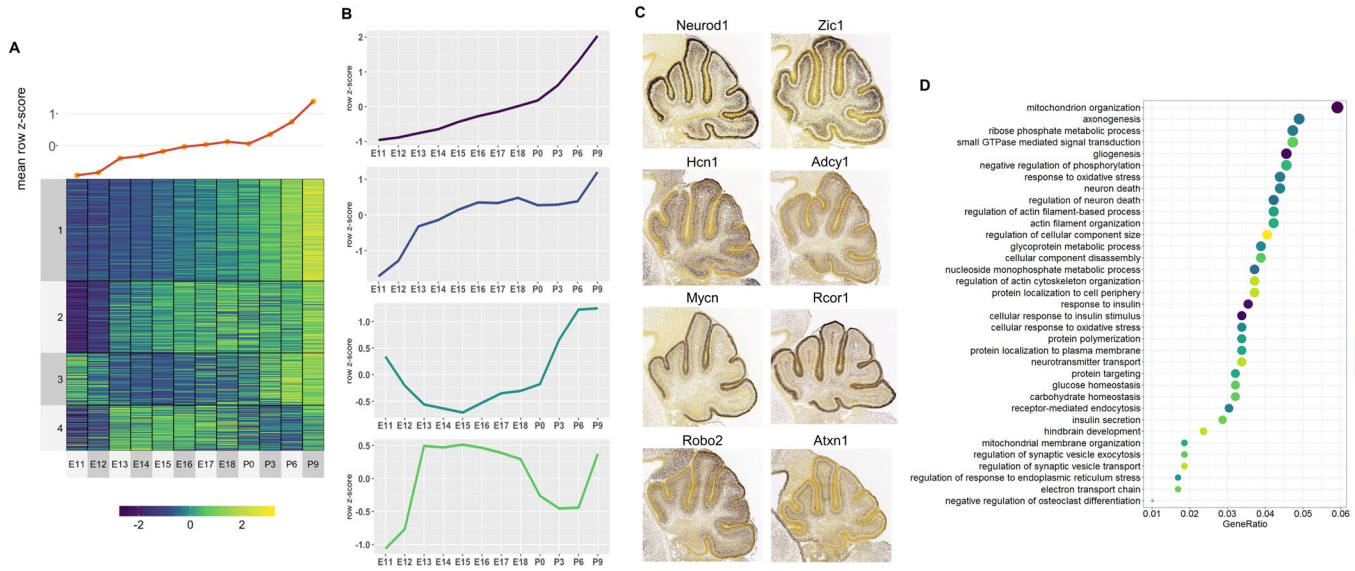

Figure 5. Co-expressed Late target genes are expressed in developing granule cells or Purkinje cells with common roles in cerebellar development. A) Line plot and heatmap showing mean z-score expression throughout the cerebellar time course. **B)** Line plot representation of expression pattern throughout time for each cluster. **C)** Known cerebellar genes in each cluster and *in situ* hybridization showing spatial expression at peak expression ages. ISH images were taken from the Developing Mouse Atlas provided by the Allen Brain Atlas conducted at P4.5 for all clusters. **D)** Gene Ontology (GO) enrichment analysis of all target genes of Late active enhancers, displaying the top enriched GO terms. Size of the dots indicates the number target genes within a given GO term and the gene ratio (x-axis) is equal to the number of genes within the GO category divided by the total number of genes in the cluster. Color scale indicates the adjusted p-value for each GO term.

A GO enrichment analysis was conducted for each Cluster; with no significantly enriched Cluster-specific GO terms. However, if all Late enhancer target genes were combined, several enriched GO terms emerged including ones involved in postnatal neuronal development,

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

such as **neuron death** (GO:0070997, p-value: 0.003), **neurotransmitter transport** (GO:0006836, p-value: 0.006) and **regulation of synaptic vesicle exocytosis** (GO:2000300, p-value: 0.005) (**Figure 5D**). Overall, this analysis provides a working framework for the placement of hundreds of genes into the overall structure of embryonic or postnatal cerebellar development

***Bhlhe22* is a novel regulator of granule cell development**

To demonstrate the utility of our results, we sought to identify target genes not previously identified in cerebellar development. We focused on Late Cluster 1, which contained target genes expressed in granule cells. We hypothesized that genes within this cluster regulated postnatal granule cell differentiation. To identify genes in this cluster regulating granule cell development, we filtered these genes relative to their interaction with Atoh1, the lineage defining molecule for granule cells and other glutamatergic neurons in the developing cerebellum (Ben-Arie et al., 1997). The genes were filtered using the following criteria: 1) Atoh1 is bound to the predicted enhancer during postnatal development (Klisch et al., 2011) and 2) the genes are differentially expressed in the Atoh1-null mouse (Klisch et al., 2011). Among the top 15 genes in the filtered list, we identified **4** novel genes and **11** genes that have previously been implicated in postnatal granule cell development (Supplementary Table 1). The known genes provided validation for our approach. The novel genes included *Bhlhe22* (also known as *Bhlhb5*), *Purb*, *Klf13* and *Sox18*. We focused particularly on *Bhlhe22* as it has previously been implicated in the differentiation of neurons in the cortex (Joshi et al., 2008). An enhancer ~2 kb upstream of the *Bhlhe22* transcriptional start site was identified and is bound by Atoh1 during postnatal development (**Supplemental Figure 3A**). This enhancer displayed H3K27ac activity highly correlated (Pearson correlatoin coefficient = 0.98) to *Bhlhe22* expression, which consistently rises throughout cerebellar development and peaks at P9.5 (**Supplemental Figure 3B**).

To attain a cellular resolution of the expression pattern for *Bhlhe22* over developmental time, a time-course of protein expression using immunofluorescent staining spanning cerebellar development was conducted. Bhlhe22 expression was observed in cells within the inner external

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

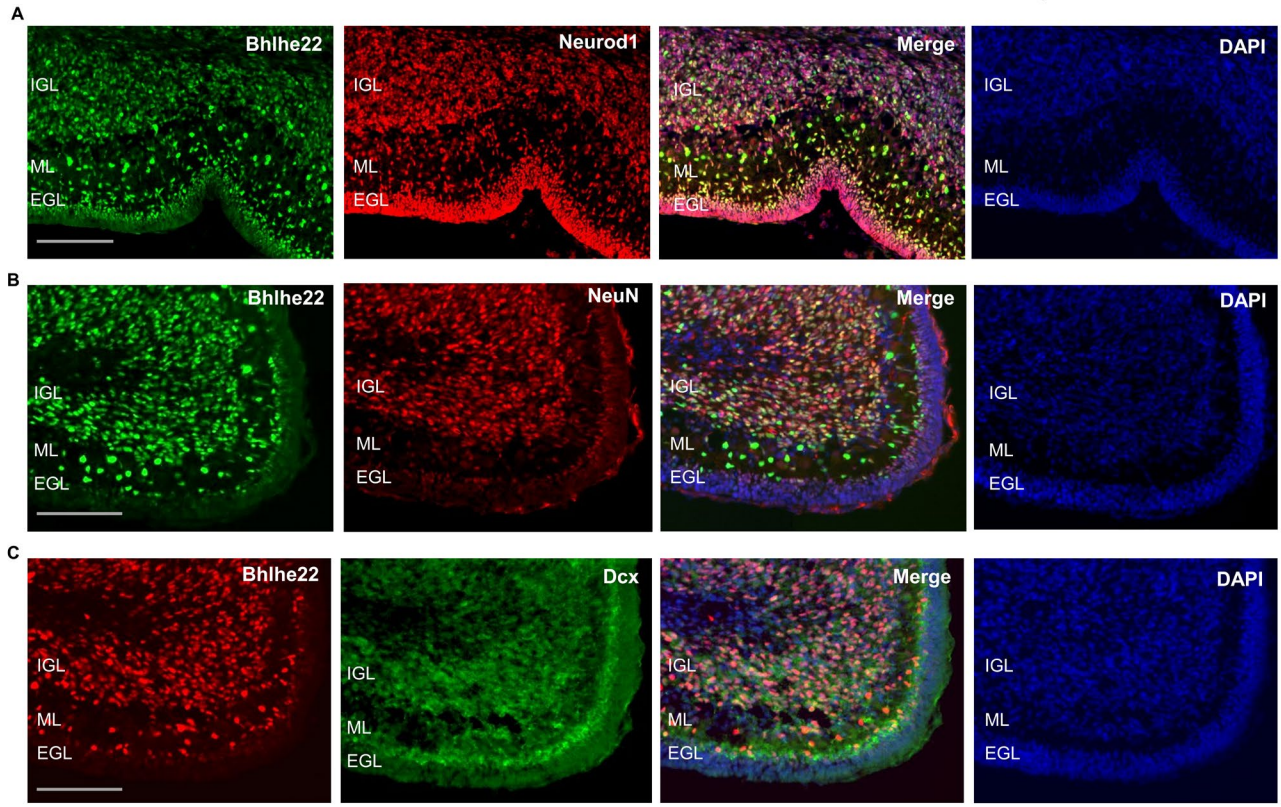

Figure 6. Bhlhe22 is expressed in differentiating granule cells in postnatal cerebellar development. A) Bhlhe22 (green) and Neurod1 (red) immunofluorescent co-staining at P9.5 of taken from a posterior lobe IX. **B)** Bhlhe22 (green) and NeuN (red) immunofluorescence co-staining at P6 taken from posterior lobe IX. **C)** Bhlhe22 (red) and Dcx (green) immunofluorescent co-staining at P6 showing the posterior lobe IX; Labels: EGL= external granular layer IGL = inner granular layer, ML = molecular layer, Scalebars = 100um.

granule layer (EGL), molecular layer and in the inner granule layer (IGL) of the postnatal cerebellum (**Figure 6A**). To identify whether Bhlhe22 is expressed in differentiating granule cells, co-staining experiments were performed with Neurod1 and NeuN which mark differentiating and mature granule cells, respectively (Miyata et al., 1999; Weyer & Schilling, 2003). At P6.5, colocalization between Bhlhe22 and Neurod1 was observed, indicating expression in differentiating and migrating granule cells (**Figure 6A**). Co-staining between Bhlhe22 and NeuN expression was also observed, indicating expression in maturing granule cells found within the IGL (**Figure 6B**). To confirm whether the Bhlhe22-positive cells within the molecular layer were migrating granule cells, we performed a double labelling experiment for a neuronal migration marker Doublecortin (Takács, Zaninetti, Vig, Vastagh, & Hámori, 2008). Colocalization between Doublecortin and Bhlhe22 was observed in cells within the inner EGL and the molecular layer, confirming Bhlhe22 expression in migrating granule cells (**Figure 6C**).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We then investigated the role that Bhlhe22 plays in postnatal granule cell development using a well-established *in vitro* system (Lee, Greene, Mason, & Manzini, 2009). Three sets of experiments were performed using isolated granule cells from P6.5 cerebella transfected with siRNA targeting Bhlhe22 transcripts (**Figure 7A**). First, to determine if Bhlhe22 expression was diminished, changes in gene expression were assessed after 3 days *in vitro* (DIV) using reverse transcriptase quantitative PCR (RT-qPCR). A 50% reduction of Bhlhe22 expression, on average, was found in treated cultures compared to controls (**Figure 7B**).

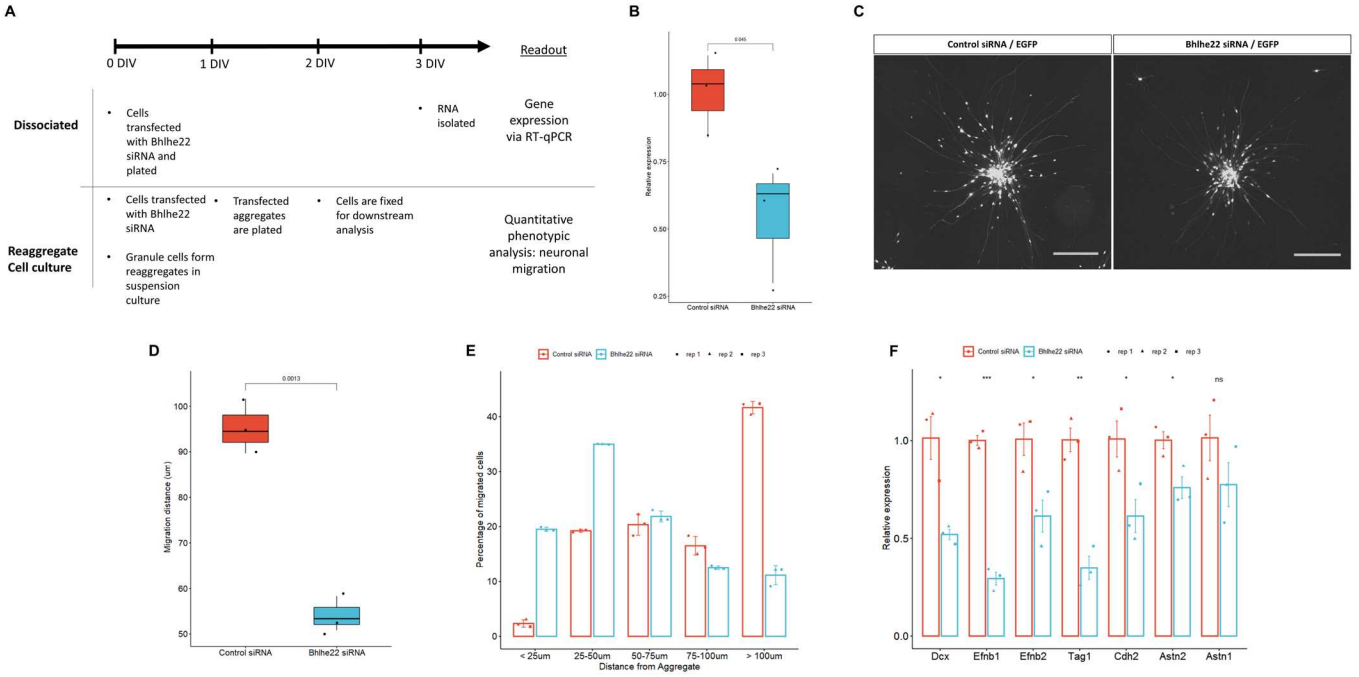

Figure 7. Knockdown of Bhlhe22 reduces migration of cultured cerebellar granule cells. A) Workflow for dissociated and reaggregate postnatal granule cell cultures. **B)** RT-qPCR analysis of Bhlhe22 gene expression in dissociated postnatal granule cell cultures after treatment with Bhlhe22 siRNA. Gene expression was normalized relative to the expression of the co-transfected EGFP protein to account for transfection variability between cultures. Data are presented as mean ± SD (n = 3). **C)** Image of cultured cerebellar granule cell reaggregates treated with control and Bhlhe22 siRNA. Shown are EGFP positive cells indicating successful transfection. Scalebars = 100um. **D)** Box plot displaying mean distance of granule cell migration from the aggregate. Value above indicates a statistical difference between control cultures and those treated with Bhlhe22 siRNA (p-value = 0.0013). **E)** Bar plot showing the percentage of cells migrated at different distances from the aggregate for control and Bhlhe22 siRNA treated cerebellar granule cell cultures. **F)** RT-qPCR analysis of gene expression of cell adhesion molecules in dissociated postnatal granule cell cultures after treatment with Bhlhe22 siRNA. Gene expression was normalized relative to the expression of the co-transfected EGFP protein to account for transfection variability between cultures. Data are presented as mean ± SD (n = 3). Symbols: *: p <= 0.05, **: p <= 0.01, ***: p <= 0.001, which indicate statistical differences observed between Bhlhe22 siRNA treated samples and controls. All error bars represent the standard error of the man

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Second, phenotypes of the transfected cells were examined; examining their neuritic outgrowths from the aggregate, and the migration of granule cells from the aggregates, within the first 24 hours of plating (Gartner, Collin, & Lalli, 2006). Neuritic outgrowth was unaffected, however there was a marked reduction in migration (**Figure 7C)**. Bhlhe22 siRNA transfected cells travelled on average 54.2um from the edge of the aggregate, a 50% reduction compared to control samples (**Figure 7D**). Examining the distribution of migrated cells from the edge of the aggregate, there was a higher percentage of Bhlhe22 siRNA transfected granule cells migrating less than 50um, while the majority of the cells in control samples migrated 100um and beyond (**Figure 7E**).

Third, changes in the expression of cell adhesion molecules that are known to be involved in granule cell development were assessed (Consalez et al., 2021; Wang et al., 2007). A significant reduction of *Efnb1, Efnb2, Tag1, Cdh2* and *Astn2* was observed in Bhlhe22 knockdown granule cell cultures compared to controls (**Figure 7F**). In addition to these genes, we also found a significant reduction in Doublecortin (*Dcx*) expression. Taken together, these *in vitro* knockdown experiments reveal a novel function for Bhlhe22, a gene that was identified by our temporal enhancer-target gene analysis and was predicted to have a critical role in postnatal granule cell development.

Active cerebellar enhancers are enriched for common and *de novo* genetic variants associated with autism spectrum disorder.

Genome wide association studies (GWAS) have revealed that the majority of variants associated with neurodevelopmental diseases are found within non-coding regulatory sequences, particularly enhancers (Visel, Rubin, & Pennacchio, 2009). Given the emerging importance of the cerebellum in the etiology of autism spectrum disorder (ASD), we tested whether ASD-associated variants are enriched in cerebellar enhancers. The software tool GREGOR (Genomic Regulatory Elements and Gwas Overlap algoRithm) was used to evaluate the enrichment of common genetic variants associated with ASD in cerebellar enhancers (Schmidt et al., 2015). ASD associated SNPs were retrieved from the GWAS Catalog (Buniello et al., 2019) and a stringent filter was applied to identify SNPs associated with the ASD (Supplementary Table 2). We examined 174 ASD-associated SNPs with a maximum p-value of 9E-06 (Buniello et al., 2019). ASD-associated SNPs were enriched in cerebellar enhancers (p-value = 2.34E-03) and in

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

H3K27ac peaks at E12, P0, and P9 (p-values of 1.29E-03, 1.05E-02 and 1.42E-04, respectively) (**Figure 8A**). For the 13 cerebellar enhancers containing ASD-associated SNPs, we identified 12 predicted target genes (Supplementary Table 3). Among these, three (*PAX6*, *TCF4*, and *ZMIZ1*) are ASD risk genes according to the Simons Foundation Autism Research Initiative (SFARI) gene database (Banerjee-Basu & Packer, 2010).

De novo mutations (DNMs) (variants present in the genome of a child but not his or her parents) have been found to play a significant role in the etiology of ASD, including those found in non-coding regions of the genome (Grove et al., 2019; Yuen et al., 2016). We hypothesized that DNMs within cerebellar enhancers would be more prevalent in ASD-affected individuals compared with their unaffected siblings. We used whole-genome sequencing data from 2,603 ASD-affected individuals and 164 unaffected siblings from the MSSNG cohort(C Yuen et al., 2017), as well as 2,340 ASD-affected individuals and 1,898 unaffected siblings from the Simons Simplex Collection (SSC) (Fischbach & Lord, 2010) to analyze the prevalence of DNMs in ASD-affected individuals compared with their unaffected siblings.

A

Dataset	Reported Number of SNPs	Expected Number of SNPs	P-value
Cerebellar enhancers	13	5.31	2.34E-03
E12 H3K27ac peaks	27	15.40	1.29E-03
P0 H3K27ac peaks	5	1.37	1.05E-02
P9 H3K27ac peaks	19	7.94	1.42E-04

B

	Number of DNMs					
Dataset	Enhancer affected	Enhancer unaffected	Non-enhancer affected	Non-enhancer unaffected	Odds ratio (confidence interval)	P-value
Cerebellar enhancers	1605	609	389636	161992	1.10 (1.00, 1.21)	0.055
E12 H3K27ac peaks	4321	1726	386920	160875	1.04 (0.98, 1.10)	0.16
P0 H3K27ac peaks	337	132	390904	162469	1.06 (0.87, 1.31)	0.58
P9 H3K27ac peaks	2041	796	389200	161805	1.07 (0.98, 1.16)	0.13
Combined Coordinates	7712	3016	383529	159585	1.06 (1.02, 1.11)	0.0043

C

Enhancer Coordindates (hg38)	Predicted Target Gene	CNV Coordinates (hg38)	CNV Type	CNV Size (bp)
chr10:13928845-13929693	*FAM107B*	chr10:13912983-13942132	Deletion	29,150
chr11:63909702-63910591	*STIP1*	chr11:63907636-63910623	Deletion	2,988
chr12:96838867-96841675	*NEDD1*	chr12:96825001-96842000	Duplication	17,000
chr14:76714160-76716619	*VASH1*	chr14:76711247-76714368	Deletion	3,122
chr14:103861041-103862026	*CDC42BPB*	chr14:103855620-103866732	Deletion	11,113
chr15:98960270-98962257	-	chr15:98937001-99031000	Duplication	94,000
chr19:47156912-47157972	*MEIS3*	chr19:47133001-47184000	Duplication	51,000

D

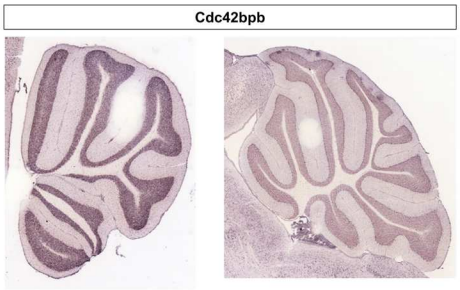

Figure 8. Cerebellar enhancers are enriched for GWAS SNPs and DNMs associated with ASD. A) Number of ASD-associated GWAS variants identified in cerebellar enhancers and H3K27ac peaks called from E12, P0, and P9 samples. **B)** Enrichment of *de novo* single nucleotide variants and indels in ASD-affected individuals compared with their unaffected siblings. Counts are not equal to the sum of the four enhancer types because some enhancers are categorized as more than one type. **C)** Gene targets for enhancers overlapped by *de novo* CNVs in the SSC cohort. **D)** *In situ* hybridization showing Cdc42bpb expression in the lateral (left) and medial (right) adult mouse cerebellum (REF mouse brain atlas). Note expression is found in granule cells, particularly those of the lateral cerebellum.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We found that DNMs (specifically *de novo* single nucleotide variants and indels) in cerebellar enhancers and H3K27ac peaks from E12, P0 and P9 were enriched in ASD-affected individuals, with odds ratios ranging from 1.04 to 1.10 (**Figure 8B**). While these differences were not statistically significant for cerebellar enhancers and peak coordinates individually, statistical significance was achieved when combined (odds ratio=1.06; p-value=0.0043). We also identified *de novo* CNVs overlapping cerebellar enhancers. Since the number of such CNVs was too small to perform statistical enrichment tests, we selected a subset of 7 of these CNVs (4 deletions and 3 duplications) for further characterization to identify candidates for association with ASD (**Figure 8C**). The most promising candidate was an ~11 kb deletion overlapping an enhancer predicted to target *CDC42BPB*, which has previously been implicated in neurodevelopmental phenotypes (Chilton et al., 2020). By visual validation in Integrative Genomics Viewer (Robinson et al., 2011), we verified that this deletion was truly *de novo* (Supplementary Figure 4**).** *CDC42BPB* is expressed in the granule cell layer of the adult mouse cerebellum (**Figure 8D**) and is expressed in the lateral aspects of the cerebellum but not the medial cerebellum.

Discussion

The current model of gene expression regulation during brain development posits that temporal and spatial transcription is under the intricate control of thousands of non-coding enhancer sequences (Nord & West, 2020). This model has emerged from the findings of several studies of enhancer activity in various parts of the developing brain (Nord et al., 2013; Pattabiraman et al., 2014; Visel et al., 2013). In our study on the cerebellum, we performed an assessment of enhancer activity through genome-wide profiling of H3K4me1 and H3K27ac deposition at three time points during embryonic and early postnatal times. These datasets were utilized to define functional enhancer elements with temporally specific activity during these developmental ages. In doing so, we establish the first catalog of predicted enhancers active during embryonic cerebellum development. Through a motif enrichment analysis, neural TF motifs were found to be enriched in cerebellar enhancers which may drive temporally specific enhancer activation. This analysis highlighted a novel regulatory role for several understudied TFs in the context of cerebellar development. These data were then integrated with a transcriptomic time course to identify predicted target genes to inform our understanding of enhancer regulation in the

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

466 developing cerebellum. Through unbiased clustering, we identified enhancer-regulated co-
467 expression gene programs with spatially distinct patterns of expression and unique biological
468 functions during embryonic and postnatal development. Further analysis of these results led to
469 the discovery of novel cell-type marker and regulator of cerebellar development and highlights
470 the importance of enhancer regulation during brain development and the etiology of ASD.

471

472 **Cerebellar enhancers regulate gene expression important for distinct stages of neuronal**
473 **development**

474 Identification of enriched TF binding sites and putative target genes indicated that
475 cerebellar enhancers likely play a regulatory role in various phases of neuronal development. In
476 agreement with our results, previous examinations of active non-coding regulatory sequences
477 revealed that neural progenitor cells and mature neurons exhibit distinct signatures of enhancer-
478 associated histone profiles, DNA methylation, chromatin conformation and enhancer-promoter
479 interactions (Bonev et al., 2017; de la Torre-Ubieta et al., 2018; Lister et al., 2013; Whyte et al., 2012).
480 Bonev et al. (2017) examined changes in enhancer-promoter interactions between transgenic cell
481 lines FACS sorted for embryonic stem cells, neural progenitors and mature neurons and
482 identified that changes in enhancer-promoter contacts are cell-state specific and correlate with
483 changes in gene expression (Bonev et al., 2017). When comparing contacts in neuro-progenitors
484 and mature neurons, a decrease in the interaction strength was observed between active domains
485 compared to an increased strength in inactive domains, indicating a shift in usage of regulatory
486 sequences. These changes were also reflected at the level of TF binding, as interactions at Pax6-
487 bound sites, a TF marking neural progenitors, were stronger in neural progenitors than in
488 neurons, while NeuroD2-bound sites, a TF marking mature neurons, were stronger in neurons
489 than NPCs (Bonev et al., 2017). This shift in enhancer usage throughout cortical development is
490 also reflected in DNA methylation profiles, where fetal enhancers are hypermethylated and
491 decommissioned in the adult brain, while enhancers regulating adult gene expression were
492 hypomethylated (Lister et al., 2013). Hypermethylation was accompanied by a decrease in
493 H3K4me1, H3K27ac and DNase hypersensitivity while the increase was observed after
494 hypomethylation (Lister et al., 2013). Our study supports the importance of temporally-specific

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

activity during different stages of neuron development *in vivo* and details the processes driven by enhancer-regulated expression during embryonic and early postnatal brain development.

Expression analysis of two genes novel to cerebellar development, Pax3 and Bhlhe22, supported the notion that enhancer profiles are specific to developmental stage. TF enrichment analysis identified Pax3 preferentially enriched in Early active enhancers and robust expression of Pax3 was localized to GABAergic interneuron progenitor cells. Interestingly, through further examination of cerebellar single-cell RNA-seq data produced by Carter et al. (2018), we found that Pax3 expression is enriched in GABAergic progenitors and differentiating GABAergic interneurons (Carter et al., 2018). Analysis of predicted gene targets of Late active enhancers identified Bhlhe22 as a novel gene expressed in postnatal differentiated granule cells, and *in vitro* knockdown experiments in primary granule cells indicated Bhlhe22 regulates granule cell migration potentially through regulation of cell adhesion molecule expression. These results are supported by findings in the developing cortex, where Bhlhe22 has been shown to regulate post-mitotic acquisition of area identity in layers II-V of the somatosensory and caudal motor cortices (Joshi et al., 2008). The contrasting expression profiles of Pax3 and Bhlhe22 highlight the wide-ranging developmental impact of enhancer-mediated gene expression regulation.

Our findings from the TF enrichment and gene target analyses generated from preferentially active postnatal enhancers indicate that many of the enhancers captured are active in the developing granule cells and Purkinje cells. We attribute this apparent bias to our whole tissue approach, as granule cells and Purkinje cells are the predominant cells in the cerebellum at that time, making it more likely to capture signals specific to these cells compared to other less abundant cell types. Our study therefore reveals that to elucidate enhancers specifically active within less abundant cell types, a more granular approach may be required through single-cell examination of chromatin accessibility, such as single cell ATAC-seq. This approach can be coupled with the abundance of scRNA-seq data that has been collected in the developing cerebellum. This strategy has seen success in the developing cortex and more recently in the cerebellum (Preissl et al., 2018; Sarropoulos et al., 2021).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Co-expressed gene targets of cerebellar enhancers display cell type-specific expression patterns

In addition to being temporally-specific, recent evidence indicates that enhancer activity is cell type specific in the brain (Blankvoort, Witter, Noonan, Cotney, & Kentros, 2018). This is highlighted in the identification of cerebellar enhancer target gene clusters for Early and Late active enhancers with cell specific patterns of expression (Figure 4 and 5). For example, distinct boundaries can be seen in gene expression from Early Clusters 3 and 4 at E11.5 between cells in the subpial stream (Cluster 4) and neuroepithelium (Cluster 3) where neural precursors of two separate lineages, the glutamatergic cerebellar nuclei and GABAergic cerebellar nuclei and Purkinje cells, are found, respectively. These sharp borders are reminiscent of the small domains of distinct enhancer activity identified in neural progenitors in the telencephalon, which were found to fate-map to specific prefrontal cortex subdivisions (Pattabiraman et al., 2014). We see a similar pattern in the more developed postnatal cerebellum, observing a spatial distinction between Late Clusters 1/3 and 2/4 delineating expression in granule cells and Purkinje cells, respectively. This cell type specific enhancer usage is demonstrated in the adult brain. Blankvoort et al (2018) (Blankvoort et al., 2018) used ChIP-seq analysis of microdissected subregions of the adult mouse cortex to reveal unique enhancer profiles pertaining to each region. Additionally, Nott et al (2019) (Nott et al., 2019) identified enhancer-promoter interactome maps specific to the major cell types in the cortex, which included neurons, microglia, oligodendrocyte and astrocytes. Enriched GO terms for each cerebellar target gene clusters were cell-type and temporally specific, highlighting enhancer specificity. Functionally annotating their respective clusters provides a working hypothesis for hundreds of genes, which can be used as a jumping point for future in-depth studies in the cerebellum. Collectively, these findings support the notion that the cell types in the cerebellum have specific enhancer signatures relevant which are reflected by the expression and functions of their target genes.

GWAS SNPs and DNMs associated with ASD are enriched in cerebellar enhancers

Having established and characterized enhancer sequences in the cerebellum, we sought to elucidate the potential involvement of these regions in the etiology of neurological disorders; imaging and quantitative data show consistent cerebellar abnormalities, particularly in cases of

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

554 individuals with autism (Limperopoulos et al., 2014) (Stoodley & Limperopoulos, 2016). Our results
555 indicate that cerebellar enhancer sequences are significantly enriched for GWAS variants and
556 DNMs associated with ASD, suggesting an important role for enhancers in contributing to the
557 condition. PAX6 was among 12 target genes of cerebellar enhancers containing ASD-associated
558 variants and is classified in the SFARI as a ASD risk gene. The deletion of Pax6 in the murine
559 cerebellum results in aberrant development of the glutamatergic cells in the cerebellum: the
560 cerebellar nuclei, unipolar brush cells and granule cells (Yeung et al., 2016). Behavioral analysis
561 of Pax6 animal models has also indicated a possible link between this gene and autistic-like
562 behavior (Umeda et al., 2010). Additionally, Pax6 has been linked with WAGR (Wilm's tumor,
563 Aniridia, Genitourinary malformations, and mental Retardation syndrome) which is co-morbid
564 for ASD. Our analysis invites future investigation these target genes and how perturbation of
565 their expression may lead to ASD phenotypes.

566 Of the target genes of the enhancers that overlapped *de novo* CNVs in the SSC cohort,
567 none have been previously associated with cerebellar development. Interestingly, one of these
568 target genes, *CDC42BPB,* has recently been associated with neurodevelopmental disorders
569 including ASD (Chilton et al., 2020). This gene is a serine/threonine protein kinase and codes for
570 MRCKβ (myotonic dystrophy-related Cdc42-binding kinase beta), a regulator of cell
571 cytoskeletal reorganization and cell migration (Pichaud, Walther, & Nunes de Almeida, 2019). Of
572 note, the CNV associated with this gene deletes the entire enhancer. CDC42BPB shows
573 expression in the granule cell layer of the lateral adult cerebellum, which has been associated
574 with cognitive functions (Koziol et al., 2014).

575 Together, our data serves as an invaluable resource for future studies, by providing
576 candidate genes involved in cerebellar development with potentially meaningful impact in the
577 etiology of ASD and other neurodevelopmental disabilities.

578

579 **Materials and Methods**

580 **Mouse strains and husbandry**

581 C57BL/6 J mice were originally purchased from JAX laboratory and maintained and bred in a
582 pathogen-free animal facility with 12/12 hour light/dark cycle and a controlled environment.
583 Embryonic ages utilized in these experiments were confirmed based upon the appearance of a

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

584 vaginal plug. The morning that a vaginal plug was detected was designated as E0.5. Pregnant
585 females were cervically dislocated and embryos were harvested from the uterus. Postnatal ages
586 were determined based upon the date of birth with the morning of the observation of newborn
587 pups considered as P0.5. All studies were conducted according to the protocols approved by the
588 Institutional Animal Care and Use Committee and the Canadian Council on Animal Care at the
589 University of British Columbia.
590
591 **Tissue preparation for chromatin immunoprecipitation**
592 C57BL/6 J mice (male and female) at E12.5, P0.5 and P9.5 (henceforth referred to as E12, P0,
593 and P9) were decapitated for dissection of cerebella. Cerebella were dissected and collected in
594 ice cold Dulbecco's PBS (DPBS) without magnesium or calcium and subsequently washed 2x
595 for 5 minutes. Samples from each litter were pooled and trypsinized in DPBS containing 0.25%
596 trypsin for 10, 15 and 30 min at room temperature for E12, P0 and P9, respectively. Following 3
597 washes with fresh DPBS, the tissue was triturated with 3 progressively smaller (1, 0.5, 0.1mm)
598 bore polished and sterile pipettes in DPBS containing 250U/ml DNase, 0.25% glucose, and
599 8mg/ml BSA. The triturated cells were diluted 1:4 with cold DPBS and passed through a cell
600 strainer (40µm mesh) to remove large cellular debris. The cells were collected by mild
601 centrifugation, washed in fresh DPBS and counted. The cells were split into 100,000 cell
602 aliquots, pelleted and snap frozen using liquid nitrogen. Cell pellets were stored at -80°C.
603
604 **Histone chromatin immunoprecipitation**
605 We performed native chromatin immunoprecipitation (ChIP) using validated antibodies against
606 H3K4me1 and H3K27ac according to previously established protocols by the International
607 Human Epigenomics Consortium (IHEC) (Lorzadeh, Lopez Gutierrez, Jackson, Moksa, & Hirst,
608 2017). Briefly, cells were lysed in mild non-ionic detergents (0.1% Triton X-100 and
609 Deoxycholate) and protease inhibitor cocktail (Calbiochem) to preserve the integrity of histones
610 harbouring epitopes of interest during cell lysis. Cells were digested by Microccocal nuclease
611 (MNase) at room temperature for 5 minutes and 0.25mM EDTA was used to stop the reaction.
612 Antibodies to H3K4me1 (Diagenode: Catalogue #C15410037, lot A1657D) and H3K27ac
613 (Hiroshi Kimura, Cell Biology Unit, Tokyo Institute of Technology) were incubated with anti-
614 IgA magnetic beads (Dynabeads from Invitrogen) for 2 hours. Digested chromatin was incubated

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

with magnetic beads alone for 1.5 hours. Digested chromatin was separated from the beads and incubated with antibody-bead complex overnight in immunoprecipitation buffer (20mM Tris-HCl pH 7.5, 2mM EDTA, 150mM NaCl, 0.1% Triton X-100, 0.1% Deoxycholate). The resulting immunoprecipitations were washed 2 times by low salt (20mM Tris-HCl pH 8.0, 2mM EDTA, 150mM NaCl, 1% Triton X-100, 0.1% SDS) and then high salt (20mM Tris-HCl pH 8.0, 2mM EDTA, 500 mM NaCl, 1% Triton X-100, 0.1% SDS) wash buffers. Immunoprecipitations were eluted in an elution buffer (1% SDS, 100 mM Sodium Bicarbonate) for 1.5 hours at 65°C. Remaining histones were digested by Protease (Invitrogen) for 30 minutes at 50°C and DNA fragments were purified using Ampure XP beads (Beckman Coulter). The library preparation was conducted by Diagenode ChIP-seq/ChIP-qPCR Profiling service (Diagenode Cat# G02010000) using the MicroPlex Library Preparation Kit v2 (Diagenode Cat. C05010013). 50-bp single-end sequencing was performed on all libraries by Diagenode (Belgium) on an Illumina HiSeq 3000 platform. Two independent biological replicates were performed for each antibody and developmental time point.

Histone modification ChIP-seq data processing

The sequencing data were uploaded to the Galaxy web platform (usegalaxy.org) for analyses (Afgan et al., 2016). 50-bp single-end ChIP-seq reads were aligned to the NCBI37/mm9 reference genome and converted to binary alignment/map (BAM) format by Bowtie2 v.2.3.4 (Langmead, Trapnell, Pop, & Salzberg, 2009) with default parameters. Duplicate reads were marked using Picard v.1.52. Peak enrichment was computed using MACS v.2.1.1 (Zhang, Y. et al., 2008) with a false discovery rate (FDR) cutoff of 0.01 (p-value < 1E-5) using input samples as a control for each replicate. bigWigs were generated and normalized by the total number of mapped reads using the BamCompare and profiles were generated from these bigWigs by calculating average coverage in 50 bp bins using Deeptools v.3.3 (Ramírez et al., 2016) for downstream analysis and visualization.

Identification of active cerebellar enhancers

We first determined consensus peaks between replicates for both H3K27ac and H3K4me1 peaks collected at E12, P0 and P9 using the *intersect* function from Bedtools v.2.28 (Quinlan & Hall, 2010). Robust active cerebellar enhancers were identified by overlapping replicated H3K27ac

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

and H3K4me1 peaks called for E12, P0 and P9 samples. The genomic coordinates of the H3K27ac peaks that overlapped with H3K4me1 enriched regions at the same age were used for our list of robust active cerebellar enhancers. We then removed peaks found within promoter sequences by eliminating any peaks found 500bp upstream or downstream of transcription start sites (TSSs) in the developing cerebellum as determined previously (Zhang et al., 2018). The resulting list of robust active cerebellar enhancer sequences at E12, P0, and P9 were used for downstream analysis.

For the comparative analysis with cerebellar postnatal enhancers previously published by Frank et al. (2015), H3K27ac and DNase-seq peak coordinates were downloaded from Gene Expression Omnibus (GEO) (GSE60731). The following sequences were downloaded from public enhancer databases: 1) enhancers downloaded from the VISTA Enhancer Browser (https://enhancer.lbl.gov/) (Visel et al., 2007) with hindbrain activity were filtered using the 'Advanced Search' tool, selecting "hindbrain" under Expression Pattern and retrieving only mouse sequences with positive signal and 2) mouse cerebellar neonate enhancer coordinates were downloaded from the Enhancer Atlas 2.0 repository (http://www.enhanceratlas.org/downloadv2.php) (Gao & Qian, 2020). For the comparisons, sequences were overlapped with our robust cerebellar enhancer peaks and H3K27ac peaks at E12, P0 and P9 using Bedtools v.2.28(Quinlan & Hall, 2010).

Differential binding analysis

Aligned read counts (BAM file format) from our H3K27ac ChIP-seq experiments mapped to our robust active cerebellar enhancers for E12, P0 and P9 samples were used as input to the package DiffBind v1.4.2 (Stark & Brown, 2011). Read counting at each genomic location was conducted, which was subsequently normalized by experimental input samples. The result of counting is a binding affinity matrix containing normalized read counts for every sample at each robust active cerebellar enhancer. For differential binding affinity analysis, three contrasts were set up in DiffBind: E12 vs P0, E12 vs P9, and P0 vs P9. Differential binding was determined by DiffBind using a negative binomial test at an FDR < 0.05 threshold. The FDRs and normalized signal difference for each contrast were plotted using the EnhancedVolcano package in R (Kevin Blighe, Sharmila Rana, & Myles Lewis, 2020).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Temporal classification of cerebellar enhancers

To determine cerebellar enhancers with embryonic-specific activity, H3K27ac signal at E12 was compared to P0 and P9. Enhancers with significantly higher signal at E12 for either contrast were considered "Early" active enhancers. A region found to be enriched for both contrasts was counted as one enhancer. To determine cerebellar enhancers with postnatal-specific activity, H3K27ac signal at P9 was compared to E12 and P0. Enhancers with significantly higher signal at P9 for either contrast were considered "Late" active enhancers. A region found to be enriched for both contrasts was counted as one enhancer. To determine cerebellar enhancers with activity specific to birth, H3K27ac signal at P0 was compared to P9 and E12. Enhancers with significantly higher signal at P0 in both contrasts would identify enhancers that peaked in activity at P0. We did not identify any enhancers that peaked in activity at P0 and conducted the remaining analysis for only Early and Late enhancers.

Transcription factor motif enrichment analysis

Transcription factor motif enrichment was calculated using the software HOMER using the script FindMotifsGenome.pl with default parameters (Heinz et al., 2010). Analyses for Early and Late active enhancers were conducted separately. Motif enrichment was statistically analyzed using a cumulative binomial distribution. Enriched motifs were aligned with known transcription factor binding sites to determine the best matches. Top known motif matches were filtered based on expression within the developing cerebellum at E12 for "Early" active enhancers and P9 for "Late" active enhancers.

Cerebellar enhancer target gene prediction and co-expression analysis

To identify possible gene targets of our robust cerebellar enhancers, the correlation between H3K27ac signal and mRNA expression of genes located in *cis* at E12, P0 and P9 was calculated. For a given enhancer, a gene located in *cis* was considered a possible target if it was positively correlated with H3K27ac signal throughout time. These genes were then filtered based on location using conserved topologically associating domains (TADs), which are areas of the genome that preferentially interact (Dixon et al., 2012). These TADs are conserved between different cell types and even across species and were established using Hi-C data generated, previously. Gene target candidates for a given enhancer were curated for those located within the

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

same TAD. Predicted gene targets were then ranked based on their Pearson correlation coefficient value. For the predicted gene targets of Early and Late active enhancers, we conducted *k-means* clustering of predicted gene targets separately. Input for this analysis was gene expression captured from cerebella at 12 embryonic and postnatal time points (Zhang et al., 2018). Briefly, gene expression was quantified using Cap Analysis of Gene Expression followed by sequencing (CAGE-seq) for mouse cerebellar samples dissected every 24 hours from E11-P0 and then every 72 hours until P9 (12 in vivo time points in total). The number of clusters for the *k-means* clustering was determined using the Elbow analysis for each classified group of enhancers: Early (n=4) and Late (n=4).

Tissue preparation for histology

Embryos harvested between E11.5 to E15.5 were fixed by immersion in 4% paraformaldehyde in 0.1M phosphate buffer (PB, pH 7.4) for 1 hour at 4°C. Postnatal mice between P0.5 to P6.5 were perfused through the heart with a saline solution followed by 4% paraformaldehyde/0.1M PBS. The brain tissues were isolated and further fixed in 4% paraformaldehyde in 0.1M PB for 1 hour at room temperature. Fixed tissues were rinsed with PBS, followed by cryoprotection with 30% sucrose/PBS overnight at 4°C before embedding in the Optimal Cutting Temperature compound (Tissue-Tek). Tissues were sectioned at 12um for immunofluorescence experiments and cryosections were mounted on Superfrost slides (Thermo Fisher Scientific), air dried at room temperature, and stored at −80°C until used. Sagittal sections were cut from one side of the cerebellum to the other (left to right, or vice versa). In all cases, observations were based on a minimum of 3 embryos per genotype per experiment.

Cerebellar immunostaining

Tissue sections were first rehydrated in PBS (3 x 5 minute washes) followed by a phosphate buffered saline with Triton X-100 (PBS-T) rinse. Sections were then incubated at room temperature for 1 hour with blocking solution (0.3% BSA, 10% normal goat serum, 0.02% sodium azide in PBS-T). Following the blocking step, the slides were incubated with primary antibody in incubation buffer (0.3% BSA, 5% normal goat serum, 0.02% sodium azide in PBS-T) at room temperature overnight in a humid chamber. Following the overnight incubation, the slides were rinsed in 3 x 10 minute PBS-T washes. The sections were then incubated with the

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

appropriate secondary antibody at room temperature for 1 hour, followed by three 0.1M PB washes and one 0.01M PB wash. Coverslips were applied to the slides using FluorSave mounting medium (345789, Calbiochem). The primary antibodies used were: rabbit anti-Bhlhe22 (1:1000, a gift from Dr. Michael Greenburg, Harvard University), mouse anti-Neurod1 (1:500, Abcam, ab60704), mouse anti-Pax3 (1:500, R&D systems, MAB2457), rabbit anti-Pax2 (1:200, Invitrogen, 71-6000), mouse anti-NeuN (1:100, Millipore, MAB377), rabbit anti-Calbindin (1:1000, Millipore, AB1778), rabbit anti-Foxp2 (1:2000, Novus Biologicals NB100-55411), chicken anti-Doublecortin: (1:100, Abcam ab153668). For immunofluorescence, secondary antibodies (Invitrogen) labeled with fluorochrome were used to recognize the primary antibodies.

Granule cell culture

Granule cells were isolated and cultured as previously described (Lee et al., 2009). Briefly, cerebella from litters of P6 mice were pooled and digested at 37 °C for 20 minutes in 10U ml−1 papain (Worthington), and 250U ml−1 DNase in EBSS using the Papain Dissociation Kit (Worthington, Cat #:LK003150). The tissue was mechanically triturated and suspended cells were isolated and resuspended in EBSS with albumin-ovomucoid inhibitor solution. Cell debris was removed using a discontinuous density gradient and cells were resuspended in HBSS, glucose and DNase. The cell suspension was then passed through a 40um cell strainer (Falcon 2340), layered on a step gradient of 35% and 65% Percoll (Sigma), and centrifuged at 2,500rpm for 12 minutes at 25°C. Granule cells were harvested from the 35/65% interface and washed in HBSS-glucose. Granule cells were then resuspended in Neurobasal medium and 10% FBS and pre-plated on lightly coated poly-D-lysine–coated dishes for 20 minutes. This step allows any heavier cells to drop and adhere to the coated surface while the granule cells are retained in the media. Granule cells in the media were then collected, washed and counted using a Hemocytomoter. The cells were then plated on 25mm or 12mm poly-D-lysine (Sigma), laminin coated coverglasses placed in 6-well plates with Neurobasal medium containing B-27 serum-free supplement, 2mm l-glutamine, 100U/ml penicillin, and 100µg/ml streptomycin (pen-strep) (Invitrogen, Grand Island, NY) and 0.45% d-glucose (Gibco). Granule cells were incubated at 37°C at 5% CO2 were incubated for 3 days in vitro (DIV).

For aggregate cultures, aggregates were allowed to form by incubating purified granule cells for 20 hours on uncoated tissue culture dishes in DMEM containing 10% FBS, 0.45% D-glucose,

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Pen-strep, 2mM L-glutamine at 4E6 cells/ml. Aggregates were then washed and cultured in Neurobasal/B27 medium on poly-d-lysine/laminin-treated chamber slides at 37°C/5% CO2. Neuronal processes extend from aggregates and most form neurite bundles. After several hours, small bipolar granule cells migrate unidirectionally away from the cell clusters along these neurites and neurite bundles by extension of processes, followed by translocation of cell bodies outside of the aggregate cell cluster margin. For immunofluorescence experiments, cells were fixed in 4% paraformaldehyde for 10 minutes and washed with calcium and magnesium-free PBS.

RNA interference

For the knockdown of Bhlhe22, we purchased ON-TARGETplus SMARTPool Mouse Bhlhe22 siRNA from Horizon Discovery (Cat ID: L-063262-01). Control samples were transfected with ON-TARGETplus Non-targeting Control Pool (Cat ID: D-001810-10). siRNA molecules were electroporated into isolated postnatal cerebellar granule cells using the Nucleofector 2b Device (Lonza, AAB-1001) as previously described (Gartner et al., 2006). Briefly, after cells were isolated (described above), 6-7E6 cells were resuspended in nucleofection solution and mixed with 3ug of pCAG-EGFP plasmid (Addgene, 89684) and 600nM of siRNA. Cuvettes were loaded with cellular solution and nucleofected using program O-03. After electroporation, cells were allowed to recover in DMEM media in a humidified 37°C/5% CO2 incubator for 90 minutes. Cells were washed and resuspended in either culture media for plating (dissociated cultures) or DMEM media for overnight incubation (aggregate cultures).

RNA isolation and reverse transcription followed by quantitative PCR (RT-qPCR)

RNA was collected from cultured granule cells using the Monarch® Total RNA Miniprep Kit (NEB). Then cDNA was reverse transcribed using SuperScript™ IV First-Strand Synthesis System (Invitrogen) using random hexamers. Quantitative PCR was conducted using the Applied Biosystems Fast SYBR Green Master Mix reagent and Applied Biosystems 7500 Real-time PCR system. PCR conditions were as follows: 95 °C for 20 seconds, 40 cycles of 95 °C for 3 seconds, and 60 °C for 30 seconds followed by 95 °C for 15 seconds, 60 °C for 1 minute, 95 °C for 15 seconds and 60 °C for 15 seconds. Three biological replicates were analyzed for each target gene. Amplification of eGFP was used as a reference gene to normalize the relative amounts of

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

successfully transfected cells between treated and control experiments. Gene specific primers are listed in Supplementary Table 4. Expression profiles for each gene were calculated using the average relative quantity of the sample using the deltaCT method (Livak & Schmittgen, 2001). For comparisons between siRNA treated and control samples, means were compared using a two-tailed t-test. Results were expressed as the average ± SE, and p-values <0.05 were considered significant.

Image analysis and microscopy

Analysis and photomicroscopy were performed with a Zeiss Axiovert 200M microscope with the Axiocam/Axiovision hardware-software components (Carl Zeiss) and downstream image analysis was conducted using the AxioVision software v.4.9.1 (Carl Zeiss). For cerebellar granule cell aggregate cultures, aggregate size determined using the tracing tool and all aggregates analyzed were within 1000 squared microns of each other. Transfected cells were identified by examining eGFP expression and for each biological replicate/experimental treatment, 20 aggregates were examined. Granule cell migration was measured by calculating the distance of migrated cells from the edge of the aggregate on captured images. Mean migration distance was calculated for each aggregate, and the average of all 20 aggregates was used for statistical analysis. The distribution of migrated cells from the aggregate was calculated for the following ranges: <25um, 25-50um, 50-75um, 75-100um, >100um. For each range, the average percentage was calculated for 20 aggregates per replicate. For comparisons between siRNA treated and control samples, means were compared using a two-tailed t-test. Results were expressed as the average ± SE, and p-values <0.05 were considered significant.

Plots and statistical methods

All plots and correlation analysis were generated in R version 3.2.3 and figures were produced using the package ggplot2. Bedtools v.2.28 (Quinlan & Hall, 2010) was used for comparing and overlapping the genomic coordinates of peaks and existing genomic features described in the manuscript. Boxplots represent the median (centre line), first and third quartiles (top and bottom of box, respectively) and confidence intervals (95%; black lines). Genome browser screenshots were taken from the IGV genome browser (Robinson et al., 2011). Bar plots results were expressed as the average and the corresponding error bars represent standard error.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

GWAS SNP enrichment analysis

Single nucleotide polymorphisms (SNPs) were retrieved from the GWAS Catalog (Buniello et al., 2019) downloaded on March 8th, 2020. The SNPs were then filtered by their associated traits. Traits containing the word "autism" were selected and from this list any traits containing the word "or" were excluded. This resulted in a final list of 8 traits (Supplementary Table 2) and the associated SNPs were used as input for our analysis. The software Genomic Regulatory Elements and Gwas Overlap algoRithm (GREGOR) v.1.4.0 (Schmidt et al., 2015), a tool to test for enrichment of an input list of trait-associated index SNPs in experimentally annotated regulatory domains, was used to identify enrichment of trait-specific disease variants within enhancers. An underlying hypothesis of GREGOR is that both trait-associated SNPs and variants in strong linkage disequilibrium (LD) may be deemed as causal. For this, we used the European population reference file (EUR; LD window size = 1 Mb; LD $r^2 \geq 0.7$) from 1000G data (Release date: May 21, 2011). The probability of an overlap of either a SNP or at least one of its LD proxies with our enhancers relative to a set of matched control variants was used to evaluate significance of overlap. The enrichment p-value is the probability that the overlap of control variants with our enhancers is greater than or equal to the overlap of the GWAS variants with our enhancers.

De novo mutation analysis

De novo mutations were detected using whole-genome sequencing data from the MSSNG (Yuen et al., 2016) and Simons Simplex Collection (SSC) (Isoda et al., 2017) cohorts using a pipeline involving DeNovoGear (Ramu et al., 2013) as previously described (C Yuen et al., 2017). To maximize data homogeneity, we included only individuals sequenced on the Illumina HiSeq X platform. Individuals having a total DNM count more than three standard deviations above the mean of the cohort were excluded. The NCBI LiftOver tool was used to convert the coordinates of cerebellar enhancers from mm9 to hg19 to hg38, and BEDTools (Quinlan & Hall, 2010) was used to identify DNMs overlapping these coordinates. Contingency tables (2x2) were generated containing counts of the number of DNMs in ASD-affected individuals and unaffected siblings either overlapping or not overlapping each dataset (cerebellar enhancer or H3K27ac peak coordinates). Fisher's exact test was used to determine statistical significance. Copy number

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

variants (CNVs) >= 1000 bp were detected from the MSSNG and SSC WGS data using a pipeline involving the algorithms ERDS (Zhu et al., 2012) and CNVnator (Abyzov, Urban, Snyder, & Gerstein, 2011) as previously described (Trost et al., 2018). A CNV was deemed to be *de novo* if it was detected by both ERDS and CNVnator in the child but by neither algorithm in both parents. We then used BEDtools (Quinlan & Hall, 2010) to identify *de novo* CNVs overlapping our cerebellar enhancers.

Competing Interest Statement

Authors have no competing interests.

Author Contributions

M.R. conducted experiments and was responsible for all major areas of concept formation, data collection, analysis and manuscript composition. Y.B. processed and analyzed ChIP-seq data and conducted the human variant enrichment analysis as well as contributed to manuscript writing. J.Y. was involved in all mouse breeding and sample collection. J.W. and E.Y. were involved in the initial profiling of Pax3 and conducting immunofluorescent experiments. B.T. and S.W.S. conducted all genome-wide sequencing and analysis for the enrichment of autism spectrum disorder variants. D.G. was the supervisory author and was involved in all areas of concept formation and manuscript edits. All authors contributed to the final drafting of the manuscript.

References

Abyzov, A., Urban, A. E., Snyder, M., & Gerstein, M. (2011). CNVnator: An approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. *Genome Research, 21*(6), 974-984. doi:10.1101/gr.114876.110

Afgan, E., Baker, D., van den Beek, M., Blankenberg, D., Bouvier, D., Čech, M., . . . Goecks, J. (2016). The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2016 update. *Nucleic Acids Research, 44*(W1), W3-W10. doi:10.1093/nar/gkw343

Aruga, J., & Millen, K. J. (2018). ZIC1 function in normal cerebellar development and human developmental pathology. *Advances in Experimental Medicine and Biology, 1046*, 249-268. doi:10.1007/978-981-10-7311-3_13

Banerjee-Basu, S., & Packer, A. (2010). SFARI gene: An evolving database for the autism research community. *Disease Models & Mechanisms, 3*(3-4), 133-135. doi:10.1242/dmm.005439

Barešić, A., Nash, A. J., Dahoun, T., Howes, O., & Lenhard, B. (2020). Understanding the genetics of neuropsychiatric disorders: The potential role of genomic regulatory blocks. *Molecular Psychiatry, 25*(1), 6-18. doi:10.1038/s41380-019-0518-x

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

898 Ben-Arie, N., Bellen, H. J., Armstrong, D. L., McCall, A. E., Gordadze, P. R., Guo, Q., . . .
899 Zoghbi, H. Y. (1997). Math1 is essential for genesis of cerebellar granule neurons. *Nature,*
900 *390*(6656), 169-172. doi:10.1038/36579

901 Blankvoort, S., Witter, M. P., Noonan, J., Cotney, J., & Kentros, C. (2018). Marked diversity of
902 unique cortical enhancers enables neuron-specific tools by enhancer-driven gene expression.
903 *Current Biology: CB, 28*(13), 2103-2114.e5. doi:10.1016/j.cub.2018.05.015

904 Bonev, B., Mendelson Cohen, N., Szabo, Q., Fritsch, L., Papadopoulos, G. L., Lubling, Y., . . .
905 Cavalli, G. (2017). Multiscale 3D genome rewiring during mouse neural development. *Cell,*
906 *171*(3), 557-572.e24. doi:10.1016/j.cell.2017.09.043

907 Buniello, A., MacArthur, J. A. L., Cerezo, M., Harris, L. W., Hayhurst, J., Malangone, C., . . .
908 Parkinson, H. (2019). The NHGRI-EBI GWAS catalog of published genome-wide
909 association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Research,*
910 *47*(D1), D1005-D1012. doi:10.1093/nar/gky1120 [doi]

911 C Yuen, R. K., Merico, D., Bookman, M., L Howe, J., Thiruvahindrapuram, B., Patel, R. V., . . .
912 Scherer, S. W. (2017). Whole genome sequencing resource identifies 18 new candidate
913 genes for autism spectrum disorder. *Nature Neuroscience, 20*(4), 602-611.
914 doi:10.1038/nn.4524

915 Calo, E., & Wysocka, J. (2013). Modification of enhancer chromatin: What, how, and why?
916 *Molecular Cell, 49*(5), 825-837. doi:10.1016/j.molcel.2013.01.038

917 Carter, R. A., Bihannic, L., Rosencrance, C., Hadley, J. L., Tong, Y., Phoenix, T. N., . . . Gawad,
918 C. (2018). A single-cell transcriptional atlas of the developing murine cerebellum. *Current*
919 *Biology: CB, 28*(18), 2910-2920.e2. doi:10.1016/j.cub.2018.07.062

920 Carullo, N. V. N., & Day, J. J. (2019). Genomic enhancers in brain health and disease. *Genes,*
921 *10*(1) doi:10.3390/genes10010043

922 Chilton, I., Okur, V., Vitiello, G., Selicorni, A., Mariani, M., Goldenberg, A., . . . Chung, W. K.
923 (2020). De novo heterozygous missense and loss-of-function variants in CDC42BPB are
924 associated with a neurodevelopmental phenotype. *American Journal of Medical Genetics.*
925 *Part A, 182*(5), 962-973. doi:10.1002/ajmg.a.61505

926 Consalez, G. G., Goldowitz, D., Casoni, F., & Hawkes, R. (2021). Origins, development, and
927 compartmentation of the granule cells of the cerebellum. *Frontiers in Neural Circuits, 14*
928 doi:10.3389/fncir.2020.611841

929 Creyghton, M. P., Cheng, A. W., Welstead, G. G., Kooistra, T., Carey, B. W., Steine, E. J., . . .
930 Jaenisch, R. (2010). Histone H3K27ac separates active from poised enhancers and predicts
931 developmental state. *Proceedings of the National Academy of Sciences of the United States*
932 *of America, 107*(50), 21931-21936. doi:10.1073/pnas.1016071107

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

de la Torre-Ubieta, L., Stein, J. L., Won, H., Opland, C. K., Liang, D., Lu, D., & Geschwind, D. H. (2018). The dynamic landscape of open chromatin during human cortical neurogenesis. *Cell, 172*(1-2), 289-304.e18. doi:10.1016/j.cell.2017.12.014

Dixon, J. R., Selvaraj, S., Yue, F., Kim, A., Li, Y., Shen, Y., . . . Ren, B. (2012). Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature, 485*(7398), 376-380. doi:10.1038/nature11082

Ebner, B. A., Ingram, M. A., Barnes, J. A., Duvick, L. A., Frisch, J. L., Clark, H. B., . . . Orr, H. T. (2013). Purkinje cell ataxin-1 modulates climbing fiber synaptic input in developing and adult mouse cerebellum. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 33*(13), 5806-5820. doi:10.1523/JNEUROSCI.6311-11.2013

Elsen, G. E., Bedogni, F., Hodge, R. D., Bammler, T. K., MacDonald, J. W., Lindtner, S., . . . Hevner, R. F. (2018). The epigenetic factor landscape of developing neocortex is regulated by transcription factors Pax6→ Tbr2→ Tbr1. *Frontiers in Neuroscience, 12*, 571. doi:10.3389/fnins.2018.00571

Fischbach, G. D., & Lord, C. (2010). The simons simplex collection: A resource for identification of autism genetic risk factors. *Neuron, 68*(2), 192-195. doi:10.1016/j.neuron.2010.10.006

Frank, C. L., Liu, F., Wijayatunge, R., Song, L., Biegler, M. T., Yang, M. G., . . . West, A. E. (2015). Regulation of chromatin accessibility and zic binding at enhancers in the developing cerebellum. *Nature Neuroscience, 18*(5), 647-656. doi:10.1038/nn.3995

Frantz, G. D., Bohner, A. P., Akers, R. M., & McConnell, S. K. (1994). Regulation of the POU domain gene SCIP during cerebral cortical development. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 14*(2), 472-485.

Fujita, E., Tanabe, Y., Shiota, A., Ueda, M., Suwa, K., Momoi, M. Y., & Momoi, T. (2008). Ultrasonic vocalization impairment of Foxp2 (R552H) knockin mice related to speech-language disorder and abnormality of purkinje cells. *Proceedings of the National Academy of Sciences of the United States of America, 105*(8), 3117-3122. doi:10.1073/pnas.0712298105

Gao, T., & Qian, J. (2020). EnhancerAtlas 2.0: An updated resource with enhancer annotation in 586 tissue/cell types across nine species. *Nucleic Acids Research, 48*(D1), D58-D64. doi:10.1093/nar/gkz980 [doi]

Gartner, A., Collin, L., & Lalli, G. (2006). Nucleofection of primary neurons. *Methods in Enzymology, 406*, 374-388. doi:S0076-6879(06)06027-7 [pii]

Goldowitz, D., & Hamre, K. (1998). The cells and molecules that make a cerebellum. *Trends in Neurosciences, 21*(9), 375-382.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

968 Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., . . . Børglum, A. D.
969 (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature*
970 *Genetics, 51*(3), 431-444. doi:10.1038/s41588-019-0344-8

971 Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., . . . Glass, C. K. (2010).
972 Simple combinations of lineage-determining transcription factors prime cis-regulatory
973 elements required for macrophage and B cell identities. *Molecular Cell, 38*(4), 576-589.
974 doi:10.1016/j.molcel.2010.05.004 [doi]

975 Heinz, S., Romanoski, C. E., Benner, C., & Glass, C. K. (2015). The selection and function of
976 cell type-specific enhancers. *Nature Reviews. Molecular Cell Biology, 16*(3), 144-154.
977 doi:10.1038/nrm3949

978 Hoshino, M., Nakamura, S., Mori, K., Kawauchi, T., Terao, M., Nishimura, Y. V., . . .
979 Nabeshima, Y. (2005). Ptf1a, a bHLH transcriptional gene, defines GABAergic neuronal
980 fates in cerebellum. *Neuron, 47*(2), 201-213. doi:10.1016/j.neuron.2005.06.007

981 Isoda, T., Moore, A. J., He, Z., Chandra, V., Aida, M., Denholtz, M., . . . Murre, C. (2017). Non-
982 coding transcription instructs chromatin folding and compartmentalization to dictate
983 enhancer-promoter communication and T cell fate. *Cell, 171*(1), 103-119.e18.
984 doi:10.1016/j.cell.2017.09.001

985 Joshi, P. S., Molyneaux, B. J., Feng, L., Xie, X., Macklis, J. D., & Gan, L. (2008). Bhlhb5
986 regulates the postmitotic acquisition of area identities in layers II-V of the developing
987 neocortex. *Neuron, 60*(2), 258-272. doi:10.1016/j.neuron.2008.08.006

988 Kevin Blighe, Sharmila Rana, & Myles Lewis. (2020). EnhancedVolcano: Publication-ready
989 volcano plots with enhanced colouring and labeling [computer software]

990 Kim, E. J., Battiste, J., Nakagawa, Y., & Johnson, J. E. (2008). Ascl1 (Mash1) lineage cells
991 contribute to discrete cell populations in CNS architecture. *Molecular and Cellular*
992 *Neurosciences, 38*(4), 595-606. doi:10.1016/j.mcn.2008.05.008 [doi]

993 Klisch, T. J., Xi, Y., Flora, A., Wang, L., Li, W., & Zoghbi, H. Y. (2011). In vivo Atoh1
994 targetome reveals how a proneural transcription factor regulates cerebellar development.
995 *Proceedings of the National Academy of Sciences, 108*(8), 3288-3293.

996 Koziol, L. F., Budding, D., Andreasen, N., D'Arrigo, S., Bulgheroni, S., Imamizu, H., . . .
997 Yamazaki, T. (2014). Consensus paper: The cerebellum's role in movement and cognition.
998 *Cerebellum (London, England), 13*(1), 151-177. doi:10.1007/s12311-013-0511-x

999 Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient
1000 alignment of short DNA sequences to the human genome. *Genome Biology, 10*(3), R25.
1001 doi:10.1186/gb-2009-10-3-r25

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1002 Lee, H. Y., Greene, L. A., Mason, C. A., & Manzini, M. C. (2009). Isolation and culture of post-
1003 natal mouse cerebellar granule neuron progenitor cells and neurons. *Journal of Visualized*
1004 *Experiments : JoVE, (23). pii: 990. doi*(23), 10.3791/990. doi:10.3791/990 [doi]

1005 Leto, K., Bartolini, A., Yanagawa, Y., Obata, K., Magrassi, L., Schilling, K., & Rossi, F. (2009).
1006 Laminar fate and phenotype specification of cerebellar GABAergic interneurons. *The*
1007 *Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 29*(21),
1008 7079-7091. doi:10.1523/JNEUROSCI.0957-09.2009

1009 Leto, K., Carletti, B., Williams, I. M., Magrassi, L., & Rossi, F. (2006). Different types of
1010 cerebellar GABAergic interneurons originate from a common pool of multipotent progenitor
1011 cells. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience,*
1012 *26*(45), 11682-11694. doi:10.1523/JNEUROSCI.3656-06.2006

1013 Limperopoulos, C., Chilingaryan, G., Sullivan, N., Guizard, N., Robertson, R. L., & du Plessis,
1014 A. J. (2014). Injury to the premature cerebellum: Outcome is related to remote cortical
1015 development. *Cerebral Cortex (New York, N.Y.: 1991), 24*(3), 728-736.
1016 doi:10.1093/cercor/bhs354

1017 Lindtner, S., Catta-Preta, R., Tian, H., Su-Feher, L., Price, J. D., Dickel, D. E., . . . Rubenstein, J.
1018 L. R. (2019). Genomic resolution of DLX-orchestrated transcriptional circuits driving
1019 development of forebrain GABAergic neurons. *Cell Reports, 28*(8), 2048-2063.e8.
1020 doi:10.1016/j.celrep.2019.07.022

1021 Lister, R., Mukamel, E. A., Nery, J. R., Urich, M., Puddifoot, C. A., Johnson, N. D., . . . Ecker, J.
1022 R. (2013). Global epigenomic reconfiguration during mammalian brain development.
1023 *Science (New York, N.Y.), 341*(6146), 1237905. doi:10.1126/science.1237905

1024 Livak, K. J., & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-
1025 time quantitative PCR and the 2(-delta delta C(T)) method. *Methods (San Diego, Calif.),*
1026 *25*(4), 402-408. doi:10.1006/meth.2001.1262 [doi]

1027 Lorzadeh, A., Lopez Gutierrez, R., Jackson, L., Moksa, M., & Hirst, M. (2017). Generation of
1028 native chromatin immunoprecipitation sequencing libraries for nucleosome density analysis.
1029 *Journal of Visualized Experiments : JoVE, (130). doi*(130), 10.3791/56085.
1030 doi:10.3791/56085 [doi]

1031 Maricich, S. M., & Herrup, K. (1999). Pax-2 expression defines a subset of GABAergic
1032 interneurons and their precursors in the developing murine cerebellum. *Journal of*
1033 *Neurobiology, 41*(2), 281-294. doi:10.1002/(sici)1097-4695(19991105)41:23.0.co;2-5

1034 Miyata, T., Maeda, T., & Lee, J. E. (1999). NeuroD is required for differentiation of the granule
1035 cells in the cerebellum and hippocampus. *Genes & Development, 13*(13), 1647-1652.
1036 doi:10.1101/gad.13.13.1647

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1037 Nord, A. S., Blow, M. J., Attanasio, C., Akiyama, J. A., Holt, A., Hosseini, R., . . . Visel, A.
1038 (2013). Rapid and pervasive changes in genome-wide enhancer usage during mammalian
1039 development. *Cell, 155*(7), 1521-1531. doi:10.1016/j.cell.2013.11.033

1040 Nord, A. S., & West, A. E. (2020). Neurobiological functions of transcriptional enhancers.
1041 *Nature Neuroscience, 23*(1), 5-14. doi:10.1038/s41593-019-0538-5

1042 Nott, A., Holtman, I. R., Coufal, N. G., Schlachetzki, J. C. M., Yu, M., Hu, R., . . . Glass, C. K.
1043 (2019). Brain cell type-specific enhancer-promoter interactome maps and disease-risk
1044 association. *Science (New York, N.Y.), 366*(6469), 1134-1139. doi:10.1126/science.aay0793

1045 Osterwalder, M., Barozzi, I., Tissières, V., Fukuda-Yuzawa, Y., Mannion, B. J., Afzal, S. Y., . . .
1046 Pennacchio, L. A. (2018). Enhancer redundancy provides phenotypic robustness in
1047 mammalian development. *Nature, 554*(7691), 239-243. doi:10.1038/nature25461

1048 Pattabiraman, K., Golonzhka, O., Lindtner, S., Nord, A. S., Taher, L., Hoch, R., . . . Rubenstein,
1049 J. L. R. (2014). Transcriptional regulation of enhancers active in protodomains of the
1050 developing cerebral cortex. *Neuron, 82*(5), 989-1003. doi:10.1016/j.neuron.2014.04.014

1051 Peng, J., Sheng, A., Xiao, Q., Shen, L., Ju, X., Zhang, M., . . . Luo, Z. (2019). Single-cell
1052 transcriptomes reveal molecular specializations of neuronal cell types in the developing
1053 cerebellum. *Journal of Molecular Cell Biology,* doi:10.1093/jmcb/mjy089

1054 Pichaud, F., Walther, R. F., & Nunes de Almeida, F. (2019). Regulation of Cdc42 and its
1055 effectors in epithelial morphogenesis. *Journal of Cell Science, 132*(10)
1056 doi:10.1242/jcs.217869

1057 Porter, F. D., Drago, J., Xu, Y., Cheema, S. S., Wassif, C., Huang, S. P., . . . Westphal, H.
1058 (1997). Lhx2, a LIM homeobox gene, is required for eye, forebrain, and definitive
1059 erythrocyte development. *Development (Cambridge, England), 124*(15), 2935-2944.

1060 Preissl, S., Fang, R., Huang, H., Zhao, Y., Raviram, R., Gorkin, D. U., . . . Ren, B. (2018).
1061 Single-nucleus analysis of accessible chromatin in developing mouse forebrain reveals cell-
1062 type-specific transcriptional regulation. *Nature Neuroscience, 21*(3), 432-439.
1063 doi:10.1038/s41593-018-0079-3

1064 Quinlan, A. R., & Hall, I. M. (2010). BEDTools: A flexible suite of utilities for comparing
1065 genomic features. *Bioinformatics (Oxford, England), 26*(6), 841-842.
1066 doi:10.1093/bioinformatics/btq033

1067 Ramírez, F., Ryan, D. P., Grüning, B., Bhardwaj, V., Kilpert, F., Richter, A. S., . . . Manke, T.
1068 (2016). deepTools2: A next generation web server for deep-sequencing data analysis.
1069 *Nucleic Acids Research, 44*(W1), 160. doi:10.1093/nar/gkw257

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1070 Ramu, A., Noordam, M. J., Schwartz, R. S., Wuster, A., Hurles, M. E., Cartwright, R. A., &
1071 Conrad, D. F. (2013). DeNovoGear: De novo indel and point mutation discovery and
1072 phasing. *Nature Methods, 10*(10), 985-987. doi:10.1038/nmeth.2611

1073 Rinaldi, A., Defterali, C., Mialot, A., Garden, D. L. F., Beraneck, M., & Nolan, M. F. (2013).
1074 HCN1 channels in cerebellar purkinje cells promote late stages of learning and constrain
1075 synaptic inhibition. *The Journal of Physiology, 591*(22), 5691-5709.
1076 doi:10.1113/jphysiol.2013.259499

1077 Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., &
1078 Mesirov JP, -. e. (2011). Integrative genomics viewer. *Nature Biotechnology, 29*(1), 24-26.
1079 doi:10.1038/nbt.1754 [doi]

1080 Sanchez-Ortiz, E., Cho, W., Nazarenko, I., Mo, W., Chen, J., & Parada, L. F. (2014). NF1
1081 regulation of RAS/ERK signaling is required for appropriate granule neuron progenitor
1082 expansion and migration in cerebellar development. *Genes & Development, 28*(21), 2407-
1083 2420. doi:10.1101/gad.246603.114

1084 Sarropoulos, I., Sepp, M., Frömel, R., Leiss, K., Trost, N., Leushkin, E., . . . Kaessmann, H.
1085 (2021). The regulatory landscape of cells in the developing mouse cerebellum. *bioRxiv,* ,
1086 2021.01.29.428632. doi:10.1101/2021.01.29.428632

1087 Schmidt, E. M., Zhang, J., Zhou, W., Chen, J., Mohlke, K. L., Chen, Y. E., & Willer, C. J.
1088 (2015). GREGOR: Evaluating global enrichment of trait-associated variants in epigenomic
1089 features using a systematic, data-driven approach. *Bioinformatics (Oxford, England),*
1090 *31*(16), 2601-2606. doi:10.1093/bioinformatics/btv201 [doi]

1091 Stark, R., & Brown, G. (2011). DiffBind: Differential binding analysis of ChIP-seq peak
1092 data.[computer software]

1093 Stoodley, C. J., & Limperopoulos, C. (2016). Structure-function relationships in the developing
1094 cerebellum: Evidence from early-life cerebellar injury and neurodevelopmental disorders.
1095 *Seminars in Fetal & Neonatal Medicine, 21*(5), 356-364. doi:10.1016/j.siny.2016.04.010

1096 Su, X., Liu, X., Ni, L., Shi, W., Zhu, H., Shi, J., . . . Huang, Q. (2016). GFAP expression is
1097 regulated by Pax3 in brain glioma stem cells. *Oncology Reports, 36*(3), 1277-1284.
1098 doi:10.3892/or.2016.4917

1099 Swanson, D. J., & Goldowitz, D. (2011). Experimental sey mouse chimeras reveal the
1100 developmental deficiencies of Pax6-null granule cells in the postnatal cerebellum.
1101 *Developmental Biology, 351*(1), 1-12. doi:10.1016/j.ydbio.2010.11.018

1102 Takács, J., Zaninetti, R., Vig, J., Vastagh, C., & Hámori, J. (2008). Postnatal expression of
1103 doublecortin (dcx) in the developing cerebellar cortex of mouse. *Acta Biologica Hungarica,*
1104 *59*(2), 147-161. doi:10.1556/ABiol.59.2008.2.2

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1105 Thompson, C. L., Ng, L., Menon, V., Martinez, S., Lee, C., Glattfelder, K., . . . Jones, A. R.
1106 (2014). A high-resolution spatiotemporal atlas of gene expression of the developing mouse
1107 brain. *Neuron, 83*(2), 309-323. doi:10.1016/j.neuron.2014.05.033

1108 Trost, B., Walker, S., Wang, Z., Thiruvahindrapuram, B., MacDonald, J. R., Sung, W. W. L., . . .
1109 Scherer, S. W. (2018). A comprehensive workflow for read depth-based identification of
1110 copy-number variation from whole-genome sequence data. *American Journal of Human*
1111 *Genetics, 102*(1), 142-155. doi:10.1016/j.ajhg.2017.12.007

1112 Umeda, T., Takashima, N., Nakagawa, R., Maekawa, M., Ikegami, S., Yoshikawa, T., . . .
1113 Osumi, N. (2010). Evaluation of Pax6 mutant rat as a model for autism. *PloS One, 5*(12),
1114 e15500. doi:10.1371/journal.pone.0015500

1115 Urbánek, P., Fetka, I., Meisler, M. H., & Busslinger, M. (1997). Cooperation of Pax2 and Pax5
1116 in midbrain and cerebellum development. *Proceedings of the National Academy of Sciences*
1117 *of the United States of America, 94*(11), 5703-5708. doi:10.1073/pnas.94.11.5703

1118 Visel, A., Minovitsky, S., Dubchak, I., & Pennacchio, L. A. (2007). VISTA enhancer browser--a
1119 database of tissue-specific human enhancers. *Nucleic Acids Research, 35*(Database issue),
1120 88. doi:10.1093/nar/gkl822

1121 Visel, A., Rubin, E. M., & Pennacchio, L. A. (2009). Genomic views of distant-acting enhancers.
1122 *Nature, 461*(7261), 199-205. doi:10.1038/nature08451

1123 Visel, A., Taher, L., Girgis, H., May, D., Golonzhka, O., Hoch, R. V., . . . Rubenstein, J. L. R.
1124 (2013). A high-resolution enhancer atlas of the developing telencephalon. *Cell, 152*(4), 895-
1125 908. doi:10.1016/j.cell.2012.12.041

1126 Wang, V. Y., & Zoghbi, H. Y. (2001). Genetic regulation of cerebellar development. *Nature*
1127 *Reviews. Neuroscience, 2*(7), 484-491. doi:10.1038/35081558

1128 Wang, W., Mullikin-Kilpatrick, D., Crandall, J. E., Gronostajski, R. M., Litwack, E. D., &
1129 Kilpatrick, D. L. (2007). Nuclear factor I coordinates multiple phases of cerebellar granule
1130 cell development via regulation of cell adhesion molecules. *Journal of Neuroscience,*
1131 *27*(23), 6115-6127.

1132 Weyer, A., & Schilling, K. (2003). Developmental and cell type-specific expression of the
1133 neuronal marker NeuN in the murine cerebellum. *Journal of Neuroscience Research, 73*(3),
1134 400-409. doi:10.1002/jnr.10655

1135 Whyte, W. A., Bilodeau, S., Orlando, D. A., Hoke, H. A., Frampton, G. M., Foster, C. T., . . .
1136 Young, R. A. (2012). Enhancer decommissioning by LSD1 during embryonic stem cell
1137 differentiation. *Nature, 482*(7384), 221-225. doi:10.1038/nature10805

1138 Wizeman, J. W., Guo, Q., Wilion, E. M., & Li, J. Y. (2019). Specification of diverse cell types
1139 during early neurogenesis of the mouse cerebellum. *eLife, 8* doi:10.7554/eLife.42388

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.28.461416; this version posted September 29, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1140 Yao, P., Lin, P., Gokoolparsadh, A., Assareh, A., Thang, M. W., & Voineagu, I. (2015).
1141 Coexpression networks identify brain region-specific enhancer RNAs in the human brain.
1142 *Nature Neuroscience, 18*(8), 1168-1174. doi:10.1038/nn.4063 [doi]

1143 Yeung, J., Ha, T. J., Swanson, D. J., & Goldowitz, D. (2016a). A novel and multivalent role of
1144 Pax6 in cerebellar development. *The Journal of Neuroscience: The Official Journal of the*
1145 *Society for Neuroscience, 36*(35), 9057-9069. doi:10.1523/JNEUROSCI.4385-15.2016

1146 Yeung, J., Ha, T. J., Swanson, D. J., & Goldowitz, D. (2016b). A novel and multivalent role of
1147 Pax6 in cerebellar development. *The Journal of Neuroscience: The Official Journal of the*
1148 *Society for Neuroscience, 36*(35), 9057-9069. doi:10.1523/JNEUROSCI.4385-15.2016

1149 Yuen, R. K. C., Merico, D., Cao, H., Pellecchia, G., Alipanahi, B., Thiruvahindrapuram, B., . . .
1150 Scherer, S. W. (2016). Genome-wide characteristics of de novo mutations in autism. *NPJ*
1151 *Genomic Medicine, 1*, 160271-1602710. doi:10.1038/npjgenmed.2016.27

1152 Zhang, D., Stumpo, D. J., Graves, J. P., DeGraff, L. M., Grissom, S. F., Collins, J. B., . . .
1153 Blackshear, P. J. (2006). Identification of potential target genes for RFX4_v3, a
1154 transcription factor critical for brain development. *Journal of Neurochemistry, 98*(3), 860-
1155 875. doi:10.1111/j.1471-4159.2006.03930.x

1156 Zhang, P. G. Y., Yeung, J., Gupta, I., Ramirez, M., Ha, T., Swanson, D. J., . . . Goldowitz, D.
1157 (2018). Discovery of transcription factors novel to mouse cerebellar granule cell
1158 development through laser-capture microdissection. *Cerebellum (London, England),*
1159 doi:10.1007/s12311-017-0912-3

1160 Zhang, Y., Liu, T., Meyer, C. A., Eeckhoute, J., Johnson, D. S., Bernstein, B. E., . . . Liu, X. S.
1161 (2008). Model-based analysis of ChIP-seq (MACS). *Genome Biology, 9*(9), R137.
1162 doi:10.1186/gb-2008-9-9-r137

1163 Zhao, Y., Kwan, K., Mailloux, C. M., Lee, W., Grinberg, A., Wurst, W., . . . Westphal, H.
1164 (2007). LIM-homeodomain proteins Lhx1 and Lhx5, and their cofactor Ldb1, control
1165 purkinje cell differentiation in the developing cerebellum. *Proceedings of the National*
1166 *Academy of Sciences of the United States of America, 104*(32), 13182-13186.
1167 doi:10.1073/pnas.0705464104

1168 Zhu, M., Need, A. C., Han, Y., Ge, D., Maia, J. M., Zhu, Q., . . . Goldstein, D. B. (2012). Using
1169 ERDS to infer copy-number variants in high-coverage genomes. *American Journal of*
1170 *Human Genetics, 91*(3), 408-421. doi:10.1016/j.ajhg.2012.07.004

1171 Ziats, M. N., Grosvenor, L. P., & Rennert, O. M. (2015). Functional genomics of human brain
1172 development and implications for autism spectrum disorders. *Translational Psychiatry, 5*,
1173 e665. doi:10.1038/tp.2015.153