

***CHD8* Suppression Impacts on Histone H3 Lysine 36 Trimethylation and Alters RNA Alternative Splicing.**

Emanuela Kerschbamer,^{1*} Takshashila Tripathi,^{1*} Serkan Erdin,^{2,3} Elisa Salviato,⁴ Francesca Di Leva,¹ Endre Sebestyén,⁴ Michele Arnoldi,¹ Matteo Benelli,⁵ James F. Gusella,^{2,3,6} Silvano Piazza,⁷ Francesca Demichelis,⁸ Michael E. Talkowski,^{2,3,6} Francesco Ferrari,^{4,9} and Marta Biagioli^{1,#}

¹ Neuroepigenetics laboratory, Department of Cellular, Computational and Integrative Biology, (CIBIO) University of Trento, Trento, Italy

² Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, United States

³ Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, United States

⁴ IFOM, the FIRC Institute of Molecular Oncology, Milan, Italy

⁵ Bioinformatics Unit, Hospital of Prato, Istituto Toscano Tumori, Prato, Italy

⁶ Department of Neurology, Harvard Medical School, Boston, MA, United States

⁷ Bioinformatic facility, Department of Cellular, Computational and Integrative Biology (CIBIO) University of Trento, Italy

⁸ Laboratory of Computational and Functional Oncology, Department of Cellular, Computational and Integrative Biology (CIBIO), University of Trento, Trento, Italy

⁹ CNR Institute of Molecular Genetics "Luigi Luca Cavalli-Sforza", Pavia, Italy

* Joint co-first authors

Corresponding Author

Running Title: CHD8 Suppression alters RNA Alternative Splicing

Key words: Autism Spectrum Disorder, CHD8, Chromatin, H3K36me3, Alternative Splicing

Abstract

Disruptive mutations in the chromodomain helicase DNA binding protein 8 (*CHD8*) have been recurrently associated with Autism Spectrum Disorders (ASD). In normal cellular physiology, CHD8 co-purifies with MLL1 and MOF transcriptional activation complex, with elongating RNAPII and directly binds to DNA promoters and enhancers regions, thus a regulatory role in transcriptional initiation and elongation could be postulated.

Here we investigated how chromatin landscape reacts to *CHD8* suppression by analyzing a panel of histone modifications in induced pluripotent stem cell-derived neural progenitors. We interrogated transcriptionally active and repressed regions, as well as active and poised enhancers.

CHD8 suppression led to significant reduction (47.82%) in histone H3K36me3 peaks at gene bodies, particularly impacting on transcriptional elongation chromatin states. H3K36me3 reduction specifically affects highly expressed, CHD8-bound genes. Histone H3K36me3 reduction associated to *CHD8*-suppression does not functionally impact on global transcriptional levels, but correlated with altered alternative splicing patterns of ~ 2000 protein coding genes implicated in “RNA splicing”, “mitotic cell cycle phase transition” and “mRNA processing”, especially affecting alternative first exon and exon skipping events.

In summary, our results point toward broad molecular consequences of *CHD8* suppression, implicating altered histone deposition/maintenance and RNA processing regulation as important regulatory processes in ASD.

Introduction

De novo truncating mutations in *CHD8* have been reported and independently validated to be a strong risk factor for autism spectrum disorders (ASD) (Neale et al. 2012; O'Roak et al. 2012a; O'Roak et al. 2012b; Talkowski et al. 2012; Iossifov et al. 2014; Parikshak et al. 2013; De Rubeis et al. 2014). *CHD8* has been classified as a high confidence ASD candidate risk factor (score 1) in the Simons Foundation Autism Research Initiative (SFARI) [<https://gene.sfari.org> (Abrahams et al. 2013)] with most *CHD8* mutations, identified from ~70 ASD probands (www.ncbi.nlm.nih.gov/clinvar/ interrogated on January 8th, 2020), considered disruptive and resulting in protein haploinsufficiency. *CHD8* defines a subclass of ASD patients, displaying evident macrocephaly, distinct faces, sleep problems and gastrointestinal complaints (Bernier et al. 2014; Yasin et al. 2019). Most of these phenotypic characteristics were recapitulated in *chd8* knock-down zebrafish (Bernier et al. 2014; Sugathan et al. 2014) and, more recently, in *Chd8* suppression mouse models (Durak et al. 2016; Katayama et al. 2016; Suetterlin et al. 2018). Indeed, *chd8*-morpholino zebrafish and *Chd8* heterozygous mice display increased brain size, due to increased numbers of proliferating cells and newborn neurons, possibly initiated by altered gene expression in the developing neocortex (Sugathan et al. 2014; Suetterlin et al. 2018). Remarkably, genome-wide transcriptomic changes that impact on ASD related genes were also detected *in vitro*, in human neural progenitor cells (NPC) with reduced *CHD8* expression (Sugathan et al. 2014). Taken together, these observations suggest that aberrant genome-wide transcription leading to altered brain development is strictly correlated to reduced levels of *CHD8* function. However, the detailed molecular mechanism through which *CHD8* regulates this process still remains obscure.

A direct effect can be proposed since CHD8 is able to bind DNA at promoters and enhancer regions in hNPCs, mouse midfetal brain and embryonic cortex (Sugathan et al. 2014; Cotney et al. 2015). However, an indirect mechanism can also be postulated since other genes, not bound by CHD8, appear also to be transcriptionally dysregulated following *CHD8* suppression (Sugathan et al. 2014). Chromatin structure is intimately related to transcription, as DNA tightly packed around nucleosomes prevents transcription, while, conversely, the exchange or removal of nucleosomes allows free access to DNA, thus correlating with active gene expression (Venkatesh and Workman 2015). Using mass spectrometry, CHD8 was shown to co-purify with components of the MLL and CoREST, SWI/SNF and NuRD ATP-dependent remodeling complexes, supporting its possible role in transcriptional initiation (Thompson et al. 2008). On the other hand, CHD1, another member of the CHD protein family, has been shown to interact with the PAF1 transcription elongation complex maintaining H3K4me3/H3K36me3 domains at actively transcribed genes (Lee et al. 2017). Reduction of CHD1 alters H3K4me3 and H3K36me3 patterns, suggesting a role for CHD1 in establishing/maintaining the boundaries of these mutually exclusive histone marks (Lee et al. 2017). The SETD2 and SETD5 histone H3, Lysine 36 methyltransferases normally associate with RNAPII, and their activity results in increased H3K36me3 toward the 3' end of active genes (Krogan et al. 2003; Kizer et al. 2005). Loss of SETD2/5-dependent H3K36me in mammals (Sessa et al. 2019) as well as loss of yeast Chd1 causes reduced Rpd3S activity (histone deacetylase complex), increased acetylation, and increased cryptic transcription within gene bodies (Selth et al. 2010). Based on its placement on the phylogenetic tree and the presence of an ATPase domain (Thompson et al. 2008), CHD8 is most likely acting as an ATP-dependent chromatin remodeling factor; thus, similar to CHD1, CHD8 loss might cause increased nucleosome turnover and alterations in co-transcriptional

processes, such as cryptic transcription within gene bodies and alternative splicing (Radman-Livaja et al. 2012; Smolle et al. 2012).

Alternative splicing (AS), one of the major contributor to protein diversity in eukaryotes, is important for highly specialized cells, such as neurons. Aberrant splicing might contribute to neuronal dysfunction and has been associated with several neurological diseases (Cuajungco et al. 2003; Mordes et al. 2006) (Wang et al. 2012; Gompers et al. 2017). Broad chromatin conformation and transcriptional kinetics are major factors in the regulation of AS. Chromatin relaxation accelerates RNAPII processing and correlates with alternative exons skipping; conversely, packed nucleosomes slow down RNAPII progression causing pausing of transcription and the inclusion of non-constitutive weak exons (Luco et al. 2010; Nilsen and Graveley 2010; Luco and Misteli 2011). As experimental evidence is pointing to dysregulated chromatin regulation as a key feature in the pathogenesis of ASD, it is tempting to hypothesize that chromatin function of CHD proteins, and CHD8 in particular, might act to regulate RNA transcription, elongation and processing thereby being responsible for the characteristic neurodevelopmental effects observed in ASD.

In order to possibly dissect this mechanism, we characterized the consequences of *CHD8* suppression on the chromatin landscape, analyzing different histone modifications using chromatin immunoprecipitation and sequencing (ChIP-seq). Specifically, we interrogated histone marks characteristic of transcriptionally active (H3K4me2, H3K4me3, H3K27ac, H3K36me3) and repressed regions (H3K27me3) as well as active enhancers (H3K4me1, H3K27ac) in control induced pluripotent stem cell (iPSC)-derived neuronal progenitors cell line and in previously characterized lines where a ~50% reduction in *CHD8* was obtained by lentiviral delivery of shRNAs (Sugathan et al. 2014). We uncovered alterations affecting the H3K36me3 histone mark

in the body of highly transcribed genes, which do not primarily affect RNA transcription, but rather alter alternative splicing of genes implicated in “histone modification”, “covalent chromatin modification” and “mitotic cell cycle phase transition”, thus unveiling altered histone deposition/maintenance and RNA processing regulation as important regulatory processes in ASD.

Results

Epigenetic consequences of CHD8 suppression in human neuronal progenitor cells

To assess the functional consequences of *CHD8* suppression in chromatin organization, we resourced to previously characterized control iPSC-derived neuronal progenitors cell line, GM8330-8 and its derivatives where ~50% reduction in *CHD8* was obtained by lentiviral-mediated delivery of short-hairpin RNAs (shRNAs) (Sugathan et al. 2014). In these model systems, we analyzed six different histone modifications using chromatin immunoprecipitation and sequencing (ChIP-seq), specifically interrogating transcriptionally active (H3K4me2, H3K4me3, H3K27ac, H3K36me3) and repressed regions (H3K27me3) as well as active enhancers (H3K4me1, H3K27ac). For each of the six histone marks, three independent shRNAs targeting the coding sequence of *CHD8* (Sh1, Sh2, Sh4) and two technical replicate controls against *GFP* sequence were used (Fig. 1A). Sh1-*CHD8*, Sh2-*CHD8*, Sh4-*CHD8* presented nearly comparable levels of CHD8 at ~ 50% of the physiological levels, thus precisely mimicking the human haploinsufficiency condition (Sugathan et al. 2014). Importantly, for *CHD8*-knock-down models as well as for *GFP* controls, genome-wide transcriptomic data and CHD8 binding sites were available (Fig. 1A) (Sugathan et al. 2014).

ChIP-seq experiments were conducted to obtain on average 40M reads for narrow marks (H3K4me3/me2/me1, H3K27ac) and 60M for broad histone marks (H3K36me3/K27me3) and for INPUT samples. After mapping and filtering, an average of 42308 peaks per sample was identified (Suppl. Fig. 1A-B), and showed the expected enrichment pattern and metagene profiles at the transcriptional start site (H3K4me3/me2/me1, H3K27ac), gene body (H3K36me3) of actively transcribed genes or on larger genomic regions spanning transcriptionally silent gene

units (H3K27me3) and the expected enrichment correlation with ENCODE public datasets (Suppl. Fig. 1C-D) (Consortium 2012; Davis et al. 2018).

Upon *CHD8* suppression, H3K4me1, H3K27ac and H3K36me3 presented a substantial decrease in the number of peaks (37.44%, 38.45% and 47.82%, respectively) compared to control (intersection of Sh-*GFP* and Sh-*GFP2*) (Fig. 1B-C). Peaks detected exclusively in knock-down replicates (intersection of Sh2-*CHD8* and Sh4-*CHD8*, the two Sh-*CHD8* presenting nearly identical (46.63-48.61%) levels of *CHD8* suppression (Fig. 2C-D)) were limited (Fig. 1C). Notably, the third biological replicate Sh1-*CHD8*, presenting lower levels of *CHD8* suppression (Fig. 2C-D) (Sugathan et al. 2014), generally confirmed previous findings and the intersection of two biological replicates - in any of the possible combinations (Sh1-Sh2 *CHD8*, Sh2-Sh4 *CHD8*, Sh1-Sh4 *CHD8*), supported the reduction in H3K4me1, H3K27ac and H3K36me3 enrichment peaks (Suppl. Fig. 2A). In addition, Sh2-*CHD8* and Sh-*GFP* independent ChIP-seq datasets (generated in different laboratory than previous set) again sustained the conclusion that *CHD8* suppression was associated with a decreased H3K4me1 and H3K36me3 enrichment (Suppl. Fig. 2B).

Genes presenting reduced peaks for H3K4me1, H3K27ac and H3K36me3 histone marks were involved in “protein targeting”, “ribosome biogenesis” and “mitotic nuclear division” pathways (Fig. 1D) and enriched for ‘essential genes’ (required for a cell’s survival, Wang et al. 2015) and ‘constrained’ genes (intolerant to loss of function mutations, gnomAD), ‘FMRP targets in brain’ (Darnell et al. 2011), ‘post synaptic genes’ (Krishnan et al. 2016), and the ‘M3 co-expression module’ (Parikshak et al. 2013), whose expression peaks early during nervous system development (Fig. 1E).

CHD8 suppression significantly affects transcriptional elongation chromatin states

By combined analysis of the histone marks enrichment through ChromHMM (Ernst and Kellis 2012), we defined 10 types of genomic regions with a specific chromatin state in control hiNPC (1. transcriptional initiation (H3K4me3, H3K4me2, H3K4me1, H3K36me3), 2. transcriptional elongation (high H3K36me3), 3. weak transcribed (low H3K36me3), 4. strong enhancer (H3K4me2, H3K4me1, H3K27ac), 5. weak/poised enhancer a (H3K4me2, H3K4me1), 6. weak/poised enhancer b (H3K4me1, H3K27ac), 7. active promoter (H3K4me3, H3K4me2, H3K4me1, H3K27ac), 8. inactive/poised promoter (H3K27me3, H3K4me3, H3K4me2, H3K4me1), 9. Polycomb repressed (H3K27me3), 10. heterochromatin/low signal (no enrichment)) (Fig. 2 A-left). To identify states affected by *CHD8* knock-down, we then compared the histone marks (peak counts for each histone mark within each chromatin states) across two conditions, i.e. controls versus *CHD8* knock-down (Sh2-*CHD8* and Sh4-*CHD8*) (Fig. 2 A-right) and detected transcriptional elongation, strong-weak/poised enhancer and active promoter as the chromatin states most affected by *CHD8* suppression.

To further dissect these results, we compared the number of peaks for each histone mark in controls and *CHD8* knock-downs within each chromatin state. While differences between biological replicates did not support a statistically significant variation at enhancer and promoter states (Supp. Fig. 3), the number of peaks decorated by histone H3K36me3 at transcriptional elongation genomic regions was significantly lower in Sh2-Sh4 *CHD8* (p value < 0.05, t test) compared to controls (Fig. 2B). Within the transcriptional elongation regions, H3K36me3 reduction affected ~50% of all expressed genes in human neuronal progenitors (out of protein coding genes with > 2 TPM: 5447 lose H3K36me3, while 6451 were not affected). Additionally, genes that lost H3K36me3 following *CHD8*-haploinsufficiency, are significantly longer (90290

vs 60902 bp) and composed of larger exon number (7.36 vs 6.20) (not shown), compared to the unaffected ones. Importantly, independent western blot quantification of total histone H3K36me3 levels in the three Sh1-Sh2-Sh4 *CHD8* knock-down clones confirmed reduction of this histone modification compared to control Sh-*GFP* (Fig. 2 C-D). Genes with reduced H3K36me3 following *CHD8* suppression at transcriptional elongation sites were strongly enriched for processes related to cell cycle and proliferation (“cell cycle”, cell cycle process”, “mitotic cell cycle”, “ mitotic cell cycle process”) and “cell response to DNA damage stimulus” biological processes (Fig. 2E). For instance, the *HCN1*, gene encoding for a hyperpolarization gated chloride channel involved in epilepsy (Nava et al. 2014; Marini et al. 2018), presents reduced H3K36me3 enrichment following *CHD8* suppression, while other actively transcribed genomic locations (i.g. *BIRC6*) showed unaltered distribution (Suppl. Fig.4).

CHD8 suppression-dependent reduction in Histone H3 Lysine 36 trimethylation impacts on CHD8-bound genes.

By overlaying human CHD8 binding sites on the previously established chromatin states (Fig. 2A), we confirmed that CHD8 was confined to active promoters (90.11%) and, less prominently, to enhancers (strong, weak/poised a and b (3.18%, 3.00%, and 1.34%)) (Suppl. Fig. 5A) (Sugathan et al. 2014). As previously reported (Wade et al. 2018), CHD8 binding correlated with higher histone H3K36me3 (Suppl. Fig. 5B-E) and H3K4me3 enrichment metagene profiles (Suppl. Fig. 5C-F), as well as elevated RNA expression levels compared to CHD8-unbound genes (Suppl. Fig. 5 D).

Upon *CHD8* suppression, stringently defined (see materials and methods) *CHD8*-bound genes appeared to be more sensitive than *CHD8*-unbound genes (988 and 4205 respectively), presenting a significantly reduced H3K36me3 enrichment profile as confirmed by the effect size analysis (Fig. 3A-B. Suppl. Fig. 6A). Notably, the reduction in histone H3K36me3 elicited by *CHD8* suppression was specific since histone H3K4me3, another histone modification enriched at transcriptional start sites (TSS) of highly expressed genes (Suppl. Fig. 7A-B), remained unaltered (Suppl. Fig. 7C-E).

Remarkably, clustering analysis based on *CHD8* binding sites enrichment in control hiNPC identified 3 different clusters: Cluster #1 composed of 1239 genes with high *CHD8* enrichment (mean $\log_2(\text{ChIP}/\text{INPUT}) = 0.27$), cluster #2 composed of 2429 genes with medium-low *CHD8* enrichment (mean $\log_2(\text{ChIP}/\text{INPUT}) = -0.04$) and cluster #3 composed of 1380 genes with negligible *CHD8* enrichment (mean $\log_2(\text{ChIP}/\text{INPUT}) = -0.31$) (Fig. 3C-D). Strikingly, clusters #1 was strongly affected by *CHD8* decline, displaying significantly reduced H3K36me3 levels across the gene body (Fig. 3E and Suppl. Fig. 6B). Cluster #2 and #3, instead, with poor *CHD8* enrichment in control hiNPC, displayed a correspondingly lower H3K36me3 enrichment, with no significant difference following *CHD8* suppression (Fig. 3F-G and Suppl. Fig. 6C-D). In conclusion, this analysis confirmed that *CHD8*-bound genes were strongly sensitive to *CHD8* reduction, presenting a substantial and specific drop in H3K36me3 histone modification (Fig. 3 C-E, Suppl. Fig.6C-D). Functional enrichment of genes from Clusters #1, i.e. *CHD8*-bound and losing H3K36me3 enrichment upon *CHD8*-suppression, highlighted GO Biological Process terms related to ‘mRNA processing’ and ‘RNA splicing’ (Suppl. Fig. 8 and Suppl. Table 1).

Histone H3 Lysine 36 trimethylation elicited by CHD8 suppression alters RNA alternative splicing.

To gauge the functional significance of H3K36me3 reduction observed following *CHD8* suppression, we then leveraged RNA-seq data from controls and *CHD8*-knockdown clones. As histone H3K36me3 seems to be correlated with high levels of RNA expression (suppl. Fig. 7A) (Wagner and Carpenter 2012), we reasoned that a decline in H3K36me3 levels would be associated with reduced RNA expression levels. Unexpectedly, however, haploinsufficient levels of *CHD8* and, impaired H3K36me3 enrichment, didn't correspond to a global difference in transcription in either *CHD8*-bound or *CHD8*-unbound genes (Fig. 4A-B).

Importantly, a significant proportion of genes losing H3K36me3 upon *CHD8*-suppression (905 genes of which 326 (36%) with *CHD8* binding sites in promoter or enhancer regions) presented altered alternative splicing profiles, as evidenced by two analysis approaches (Fig. 4C-D and Suppl. Fig. 9) (Trincado et al. 2018; Shen et al. 2014). We noted an almost equal distribution of the H3K36me3 peaks lost following *CHD8* suppression between exonic and intronic regions (data not shown). While a significant correlation between reduction in H3K36me3 and alternative splicing was observed, however, the direct comparison of H3K36me3 lost peaks and differential splicing genomic coordinates revealed poor intersection. The vast majority of alternative splicing events with associated reduced H3K36me3 profiles were categorized as 'alternative first exon' (1892 (46.60%)) or 'exon skipping' (759 (18.69%)) (Fig. 4E,F and G). For the roughly ~2000 differential splicing events detected by SUPPA (Fig. 4C,D), the proportion of genes presenting positive or negative ΔPSI ($\Delta\text{PSI} = \text{PSI}_{\text{ctrl}} - \text{PSI}_{\text{KD}}$) (positive values: exon inclusion, negative values: exons skipping) remained pretty similar, however, for 'exon skipping' and 'alternative 5' events' types (SE and A5), the majority (60%)

presented a negative Δ PSI, suggesting that reduction of H3K36me3 could correlate with reduced exon inclusion/increased exon skipping. Among genes characterized by a reduction in H3K36me3 and presenting an altered splicing pattern following *CHD8* suppression, overrepresentation of GO terms and pathways related to “histone modification”, “covalent chromatin modification” and “mitotic cell cycle phase transition” (Suppl. Table 1) was observed.

Discussion

Disruption of *CHD8* from *de novo* protein truncating variants and structural variants is well established as a highly penetrant risk factor for ASD (Neale et al. 2012; Satterstrom et al. 2020; O'Roak et al. 2012a; Talkowski et al. 2012). Chromodomain helicase DNA-binding protein 8 (CHD8), initially described as interacting with β -catenin as negative regulator of WNT signaling (Sakamoto et al. 2000; Nishiyama et al. 2004; Nishiyama et al. 2009; Nishiyama et al. 2012), has important roles during nervous system development. Mice completely lacking *Chd8* exhibit early embryonic lethality (between E5.5 and E7.5) (Nishiyama et al. 2009), while *Chd8* hypomorphic mutations are associated with perinatal mortality, pronounced brain hyperplasia in surviving animals and altered transcriptional expression (Hurley et al. 2018; Suetterlin et al. 2018). CHD8, a member of the chromo-helicases, has the ability to directly bind DNA (Sugathan et al. 2014; Cotney et al. 2015) and to slide nucleosomes (Manning and Yusufzai 2017). ChIP-sequencing has revealed ~7000 CHD8 binding sites located at H3K27ac and H3K4me3 enriched regions, highly transcriptionally active promoters and enhancers (Sugathan et al. 2014; Cotney et al. 2015). However, with both direct and indirect transcriptional effects observed following *CHD8* suppression, its molecular mode of action in ASD remains unclear (Sugathan et al. 2014). CHD8 recruits histone KMT2/MLL methyltransferase complexes to induce mono-, di-, and trimethylation at lysine 4 of histone H3 and directly interacts with KMT2 components ASH2L and WDR5 (Thompson et al. 2008; Zhao et al. 2018). Reduced H3K27me3 enrichment was observed following *CHD8* suppression, which might reflect alteration to the core PRC2 methyltransferase, Ezh2 (Durak et al. 2016). However, a myriad of functions of CHD8 complicates this interpretation as it is associated with elongating RNAPII, so its role in

transcriptional elongation has to be considered, especially for highly expressed genes, that are densely decorated by histone H3K36me3 (Rodriguez-Paredes et al. 2009).

In our ASD-relevant, human neuronal progenitor model system, we observed that *CHD8* suppression is prominently associated with a depletion rather than a gain, in H3K4me1, H3K27ac and H3K36me3 histone modifications. Genes displaying reduction in these histones marks are involved in ‘Ribosome Biogenesis’, ‘Mitotic cell Division’, ‘Methyltransferase Activity’ and strongly enriched for ‘Essential, Intolerant to LoF genes’, whose expression peaks early during nervous system development (M2, M3, post conception week 10-12) or increases during early cortical development (M16, post conception week 25-26) (Parikshak et al. 2013). The observed reduction in histone marks, while appreciable in all cases, is subtle and may suggest that CHD8 is unlikely to be a core component or a crucial recruiting factor for KMT2/COMPASS complex or SETD2/SETD5 methyltransferases. Rather it might act as a facilitator of histone methyltransferase activity. Using a combinatorial analysis of histone marks enrichment (Ernst and Kellis 2012), we dissected the differences elicited by *CHD8* suppression into 10 chromatin states. Stringent statistical criteria to focus on the strongest and most reliable epigenetic changes supported a drastic depletion of H3K36me3 peaks in the transcription elongation chromatin state. While an effect at enhancers and promoters could not be completely ruled out, CHD8 seems to facilitate the methyltransferase activity leading to H3K36 trimethylation. Given the dynamic balance between the activity of histone methyltransferases and demethylases, we also cannot exclude the hypothesis that CHD8 might work as an inhibitor of KDM2B (He et al. 2013). GO terms associated with ‘cell cycle progression’ and ‘mitosis’ are enriched among genes losing H3K36me3 at transcriptional elongation states. Interestingly, SETD2-5/KDM2B dysregulation has been correlated to the change of cell-cycle regulators

(Kawakami et al. 2015), with altered H3K36 trimethylation associated with aberrant cell proliferation and neoplastic transformations (Wang et al. 2007; Jaffe et al. 2013; Simon et al. 2014). However, the modulation of H3K36 methylation is connected with the cell cycle not only in the context of tumorigenicity, but also during neuronal development and differentiation. Mutations in *SETD2* have been described in Sotos syndrome, childhood overgrowth condition with macrocephaly (Luscan et al. 2014) and in an ASD proband, also presenting macrocephaly (O'Roak et al. 2012a; Lumish et al. 2015), while disruptive mutations in *SETD5* – a newly described H3K36me3 methyltransferase (Sessa et al. 2019) – are associated with ID/ASD (Kuechler et al. 2015; Rauch et al., 2012; De Rubeis et al. 2014; Szczaluba et al. 2016; Parenti et al. 2017; Rawlins et al. 2017) and 3p25.3 microdeletion syndrome (Kellogg et al. 2013). Thus, it is possible that *CHD8* suppression, through its ability to modulate H3K36me3 levels, might lead to aberrant cell cycle regulation and macrocephaly as observed in animal models and in *CHD8*-autistic subjects (Bernier et al. 2014; Suetterlin et al. 2018). *CHD8* direct targets, transcriptionally upregulated upon *CHD8* suppression (i.e. PML, HDAC7, CDK6 (Sugathan et al. 2014)) and implicated in cell cycle progression were indeed significantly enriched (Fisher's exact test p value: 1.13×10^{-7}) among genes that lose H3K36me3, suggesting a functional interplay between cell-cycle, *CHD8* and H3K36me3.

But how can *CHD8* modulate H3K36me3? From RNA sequencing data, *CHD8* suppression doesn't seem to correlate with a direct reduction in *SETD2*/*SETD5* levels nor to upregulation in H3K36me3 KDM2B methyltransferase (not shown). Based on our data, genes bound at their promoters/enhancers by *CHD8* specifically present a significant depletion in H3K36me3, thus suggesting a possible direct interplay between the chromodomain and the H3K36 tri-methylases. Indeed, *SETD2* and *CHD8* display similar temporal expression patterns

during human brain development (data from BrainSpan Atlas) (Bernier et al. 2014), though direct binding between CHD8 and SETD2 or SETD5 still needs to be demonstrated. The H3K36me3 depletion, observed upon *CHD8* suppression, is reminiscent of the function of another member of the CHD family, CHD1, which remodels nucleosomes within the gene body of actively transcribed gene (Petty and Pillus 2013). In fact, *chd1* loss alters H3K4me3 and H3K36me3 patterns throughout the yeast genome and loss of *chd1* causes increased cryptic transcription, altered splicing and nucleosome turnover within gene bodies (Petty and Pillus 2013; Lee et al. 2017). On the other hand, KIS-L in *Drosophila* is associated with all sites of transcriptionally active chromatin in a pattern that largely overlaps that of RNA Polymerase II (Pol II). Moreover, *kismet* mutant larvae, present severe reduction in the levels of elongating Pol II, suggesting that Kismet is also required for transcription elongation (Srinivasan et al. 2005). Similarly, it is tempting to hypothesize that CHD8, recruited to highly transcribed genes and in concert with SET2, might play crucial roles in nucleosome stability during elongation (Krogan et al. 2003) (Huang and Zhu 2018).

Trimethylation of H3K36 demarcates body regions of actively transcribed genes, providing signals for modulating transcription fidelity, mRNA splicing and DNA damage repair (Wagner and Carpenter 2012). Aberrant reduction in *SETD2* and reduced levels of H3K36me3 in our data are not directly causative of transcriptional differences and this is coherent with previous findings (Simon et al. 2014). Rather, reduced H3K36me3 correlates with altered alternative splicing. This is likely linked to the kinetics of PolII progression, since, H3K36me3 decorated nucleosomes act as intrinsic pause sites for elongating RNAPII. Altered H3K36me3 enrichment, then, could alter splice site recognition and change exon inclusion (Wagner and Carpenter 2012). In any event, alternative splicing differences detected in this study as a

consequence of *CHD8* suppression, are particularly relevant in the context of nervous system development and function, especially for genes correlated with “RNA splicing” and “mRNA processing”. Interestingly, aberrant neuronal splicing has been previously related to *Chd8* haploinsufficiency in a mouse model of ASD (Gompers et al. 2017). In this context, our work establishing, for the first time, a functional link between H3K36me3 and *CHD8*, provides a molecular mechanism that warrants further investigations especially in neurodevelopmental syndromes and ASD in particular.

Finally, elongation-coupled H3K36 methylation usually serves also as a docking site for the histone deacetylase complex Rpd3S, which restores the repressive chromatin environment following Pol II passage to prevent cryptic transcription initiation (Huang and Zhu 2018). According to this hypothesis, *CHD8* suppression-related H3K36me3 reduction might also cause reduced repression in gene bodies, possibly resulting in increased uncontrolled, cryptic transcription. Although not possible to address in this study, which relied on previously generated poly-A plus RNA libraries, the analysis of spurious transcription as a consequence of H3K36me3 reduction, represents the next challenge for understanding chromatin-linked consequences in neurodevelopmental disorders and ASD in particular.

Methods

Cellular model

Human iPSC-derived NPC line GM8330-8 (Sheridan et al. 2011), Sh1-*CHD8*, Sh2-*CHD8*, Sh4-*CHD8* and Sh-*GFP* lines, previously generated by lentiviral delivery of shRNAs targeting *CHD8* and *GFP* coding sequences respectively (Sugathan et al. 2014), were kindly provided by Dr. Stephen Haggarty laboratory (Massachusetts General Hospital and Harvard Medical School, Boston, MA, U.S.A.).

Cells were cultured on poly-L-ornithine hydrobromide (20 µg/mL, Sigma)/laminin (3 µg/mL, Life Technologies) - coated plates in hiNPC medium (70% v/v DMEM (Life Technologies) completed with 30% v/v HAM F12 (Euroclone), 2% v/v B27 (Life Technologies), 1% v/v Penicillin-Streptomycin solution (Life Technologies) and 1% v/v L-Glutamine (Corning) and supplemented with EGF (20 ng/mL, Sigma), bFGF (20 ng/mL, R&D) and Heparin (5 µg/mL, Sigma)). Semi-confluent monolayers of hiNPCs were maintained in 5% CO₂, 37°C humidified incubator.

Protein extraction and western blot analysis

hiNPC cells were washed with PBS and resuspended in 100 ml of extraction buffer (10mM Hepes pH=8, 10mM KCl, 0.1mM MgCl₂, 0.1mM EDTA pH=8, 0.1mM DTT and halt protease & phosphatase inhibitor cocktail (Life Technology)). Samples were centrifuged at 5.000 rpm for 10 minutes at 4°C to remove cytosolic fraction. Nuclear pellets were resuspended in HCl 0.2N and put in rotation at 4°C overnight. After centrifugation at 4.000 rpm for 10 minutes at 4°C, supernatants containing nuclear proteins were recovered. Proteins were quantified by Bradford Protein Assay Kit (Sigma-Aldrich). Protein samples were separated by 4-12% Bis-Tris Protein

Gels (Thermo Fisher) and transferred on Amersham™ Protran™ 0.45μm Nitrocellulose (GE-healthcare) membrane. Membranes were blocked with 5% w/v nonfat dried milk and incubated with the following primary antibodies: anti-CHD8 (Novus Biologicals NB-10060417) (1:1000), anti-HSP90 (Cell Signaling Tech. #4874S) (1:5000), anti-Histone H3 (1:1.000) (Cell Signaling Tech, #4499), anti-Histone H3K36me3 (1:1.000) (Abcam #ab9050). Proteins were detected using horseradish peroxidase conjugated secondary antibodies anti-rabbit IgG 1:7.500 (GeneTex, #GTX213110-01), visualized by ECL Select WB detection reagent (GE Healthcare) following manufacturer's instructions. Signal quantification was performed with Imagelab software (BIORAD).

Statistical Analysis

The analysis of target proteins differences in western blotting experiments was evaluated by performing an unpaired, one-tailed t test. In all t tests, the significance level was set to 0.05. Data were represented as mean ± Standard Deviation (SD). The significance level was reported as: not significant (NS) $p > 0.05$, * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

Chromatin immunoprecipitation and sequencing (ChIP-seq)

Chromatin immunoprecipitation was performed using the protocol described by (Bernstein et al. 2006); with minor modifications. Briefly, ~25 million iPSC-derived neural progenitor cells, controls and *CHD8*-Shs were fixed with 1% formaldehyde incubated for 10 min at RT with rotation. The crosslinking was quenched by adding 1.1 ml 2.5M Glycine, incubated 5 min at RT with rotation. The cells were pelleted at 1000 rpm and resuspended ice-cold PBS/protease inhibitor (PI) and spin for 5 minutes at 1000 rpm, washed with ice-cold PBS twice, harvested, pelleted and directly resuspended in 300ul lysis buffer/PI (50 mM Tris-HCl (pH8.1), 1% SDS,

10 mM EDTA), kept on ice for 10 minutes rotating occasionally and vortexed vigorously for 15 seconds every 3 minutes. Sonication of the samples 200-700 bps smear was accomplished using a Bioruptor sonicator (Diagenode), for a total of 45 minutes sonication at full power and sonication cycles of 30"ON/30"OFF. Samples were centrifuged at max speed for 10 minutes at 4°C. Then, sheared-chromatin was diluted 10 fold in ChIP dilution buffer (16.7 mM Tris-HCl (pH 8.1), 167 mM NaCl, 0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA), supplemented with protease inhibitor. 50 µl of sheared chromatin was removed and stored at 4°C as control aliquot (INPUT). Each sample was incubated at 4°C overnight with antibodies (20 ug/ChIP) of interest. The following primary antibodies were used: H3K27me3 (07-449, Millipore), H3K4me3 (AB8580, Abcam), H3K36me3 (Ab9050, Abcam), H3K4me2 (AB7766, Abcam), H3K4me1 (AB8895, Abcam) and H3K27ac (AB4729, Abcam). Chromatin-Antibody complexes were precipitated with Dynabeads Protein A beads (Invitrogen) and washed sequentially with low-salt (20 mM Tris-HCl (pH 8.1), 150 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA), high-salt (20 mM Tris-HCl (pH 8.1), 500 mM NaCl, 0.1% SDS, 1% Triton X-100, 2 mM EDTA), LiCl (10 mM Tris-HCl (pH 8.1), 0.25 M LiCl, 1% NP40, 1% sodium deoxycholate, 1 mM EDTA), and TE wash buffers (10 mM Tris-HCl (pH 8.0), 1 mM EDTA). Immunoprecipitated chromatin and INPUT samples were then eluted in elution buffer (TE plus 1% SDS, 150 mM NaCl, 5 mM DTT), de-crosslinked at 65°C for overnight and treated with proteinase K. DNA isolation was performed by phenol:chloroform:isoamyl alcohol. DNA was precipitated with 200 mM NaCl, supplemented with 30 µg glycogen, washed with EtOH and then treated with RNase I (Invitrogen). Finally, DNA was purified with MinElute Kit (Qiagen). Quantification of ChIP and INPUT DNA was accomplished using Qubit 2.0 Fluorometer system (Invitrogen). ChIP-seq libraries were prepared using NEBNext UltraII DNA Library preparation kit (Illumina)

following manufacturer's instruction with no modification. ChIP-seq libraries were prepared starting from 5ng of fragmented DNA using NEBNext UltraII DNA Library preparation kit (Illumina) following manufacturer's instruction with no modifications. In order to obtain enough material for sequencing, eight cycles of PCR amplification were performed on adaptor ligated fragments.

Histone marks ChIP-seq Data Analysis

ChIP-seq reads were aligned on the human genome reference assembly GRCh38 using BWA (version 0.7.15) (Li and Durbin 2009). Aligned reads were filtered to discard unmapped, multiply mapped, PCR duplicates reads (Picard tools MarkDuplicates version: 2.3.0, Picard Toolkit." 2019. Broad Institute, GitHub Repository. <http://broadinstitute.github.io/picard/>) along with low quality alignments (samtools view -q 1, samtools version 1.7) (Li et al. 2009). Peak calling was performed with MACS2 (version 2.1.0) using a minimum FDR threshold 0.00001. The same settings with the addition of the --broad option were used for H3K27me3 and H3K36me3 marks (Zhang et al. 2008). Peaks localized to blacklisted regions (<http://mitra.stanford.edu/kundaje/akundaje/release/blacklists/>) or unplaced contigs were filtered. Narrow peaks (for H3K27ac, H3K4me1, H3K4me2, H3K4me3 histone marks) closer than 350 bp were merged into a single peak. Peaks were considered common between replicates if they overlapped by at least 50% of the length of the shortest peak, then extended coordinates were maintained and used in downstream analyses. CHD8 ChIP-seq reads from Sugathan et al. 2014 were realigned to the reference GRCh38 and peaks were recalled with the same procedure used for the narrow histone marks, with a default FDR threshold 0.05 as only peaks identified by all three antibodies were retained. GENCODE v.26 was used for peaks annotation (Frankish et al. 2019). Genes losing H3K4me1, H3K27ac and H3K36me3 peaks in *CHD8* knock-down were

tested for enrichment of GO terms with the enricher function from clusterProfiler (version 3.10.1, Yu et al. 2012) with Benjamini-Hochberg adjusted p value cutoff of 0.05. The full list of slimGO terms used was downloaded from Ensembl BioMart on September 9th, 2019. Enrichment of the same gene list was assessed on custom gene sets by using one-tailed Fisher's exact test. Custom gene sets were derived from public databases and publications of interest (Supplementary Table 1). Only non-redundant gene lists enriched at p value < 0.01 were shown in Fig. 1E, complete enrichment results are available in Supplementary Table 1. Combining the histone marks enrichment patterns over the genome, 10 chromatin states were identified for control (Sh-*GFP*, Sh-*GFP2*) using ChromHMM (version 1.14) (Ernst and Kellis 2012). The states were manually annotated according to the literature. The number of peaks called in these regions in both replicates for each histone mark was counted using BEDTools intersect, version 2.25.0 (Quinlan and Hall 2010), and the difference between control and *CHD8* knock-down was calculated. The total number of peaks for each histone marks in percentage was plotted as heat map in control cells (Fig. 2A), the difference (Fig. 2B) refers to the percentage of control peaks – *CHD8* KD peaks. The number of peaks per mark called in each chromatin state was tested with a two-sided t test considering two replicates in both control and *CHD8* knock-down. To generate the metagene profiles with deepTools, version 3.2.1, (Ramirez et al. 2016), ChIP-seq samples were normalized to INPUT with the SES method (Diaz et al. 2012). Enrichment was calculated in 10 bp bins over the gene body, scaled to 5 Kbp, and 2 Kbp up- and down-stream of the gene. In order to minimize noise in the metagene profile plots, a stringent set of filters was applied to protein coding genes before plotting: a minimum length of 2 Kbp, a minimum distance of 4 Kbp from other genes and absence of other features on the opposite strand, leading to a set of 9442 protein coding genes. For each histone mark, only genes with enrichment were plotted (at least

one non-zero bin). These genes were divided into three groups based on *CHD8* binding enrichment pattern via k-means with deepTools plotProfile command. To complement the statistical hypothesis testing, paired Cohen's d effect size statistics were calculated between groups along the entire region (Gibbons 1993). 99% simultaneous confidence intervals were constructed controlling FWER (Bonferroni correction). CHD8 binding sites profiles were calculated with deepTools computeMatrix reference-point and plotted with deepTools plotProfile. Visualization of enrichment tracks for chromatin was performed with the Integrative Genomic Viewer (IGV, version 2.4.9 (Robinson et al. 2011)).

RNA-seq Data Analysis

Raw reads obtained from Sugathan et al., 2014 for corresponding samples (Sh-*GFP*, Sh-*GFP2*, Sh2-*CHD8*, Sh4-*CHD8*) were used to calculate transcripts abundance by kallisto (version 0.44.0) (Bray et al. 2016) on GENCODE v.26 transcripts. Transcripts per Million (TPM) were used to plot expression levels. SUPPA (version 2.3) was used to calculate the percent spliced-in (PSI) value per splicing event with an empirical method (Trincado et al. 2018). The same raw reads from Sugathan et al. 2014 were aligned with STAR (version 2.6) (Dobin et al. 2013) with default parameters on the GRCh38 reference, to analyze the alternative splicing with rMATS (version 3.1.0, Shen et al. 2014). As the two alternative splicing analysis methods are complementary (<https://github.com/comprna/SUPPA/issues/47>), both were included in the analysis. To check the overlap of splicing events with chromatin marks, the spliced in/out exon coordinates were intersected with the histone mark peaks' coordinates. Alternative splicing events were represented in sashimi plots via ggsashimi (version 0.4.0) (Garrido-Martin et al. 2018) using GENCODE annotation v.33 as a reference (Frankish et al. 2019).

Acknowledgments

We are indebted to the Biagioli, Ferrari and Talkowski lab members for helpful discussions, to Dr. Albert Basson (King's College, UK) and Dr. Fulvio Chiacchiera (Department CIBIO, University of Trento) for critical reading of the manuscript. We are grateful to Dr. Alessandro Quattrone (Director, Department CIBIO, University of Trento) for the scientific support to CIBIO Core Facilities research. We thank Drs. Veronica De Sanctis and Roberto Bertorelli for technical support and assistance with preparation and sequencing of libraries. We thank Dr. Stephen Haggarty (Massachusetts General Hospital and Harvard Medical School, Boston, MA, U.S.A.) for sharing the human neuronal progenitor cell lines. This work was supported by Department CIBIO Institutional funding to M.B. and S.P., grants from Autism Speaks and the Simons Foundation Autism Research Initiative to J.F.G. and the National Institutes of Health GM061354 and NS093200 to J.F.G and M.E.T. E.S. was supported by the Structured International Post-doc Program of SEMM (SIPOD) and the AFM-TELETHON fellowship n. 21835.

Author contributions: E.K, T.T., S.E., E.Salviato, F.D.L., E.S., M.A., performed experiments and analyzed the data. E.Salviato assisted with statistical analysis. E.S. and M.Benelli assisted with data analysis of alternative splicing and histone modifications. E.K. and M.Biagioli designed the study and wrote the paper. J.F.G., S.P., F.D., F.F., and M.E.T. provided intellectual guidance and assistance with bioinformatic analyses. M. Biagioli, M.E.T. and F.F. supervised this study. All authors discussed the results and commented on the manuscript.

References

- Abrahams BS, Arking DE, Campbell DB, Mefford HC, Morrow EM, Weiss LA, Menashe I, Wadkins T, Banerjee-Basu S, Packer A. 2013. SFARI Gene 2.0: a community-driven knowledgebase for the autism spectrum disorders (ASDs). *Molecular autism* **4**: 36.
- Bernier R, Golzio C, Xiong B, Stessman HA, Coe BP, Penn O, Witherspoon K, Gerdt J, Baker C, Vulto-van Silfhout AT et al. 2014. Disruptive CHD8 mutations define a subtype of autism early in development. *Cell* **158**: 263-276.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315-326.
- Bray NL, Pimentel H, Melsted P, Pachter L. 2016. Near-optimal probabilistic RNA-seq quantification. *Nature biotechnology* **34**: 525-527.
- Consortium EP. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**: 57-74.
- Cotney J, Muhle RA, Sanders SJ, Liu L, Willsey AJ, Niu W, Liu W, Klei L, Lei J, Yin J et al. 2015. The autism-associated chromatin modifier CHD8 regulates other autism risk genes during human neurodevelopment. *Nat Commun* **6**: 6404.
- Cuajungco MP, Leyne M, Mull J, Gill SP, Lu W, Zagzag D, Axelrod FB, Maayan C, Gusella JF, Slaugenhaupt SA. 2003. Tissue-specific reduction in splicing efficiency of IKBKAP due to the major mutation associated with familial dysautonomia. *Am J Hum Genet* **72**: 749-758.

- Darnell JC, Van Driesche SJ, Zhang C, Hung KY, Mele A, Fraser CE, Stone EF, Chen C, Fak JJ, Chi SW et al. 2011. FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**: 247-261.
- Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, Hilton JA, Jain K, Baymuradov UK, Narayanan AK et al. 2018. The Encyclopedia of DNA elements (ENCODE): data portal update. *Nucleic acids research* **46**: D794-D801.
- De Rubeis S, He X, Goldberg AP, Poultney CS, Samocha K, Cicek AE, Kou Y, Liu L, Fromer M, Walker S et al. 2014. Synaptic, transcriptional and chromatin genes disrupted in autism. *Nature* **515**: 209-215.
- Diaz A, Park K, Lim DA, Song JS. 2012. Normalization, bias correction, and peak calling for ChIP-seq. *Statistical applications in genetics and molecular biology* **11**: Article 9.
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**: 15-21.
- Durak O, Gao F, Kaeser-Woo YJ, Rueda R, Martorell AJ, Nott A, Liu CY, Watson LA, Tsai LH. 2016. Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. *Nature neuroscience* **19**: 1477-1488.
- Ernst J, Kellis M. 2012. ChromHMM: automating chromatin-state discovery and characterization. *Nature methods* **9**: 215-216.
- Frankish A, Diekhans M, Ferreira AM, Johnson R, Jungreis I, Loveland J, Mudge JM, Sisu C, Wright J, Armstrong J et al. 2019. GENCODE reference annotation for the human and mouse genomes. *Nucleic acids research* **47**: D766-D773.

- Garrido-Martin D, Palumbo E, Guigo R, Breschi A. 2018. ggsashimi: Sashimi plot revised for browser- and annotation-independent splicing visualization. *PLoS computational biology* **14**: e1006360.
- Gibbons ED, Hedeker, D.R. Devis, J.M. 1993. Estimation of Effect Size From a Series of Experiments Involving Paired Comparisons. *Journal of Educational Statistics* **18**: 271-279.
- Gompers AL, Su-Feher L, Ellegood J, Copping NA, Riyadh MA, Stradleigh TW, Pride MC, Schaffler MD, Wade AA, Catta-Preta R et al. 2017. Germline Chd8 haploinsufficiency alters brain development in mouse. *Nature neuroscience* **20**: 1062-1073.
- He J, Shen L, Wan M, Taranova O, Wu H, Zhang Y. 2013. Kdm2b maintains murine embryonic stem cell status by recruiting PRC1 complex to CpG islands of developmental genes. *Nature cell biology* **15**: 373-384.
- Huang C, Zhu B. 2018. Roles of H3K36-specific histone methyltransferases in transcription: antagonizing silencing and safeguarding transcription fidelity. *Biophysics reports* **4**: 170-177.
- Hurley S, Mohan C, Suetterlin P, Ellegood J, Rudari F, Lerch JP, Fernandes C, Basson MA. 2018. Non-monotonic regulation of gene expression, neural progenitor fate and brain growth by the chromatin remodeller CHD8. *bioRxiv*: 469031.
- Iossifov I, O'Roak BJ, Sanders SJ, Ronemus M, Krumm N, Levy D, Stessman HA, Witherspoon KT, Vives L, Patterson KE et al. 2014. The contribution of de novo coding mutations to autism spectrum disorder. *Nature* **515**: 216-221.

- Jaffe JD, Wang Y, Chan HM, Zhang J, Huether R, Kryukov GV, Bhang HE, Taylor JE, Hu M, Englund NP et al. 2013. Global chromatin profiling reveals NSD2 mutations in pediatric acute lymphoblastic leukemia. *Nature genetics* **45**: 1386-1391.
- Katayama Y, Nishiyama M, Shoji H, Ohkawa Y, Kawamura A, Sato T, Suyama M, Takumi T, Miyakawa T, Nakayama KI. 2016. CHD8 haploinsufficiency results in autistic-like phenotypes in mice. *Nature* **537**: 675-679.
- Kawakami E, Tokunaga A, Ozawa M, Sakamoto R, Yoshida N. 2015. The histone demethylase Fbxl11/Kdm2a plays an essential role in embryonic development by repressing cell-cycle regulators. *Mechanisms of development* **135**: 31-42.
- Kellogg G, Sum J, Wallerstein R. 2013. Deletion of 3p25.3 in a patient with intellectual disability and dysmorphic features with further definition of a critical region. *American journal of medical genetics Part A* **161A**: 1405-1408.
- Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD. 2005. A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Molecular and cellular biology* **25**: 3305-3316.
- Krishnan A, Zhang R, Yao V, Theesfeld CL, Wong AK, Tadych A, Volfovsky N, Packer A, Lash A, Troyanskaya OG. 2016. Genome-wide prediction and functional characterization of the genetic basis of autism spectrum disorder. *Nature neuroscience* **19**: 1454-1462.
- Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C et al. 2003. Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Molecular and cellular biology* **23**: 4207-4218.

- Kuechler A, Zink AM, Wieland T, Ludecke HJ, Cremer K, Salviati L, Magini P, Najafi K, Zweier C, Czeschik JC et al. 2015. Loss-of-function variants of SETD5 cause intellectual disability and the core phenotype of microdeletion 3p25.3 syndrome. *European journal of human genetics : EJHG* **23**: 753-760.
- Lee Y, Park D, Iyer VR. 2017. The ATP-dependent chromatin remodeler Chd1 is recruited by transcription elongation factors and maintains H3K4me3/H3K36me3 domains at actively transcribed and spliced genes. *Nucleic acids research* **45**: 7180-7190.
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**: 1754-1760.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Genome Project Data Processing S. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078-2079.
- Luco RF, Misteli T. 2011. More than a splicing code: integrating the role of RNA, chromatin and non-coding RNA in alternative splicing regulation. *Current opinion in genetics & development* **21**: 366-372.
- Luco RF, Pan Q, Tominaga K, Blencowe BJ, Pereira-Smith OM, Misteli T. 2010. Regulation of alternative splicing by histone modifications. *Science* **327**: 996-1000.
- Lumish HS, Wynn J, Devinsky O, Chung WK. 2015. Brief Report: SETD2 Mutation in a Child with Autism, Intellectual Disabilities and Epilepsy. *Journal of autism and developmental disorders* **45**: 3764-3770.
- Luscan A, Laurendeau I, Malan V, Francannet C, Odent S, Giuliano F, Lacombe D, Touraine R, Vidaud M, Pasmant E et al. 2014. Mutations in SETD2 cause a novel overgrowth condition. *Journal of medical genetics* **51**: 512-517.

- Manning BJ, Yusufzai T. 2017. The ATP-dependent chromatin remodeling enzymes CHD6, CHD7, and CHD8 exhibit distinct nucleosome binding and remodeling activities. *The Journal of biological chemistry* **292**: 11927-11936.
- Marini C, Porro A, Rastetter A, Dalle C, Rivolta I, Bauer D, Oegema R, Nava C, Parrini E, Mei D et al. 2018. HCN1 mutation spectrum: from neonatal epileptic encephalopathy to benign generalized epilepsy and beyond. *Brain : a journal of neurology* **141**: 3160-3178.
- Mordes D, Luo X, Kar A, Kuo D, Xu L, Fushimi K, Yu G, Sternberg P, Jr., Wu JY. 2006. Pre-mRNA splicing and retinitis pigmentosa. *Mol Vis* **12**: 1259-1271.
- Nava C, Dalle C, Rastetter A, Striano P, de Kovel CG, Nabbout R, Cances C, Ville D, Brilstra EH, Gobbi G et al. 2014. De novo mutations in HCN1 cause early infantile epileptic encephalopathy. *Nature genetics* **46**: 640-645.
- Neale BM, Kou Y, Liu L, Ma'ayan A, Samocha KE, Sabo A, Lin CF, Stevens C, Wang LS, Makarov V et al. 2012. Patterns and rates of exonic de novo mutations in autism spectrum disorders. *Nature* **485**: 242-245.
- Nilsen TW, Graveley BR. 2010. Expansion of the eukaryotic proteome by alternative splicing. *Nature* **463**: 457-463.
- Nishiyama M, Nakayama K, Tsunematsu R, Tsukiyama T, Kikuchi A, Nakayama KI. 2004. Early embryonic death in mice lacking the beta-catenin-binding protein Duplin. *Molecular and cellular biology* **24**: 8386-8394.
- Nishiyama M, Oshikawa K, Tsukada Y, Nakagawa T, Iemura S, Natsume T, Fan Y, Kikuchi A, Skoultschi AI, Nakayama KI. 2009. CHD8 suppresses p53-mediated apoptosis through histone H1 recruitment during early embryogenesis. *Nature cell biology* **11**: 172-182.

- Nishiyama M, Skoultschi AI, Nakayama KI. 2012. Histone H1 recruitment by CHD8 is essential for suppression of the Wnt-beta-catenin signaling pathway. *Molecular and cellular biology* **32**: 501-512.
- O'Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, Carvill G, Kumar A, Lee C, Ankenman K et al. 2012a. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* **338**: 1619-1622.
- O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, Levy R, Ko A, Lee C, Smith JD et al. 2012b. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**: 246-250.
- Parenti I, Teresa-Rodrigo ME, Pozojevic J, Ruiz Gil S, Bader I, Braunholz D, Bramswig NC, Gervasini C, Larizza L, Pfeiffer L et al. 2017. Mutations in chromatin regulators functionally link Cornelia de Lange syndrome and clinically overlapping phenotypes. *Human genetics* **136**: 307-320.
- Parikshak NN, Luo R, Zhang A, Won H, Lowe JK, Chandran V, Horvath S, Geschwind DH. 2013. Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell* **155**: 1008-1021.
- Petty E, Pillus L. 2013. Balancing chromatin remodeling and histone modifications in transcription. *Trends in genetics : TIG* **29**: 621-629.
- Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**: 841-842.
- Radman-Livaja M, Quan TK, Valenzuela L, Armstrong JA, van Welsem T, Kim T, Lee LJ, Buratowski S, van Leeuwen F, Rando OJ et al. 2012. A key role for Chd1 in histone H3 dynamics at the 3' ends of long genes in yeast. *PLoS Genet* **8**: e1002811.

- Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic acids research* **44**: W160-165.
- Rawlins LE, Stals KL, Eason JD, Turnpenny PD. 2017. De novo SETD5 nonsense mutation associated with diaphragmatic hernia and severe cerebral cortical dysplasia. *Clinical dysmorphology* **26**: 95-97.
- Robinson JT, Thorvaldsdottir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. *Nature biotechnology* **29**: 24-26.
- Rodriguez-Paredes M, Ceballos-Chavez M, Esteller M, Garcia-Dominguez M, Reyes JC. 2009. The chromatin remodeling factor CHD8 interacts with elongating RNA polymerase II and controls expression of the cyclin E2 gene. *Nucleic acids research* **37**: 2449-2460.
- Sakamoto I, Kishida S, Fukui A, Kishida M, Yamamoto H, Hino S, Michiue T, Takada S, Asashima M, Kikuchi A. 2000. A novel beta-catenin-binding protein inhibits beta-catenin-dependent Tcf activation and axis formation. *The Journal of biological chemistry* **275**: 32871-32878.
- Satterstrom FK, Kosmicki JA, Wang J, Breen MS, De Rubeis S, An JY, Peng M, Collins R, Grove J, Klei L et al. 2020. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell* **180**: 568-584 e523.
- Selth LA, Sigurdsson S, Svejstrup JQ. 2010. Transcript Elongation by RNA Polymerase II. *Annu Rev Biochem* **79**: 271-293.
- Sessa A, Fagnocchi L, Mastrototaro G, Massimino L, Zaghi M, Indrigo M, Cattaneo S, Martini D, Gabellini C, Pucci C et al. 2019. SETD5 Regulates Chromatin Methylation State and

- Preserves Global Transcriptional Fidelity during Brain Development and Neuronal Wiring. *Neuron* **104**: 271-289 e213.
- Shen S, Park JW, Lu ZX, Lin L, Henry MD, Wu YN, Zhou Q, Xing Y. 2014. rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. *Proc Natl Acad Sci U S A* **111**: E5593-5601.
- Sheridan SD, Theriault KM, Reis SA, Zhou F, Madison JM, Daheron L, Loring JF, Haggarty SJ. 2011. Epigenetic characterization of the FMR1 gene and aberrant neurodevelopment in human induced pluripotent stem cell models of fragile X syndrome. *PloS one* **6**: e26203.
- Simon JM, Hacker KE, Singh D, Brannon AR, Parker JS, Weiser M, Ho TH, Kuan PF, Jonasch E, Furey TS et al. 2014. Variation in chromatin accessibility in human kidney cancer links H3K36 methyltransferase loss with widespread RNA processing defects. *Genome research* **24**: 241-250.
- Smolle M, Venkatesh S, Gogol MM, Li H, Zhang Y, Florens L, Washburn MP, Workman JL. 2012. Chromatin remodelers Isw1 and Chd1 maintain chromatin structure during transcription by preventing histone exchange. *Nat Struct Mol Biol* **19**: 884-892.
- Srinivasan S, Armstrong JA, Deuring R, Dahlsveen IK, McNeill H, Tamkun JW. 2005. The *Drosophila* trithorax group protein Kismet facilitates an early step in transcriptional elongation by RNA Polymerase II. *Development* **132**: 1623-1635.
- Suetterlin P, Hurley S, Mohan C, Riegman KLH, Pagani M, Caruso A, Ellegood J, Galbusera A, Crespo-Enriquez I, Michetti C et al. 2018. Altered Neocortical Gene Expression, Brain Overgrowth and Functional Over-Connectivity in Chd8 Haploinsufficient Mice. *Cereb Cortex* **28**: 2192-2206.

- Sugathan A, Biagioli M, Golzio C, Erdin S, Blumenthal I, Manavalan P, Ragavendran A, Brand H, Lucente D, Miles J et al. 2014. CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. *Proc Natl Acad Sci U S A* **111**: E4468-4477.
- Szczaluba K, Brzezinska M, Kot J, Rydzanicz M, Walczak A, Stawinski P, Werner B, Ploski R. 2016. SETD5 loss-of-function mutation as a likely cause of a familial syndromic intellectual disability with variable phenotypic expression. *American journal of medical genetics Part A* **170**: 2322-2327.
- Talkowski ME, Rosenfeld JA, Blumenthal I, Pillalamarri V, Chiang C, Heilbut A, Ernst C, Hanscom C, Rossin E, Lindgren AM et al. 2012. Sequencing chromosomal abnormalities reveals neurodevelopmental loci that confer risk across diagnostic boundaries. *Cell* **149**: 525-537.
- Thompson BA, Tremblay V, Lin G, Bochar DA. 2008. CHD8 is an ATP-dependent chromatin remodeling factor that regulates beta-catenin target genes. *Molecular and cellular biology* **28**: 3894-3904.
- Trincado JL, Entizne JC, Hysenaj G, Singh B, Skalic M, Elliott DJ, Eyraas E. 2018. SUPPA2: fast, accurate, and uncertainty-aware differential splicing analysis across multiple conditions. *Genome biology* **19**: 40.
- Venkatesh S, Workman JL. 2015. Histone exchange, chromatin structure and the regulation of transcription. *Nature reviews Molecular cell biology* **16**: 178-189.
- Wade AA, Lim K, Catta-Preta R, Nord AS. 2018. Common CHD8 Genomic Targets Contrast With Model-Specific Transcriptional Impacts of CHD8 Haploinsufficiency. *Frontiers in molecular neuroscience* **11**: 481.

- Wagner EJ, Carpenter PB. 2012. Understanding the language of Lys36 methylation at histone H3. *Nature reviews Molecular cell biology* **13**: 115-126.
- Wang ET, Cody NA, Jog S, Biancolella M, Wang TT, Treacy DJ, Luo S, Schroth GP, Housman DE, Reddy S et al. 2012. Transcriptome-wide regulation of pre-mRNA splicing and mRNA localization by muscleblind proteins. *Cell* **150**: 710-724.
- Wang GG, Cai L, Pasillas MP, Kamps MP. 2007. NUP98-NSD1 links H3K36 methylation to Hox-A gene activation and leukaemogenesis. *Nature cell biology* **9**: 804-812.
- Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM. 2015. Identification and characterization of essential genes in the human genome. *Science* **350**: 1096-1101.
- Yasin H, Gibson WT, Langlois S, Stowe RM, Tsang ES, Lee L, Poon J, Tran G, Tyson C, Wong CK et al. 2019. A distinct neurodevelopmental syndrome with intellectual disability, autism spectrum disorder, characteristic facies, and macrocephaly is caused by defects in CHD8. *Journal of human genetics* **64**: 271-280.
- Yu G, Wang LG, Han Y, He QY. 2012. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology* **16**: 284-287.
- Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W et al. 2008. Model-based analysis of ChIP-Seq (MACS). *Genome biology* **9**: R137.
- Zhao C, Dong C, Frahm M, Deng Y, Marie C, Zhang F, Xu L, Ma Z, Dong X, Lin Y et al. 2018. Dual Requirement of CHD8 for Chromatin Landscape Establishment and Histone Methyltransferase Recruitment to Promote CNS Myelination and Repair. *Developmental cell* **45**: 753-768 e758.

Figures

Kerschbamer_Fig1

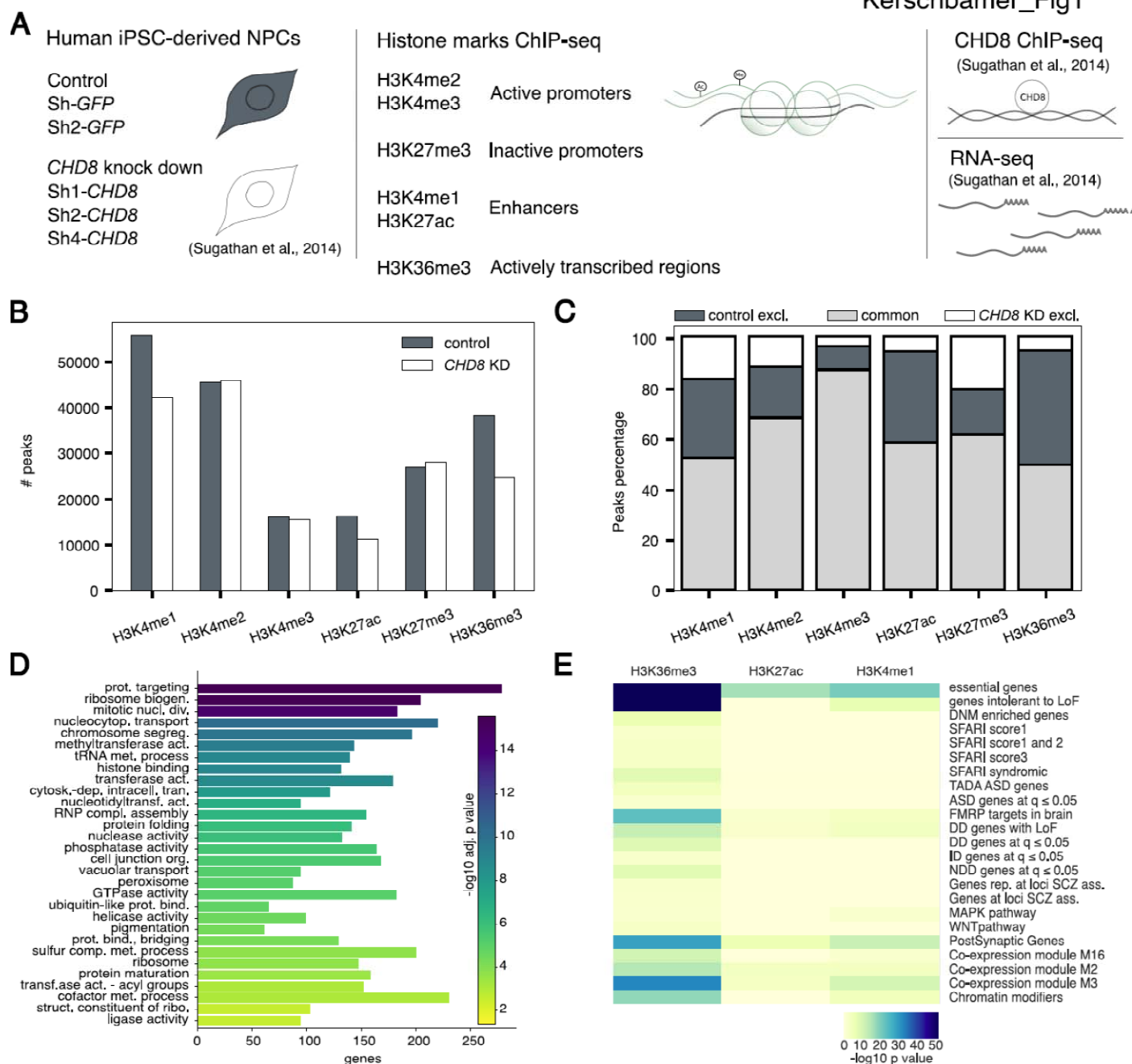


Figure 1 Altered chromatin landscape upon *CHD8*-suppression: reduction of histone *H3K27ac*, *H3K36me3* and *H3K4me1*.

A. Schematic representation of the study design and integrative approach used in this work. Human iPSC-derived NPCs (hiNPC) knocked-down for *CHD8* (Sh1-, Sh2- and Sh4-*CHD8*) and control hiNPCs (Sh-*GFP*, Sh-*GFP2*) (Sugathan et al. 2014), were analyzed via ChIP-seq for six histone marks representative of different chromatin regions - active promoters (H3K4me2, H3K4me3), inactive promoters (H3K27me3), enhancers (H3K4me1, H3K27ac) and actively transcribed regions (H3K36me3). ChIP-seq results were subsequently integrated with *CHD8* binding sites and available transcriptomics (RNA-seq) datasets obtained from the same model system (Sugathan et al. 2014).

B. The bar plot reports the total number of peaks identified in controls (grey) and *CHD8* knock-downs (white) for each histone mark analyzed. The y-axis presents the number of peaks resulting from the intersection of two replicates for each experimental condition: controls (intersection between Sh-*GFP*, Sh-*GFP2*), *CHD8*-knock-down (intersection between Sh2 and Sh4 *CHD8*). Histone marks H3K4me1, H3K27ac and H3K36me3 were the mostly affected by *CHD8* knock-down, displaying, in all cases, decreased number of peaks in *CHD8* suppression hiNPC models (19388, 6445 and 20507 peaks were lost, respectively).

C. The bar plot highlights the percentage of peaks in common between controls (intersection between Sh-*GFP*, Sh-*GFP2*) and *CHD8*-knock-down (intersection between Sh2- and Sh4-*CHD8*) (light gray). Percentage of peaks exclusively present in controls (intersection between Sh-*GFP*, Sh-*GFP2*) or in *CHD8* knock-down (intersection between Sh2- and Sh4-*CHD8*) are presented in grey and white, respectively. The total number of peaks (100%) is the sum of the peaks common and exclusive to each experimental condition. Upon *CHD8* suppression, H3K4me1, H3K27ac and H3K36me3 confirm extensive peak loss compared to controls, with a reduction of 1.22%, 36.23% and 45.20%, respectively.

D. Graphical representation of slim GO functional annotation enrichment for all genes interested by H3K4me1, H3K27ac and H3K36me3 peaks depletion following *CHD8* knock-down. Top 30 GO terms are shown in figure, while full list of significant terms is reported in Suppl. Table 1. Color-coded bar plot according to adjusted p values in $-\log_{10}$ scale, displays statistically significant slim GOs. The size of the bar (x-axis) reports the number of genes for each slim GO term.

E. The heatmap represents gene set enrichment p values in $-\log_{10}$ scale for all genes losing H3K4me1, H3K27ac and H3K36me3 peaks following *CHD8* knock-down. Genes' lists related to ASD, neurodevelopment, co-expression modules in brain and intolerance to loss of function, were tested for enrichment as described in Materials and Methods. Full gene list description and enrichment results are available in Supplementary Table 1.

Kerschbamer_Fig2

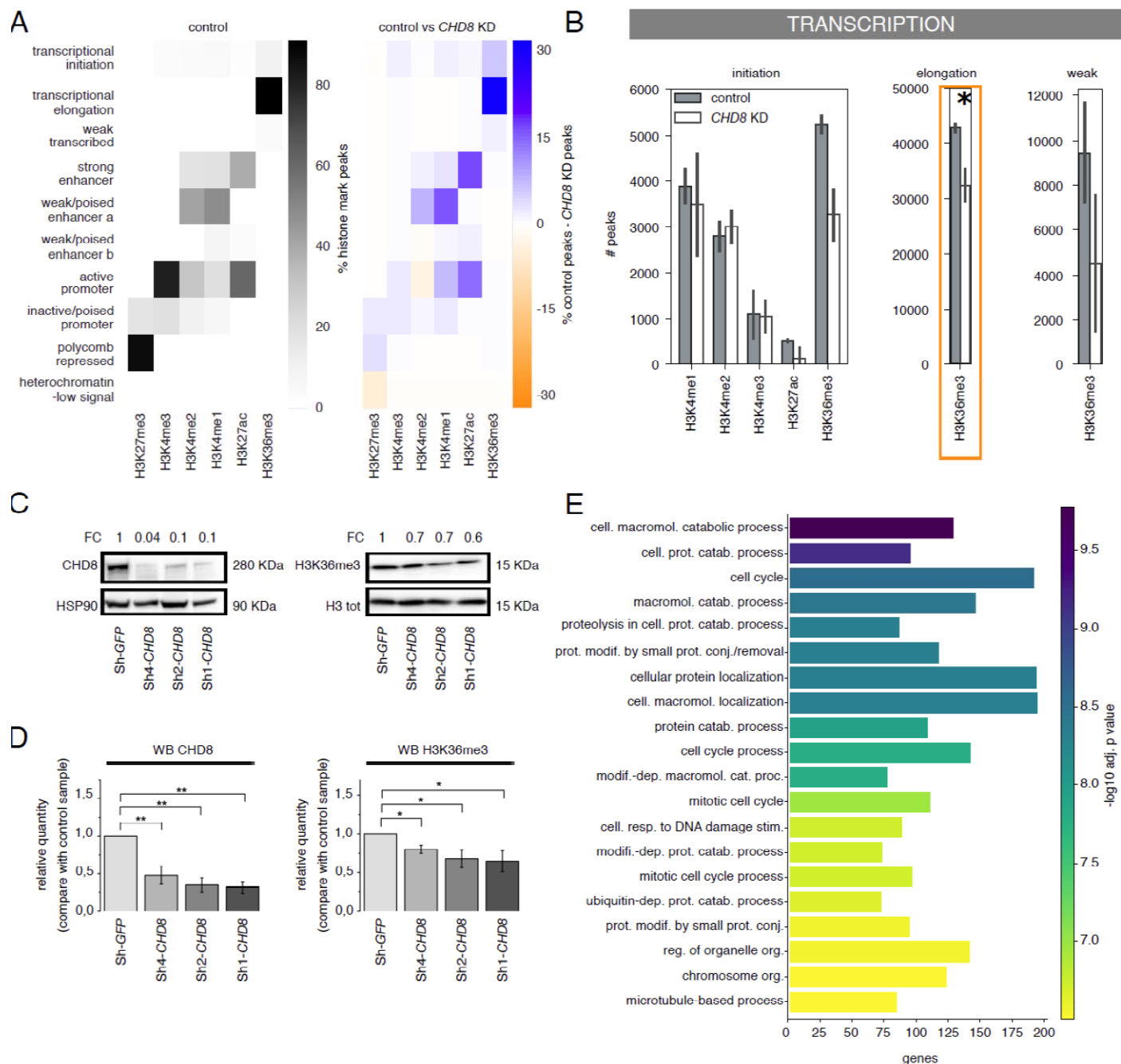


Figure 2 CHD8-suppression significantly impacts on histone H3K36me3 enrichment at transcriptional elongation sites.

A. The heatmaps represent 10 different chromatin states (1. transcriptional initiation, 2. transcriptional elongation, 3. weakly transcribed, 4. strong enhancer, 5. weak/poised enhancer a,

6. weak/poised enhancer b, 7. active promoter, 8. inactive/poised promoter, 9. polycomb repressed, 10. heterochromatin/low signal), defined by the combination of different histone marks in control hiNPC as defined by ChromHMM(Ernst and Kellis 2012). The distribution of histone marks peaks across different chromatin states (see materials and methods for details) is presented as percentage of the total and color-coded in the heatmap (left).

On the right, the heatmap describes the difference in number of peaks between two experimental conditions (controls versus *CHD8* knock-down). In blue are indicated chromatin states enriched in control, in orange chromatin states enriched in *CHD8* knock-down. H3K36me3 in transcriptional elongation is identified as the most affected chromatin state.

B. The bar plots represent the number of peaks for each histone mark identified at transcriptional initiation (left), elongation (center) and weakly transcribed (right) genomic regions. Grey bars indicate controls (n=2, Sh-*GFP*, Sh-*GFP2*) and white bars refer to *CHD8* knock-down (n=2, Sh2-*CHD8*, Sh4-*CHD8*). H3K36me3 peak loss upon *CHD8* suppression was significant at the transcriptional elongation states (two biological replicates, T-test, $p < 0.05$).

C. The image describes a representative western blot illustrating *CHD8* (left) and total histone H3K36me3 (right) levels across control (Sh-*GFP*) and *CHD8* knock-down clones (Sh1-*CHD8*, Sh2-*CHD8*, Sh4-*CHD8*). Down-regulation of *CHD8* and total H3K36me3 amount is observed following *CHD8* suppression. Levels of *CHD8* and H3K36me3 reduction are indicated in the top part of the panel as fold change (FC) compared to control Sh-*GFP*. Comparable amounts of total protein were loaded and HSP90 and total histone H3 was used as loading control.

D. The bar chart reports fold change differences in *CHD8* (left) and histone H3K36me3 (left) levels comparing control (Sh-*GFP*) and *CHD8* knock-down clones (Sh1-*CHD8*, Sh2-*CHD8*, Sh4-*CHD8*) in western blotting experiments. The bars represent normalized (HSP90 and total

histone H3 were used as loading control) *CHD8* and H3K36me3 values relative to *Sh-GFP* controls. Mean values \pm s.e. from independent biological replicates (n=4 for *Sh4-CHD8* and n=6 for the other samples) are plotted. t test for two mean population was performed. * $p \leq 0.05$.

E. Bar plot represents GO biological process terms significantly enriched in genes with a reduced level of H3K36me3 at transcriptional elongation sites upon *CHD8* suppression. The bars are ordered according to adjusted p values in -log10 scale, the x-axis represents the number of genes enriched for each term. Top 20 GO terms are shown in figure, while full list of significant terms is reported in Suppl. Table 1.

Kerschbamer_Fig3

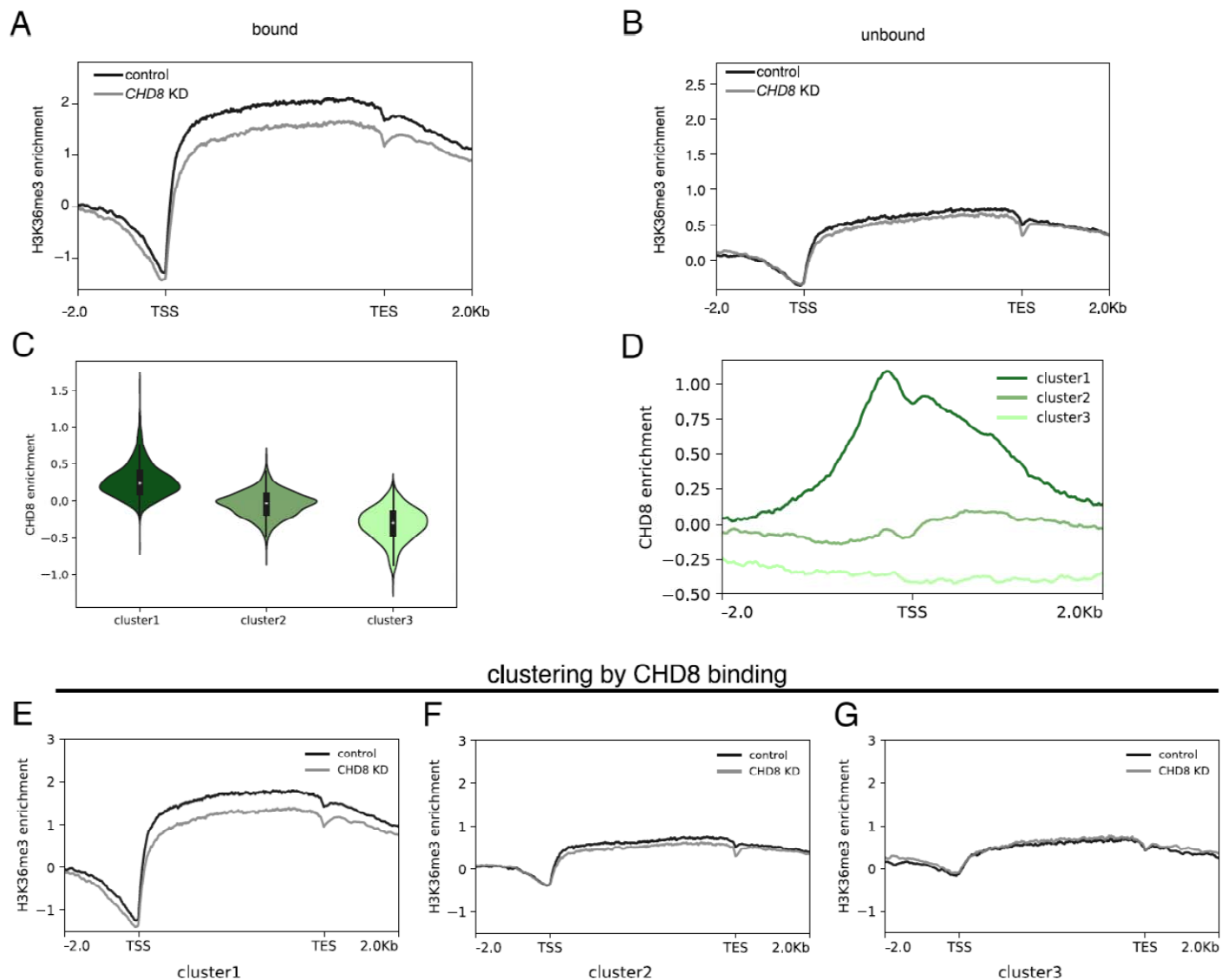


Figure 3 CHD8-suppression correlates with reduced H3K36me3 enrichment preferentially at CHD8-bound genes.

A. B. Metagenes profiles display the average of histone H3K36me3 enrichment (scaled log₂ ratio of normalized ChIP value over INPUT control - see also Materials and Methods) in a region of ± 2 Kbp upstream the transcriptional start site (TSS) and downstream the transcriptional end site (TES), calculated for control (black line) and CHD8 knock-down (grey line) hiNPC and for

CHD8-bound (#988) (A) and CHD8-unbound genes (#4205) (B). The difference between H3K36me3 enrichment in control and *CHD8* knock-down is significant for CHD8-bound genes, but not for CHD8 unbound genes (paired Cohen's d effect size statistics).

C. The violin plots represent the level of CHD8 binding enrichment ($\log_2(\text{ChIP/INPUT})$) for 3 groups of genes as clustered by k-means. Cluster #1 composed of 1239 genes shows high CHD8 enrichment (mean $\log_2(\text{ChIP/INPUT}) = 0.27$), cluster #2 composed of 2429 genes with medium-low CHD8 enrichment (mean $\log_2(\text{ChIP/INPUT}) = -0.04$) and cluster #3 composed of 1380 genes displays negligible CHD8 enrichment (mean $\log_2(\text{ChIP/INPUT}) = -0.31$).

D. Metagene profiles show the average of CHD8 binding enrichment in a region of ± 2 Kbp around the TSS, calculated for the three clusters #1, #2, #3 as described in C.

E. F. G. Metagene profiles display the average of histone H3K36me3 enrichment ($\log_2(\text{ChIP/INPUT})$) in a region of ± 2 Kbp upstream the TSS and downstream the TES, calculated for control (black line) and *CHD8* knock-down (grey line) hiNPC for each of the 3 clusters identified in Fig. 3C. The difference between H3K36me3 enrichment in control and *CHD8* knock-down is significant for Cluster#1, but not for Cluster#2 or Cluster#3 (paired Cohen's d effect size statistics, see suppl. Fig. 6B-D).

Kerschbamer_Fig4

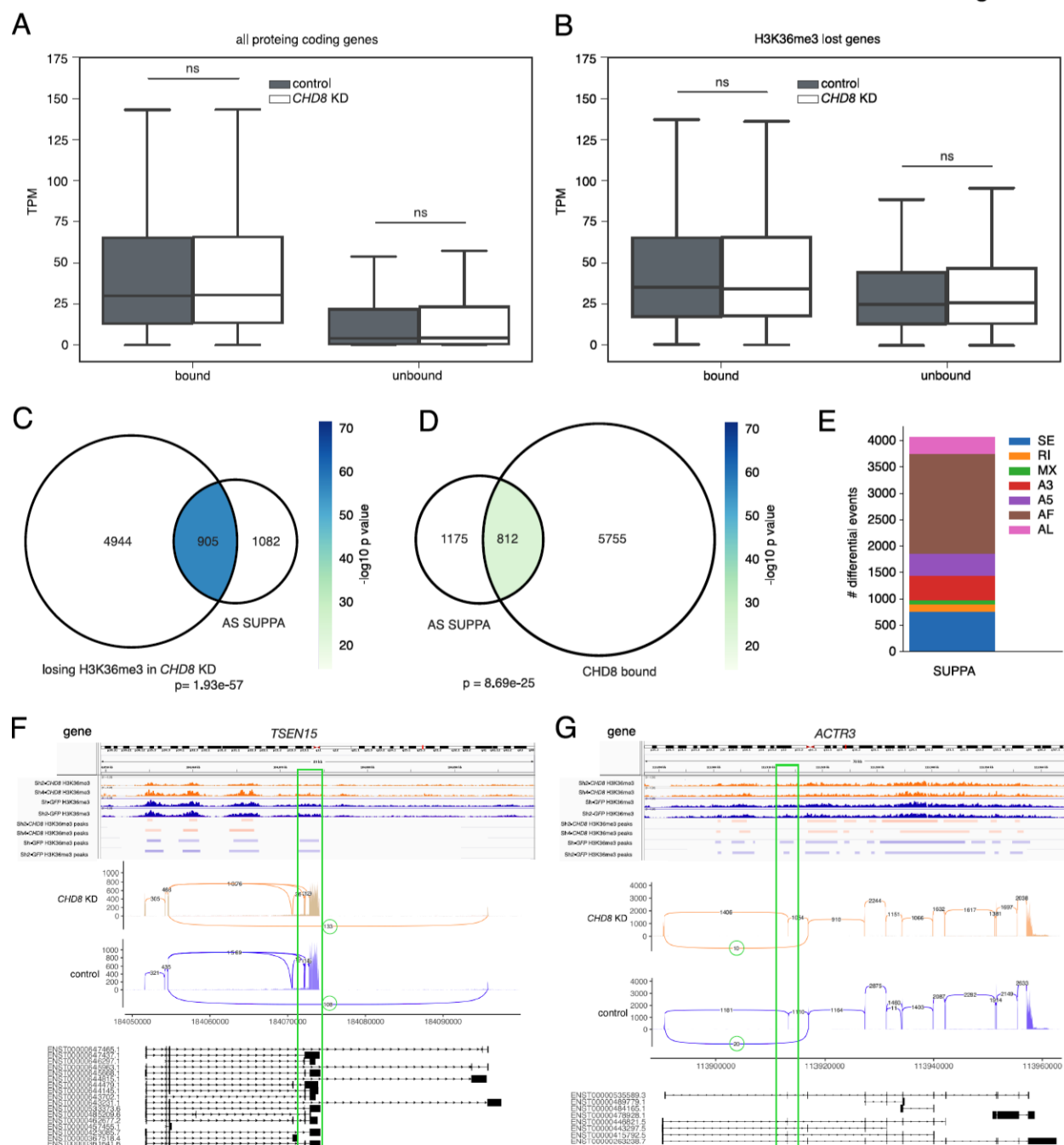


Figure 4 *CHD8-suppression elicited reduction in H3K36me3 correlates with significant alterations in RNA alternative splicing.*

A. B. The box plots show the RNA expression level average of protein coding genes (A) and for genes that lose H3K36me3 peaks following *CHD8* knock-down (B). Control condition is shown in dark grey, *CHD8* knock-down in white; genes bound (left) and unbound by CHD8 (right) are also indicated. TPM: Transcripts Per Kilobase Million. Extreme outliers not shown. ns = not significant, t test statistic.

C. D. Venn diagrams represent the overlap between genes losing H3K36me3 peaks following *CHD8* knock-down (losing H3K36me3 in *CHD8* knock-down) and genes presenting altered alternative splicing events as detected by SUPPA (AS SUPPA) (C), and the overlap between genes bound by CHD8 (CHD8-bound) and genes presenting altered alternative splicing events as detected by SUPPA (AS SUPPA). Number of genes for each condition is indicated. The enrichment significance for each intersection is computed by Fisher's exact test and represented by colors. Color coded legend: $-\log_{10}(\text{p value})$.

E. Stacked bar plot represents the 1987 differential alternative splicing events detected by SUPPA, distributed by event type. SE, skipped event; RI, retained intron; MX, mixed event; A3, alternative 3'; A5, alternative 5'; AF, alternative first exon; AL, alternative last exon.

F. G. Data from two representative genes are shown along with their genomic sites, *TSEN15* (G) and *ACTR3* (F), including H3K36me3 ChIP-seq tracks for *CHD8* knock-down (Sh2-*CHD8* and Sh4-*CHD8* in orange) and control (Sh-*GFP* and Sh-*GFP2* in blue), with the corresponding peaks identified by MACS2 (top). At the bottom, sashimi plots display the alternative last (AL) event for *TSEN15* (G) and skipped exon (SE) for *ACTR3* (F), corresponding to the H3K36me3 peak

lost after *CHD8* suppression (highlighted by green rectangles), in orange for *CHD8* knock-down and in blue for control.