Article

Zeb2 DNA-Binding Sites in Neuroprogenitor Cells Reveal Autoregulation and Affirm Neurodevelopmental Defects, Including in Mowat-Wilson Syndrome

Judith C. Birkhoff [1,†] [iD], Anne L. Korporaal [1], Rutger W. W. Brouwer [2,†] [iD], Karol Nowosad [1,3,4,†] [iD], Claudia Milazzo [1,†], Lidia Mouratidou [1], Mirjam C. G. N. van den Hout [2], Wilfred F. J. van IJcken [1,2] [iD], Danny Huylebroeck [1,5,‡] [iD] and Andrea Conidi [1,*,‡]

1 Department of Cell Biology, Erasmus University Medical Center, 3015 Rotterdam, The Netherlands
2 Center for Biomics-Genomics, Erasmus University Medical Center, 3015 Rotterdam, The Netherlands
3 Department of Biochemistry and Molecular Biology, Medical University of Lublin, 20-093 Lublin, Poland
4 The Postgraduate School of Molecular Medicine, Medical University of Warsaw, 02-091 Warsaw, Poland
5 Department of Development and Regeneration, KU Leuven, 3000 Leuven, Belgium
* Correspondence: a.conidi@erasmusmc.nl; Tel.: +31-10-7043169
† Current address: J.C.B.: Institute of Translational Stem Cell Research, Helmholtz Munich, 85764 Neuherberg, Germany; K.N.: Genome Biology Unit, European Molecular Biology Laboratory, 69117 Heidelberg, Germany; R.W.W.B.: Future Facts, 1013 Amsterdam, The Netherlands; C.M.: Department Neuroscience, Erasmus University Medical Center, 3015 Rotterdam, The Netherlands; L.M.: First Department of Psychiatry, Medical School, National and Kapodistrian University of Athens, 157 72 Athens, Greece, and University Mental Health, Neurosciences and Precision Medicine Research Institute "Costas Stefanis", 115 27 Athens, Greece; A.C.: Cergentis B.V., 3584 Utrecht, The Netherlands, andrea.conidi@cergentis.com.
‡ Shared senior authors.

Abstract: Functional perturbation and action mechanism studies have shown that the transcription factor Zeb2 controls cell fate decisions, differentiation, and/or maturation in multiple cell lineages in embryos and after birth. In cultured embryonic stem cells (ESCs), Zeb2's mRNA/protein upregulation is necessary for the exit from primed pluripotency and for entering general and neural differentiation. We edited mouse ESCs to produce Flag-V5 epitope-tagged Zeb2 protein from one endogenous allele. Using chromatin immunoprecipitation coupled with sequencing (ChIP-seq), we mapped 2432 DNA-binding sites for this tagged Zeb2 in ESC-derived neuroprogenitor cells (NPCs). A new, major binding site maps promoter-proximal to *Zeb2* itself. The homozygous deletion of this site demonstrates that autoregulation of *Zeb2* is necessary to elicit the appropriate Zeb2-dependent effects in ESC-to-NPC differentiation. We have also cross-referenced all the mapped Zeb2 binding sites with previously obtained transcriptome data from Zeb2 perturbations in ESC-derived NPCs, GABAergic interneurons from the ventral forebrain of mouse embryos, and stem/progenitor cells from the post-natal ventricular-subventricular zone (V-SVZ) in mouse forebrain, respectively. Despite the different characteristics of each of these neurogenic systems, we found interesting target gene overlaps. In addition, our study also contributes to explaining developmental disorders, including Mowat-Wilson syndrome caused by *ZEB2* deficiency, and also other monogenic syndromes.

Keywords: chromatin immunoprecipitation sequencing; embryonic stem cells; Mowat-Wilson syndrome; neural differentiation; neurodevelopmental disorder; syndromes; target genes; transcription factor; transcriptomics; Zeb2

1. Introduction

Zeb2 (also named Sip1/Zfhx1b) and Zeb1 (δEF1/Zfhx1a), the two members of the small family of Zeb transcription factors (TFs) in vertebrates, bind to DNA to two separated E-box like sequences, as determined by in vitro binding to double-stranded oligonucleotides and/or Zeb-mediated repression of transfected reporter DNA constructs [1–4].

Spaced bi-partite CACCT sequences, often present as CACCTG E2-box, and sometimes CA-CANNT(G) sequences [1,2], are bound via two (between Zeb1 and Zeb2) highly conserved, separated clusters of zinc fingers [3]. Mutations in *ZEB2* cause Mowat-Wilson Syndrome (MOWS, OMIM#235730) [5–7], a rare congenital disease. MOWS patients display intellectual disability, epilepsy/seizures, typical facial dimorphism, and often Hirschsprung disease (HSCR), as well as multiple other defects [8–10]. Typical are also the delay in developmental milestones such as motoric development, and anomalies of eyes and teeth. Other features include specific craniofacial malformation and sensorineural deafness, which together with HSCR originate from defects in the ZEB2-positive (+) cells of the embryonic neural crest cell lineage. Meanwhile, mutant *ZEB2* alleles have been determined for about 350 patients [11–16]. Recent reports have described malformations in the central nervous system of MOWS patients over a broad age range, including defects of the corpus callosum and/or hippocampus, and can be seen by neuroimaging. These reports also followed up on electro-clinical defects, such as focal seizures, with MOWS patients [16–18].

Zeb2's multiple functions, its action mechanisms and many partner proteins, the still few proven or candidate direct target genes, and lists of genes whose normal expression depends on intact Zeb2 levels, have been documented in various cell types [19–29]. For a recent review, see [30]). Such combinations of studies have allowed for explanations of specific phenotypes caused by Zeb2 perturbation. Zeb2 DNA-binding to candidate target genes helped to explain *Zeb2* loss-of-function phenotypes in embryonic stem cells (ESCs) and initial cells of early and late embryos, later followed by post-natal and adult mice. These intact-Zeb2 dependent and/or Zeb2 target genes are involved in pluripotency (*Nanog*, *Sox2*), cell differentiation (*Id1*, *Smad7*), and embryonic brain cortical and adult neurogenesis (*Ntf3*, *Sox6*), as well as in epithelial-to-mesenchymal transition (EMT) (*Cdh1*) [19–26]. In reverse, subtle mutagenesis of Zeb2 DNA-binding sites in demonstrated target genes also confirmed Zeb DNA-binding, including its repressive activity on mesodermal X*Bra* in *Xenopus* embryos [27] and *Cdh1* in epithelial cells [19].

Despite its critical functions in the precise spatial-temporal regulation of expression of many system/process-specific relevant genes during embryogenesis and post-natal development, and more recently adult tissue homeostasis, stem cell-based repair, and acute and chronic disease [28–30], data from chromatin immunoprecipitation followed by sequencing (ChIP-seq) for Zeb2 has been obtained in very few cases only. A major reason is that ChIP-seq grade antibodies specific for Zeb2 are not readily available. ChIP-seq data have been published for high-Zeb2 hepatocellular carcinoma and leukemia cell lines, or cultured cells that overproduce an epitope-tagged Zeb2 (tag-Zeb2) from cDNA-containing episomal vectors or the safe-harbor *Rosa26* locus [25,31,32]. Neither of these represent normal endogenous Zeb2 levels and dynamics. Furthermore, many anti-Zeb2 antibodies cross-react with Zeb1, so do not discriminate between both proteins when their presence overlaps or succeeds one another in a given cell type. However, these TFs compete for the same target genes, which for the individual proteins in any case, depends on cell identity/state, extrinsic stimulation of the cells, or cellular context (e.g., as demonstrated in somitogenesis [33] and melanoma [34]).

In undifferentiated mouse (m) ESCs, Zeb2 mRNA/protein is undetectable, whereas during neural differentiation (ND) of ESCs its strong upregulation accompanies efficient conversion of naïve ESCs into epiblast stem cell like cells (EpiLSCs). This is essential for the subsequent exit of ESCs from primed pluripotency and the onset of ESC differentiation, including progression to neuroprogenitor cells (NPCs) [25]. Here, we have edited one *Zeb2* allele of mESCs by inserting a Flag-V5 epitope tag just before the Zeb2 stop codon, in-frame with the last exon (ex9 of mouse *Zeb2* [35]). These Zeb2-V5 mESCs were then differentiated into NPCs and the Zeb2 DNA-binding sites were determined by V5-tag-based ChIP-seq. Doing so, we identified 2432 binding sites for Zeb2 in NPCs, of which 2294 map to 1952 protein-encoding genes. We then cross-referenced these ChIP-positive (+) target genes with RNA-sequencing data of differentially expressed genes (DEGs) in cell-type specific, neurodevelopment-relevant Zeb2 perturbations [23,36,37]. Although we compare

non-identical systems, the overall approach still revealed a number of interesting overlaps of target genes, as well as Zeb2's role in regulating critical targets in neurodevelopment, including its own gene promoter. Taken together, for the first time we report the identification of Zeb2's genome-wide binding sites (GWBS) in ESC-derived NPCs at normal Zeb2 level.

2. Materials and Methods

2.1. ESC Culture Conditions and Differentiation

CGR8 (strain 129) wild-type (WT) and Zeb2-Flag-V5+ (in brief, Zeb2-V5) mouse (m) ESCs were cultured and differentiated towards the neural lineage [38], with few modifications). Briefly, the mESCs were cultured on 0.1% gelatin-coated plates in ESC-medium (DMEM supplemented with 15% heat-inactivated (HI) foetal bovine serum (FBS), 2 mM L-Glutamine, $1\times$ Non-Essential Amino Acids (NEAA), 143 µM β-mercapto-ethanol (β-EtSH) (all ThermoFisher Scientific, TFS, Waltham, MA, USA) and leukemia inhibitory factor (LIF) at 10^3 U/mL).

For neural differentiation (ND), 4×10^6 cells were plated on non-adherent 10-cm dishes (Greiner, Kremsmünster, Austria) and allowed to form cellular aggregates (CAs) in 10 mL CA-medium (DMEM, 10% HI-FBS, 2 mM L-Glutamine, $1\times$ NEAA, and 143 µM β-EtSH). From day (D) 4 of ND, cells were grown in CA-medium supplemented with 5 µM retinoic acid (RA). During the aggregation stages of ND, the medium was changed every other day by carefully collecting the aggregates with a 10-mL pipet and transferring them to a 15-mL conical tube. The CAs were allowed to sink to the bottom of the tube where, after the previous medium was carefully discarded, the CAs were then gently resuspended in fresh medium and transferred back to the dishes. At D8 of ND, the aggregates were harvested and dissociated by resuspension in 1 mL Accutase (TFS) and pipetting them up and down using a 1-mL pipet, after shaking them in a 37 °C water bath for 5 min. The Accutase was deactivated by adding 9 mL of fresh N2-medium (DMEM with 2 mM L-Glutamine, 50 µg BSA/mL, and $1\times$ N2-supplement) to the dissociated cells, and pelleting the cells gently for 5 min at 200 g. The cells were then resuspended in fresh N2-medium. To ensure single-cell suspension, the cells were filtered by passing them through a 40-µm nylon cell strainer (Corning, Corning, NY, USA); 2.5×10^5 cells/cm^2 were then plated on poly-DL-ornithine hydrobromide/laminin (both from Sigma-Aldrich, St. Louis, MI, USA) coated plates. Cells were harvested at D8 or D10 of ND.

2.2. Western Blots

To check for Zeb2-V5 protein, the 2BE3-clone ESCs were subjected to ND till D8. Cytoplasmic and nuclear fractions were split using the NePer-kit® (TFS). Protein concentrations were measured using the Bradford BCA (TFS), and equal quantities of protein lysates were loaded on 6% polyacrylamide gels (with SDS) and thereafter cut, according to relative molecular mass. Gels were then transferred onto nitrocellulose membranes (Amersham Bioscience, Amersham, UK), which were incubated overnight with anti-Zeb2 [20] and anti-V5 (Life Technologies, Carlsbad, CA, USA) antibodies, followed by incubation at room temperature with horse radish peroxidase (HRP) conjugated secondary anti-rabbit and anti-mouse antibodies (Jackson ImmunoResearch, West Grove, PA, USA). Protein bands corresponding to Zeb2 or Zeb2-V5 were visualized on an AI-600 digital imager (Amersham Bioscience). As loading control, we used Valosin-containing Protein (VCP) and anti-VCP antibody (Santa Cruz, Dallas, TX, USA, sc-57492, mouse).

2.3. RNA Extraction and RT-qPCR Analysis

Total RNA was extracted from ESCs using TRI Reagent (Sigma), and used for cDNA synthesis with RevertAid RT Kit (from TFS) with oligodT-primers. RT-qPCR was performed using SybrGreen dye (BioRad, Hercules, CA, USA) on a CFX96 T1000 thermal cycler (BioRad). All data shown are averages of three independent biological replicates and three technical replicates, normalized to β-Actin mRNA levels. Primers are listed in Table 1.

Analysis and data visualization was performed in R environment for statistical computing version 3.5.3, implemented with the tidyverse v1.3 package (https://github.com/tidyverse, accessed on 1 November 2019).

Table 1. List of primers used in the study.

Primer Name	Sense/Antisense	Sequence (5′ → 3′)	Application
FlV5mZeb2Ex9_Fwd	Sense	GGCTTACCTGCAGAGCATCA	genotyping
FlV5mZeb2Ex9_Rev	Antisense	CTCCATCTAACTCTGTCTTGGC	genotyping
FlV5_Fwd	Sense	CTACTCGCAGCACATGAATC	genotyping
FlV5_Rev	Antisense	GAGAGGGTTAGGGATAGGC	genotyping
ΔZP_P1_Fwd	Sense	GTCAGTCCGTCCCCAGGTTT	genotyping
ΔZP_P2_Rev	Antisense	GGCATGCTAGCTGGGCTGGT	genotyping
LN249_Fwd	Sense	GGAGCAAACTGAACAAAACCTCGCC	genotyping
LN249_Rev	Antisense	GGCGAGGTTTTGTTCAGTTTGCTCC	genotyping
LN209_Fwd	Sense	AGCGGATCAGATGGCAGTTCGCATG	genotyping
LN209_Rev	Antisense	CATGCGAACTGCCATCTGATCCGCT	genotyping
Zeb2_Fwd	Sense	CAATGCAGCACTTAGGTGTA	qPCR
Zeb2_Rev	Antisense	TTGCCTAGAAACCGTATTGT	qPCR
Zeb2V5_Fwd	Sense	GAAACGATACGGGATGAGGA	qPCR
Zeb2V5_Rev	Antisense	AGGAGAGGGTTAGGGATAGG	qPCR
Nanog_Fwd	Sense	TCTTCCTGGTCCCCACAGTTT	qPCR
Nanog_Rev	Antisense	GCAAGAATAGTTCTCGGGATGAA	qPCR
Pou5f1_Fwd	Sense	AGAGGATCACCTTGGGGTACA	qPCR
Pou5f1_Rev	Antisense	CGAAGCGACAGATGGTGG TC	qPCR
Sox2_Fwd	Sense	GCGGAGTGGAAACTTTTGTCC	qPCR
Sox2_Rev	Antisense	CGGGAAGCGTGTACTTATCCTT	qPCR
Pax6_Fwd	Sense	ACATCTTTTACCCAAGAGCA	qPCR
Pax6_Rev	Antisense	GGCAAACACATCTGGATAAT	qPCR
Acrv1b_Fwd	Sense	CTGCCTACAGACCAACTACACC	qPCR
Acrv1b_Rev	Antisense	CCACGCCATCCAGGTTAAAGA	qPCR
Lhx5_Fwd	Sense	AGAACCGAAGGTCCAAAGAA	qPCR
Lhx5_Rev	Antisense	TCACTTTGGTAGTCTCCGTA	qPCR
Ntng2_Fwd	Sense	CAAGGACTCTACGCTTTTCG	qPCR
Ntng2_Rev	Antisense	AGCACTCGCAGTCTTGAAAT	qPCR
Sema3f_Fwd	Sense	CTACACAGCATCCTCCAAGA	qPCR
Sema3f_Rev	Antisense	ACGGCATTCTTGTTTGCATT	qPCR
Smad1_Fwd	Sense	TACTATGAGCTCAACAACCG	qPCR
Smad1_Rev	Antisense	GAAGCGGTTCTTATTGTTGG	qPCR
Smad3_Fwd	Sense	CACGCAGAACGTGAACACC	qPCR
Smad3_Rev	Antisense	GGCAGTAGATAACGTGAGGGA	qPCR
Sox13_Fwd	Sense	CTTACAGGAGGTTGTGCCA	qPCR
Sox13_Rev	Antisense	TCCTTAGCTTCCACATTGCT	qPCR
Stat3_Fwd	Sense	CAATACCATTGACCTGCCGAT	qPCR

Table 1. *Cont.*

Primer Name	Sense/Antisense	Sequence (5′ → 3′)	Application
Stat3_Rev	Antisense	GAGCGACTCAAACTGCCCT	qPCR
Tcf4_Fwd	Sense	TTGAAGATGTTTTCGCCTCC	qPCR
Tcf4_Rev	Antisense	CCTGCTAGTCATGTGGTCAT	qPCR
Tgfbr2_Fwd	Sense	GAAGGAAAAGAAAAGGGCGG	qPCR
Tgfbr2_Rev	Antisense	TGCTGGTGGTGTATTCTTCC	qPCR
Amylase_Fwd	Sense	GGCTGAGTGTTCTGGGAT	ChIP-qPCR
Amylase_Rev	Antisense	CACGGTGCTCTGGTAGAT	ChIP-qPCR
Cdh1_R1_Fwd	Sense	GCTAGGCTAGGATTCGAACGAC	ChIP-qPCR
Cdh1_R1_Rev	Antisense	TGCAGGGCCCTCAACTT	ChIP-qPCR

2.4. Tag-Zeb2 Mouse ESCs

gRNAs (Table 2) targeting *Zeb2*-ex9, and tracrRNA (Integrated DNA Technologies, IDT, Coralville, IA, USA), were diluted to 125 ng/µL in duplex buffer (from IDT). gRNAs were annealed to tracrRNA at a 1:1 ratio at 95 °C for 5 min and cooling the samples to room temperature. 250 ng of these annealed gRNAs were transfected in 350,000 mESCs together with 2 µg pX459-Cas9-puro vector and 1 µg ssDNA oligo of the donor template containing the FlagV5-tag sequence (Table 2). Transfection was done in a gelatin-coated 6-well plate using DNA: Lipofectamine-2000 (ratio of 1:2). Six hours after transfection the medium was refreshed, and at 24 h the cells were selected in puromycin (2 µg/mL). After two days, the remaining cells were transferred to gelatin-coated 10-cm dishes and given fresh ESC medium (see below). Per dish 1000; 1500; or 2000 cells were plated and allowed to form colonies. The medium was changed every other day. Colonies were picked, expanded, and genotyped by PCR (both outer and inner primer sets were used (Table 1; Figure S1). All candidate clones were validated by Sanger-sequencing; correct clones were expanded and validated by western blot. Mouse ESCs genome-editing was performed under the GGO (genetically modified organisms) institutional licenses 95-053 and 99-164 assigned to the Erasmus University Medical Center.

Table 2. gRNAs and donor template used for CRISPR/Cas9-mediated Zeb2 editing.

Name	Sequence	CRISPR/Cas9
Flag-V5 donor template [1]	aaaatggaaaccaaatcagaccacgaagaagacaatatgg aagatggcatcgaaGACTACAAAGACGATGAC GACAAG*gatatc*GGTAAGCCTATCCCTAA CCCTCTCCTCGGTCTCGATTCTACG**TAA** actactgcattttaagcttcctatttttttttttccagtagtattgtt	in-frame knock-in of Flag-V5 tag
gRNA_ex9_1	GGAAACCAAATCAGACCACGAGG	
gRNA_ΔZP1 gRNA_ΔZP2 gRNA_ΔZP3 gRNA_ΔZP4	CCCGCGCGCGTTTCAATGGGCGC CCCTCGCGAGTGCAACACACCAA GGGCTCGGAGCGCTGCCGATCGG CCGCTGGACCGGGGGGGAGTTGA	*Zeb2 ΔP* deletion

[1] Flag-V5 donor template (from top left to bottom right): in lowercase: homology arms located in exon9 and 3′ UTR of *Zeb2*, respectively in underlined lowercase: mutated PAM sequences; in uppercase: Flag-encoding sequence; in lowercase italics and bold: *Eco*RI restriction site; in underlined uppercase: V5-encoding sequence; in bold uppercase: STOP codon.

2.5. CRISPR/Cas9-Mediated Deletion of the Zeb2 Binding Site Located at chr2:45109746-45110421

Oligonucleotides for gRNAs (Table 2) with target outside of this chr2-region were cloned into *Bbs*I-digested pX330-hspCas9-T2A-eGFP plasmid. All resulting plasmids

used further were sequenced. 4 μg of gRNA-plasmids (1 μg each) were transfected in 350,000 mESCs and selected (see above). After 24 h these cells were sorted as GFP+ cells (LSR Fortessa, Becton-Dickinson (BD), Franklin Lakes, NJ, USA). Per well of a 6-well plate 1000; 1500; or 2000 GFP+ cells were plated and colonies were allowed to form, picked (see above), and genotyped by PCR using primers flanking this deletion, and within and outside of it. Clones showing a possible heterozygous or homozygous deletion, as concluded from the PCR analysis, were subjected to ND. At D8, they were harvested, RNA was isolated and cDNA synthesized (see below), and amplified (for the primers, see Table 1). All candidate clones were validated by Sanger-sequencing.

2.6. Chromatin Immunoprecipitation (ChIP)

ChIP 2×10^8 cells were harvested at ND-D8 in 10 mL of PBS and cross-linked using 1% formaldehyde (Sigma Aldrich) for 15 min, rotating at room temperature. Quenching followed with 125 mM glycine for 5 min, again rotating at room temperature. Cross-linked cells were washed twice with ice-cold PBS (5 min; 1500 rpm (240 rcf), 4 °C) before the pelleted cells were snap-frozen and stored at −80 °C. For sonication, the cell pellets were thawed on ice, resuspended in 1 mL sonication buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 0.5 mM EGTA) supplemented with protease and phosphatase inhibitors (PPI, from Roche, Basel, Switzerland), and incubated on ice for 10 min. DNA was sheared by sonicating the cells using a probe sonicator (32 cycles, 30 sec-on amplitude 9, and 30 sec-off). These samples were centrifuged at 13,200 rpm (17,000 rcf) for 10 min at 4 °C. Chromatin pellets were snap-frozen and stored at −80 °C. To check sonication efficiency, 50 μL of sample was de-crosslinked overnight by adding NaCl (final concentration 5 mM) at 65 °C and constantly shaken (950 rpm, ThermoMixer C, Eppendorf, Hamburg, Germany). The next morning 5, 10, and 20 μL of sample were loaded on a 2% agarose gel, revealing ideally a DNA-smear around 300 bp. A 50-μL sample was used as a control input.

For immunoprecipitation, the chromatin of 10^7 cells was diluted in ChIP-dilution buffer (17 mM Tris-HCl pH 8.0, 170 mM NaCl, 1.2 mM EDTA, 0.01% SDS, 1.1% Triton X-100, with $1\times$ PPI) to a final volume of 1 mL. Samples were pre-cleared by adding pre-washed Protein A/G agarose beads (Santa Cruz) and further incubation for 1 h, rotating at 4 °C. Then, samples were centrifuged for 1 min (1000 rpm; 106 rcf) at 4 °C, and the pre-cleared chromatin (supernatant) was transferred to a new low-binding 1.5-mL tube and incubated with 50 μL of pelleted V5-Agarose beads (Sigma Aldrich), rotating at 4 °C overnight. Before the addition of V5-agarose beads, the beads were washed 5 times (5 min each) in PBS by rotating them. As a negative control, half of the sample was incubated with Protein A/G beads (Santa Cruz, sc-2003).

The following day the beads were pelleted (1000 rpm; 1 min) and washed as follows: once with lower-salt buffer (20 mM Tris-HCl pH 8.0, 150 mM NaCl, 2 mM EDTA, 0.1% SDS, 1% Triton X-100), transferred to non-stick low-binding 1.5-mL tubes and then washed once with high-salt buffer (i.e., lower-salt buffer, but now 500 mM NaCl), once washed with LiCl buffer (which is 10 mM Tris-HCl pH 8.0, 250 mM LiCl, 1 mM EDTA, 1% NP-40, 1% sodium deoxycholate (DOC), and twice washed with 10 mM Tris-HCl pH 8.0, 1 mM EDTA (each incubation for 5 min, rotating at 4 °C, followed by gently spinning down.

The protein-chromatin was then eluted from the beads by adding 250 μL of elution buffer (1% SDS, 100 mM NaHCO$_3$), rotating for 1 h at room temperature twice, and combining the eluates from both steps. To the input sample, 450 μL of elution buffer was also added, and all samples were de-crosslinked through the addition of 5 mM NaCl at 65 °C overnight, shaking at 950 rpm. The day after, 2 μL of proteinase-K (from 10 mg/mL stock), 20 mM (final concentration) Tris-HCl pH 6.5, 5 mM (final concentration) EDTA pH 8.0 and 10 mg/mL RNase-A (Sigma-Aldrich) were added to each sample and incubated for 1 h at 45°C while shaking (700 rpm). DNA was extracted from the samples using the PCI method and diluted in water. Five independent ChIPs were performed, for a total of 10^8 mESCs used per condition, and pulled-down chromatin was pooled. ChIP efficiency

was assessed by qPCR using primers amplifying *Cdh1* promoter sequences bound by Zeb2 [25]. All primers used are listed in Table 1.

2.7. ChIP-Sequencing

DNA libraries from input (i.e., control) and V5 ChIPs were prepared using ThruPLEX DNA protocol (TakaraBio, Kusatsu, Shiga, Japan) specific for low amounts of DNA and sequenced on Illumina HiSeq-2500, and single reads of 50 bp were generated. Adapter sequences were trimmed from the $3'$-end of the reads, after which the reads were aligned to the mm10/GRCm38 genome using HISAT2 [39]. From the alignments, secondary or supplementary, low-quality, and fragmented alignments (fragments longer than 150 bp) were filtered away. Peaks were called with MACS [40], and coverage was determined. 42 and 25 million reads were generated for input and V5 ChIP, respectively.

2.8. ChIP-Sequencing Data Analysis

Peak calling was performed with MACS2 (Galaxy version 2.1.1.20160309.6) [40,41], with default parameters (narrow peak calling, Mm1.87e9, FDR < 0.05) using the input sample as background. The No model parameter was used, and the extension size was set at 210 bp based on the predicted fragment lengths from the alignments (MACS2 predict-tool, Galaxy version 2.1.1.20160309.1) [40,41]. The distance of the aligned reads from the TSS of the gene was analyzed using ComputeMatrix (Galaxy version 3.3.2.0.0) and PlotHeatmap Galaxy version 3.3.2.0.1; the used matrix is based on the log2ratio of the aligned ChIP peaks over the input, calculated using BamCompare (Galaxy version 3.3.2.0.0) [42].

2.9. Transcription Factors Motif Enrichment Analysis

To identify the transcription factor binding sites (TFBS) in Zeb2-binding regions associated with DEGs, we first extracted unique Zeb2-peaks located 10 kb $-/+$ from the transcription start site (TSS). Next, we analyzed the TFBS enrichment using a UniBind enrichment tool with motifs from the UniBind database [43] (using reference genome GRCm38/mm10). As a background for the analysis, all Zeb2-peaks were used. The *p*-value from Fisher's exact test after multitest adjustments was used to identify significantly enriched TFBS. Further, the max rank index calculated based on the odds ratio, *p*-value from Fisher's exact test, and the number of overlapping regions, was applied to rank the top enriched motifs.

2.10. RNA-Sequencing

The quality of total RNA (of biologically independent triplicates) of wild-type mESCs at D0, and at ND D4, D6, and D8, was checked on Agilent Technologies-2100 Bioanalyzer, using an RNA nano-assay. All samples had RIN values of 9.8 or higher. Triplicate RNA-seq libraries were prepared (Illumina, San Diego, CA, USA TruSeq stranded mRNA protocol; www.illumina.com, accessed on 1 January 2020). Briefly, 200 ng of total RNA was purified using polyT-oligo-attached magnetic beads for ending with polyA-RNA. The polyA-tailed RNA was fragmented, and cDNA synthesized (SuperScript II, Invitrogen, Waltham, MA, USA, random primers, in the presence of Actinomycin D). cDNA fragments were end-repaired, purified (AMPureXP beads), and A-tailed using Klenow exo-enzyme and dATP. Paired-end adapters with dual index (Illumina) were ligated to the A-tailed cDNA fragments and purified (AMPureXP beads).

The resulting adapter-modified cDNAs were enriched by PCR (using Phusion polymerase) as follows: 30 s at 98 °C, 15 cycles of (10 s at 98 °C, 30 s at 60 °C, 30 s at 72 °C), 5 min at 72 °C. PCR products were purified (AMPureXP beads) and eluted in 30 μL resuspension buffer. One μL was loaded on an Agilent 2100 Bioanalyzer using a DNA-1000 assay to determine the concentration and for a quality check. Cluster generation was performed according to the Illumina TruSeq SR Rapid Cluster kit v2 Reagents Preparation Guide (www.illumina.com, accessed on 1 January 2020). After the hybridization of the sequencing primer, sequencing-by-synthesis was performed using a HiSeq-2500 with a single-read

50-cycle protocol followed by dual index sequencing. Illumina adapter sequences have been trimmed off the reads, which were subsequently mapped against the GRCm38 mouse reference (using HiSat2 version 2.1.0) [39]. Gene expression values were called using HTSeq-count version 0.9.1 [44] and Ensembl released 84 gene and transcript annotation. Sample QC and DEG analysis have been performed in the R environment for statistical computing (version 3.5.3, using DESeq2 version 1.22.1 [45] and Tidyverse version 1.2.1 (https://github.com/tidyverse;https://www.r-project.org/ from R Core Team, accessed on 1 January 2019).

2.11. Meta-Analysis, Pathways Enrichment, Gene Ontology, Function Analysis, and Gene to Disease Association

RNA-seq datasets (as DEG tables, from [23,36], were downloaded from GEO (https://www.ncbi.nlm.nih.gov/geo/, GSE35616, and GSE103003, respectively, accessed on 1 March 2019). Cross-referencing and visualization were performed in R using Tidyverse, VennDiagram, and pheatmap packages. The remaining analyses were performed with the StringDB package for R [46], while for Gene-to-Disease association Disgenet2R for R was used [47].

2.12. Zeb2 Short Hairpin (sh) RNA-Mediated Knock-Down

Zeb2 knock-down (KD) was carried out by transfecting *Zeb2*-shRNAs into ESCs, at ND-D8. For this, the CAs were dissociated (see above), and single-cell suspensions were transfected using Amaxa Nucleofector II (using kit V, program A-33). Table 3 lists the shRNAs used in this study; the *Zeb2* target sequence is indicated in bold. In total 4 μg of shRNA was used for the transfection of 4.5×10^6 cells. After transfection, the cells were plated in 5 mL of N2-medium on a poly-ornithine/laminin-coated 6-cm cell culture dish. Two hours post-transfection, the medium was refreshed, and 24 h after transfection was changed to N2-medium in the presence of puromycin (see above) for 48 h. The cells were then harvested, and KD efficiencies were examined using RT-qPCR. As a control, scrambled shRNA was used.

Table 3. shRNAs used.

Name	Sequence
shZeb2_1	CCGG**CCGAATGAGAAACAATATCAA**CTCGAGTTGATATTGTTTCTCATTCGGTTTTTG
shZeb2_2	CCGG**CCTCAGGAATTTGTGAAGGAA**CTCGAGTTCCTTCACAAATTCCTGAGGTTTTTG
shZeb2_3	CCGG**CCAGTGTCAGATTTGTAAGAA**CTCGAGTTCTTACAAATCTGACACTGGTTTTTG
shZeb2_4	CCGG**CCCATTTAGTGCCAAGCCTTT**CTCGAGAAAGGCTTGGCACTAAATGGGTTTTTG
shCTRL	CCGGCAACAAGATGAAGAGCACCAACTCGAGTTGGTGCTCTTCATCTTGTTGTTTTT

3. Results

3.1. Heterozygous Zeb2-V5 ESCs Differentiate as Wild-Type Cells

The addition of short epitope(s) at the N- or C-terminus, as well as activation/repression domains of heterologous transcription factors (TFs) at the Zeb2 C-terminus was previously shown not to interfere with Zeb2's DNA-binding (as tested in *Xenopus* embryos [48], heterologous cells [49], mouse forebrain [36] and mESCs [25]). Here, we have used a CRISPR/Cas9 approach (see Materials and Methods) to insert an in-frame Flag-V5-tag encoding sequence in *Zeb2*-ex9 of mESCs (ESC clone 2BE3; Figure S1). Allele-specific RT-qPCR, using primers that amplify sequences between the ex9 and the V5-tag, showed mRNA expression from the tagged allele in ESC culture at day (D) 0, 4, 6, and 8 of ND as compared to the parental wild-type (WT) mESC line (Figure 1A).
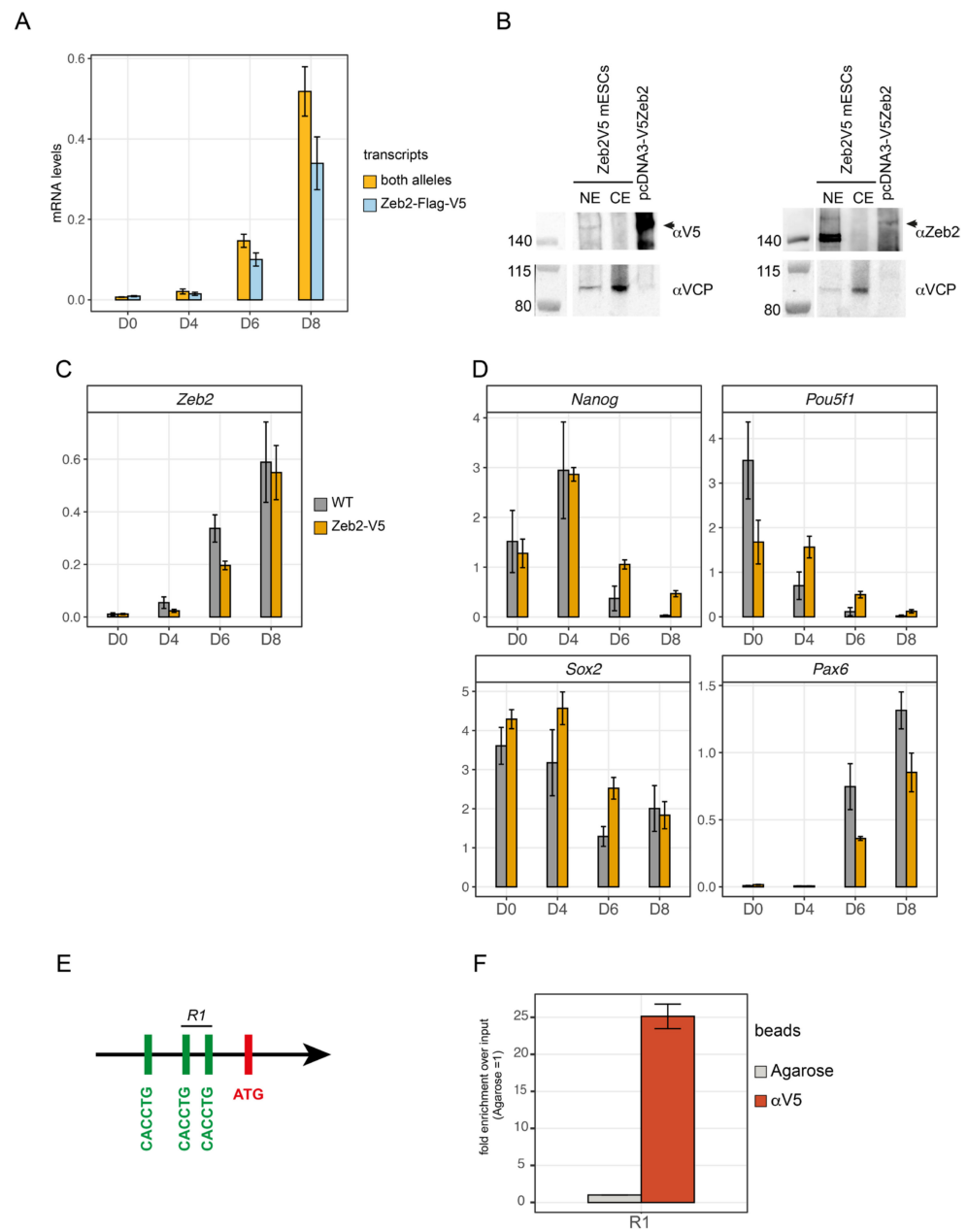
Figure 1. Characterization of the heterozygous Zeb2-Flag-V5 mESC line and ChIP-qPCR validation. (**A**) Allele-specific RT-qPCR using two sets of primers located either in exon7 (and therefore able to detect the whole Zeb2 mRNA produced by both alleles; orange bar) or located in exon9 and the V5-tag (thus recognizing specifically the knocked-in tagged allele; light blue bar); (**B**) Western blot analysis showing V5 epitope containing Zeb2 in ESC-derived NPCs (at D8 of neural differentiation, ND) in nuclear extracts (NE) and cytoplasmic extracts (CE). Membranes were blotted with anti-V5 antibody (left panel, αV5) or anti-Zeb2 antibody (right panel, αZeb2, [20]). As a control, a fraction of Zeb2-rich extract obtained from HeLa cells transfected with a pcDNA3-V5Zeb2 vector was also separated in the same gel; (**C**) Zeb2 mRNA levels in wild-type (WT, grey bar) and Zeb2-V5 (orange bar, clone 2BE3, indicated as Zeb2-V5) mESCs during ND, as determined by RT-qPCR; (**D**) RT-qPCR for ND, using marker genes: pluripotency marker genes *Nanog*, *Pou5f1* (*Oct4*) and *Sox2* are downregulated in Zeb2-V5 mESCs similarly to WT. The neuronal marker gene *Pax6* is also significantly upregulated during differentiation, such as in WT mESCs; (**E**) Scheme of the mouse *Cdh1* promoter showing the three E-boxes located upstream of the ATG start codon. Zeb2 binds specifically to only two of these, indicated as R1 [25]; (**F**) ChIP-qPCR showing enrichment for Zeb2-V5 binding to the R1 region of the *Cdh1* promoter. Agarose beads were used as negative control (in grey).

Western blot analysis in nuclear extracts of ND-ESCs at D8 (thus NPCs, [25]) confirmed the presence of Zeb2 of expected molecular mass, using either anti-V5 (αV5) or anti-Zeb2 antibodies (Figure 1B). Both the Zeb2-V5 and wild-type (WT) ESCs were then also verified during ND differentiation for temporal expression of *Zeb2*, core pluripotency genes (*Pou5f1*, *Nanog*, both downregulated upon ND, and *Sox2*, also an NPC TF) and an acknowledged NPC marker (*Pax6)* (Figure 1C,D).

The untagged and tagged Zeb2 ESC lines displayed comparable expression dynamics of *Zeb2*, indicating that Zeb2-V5 NPCs at D8 of ND can be used for chromatin immunoprecipitation sequencing (ChIP-seq). Further confirmation came from the selective pull-down of Zeb2 on the known target *Cdh1*, using ChIP-qPCR. Zeb2 binds to two of three E-boxes in the mouse *Cdh1* promoter (Figure 1E), which it represses during epithelial-to-mesenchymal transition (EMT) [19,25]. A ±25-fold enrichment for Zeb2-V5 was obtained when probing this *Cdh1* region using anti-V5 antibody (αV5) conjugated beads compared to agarose beads as negative control (Figure 1F). Hence, Zeb2-V5 binds to known Zeb2 target sites, and the NPCs are suitable for endogenous mapping of the Zeb2 genome-wide binding sites (GWBS).

3.2. One-Third of 2432 Zeb2 DNA-Binding Sites Map Close to the Transcription Start Site of System-Relevant Expressed and Protein-Encoding Genes, Including the Zeb2 Gene Itself

αV5-precipitated samples from upscaled Zeb2-V5 NPCs were used for ChIP-seq (see materials and methods), followed by analysis with Galaxy Software [50]. Of the 2432 total significant peaks, 2294 peaks (94% of total) mapped to 1952 loci that encode proteins, while 125 peaks (5% of total) mapped to micro-RNA (miRNA) genes, and 1% to regions that lack annotation (NA, using ENSEMBL-GRCm38.99; Figure 2A; Table S1).

About 37.5% of all binding sites of Zeb2-V5 are located within −10/+10 kb of annotated transcription start sites (TSS, Figure 2B). Gene ontology (GO) pathway enrichment analysis of the aforementioned 1952 loci revealed binding of Zeb2 to classes of genes annotated to signaling by Wnt, integrin, chemokine/cytokine (predominantly as defined in inflammation) and cadherin, respectively, as well as to developmental signaling by EGF, VEGF, TGFβ, and FGF family pathways (Figure 2C). Among these 1952 loci, those for genes encoding transcription regulatory proteins, and post-translational modification as well as metabolic enzymes, are well-represented (Figure 2D).

In parallel, we applied bulk temporal RNA-seq of WT mESCs at D0 (undifferentiated), D4 (induction of ND), D6 (early NPCs), and D8 (NPCs) and checked the expression dynamics of the 1952 Zeb2-bound genes (from the D8 ChIP-seq sample). Among these, 1244 changed in steady-state transcription levels between D4-8 as compared to D0 (Figure 2E; log2FoldChange < −0.5 or >0.5 and *p*-value < 0.05; low-stringency analysis was opted to assess also small differences in mRNA of Zeb2-bound genes). Further, 335 of these genes, including Zeb2 itself, are commonly expressed between D4-6-8, but at different levels (for lists of all DEGs, see Table S2). Figure S2A depicts the D4, D6, and D8 transcriptomes of ND-mESCs, each compared to D0, with an indication of whether the genes are bound or not by Zeb2, as determined by our Zeb2-V5 ChIP-seq. At each of these respective time points, hence at different Zeb2 mRNA levels, about 11–14% of the up-/down-regulated DEGs are bound by Zeb2 (Figure S2B).
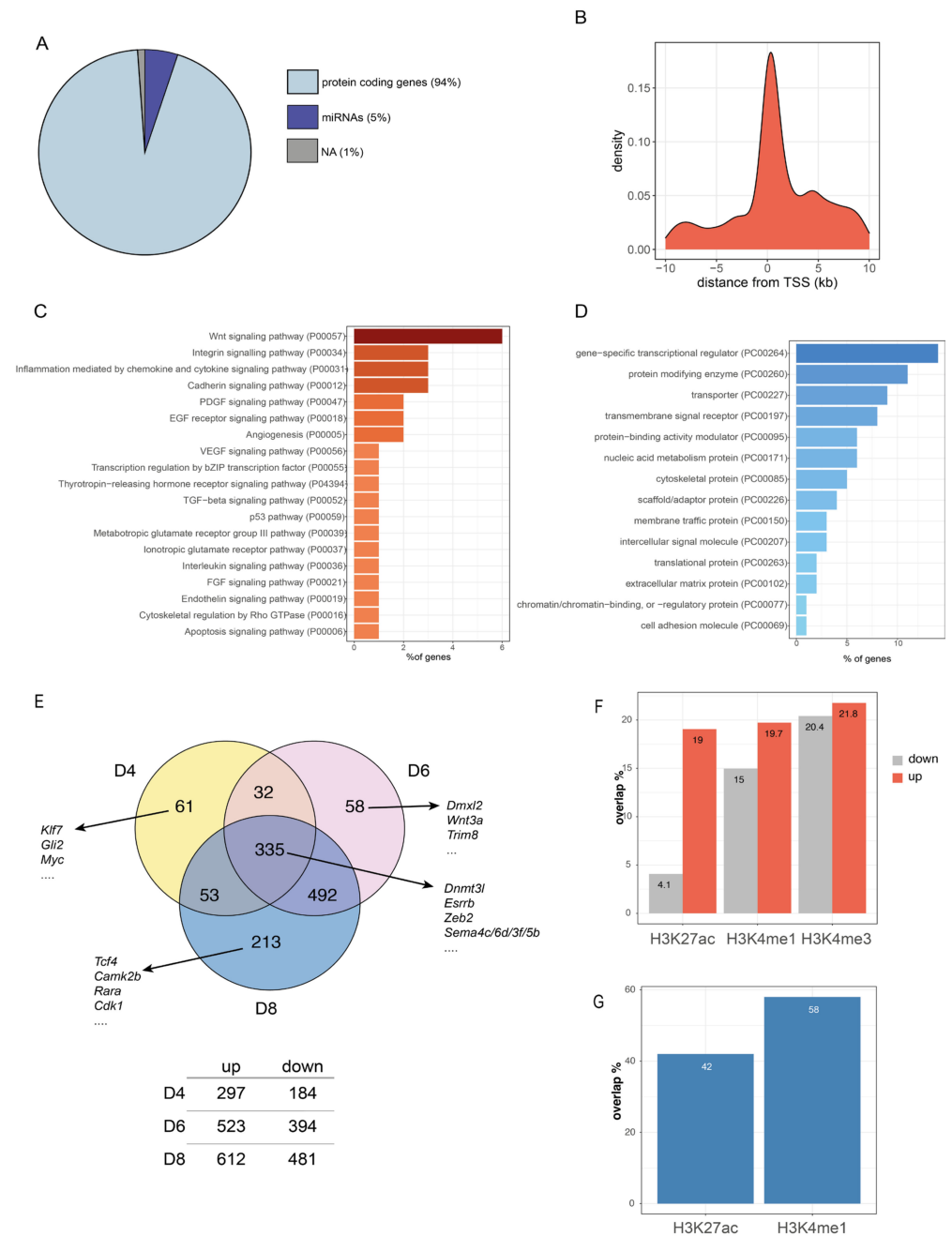
Figure 2. Zeb2-V5 protein is recruited at the TSS of transcriptional regulator encoding genes, predominantly those classified in Wnt signaling. (**A**) 2432 peaks were selected from our ChIP-seq data set (see Section 2). Of these, 94% are associated with protein-coding loci, 5% with miRNAs, and the remaining 1% map to regions without functional annotation (NA); (**B**) Frequency plot showing the binding of Zeb2-V5 at and around (−10 to +10 kb) the TSS; (**C**) The 2294 peaks map to 1952 protein-encoding genes, many of which operate in Wnt signaling (Table S1) or are (as shown in panel (**D**)) transcriptional regulators; (**E**) Of the 1952 protein-encoding genes, 1244 are differentially expressed during ND when compared to the undifferentiated state (D0). Of these 1244 genes, 335 are differentially expressed at all three time points of ND. We also list a few examples of DEGs uniquely expressed at one time point, as well as these that are shared among three time points; a full list is provided in Table S2; (**F**) Overlap of the Zeb2-bound regions with H3K27ac, H3K4me1, and H3K4me3 histone marks in the −10/+10 kb from the TSS of up or downregulated genes at D8 of mESCs differentiation; (**G**) Overlap of the Zeb2-bound regions outside the −10/+10 kb region from the TSS with histone marks.

Among the Zeb2-bound genes that are normally down-regulated, *Dnmt3l* and *Esrrb* are present, suggesting that upregulation of *Zeb2* in ND-ESCs (D6 and D8) directly causes downregulation of these two genes accordingly (Figure S2C; Table S2). Zeb2 has been suggested as a direct repressor of *Dnmt3l* and *Esrrb*, facilitating the switch from self-renewal of ESCs to their exit from pluripotency, and promoting differentiation, since expression levels of all *Dnmt3* genes remained higher in *Zeb2*-knockout (KO) ND-ESCs [25]. However, these *Zeb2*-KO cells also convert very inefficiently into EpiLSCs and fail to exit from primed pluripotency. Importantly, among the Zeb2-binding genes whose mRNA levels increased during ND, *Zeb2* itself is also present (yellow dot, Figure S2C), indicating autoregulation. In fact, in this ND model, the highest recruitment of Zeb2-V5 in ChIP-seq data was mapped upstream of the TSS of Zeb2 (Table S1).

Out of the 1244 Zeb2-bound genes that significantly changed steady-state mRNA levels in our D4 to D8 transcriptome data sets, 213 are exclusive DEGs in D8 NPCs (Figures 2E and S2D). Among these, *Tcf4* is bound by Zeb2 and increases in expression in NPCs (Figures 2E and S2D). Tcf4 is a ubiquitous basic helix-loop-helix (bHLH) type TF that binds to E-boxes; its many isoforms [51,52] cooperate with cell-type specific bHLH TFs in heterodimers, which are active during CNS development [53,54] (for a review, see [55]). In oligodendrocyte precursors (OPCs), Tcf4 is essential for their subsequent differentiation. It dimerizes with the lineage-specific bHLH-TF Olig2, further promoting their differentiation and maturation [56], while Zeb2, together with upstream Olig1/2, is essential for myelinogenesis in the embryonic CNS [21]. Here, Zeb2 generates anti-BMP(-Smad)/anti-Wnt(-β-catenin) activities, which is crucial for CNS myelinogenesis by differentiation of OPCs. The regulatory action of Zeb2 on the *Tcf4* target gene, as found in mouse cells by our ChIP-seq, may underpin phenotypic similarities between MOWS and Pitt-Hopkins syndrome patients (PTHS, OMIM #610954; for a recent discussion, see [57]) the latter caused by mutations in *TCF4* [58], making us speculate that TCF4 may be deregulated in neural cells in MOWS.

3.3. Zeb2 Peaks Overlap with Active Enhancers and Promoters

To assess whether Zeb2-peaks are present in the regulatory regions of up or down-regulated genes from our transcriptome data, we cross-referenced the coordinates of the Zeb2 broad peaks within −10/+10 kb from the TSS with the mouse ChIP-seq datasets available in ENCODE for nervous systems (cerebellum, cortical plate, olfactory bulb, forebrain, midbrain, hindbrain, neural tube, and olfactory bulb, respectively). We found that, of these datasets from histone ChIP-seq available in ENCODE, the respective H3K27ac, H3K4me1, and H3K4me3 marks were overlapping with our ChIP-seq data (Figure 2F).

The H3K27ac signature strongly overlaps with the Zeb2 peaks in genes upregulated at D8 (19% in upregulated genes vs. 4% in downregulated genes). For H3K4me1 and H4K4me3 marks, no big difference in overlap between up and down-regulated genes was observed. While the H3K27ac mark is associated with active enhancers, H3K4me1 is associated with primed enhancers, and H3K4me3 is considered a "promoter" marker [59]. Taken together, these data suggest an activating role for Zeb2 here. Outside the −10/+10kb considered range, about 48% of the identified peaks had a 42% overlap with H3K27Ac and a 58% overlap with H3K4me1 histone marks (Figure 2G).

We then did motif enrichment analysis using UniBind (https://unibind.uio.no, 1 May 2022) for TFs that could bind the Zeb2-bound peaks or could do so in proximity to up or down-regulated genes at D8 of mESCs differentiation (Figure S3A,B). In those peaks close to the TSS of upregulated genes, binding motifs for Sox2, Gata2, and Tcf3 are very abundant. These TFs are known to function during NPC or ND. For example, Sox2 is an acknowledged marker for neurogenesis [60]. It has been demonstrated that Zeb2 is needed to elicit anti-Sox2 activities in (re)myelination by adult Schwann cells in the PNS, needed for normal progression of commitment, differentiation, and maturation in this glial cell lineage [21,24,61]. Tcf3 (also known as E2A) plays a role in stem cell self-renewal [62], but is also important during neural fate commitment and possibly repressing Nodal signaling

during ND [63]. Gata2 has been associated with negative regulation of proliferation in NPCs and, as a result, further differentiation of these cells [64]. However, how Zeb2 acts upon or together with these TFs during NPC differentiation is not fully known yet.

In the peaks close to the TSS of downregulated genes there is a prevalence for CTCF, Fos, Myc, and Stat5a binding. CTCF is of interest because it acts as a link between 3D genome architecture and gene expression regulation. During NPC differentiation however, it was observed that 40% of the NPC-specific DNA loops were not CTCF-dependent, whereas, in other cell-state specific loops, this was only 10%, indicating a less important role for CTCF in the regulation of NPC differentiation compared to other cell lineages [65]. This might indicate an interesting role for Zeb2 in binding and possibly regulating CTCF mRNA levels during NPC differentiation, supporting the subsequent activation of NPC-specific genes arising from e.g., repressing CTCF. Also here, more studies are required to get more insights into the cooperativity or counteracting actions of Zeb2 with candidate TFs in the regulation of the candidate Zeb2 targets.

3.4. Meta-Analysis of Identified Binding Sites and Perturbed-Zeb2 RNA-Seq Data Reveal Overlapping Zeb2 Target Genes

We performed a meta-analysis of three published transcriptome data sets from control and *Zeb2*-KO mice: sorted E14.5 mouse ventral forebrain interneurons (Nkx2.1-Cre driven *Zeb2*-KO; [36]) and sorted (at P2) progenitors of the ventricular-subventricular zone (V-SVZ), an adult neurogenic niche in the forebrain (Gsh2-Cre; [23,36]). In addition, we used high-throughput RT-qPCR data generated on a Fluidigm platform and obtained after esiRNA-based knockdown (KD) of Zeb2, as part of a systems-biology study in ND-mESCs [37] (with the Zeb2 KD data subset kindly provided by R. Dries, Boston University). From these respective datasets, the DEGs upon the Zeb2 perturbations (p-value < 0.05; log2FoldChange < −1 and >1) were filtered. This identified 108 genes in total, and these depend on normal Zeb2 levels for their (i) downregulation/repression (if directly by Zeb2, as a repressor) or (ii) other genes that depend on Zeb2 for their upregulation/activation (if directly by Zeb2, as an activator) (for Zeb2 as dual TF [29,30,66]. In parallel, the 2294 Zeb2-V5 sites mapping to the 1952 protein-encoding genes were filtered from the complete ChIP-seq dataset and then used as references for the RNA data sets (Figure 3A).

Thus, this cross-referencing identified 108 protein-encoding genes among the three transcriptomic data sets and the ChIP-seq data set (Figure 3A, Table S3). Figure S4 shows a heatmap of the changes in mRNA levels of these 108 genes during ND of wild-type ESCs and their correlation with the analyzed datasets. Noteworthy, only *Cxcr4* was common to all RNA data sets. This is likely due to the fact that two RNA-seq sets are generated in different brain/neuron cell-type in vivo mouse models, while the other steady-state RNA level data documented the effects of Zeb2-KD on mRNA levels of (only 96 in total) TGFβ/BMP-system components [37], so the timing does not completely overlap with our ChIP-seq dataset. However, Cxcr4 and its ligand Cxcl12/Sdf-1 are crucial for migration of interneurons from the ventral forebrain to the neocortex [67,68], processes co-controlled by Zeb2 as shown in cell-type specific KO mice [36]. Furthermore, the identified 108 genes are involved in the regulation of stem cell pluripotency, signaling by TGFβ, FoxO, and Hippo, and in axon guidance. Taken together these data further confirm the pivotal role of Zeb2 in these processes. Further Gene Ontology (GO) analysis reveals that these 108 genes cluster as important regulators of developmental processes, cell locomotion, and signaling (Figure 3C). These processes are affected in human conditions following *ZEB2* heterozygosity, as in the case of MOWS.
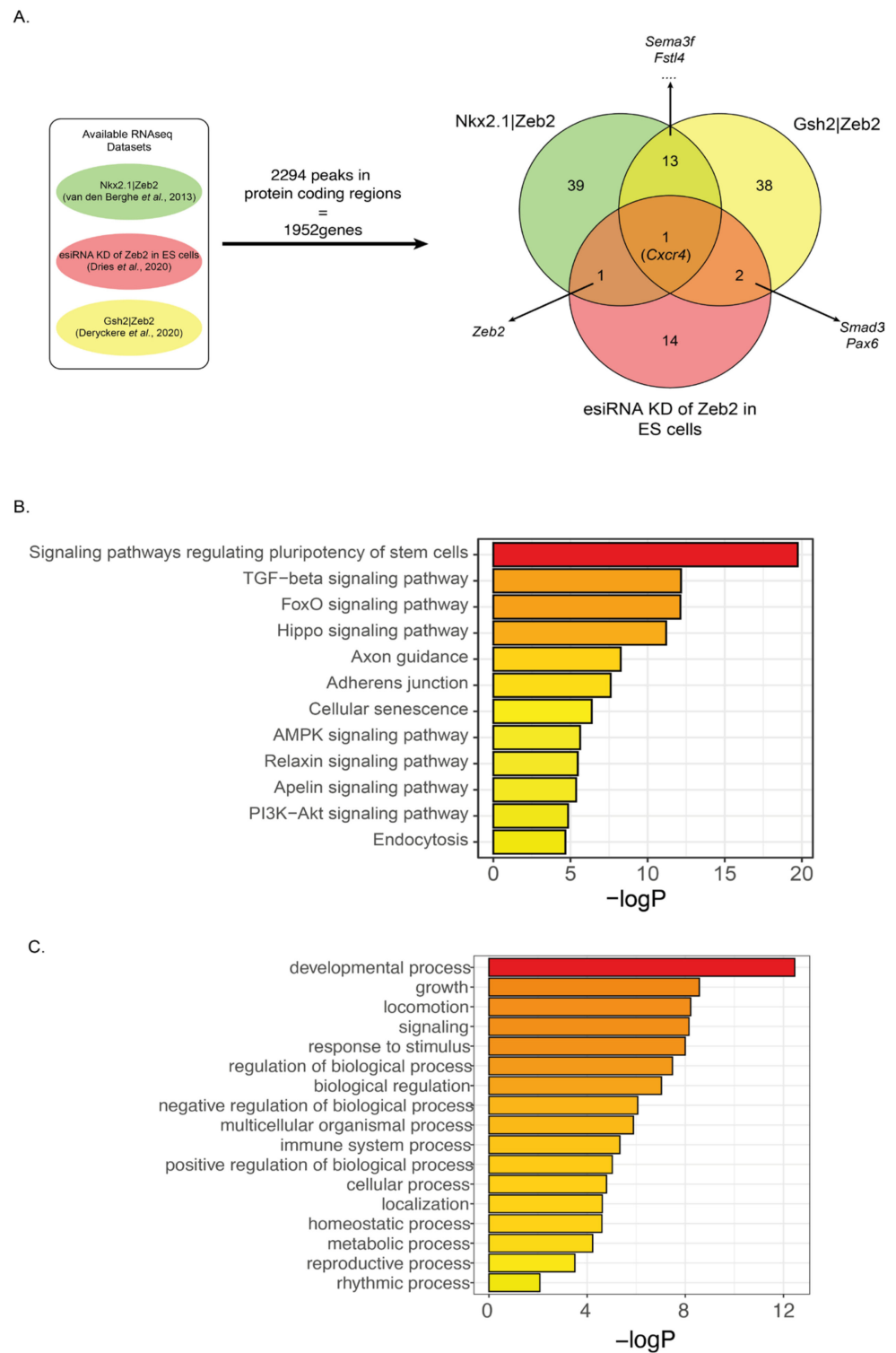
Figure 3. Schematic representation of the meta-analysis of Zeb2−bound genes versus RNA−seq datasets. (**A**) 108 genes bound by Zeb2 are also differentially expressed in the three data sets from other studies in mouse models and ESCs (see main text for details, [23,36,37]). *Cxcr4* is the only DEG bound by Zeb2 and common among the three data sets (for discussion, see main text); (**B**) These 108 genes mainly map to signaling pathways regulating stem cell pluripotency, and effects of TGFβ family, FoxO, and Hippo signaling/activity; (**C**) The 108 genes cluster as important regulators of developmental processes, locomotion, and signaling.

3.5. Zeb2 Directly Controls TGFβ/BMP-System Component and Neuronal Differentiation/Migration Genes

We then validated 14 out of the 108 cross-referenced target genes, selected based on either being already known as a target of Zeb2 (*Nanog*, [25]), or as TGFβ/BMP-system component (*Bmp7*, *Tgfbr2*, *Smad1*, *Smad2*, *Smad3*, *Id2*, *Cited2*), or having a crucial role in neurogenesis and neuronal maturation (*Sema3f*, *Cxcr4*, *Lhx5*, *Ntng2*, *Pax6*, *Tcf4*; their mRNA levels in ND-ESCs are highlighted in the heatmap in Figure S4). Because *Zeb2*-KO ESCs do not exit from primed pluripotency and thus cannot differentiate [25], we validated our findings using shRNA-mediated Zeb2-KD at ND-D8 and analyzed these aforementioned 14 genes two days later (D10 NPCs) (Figure 4A). At this read-out time point, >50% reduction of Zeb2 mRNA expression was obtained (Figure 4B).
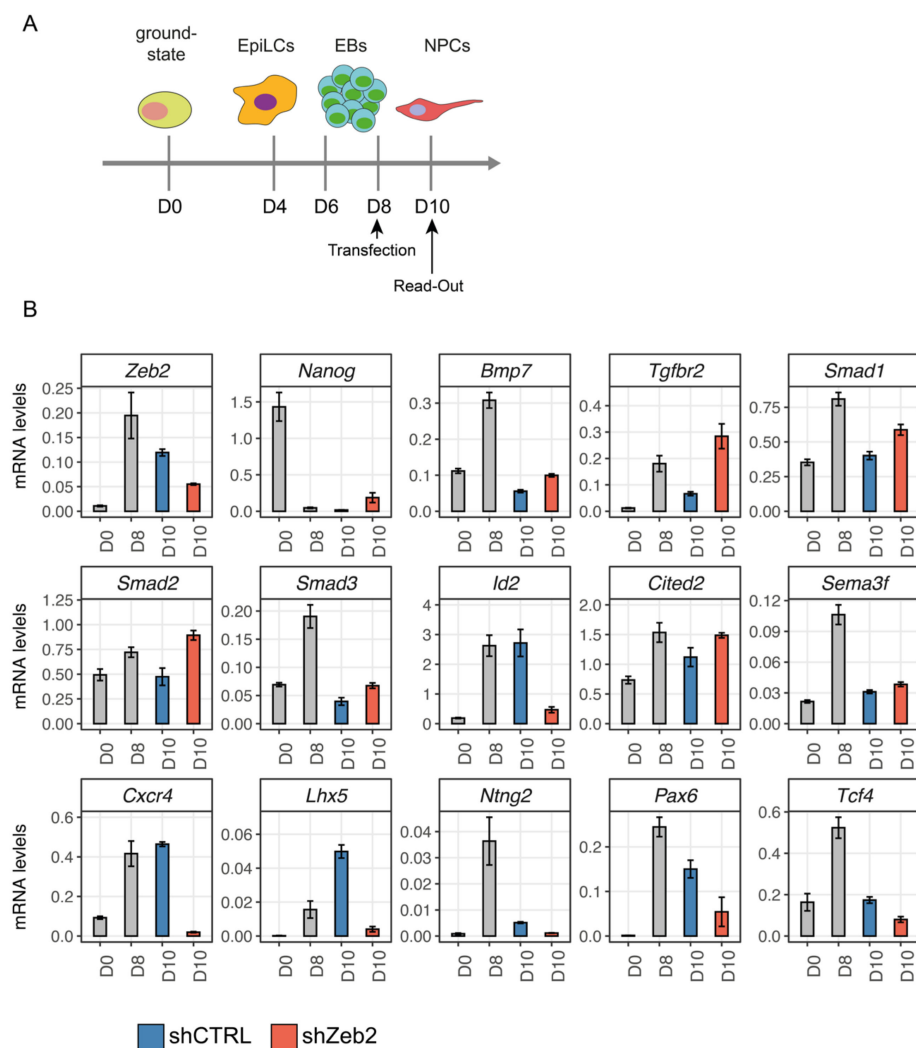
Figure 4. shRNA-mediated KD of Zeb2 discriminates between primary and secondary target genes. (**A**) Schematic overview of the shRNA transfection targeting Zeb2 (shZeb2) and read-out of the effect. Cellular aggregates at D8 of ND are dissociated and transfected with shZeb2 or against a scrambled, control sequence (shCTRL). Read-out is done two days after the start of shRNA addition. The list of shRNAs is given in Table 3. (**B**) RT-qPCR measurements: Zeb2 levels after KD were reduced to 40–50% of their normal level (shZeb2, orange bars) compared to shCTRL (blue bars). *Bmp7*, *Cited2*, *Nanog*, *Sema3f*, *Smad1*, *Smad2*, *Smad3*, and *Tgfbr2* were upregulated following Zeb2 KD, whereas genes encoding for neuronal specification and migration (*Cxcr4*, *Lhx5*, *Ntng2*, *Pax6*, and *Tcf4*) were downregulated.

Zeb2-KD resulted in reduced mRNA levels of *Cxcr4*, *Ntng2*, and *Pax6* (Figure 4B), genes that are each involved in neuron specification and migration [69,70]. Zeb2-KD also caused down-regulation of *Lhx5*, involved in the differentiation of interneurons, including cytoskeletal rearrangements during dendritogenesis [71], and of *Tcf4*, which acts in neurogenesis [54,55]. Sema3f is a cue for axon outgrowth and neuron migration guidance, and its gene was slightly upregulated (Figure 4B). These results confirm the regulation by Zeb2 of its direct targets in later phases of neuronal differentiation/migration.

The expression of Nanog, the promoter of which binds Zeb2 as a repressor [25], was increased in the Zeb2-KD cells (Figure 4B). Zeb2-KD caused increased mRNA of *Bmp7*, *Tgfbr2*, *Smad1*, *Smad2*, *Smad3*, and *Cited2* (Figure 4B), fitting with the normal levels of Zeb2 that mount anti-TGFβ/BMP family effects [29]. In contrast, *Id2* is strongly downregulated in shZeb2-treated ESCs (Figure 4B). *Id2* is normally activated by BMP-Smads and, together with other Id proteins (Id1, Id3, and Id4), inhibits cell differentiation, e.g., Zeb2 represses *Id2* in immune cells to promote differentiation [22]. However, *Id2* as well as other *Id* genes [72–74] is, such as *Zeb2* [20], also expressed in the developing forebrain.

Taken together, these data suggest an active and direct role for Zeb2 in repressing genes regulating stem cell pluripotency as well as a number of TGFβ/BMP-system components (*Bmp7*, *Smad1/2/3*), but also in activating genes during neurogenesis (*Cxcr4*, *Ntng2*, *Lhx5*).

3.6. Zeb2 Potentiates Its Own Gene Expression, Which Is Crucial for Proper Control of Some of Its Direct Target Genes

Strikingly, in our ChIP-seq dataset, the peak with the highest enrichment (~200-fold) mapped 232 bp upstream of the *Zeb2* TSS (Figure 5A, Table S1). For further functional studies of this site, we deleted the encompassing region (chr2:45109746-45110421) using CRISPR/Cas9 in wild-type mESCs, thereby obtaining $Zeb2^{\Delta P/\Delta P}$ ESCs (Figure S5 see Section 2). Figure 5B shows that the Zeb2 mRNA levels in the homozygous ΔP clone stayed strikingly low during ND, already from D8 onwards, compared to control cells. We then used these $Zeb2^{\Delta P/\Delta P}$ mESCs to read out the same genes that depend on intact Zeb2 levels and are Zeb2 ChIP+ (see Figure 4B). Levels of Zeb2 mRNA stayed abnormally low at D10 in $Zeb2^{\Delta P/\Delta P}$ ND-mESCs, whereas *Nanog* was still expressed and remained higher than in control WT cells (Figure 5B). Hence, Zeb2 levels are critical, albeit to a different degree for sets of genes. The latter include neuronal-relevant genes such as *Cxcr4*, *Lhx5*, *Ntng2*, *Pax6*, and *Tcf4* (Figure 5B). Among the TGFβ/BMP-system components (see Figure 4) we observed a limited reduction of *Bmp7*, *Smad1*, and *Smad3*, whereas *Tgfbr2*, *Smad2*, *Cited2*, and *Sema3f* expression was not affected in $Zeb2^{\Delta P/\Delta P}$ ND-mESCs. Based on these results, we speculate that Zeb2, the precise amounts of Zeb2, and in a critical stage also its autoregulation, are crucial in discriminating genes where Zeb2 plays the aforementioned primary, active role (as for *Cxcr4*, *Lhx5*, *Ntng2*, etc.). For these genes, ~50% reduction or mutation of the autoregulatory binding sequence is sufficient to strongly deregulate them, but other genes' expression is either not or just slightly affected (*Sema3f*, *Smad2*, *Cited2* vs. *Bmp7*, *Tgfbr2*, *Smad1*, *Smad3*).

Zeb2 also binds phospho(p)-Smads [4,21,23,29]. Therefore, we also scanned the Zeb2 ChIP+ direct target genes deregulated upon Zeb2-KD and/or in the $Zeb2^{\Delta P/\Delta P}$ cells during ND (i.e., without Smad activation) for the presence of (i) the Zeb half-sites CACCT(G) [3] and (ii) candidate p-Smad binding and responsive genes (using $GTC(^T/_G)CT(^T/_G)(^A/_C)GCC$ for p-Smad1/Smad5, GTCTAGAC for p-Smad2/3) and (iii) the co-Smad Smad4 ($C(^C/_T)AGAC$), using the Jaspar database (for a review on Smad target sites, see [75]; see also Section 2). Figure S6 shows the distribution of such identified Zeb and Smad-binding motifs (threshold score > 85%) in those genes strongly affected by Zeb2-KD and/or in $Zeb2^{\Delta P/\Delta P}$ cells. Interestingly, in the regions where Zeb2 binds close to the TSS (*Zeb2*, *Ntng2*, *Lhx5*, *Nanog*), the p-Smad and Smad4 binding elements are sometimes present in very close proximity to the ChIP+ Zeb2-bound E-box, indicating a possible cross-talk between receptor-activated Smads and Zeb2 in regulating target genes.

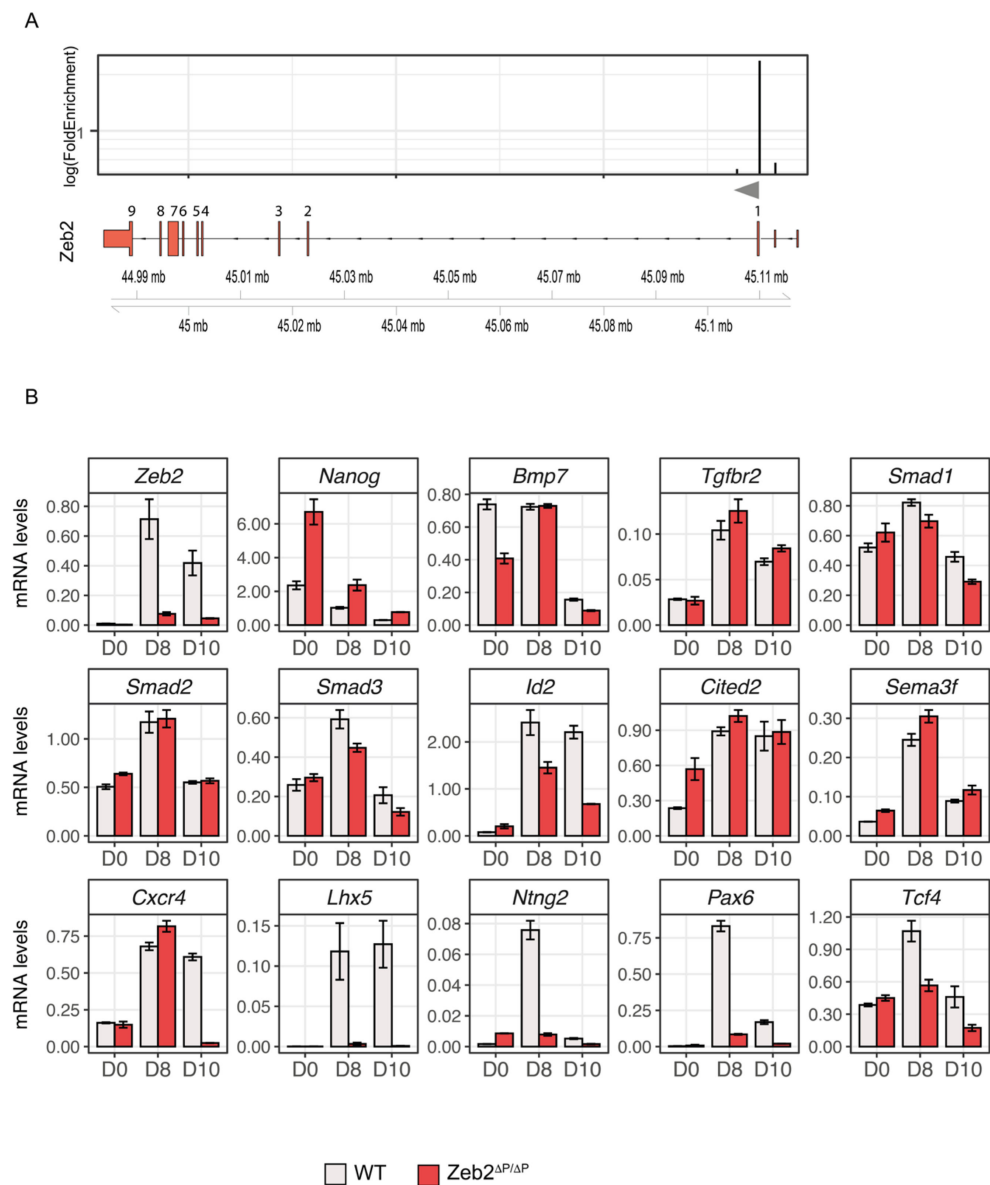**Figure 5.** Deletion of the Zeb2-binding autoregulatory site (chr2:45109746-45110421) impairs Zeb2 mRNA levels and neuronal markers. (**A**) Schematic overview of the log (FoldEnrichment) of the peaks identified by ChIP-seq located in the mouse *Zeb2* locus, and localization on top of *Zeb2* intron/exon structure. The highest peak is located 232 bp upstream of the first translated exon. The grey arrow indicates the TSS of *Zeb2*; (**B**) RT-qPCR: Zeb2 mRNA levels are strongly reduced in the *Zeb2*$^{\Delta P/\Delta P}$ clone. For the target genes validated with shRNA (see also Figure 4), most, but not all of the genes found to be affected following Zeb2 KD are also deregulated in *Zeb2*$^{\Delta P/\Delta P}$ mESCs, in particular Id2 and the neuronal markers *Cxcr4*, *Lhx5*, *Ntng2*, *Pax6*, and *Tcf4*.

3.7. Extrapolation of Zeb2 ChIP-Seq Data to Cell-Based Clinical Manifestation of MOWS

Despite the unprecedented nature of our Zeb2 ChIP-seq data obtained in ESC-derived NPCs in this study, an ideal extrapolation to MOWS, and in particular its clinical manifestation, is not straightforward and remains speculative (see Section 4). In this respect, the selected three experimental models of Zeb2 perturbation (including the two Zeb2-cKO models with cellular phenotypes that reveal underlying defects in MOWS [23,36]) present a more suitable intermediate hold, considering our list of 108 target genes for Zeb2. These options keep in mind the molecular and cellular consequences (e.g., gene expression profiles, cell differentiation states) of Zeb2 deficiency in these mouse models (for a recent review of most mouse models, see [30]). In addition, we considered deficiencies (with a focus

on neurodevelopmental defects) in MOWS patients as documented by clinicians [8–18]. The realistic options in trying to correlate Zeb2 ChIP-seq and Zeb2-perturbation RNA-seq data with MOWS in the clinic are presently three-fold. They each consider the aberrant expression of putative Zeb2-dependent direct target genes in (i) tissues/cells from MOWS patients, including MOWS iPSCs and derived neurons. However, this work in the MOWS field is only starting now. Therefore, we did this (ii) first for each of the 108 identified genes (Table S4), and then (iii) repeated this exercise for our entire list of genes to which Zeb2 binding sites were assigned (Table S1). Again, in both exercises, we tried to relate genes to clinical consequences in MOWS.

The Zeb2-dependent direct target genes within the list of 108 genes from this study were first grouped based on cellular functions that may relate to aspects of MOWS, and within these groups, relevant gene candidates for further investigation in extra cell systems in the future were then proposed, such as MOWS iPSCs that have been driven into neural differentiation and possibly neuron subtypes (option (i) above). Table S4 shows such genes for (calcium) ion-binding and channels (with *Syt13*, *Pcdh9*, *Nalcn*, and *Kcnj6* as candidate relevant genes); or operate in synapse biology (with again *Pcdh9*); axon outgrowth, guidance, and connectivity (with many genes, including *Epha5* and *Sema3f*); neuron subtype specification (including *Pax6*, *Klf7*, again *Efna5*, and *Lhx5*, *Hmx2*); motoric capabilities (with again *Nalcn* and *Lhx5*); and genes related to neurodevelopmental disorders and behavior (such as *Npas1*, *Cacng5*, *Foxp1*, *Wdr62*, and again *Kcnj6*).

Interestingly, some of these candidate genes for priority inclusion in future MOWS cell-based studies also emanated from independently overlooking all mapped GWBS for Zeb2 from this study (in Table S1) with an eye for anticipated MOWS cell biological defects, focusing mainly on neural and glial cells. These genes include *Bcl11* (developmental intellectual disorders, agenesis of corpus callosum), *Caln1*, *Efna5*, *Galnt5/6*, *Gng4/7*, *Isl1* (a TF that regulates expression of *Slit* and *Robo* genes), *Klf2/7/14*, *Nalcn*, *Pax6*, *Pcdh9/20*, *Pipox*, *Pou3f/4f* members (one also known as *Brn3a*), *Ror2*, *Sema3*, *Slc14a2*, *Sox1*, *Tcf4* (see [57] for a detailed discussion), *Tubb3/6*, and of course *Zeb2*. From Table S1, we would also prioritize genes for adhesion G-coupled receptors (e.g., *Adgre5*, *Adgrl2*), adherens junctions (*Ajap1*, *Frmd4a*, *Jam3*), several *Cdh* genes, genes involved in Wnt (*Axin2*, *Kremen1*) or BMP signaling (*Rgmb*, *Ror2*), sulfotransferase-encoding genes (*Chst2/7*, *Hs3st3a1*, *Ndst1*), *Camkk1* (for MOWS patients have CAMK deficiency), chemokine receptor genes (*Ccr1/7*), *Ddx10/18*, *Efhb*, *Fgf14*, *Lrrc4c* (encoding a binding partner of long-range guidance cue Netrin G1), *Nrn1* (encoding a neuritin, involved in neuronal plasticity), *Pitx2*, *Plk2* (encoding a kinase that links to epilepsy), *Prex1*, *Prox1*, *Robo2*, *Six2*, *Snai3*, *Sox5*, *Tenm3* (for proper connectivity in the nervous system), and Tox3 (chromatin bending).

4. Discussion

We report for the first time the endogenous genome-wide binding sites (GWBS) for Zeb2, in ESC-derived NPCs. In previous work, we have used ESCs established from *Zeb2*$^{\Delta ex7/\Delta ex7}$-KO [76] pre-implantation embryos and, for rescue purposes, such KO ESCs cells in which Flag$_3$-Strep-Zeb2 was produced from (a Cre-controllable) *Rosa26* locus [25]. The latter cells are different from the mESCs that were established here, since *Zeb2* is not subjected to its normal temporal regulation during cell differentiation, in contrast to the Flag-V5 mouse (m) ESCs obtained here.

Precise dosage of Zeb2 is however a critical factor in vivo (for a recent discussion, see [30]). This is concluded from transgenic Zeb2 cDNA-based rescues in *Zeb2*-KO ESCs and similar genetic rescues in Zeb2-mutant cells in mice, which via heterozygous/homozygous combinations create an elegant and large panel of Zeb2 mRNA levels (in interneurons [36], NK cells [77], ESCs [25]). Another illustration of the relevance of fine-tuned control of Zeb2 levels are miRs that target Zeb2, and lncRNAs that regulate these miRs [78–81], with Zeb2 in its turn also controlling some of its own miR-encoding genes or clusters [82–84]. In our ChIP-seq, we find 125 peaks (~5% of the total) that correspond to the TSSs of 98 miR genes (Table S1). Among these miR genes, Zeb2 binds to loci encoding miR-144, miR-148a,

miR-9, and miR-153, known to target Zeb2 in the context of e.g., tumor EMT and tumor progression [85–88]. We recently added the identification, in human iPSCs subjected to ND, of ZEB2 distant (~600 kb upstream) enhancers, which act through DNA-looping to the ZEB2 promoter-proximal region [89]. In addition, we have documented dynamic expression patterns of Zeb2 in early embryos [27,28,33,90–92]. We have also shown that cDNA-based expression of various tag-Zeb2 proteins is compatible with functional embryology-type and action mechanism studies [25,48,49]. Importantly, our *Zeb2-V5* allele steers normal production of tag-Zeb2 from its endogenous locus.

Only two studies present ZEB2 ChIP-seq data in human cells, i.e., SNU398 hepatocellular carcinoma and K562 erythroleukemia cells, respectively. In K562 cells ZEB2 binds to the promoters of *NR4A2*, *NEUROG2*, and *PITX3*, expressed in midbrain dopaminergic neurons, wherein—in mice—Zeb2 negatively regulates axon growth and target innervation [93]. In SNU398 cells, ZEB2 represses *GALNT3*, which is normally expressed in epithelial cells. This repression coincides with the acquisition of a mesenchymal phenotype, linking ZEB2 here again to an EMT-like process. These two valuable studies also present limitations. The use of cancer cell lines of a genomica unstable nature may create possible bias in ChIP-seq, and in any case, they overproduce ZEB2. Our ChIP-seq identifies >2400 peaks for Zeb2-V5 in mESCs at ND-D8; 37.5% of the Zeb2 sites map close to TSSs (when defined as −10/+10 kb). Most of these genes function in growth factor or cytokine signaling and/or encode transcriptional regulators. The latter suggests that Zeb2 orchestrates other cooperating TFs driving the transcriptomic signature of NPCs. The regulation of Wnt signaling by Zeb2 is in line with observations that inhibition of the Wnt-βcatenin pathway suppresses ND in vitro and in vivo and that Wnt (and Zeb2)-controlled *Tcf4* expression promotes neurogenesis and is required for normal brain development [94–98].

The 2294 Zeb2 peaks map to 1952 protein-coding genes, of which 1244 are DEGs in NDmESCs. Strikingly, the strongest enrichment of Zeb2 occurs on the *Zeb2* promoter itself, leading to the identification of a novel self-regulatory mechanism where Zeb2 binds upstream of its TSS to maintain its levels sufficiently high, at least during ND. While this autoregulation needs further investigation in Zeb2-dependent differentiation and/or maturation of other cell types (e.g., in cKO mouse models or in ND-iPSCs derived from appropriate MOWS patient cells), we propose that lower Zeb2 levels might compromise this autoregulatory loop. Deletion of the autoregulatory site from both Zeb2 alleles (in the $\Delta P/\Delta P$ cells), results in a significant decrease of Zeb2 mRNA levels but *Zeb2* is still partially expressed, and these cells can still exit from pluripotency and differentiate. A number of genes, which are mainly linked to neuron maturation, are affected in $Zeb2^{\Delta P/\Delta P}$ ESCs, whereas TGFβ/BMP-system component genes are not deregulated. Zeb2 dosage might thus underlie this difference in regulating its direct, ChIP+ genes in our ESCs. Zeb2 might be key to maintaining expression of neuronal genes, while for TGFβ/BMP system genes Zeb2 may cooperate with other TFs (including p-Smads) or DNA-modifying enzymes to regulate the expression of target genes.

Zeb2 binds to TGFβ/BMP family receptor-activated phospho-Smads (pSmads), and several studies indicate its negative regulation of BMP-Smad activation of specific target genes, although Zeb2 also has Smad-independent functions [23,29]. BMP-pSmads bind to GGCGCC with high affinity [99]. Morikawa and co-workers [100] have confirmed these results using ChIP-seq, and also identified a lower-affinity (so, higher BMP-doses required) BMP-Smad element (GGAGCC). To achieve full responsiveness, it was proposed that the GG(^{A}/$_C$)GCC element needs to be coupled with a Smad4 site, optimally located 5 bp away [100]. We find that in primary targets affected by varying levels of Zeb2, E-boxes are located close to Smad-binding motifs. However, whether Zeb2 and Smads are co-present in target regions requires further experiments, such as ChIP-on-chip assays, and (non-neural) differentiation protocols (involving stimulation of the cells by addition of BMP and/or Nodal). However, these studies may be further complicated because of the post-translational modification status of Zeb2, nuclear p-Smads, and/or Smad4 [101–104].

The 1093 Zeb2-binding DEGs at D8 are also striking. When we performed a gene-to-disease association using the human orthologues of these D8-DEGs, we found a clear association with several disorders (Figure S6). These include neurodevelopmental, mental, and ocular defects, which occur in MOWS. Altogether, our data may provide novel insights into MOWS due to suboptimal ZEB2 amounts in patients, and from now includes a gene autoregulation aspect, as well as *ZEB2* as a putative modifier gene for many other congenital disorders.

Several *Zeb2*-cKO mouse models have been generated, and many bulk RNA-seq data are available, from which here we selected two such data sets [23,36,105]. In addition, similar data were obtained from cultured mESCs, either *Zeb2*-KO cells (25) or cells submitted to ND wherein e.g., Zeb2-KD was performed [37]. We could not include *Zeb2*-KO mESCs in these comparisons, for they convert dramatically less efficiently to EpiLSCs and they fail to differentiate beyond this EpiLC state [25]. The meta-analysis of the three different used data sets, overlaid with the 1952 Zeb2-V5 ChIP+ loci/genes, show therefore a limited number of common targets, *Cxcr4* being the only one common to all data sets. Throughout all three different datasets, we could narrow down the target gene collection to 108 Zeb2-bound genes in total. However, interestingly, these 108 genes enrich GO terms such as pluripotency of stem cells, signaling by TGFβ and Wnt, cell fate commitment, and neuron differentiation, all processes wherein Zeb2 plays a crucial role.

Out of these 108 genes, we selected 14 covering TGFβ/BMP signaling, pluripotency, neuron migration, and differentiation/maturation, and checked their levels two days after Zeb2-KD at ND-D8. Most of these 14 genes relevant to NPC status were shown to be critically dependent on intact levels of Zeb2. They may help to explain why the defects caused by MOWS are observed later after birth and why (the few) missense mutations in MOWS (besides the more abundant significant deletions) present with milder syndromic manifestation. Both the cross-reference of Zeb2-ChIP+ genes with the transcriptome of ND-ESCs and the meta-analysis identify a number of common genes, such as *Bmp7*, *Tgfbr2*, *Tcf4*, *Smad1/2/3*, and *Sema3f* (Table S3; Figure S4), making Zeb2 a likely direct regulator of these genes. It is also intriguing that Zeb2 is recruited to and controlling *Tcf4* at D8, and that *Tcf4* is deregulated upon Zeb2-KD (using esiRNA [37], and here shRNA). Mutations in *TCF4* cause PTHS, a rare neurodevelopmental disorder with some defects overlapping with MOWS (for an extensive discussion, see [57]). The binding of Zeb2 to *Tcf4* opens new attractive roads to further investigate the crosstalk between these two TFs and their role in regulating crucial aspects of neurodevelopment.

Despite the novelty of our mapped Zeb2 GWBS, the extrapolation of this work to the clinical manifestation of MOWS itself is interesting but remains speculative. First, all Zeb2-dependent genes whose mRNA levels change upon Zeb2 perturbation in vivo, and that are relevant to explain aspects of MOWS (for example, neurodevelopmental deficiencies; see [30]), have been identified in homozygous *Zeb2*-KO cell types from mouse models, whereas MOWS in humans is caused by heterozygous *ZEB2* mutations [12–16,30]. At present, we do not know what such heterozygous MOWS mutation means with regard to *ZEB2* autoregulation in patient cells. However, the *Zeb2*-KO mouse models, if ideally including phenotyping by RNA-seq, have allowed for the proposal of cellular mechanisms underlying major MOWS-related defects in humans. Examples would be the neurocristopathies that impact craniofacial and enteric nervous system development, leading to Hirschsprung disease in (most) MOWS patients [9,10,16]. Another defect, intriguingly but largely clinically unexplored, concerns the development of dorsal root ganglia, relating to possible pain deficiencies in MOWS patients. Key examples may be nervous system deficiencies, such as those found in the timing of neuro-/gliogenesis in corticogenesis in KO mice, relating to intellectual disability in the patients [20,23,36], and those in guided migration and connectivity of forebrain interneurons in KO mice, relating to seizures and epilepsy in MOWS patients [36]. Confirmatory gene expression profiling studies have, to our knowledge, not been carried out on tissues/cells from MOWS patients. Furthermore, this approach would be restricted to cells from post-natal and adult tissues and would focus

on fully differentiated cells, e.g., mature neurons. In reverse, our ESC-to-NPC-based cellular model has not been explored yet towards differentiation of specific neuron subtypes, and their maturation in vitro. Last but not least, the establishment of iPSCs from MOWS patients directly or as the result of *ZEB2* (e.g., heterozygous inactivation) gene-editing in wild-type iPSCs is still in a very early phase. Such iPSCs would even be more ideal if the single remaining, intact wild-type *ZEB2* allele would also be epitope-tagged—as carried out here in mESCs—for repeating the ChIP-seq experiments upon neural differentiation. With these careful considerations in mind, and in an attempt to correlate this Zeb2 ChIP-seq work with the clinical manifestation of MOWS, we nonetheless tentatively propose a number of candidate relevant genes for further investigation that might be relevant to MOWS as it presents in the clinic (see Results Section 3.7). These speculations do not yet include other crucial Zeb2 functions—as documented mainly in mouse models [30], such as in (re)myelination [21,24,61], diverse cell types of the immune system [22,34,77], and perhaps the formation of the immune synapse, and fibrosis [106]. Some of these novel phenotypes in mice prompt many clinicians to start longitudinal follow-up in MOWS patients also from these perspectives.

References

1.   Funahashi, J.; Kamachi, Y.; Goto, K.; Kondoh, H. Identification of nuclear factor delta EF1 and its binding site essential for lens-specific activity of the delta 1-crystallin enhancer. *Nucleic Acids Res.* **1991**, *19*, 3543–3547. [CrossRef]
2.   Sekido, R.; Murai, K.; Funahashi, J.; Kamachi, Y.; Fujisawa-Sehara, A.; Nabeshima, Y.; Kondoh, H. The delta-crystallin enhancer-binding protein delta EF1 is a repressor of E2-box-mediated gene activation. *Mol. Cell. Biol.* **1994**, *14*, 5692–5700. [CrossRef]
3.   Remacle, J.E.; Kraft, H.; Lerchner, W.; Wuytens, G.; Collart, C.; Verschueren, K.; Smith, J.C.; Huylebroeck, D. New mode of DNA binding of multi-zinc finger transcription factors: DeltaEF1 family members bind with two hands to two target sites. *EMBO J.* **1999**, *18*, 5073–5084. [CrossRef]
4.   Verschueren, K.; Remacle, J.E.; Collart, C.; Kraft, H.; Baker, B.S.; Tylzanowski, P.; Nelles, L.; Wuytens, G.; Su, M.T.; Bodmer, R.; et al. SIP1, a novel zinc finger/homeodomain repressor, interacts with Smad proteins and binds to 5′-CACCT sequences in candidate target genes. *J. Biol. Chem.* **1999**, *274*. [CrossRef]
5.   Mowat, D.R.; Wilson, M.J.; Goossens, M. Mowat-Wilson syndrome. *J. Med. Genet.* **2003**, *40*, 305–310. [CrossRef]
6.   Cerruti Mainardi, P.; Pastore, G.; Zweier, C.; Rauch, A. Mowat-Wilson syndrome and mutation in the zinc finger homeo box 1B gene: A well-defined clinical entity. *J. Med. Genet.* **2004**, *41*, e16. [CrossRef]
7.   Ishihara, N.; Yamada, K.; Yamada, Y.; Miura, K.; Kato, J.; Kuwabara, N.; Hara, Y.; Kobayashi, Y.; Hoshino, K.; Nomura, Y.; et al. Clinical and molecular analysis of Mowat-Wilson syndrome associated with ZFHX1B mutations and deletions at 2q22-q24.1. *J. Med. Genet.* **2004**, *41*, 387–393. [CrossRef]
8.   Mowat, D.R.; Croaker, G.D.; Cass, D.T.; Kerr, B.A.; Chaitow, J.; Adès, L.C.; Chia, N.L.; Wilson, M.J. Hirschsprung disease, microcephaly, mental retardation, and characteristic facial features: Delineation of a new syndrome and identification of a locus at chromosome 2q22-q23. *J. Med. Genet.* **1998**, *35*, 617–623. [CrossRef]
9.   Cacheux, V.; Dastot-Le Moal, F.; Kääriäinen, H.; Bondurand, N.; Rintala, R.; Boissier, B.; Wilson, M.; Mowat, D.; Goossens, M. Loss-of-function mutations in SIP1 Smad interacting protein 1 result in a syndromic Hirschsprung disease. *Hum. Mol. Genet.* **2001**, *10*, 1503–1510. [CrossRef]
10. Wakamatsu, N.; Yamada, Y.; Yamada, K.; Ono, T.; Nomura, N.; Taniguchi, H.; Kitoh, H.; Mutoh, N.; Yamanaka, T.; Mushiake, K.; et al. Mutations in SIP1, encoding Smad interacting protein-1, cause a form of Hirschsprung disease. *Nat. Genet.* **2001**, *27*, 369–370. [CrossRef]
11. Yamada, K.; Yamada, Y.; Nomura, N.; Miura, K.; Wakako, R.; Hayakawa, C.; Matsumoto, A.; Kumagai, T.; Yoshimura, I.; Miyazaki, S.; et al. Nonsense and frameshift mutations in ZFHX1B, encoding Smad-interacting protein 1, cause a complex developmental disorder with a great variety of clinical features. *Am. J. Hum. Genet.* **2001**, *69*, 1178–1185. [CrossRef]
12. Wilson, M.; Mowat, D.; Dastot-Le Moal, F.; Cacheux, V.; Kääriäinen, H.; Cass, D.; Donnai, D.; Clayton-Smith, J.; Townshend, S.; Curry, C.; et al. Further delineation of the phenotype associated with heterozygous mutations in ZFHX1B. *Am. J. Med. Genet. A* **2003**, *119*, 257–265. [CrossRef]
13. Zweier, C.; Thiel, C.T.; Dufke, A.; Crow, Y.J.; Meinecke, P.; Suri, M.; Ala-Mello, S.; Beemer, F.; Bernasconi, S.; Bianchi, P.; et al. Clinical and mutational spectrum of Mowat-Wilson syndrome. *Eur. J. Med. Genet.* **2005**, *48*, 97–111. [CrossRef]
14. Zweier, C.; Horn, D.; Kraus, C.; Rauch, A. Atypical ZFHX1B mutation associated with a mild Mowat-Wilson syndrome phenotype. *Am. J. Med. Genet. A* **2006**, *140*, 869–872. [CrossRef]
15. Garavelli, L.; Zollino, M.; Mainardi, P.C.; Gurrieri, F.; Rivieri, F.; Soli, F.; Verri, R.; Albertini, E.; Favaron, E.; Zignani, M.; et al. Mowat-Wilson syndrome: Facial phenotype changing with age: Study of 19 Italian patients and review of the literature. *Am. J. Med. Genet. A* **2009**, *149*, 417–426. [CrossRef]
16. Ivanovski, I.; Djuric, O.; Caraffi, S.G.; Santodirocco, D.; Pollazzon, M.; Rosato, S.; Cordelli, D.M.; Abdalla, E.; Accorsi, P.; Adam, M.P.; et al. Phenotype and genotype of 87 patients with Mowat-Wilson syndrome and recommendations for care. *Genet. Med.* **2018**, *20*, 965–975. [CrossRef]
17. Garavelli, L.; Ivanovski, I.; Caraffi, S.G.; Santodirocco, D.; Pollazzon, M.; Cordelli, D.M.; Abdalla, E.; Accorsi, P.; Adam, M.P.; Baldo, C.; et al. Neuroimaging findings in Mowat-Wilson syndrome: A study of 54 patients. *Genet. Med.* **2017**, *19*, 691–700. [CrossRef]
18. Ricci, E.; Fetta, A.; Garavelli, L.; Caraffi, S.; Ivanovski, I.; Bonanni, P.; Accorsi, P.; Giordano, L.; Pantaleoni, C.; Romeo, A.; et al. Further delineation and long-term evolution of electroclinical phenotype in Mowat Wilson Syndrome. A longitudinal study in 40 individuals. *Epilepsy Behav.* **2021**, *124*, 108315. [CrossRef]
19. Comijn, J.; Berx, G.; Vermassen, P.; Verschueren, K.; van Grunsven, L.; Bruyneel, E.; Mareel, M.; Huylebroeck, D.; van Roy, F. The two-handed E box binding zinc finger protein SIP1 downregulates E-cadherin and induces invasion. *Mol. Cell* **2001**, *7*, 1267–1278. [CrossRef]
20. Seuntjens, E.; Nityanandam, A.; Miquelajauregui, A.; Debruyn, J.; Stryjewska, A.; Goebbels, S.; Nave, K.A.; Huylebroeck, D.; Tarabykin, V. Sip1 regulates sequential fate decisions by feedback signalling from postmitotic neurons to progenitors. *Nat. Neurosci.* **2009**, *12*, 1373–1380. [CrossRef]
21. Weng, Q.; Chen, Y.; Wang, H.; Xu, X.; Yang, B.; He, Q.; Shou, W.; Chen, Y.; Higashi, Y.; van den Berghe, V.; et al. Dual-mode modulation of Smad signalling by Smad-interacting protein Sip1 is required for myelination in the central nervous system. *Neuron* **2012**, *73*, 713–728. [CrossRef]

22. Scott, C.L.; Soen, B.; Martens, L.; Skrypek, N.; Saelens, W.; Taminau, J.; Blancke, G.; Van Isterdael, G.; Huylebroeck, D.; Haigh, J.; et al. The transcription factor Zeb2 regulates development of conventional and plasmacytoid DCs by repressing Id2. *J. Exp. Med.* **2016**, *213*, 897–911. [CrossRef]

23. Deryckere, A.; Stappers, E.; Dries, R.; Peyre, E.; van den Berghe, V.; Conidi, A.; Zampeta, F.I.; Francis, A.; Bresseleers, M.; Stryjewska, A.; et al. Multifaceted actions of Zeb2 in postnatal neurogenesis from the ventricular-subventricular zone to the olfactory bulb. *Development* **2020**, *147*, dev184861. [CrossRef]

24. Wu, L.M.; Wang, J.; Conidi, A.; Zhao, C.; Wang, H.; Ford, Z.; Zhang, L.; Zweier, C.; Ayee, B.G.; Maurel, P.; et al. Zeb2 recruits HDAC-NuRD to inhibit Notch and controls Schwann cell differentiation and remyelination. *Nat. Neurosci.* **2016**, *19*, 1060–1072. [CrossRef]

25. Stryjewska, A.; Dries, R.; Pieters, T.; Verstappen, G.; Conidi, A.; Coddens, K.; Francis, A.; Umans, L.; van IJcken, W.F.; Berx, G.; et al. Zeb2 Regulates Cell Fate at the Exit from Epiblast State in Mouse Embryonic Stem Cells. *Stem Cells* **2017**, *35*, 611–625. [CrossRef]

26. Menuchin-Lasowski, Y.; Dagan, B.; Conidi, A.; Cohen-Gulkar, M.; David, A.; Ehrlich, M.; Giladi, P.O.; Clark, B.S.; Blackshaw, S.; Shapira, K.; et al. Zeb2 regulates the balance between retinal interneurons and Müller glia by inhibition of BMP-Smad signalling. *Dev. Biol.* **2020**, *468*, 80–92. [CrossRef]

27. Lerchner, W.; Latinkic, B.V.; Remacle, J.E.; Huylebroeck, D.; Smith, J.C. Region-specific activation of the Xenopus brachyury promoter involves active repression in ectoderm and endoderm: A study using transgenic frog embryos. *Development* **2000**, *127*, 2729–2739. [CrossRef]

28. Van Grunsven, L.A.; Taelman, V.; Michiels, C.; Opdecamp, K.; Huylebroeck, D.; Bellefroid, E.J. deltaEF1 and SIP1 are differentially expressed and have overlapping activities during Xenopus embryogenesis. *Dev. Dyn.* **2006**, *235*, 1491–1500. [CrossRef]

29. Conidi, A.; Cazzola, S.; Beets, K.; Coddens, K.; Collart, C.; Cornelis, F.; Cox, L.; Debruyn, J.; Dobreva, M.P.; Dries, R.; et al. Few Smad proteins and many Smad-interacting proteins yield multiple functions and action modes in TGFβ/BMP signalling in vivo. *Cytokine Growth Factor Rev.* **2011**, *22*, 287–300. [CrossRef]

30. Birkhoff, J.; Huylebroeck, D.; Conidi, A. ZEB2, the Mowat-Wilson Syndrome transcription factor: Confirmations, novel functions, and continuing surprises. *Genes* **2021**, *12*, 1037. [CrossRef]

31. Balcik-Ercin, P.; Cetin, M.; Yalim-Camci, I.; Odabas, G.; Tokay, N.; Sayan, A.E.; Yagci, T. Genome-wide analysis of endogenously expressed ZEB2 binding sites reveals inverse correlations between ZEB2 and GalNAc-transferase GALNT3 in human tumors. *Cell. Oncol.* **2018**, *41*, 379–393. [CrossRef]

32. Yang, S.; Toledo, E.M.; Rosmaninho, P.; Peng, C.; Uhlén, P.; Castro, D.S.; Arenas, E. A Zeb2-miR-200c loop controls midbrain dopaminergic neuron neurogenesis and migration. *Commun. Biol.* **2018**, *1*, 75. [CrossRef]

33. Miyoshi, T.; Maruhashi, M.; Van de Putte, T.; Kondoh, H.; Huylebroeck, D.; Higashi, Y. Complementary expression pattern of Zfhx1 genes Sip1 and deltaEF1 in the mouse embryo and their genetic interaction revealed by compound mutants. *Dev. Dyn.* **2006**, *235*, 1941–1952. [CrossRef]

34. Vandamme, N.; Denecker, G.; Bruneel, K.; Blancke, G.; Akay, Ö.; Taminau, J.; De Coninck, J.; De Smedt, E.; Skrypek, N.; Van Loocke, W.; et al. The EMT Transcription Factor ZEB2 Promotes Proliferation of Primary and Metastatic Melanoma While Suppressing an Invasive, Mesenchymal-Like Phenotype. *Cancer Res.* **2020**, *80*, 2983–2995. [CrossRef]

35. Nelles, L.; Van de Putte, T.; van Grunsven, L.; Huylebroeck, D.; Verschueren, K. Organization of the mouse Zfhx1b gene encoding the two-handed zinc finger repressor Smad-interacting protein-1. *Genomics* **2003**, *82*, 460–469. [CrossRef]

36. Van den Berghe, V.; Stappers, E.; Vandesande, B.; Dimidschstein, J.; Kroes, R.; Francis, A.; Conidi, A.; Lesage, F.; Dries, R.; Cazzola, S.; et al. Directed migration of cortical interneurons depends on the cell-autonomous action of Sip1. *Neuron* **2013**, *77*, 70–82. [CrossRef]

37. Dries, R.; Stryjewska, A.; Coddens, K.; Okawa, S.; Notelaers, T.; Birkhoff, J.; Dekker, M.; Verfaillie, C.M.; Del Sol, A.; Mulugeta, E.; et al. Integrative and perturbation-based analysis of the transcriptional dynamics of TGFβ/BMP system components in transition from embryonic stem cells to neural progenitors. *Stem Cells* **2020**, *38*, 202–217. [CrossRef]

38. Bibel, M.; Richter, J.; Lacroix, E.; Barde, Y.A. Generation of a defined and uniform population of CNS progenitors and neurons from mouse embryonic stem cells. *Nat. Protoc.* **2007**, *2*, 1034–1043. [CrossRef]

39. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A fast spliced aligner with low memory requirements. *Nat. Methods* **2015**, *12*, 357–360. [CrossRef]

40. Zhang, Y.; Liu, T.; Meyer, C.A.; Eeckhoute, J.; Johnson, D.S.; Bernstein, B.E.; Nusbaum, C.; Myers, R.M.; Brown, M.; Li, W.; et al. Model-based analysis of ChIP-Seq (MACS). *Genome Biol.* **2008**, *9*, R137. [CrossRef]

41. Feng, J.; Liu, T.; Qin, B.; Zhang, Y.; Liu, X.S. Identifying ChIP-seq enrichment using MACS. *Nat. Protoc.* **2012**, *7*, 1728–1740. [CrossRef]

42. Ramirez, F.; Ryan, D.P.; Gruning, B.; Bhardwaj, V.; Kilpert, F.; Richter, A.S.; Heyne, S.; Dundar, F.; Manke, T. deepTools2: A next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **2016**, *44*, W160–W165. [CrossRef]

43. Puig, R.R.; Boddie, P.; Khan, A.; Castro-Mondragon, J.A.; Mathelier, A. UniBind: Maps of high-confidence direct TF-DNA interactions across nine species. *BMC Genom.* **2021**, *22*, 482. [CrossRef]

44. Anders, S.; Pyl, P.T.; Huber, W. HTSeq—A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2015**, *31*, 166–169. [CrossRef]

45. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]

46. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [CrossRef]

47. Pinero, J.; Ramirez-Anguita, J.M.; Sauch-Pitarch, J.; Ronzano, F.; Centeno, E.; Sanz, F.; Furlong, L.I. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res.* **2020**, *48*, D845–D855. [CrossRef]

48. Papin, C.; van Grunsven, L.A.; Verschueren, K.; Huylebroeck, D.; Smith, J.C. Dynamic regulation of Brachyury expression in the amphibian embryo by XSIP1. *Mech. Dev.* **2002**, *111*, 37–46. [CrossRef]

49. Verstappen, G.; van Grunsven, L.A.; Michiels, C.; Van de Putte, T.; Souopgui, J.; Van Damme, J.; Bellefroid, E.; Vandekerckhove, J.; Huylebroeck, D. Atypical Mowat-Wilson patient confirms the importance of the novel association between ZFHX1B/SIP1 and NuRD corepressor complex. *Hum. Mol. Genet.* **2008**, *17*, 1175–1183. [CrossRef]

50. Afgan, E.; Baker, D.; Batut, B.; van den Beek, M.; Bouvier, D.; Cech, M.; Chilton, J.; Clements, D.; Coraor, N.; Gruning, B.A.; et al. The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **2018**, *46*, W537–W544. [CrossRef]

51. Sepp, M.; Pruunsild, P.; Timmusk, T. Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. *Hum. Mol. Genet.* **2012**, *21*, 2873–2888. [CrossRef]

52. Teixeira, J.R.; Szeto, R.A.; Carvalho, V.M.A.; Muotri, A.R.; Papes, F. Transcription factor 4 and its association with psychiatric disorders. *Transl. Psychiatry* **2021**, *11*, 19. [CrossRef]

53. Corneliussen, B.; Thornell, A.; Hallberg, B.; Grundstrom, T. Helix-loop-helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers. *J. Virol.* **1991**, *65*, 6084–6093. [CrossRef]

54. Forrest, M.P.; Hill, M.J.; Quantock, A.J.; Martin-Rendon, E.; Blake, D.J. The emerging roles of TCF4 in disease and development. *Trends Mol. Med.* **2014**, *20*, 322–331. [CrossRef]

55. Wang, L.H.; Baker, N.E. E Proteins and ID Proteins: Helix-Loop-Helix Partners in Development and Disease. *Dev. Cell* **2015**, *35*, 269–280. [CrossRef]

56. Fu, H.; Cai, J.; Clevers, H.; Fast, E.; Gray, S.; Greenberg, R.; Jain, M.K.; Ma, Q.; Qiu, M.; Rowitch, D.H.; et al. A genome-wide screen for spatially restricted expression patterns identifies transcription factors that regulate glial development. *J. Neurosci.* **2009**, *29*, 11399–11408. [CrossRef]

57. Meert, L.; Birkhoff, J.C.; Conidi, A.; Poot, R.A.; Huylebroeck, D. Different E-box binding transcription factors, similar neuro-developmental defects: ZEB2 (Mowat-Wilson syndrome) and TCF4 (Pitt-Hopkins syndrome). *Rare Dis. Orphan Drugs J.* **2022**, *1*, 8. [CrossRef]

58. Brockschmidt, A.; Todt, U.; Ryu, S.; Hoischen, A.; Landwehr, C.; Birnbaum, S.; Frenck, W.; Radlwimmer, B.; Lichter, P.; Engels, H.; et al. Severe mental retardation with breathing abnormalities (Pitt-Hopkins syndrome) is caused by haploinsufficiency of the neuronal bHLH transcription factor TCF4. *Hum. Mol. Genet.* **2007**, *16*, 1488–1494. [CrossRef]

59. Zhang, T.; Zhang, Z.; Dong, Q.; Xiong, J.; Zhu, B. Histone H3K27 acetylation is dispensable for enhancer activity in mouse embryonic stem cells. *Genome Biol.* **2020**, *21*, 45. [CrossRef]

60. Suh, H.; Consiglio, A.; Ray, J.; Sawai, T.; D'Amour, K.A.; Gage, F.H. In vivo fate analysis reveals the multipotent and self-renewal capacities of Sox2+ neural stem cells in the adult hippocampus. *Cell Stem Cell* **2007**, *1*, 515–528. [CrossRef]

61. Quintes, S.; Brinkmann, B.G.; Ebert, M.; Fröb, F.; Kungl, T.; Arlt, F.A.; Tarabykin, V.; Huylebroeck, D.; Meijer, D.; Suter, U.; et al. Zeb2 is essential for Schwann cell differentiation, myelination and nerve repair. *Nat. Neurosci.* **2016**, *19*, 1050–1059. [CrossRef]

62. Dunn, S.J.; Martello, G.; Yordanov, B.; Emmott, S.; Smith, A.G. Defining an essential transcription factor program for naïve pluripotency. *Science* **2014**, *344*, 1156–1160. [CrossRef]

63. Rao, C.; Malaguti, M.; Mason, J.O.; Lowell, S. The transcription factor E2A drives neural differentiation in pluripotent cells. *Development* **2020**, *147*, dev184093. [CrossRef]

64. El Wakil, A.; Francius, C.; Wolff, A.; Pleau-Varet, J.; Nardelli, J. The GATA2 transcription factor negatively regulates the proliferation of neuronal progenitors. *Development* **2006**, *133*, 2155–2165. [CrossRef]

65. Beagan, J.A.; Duong, M.T.; Titus, K.R.; Zhou, L.; Cao, Z.; Ma, J.; Lachanski, C.V.; Gillis, D.R.; Phillips-Cremins, J.E. YY1 and CTCF orchestrate a 3D chromatin looping switch during early neural lineage commitment. *Genome Res.* **2017**, *27*, 1139–1152. [CrossRef]

66. Hegarty, S.V.; Sullivan, A.M.; O'Keeffe, G.W. Zeb2: A multifunctional regulator of nervous system development. *Prog. Neurobiol.* **2015**, *132*, 81–95. [CrossRef]

67. Stumm, R.K.; Zhou, C.; Ara, T.; Lazarini, F.; Dubois-Dalcq, M.; Nagasawa, T.; Hollt, V.; Schulz, S. CXCR4 regulates interneuron migration in the developing neocortex. *J. Neurosci.* **2003**, *23*, 5123–5130. [CrossRef]

68. Nash, B.; Meucci, O. Functions of the chemokine receptor CXCR4 in the central nervous system and its regulation by mu-opioid receptors. *Int. Rev. Neurobiol.* **2014**, *118*, 105–128. [CrossRef]

69. Georgala, P.A.; Carr, C.B.; Price, D.J. The role of Pax6 in forebrain development. *Dev. Neurobiol.* **2011**, *71*, 690–709. [CrossRef]

70. Dias, C.M.; Punetha, J.; Zheng, C.; Mazaheri, N.; Rad, A.; Efthymiou, S.; Petersen, A.; Dehghani, M.; Pehlivan, D.; Partlow, J.N.; et al. Homozygous Missense Variants in NTNG2, Encoding a Presynaptic Netrin-G2 Adhesion Protein, Lead to a Distinct Neurodevelopmental Disorder. *Am. J. Hum. Genet.* **2019**, *105*, 1048–1056. [CrossRef]

71. Lui, N.C.; Tam, W.Y.; Gao, C.; Huang, J.D.; Wang, C.C.; Jiang, L.; Yung, W.H.; Kwan, K.M. Lhx1/5 control dendritogenesis and spine morphogenesis of Purkinje cells via regulation of Espin. *Nat. Commun.* **2017**, *8*, 15079. [CrossRef]

72. Tzeng, S.F.; de Vellis, J. Id1, Id2, and Id3 gene expression in neural cells during development. *Glia* **1998**, *24*, 372–381. [CrossRef]

73. Blomfield, I.M.; Rocamonde, B.; Masdeu, M.D.M.; Mulugeta, E.; Vaga, S.; van den Berg, D.L.; Huillard, E.; Guillemot, F.; Urban, N. Id4 promotes the elimination of the pro-activation factor Ascl1 to maintain quiescence of adult hippocampal stem cells. *eLife* **2019**, *8*, e48561. [CrossRef]

74. Havrda, M.C.; Harris, B.T.; Mantani, A.; Ward, N.M.; Paolella, B.R.; Cuzon, V.C.; Yeh, H.H.; Israel, M.A. Id2 is required for specification of dopaminergic neurons during adult olfactory neurogenesis. *J. Neurosci.* **2008**, *28*, 14074–14086. [CrossRef]

75. Hill, C.S. Transcriptional Control by the SMADs. *Cold Spring Harb. Perspect. Biol.* **2016**, *8*, a022079. [CrossRef]

76. Higashi, Y.; Maruhashi, M.; Nelles, L.; Van de Putte, T.; Verschueren, K.; Miyoshi, T.; Yoshimoto, A.; Kondoh, H.; Huylebroeck, D. Generation of the floxed allele of the SIP1 (Smad-interacting protein 1) gene for Cre-mediated conditional knockout in the mouse. *Genesis* **2002**, *32*, 82–84. [CrossRef]

77. Van Helden, M.J.; Goossens, S.; Daussy, C.; Mathieu, A.L.; Faure, F.; Marcais, A.; Vandamme, N.; Farla, N.; Mayol, K.; Viel, S.; et al. Terminal NK cell maturation is controlled by concerted actions of T-bet and Zeb2 and is essential for melanoma rejection. *J. Exp. Med.* **2015**, *212*, 2015–2025. [CrossRef]

78. Guan, J.; Liu, P.; Wang, A.; Wang, B. Long noncoding RNA ZEB2AS1 affects cell proliferation and apoptosis via the miR1225p/PLK1 axis in acute myeloid leukemia. *Int. J. Mol. Med.* **2020**, *46*, 1490–1500. [CrossRef]

79. Jiang, Y.; Wu, K.; Cao, W.; Xu, Q.; Wang, X.; Qin, X.; Wang, X.; Li, Y.; Zhang, J.; Chen, W. Long noncoding RNA KTN1-AS1 promotes head and neck squamous cell carcinoma cell epithelial-mesenchymal transition by targeting miR-153-3p. *Epigenomics* **2020**, *12*, 487–505. [CrossRef]

80. Yao, H.; Hou, G.; Wang, Q.Y.; Xu, W.B.; Zhao, H.Q.; Xu, Y.C. LncRNA SPRY4IT1 promotes progression of osteosarcoma by regulating ZEB1 and ZEB2 expression through sponging of miR101 activity. *Int. J. Oncol.* **2020**, *56*, 85–100. [CrossRef]

81. Cheng, H.; Zhao, H.; Xiao, X.; Huang, Q.; Zeng, W.; Tian, B.; Ma, T.; Lu, D.; Jin, Y.; Li, Y. Long Non-coding RNA MALAT1 Upregulates ZEB2 Expression to Promote Malignant Progression of Glioma by Attenuating miR-124. *Mol. Neurobiol.* **2021**, *58*, 1006–1016. [CrossRef]

82. Brabletz, S.; Brabletz, T. The ZEB/miR-200 feedback loop–a motor of cellular plasticity in development and cancer? *EMBO Rep.* **2010**, *11*, 670–677. [CrossRef]

83. Exposito-Villen, A.; Aránega, E.A.; Franco, D. Functional Role of Non-Coding RNAs during Epithelial-To-Mesenchymal Transition. *Noncoding RNA* **2018**, *4*, 14. [CrossRef]

84. Gregory, P.A.; Bracken, C.P.; Bert, A.G.; Goodall, G.J. MicroRNAs as regulators of epithelial-mesenchymal transition. *Cell Cycle* **2008**, *7*, 3112–3118. [CrossRef]

85. Guan, H.; Liang, W.; Xie, Z.; Li, H.; Liu, J.; Liu, L.; Xiu, L.; Li, Y. Down-regulation of miR-144 promotes thyroid cancer cell invasion by targeting ZEB1 and ZEB2. *Endocrine* **2015**, *48*, 566–574. [CrossRef]

86. Pan, Y.; Zhang, J.; Fu, H.; Shen, L. miR-144 functions as a tumor suppressor in breast cancer through inhibiting ZEB1/2-mediated epithelial mesenchymal transition process. *Onco Targets Ther.* **2016**, *9*, 6247–6255. [CrossRef]

87. Nourmohammadi, B.; Tafsiri, E.; Rahimi, A.; Nourmohammadi, Z.; Daneshvar Kakhaki, A.; Cho, W.; Karimipoor, M. Expression of miR-9 and miR-200c, ZEB1, ZEB2 and E-cadherin in Non-Small Cell Lung Cancers in Iran. *Asian Pac. J. Cancer Prev.* **2019**, *20*, 1633–1639. [CrossRef]

88. Wahab, N.A.; Othman, Z.; Nasri, N.W.M.; Mokhtar, M.H.; Ibrahim, S.F.; Hamid, A.A.; Raja Ali, R.A.; Mokhtar, N.M. Inhibition of miR-141 and miR-200a Increase DLC-1 and ZEB2 Expression, Enhance Migration and Invasion in Metastatic Serous Ovarian Cancer. *Int. J. Environ. Res. Public Health* **2020**, *17*, 2766. [CrossRef]

89. Birkhoff, J.C.; Brouwer, R.W.W.; Kolovos, P.; Korporaal, A.L.; Bermejo-Santos, A.; Boltsis, I.; Nowosad, K.; van den Hout, M.; Grosveld, F.G.; van IJcken, W.F.J.; et al. Targeted chromatin conformation analysis identifies novel distal neural enhancers of ZEB2 in pluripotent stem cell differentiation. *Hum. Mol. Genet.* **2020**, *29*, 2535–2550. [CrossRef]

90. Van Grunsven, L.A.; Papin, C.; Avalosse, B.; Opdecamp, K.; Huylebroeck, D.; Smith, J.C.; Bellefroid, E.J. XSIP1, a Xenopus zinc finger/homeodomain encoding gene highly expressed during early neural development. *Mech. Dev.* **2000**, *94*, 189–193. [CrossRef]

91. Van de Putte, T.; Maruhashi, M.; Francis, A.; Nelles, L.; Kondoh, H.; Huylebroeck, D.; Higashi, Y. Mice lacking ZFHX1B, the gene that codes for Smad-interacting protein-1, reveal a role for multiple neural crest cell defects in the etiology of Hirschsprung disease-mental retardation syndrome. *Am. J. Hum. Genet.* **2003**, *72*, 465–470. [CrossRef]

92. Takagi, T.; Nishizaki, Y.; Matsui, F.; Wakamatsu, N.; Higashi, Y. De novo inbred heterozygous Zeb2/Sip1 mutant mice uniquely generated by germ-line conditional knockout exhibit craniofacial, callosal and behavioral defects associated with Mowat-Wilson syndrome. *Hum. Mol. Genet.* **2015**, *24*, 6390–6402. [CrossRef]

93. Hegarty, S.V.; Wyatt, S.L.; Howard, L.; Stappers, E.; Huylebroeck, D.; Sullivan, A.M.; O'Keeffe, G.W. Zeb2 is a negative regulator of midbrain dopaminergic axon growth and target innervation. *Sci. Rep.* **2017**, *7*, 8568. [CrossRef]

94. Hirabayashi, Y.; Itoh, Y.; Tabata, H.; Nakajima, K.; Akiyama, T.; Masuyama, N.; Gotoh, Y. The Wnt/β-catenin pathway directs neuronal differentiation of cortical neural precursor cells. *Development* **2004**, *131*, 2791–2801. [CrossRef]

95. Slawny, N.A.; O'Shea, K.S. Dynamic changes in Wnt signalling are required for neuronal differentiation of mouse embryonic stem cells. *Mol. Cell. Neurosci.* **2011**, *48*, 205–216. [CrossRef]

96. Li, M.; Santpere, G.; Imamura Kawasawa, Y.; Evgrafov, O.V.; Gulden, F.O.; Pochareddy, S.; Sunkin, S.M.; Li, Z.; Shin, Y.; Zhu, Y.; et al. Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science* **2018**, *362*, eaat7615. [CrossRef]

97. Li, H.; Zhu, Y.; Morozov, Y.M.; Chen, X.; Page, S.C.; Rannals, M.D.; Maher, B.J.; Rakic, P. Disruption of TCF4 regulatory networks leads to abnormal cortical development and mental disabilities. *Mol. Psychiatry* **2019**, *24*, 1235–1246. [CrossRef]

98. Mesman, S.; Bakker, R.; Smidt, M.P. Tcf4 is required for correct brain development during embryogenesis. *Mol. Cell. Neurosci.* **2020**, *106*, 103502. [CrossRef]

99. BabuRajendran, N.; Palasingam, P.; Narasimhan, K.; Sun, W.; Prabhakar, S.; Jauch, R.; Kolatkar, P.R. Structure of Smad1 MH1/DNA complex reveals distinctive rearrangements of BMP and TGF-β effectors. *Nucleic Acids Res.* **2010**, *38*, 3477–3488. [CrossRef]

100. Morikawa, M.; Koinuma, D.; Tsutsumi, S.; Vasilaki, E.; Kanki, Y.; Heldin, C.-H.; Aburatani, H.; Miyazono, K. ChIP-seq reveals cell type-specific binding patterns of BMP-specific Smads and a novel binding motif. *Nucleic Acids Res.* **2011**, *39*, 8712–8727. [CrossRef]

101. Liu, S.; Long, J.; Yuan, B.; Zheng, M.; Xiao, M.; Xu, J.; Lin, X.; Feng, X.H. SUMO Modification Reverses Inhibitory Effects of Smad Nuclear Interacting Protein-1 in TGF-β Responses. *J. Biol. Chem.* **2016**, *291*, 24418–24430. [CrossRef]

102. Xu, P.; Lin, X.; Feng, X.H. Posttranslational Regulation of Smads. *Cold Spring Harb. Perspect. Biol.* **2016**, *8*, a022087. [CrossRef]

103. Kim, Y.J.; Kang, M.J.; Kim, E.; Kweon, T.H.; Park, Y.S.; Ji, S.; Yang, W.H.; Yi, E.C.; Cho, J.W. O-GlcNAc stabilizes SMAD4 by inhibiting GSK-3beta-mediated proteasomal degradation. *Sci. Rep.* **2020**, *10*, 19908. [CrossRef]

104. Lin, X.; Wang, Y.; Jiang, Y.; Xu, M.; Pang, Q.; Sun, J.; Yu, Y.; Shen, Z.; Lei, R.; Xu, J. Sumoylation enhances the activity of the TGF-β/SMAD and HIF-1 signalling pathways in keloids. *Life Sci.* **2020**, *255*, 117859. [CrossRef]

105. McKinsey, G.L.; Lindtner, S.; Trzcinski, B.; Visel, A.; Pennacchio, L.A.; Huylebroeck, D.; Higashi, Y.; Rubenstein, J.L. Dlx1&2-dependent expression of Zfhx1b (Sip1, Zeb2) regulates the fate switch between cortical and striatal interneurons. *Neuron* **2013**, *77*, 83–98. [CrossRef]

106. De Haan, W.; Dheedene, W.; Apelt, K.; Décombas-Deschamps, S.; Vinckier, S.; Verhulst, S.; Conidi, A.; Deffieux, T.; Staring, M.W.; Vandervoort, P.; et al. Endothelial Zeb2 preserves the hepatic angioarchitecture and protects against liver fibrosis. *Cardiovasc. Res.* **2022**, *118*, 1262–1275. [CrossRef]