

# Evidence for a Potential Common Gene Network of Smith-Magenis and Potocki-Lupski Syndromes, DAND (DEAF1-Associated Neurodevelopmental Disorder) and ZEB1-Associated Neurodevelopment Disorder

Li Chen (✉ [chenli2008@fudan.edu.cn](mailto:chenli2008@fudan.edu.cn))

Fudan University <https://orcid.org/0000-0002-2429-7916>

Philip J Jensik

Southern Illinois University Carbondale

Xi Yuan

Fudan University

Muyin Liu

Fudan University

David Saffen

Fudan University

Sarah H Elsea

Baylor College of Medicine

---

## Research

**Keywords:** deformed epidermal autoregulatory factor-1 (DEAF1), retinoic acid induced-1 (RAI1), zinc finger E-box binding homeobox 1 (ZEB1), DEAF1-associated neurodevelopmental disorder (DAND), Smith-Magenis syndrome (SMS), Potocki-Lupski Syndrome (PTLS)

**Posted Date:** December 29th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-135772/v1>

**License:** © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

**Evidence for a potential common gene network of Smith-Magenis and Potocki-Lupski syndromes, DAND (DEAF1-associated neurodevelopmental disorder) and ZEB1-associated neurodevelopment disorder**

Li Chen<sup>\*,1</sup>, Philip J Jensik<sup>2</sup>, Xi Yuan<sup>1</sup>, Muyin Liu<sup>1</sup>, David Saffen<sup>1,3,4</sup>, and Sarah H Elsea<sup>\*,5</sup>

<sup>1</sup>Department of Cellular and Genetic Medicine, School of Basic Medical Sciences, Fudan University, Shanghai 200032, China

<sup>2</sup>Department of Physiology, Southern Illinois University, Carbondale, IL USA

<sup>3</sup>Institutes of Brain Science, Fudan University, Shanghai, China.

<sup>4</sup>State Key Laboratory for Medical Neurobiology, Fudan University, Shanghai, China.

<sup>5</sup>Department of Molecular and Human Genetics, Baylor College of Medicine, Houston, TX, USA

\*Correspondence should be addressed to:

Sarah H. Elsea, Ph.D, FACMG, Department of Molecular and Human Genetics, One Baylor Plaza, BCM225, Baylor College of Medicine, Houston, TX 77030, USA, Email: [elsea@bcm.edu](mailto:elsea@bcm.edu)

Li Chen, Ph.D, Department of Cellular and Genetic Medicine School of Basic Medical Sciences, Fudan University, Shanghai 200032, China, Email: [chenli2008@fudan.edu.cn](mailto:chenli2008@fudan.edu.cn)

## **Abstract**

**Background:** Neuropsychiatric disorders are highly heterogeneous and often display overlapping phenotypes, suggesting possible defects in common genetic networks. We previously reported pathogenic variants in deformed epidermal autoregulatory factor-1 (DEAF1) that contribute to *DEAF1*-associated neurodevelopmental disorder (DAND) and predicted that DEAF1 regulates expression of *retinoic acid induced 1 (RAI1)*, the causative gene of Smith-Magenis syndrome (SMS) and Potocki-Lupski Syndrome (PTLS), through a putative DEAF1 binding sequence (DBS) in *RAI1* intron 2, suggesting a common genetic network of these disorders that needs further study.

**Methods:** In this study, we tested the DEAF1 binding and transcriptional regulation of *RAI1* using luciferase reporter gene assay, EMSA and CHIP. We explored the upstream regulator of *DEAF1* using bioinformatics analysis, qPCR and CHIP, and generated a potential network involving these genes based on patients' phenotype and bioinformatics analysis.

**Results:** We demonstrated that DEAF1 binds to the *RAI1* DBS and regulates *RAI1* expression *in vitro*, and a pathogenic variant in the *DEAF1* SAND domain was defective in binding the *RAI1* DBS. We also obtained evidence that *DEAF1* expression is regulated by the transcription factor ZEB1 via predicted binding sites in *DEAF1* intron 1. Microdeletions and variants in *ZEB1* have been implicated in brain developmental disorders, including intellectual disability (ID), autism, agenesis of corpus callosum (ACC) and corneal dystrophy. Patients harboring deletions or pathogenic variants of *ZEB1*, *DEAF1* or *RAI1* display partially overlapped symptoms, including ID, autistic features, speech impairment, developmental delay and dysmorphologies.

**Conclusions:** Together, these results provide evidence for a common molecular network in *ZEB1*-associated neurodevelopment disorder, DAND, SMS and PTLs.

**Keywords:** deformed epidermal autoregulatory factor-1 (DEAF1); retinoic acid induced-1 (RAI1); zinc finger E-box binding homeobox 1 (ZEB1); *DEAF1*-associated neurodevelopmental disorder (DAND); Smith-Magenis syndrome (SMS); Potocki-Lupski Syndrome (PTLS)

## Introduction

Neuropsychiatric disorders are highly heterogeneous, often co-occur with developmental delay, autism and intellectual disability (ID) [1], suggesting the involvement of common genetic networks. We previously reported that pathogenic variants of deformed epidermal autoregulatory factor-1 (DEAF1) with impaired transcriptional functions contribute to a spectrum of symptoms defined as *DEAF1*-associated neurodevelopmental disorder (DAND) [2]. DEAF1 is a transcriptional factor important in early embryonic patterning in *Drosophila* [3], neural tube development and neurobehaviors in mouse [4, 5], and in human, several pathogenic variants in DEAF1 have been reported to be associated with ID, speech impairment, autism, and developmental delay [4, 6-16], which together comprising the DAND phenotypic spectrum.

DEAF1 activates or represses expression of various genes via binding to promoter sequences containing TTCG motifs [17]. Recently, we identified a possible DEAF1 binding sequence (DBS) in intron 2 of *retinoic acid induced 1 (RAI1)* [18]. *RAI1* is a dosage-sensitive gene for which deletion or mutation causes Smith-Magenis syndrome (SMS; OMIM 182290) and duplication causes Potocki-Lupski syndrome (PTLS; OMIM 610883). Symptoms of SMS include ID, autism, speech and motor delays, stereotypic, aggressive and self-injurious behaviors, temper tantrums, hypotonia, sleep disturbance, obesity, a high threshold for pain, and craniofacial and skeletal anomalies [19], while PTLS symptoms include ID, autism, speech and motor delays, infantile hypotonia, failure to thrive, and cardiovascular malformations [20]. *RAI1* is a transcriptional factor important in early embryonic development of the nervous system, regulating the expression of genes involved in development, lipid

metabolism, circadian rhythms, and behavior [21-24], and functions as a chromatin modifier, recognizing specific histone modifications and recruiting additional transcription factors [25, 26].

In this study, we demonstrate that DEAF1 binds and activates *RAI1* expression, pathogenic variant in *DEAF1* results in impaired binding to *RAI1*. We also identify the transcription factor ZEB1 as a potential upstream regulator of *DEAF1*. ZEB1 plays an important role in the neuronal development and epithelium differentiation [27], and has been implicated with autism, intellectual disability (ID), agenesis of corpus callosum abnormality (ACC) and corneal dystrophy. The phenotypic overlaps among the above neuropsychiatric disorders caused by sequence variants, deletion or duplication of these genes combined with network analysis provide further evidence for a common molecular network in these neuropsychiatric disorders.

## **Methods and Materials**

### **Cell culture**

Human embryonic kidney-derived (HEK293T) cells were cultured in Eagle's Minimum Essential Medium (MEM, Invitrogen) containing 10% (v/v) FBS, 2 mM L-glutamine, 1% non-essential amino acids (NEAA, Invitrogen) and 1X antibiotic-antimycotic (Invitrogen) at 37 °C under 5% CO<sub>2</sub>. Human neuroblastoma-derived SH-SY5Y cells were cultured in 1:1 mixtures of Dulbecco's Modified Eagle's Minimum Essential Medium (DMEM) and F12 Medium with 10% (v/v) fetal bovine serum (FBS) and 2 mM L-glutamine and were maintained at 37 °C in a 5% CO<sub>2</sub> incubator.

## Plasmid construction

Construction of the wildtype *DEAF1* expression plasmid in pcDNA3 (*DEAF1*-pcDNA3) and site-directed mutagenesis to introduce the c.737G>C (p.Arg246Thr) mutation in *DEAF1* expression plasmids have been previously described [6]. A DNA segment encoding the *DEAF1* open reading frame (GenBank accession number NM\_021008) with an HA peptide inserted at the N-terminal after ATG translation start site was cloned into pcDNA3.1 (Invitrogen) to generate the HA-*DEAF1* plasmid.

Luciferase reporter plasmids *RAI1*-Luc plasmid was generated by PCR amplified the *DEAF1* binding site segment (DBS, 81 bp) in human *RAI1* intron 2, digested with *KpnI* and *XhoI* (New England Biolabs), gel extracted using a Zymoclean gel DNA recovery kit (Zymo Research), and ligated into pGL4.23 (Promega Corporation, luc2/minP) using T4 DNA ligase (Promega Corporation). *RAI1*-DBS plasmid was constructed by cloning the *RAI1* DBS segment (407 bp) into pGL4.23 vector.

All inserts were confirmed by Sanger sequencing. Sequences of oligonucleotide primers used for the construction of plasmids are listed in **Table S1**.

## Cell transfection

Approximately  $5 \times 10^5$  HEK293T cells were plated in 2 ml growth medium without antibiotics in each well of 6-well culture plates 24 h prior to transfection. A total of 4  $\mu$ g total plasmid DNA was transfected into the cells in each well using 5  $\mu$ l of Lipofectamine 2000 (Invitrogen). pUC19 plasmid was used as filler DNA to equalize the transfection efficiencies

across all groups. Both the plasmids and Lipofectamine 2000 were diluted in 250  $\mu$ l OptiMEM Reduced Serum Medium (Invitrogen), mixed and incubated for 20 min at room temperature before addition to wells. Following transfection, the cells were incubated at 37°C under 5% CO<sub>2</sub> for 6 hours, at which time the culture medium was replaced. The cells were then incubated under the same conditions for an additional 24 h.

For siRNA transfection, 6 pmol siRNA (GenePharma Co. Ltd) were diluted in 100  $\mu$ l OptiMEM Reduced Serum Medium (Invitrogen) for each well, after which 1  $\mu$ l of Lipofectamine RNAiMAX (Invitrogen) was added and the mixture incubated at room temperature for 10-20 minutes. The mixture was then added to SH-SY5Y cells at a final concentration of 10 nM. Sequences of *ZEB1* and control siRNAs are listed in **Table S1**.

### **Luciferase assays**

Transcription assays using the p*DEAF1*-pro-luciferase construct have been previously described [6]. Briefly, cells transfected with *RAI1*-Luc luciferase reporter plasmids were washed with 2 ml Dulbecco's phosphate-buffered saline (DPBS, Invitrogen), lysed, and scraped with 500  $\mu$ l of Tropix lysis solution (ABI) in each well. Lysates were collected and centrifuged at 12,000 RPM for 2 min to pellet cellular debris. For each sample, 50  $\mu$ l supernatant was transferred to each of the four wells in a 96-well luminometer plate. Two wells were treated with 70  $\mu$ l diluted Tropix Galacton substrate (1:100, Tropix Galacton reaction buffer diluent with Tropix Galacton substrate, ABI), incubated for 30 min, after which 100  $\mu$ l of Tropix light emission accelerator (ABI) was added. The other two wells were treated with 100  $\mu$ l Steady-Glo luciferase substrate (Promega Corporation). Synergy 2 (Biotek) was

used to quantify the luminescence. Relative luciferase activity was calculated by dividing the average number of light units from the wells containing the Steady-Glo luciferase substrate (Promega Corporation) by the average number of light units from the wells containing the Galacton substrate and Accelerator. Wells containing pUC19, psvb-Gal, and *RAI1* luciferase constructs were used as baseline luciferase activity and normalized to 1. Each experiment was performed in triplicate biological replicates and p-values derived from ANOVA and *post hoc* Tukey's tests.

### **Quantitative polymerase chain reaction (qPCR)**

The cDNA synthesis and quantitative PCR were carried out as described previously [18]. Sequences of *DEAF1*, *ZEB1* and *GAPDH* primers are listed in **Table S1**. Gene expression levels of control samples were normalized to *GAPDH* and set to 1.

### **Electrophoretic mobility shift assays (EMSA)**

Epitope FLAG-tagged WT protein was purified from transfected HEK293T cells and nuclear extracts was prepared as previously described [6]. EMSAs were performed using double-stranded (ds) DNA probes containing the putative DBS within *RAI1* exon 2, a DEAF1 binding site in *DEAF1* 5' region [28], putative ZEB1 binding sites within *DEAF1* 5' region. Sequences of the synthetic oligonucleotide DNA probes used in these experiments are listed in **Table S1**.

### **Chromatin immunoprecipitation (ChIP)**

ChIP assays were performed following the Cold Spring Harbor protocol with minor modifications [29]. Briefly, HEK293T cells co-transfected with a HA-DEAF1 expression plasmid and *RAI1* DBS plasmid were fixed in 1% formaldehyde and collected 48 h after transfection. Chromatin was digested using micrococcal nuclease (Pierce), as recommended by the manufacturer. Each immunoprecipitation reaction mix contained 50 µg chromatin and 5% of this amount (2.5 µg) chromatin was collected separately to quantify the amount of input target DNA present in each immunoprecipitation reaction mix. Five µg mouse anti-HA-Tag (6E2) monoclonal antibody (Cell Signaling Technologies, #2367; IgG1 isotype) was used for DEAF1 immunoprecipitation, 5 µg ZEB1 monoclonal antibody (Sigma, HPA027524) was used for DEAF1 immunoprecipitation. Immunoprecipitation with 5 µg mouse IgG (Sigma, I5318) used as a negative control. DNA was purified using MinElute reaction cleanup kits (Qiagen) and subjected to PCR using primers designed to amplify the DBS-containing fragment in *RAI1* 5' region. PCR products were quantified using qPCR and quantitative scanning of the bands in the gels using ImageJ program. The sequences of primers used to amplify the *RAI1* DBS are listed in **Table S1**.

### **Bioinformatics and statistical analysis**

DEAF1 consensus binding sequences were predicted using the JASPAR database (<http://jaspar.genereg.net/>). Genome-wide DEAF1 consensus binding sequences were analyzed using Regulatory Sequence Analysis Tools (RSAT, <http://www.rsat.eu/>). Gene function was annotated and classified using Database for Annotation, Visualization, and Integrated Discovery (DAVID, <http://david.abcc.ncifcrf.gov/>). OMIM (Online Mendelian

Inheritance in Man, <http://www.omim.org/>) and National Institutes of Health (NIH) Genetic Association Database (<http://geneticassociationdb.nih.gov/>) were used to determine gene-phenotype relationships in human diseases. SNP eQTLs were analyzed using the LIBD database (<http://eqtl.brainseq.org>) and functional SNPs scores were obtained from the Regulome database ([www.regulomedb.org](http://www.regulomedb.org)). The pairwise linkage disequilibrium (LD) constants  $D'$  and  $R^2$  between each two SNPs were calculated using LDlink database (<https://ldlink.nci.nih.gov/>). Histone III lysine 27 acetylation and DNase sensitive clusters near eQTLs were analyzed using UCSC Genome Browser (<http://genome.ucsc.edu/>). DNA motif is analyzed using MEME suite ([meme-suite.org/](http://meme-suite.org/)). Gene co-expression and network analysis was performed using String database (<https://version11.string-db.org>) and HumanBase database (<https://hb.flatironinstitute.org/>). Plots and statistical analysis were generated using GraphPad Prism 6.

## Results

### DEAF1 is a potential upstream regulator of *RAI1*

DEAF1 is a transcriptional factor recently implicated in ID, autism, and development delay [2]. It binds to specific sequences within gene regulatory regions and stimulates or represses mRNA expression [2]. To identify novel candidate target genes, we performed a genome-wide analysis of potential DEAF1 response elements in human gene transcriptional regions (~45 kb) in the Regulatory Sequence Analysis Tools (RSAT) database [30] using a preferred degenerate DEAF1 binding consensus sequence, 5' YtcggNNNNYtccg 3' [17, 28], and identified 318 candidate genes, representing a subset of possible DEAF1-regulated

genes (**Table S2**). Among these, Database for Annotation, Visualization, and Integrated Discovery (DAVID) database identified 66 genes related to human diseases (**Table S3**). Further analyses revealed that 15 genes are implicated in psychiatric, neurological, and developmental abnormalities related to DAND phenotypes, including *RAI1*, the causative gene for SMS and PTL5, based upon information from the Online Mendelian Inheritance in Man database (OMIM) (<http://www.omim.org/>) and the NIH Genetic Association Database (<http://geneticassociationdb.nih.gov/>) (**Figure S1, Table S4**) [31, 32]. We have previously provided evidence that DEAF1 regulates *RAI1* expression in human brain via a putative DEAF1 binding sequence (DBS) located within *RAI1* intron 2 [18], suggesting that DEAF1 and *RAI1* may function within a common molecular network.

#### **DEAF1 binds and regulates *RAI1* in vitro**

To obtain evidence for DEAF1 regulation of *RAI1* expression, we measured luciferase activity in HEK293T cells transfected with an expression plasmid (*RAI1*-Luc) containing a fragment from *RAI1* intron 2 that includes the putative DBS positioned upstream from a minimal promoter firefly luciferase. A statistically significant increase in luciferase activity was observed in cells co-transfected with a WT DEAF1 expression plasmid, compared to cells transfected with *RAI1*-luc alone. (**Figure 1A**).

To determine whether DEAF1 directly binds to the *RAI1* intron 2 DBS, EMSAs were performed using a double-stranded (ds) DNA probe containing this sequence. A previously published dsDNA probe containing a DEAF1 binding sequence in the human DEAF1

promoter [33] was used as a positive control. As predicted, purified FLAG-tagged WT DEAF1 protein DEAF1 bound to the putative DBS in *RAI1* intron 2 (**Figure 1B**).

To confirm the binding of DEAF1 to *RAI1* intron 2, chromatin immunoprecipitation (ChIP) assays were performed using nuclear extracts of HEK293T cells co-transfected with HA-DEAF1 expression and *RAI1*-DBS plasmid. DNA fragments co-immunoprecipitated with anti-HA antibodies were amplified by PCR (**Figure 1C**). PCR products were quantified by quantitative scanning of the bands in the gels using ImageJ program (**Figure 1D**).

### **DEAF1 variants are impaired in DNA binding of *RAI1***

DEAF1 comprises a N-terminal DNA binding domain containing Sp-100, AIRE, NucP41/75, and DEAF1 (SAND) motifs, a zinc-binding motif, a nuclear localization signal (NLS), a leucine-rich nuclear export signal (NES), and a C-terminal cysteine-rich protein interaction domain termed Myeloid translocation protein 8, Nervy, and DEAF1 (MYND) domain [4]. A SAND domain c.737G>C (p.Arg246Thr) has been reported to associate with ID, speech impairment, autism, and developmental delay [6]. We tested the influence of the SAND domain *DEAF1* variant c.737G>C (p.Arg246Thr) on binding to the double-stranded DNA probe containing the *RAI1* intron2 DBS. The results show that the c.737G>C (p.Arg246Thr) variant impairs the ability of DEAF1 to bind to the *RAI1* intron 2 DBS (**Figure S2**).

### **ZEB1 regulates *DEAF1* expression**

To further explore the genes involved in the *DEAF1-RAI1* network, we used BrainSeq

Consortium database (<http://eqtl.brainseq.org>) to look for potential upstream regulators of *DEAF1*. We identified 106 expression quantitative trait loci (eQTLs) in the dorsal lateral prefrontal cortex (DLPFC) of a 412 subjects cohort, and 66 eQTLs in a 237 subjects cohort which are significantly associated with *DEAF1* mRNA expression based on RNA sequencing (RNA-seq) and genotype data in BrainSeq Consortium database. The eQTL modeling in the database was adjusted for multiple covariates, including sex, age (>13 years), and ancestry (multidimensional scaling components). Significant eQTLs were defined as SNP/*DEAF1* mRNA expression pairs with a false discovery rate (FDR) less than 1% that replicated with the same allelic direction and  $p < 0.01$  in two independent sets of DLPFC samples from the CommonMind Consortium (<https://www.nimhgenetics.org/resources/commonmind>) and the GTEx project (<https://www.gtexportal.org>) databases. Among the 52 overlapped significant eQTLs identified from the two cohorts, we screened for SNPs that had Regulome database scores  $\geq 2b$ , which denotes evidence for transcription factors (TF) binding, a known TF motif, a DNase footprint and location within a DNase-sensitive site. Three SNPs (rs10751662, rs6597991, rs71464105) had Regulome database scores of 2b, suggesting possible transcriptional factors that regulate *DEAF1* mRNA expression (**Table S5**). Low- and high-expression alleles of these SNPs for *DEAF1* mRNA expression in human dorsolateral prefrontal cortex (DLPF) are shown in **Figure S3**. Pairwise linkage disequilibrium (LD) coefficients  $D'$  and  $r^2$  calculated using the LDlink database (<https://ldlink.nci.nih.gov/>) showed that all three SNPs are in high LD in the Caucasian population (**Table S6**). We compared the locations of histone III lysine<sup>27</sup> acetylation (H3K27Ac) and DNase sensitive clusters near these eQTLs using UCSC Genome Browser (<http://genome.ucsc.edu/>) and found strong

signals for H3K27Ac and DNase sensitive clusters near SNP rs6597991, consistent with chromatin remodeling and active transcription in this region. Based on the Regulome database, SNP rs6597991 is located within a predicted ZEB1 binding site, suggesting that the ZEB1 transcription factor may contribute to the regulation of *DEAF1* transcription (**Figure 2**).

ZEB1 (zinc finger E-box binding homeobox 1; OMIM 189909) is strongly expressed in early development of the central nervous system (CNS) [27], and has been implicated in patients with autism, intellectual disability (ID), agenesis of corpus callosum abnormality (ACC) [63, 65], posterior polymorphous corneal dystrophy 3 (PPCD 3, OMIM 609141) [34, 35] and Fuchs' endothelial corneal dystrophy (FECD6, OMIM 613270) [36]. ZEB1 recognizes a single or bipartite motif comprising one CACCT sequence and one CACCTG E-box on its target gene with various orientations and spacing [37]. We analyzed the sequence in the *DEAF1* promoter region using the Find Individual Motif Occurrences (FIMO) program in the MEME suite ([meme-suite.org/](http://meme-suite.org/)), and identified a potential bipartite ZEB1 binding sequence, CATGT and CAGGTG (ACATG and CACCTG in the complementary DNA strand) separated by 30 nt within *DEAF1* intron 1, with SNP rs6597991 located within the CAGG(T/C)G sequence (**Table S7**). DNA binding activity to a dsDNA probe containing the putative rs6597991-centered ZEB1 binding site in *DEAF1* intron 1 was detected using HEK293T nuclear extracts in EMSAs (**Figure 3A**). We carried out chromatin immunoprecipitation (ChIP) assays in HEK293T cells using mouse monoclonal antibodies against human ZEB1 and PCR to amplified a 118 bp DNA fragment centered on rs6597991 (**Figure 3B**). Quantitative PCR confirmed statistically significant differences between the amounts of targeted DNA

precipitated by anti-ZEB1 antibodies or by IgG (**Figure 3C**).

We noticed that the sequence CATGT is less conserved than the canonical AGGTG sequence in the middle two nucleotides. Scanning a wider region in the 5' regulatory sequence of *DEAF1* identified a cluster of potential ZEB1 binding sites within *DEAF1* intron 1, located ~900 bp upstream of SNP rs6597991 (chr11:693,829-695,048, 1219 bp, **Figure 4**). Using the FIMO program in MEME suite, we identified clusters of possible bipartite ZEB1 binding sites with more conservation in the consensus sequence [37]. E-box sequences in this region predicted by the FIMO program are listed in **Table S8**. We carried out CHIP assays and confirmed that ZEB1 binds to this 5' region of the *DEAF1* gene (**Figure 5A**). Quantitative PCR detected significant differences between immunoprecipitation of the target DNA fragment compared to chromosomal DNA immunoprecipitated by IgG. Examination of the literature also confirmed that these sites were detected in a genome-wide, CHIP-based scan of potential ZEB1 binding sites [38]. And an independent study from the ENCODE Transcription Factor Targets dataset by CHIP-seq also confirmed that ZEB1 binds to *DEAF1* [39].

Knock down of *ZEB1* mRNA expression in neuronal SHSY5Y cells using *ZEB1* siRNA increased *DEAF1* mRNA expression by ~50%, compared to cells treated with control siRNA (**Figure 6**). Because ZEB1 encodes a zinc finger transcription factor that is likely to play a role in transcriptional repression of its target genes when bound to its consensus binding site [33], our data suggest that ZEB1 may function as a potential upstream repressor of *DEAF1* in neuronal cells by binding to multiple bipartite ZEB1 binding sites within *DEAF1* intron 1.

## **RAI1, DEAF1 and ZEB1 may function in a network common in neuropsychiatric disorders**

ZEB1 has been reported to function within the TGF- $\beta$ 1-induced epithelial-to-mesenchymal transition downstream from the AKT/mTORC1/GSK3 $\beta$  pathway [40, 41], an important intracellular signaling pathway involved in cell proliferation and neuropsychiatric disorders [42, 43].

To identify possible networks common to RAI1, DEAF1, ZEB1 and mTOR, we explored current scientific literature and identified 14 genes, including *MBD5*, *HDAC4*, *CLOCK*, *BDNF*, *GSK3 $\beta$* , *LDB1*, *LMO4*, *CHD8*, *CTNNB1*, *5-HTR1A*, *AKT1*, *GSK3 $\alpha$* , *RPTOR* and *AKT3*, that associate with our four input genes during neurogenesis and development. Associations between pairs of genes are presented using the Sting database (**Figure 7**). Enrichments for these genes in diseases such as ID, developmental disorders and ovarian cancer, biological process such as regulation of gene expression and binding, CNS development, cell growth and differentiation were demonstrated using the HumanBase database (**Table S9, S10**). It's worth mentioning that all the 18 genes have high confidence with autism, with average confidence of 0.605. Tissue-specific expressions of these genes were demonstrated using the HumanBase database (**Figure S4**). And we further tested if there's tissue-specific gene network by investigating co-expression-based molecular interactions among these genes (**Figure 8**). In neurons, DEAF1, ZEB1 and GSK3 $\beta$  have more high-confidence interactors than other genes inquired, suggesting that these three genes play crucial roles in neuronal growth and differentiation. In all tissues, CTNNB1 and AKT1, but not *DEAF1*, *ZEB1* and *RAI1* have more high-confidence interactors than other genes inquired, implies that ZEB1, DEAF1

and GSK3 $\beta$  comprise a neuronal functional network compared to other tissues, that possibly contributes to psychiatric and development disorders.

Evidence for a common network in neurons suggests that pathogenic variants in these genes may produce overlapping defects in neurogenesis and development. To test this hypothesis, we compared the phenotypes of *RAI1*-associated SMS/PTLS, *DEAF1*-associated DAND and *ZEB1*-associated neurodevelopment disorder (**Table 1**). All of these disorders display mild-to-severe ID, speech impairment, developmental delay, hypotonia, seizures, autistic behaviors, sleep disturbances, brain abnormalities including microcephaly, brachycephaly and macrocephaly, genitourinary abnormalities, craniofacial and skeletal dysmorphisms, hyperactivity and hearing loss. Temper tantrums, aggressive behaviors, abnormalities detected by magnetic resonance imaging (MRI), cardiovascular anomalies, eye hypoplasia, feeding difficulties, digestive and metabolic problems are present in three out of four disorders. Two out of four disorders display abnormal gait, pain insensitivity, abnormalities detected by electroencephalogram (EEG), compulsive and self-injurious behaviors, and recurrent infections. The most frequent common phenotypes observed in SMS, PTLS, and DAND patients include: moderate-to-severe ID (> 91%), speech impairments (>75%), developmental delay (>66%) and autistic behaviors (>60%). The existence of partially overlapping phenotypes supports the hypothesis that *ZEB1*, *DEAF1* and *RAI1* function within common networks and/or interact with each other.

## **Discussion**

In this study, we demonstrated that the transcription factor DEAF1 binds to a DBS within

intron 2 of *RAI1* and stimulates *RAI1* mRNA expression *in vitro*, providing experimental confirmation of previous bioinformatic-based predictions of *RAI1* regulation in human brain [18]. These observations provide the first experimental evidence that DEAF1 and RAI1 function within the same intracellular network. Our study also provides bioinformatic and experimental evidence that *DEAF1* expression is regulated by the transcription factor ZEB1, which is a downstream target for regulation by the kinase mTOR. Together, these observations in combination with the results of previously published studies support the existence of a biological signaling network comprising mTOR, ZEB1, DEAF1 and RAI1.

*RAI1*, *DEAF1* and *ZEB1* are associated with multiple neuropsychiatric disorders with overlapped phenotypes, including ID, autism, developmental delay and malformations, and linked to genes that also play important roles in neuronal development. *RAI1*, a causative gene of SMS and PTL5 showed reduced expression in 2q23.1 deletion syndrome (OMIM 156200) [24], a disorder caused by haploinsufficiency of a histone reader, *MBD5* [44]. Deletion or mutation of *HDAC4*, a histone deacetylase eraser associated with brachydactyly mental retardation syndrome (BDMR, OMIM 600430) and 2q37 deletion syndrome [44], also results in reduced expression of *RAI1* [45]. Reporter studies revealed that RAI1 directly regulates *BDNF* expression, a gene associated with neurodevelopment and behavior problems [22]. RAI1 positively regulates *CLOCK*, a key component of the mammalian circadian oscillator, and was associated with mood problems. Haploinsufficiency of *RAI1* in SMS fibroblasts and in mouse hypothalamus results in dysregulation of multiple circadian genes including *CLOCK* [21, 46].

Whole exome sequencing identified several pathogenic variants in DEAF1 producing a

spectrum of symptoms recently designated *DEAF1*-associated neurodevelopmental disorder (DAND) [5,7-8], and a subset of these symptoms are also observed in SMS/PTLS. The SAND domain of *DEAF1* is crucial for multimerization and DNA-binding as well as protein–protein interactions [47]. It represses the *DEAF1* promoter through autoregulation [48], and recognizes core-binding motifs (TTCGGGNNTTCCGG or flexibly spaced TTCGGN<sub>3-8</sub>TTCGG) in other gene promoters [17, 33]. It also binds with lower affinity to a single half-site motif sequence TTCG in *5-HTR1A* promoters, a gene associated with major depression, anxiety, suicidal tendencies, and panic disorder [49]. We tested the DNA binding affinity of pathogenic *DEAF1* variants, and showed that the SAND variant c.737G>C (p.Arg246Thr) [6, 9, 14] had reduced binding affinity for the putative *DEAF1* consensus sequence in *RAI1*. Exploring how *RAI1* expression is regulated by *DEAF1* helps to elucidate the genetic mechanisms underlying these disorders, and may provide helpful information for identifying potential drug targets for their treatment or prevention.

We also showed evidence that *DEAF1* expression is regulated by the transcription factor ZEB1. Patients harboring microdeletions encompassing *ZEB1* or intragenic *ZEB1* deletions show agenesis of corpus callosum (ACC), autism, ID, and corneal dystrophy. Although the microdeletions in three of the patients also include *WAC* (WW Domain Containing Adaptor With Coiled-Coil) [50], a gene associated with DeSanto-Shinawi syndrome (DESSH, OMIM 616708), which is characterized by developmental delay and ocular abnormalities, but *WAC* was not associated with ACC observed in all three patients. And patients who carried pathogenic variants or deletions within *ZEB1* gene [34, 63, 65] further support the fact that disruption of *ZEB1* may lead to neurodevelopment defects such as ACC, autism, ID. ZEB1

recognizes a single or bipartite motif comprising one CACCT sequence and one CACCTG E-box on its target gene with various orientations and spacing [37]. It binds with relatively low affinity to a single E-box [51], and with high affinity to a bipartite motif as a monomer with its N-terminal zinc finger cluster attached to one E-box and its C-terminal zinc finger cluster attached to another E-box/E-box-like sequence [37]. It also regulates target gene expression indirectly by interacting with transcription factors that bind chromosomal DNA at sites unrelated to the canonical ZEB1 binding site [52]. ZEB1 is a downstream gene in the AKT/mTORC1/GSK3 $\beta$  pathway, involved in the TGF- $\beta$ 1-induced epithelial-to-mesenchymal transition (EMT) [40], a physiologic process occurs largely during embryonic development but is aberrantly reactivated in pathologic situations [53].

Mechanistic target of rapamycin (mTOR), a member of phosphoinositide 3-kinase (PI3K)-related kinase family, plays important roles in neurogenesis, cellular proliferation, apoptosis, metabolism and development. Altered activity of mTOR has been implicated in many neurological and neuropsychiatric disorders, including autism, intellectual disability and epilepsy [42]. In the AKT/mTORC1/GSK3 $\beta$  pathway, AKT3 are predominantly expressed in the CNS, and regulates axon regeneration in CNS, including in retinal ganglion cells, via interactions between mTOR and GSK3 $\beta$  [54]. Depletion of RPTOR, a component of mTOR1 complex, significantly decreased AKT3-induced axon regeneration [54]. AKT1 also interacts with GSK3 $\beta$  to mediate astroglial autophagy [43]. Interestingly, DEAF1 has been identified as an interactor and *in vitro* substrate of glycogen synthase kinase-3 (GSK3 $\alpha$  and GSK3 $\beta$ ) which also interact with the PI3K/mTOR pathway [55]. Yeast two-hybrid assays demonstrated that DEAF1 and LDB1 binds with LMO4 through LIM domains [56]. LDB1 interacts with CHD8,

which is regulated by CTNNB1, a key downstream gene of *GSK3B*. These results implied that DEAF1 may act downstream of the mTOR/GSK3 $\beta$  pathway. RAI1 was also predicted to interact with GSK3 $\beta$  based on yeast two-hybrid assays [57]. Circadian and mTOR signaling pathways were significantly altered in both *MBD5* and *RAI1* siRNA-mediated knockdown SH-SY5Y cells in microarray analysis, and levels of mTOR mRNA levels were significantly increased in SMS patient-derived lymphoblastoid cells (LCLs) compared to LCLs derived from controls [24]. These results suggest that RAI1 may function in the feedback loop of the mTOR network in neuropsychiatric disorders.

Through network analysis together with the above literature integration, our results show that DEAF1 binds and transcriptionally regulates *RAI1*, and may function downstream of ZEB1. The common molecular network comprised these genes may help to explain overlapping symptoms in *ZEB1*-associated neurodevelopment disorder, DAND, SMS, PTLs, as well as other neuropsychiatric disorders.

## **Limitations**

In this study, molecular mechanism for the potential common gene network of SMS, PTLs, DAND and *ZEB1*-associated neurodevelopment disorder are mostly *in vitro* evidence. To understand how the dysregulation of this gene network contributes to the overlapped phenotypes observed in these disorders *in vivo*, future studies using genetic engineering animals and/or recruiting patients with multiple variants in these genes need to be carried out. Moreover, of a potential upstream regulator, mTOR regulation of *RAI1*, *DEAF1* or *ZEB1* need to be explored in future studies. These studies would help to elucidate the roles of these

genes in neurogenesis and defects in general, including autism, ID and rare neuropsychiatric syndromes.

## **Conclusions**

In this study, our results show that DEAF1 regulates *RAI1* mRNA expression by binding to the *RAI1* DBS and may function downstream of ZEB1. Pathogenic variants in *RAI1*, *DEAF1* or *ZEB1* produce overlapped phenotypes in SMS, PTLS, DAND and ZEB1-associated neurodevelopmental disorder, suggest that these genes may function in neuronal intracellular signaling networks associate with multiple neuropsychiatric disorders. Our findings help to elucidate the genetic mechanisms underlying the overlapping symptoms in these disorders and provide evidence for a common molecular network across neurodevelopmental and psychiatric syndromes.

## **Abbreviations**

DEAF1: deformed epidermal autoregulatory factor-1

RAI1: retinoic acid induced-1

ZEB1: zinc finger E-box binding homeobox

DAND: *DEAF1*-associated neurodevelopmental disorder

SMS: Smith-Magenis syndrome

PTLS: Potocki-Lupski Syndrome

DBS: DEAF1 binding sequence

ID: intellectual disability

ACC: agenesis of corpus callosum

DMEM: Dulbecco's Modified Eagle's Minimum Essential Medium

FBS: fetal bovine serum

qPCR: Quantitative polymerase chain reaction

EMSA: electrophoretic mobility shift assays

ChIP: chromatin immunoprecipitation

RSAT: Regulatory Sequence Analysis Tools

DAVID: Database for Annotation, Visualization, and Integrated Discovery

OMIM: Online Mendelian Inheritance in Man

SAND: Sp-100, AIRE, NucP41/75, and DEAF1

NLS: nuclear localization signal

NES: nuclear export signal

MYND: Myeloid translocation protein 8, Nery, and DEAF1

eQTLs: expression quantitative trait loci

TF: transcription factors

PPCD 3: posterior polymorphous corneal dystrophy 3

FECD6: Fuchs' endothelial corneal dystrophy

FIMO: Find Individual Motif Occurrences

MRI: magnetic resonance imaging

EEG: electroencephalogram

## **Declarations**

**Ethics approval and consent to participate**

Not applicable.

**Consent for publication**

Not applicable.

**Availability of data and materials**

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

**Competing interests**

The authors do not declare financial or other potential conflicts of interest.

**Funding**

This work was supported by the Smith-Magenis Syndrome Research Foundation (SHE), National Science Foundation-China grants 31200937 (LC), 81571090 (DS), Shanghai Health and Family Planning Commission grant 20144Y0106 (LC), Key Laboratory of Medical Molecular Virology (MOE & NHC) funding 2020, School of Basic Medical Sciences, Shanghai Medical College, Fudan University (LC) and National Institutes of Health Grant NINDS 5R21NS091724 (PJ). The funders had no role in study design, data collection, analysis, decision to publish, or preparation of the manuscript.

**Authors' contributions**

LC conceived the study, carried out the experiments and analyzed data, and drafted the manuscript. PJ carried out the EMSA experiments and participated in the manuscript revision. XY carried out the ChIP assays. ML carried out the qPCR experiments. DS and SHE provided technical and conceptual support and contributed to the manuscript revision. All

authors read and approved the final manuscript.

### **Acknowledgements**

We thank Miaomiao Pan for assisting the molecular assays and data analysis.

### **Authors' information**

LC, XY, ML, Department of Cellular and Genetic Medicine, School of Basic Medical Sciences,

Fudan University, Shanghai 200032, China

PJ, Department of Physiology, Southern Illinois University, Carbondale, IL USA

DS, Department of Cellular and Genetic Medicine, School of Basic Medical Sciences,

Institutes of Brain Science, State Key Laboratory for Medical Neurobiology, Fudan University,

Shanghai, China.

SHE, Department of Molecular and Human Genetics, Baylor College of Medicine, Houston,

TX, USA

## References

1. Hanly C, Shah H, Au PYB, Murias K. Description of neurodevelopmental phenotypes associated with 10 genetic neurodevelopmental disorders: A scoping review. *Clin Genet*. 2020. doi: 10.1111/cge.13882
2. Chen L, Jensik PJ, Alaimo JT, Walkiewicz M, Berger S, Roeder E, et al. Functional analysis of novel DEAF1 variants identified through clinical exome sequencing expands DEAF1-associated neurodevelopmental disorder (DAND) phenotype. *Hum Mutat*. 2017;38(12):1774-85.
3. Veraksa A, Kennison J, McGinnis W. DEAF-1 function is essential for the early embryonic development of *Drosophila*. *Genesis*. 2002;33(2):67-76.
4. Vulto-van Silfhout AT, Rajamanickam S, Jensik PJ, Vergult S, de Rocker N, Newhall KJ, et al. Mutations affecting the SAND domain of DEAF1 cause intellectual disability with severe speech impairment and behavioral problems. *Am J Hum Genet*. 2014;94(5):649-61.
5. McGee SR, Rose GM, Jensik PJ. Impaired memory and marble burying activity in deformed epidermal autoregulatory factor 1 (Deaf1) conditional knockout mice. *Behav Brain Res*. 2020;380:112383.
6. Chen L JP, Alaimo JT, Walkiewicz M, Berger S, Roeder E, Faqeih EA, Bernstein JA, Smith ACM, Mullegama SV, Saffen DW, Elsea SH. Functional analysis of novel DEAF1 variants identified through clinical exome sequencing expands DEAF1-associated neurodevelopmental disorder (DAND) phenotype. *Hum Mutat* 2017;38(12):1774-85.
7. Vissers LE, de Ligt J, Gilissen C, Janssen I, Steehouwer M, de Vries P, et al. A de novo paradigm for mental retardation. *Nat Genet*. 2010;42(12):1109-12.
8. Berger SI, Ciccone C, Simon KL, Malicdan MC, Vilboux T, Billington C, et al. Exome analysis of Smith-Magenis-like syndrome cohort identifies de novo likely pathogenic variants. *Hum Genet*. 2017;doi: 10.1007/s00439-017-1767-x.
9. Wenger AM, Guturu H, Bernstein JA, Bejerano G. Systematic reanalysis of clinical exome data yields additional diagnoses: implications for providers. *Genet Med*. 2017;19:5.
10. Rauch A, Wieczorek D, Graf E, Wieland T, Endeley S, Schwarzmayr T, et al. Range of genetic mutations associated with severe non-syndromic sporadic intellectual disability: an exome sequencing study. *Lancet*. 2012;380(9854):1674-82.
11. Li SJ, Yu SS, Luo HY, Li X, Rao B, Wang Y, et al. Two de novo variations identified by massively parallel sequencing in 13 Chinese families with children diagnosed with autism spectrum disorder. *Clin Chim Acta*. 2018;479:144-7.
12. Redin C, Gerard B, Lauer J, Herenger Y, Muller J, Quartier A, et al. Efficient strategy for the molecular diagnosis of intellectual disability using targeted high-throughput sequencing. *J Med Genet*. 2014;51(11):724-36.
13. Gund C, Powis Z, Alcaraz W, Desai S, Baranano K. Identification of a syndrome comprising microcephaly and intellectual disability but not white matter disease associated with a homozygous c.676C>T p.R226W DEAF1 mutation. *Am J Med Genet A*. 2016;170(5):1330-2.
14. Faqeih EA, Al-Owain M, Colak D, Kenana R, Al-Yafee Y, Al-Dosary M, et al. Novel homozygous DEAF1 variant suspected in causing white matter disease, intellectual disability, and microcephaly. *Am J Med Genet A*. 2014;164A(6):1565-70.
15. Rajab A, Schuelke M, Gill E, Zwirner A, Seifert F, Morales Gonzalez S, et al. Recessive DEAF1 mutation associates with autism, intellectual disability, basal ganglia dysfunction and epilepsy. *J Med Genet*. 2015.
16. Nabais Sa MJ, Jensik PJ, McGee SR, Parker MJ, Lahiri N, McNeil EP, et al. De novo and biallelic DEAF1 variants cause a phenotypic spectrum. *Genet Med*. 2019;21(9):2059-69.
17. Huggenvik JJ, Michelson RJ, Collard MW, Ziembra AJ, Gurley P, Mowen KA. Characterization of a nuclear

deformed epidermal autoregulatory factor-1 (DEAF-1)-related (NUDR) transcriptional regulator protein. *Mol Endocrinol*. 1998;12(10):1619-39.

18. Chen L, Tao Y, Song F, Yuan X, Wang J, Saffen D. Evidence for genetic regulation of mRNA expression of the dosage-sensitive gene retinoic acid induced-1 (RAI1) in human brain. *Sci Rep*. 2016;6:19010.

19. Elsea SH, Girirajan S. Smith-Magenis syndrome. *Eur J Hum Genet*. 2008;16(4):412-21.

20. Potocki L, Neira-Fresneda J, Yuan B. Potocki-Lupski Syndrome. In: Adam MP, Ardinger HH, Pagon RA, Wallace SE, Bean LJH, Stephens K, et al., editors. *GeneReviews*((R)). Seattle WA: University of Washington, Seattle. GeneReviews is a registered trademark of the University of Washington, Seattle; 1993.

21. Williams SR, Zies D, Mullegama SV, Grotewiel MS, Elsea SH. Smith-Magenis syndrome results in disruption of CLOCK gene transcription and reveals an integral role for RAI1 in the maintenance of circadian rhythmicity. *Am J Hum Genet*. 2012;90(6):941-9.

22. Burns B, Schmidt K, Williams SR, Kim S, Girirajan S, Elsea SH. Rai1 haploinsufficiency causes reduced Bdnf expression resulting in hyperphagia, obesity and altered fat distribution in mice and humans with no evidence of metabolic syndrome. *Hum Mol Genet*. 2010;19(20):4026-42.

23. Smith ACM, Morse RS, Introne W, Duncan WC, Jr. Twenty-four-hour motor activity and body temperature patterns suggest altered central circadian timekeeping in Smith-Magenis syndrome, a neurodevelopmental disorder. *Am J Med Genet A*. 2019;179(2):224-36.

24. Mullegama SV, Pugliesi L, Burns B, Shah Z, Tahir R, Gu Y, et al. MBD5 haploinsufficiency is associated with sleep disturbance and disrupts circadian pathways common to Smith-Magenis and fragile X syndromes. *Eur J Hum Genet*. 2015;23(6):781-9.

25. Darvekar S, Rekdal C, Johansen T, Sjøttem E. A phylogenetic study of SPBP and RAI1: evolutionary conservation of chromatin binding modules. *PLoS One*. 2013;8(10):e78907.

26. Garay PM, Wallner MA, Iwase S. Yin-yang actions of histone methylation regulatory complexes in the brain. *Epigenomics*. 2016;8(12):1689-708.

27. Singh S, Howell D, Trivedi N, Kessler K, Ong T, Rosmaninho P, et al. Zeb1 controls neuron differentiation and germinal zone exit by a mesenchymal-epithelial-like transition. *Elife*. 2016;5.

28. Jensik PJ, Vargas JD, Reardon SN, Rajamanickam S, Huggenvik JI, Collard MW. DEAF1 binds unmethylated and variably spaced CpG dinucleotide motifs. *PLoS One*. 2014;9(12):e115908.

29. Wagschal A, Delaval K, Pannetier M, Arnaud P, Feil R. PCR-Based Analysis of Immunoprecipitated Chromatin. *CSH Protoc*. 2007;2007:pdb prot4768.

30. Medina-Rivera A, Defrance M, Sand O, Herrmann C, Castro-Mondragon JA, Delerce J, et al. RSAT 2015: Regulatory Sequence Analysis Tools. *Nucleic Acids Res*. 2015.

31. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009;4(1):44-57.

32. Huang da W, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res*. 2009;37(1):1-13.

33. Jensik PJ, Huggenvik JI, Collard MW. Identification of a nuclear export signal and protein interaction domains in deformed epidermal autoregulatory factor-1 (DEAF-1). *J Biol Chem*. 2004;279(31):32692-9.

34. Heide S, Keren B, Billette de Villemeur T, Chantot-Bastarud S, Depienne C, Nava C, et al. Copy Number Variations Found in Patients with a Corpus Callosum Abnormality and Intellectual Disability. *J Pediatr*. 2017;185:160-6 e1.

35. Bulayeva K, Lesch KP, Bulayev O, Walsh C, Glatt S, Gurganova F, et al. Genomic structural variants are linked with intellectual disability. *J Neural Transm (Vienna)*. 2015;122(9):1289-301.

36. Lechner J, Dash DP, Muszynska D, Hosseini M, Segev F, George S, et al. Mutational spectrum of the ZEB1

gene in corneal dystrophies supports a genotype-phenotype correlation. *Invest Ophthalmol Vis Sci*. 2013;54(5):3215-23.

37. Remacle JE, Kraft H, Lerchner W, Wuytens G, Collart C, Verschueren K, et al. 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(18):5073-84.

38. Maturi V, Enroth S, Heldin CH, Moustakas A. Genome-wide binding of transcription factor ZEB1 in triple-negative breast cancer cells. *J Cell Physiol*. 2018;233(10):7113-27.

39. The ENCODE (ENCyclopedia Of DNA Elements) Project. *Science*. 2004;306(5696):636-40.

40. Song Y, Chen Y, Li Y, Lyu X, Cui J, Cheng Y, et al. Metformin inhibits TGF-beta1-induced epithelial-to-mesenchymal transition-like process and stem-like properties in GBM via AKT/mTOR/ZEB1 pathway. *Oncotarget*. 2018;9(6):7023-35.

41. Kim EY, Kim A, Kim SK, Kim HJ, Chang J, Ahn CM, et al. Inhibition of mTORC1 induces loss of E-cadherin through AKT/GSK-3beta signaling-mediated upregulation of E-cadherin repressor complexes in non-small cell lung cancer cells. *Respir Res*. 2014;15:26.

42. Switon K, Kotulska K, Janusz-Kaminska A, Zmorzynska J, Jaworski J. Molecular neurobiology of mTOR. *Neuroscience*. 2017;341:112-53.

43. Kim JE, Ko AR, Hyun HW, Min SJ, Kang TC. P2RX7-MAPK1/2-SP1 axis inhibits MTOR independent HSPB1-mediated astroglial autophagy. *Cell Death Dis*. 2018;9(5):546.

44. Fahrner JA, Bjornsson HT. Mendelian disorders of the epigenetic machinery: tipping the balance of chromatin states. *Annu Rev Genomics Hum Genet*. 2014;15:269-93.

45. Williams SR, Aldred MA, Der Kaloustian VM, Halal F, Gowans G, McLeod DR, et al. Haploinsufficiency of HDAC4 causes brachydactyly mental retardation syndrome, with brachydactyly type E, developmental delays, and behavioral problems. *Am J Hum Genet*. 87(2):219-28.

46. Girirajan S, Elsas LJ, 2nd, Devriendt K, Elsea SH. RAI1 variations in Smith-Magenis syndrome patients without 17p11.2 deletions. *J Med Genet*. 2005;42(11):820-8.

47. Walzl S. Intellectual disability: novel mutations in DEAF1 cause speech impairment and behavioral problems. *Clin Genet*. 2014;86(6):507-8.

48. Bottomley MJ, Collard MW, Huggenvik JI, Liu Z, Gibson TJ, Sattler M. The SAND domain structure defines a novel DNA-binding fold in transcriptional regulation. *Nat Struct Biol*. 2001;8(7):626-33.

49. Czesak M, Le Francois B, Millar AM, Deria M, Daigle M, Visvader JE, et al. Increased serotonin-1A (5-HT1A) autoreceptor expression and reduced raphe serotonin levels in deformed epidermal autoregulatory factor-1 (Deaf-1) gene knock-out mice. *J Biol Chem*. 2012;287(9):6615-27.

50. Wentzel C, Rajcan-Separovic E, Ruivenkamp CA, Chantot-Bastaraud S, Metay C, Andrieux J, et al. Genomic and clinical characteristics of six patients with partially overlapping interstitial deletions at 10p12p11. *Eur J Hum Genet*. 2011;19(9):959-64.

51. Manavella PA, Roqueiro G, Darling DS, Cabanillas AM. The ZFH1A gene is differentially autoregulated by its isoforms. *Biochem Biophys Res Commun*. 2007;360(3):621-6.

52. Sánchez-Tilló E, Siles L, de Barrios O, Cuatrecasas M, Vaquero EC, Castells A, et al. Expanding roles of ZEB factors in tumorigenesis and tumor progression. *Am J Cancer Res*. 2011;1(7):897-912.

53. Mikaelian I, Malek M, Gadet R, Viallet J, Garcia A, Girard-Gagnepain A, et al. Genetic and pharmacologic inhibition of mTORC1 promotes EMT by a TGF-beta-independent mechanism. *Cancer Res*. 2013;73(22):6621-31.

54. Miao L, Yang L, Huang H, Liang F, Ling C, Hu Y. mTORC1 is necessary but mTORC2 and GSK3β are inhibitory for AKT3-induced axon regeneration in the central nervous system. *Elife*. 2016;5:e14908.

55. Pilot-Storck F, Chopin E, Rual JF, Baudot A, Dobrokhotov P, Robinson-Rechavi M, et al. Interactome mapping of the phosphatidylinositol 3-kinase-mammalian target of rapamycin pathway identifies deformed epidermal autoregulatory factor-1 as a new glycogen synthase kinase-3 interactor. *Mol Cell Proteomics*. 2010;9(7):1578-93.
56. Joseph S, Kwan AH, Stokes PH, Mackay JP, Cubeddu L, Matthews JM. The structure of an LIM-only protein 4 (LMO4) and Deformed epidermal autoregulatory factor-1 (DEAF1) complex reveals a common mode of binding to LMO4. *PLoS One*. 2014;9(10):e109108.
57. Vinayagam A, Stelzl U, Foulle R, Plassmann S, Zenkner M, Timm J, et al. A directed protein interaction network for investigating intracellular signal transduction. *Sci Signal*. 2011;4(189):rs8.
58. Elsea SH, Williams SR. Smith-Magenis syndrome: haploinsufficiency of RAI1 results in altered gene regulation in neurological and metabolic pathways. *Expert Rev Mol Med*. 2011;13:e14.
59. Soler-Alfonso C, Motil KJ, Turk CL, Robbins-Furman P, Friedman EM, Zhang F, et al. Potocki-Lupski syndrome: a microduplication syndrome associated with oropharyngeal dysphagia and failure to thrive. *J Pediatr*. 2011;158(4):655-9 e2.
60. Potocki L, Bi W, Treadwell-Deering D, Carvalho CM, Eifert A, Friedman EM, et al. Characterization of Potocki-Lupski syndrome (dup(17)(p11.2p11.2)) and delineation of a dosage-sensitive critical interval that can convey an autism phenotype. *Am J Hum Genet*. 2007;80(4):633-49.
61. Andoni T, Ellard S, Kapadia J, Wakeling E. A novel autosomal recessive DEAF1 nonsense variant: expanding the clinical phenotype. *Clin Dysmorphol*. 2019.
62. Sharma P, Gambhir PS, Phadke SR, Mandal K. Expanding the phenotype in autosomal dominant mental retardation-24: a novel variation in DEAF1 gene. *Clin Dysmorphol*. 2019;28(2):94-7.
63. Jang MS, Roldan AN, Frausto RF, Aldave AJ. Posterior polymorphous corneal dystrophy 3 is associated with agenesis and hypoplasia of the corpus callosum. *Vision Res*. 2014;100:88-92.
64. Shahdadpuri R, de Vries B, Pfundt R, de Leeuw N, Reardon W. Pseudoarthrosis of the clavicle and copper beaten skull associated with chromosome 10p11.21p12.1 microdeletion. *Am J Med Genet A*. 2008;146A(2):233-7.
65. Chaudhry A, Chung BH, Stavropoulos DJ, Araya MP, Ali A, Heon E, et al. Agenesis of the corpus callosum, developmental delay, autism spectrum disorder, facial dysmorphism, and posterior polymorphous corneal dystrophy associated with ZEB1 gene deletion. *Am J Med Genet A*. 2017;173(9):2467-71.

## Tables

**Table 1.** Characteristic features of SMS, PTLs, DAND, and *ZEB1* associated neurodevelopmental disorder.

Diseases and phenotypes		SMS (n > 100) [19, 58]	PTLS (n > 50) [59, 60]	DAND (n =45) [4, 6, 8, 11-16, 61, 62]	<i>ZEB1</i> associated neurodevelopmental disorder (n=7) [34, 50, 63-65]
Gene		<i>RAI1</i>	<i>RAI1</i>	<i>DEAF1</i>	<i>ZEB1</i>
<b>Cognitive functions</b>	<b>Intellectual disability</b>	<100% mild to moderate intellectual disability	100% intellectual disability and cognitive impairment	91% moderate to severe Intellectual disability, 4% severe cognitive delay	14% mild intellectual disability
	<b>Speech problems</b>	75-90%* speech delay	95% speech and language impairment, 95% articulation	91% severe speech delay, severely affected expressive speech	43% speech delay
	<b>Other</b>	decreased ability in sequential processing and short-term memory			
<b>Developmental abnormalities</b>	<b>Motor problems</b>	66-90% motor delay	100% developmental delay	89% mild to moderate motor milestone delay	71% developmental delay, 29% motor delay
	<b>Other</b>	failure to thrive	71% failure to thrive in infancy or early childhood, 8% prenatal growth delay, 8% small for gestational age at birth	9% regression, 4% prenatal growth retardation	
<b>Neurological abnormalities</b>	<b>Hypotonia</b>	63-90% hypotonia	88% infantile hypotonia	58% hypotonia	29% hypotonia
	<b>Movement abnormalities</b>	abnormal gait, balance problems		51% abnormal walking pattern, ataxia, involuntary movements, tremor, etc.	
	<b>Seizures</b>	11-30% seizures	5% epilepsy	51% seizures	14% seizure

	<b>Sleep disturbances</b>	70-100% sleeping problems	31% subjective sleep disturbance, 89% central and/or obstructive sleep apnea	29% sleeping problems	14% sleeping problems
	<b>Abnormality detected by EEG</b>		79% abnormal EEG	7% abnormal EEG	
	<b>Abnormality detected by MRI</b>		50% abnormal CNS by MRI	24% abnormal brain MRI, including 16% white matter abnormalities, 13% corpus callosum, brainstem, Basal ganglia abnormalities, 7% symmetric T2 lesions in the basal ganglia	100% agenesis of corpus callosum, 43% ventriculomegaly including dilatation of the lateral ventricles, 29% plexus choroid cysts, 29% cerebellar hypoplasia, 29% mild white matter abnormality
	<b>Pain insensitivity</b>	~80% reduced pain sensitivity		9% high pain threshold	
	<b>Hearing loss</b>	25-68%	6% hearing impairment	9% hearing loss	29% hearing loss
	<b>Ophthalmological</b>	53-60% myopia, 50% strabismus	60% hypermetropia		100% corneal dystrophy
	<b>Other</b>		80% oral-pharyngeal dysphasia		
<b>Behavioral problems</b>	<b>Autistic phenotypes</b>	>80% autistic-like phenotype and stereotypic behavior	83% autism spectrum disorder, 88% echolalia, 63% mannerisms	60% including 18% poor eye contact, 11% repetitive behavior	14% autism
	<b>Aggressive behaviors</b>	40%-80% aggressive behavior		24% aggressive behavior	29% aggressive behavior
	<b>Maladaptive behaviors</b>	>80% temper tantrums	100% low adaptive function	31% mood swing and excessive irritability	
	<b>Compulsive behaviors</b>	70-100% self-hugging or hand wringing		7% compulsive behavior	

	<b>Self-injurious behavior</b>	78-96% self-injurious behaviors, including 60-90% head banging/face slapping, 60-80% hand-biting/self-biting		24%, including head banging, hits / bites self	
	<b>Hyperactivity</b>	80-100% attention seeking,	hyperactivity	9% hyperactivity,	14% hyperactivity, 14% attention deficit hyperactivity disorder
	<b>Other</b>	25-85% onychotillomania, 25-90% polyembolokoilamania		9% self-stimulatory behavior, 9% happy predisposition, 4% fascinations, 2% anxiety	
<b>Dysmorphologies</b>	<b>Head size</b>	85-89% brachycephaly	8% microcephaly,	16% microcephaly and/or brachycephaly, 9% macrocephaly, etc.	14% abnormal occipito-frontal circumference (OFC)
	<b>Craniofacial abnormalities</b>	77-93% midface hypoplasia, 73-92% tented upper lip, 81-92% broad square face, 53-73% prognathism, 33-62% synophrys, 10% cleft lip/palate	58% micrognathia, 29% high arched palate, 8% submucosal cleft palate, 8% bifid uvula, broad forehead, gentle down-slant of the palpebral fissures, triangular face, asymmetric smile	20% lip dysmorphology including tented lip, 18% abnormal cupid's bow, 16% Straight eyebrows, 16% synophrys, 11% facial dysmorphology including prominent chin, 11% high arched palate, 7% epicanthus, up-slant, 4% low-set hairline	57% bulbous nasal tip, 57% full cheek 43% deep set eyes, 29% synophrys and low-set eyebrows, 29% upper lip, 29% dysplastic ears, 29% down-slanting palpebral fissures, 14% short neck, 14% without a cupid bow, 14% narrow palpebral fissures, 14% epicanthic folds, 14% prominent eyes, 14% hypertelorism, 14% deep set ears
	<b>Skeletal abnormalities</b>	83-85% brachydactyly, 23-69% short stature, 49-67% scoliosis,	24% short stature, 31% scoliosis > 10 degrees	18% joint hypermobility, 16% skin syndactyly in toes 2 and 3, 9% sacral dimple, 7% pes planus, 4% brachydactyly, 4% fetal finger pads, 4% clinodactyly, 4% sandal gap, 2% 4th toe overlaps	29% bilateral syndactyly of 2-3 toes, 29% hypermobile joints, spondyloarthropathy, 14% camptodactyly of toes, 14% Small and thick fingers, 14% pseudoarthrosis of the clavicle and copper beaten skull, 14% bilateral 5th finger clinodactyly, 14% sandal gap
	<b>Cardiovascular abnormalities</b>	30-40% cardiovascular abnormalities,	50% cardiovascular abnormalities		29% ventricular septal defect, 14% coarctation aorta, 14% Interventricular communications, 14% multiple

		electrocardiographic abnormalities			small ventricular septal defects with left ventricular hypertrophy, right ventricular hypertrophy and a bicuspid aortic valve with moderate aortic coarctation.
<b>Genitourinary anomalies</b>	15-30% renal/urinary tract abnormality		15% structural renal anomaly	7% cryptorchidism and/or micropenis	14% a bicornuate uterus, unilateral cryptorchidism, 14% cryptorchidy,
<b>Other</b>	88-100% hoarse, deep voice, >90% dental anomalies			18% thin fair hair, 4% dry skin, 2% stridor laryngomalacia, 2% telecanthus, 2% short webbed neck	14% lacrimal duct stenosis, bilateral inguinal hernia, 14% bifid tongue, 14% bilateral vesicoureteric reflux
<b>Additional phenotypes</b>	<b>Infections</b>	55-85% chronic ear infections		22% report recurrent infections including otitis media, respiratory tract infections	
	<b>Feeding difficulties</b>	feeding difficulties	92% poor feeding in neonatal period and infancy, 78% residual component after oral phase completion, 72% residual component after pharyngeal phase completion, 67% abnormal lingual function, 61% mild delay in swallow, 39% abnormal mastication, 39% laryngeal penetration, 29% enteral feeding tube,	15% report feeding difficulties	
	<b>Metabolic abnormalities</b>	13-67% obesity, hyperphagia, impaired satiety, high cholesterol	73% gastroesophageal reflux, 46% chronic constipation, 30% low total cholesterol and low LDL	15% digestive abnormality, including gastroesophageal reflux, and constipation	

\* Statistics including patients carrying *RAI1* mutations or deletions.

## Figure legends

**Figure 1.** DEAF1 regulates *RAI1* mRNA expression via a DEAF1 binding site (DBS) in *RAI1* intron 2. (A) A luciferase reporter plasmid containing *RAI1* DBS in intron 2 (*RAI1*-Luc) showed a significant increase in luciferase activity when co-transfected with a WT DEAF1 (DEAF1) expression plasmid (GenBank accession number NM\_021008) compared to cells transfected with the *RAI1*-Luc reporter plasmid alone. Each experiment was performed in triplicate. Error bars represent standard error of mean (SEM), and p-values were based on ANOVA and *post hoc* Tukey's tests (\*\*indicates p-value < 0.01). (B) Electrophoretic mobility shift assays (EMSAs) confirmed that DEAF1 binds to the DBS in *RAI1* intron 2 *in vitro*. A 35 bp double-stranded (ds) DNA oligonucleotide containing the *RAI1* intron 2 DBS was used as the experimental probe to test binding of purified FLAG-tagged WT DEAF1 protein. A ds DNA oligonucleotide probe containing the previously validated DEAF1 binding site in human *DEAF1* promoter was used as a positive control. (C) Chromatin immunoprecipitation (ChIP) assays showed that N-terminal hemagglutinin (HA)-tagged DEAF1 binds to chromatin fragments containing the *RAI1*-DBS in HEK293T cells. Mouse monoclonal anti-HA antibodies (IgG1 subtype) were used to precipitate exogenously expressed HA-DEAF1 and mouse IgG was used as a negative control. A 219 bp DNA fragment containing the *RAI1*-DBS was PCR amplified. No Ab: no antibody added to the IP reaction mixture. Negative control for PCR (-): deionized water was added instead of chromosomal DNA template. (D) ChIP PCR products were quantified by quantitative scanning of the bands in the gels. Statistical significance was assessed by one-way ANOVA and *post hoc* Turkey's multiple comparison, \*\*\*\* indicates p-value < 0.0001.

**Figure 2.** Location of the putative ZEB1 binding site in *DEAF1*. (top) Screen shot from the UCSC Genome Browser for the indicated region of chromosome 11 showing the positions of *DEAF1*, histone III lysine 27 acetylation, DNase sensitive clusters and location of rs6597991 within *DEAF1* intron 1. (bottom) Location of rs6597991 within a predicted binding site for the transcription factor ZEB1 within *DEAF1* intron 1, the underlined consensus binding sequences, and DNA sequences protected by transcription factor binding in DNase footprint assays as listed in RegulomeDB database. The consensus “E-box” sequence CAGGTG [37] and a possible weaker, second binding site that is required for high-affinity ZEB 1 binding as predicted by FIMO program are underlined.

**Figure 3.** ZEB1 binds to a putative *DEAF1* 5' regulatory region that includes SNP rs6597991. (A) EMSAs detect binding to putative *DEAF1* intron1 ZEB1 binding sequence using HEK293T nuclear extracts. A 38-bp double-stranded (ds) 3'-biotinylated DNA oligonucleotide containing the ZEB1 binding site in *DEAF1* 5' region was used as the experimental probe. Specificity of binding was demonstrated by addition of excess of non-biotinylated experimental probes to the binding assay. (B) ChIP assays showed that ZEB1 binds to chromosome segments containing the putative ZEB1 binding sequence in the *DEAF1* promoter in HEK293T cells. Mouse monoclonal anti-ZEB1 antibody was used to precipitate chromosome segments bound to ZEB1, with mouse IgG serving as a negative control. No Ab: no antibodies added to the IP mixture. (C) Quantitative PCR confirmed statistically significant differences between the amounts of targeted DNA precipitated by anti-ZEB1 antibodies or by IgG. Statistical

significance was assessed by one-way ANOVA and post hoc Turkey's multiple comparison, \* indicates p-value < 0.05, \*\* indicates p-value < 0.01.

**Figure 4.** Additional predicted ZEB1 binding sites in *DEAF1* 5' regulatory region. (top) Screen shot from the UCSC Genome Browser for the indicated region of chromosome 11 showing the positions of *DEAF1*, histone III lysine 27 acetylation, DNase sensitive clusters, the location of rs6597991 and additional ZEB1 binding sites within *DEAF1* intron 1. (bottom) Sequence of MACS\_peak\_517 (chr 11: 693,829-695,048, 1219 bp) [38] and potential ZEB1 binding sites within *DEAF1* intron 1, the underlined consensus binding sequences, and DNA sequences protected by transcription factor binding in DNase footprint assays predicted by FIMO program. Each yellow high-lighted sequence contains a consensus "E-box", red in + strand, blue in - strand.

**Figure 5.** ZEB1 binds to the MACS\_peak\_517 region in *DEAF1* intron 1. (A) ChIP assays showed that ZEB1 binds to chromosome segments containing the putative ZEB1 binding sequence in the *DEAF1* promoter in HEK293T cells. Mouse anti-ZEB1 antibodies were used to precipitate chromosome segments bound to ZEB1, with mouse IgG serving as a negative control. No Ab: no antibody added to the IP reaction mixture. (B) Quantitative PCR confirmed significant difference between the amounts of targeted DNA precipitated by anti-ZEB1 antibodies or by IgG. Statistical significance was assessed by one-way ANOVA and post hoc Turkey's multiple comparison, \* indicates p-value < 0.05.

**Figure 6.** ZEB1 negatively regulates *DEAF1* expression in neuronal cells. Short interfering (si) RNA-mediated knock-down of endogenous *ZEB1* expression in human SHSY5Y cells increases *DEAF1* mRNA expression using qPCR, compared to cells transfected with control siRNA, which is normalized to 1. Statistical significance was assessed by two-tailed T-test, \* indicates p-value < 0.05, \*\* indicates p-value < 0.01.

**Figure 7.** *RAI1*, *DEAF1*, *ZEB1*, and/or *mTOR* associated gene network. Gene network was generated using the STRING database. Each network node represents a protein-coding gene locus containing all known proteins isoforms and post-translational modifications. Proteins 3D structures are shown in the nodes. Edges represent protein-protein associations including those based on information from i) curated databases (cobalt blue) or experimentally determined (mauve) and ii) predicted interactions based on gene neighborhoods (green), gene fusions (red) or gene co-occurrence (navy blue). Additional interactions include those based on text-mining (yellow), co-expression (black) and protein homology (purple).

**Figure 8.** Co-expression-based tissue-specific functional network among *RAI1*, *DEAF1*, *ZEB1*, and/or *mTOR* associated genes. Gene network was built using HumanBase database, based on co-expression-based molecular interactions and tissue-specific functional interaction, in “all tissues” (A) versus “neuron” (B), with minimum relationship confidence = 0.45 and maximum number of genes other than query = 12. The color bars represent gene associations with confidences from 0 to 1. Yellow circles represent the genes which have the

most interactions in each network.

# Figures

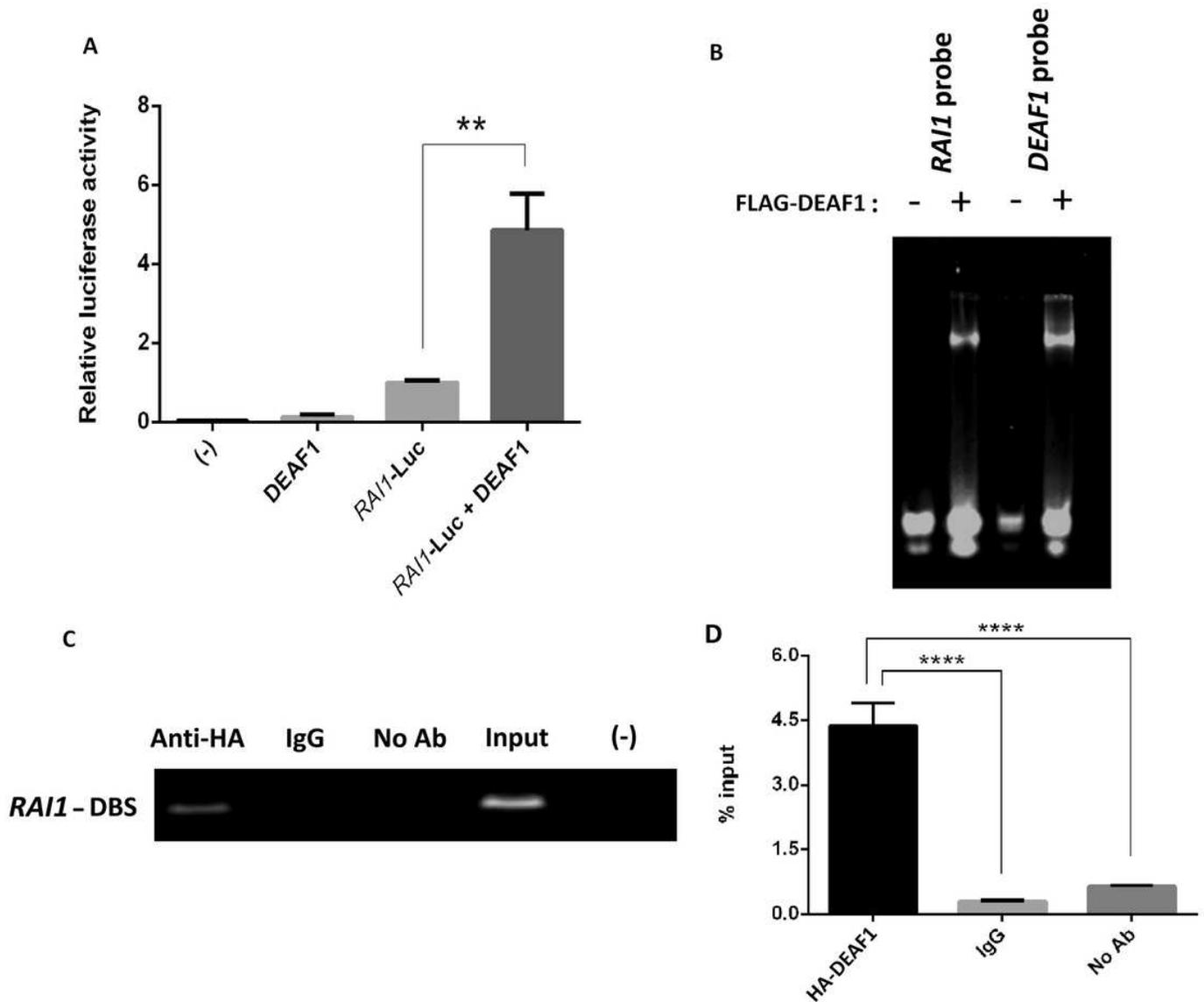
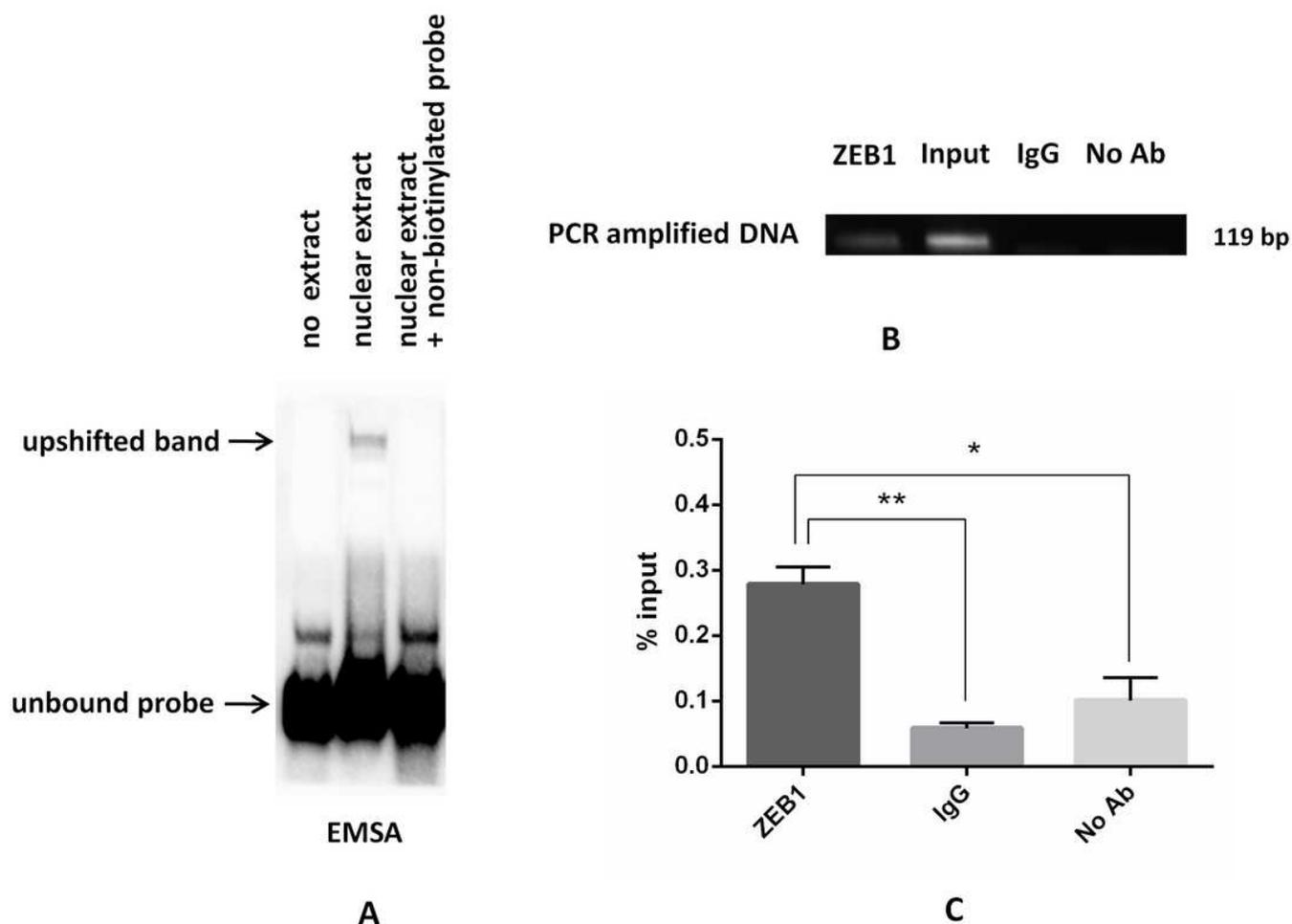


Figure 1

DEAF1 regulates RAI1 mRNA expression via a DEAF1 binding site (DBS) in RAI1 intron 2. (A) A luciferase reporter plasmid containing RAI1 DBS in intron 2 (RAI1-Luc) showed a significant increase in luciferase activity when co-transfected with a WT DEAF1 (DEAF1) expression plasmid (GenBank accession number NM\_021008) compared to cells transfected with the RAI1-Luc reporter plasmid alone. Each experiment was performed in triplicate. Error bars represent standard error of mean (SEM), and p-values were based on ANOVA and post hoc Tukey's tests (\*\*indicates p-value < 0.01). (B) Electrophoretic mobility shift assays (EMSA) confirmed that DEAF1 binds to the DBS in RAI1 intron 2 in vitro. A 35 bp double-stranded (ds) DNA oligonucleotide containing the RAI1 intron 2 DBS was used as the experimental probe to test binding of purified FLAG-tagged WT DEAF1 protein. A ds DNA oligonucleotide probe containing the previously validated DEAF1 binding site in human DEAF1 promoter was used as a positive control. (C)



[37] and a possible weaker, second binding site that is required for high-affinity ZEB 1 binding as predicted by FIMO program are underlined.



**Figure 3**

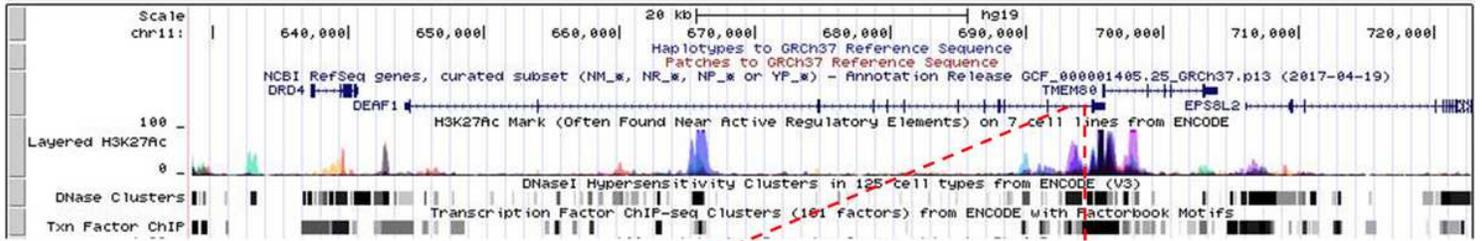
ZEB1 binds to a putative DEAF1 5' regulatory region that includes SNP rs6597991. (A) EMSAs detect binding to putative DEAF1 intron1 ZEB1 binding sequence using HEK293T nuclear extracts. A 38-bp double-stranded (ds) 3'-biotinylated DNA oligonucleotide containing the ZEB1 binding site in DEAF1 5' region was used as the experimental probe. Specificity of binding was demonstrated by addition of excess of non-biotinylated experimental probes to the binding assay. (B) ChIP assays showed that ZEB1 binds to chromosome segments containing the putative ZEB1 binding sequence in the DEAF1 promoter in HEK293T cells. Mouse monoclonal anti-ZEB1 antibody was used to precipitate chromosome segments bound to ZEB1, with mouse IgG serving as a negative control. No Ab: no antibodies added to the IP mixture. (C) Quantitative PCR confirmed statistically significant differences between the amounts of targeted DNA precipitated by anti-ZEB1 antibodies or by IgG. Statistical significance was assessed by one-way ANOVA and post hoc Turkey's multiple comparison, \* indicates p-value < 0.05, \*\* indicates p-value < 0.01.

# UCSC Genome Browser on Human Feb. 2009 (GRCh37/hg19) Assembly

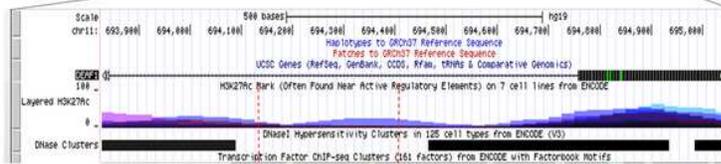
move <<< << < > >> >>> zoom in 1.5x 3x 10x base zoom out 1.5x 3x 10x 100x

chr11:628,520-722,796 94,277 bp. enter position, gene symbol, HGVS or search terms go

chr11 (p15.5) p15.4 15.1 p13 11p12 11.2 11.1 11.2 11.3 11.4 11q14.1 q21 q22.1 q22.3 q23.3 q25



SNP rs6597991 MACS\_peak\_517 (693,829 – 695,048: 1,219 bp)



```

GGGCAGGTGTGTCAGGGCAGGGCCAGGTCAGGTGAGGCTAGGCAGGTACC
AGGGGCAGGTGTGCGGGCAGGTGTGGGGTAGGTGGGCAGGCTGTGTAC
CTGGGGCAGGTATGTGGGCAGGTGTGCGGGCAGGTAGGGCAGGCCGATCA
CGTGGGGCAGGTGGGAAGGCTGGAACCTCGAGCAGGTGTGCGAGGCAGG
TATTTGTGGCAGGTGGGGCAGGCTGTGTCACCTGGAGCAGGTACGTGGGGAAA
GTGTGAGGGCGGGTGGGCAGGTGTGAGAGGCAGATGTGAGGGGTGAGCT
AGTCAGTGGGCAGGTGTGAGGGCGGGTGGGGCGGGCGGGTACAGTGG
GAGGAGGTGTGAGGGGCAGGTGTGAGGGCGGGTGGAGCGGGCTGGTACAGT
GAGCAGGTGTGAGGGCAGGTGTGCGGGCTGTGTCACCTGGGA

CCACAGGTGCCACCCACCCAGGGCTTCTCCAGGACAGGTGGG
    
```

Figure 4

Additional predicted ZEB1 binding sites in DEAF1 5' regulatory region. (top) Screen shot from the UCSC Genome Browser for the indicated region of chromosome 11 showing the positions of DEAF1, histone III lysine 27 acetylation, DNase sensitive clusters, the location of rs6597991 and additional ZEB1 binding sites within DEAF1 intron 1. (bottom) Sequence of MACS\_peak\_517 (chr 11: 693,829-695,048, 1219 bp) [38] and potential ZEB1 binding sites within DEAF1 intron 1, the underlined consensus binding sequences, and DNA sequences protected by transcription factor binding in DNase footprint assays predicted by FIMO program. Each yellow high-lighted sequence contains a consensus "E-box", red in + strand, blue in - strand.

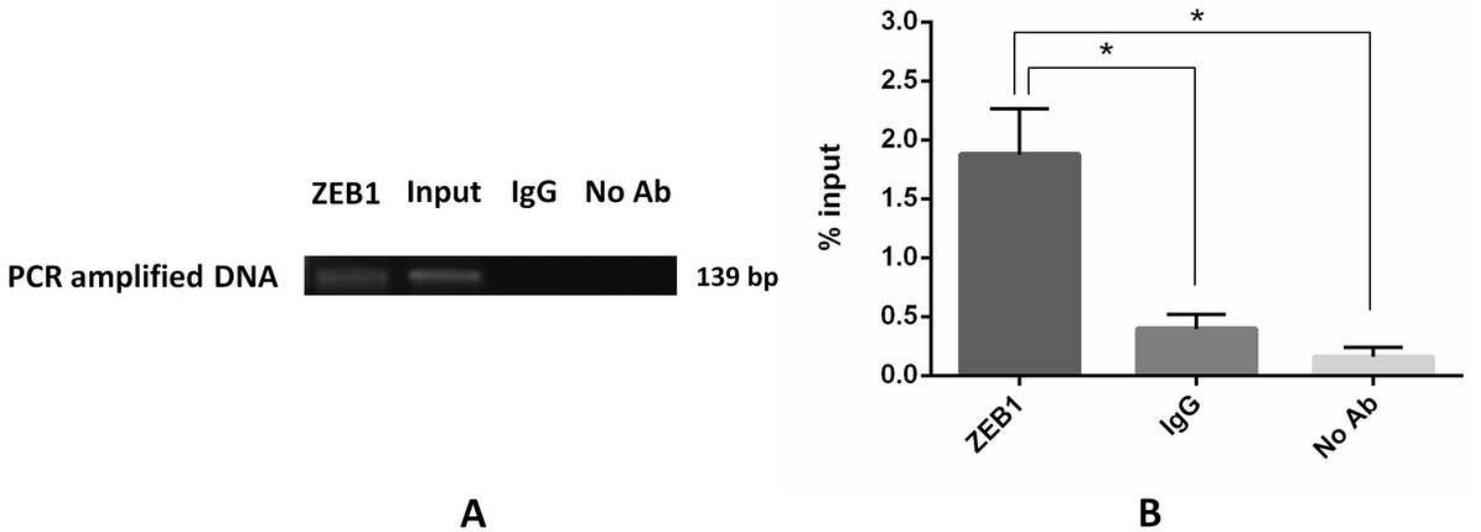


Figure 5

ZEB1 binds to the MACS\_peak\_517 region in DEAF1 intron 1. (A) ChIP assays showed that ZEB1 binds to chromosome segments containing the putative ZEB1 binding sequence in the DEAF1 promoter in HEK293T cells. Mouse anti-ZEB1 antibodies were used to precipitate chromosome segments bound to ZEB1, with mouse IgG serving as a negative control. No Ab: no antibody added to the IP reaction mixture. (B) Quantitative PCR confirmed significant difference between the amounts of targeted DNA precipitated by anti-ZEB1 antibodies or by IgG. Statistical significance was assessed by one-way ANOVA and post hoc Turkey's multiple comparison, \* indicates p-value < 0.05.

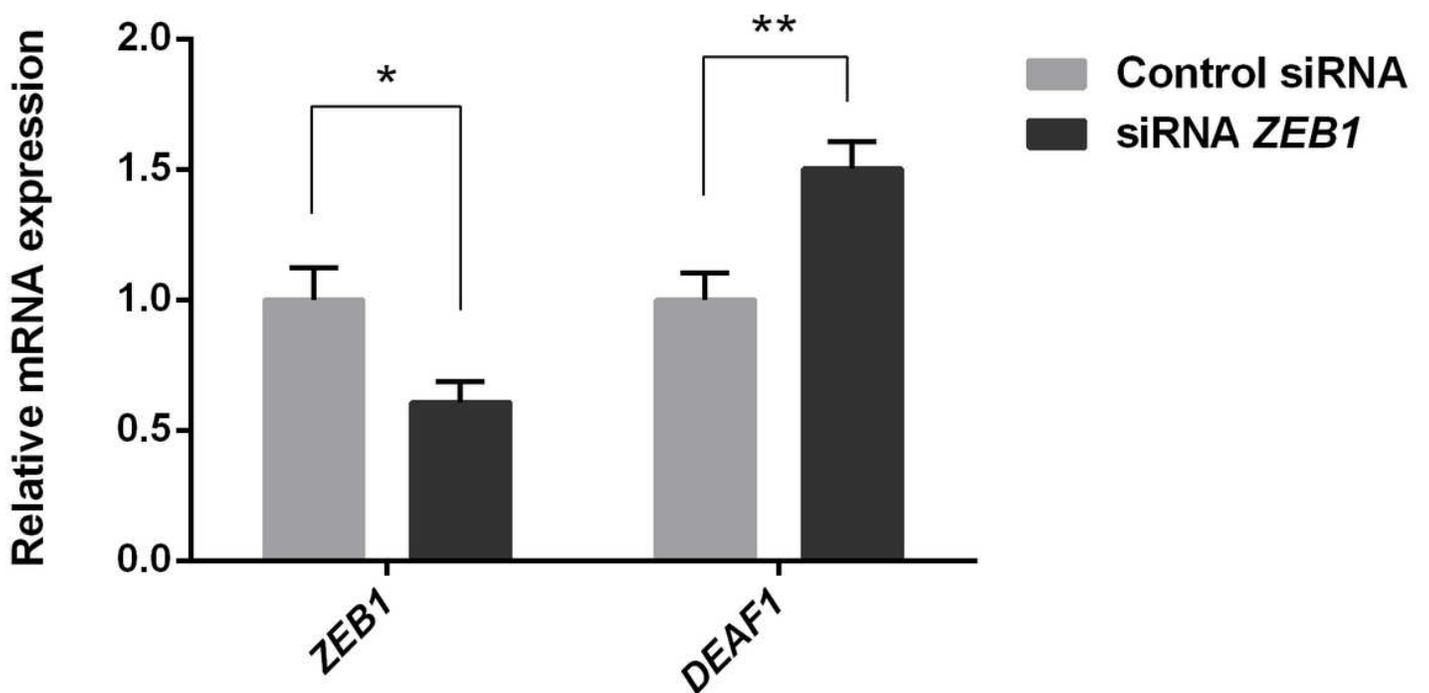
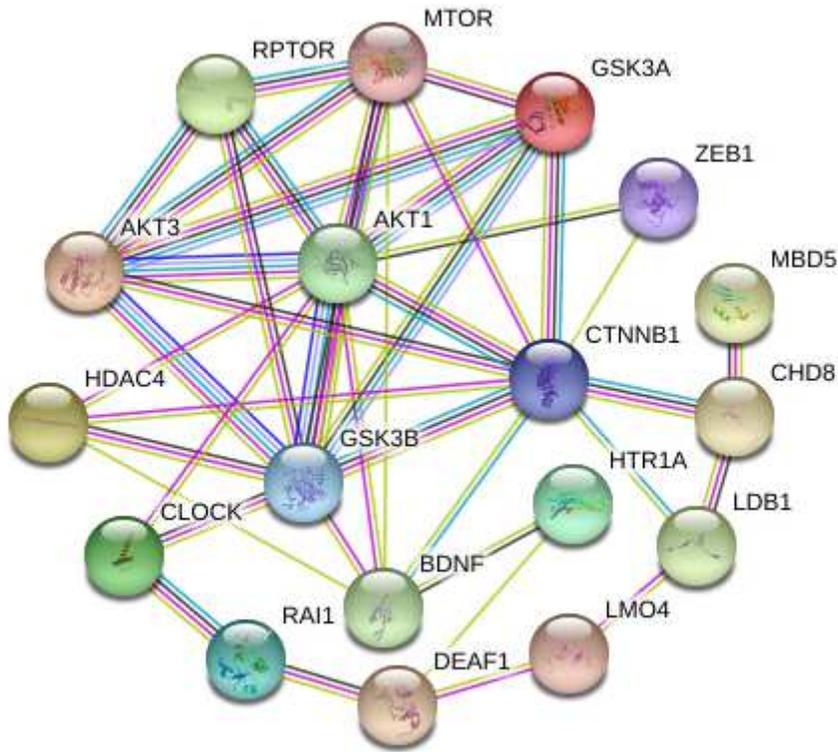


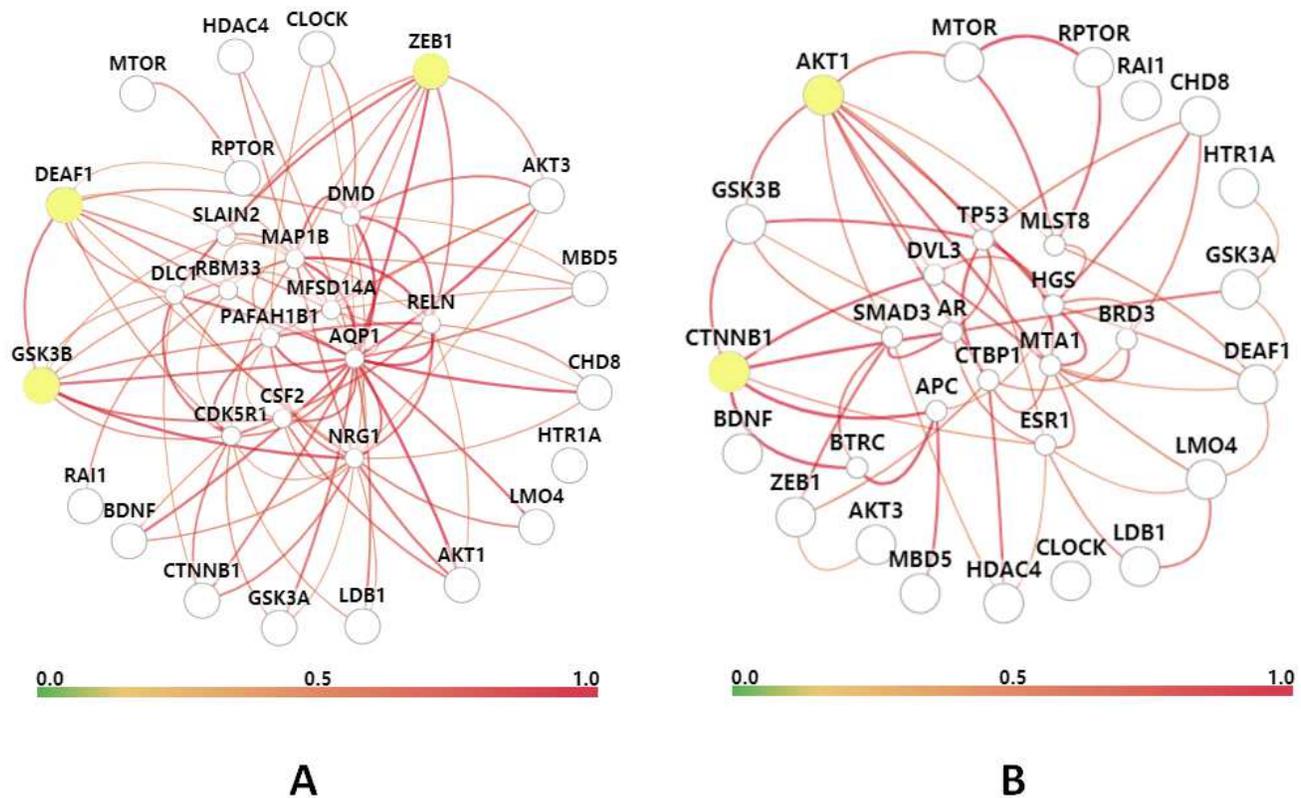
Figure 6

ZEB1 negatively regulates DEAF1 expression in neuronal cells. Short interfering (si) RNA-mediated knock-down of endogenous ZEB1 expression in human SHSY5Y cells increases DEAF1 mRNA expression using qPCR, compared to cells transfected with control siRNA, which is normalized to 1. Statistical significance was assessed by two-tailed T-test, \* indicates p-value < 0.05, \*\* indicates p-value < 0.01.



**Figure 7**

RAI1, DEAF1, ZEB1, and/or mTOR associated gene network. Gene network was generated using the STRING database. Each network node represents a protein-coding gene locus containing all known proteins isoforms and post-translational modifications. Proteins 3D structures are shown in the nodes. Edges represent protein-protein associations including those based on information from i) curated databases (cobalt blue) or experimentally determined (mauve) and ii) predicted interactions based on gene neighborhoods (green), gene fusions (red) or gene co-occurrence (navy blue). Additional interactions include those based on text-mining (yellow), co-expression (black) and protein homology (purple).



**Figure 8**

Co-expression-based tissue-specific functional network among RAI1, DEAF1, ZEB1, and/or mTOR associated genes. Gene network was built using HumanBase database, based on co-expression-based molecular interactions and tissue-specific functional interaction, in "all tissues" (A) versus "neuron" (B), with minimum relationship confidence = 0.45 and maximum number of genes other than query = 12. The color bars represent gene associations with confidences from 0 to 1. Yellow circles represent the genes which have the most interactions in each network.

## Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.tif](#)
- [FigureS2.tif](#)
- [FigureS3.tif](#)
- [FigureS4.tif](#)
- [Supplementarymaterials201210.pdf](#)