

Nucleocytoplasmic Transport of RNA-Binding Proteins Regulates Neural Stem Cell Fates

Melissa J. MacPherson

University of Alberta

Sarah L Erickson

University of Calgary

Drayden Kopp

University of Calgary

Pengqiang Wen

University of Calgary

Mohammadreza Aghanoori

University of Calgary

Antonio Vitobello

Université de Bourgogne-Franche Comté

Quentin Thomas

Université de Bourgogne-Franche Comté

Nina Gold

Boston Children's Hospital

William Brucker

Boston Children's Hospital

Louise Amlie-Wolf

Nemours/Alfred I. duPont Hospital for Children

Karen W. Gripp

Nemours/Alfred I. duPont Hospital for Children

Olaf Bodamer

Boston Children's Hospital

Laurence Faivre

Université de Bourgogne-Franche Comté

Mikko Muona

Folkhälsan Research Center

Xing-Chang Wei

University of Calgary

Quan Long

University of Calgary

A. Micheil Innes (✉ micheil.innes@albertahealthservices.ca)

University of Calgary
Guang Yang (✉ guang.yang2@ucalgary.ca)
University of Calgary

Research Article

Keywords: Celf2, neural stem cell, nucleocytoplasmic shuttling, neurogenesis, fate decision, cortical development, neurodevelopmental disorder

Posted Date: November 30th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-115227/v1>

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

[Read Full License](#)

Nucleocytoplasmic Transport of RNA-Binding Proteins Regulates Neural Stem Cell Fates

Melissa J. MacPherson^{1,2,3,#}, Sarah L. Erickson^{1,2,#}, Drayden Kopp^{2,4,#}, Pengqiang Wen^{1,2},
Mohammadreza Aghanoori^{1,2}, Antonio Vitobello^{6,7}, Quentin Thomas^{6,8}, Nina Gold⁹, William Brucker⁹,
Louise Amlie-Wolf⁵, Karen W. Gripp⁵, Olaf Bodamer⁹, Laurence Faivre^{6,8}, Mikko Muona^{10,11}, Xing-
Chang Wei¹², Quan Long^{1,2,4}, A. Micheil Innes^{1,2,*}, Guang Yang^{1,2,4,*}

1. Department of Medial Genetics, Cumming School of Medicine, University of Calgary, Calgary, AB T2N 4N1
2. Alberta Children's Hospital Research Institute, Calgary, AB T2N 4N1
3. Department of Medical Genetics, University of Alberta, Edmonton, AB T6G 2H7
4. Department of Biochemistry and Molecular Biology, Cumming School of Medicine, University of Calgary, Calgary, AB T2N 4N1
5. Nemours/Alfred I. duPont Hospital for Children 1600 Rockland Road Wilmington, DE 19803
6. UMR1231 GAD (Genetics of Developmental Disorders) INSERM et FHU TRANSLAD, Université de Bourgogne-Franche Comté, Dijon, France
7. Unité Fonctionnelle D'Innovation en Diagnostique Génomique Des Maladies Rares, Pôle de Biologie, CHU Dijon Bourgogne, 21000, Dijon, France
8. Centre de Génétique, Hôpital d'Enfants, CHU, 14 rue Paul Gaffarel, BP 77908, 21079 Dijon cedex, France
9. Division of Genetics and Genomics 300 Longwood Avenue, Enders, 5th Floor Boston, Boston Children's Hospital, MA 02115

10. Blueprint Genetics, Finland

11. Folkhälsan Research Center, Helsinki, Finland

12. Diagnostic Imaging, Alberta Children's Hospital, 28 Oki Drive NW, Calgary, AB T3B 6A8

These authors contributed equally.

*Corresponding authors: Guang Yang, guang.yang2@ucalgary.ca

A. Micheil Innes, micheil.innes@albertahealthservices.ca

Abstract

The formation of the cerebral cortex requires balanced expansion and differentiation of neural progenitor cells, the fate choice of which requires regulation at many steps of gene expression. As progenitor cells often exhibit transcriptomic stochasticity, the ultimate output of cell fate-determining genes must be carefully controlled at the post-transcriptional level, but how this is orchestrated is poorly understood. Here we report that *de novo* missense variants in an RNA-binding protein CELF2 cause human cortical malformations and perturb neural progenitor cell fate decisions in mice by disrupting the nucleocytoplasmic transport of CELF2. In self-renewing neural progenitors, CELF2 is localized in the cytoplasm where it binds and coordinates mRNAs that encode cell fate regulators and neurodevelopmental disorder-related factors. The translocation of CELF2 into the nucleus releases mRNAs for translation and thereby triggers neural progenitor differentiation. Our results reveal a mechanism by which transport of CELF2 between discrete subcellular compartments orchestrates an RNA regulon to instruct cell fates in cerebral cortical development.

Introduction:

The cerebral cortex is the most advanced part of the human brain, forming the characteristic gyri to accommodate a growing number of neurons for higher-order functions (Rakic, 2009; Taverna et al., 2014). The foundation for the development of this structure is neural stem/progenitor cells (NPCs), which possess the ability to retain their identity while giving rise to adequate daughter cells that commit to a more differentiated neuronal fate (Shitamukai and Matsuzaki, 2012; Taverna et al., 2014).

Disrupting this balance of NPC self-renewal and differentiation often causes cortical malformations (e.g. simplified gyral pattern) and cognitive deficits (e.g. speech delay) seen in human neurodevelopmental disorders (Sun and Hevner, 2014). For example, uncontrolled expansion of the NPC pool can lead to brain overgrowth and aberrant cortical folds due to overproduction of neurons (van Bon et al., 2016; Groszer et al., 2006; Park and Chung, 2013; Veleva-Rotse and Barnes, 2014). On the other hand, impaired NPC proliferation and premature differentiation can result in a global reduction in cortical volume with disruption of gyral patterns (Bae et al., 2015; Gilmore and Walsh, 2013; Gruber et al., 2011). Given these close relationships, abnormalities in the size and extent of folding of the cortex in human patients, therefore, frequently inform defects in NPC fate decisions that alter the number of neurons being produced (Bae et al., 2014; Gilmore and Walsh, 2013; Subramanian et al., 2019).

The decision of NPCs to self-renew or differentiation is tightly controlled by the expression of cell fate determination genes (Taverna et al., 2014). Interestingly, the genes that promote neurogenesis are surprisingly actively transcribed in NPCs even before they begin to differentiate (Han et al., 2020; Johnson et al., 2015; Yang et al., 2014; Zahr et al., 2018). This suggests that NPCs must therefore carefully control the actual protein output from cell fate determination genes at other levels in addition

to transcription. Indeed, recent studies reveal that NPCs are strongly regulated by mechanisms that control gene expression at the translational level, through the function of target-specific RNA-binding proteins (RBPs) (Alsina and Silver, 2020; Mohammad et al., 2019; Rodrigues et al., 2020). For example, in the murine forebrain, Pumilio2 and Smaug2 bind and repress the translation of pro-neurogenic mRNAs in cytoplasmic processing bodies (P-bodies), and in doing so, maintain the self-renewing state of cortical NPCs (Amadei et al., 2015; Yang et al., 2014; Zahr et al., 2018). Accordingly, loss of these RBPs causes premature expression of pro-neurogenic mRNAs, leading to precocious differentiation of NPCs. Despite their critical role in deciding NPC fates, these and other RBPs are often maintained at constant levels in NPCs during neurogenesis (Amadei et al., 2015; Yang et al., 2014; Zahr et al., 2018). This raises an interesting question of how RBPs achieve a timely coordination of the repression and release of pro-neurogenic mRNA in NPCs to allow accurate fate decision-making.

In this study, we demonstrated that the dynamic shuttling of the RBP CELF2 between the cytoplasm and nucleus is a molecular on and off switch that coordinates the self-renewing and differentiation states of NPCs. In the cytoplasm, CELF2 binds and represses mRNAs that promotes neurogenesis, and in so doing, maintains NPC self-renewal. The translocation of CELF2 into the nucleus, which releases these mRNAs for translation, induces the neurogenic differentiation of NPCs. Pathological variants that perturb CELF2 nuclear translocation causes cortical malformations and a rare neurodevelopmental disorder. Thus, our results reveal that transport of RBPs between discrete subcellular compartments is a critical mechanism that orchestrates mRNAs to instruct NPC fate decision in brain development.

Results:

Altered Nucleocytoplasmic Localization of CELF2 Causes Cortical Malformations in Humans. To address the underlying RNA-based mechanisms of NPC fate decision, we searched for pathogenic variants that disrupt mRNA regulators in patients with brain malformations. In a 5 year old female with a simplified gyral pattern on MRI (individual 1), we identified a heterozygous *de novo* missense variant (GeneBank: NM_001025077.2, c.1478G>A, p.Arg493His) in the CUGBP Elav-like family member 2 (*CELF2*) gene by clinical trio exome sequencing (**Fig. 1A-i, 1A-i', S1A, B**). *CELF2* belongs to a family of RNA-binding proteins (*CELF1-6*) that shuttle between the nucleus and cytoplasm to regulate mRNA splicing and translation (Dasgupta and Ladd, 2012; Subramaniam et al., 2008). Although *CELF4* and *CELF6* were recently implicated in autism spectrum disorders (ASD) and encephalopathy, the roles of *CELF* family genes in NPC fate decision during brain development are largely unknown (Barone et al., 2017; Dougherty et al., 2013).

Through international collaboration via the GeneMatcher platform (Sobreira et al., 2015), we subsequently identified four additional unrelated individuals (ages ranging from 2 to 5 years) with *de novo* heterozygous missense variants in *CELF2* (**Fig. 1A, B, S1A**). This cohort of five patients with *CELF2* variants had remarkably similar neurodevelopmental features (**Table S1**). They all exhibited global developmental delay with moderate to severe impairment of speech and language capacities. Three individuals displayed stereotypic movements and/or aggressive behaviours and one was diagnosed with ASD. Epilepsy was also frequently observed; three individuals were affected with infantile spasms, an epileptic syndrome of infancy and early childhood. Magnetic resonance imaging (MRI) revealed brain malformations in all individuals, which include simplified gyral patterning that

was predominantly in the bitemporal and bifrontal cortical areas in one patient and involved the entire cortex in three of the patients. Two patients had absence of anterior commissure. Macrocephaly and ventriculomegaly were also observed in some of the patients. Thus, each of the five patients with *CELF2* variants displayed neurodevelopmental defects and cortical abnormalities that are often linked to alterations in NPC fate decision and neurogenesis, which suggests that *CELF2* variants alter human cortical development possibly by perturbing NPC fate decision (Bae et al., 2014; Gilmore and Walsh, 2013; Subramanian et al., 2019). Consistently, gene expression analysis of the human fetal brain using a published single-cell RNA-seq dataset showed that only *CELF2*, and to a much lesser extent *CELF1*, were expressed in cortical NPCs (*SOX2+*, *PAX6+* cells), corroborating with their robust expression in the mouse cortex throughout the neurogenic period (**Fig. 1C, D**)(Pollen et al., 2015).

Further examination of *CELF2* missense variants showed that they were all located in the last exon of *CELF2*, which encodes an RNA-recognition motif (RRM3) (**Fig. 1B**). Of note, the variants in these patients all affect either Arg 493 or Pro 507. This region is highly conserved throughout the metazoa and required for CELF2 splicing activity and nuclear localization (**Fig. S1C**), which raises the possibility that the pathogenic missense variants might perturb these molecular features of CELF2 (Ladd and Cooper, 2004). To test this, we expressed wild-type (WT) and mutant *CELF2* in human embryonic kidney (HEK) 293 cells and analyzed the alternative exon usage of known CELF2 splicing targets, *PPP1R12A* and *TRAF3* (Schultz et al., 2017). CELF2 mutants induced exon skipping of both mRNAs to the same levels as wildtype (WT), which suggests that the pathogenic variants do not compromise the splicing activity of CELF2 (**Fig. 1E, F**). We next asked whether the mutants were defective in nuclear localization. As expected, transfected EGFP-tagged WT CELF2 was predominantly localized in the

nucleus. However, all mutants showed a remarkably similar pattern of distribution, where EGFP-tagged CELF2 was largely excluded from the nucleus and accumulated in the cytoplasm, suggesting that disruption of proper nucleocytoplasmic localization underlies the pathological impact of *CELF2* variants on human cortical development (**Fig. 1G**).

Nuclear Translocation of CELF2 Correlates with NPC Differentiation. Our analysis of this human neurodevelopmental condition points to a potential role of CELF2 transport in regulating NPC fates. To explore CELF2-mediated regulation of NPC fate decision, we first used CELF2-specific antibodies to examine the CELF2 localization in cultured NPCs isolated from the developing mouse cortex at embryonic day (E) 12.5 when the cortex mostly consists of NPCs (**Fig. S2A, B**). Many cultured NPCs (Sox2+) showed nuclear CELF2, but in some proliferating NPCs (Ki67+), CELF2 accumulated predominately in the cytoplasm, forming distinct foci (**Fig. 2A, S2C**). Immunostaining of cortical sections and subcellular fractionation analysis further confirmed the abundant presence of CELF2 in the cytoplasm of NPCs in the developing cortex (**Fig. 2B, C, S2D**). This distribution is not simply a consequence of nuclear membrane breakdown during mitosis and the diffusion of nuclear proteins into the cytoplasm, because cytoplasmic CELF2 was also evident in NPCs that were outside of mitosis (negative for mitosis-specific phospho-histone H3/pH3) and had an intact nuclear envelope (marked by Lap2) in the ventricular zone (VZ) (**Fig. 2D, S2E**) (Smoyer and Jaspersen, 2014). These data indicate that cortical NPCs actively maintain CELF2 in the cytoplasm regardless of cell cycle stages.

Interestingly, we found that newborn neurons (Satb2+), the postmitotic daughter cells of NPCs in the cortical plate (CP), had predominantly nuclear CELF2, in sharp contrast from that in NPCs in the VZ

(Fig. 2D, E). To examine how CELF2 changes its localization from the cytoplasm to the nucleus upon NPC differentiation into neurons, we examined cortical sections at E15.5, the peak of neurogenesis (Taverna et al., 2014). Although CELF2 was expressed throughout the cortex, nuclear accumulation of CELF2 became evident in cells moving away from the VZ toward the CP (**Fig. S2F**). Quantification of cells with enriched cytoplasmic CELF2 in eight equal-sized bins spanning the cortex showed that the transition of CELF2 distribution occurred around the interface of the subventricular zone (SVZ) and intermediate zone (IZ), where NPCs commit to a neuronal fate (**Fig. 2F**) (Taverna et al., 2014). Using the early neuronal marker β III-tubulin, we found that nearly all newly born neurons in the VZ/SVZ had CELF2 in the nucleus, which suggests that CELF2 becomes nuclear-localized as soon as the neuron is born (**Fig. 2G**). However, a strong nuclear CELF2 signal was also present in many non-neuronal cells (β III-negative) in the SVZ. These cells were immunopositive for Tbr2, a marker of intermediate progenitors (IPs) that are transient amplifying cells with a restricted fate of becoming neurons in one or two rounds of divisions (**Fig. 2H**) (Taverna et al., 2014). Over 85% of IPs had CELF2 accumulated in the nucleus, similar to that seen in β III-positive or Satb2-positive neurons but markedly different from Sox2-positive non-committed NPCs (**Fig. 2I**). These results suggest that CELF2 undergoes dynamic nucleocytoplasmic transport in NPCs and their nuclear translocation marks NPCs that undergo neurogenic differentiation (**Fig. 2J**).

Cytoplasmic CELF2 Orchestrates mRNAs Encoding Stem Cell Regulators and Disease-Related

Factors. In the cytoplasm, CELF2 formed distinct foci reminiscent of processing bodies (PBs), the ribonucleoprotein structures that store translationally repressed mRNAs (**Fig. S3A**) (Standart and Weil, 2018). PBs are critical for NPC fate decision, as they sequester mRNAs that encode pro-differentiation

factors (e.g. Neurogenins, NeuroDs) in self-renewing NPCs and release these mRNAs for translation when NPCs commit to neurogenic differentiation (Yang et al., 2014). Disruption of PBs causes premature neurogenesis (Yang et al., 2014). We found that CELF2 colocalized with the PB marker DCP1a (**Fig. 3A**). Proximity ligation assay (PLA) that detects protein-protein interactions showed the robust interactions of CELF2 with DCP1a and another PB core component 4E-T, but not with the negative control protein Ki67 in NPCs (**Fig. 3B, S3B, C**). These results confirm that CELF2 is part of PBs, which is consistent with its known role in translational repression and raise the possibility that CELF2 translocation is a timing controller of NPC differentiation by coordinating its target mRNAs in PBs (Subramaniam et al., 2008).

We thus sought to determine the target mRNAs of cytoplasmic CELF2, by performing RNA-immunoprecipitation (RIP)-sequencing with cytoplasmic RNA isolated from the E12.5 embryonic mouse cortex from three biological replicates (**Fig. S3D, E**). Distribution analysis of the input and CELF2 RIP groups showed that mapped reads were mainly enriched in exon regions with similar subgenic patterns, consistent with the isolation of mature mRNA in cytoplasmic fractions (**Fig. S3F**). Compared to the control IgG RIPs, CELF2 RIPs enriched *mcl1* mRNA, a known target of CELF2-mediated translational repression (**Fig. 3C**) (Subramaniam et al., 2008). Moreover, mRNAs encoding proneural differentiation factors that are translationally repressed in PBs, such as *neurog2*, *neurod1*, and *tbr2*, were also significantly enriched in CELF2 RIPs (**Fig. 3C**) (Yang et al., 2014). To define the full landscape of mRNAs coordinated by CELF2, we focused on a group of 2564 target mRNAs that were enriched more than 1.5-fold in CELF2 RIPs versus input and control IgG RIPs with a false-discovery rate (FDR) <0.05 (**Fig. S3G, Table S2**). We found that the CELF2-binding motif “UGUGUG” occurred

significantly more frequently in CELF2 targets versus the control IgG background, which supports the validity of our identified targets (**Fig. S3H**) (Ajith et al., 2016). Gene ontology (GO) enrichment analysis of these CELF2 targets revealed their tight association with neuron generation, supporting its role in NPC differentiation (**Fig. 3D**).

Other enriched molecular processes directly linked to neurogenesis are transcription and RNA metabolic process, kinase activity, and Wnt signaling. Moreover, a handful of CELF2 targets were related to histone modification, nuclear transport and protein ubiquitination. All of these processes are important regulatory nodes for gene expression, which suggests that CELF2 translocation may trigger a rapid and profound change in the expressed genome to drive NPC differentiation into neurons. Indeed, many CELF2 targets enriched in these GO categories are known regulators of neurogenesis and have been linked to human neurodevelopmental disorders that share similar clinical features seen in the patients with *CELF2* variants (**Fig. 3E**). For example, disruption of *FMRI* and *DDX3X* that are involved in RNA metabolisms leads to impaired neurogenesis and is linked to intellectual disability (ID) and cortical malformations (Lennox et al., 2020; Mila et al., 2018; Snijders Blok et al., 2015; Tervonen et al., 2009). *DYRK1* and *PTEN* participate in protein phosphorylation and are known regulators of neurogenesis and associated with cortical malformations (van Bon et al., 2016; Crino, 2015; Park and Chung, 2013; Veleva-Rotse and Barnes, 2014). Mutations in transcription and chromatin regulators *FOXP1* and *SETBP1* perturb neurogenesis and cause speech impairment (Filges et al., 2011; Mitter et al., 2018; Piazza et al., 2018). This suggests that alterations in CELF2 function might have a broad impact on disease-related mRNAs, which may converge on neurogenesis dysregulation and collectively contribute to a series of clinical manifestations in our cohort of patients. To evaluate this potential impact, we

examined known disease genes associated with ASD (SFARI genes and Geisinger ASD genes that are not associated with ID), speech and language capacities, ID, and gyral pattern abnormality. The mRNAs of these disease genes all showed a significantly higher enrichment in CELF2 RIPs, compared to control mRNAs randomly selected from the transcriptome (**Fig. 3F, G, S3I, Table S3**). Remarkably, 59.7% of highest confidence ASD genes (category 1 in the SFARI database) and 47.6% of speech and language-related genes that are expressed in NPCs corresponded to CELF2 targets. This enrichment of neurodevelopmental disorder mRNAs was specific because mRNAs associated with chronic neurological and non-neurological conditions (e.g. Parkinson's disease, liver disease) were not enriched in CELF2 RIPs (**Fig. S3J**). Thus, cytoplasmic CELF2 coordinates a group of mRNAs that are important for NPC fate decisions and translationally suppressed in PBs and that when perturbed cause abnormal brain development.

Cytoplasmic CELF2 Regulates NPC Fate Decision. Our results point to a potential role of CELF2 as a molecular switch to control NPC differentiation by translocating between the nucleus and cytoplasm. To test whether CELF2 has a functional effect on NPC fate decision, we first knocked down *CELF2* in cultured NPCs isolated from the E12.5 cortex with small hairpin RNAs (shRNAs) and cotransfected with an EGFP plasmid to label NPCs and the neurons they give rise to (**Fig. S4A**). Along with reduced *CELF2* expression in EGFP-positive cells, *CELF2* knockdown caused a twofold decrease in the proportion of NPCs (EGFP+, Sox2+), with a concomitant increase in neuronal numbers (EGFP+, β III-tubulin+) (**Fig. 4A-D, S4B, C**). These results indicate that CELF2 regulates NPC maintenance by preventing NPC differentiation into neurons in cultures. We next assessed whether CELF2 also regulates NPC fate decision *in vivo*. To this end, we performed *in utero* electroporation in the mouse cortex at

E13.5 with plasmids expressing EGFP and control or *CELF2* shRNAs and examined the impact on NPC self-renewal versus differentiation into neurons that migrate toward the CP. Control cortical sections showed EGFP-positive cells throughout the cortex, with about 25% of them remaining in the VZ NPC zone (**Fig. 4E, F, S4D, E**). In contrast, *CELF2* knockdown produced a three-fold reduction of EGFP-positive cells in the VZ, in agreement with a 50% decrease in the proportion of radial precursors (EGFP+, Pax6+) (**Fig. 4G, H**). *CELF2* knockdown also resulted in a concomitant increase of EGFP+ cells in the IZ, where migrating newborn neurons reside. To determine if the change in cell distribution reflected increased neurogenic differentiation of NPCs, we analyzed cells immunostained for the neuronal marker *Satb2* and the IP marker *Tbr2*, since NPCs at E13.5 mainly produce neurogenic IPs that give rise to neurons expressing *Satb2* (Yang et al., 2014). *CELF2* knockdown significantly increased EGFP-positive, *Tbr2*-positive IPs as well as *Satb2*-positive neurons (**Fig. 4I-L**). Together, these results indicate that *CELF2* maintains NPC identity by preventing neurogenic differentiation.

To ask if the cytoplasmic function of *CELF2* is critical for NPC maintenance, we used a knockdown rescue approach, re-introducing either WT or a mutant version of shRNA-resistant *CELF2* that was fused to the nuclear export signal (NES). The addition of the NES to *CELF2* resulted in its exclusion from the nucleus to the cytoplasm (**Fig. S4F**). We found that both *CELF2* WT and NES mutant were sufficient to rescue the reduction in NPCs and the increase in neurons induced by *CELF2* knockdown (**Fig. 4M, N**). Moreover the expression of *CELF2-NES* further increased the proportion of NPCs over neurons. These data suggest that cytoplasmic *CELF2* maintains NPC identity and the defective nuclear translocation of *CELF2* disrupts NPC differentiation into neurons.

Discussion:

The self-renewal and differentiation of NPCs rely on the coordinated regulation of mRNAs, and our findings reveal a mechanism for how this coordination is achieved to instruct NPC fate decisions.

Cytoplasmic CELF2 binds and represses specific mRNAs in PBs to maintain NPCs, and its translocation to the nucleus releases these mRNAs for translation, which induces neurogenic differentiation. Loss of CELF2 in the cytoplasm compromises NPC self-renewal, whereas aberrant accumulation of cytoplasmic CELF2 impairs neurogenesis, causing defects in cortical development. It is conceivable that the nucleocytoplasmic transport could couple CELF2's multiple functions in different subcellular compartments. For example, in the nucleus, CELF2 participates in pre-mRNA alternative splicing, a prevalent process that also contributes to NPC fate decisions (Raj and Blencowe, 2015). The combinatorial functions of CELF2 and its interactions with other RBPs (e.g. RBFOX) might allow an elaborate regulation of NPC fates through the precise control of mRNAs at different stages of gene expression (Gazzara et al., 2017; Zhang et al., 2016).

In NPCs, cytoplasmic CELF2 coordinates mRNAs of known disease genes whose altered expression results in neurodevelopmental disorders with features such as ASD, speech delay, ID, epilepsy or cortical malformations. Disruption of CELF2 nucleocytoplasmic transport may therefore cause widespread and pathogenic changes in the expression of these genes at the mRNA level, collectively contributing to the clinical manifestations of related symptoms. Indeed, our results show that individuals with a rare neurodevelopmental disorder in which CELF2 transport is perturbed due to a deleterious genetic change present with clinical features including epilepsy, ASD, speech disorder, ID and cortical malformations. By extension, alterations in CELF2 and other RBPs due to non-genetic perturbations

may produce similar pathogenic impact, thus contributing more broadly to idiopathic ASD and other neurodevelopmental disorders (Parras et al., 2018). Continued studies on modulators of CELF2 transport might provide insights into therapeutic strategies for these patients.

Acknowledgement:

We thank the patients and their families for participating in our study. This work was funded by CIHR Project Grant (PJT-165821 to G.Y.), Mechanism Network Modeling Rare Disease research grant (27R21814, to G.Y., M.I., M.J.M.). G.Y. is a Canada Research Chair. We thank Drs. Paul Mains and Deborah Kurrasch for their critical reading of the manuscript and suggestions.

Contribution

M.J.M. collected and analyzed clinical data, performed and analyzed some experiments and co-wrote the paper. S.L.E. performed and analyzed experiments and co-wrote the paper. D.K. performed bioinformatics analyses of RIP-seq data and co-wrote the paper. P.W. performed and analyzed experiments. M.A. and Q.L. helped with bioinformatics analysis. A.V., Q.T., N.G., W.B., L. A-W, K.G., O.B. and L.F. collected and analyzed clinical data. X-C.W. analyzed clinical MRI data. M.I. conceptualized the project and collected and analyzed clinical data. G.Y. conceptualized the project, designed, performed, and analyzed some experiments and co-wrote the paper.

Competing Interests

M.M. is employed by Blueprint Genetics. The other authors declare no competing interests.

Methods

Ethical statement: Verbal and written consent for publication has been obtained from all study subjects or their parents in accordance with the various local institutional requirements.

Patients and Clinical Investigations: The patients were all evaluated by a clinical geneticist and underwent trio exome sequencing as part of their diagnostic workup for a suspected genetic condition associated with significant neurological impairment. Heterozygous *de novo* variants were identified in the candidate gene *CELF2* in all patients. The cohort of five patients was assembled through the GeneMatcher program (Sobreira et al., 2015). Clinical information regarding the patient phenotypes was provided by the patients' physicians. Where possible, the MRI scans of the brain were reviewed by a single pediatric neuroradiologist to consistently delineate the brain malformations.

Animals: All animal use was approved by the Animal Care Committee at the University of Calgary and conformed to the guidelines set forth by the Canadian Council of Animal Care. Female CD1 mice (8–12 weeks old), purchased from Charles River Laboratory, were used for all animal experiments. Mice were housed in groups of one to five per cage, with a standard 12-hour light-dark cycle in a room maintained at 24°C with free access to water and chow. The morning of vaginal plug detection was designated as embryonic day (E) 0.5.

Plasmids and reagents: The pEF1 α -EGFP plasmid expressing nuclear EGFP has been previously reported (Yang et al., 2014). Mouse CELF1, CELF2 and CELF3 isoforms were cloned into pEGFP-C1 and pCAGIG vector (addgene #11159) using cDNA prepared from mouse cortical tissues with the

following primers:

CELF1, forward-AATAAGCTTCGATGGCTGCGTTTAAGTTGGATTTCC, reverse-

AATGGATCCTCAGTAGGGCTTACTATCATTCTTCGAAC; *CELF2*, forward-

ATTCTCGAGCTATGCGCTGTCCCAAATCCG, reverse-

ATTGGATCCTCAGTAAGGTTTGCTGTCGTTTTTGG; *CELF3*, forward-

ATTCTCGAGCTATGAAGGAGCCAGATGCCATC, reverse-

ATTGGATCCTTAGTAGGGCCTGTTTGCATCC. Patient-derived pathogenic variants in *CELF2* were

generated using the site-directed mutagenesis kit (NEB). To generate *CELF2*-NES, the nuclear exclusion

signal derived from PKI α (LALKLAGLDI) was fused to *CELF2* by PCR and sub-cloned into pEGFP-

C1 and pCAGIG. To generate shRNAs against *CELF2*, double stranded oligonucleotides encoding

shRNA sequences (from Life Technologies) were cloned into the pSUPER vector. The shRNA

sequences used were *CELF2* shRNA-1 (GCAATGGCACAGAATGCAA), *CELF2* shRNA-2

(GGTTGTTGTTTCGTAACAT), *CELF2* shR-3 (CCTCTTTATTTACCACCTT). All clones were

verified by sequencing.

Cell culture and transfections: HEK 293T cells were cultured in Dulbecco's modified Eagle Medium (DMEM) supplemented with 10% fetal bovine serum (FBS; Wisent) and 1% penicillin-streptomycin (P/S) antibiotic. Cells were cultured in a humidified incubator at 37°C with 5% CO₂. For transfection, HEK293T cells were transfected at 85% confluency using polyethylenimine (PEI). For the validation of shRNAs, pEGFP-*CELF2* were cotransfected with control or corresponding shRNAs (pSUPER) at 1:1 ratio, followed by Western blot analysis 24 hours later. For NPC cultures, E12.5 cortical precursors were cultured in Neurobasal Medium (Thermo Fisher), supplemented with FGF2, B27, P/S, L-Glutamine at a

density of 250,000 cells/mL on coverslips coated with laminin (5mg/mL; Corning) and poly-D-lysine (1mg/mL; Sigma). For NPC transfection, cells were transfected with Lipofectamine Stem Transfection Reagent (Thermo Fisher) according to the manufacturer's protocol.

Antibodies: The primary antibodies used were goat anti-Sox2 (R&D Systems #AF2018—1:500), mouse anti- β III-tubulin (Biolegend #801201—1:1000), rabbit anti-CELF2 (Abcam #ab186430—1:1000), mouse anti-CELF2 (Santa Cruz Biotechnology #sc47731—1:100), mouse anti-Ki67 (BD Biosciences #550609—1:200), rabbit anti-Csde1 (Abcam #ab201688—1:1000), mouse anti-LaminB1 (Abcam #ab8982—1:200), mouse anti-Lap2 (BD Biosciences #611000—1:200), mouse anti-Satb2 (Abcam #ab51502—1:400), rat anti-Tbr2 (Invitrogen #14-4875-80—1:500), rabbit anti-Pax6 (Biolegend #901301—1:2000), mouse anti-pH3 (Abcam—1:500), mouse anti-GFP (Abcam #ab290—1:5000), chicken anti-GFP (EMD Millipore #AB16901—1:1000), rabbit anti-GAPDH (Sigma #G9545—1:5000), mouse anti-Dcp1a (Abnova #H00055802-M06—1:100), rabbit anti-4E-T (Cell signalling #2297S—1:100). Donkey anti-mouse, anti-rabbit, anti-rat and anti-chicken Alexa Fluor 488, 555 and 647-conjugated secondary antibodies were obtained from Thermo Fisher and used at 1:1000 dilutions. HRP-conjugated goat anti-mouse and anti-rabbit secondary antibodies were purchased from Thermo Fisher and used at a concentration of 1:5000.

Quantitative real time-PCR and reverse transcription-PCR: Total RNA was extracted using Trizol (Thermo Fisher) as per the manufacturer's protocol. Total RNA was reverse transcribed using 1 μ g of RNA with the Maxima H Minus cDNA Synthesis Kit (Thermo Fisher). For reverse transcription-PCR (RT-PCR), 1ng of cDNA was used with Phusion High Fidelity Master Mix (Thermo Fisher) with

amplification for 30 cycles and an annealing temperature of 60°C. Quantitative real-time PCR (qPCR) was performed using PerfeCTa SYBR Green FastMix (Quanta Bio) with amplification for 40 cycles and an annealing temperature of 60°C, using the BioRad CFX system. Threshold cycle (Ct) values were obtained from the amplification plots and used to calculate fold change using the $\Delta\Delta C_t$ method. In all samples, *GAPDH* was used as the endogenous control. The primers used in this study are: Human *PPP1R12A*, forward- CCAAGCACCATCAACACCAACAG, reverse- CTGTGTTGATCTTCTAGATTGTCTTGC; Human *TRAF3*, forward- CACGAAGACACCGACTGTCCCTGC, reverse-CGAGCGAGTTGCTCCACTCCTTCAGC; Mouse *CELF1*, forward-ACAGATGAAGCCTGCTGACA, reverse-CTCTGCTCAAGCCATCAGGT; Mouse *CELF2*, forward-GAGCCACTGTCCGATTGAATA, reverse-GCGCCAAGTCCTCCATTC.

Immunoblotting: For transfected HEK293T cells, cells were lysed with 1x sample buffer containing 1mM dithiothreitol (DTT), boiled for 8 minutes, and analyzed with SDS-PAGE analysis. For the isolation of nuclear and cytoplasmic fractions from brain tissues, E12 cortical tissues were lysed with the cytoplasmic lysis buffer (150 mM KCl, 25 mM Tris-HCl pH 7.4, 1 mM MgCl₂ and 1 mM DTT), and incubated on ice for 15 minutes with frequent vortex, followed by centrifuged at 800 g at 4°C for 5 minutes to collect the cytoplasmic contents in the supernatant. The pellets that contain the nuclear contents were washed with the lysis buffer for three times and boiled in 1x sample buffer with DTT for 8 minutes followed by SDS-PAGE analysis. After electrophoresis, proteins were transferred to a nitrocellulose membrane (GE HealthCare), which were blocked with 5% BSA (BioShop) in TBS-T for 1 hour at room temperature, probed with the indicated primary antibodies in 5% BSA in TBS-T overnight at 4°C and HRP-conjugated secondary antibodies in TBS-T for 1 hour at room temperature. Membranes

were imaged using the Amersham Imager 600 (GE HealthCare), and quantified using ImageJ Software (NIH).

Immunofluorescence: For immunostaining of cortical sections, embryonic brains were dissected in ice-cold HBSS, fixed in 4% paraformaldehyde at 4°C overnight, cryopreserved with 30% sucrose and cryosectioned coronally at 16 µm. Sections were blocked at room temperature with 5% BSA (Jackson ImmunoResearch) in PBS with 0.3% Triton X-100, and incubated with primary antibodies in blocking buffer overnight at 4°C, followed by incubation with the appropriate secondary antibodies in PBS at room temperature for 1 hour. Nuclei were stained with Hoechst 33258 (Sigma). For immunostaining of cultured HEK293T and NPCs, cells on glass coverslips were fixed with 4% paraformaldehyde for 15 minutes at room temperature, followed by permeabilization with 0.3% Triton X-100 for 10 minutes at room temperature. Cells were incubated with primary antibodies in PBS for 1 hour at room temperature, followed by incubation with the appropriate secondary antibodies diluted in PBS for 1 hour. Nuclei were stained with Hoechst 33258 (Sigma). Coverslips were mounted on glass slides with Aqua-Poly/Mount (Polysciences).

Proximity ligation assay: The Duolink In Situ Orange Starter Kit Mouse/Rabbit (Sigma) was used and the manufacturer's protocol was followed, with modifications. Briefly, embryonic brain sections or cultured NPC's (as described above) were blocked with 5% normal donkey serum (NDS) with 0.3% Triton X-100 for 1 hour at room temperature. Primary antibodies were added in 5% NDS and incubated for 1 hour at room temperature (cultured NPC) or overnight at 4°C (embryonic brain sections). PLA

probes were then added, followed by ligation, amplification and mounting, according to the manufacturer's protocol. All steps were performed in a humidified chamber.

In utero electroporation: Expression constructs for nuclear EGFP or pCAGIG-CELF2 were co-electroporated with shRNAs constructs at a 1:3 ratio. Plasmids were mixed with 0.5% trypan blue in ddH₂O at 4 μg/μL and injected into lateral ventricles using mouth pipetting. Following injection, the square electroporator CUY21 EDIT (TR Tech, Japan) was used to deliver five 50 ms pulses of 40-50 V with 950 ms intervals per embryo as described previously (Yang et al., 2014). Brains were dissected and analyzed three days post electroporation.

Microscopy and quantification: Transfected HEK293T and NPCs were immunostained and imaged with a 60X objective on an Olympus FV3000 confocal microscope or an SD-OSR microscope. For cell cultures, representative images were selected from at least three independent experiments that were used for quantifications. For the analysis of embryonic brains with *in utero* electroporation, three to four anatomically matched sections per brain from three embryos of two independent mothers for each group were imaged with a 20X on the confocal microscope. For quantification of CELF2 localization in marker-positive cells, the immunointensity in the nucleus (stained by Hoechst 33258) and the surrounding region (cytoplasm) were measured as shown in figure S2H, and cells with higher intensity in the surrounding region were considered as those with more cytoplasmic CELF2; cortical sections from four embryos were used for the analysis. For quantification of CELF2 levels in cultured NPCs, EGFP-positive cells were categorized into three groups based on CELF2 immunointensity when comparing to the surrounding EGFP-negative cells; EGFP-positive cells that had equivalent intensity

were grouped as high intensity. For measurement of CELF2 levels in electroporated cortical sections, intensity of CELF2 immunostaining in EGFP-positive and negative cells (transfected versus non-transfected, 25 cells each) were measured using ImageJ Software (arbitrary unit/a.u.). For analysis of CELF2 granules, colocalization and PLA, z-stacked images were collected at 60X magnification on the FV3000 confocal microscope and projected with maximum intensity using ImageJ Software. The colocalized foci were verified in each section of the stack. All images were processed with Fluoview software (Olympus), ImageJ (NIH), and Adobe Photoshop. Pax6, Tbr2, and Hoechst staining were used to define the VZ, SVZ, and CP, respectively.

RNA immunoprecipitation-sequencing and processing: E12.5 cortices were used for RNA-RIP-seq using the EZ-Magna RIP-RNA Immunoprecipitation Kit (Millipore) according to the manufacturer's instructions with modifications. Cortical tissues were lysed with the cytoplasmic lysis buffer supplemented with the protease and RNase inhibitors (from kit) and 100 U/ml SUPERasine-IN (Ambion), followed by centrifugation to isolate the cytoplasmic content in the supernatant. A portion of the cytoplasmic fraction was directly used for RNA extraction and served as "Input". The rest of the supernatant was precleared with protein A/G beads and incubated with 5 µg of anti-CELF2 (rabbit, Abcam) or normal rabbit IgG for 3 hours at 4°C. Total RNA from three biological replicates of Input, IgG control and CELF2 immunoprecipitates was isolated, extracted with phenol/chloroform, subjected to rRNA depletion using the NEBNext rRNA depletion kit (E6350S), and converted into cDNA libraries using the NEBNext Ultra II Directional RNA library prep kit (E7760L), followed by Illumina sequencing with NEXTseq 500 (1 x 75bp reads).

These files were pseudo aligned with the mouse transcriptome (GRCm38 – mm10) using Pachter lab's Kallisto (version 0.46.1) (Pimentel et al., 2017). Sequence alignment used bias correction and bootstrapping was done 100 times for validation. To fit the sequencing method, alignment was conducted as single end reads, with a fragment length of 181 base-pairs and a standard deviation of 20 base pairs. Differential expression and quality control was conducted with Pachter lab's Sleuth (version 0.30.0) (O'Leary et al., 2016). The Sleuth filter function was left as default with the minimum reads set to 5 and the minimum proportion of reads set to 0.47. When using the sleuth preparation, gene isoforms were aggregated with a target map file derived from the RefSeq mm10 transcriptome. For quality control, principle component analysis (PCA) was conducted to assess the reproducibility between samples and distinction between conditions. Pearson correlation analysis was used to further assess the quality control by assessing similarity between samples and allowing for a graphical heatmap comparison.

Differential expression and bioinformatics: Sleuth was used to analyze the three conditions (CELF2 RIP, IgG RIP and Total RNA-seq) as covariates. TPM frequency was analyzed graphically to determine 10976 genes with a natural log value >1.52 as robustly expressed genes. Genes with a fold change threshold of 1.5 (both CELF2 RIP over Input and CELF2 RIP over IgG RIP) and False Discovery Rate (FDR) <0.05 were defined as CELF2 targets. CELF2 target genes were analyzed for enriched GO groups (Biological Process-EBI-UniProt-GOA, 2020-02-11) using the ClueGO (v2.5.4) and CluePedia (v1.5.4) plugins of Cytoscape (v3.7.1). GO terms with $p < 0.0001$ (Bonferroni step down correction) were used for plotting. To determine the feature counts of the RNA-seq reads, the sample files were aligned using STAR (version 2.7.3a) to the mouse genome (GRCm38-mm10) (Dobin et al., 2013). The RSeQC

package's read distribution script (version 3.0.1) was used to obtain the gene features for the aligned genes (Wang et al., 2012). The RSeQC reference would be the same mm10 mouse genome that was used for the STAR alignment. This gave the tags per kilobase for the CDS exons, 5' UTR exons, 3' UTR exons, introns, TSS upstream 1 kilobase, TSS upstream 5 kilobases, TSS upstream 10 kilobases, TES downstream 1 kilobase, TES downstream 5 kilobases and TES downstream 10 kilobases. The TTS values were aggregated due to their tags per kilobase being less than 2 – making it difficult to distinguish on a graph, the same is true for the TES values. After these values were obtained for each sample the mean tags per kilobase was calculated for each condition. To better understand the features for the CELF2 target genes, the gene list acquired via differential expression analysis were obtained from the USCS genome browser and used to filter the CELF2 genome data. The read distribution script was conducted using the filtered genes for all three samples and described as CELF2 targets.

Disease gene enrichment analysis: ASD candidate genes were obtained from the Simons Foundation Autism Research Initiative (SFARI) database, referred as SFARI categories 1-3. A group of speech and language related genes were summarized by Graham and Fisher (Graham and Fisher, 2013). Intellectual disability (ID) related genes and a list of ASD genes that are not associated with ID were obtained from Geisinger Developmental Brain Disorder database. Genes related to speech delay, gural pattern abnormality, Parkinson's disease and liver disease were obtained from OMIM (see **table S3** for terms used in search). We also used a group of ID-related genes from the Geisinger Developmental Brain Disorder Genes database. Genes with an expression level above natural log value 1.52 were used for plotting the enrichment cumulative curves. For comparison, lists of 500 genes randomly chosen from the transcriptome were used as controls. P values were calculated with the Mann-Whitney U test.

Statistics: All data were expressed as the mean plus or minus the standard error of the mean (SEM).

Statistical analyses were performed with a two-tailed Student's t-test or, where relevant, ANOVA with Tukey or Dunnett's post-hoc tests, unless otherwise indicated.

References:

- Ajith, S., Gazzara, M.R., Cole, B.S., Shankarling, G., Martinez, N.M., Mallory, M.J., and Lynch, K.W. (2016). Position-dependent activity of CELF2 in the regulation of splicing and implications for signal-responsive regulation in T cells. *RNA Biol* 13, 569–581.
- Alsina, F.C., and Silver, D.L. (2020). Posttranscriptional and translational control of neurogenesis. In *Patterning and Cell Type Specification in the Developing CNS and PNS*. Neuroscience 0, 499–528.
- Amadei, G., Zander, M.A., Yang, G., Dumelie, J.G., Vessey, J.P., Lipshitz, H.D., Smibert, C.A., Kaplan, D.R., and Miller, F.D. (2015). A smaug2-based translational repression complex determines the balance between precursor maintenance versus differentiation during mammalian neurogenesis. *J. Neurosci.*, 35,15666-15681
- Bae, B. II, Jayaraman, D., and Walsh, C.A. (2015). Genetic changes shaping the human brain. *Dev Cell*. 32, 423–434.
- Bae, B.I., Tietjen, I., Atabay, K.D., Evrony, G.D., Johnson, M.B., Asare, E., Wang, P.P., Murayama, A.Y., Im, K., Lisgo, S.N., et al. (2014). Evolutionarily dynamic alternative splicing of GPR56 regulates regional cerebral cortical patterning. *Science* 343, 764–768.
- Barone, R., Fichera, M., De Grandi, M., Battaglia, M., Lo Faro, V., Mattina, T., and Rizzo, R. (2017). Familial 18q12.2 deletion supports the role of RNA-binding protein CELF4 in autism spectrum disorders. *Am J Med Genet A* 173, 1649–1655.
- van Bon, B.W., Coe, B.P., Bernier, R., Green, C., Gerds, J., Witherspoon, K., Kleefstra, T., Willemsen, M.H., Kumar, R., Bosco, P., et al. (2016). Disruptive de novo mutations of DYRK1A lead to a syndromic form of autism and ID. *Mol Psychiatry* 21, 126–132.
- Crino, P.B. (2015). The enlarging spectrum of focal cortical dysplasias. *Brain* 138, 1446–1448.

Dasgupta, T., and Ladd, A.N. (2012). The importance of CELF control: Molecular and biological roles of the CUG-BP, Elav-like family of RNA-binding proteins. *Wiley Interdiscip. Rev. RNA* 3, 104-21.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.

Dougherty, J.D., Maloney, S.E., Wozniak, D.F., Rieger, M.A., Sonnenblick, L., Coppola, G., Mahieu, N.G., Zhang, J., Cai, J., Patti, G.J., et al. (2013). The disruption of *Celf6*, a gene identified by translational profiling of serotonergic neurons, results in autism-related behaviors. *J Neurosci* 33, 2732–2753.

Filges, I., Shimojima, K., Okamoto, N., Röthlisberger, B., Weber, P., Huber, A.R., Nishizawa, T., Datta, A.N., Miny, P., and Yamamoto, T. (2011). Reduced expression by SETBP1 haploinsufficiency causes developmental and expressive language delay indicating a phenotype distinct from Schinzel-Giedion syndrome. *J Med Genet* 48, 117–122.

Gazzara, M.R., Mallory, M.J., Roytenberg, R., Lindberg, J.P., Jha, A., Lynch, K.W., and Barash, Y. (2017). Ancient antagonism between CELF and RBFOX families tunes mRNA splicing outcomes. *Genome Res* 27, 1360–1370.

Gilmore, E.C., and Walsh, C.A. (2013). Genetic causes of microcephaly and lessons for neuronal development. *Wiley Interdiscip. Rev. Dev. Biol.* 2, 461-78.

Graham, S.A., and Fisher, S.E. (2013). Decoding the genetics of speech and language. *Curr Opin Neurobiol* 23, 43–51.

Groszer, M., Erickson, R., Scripture-Adams, D.D., Dougherty, J.D., Le Belle, J., Zack, J.A., Geschwind, D.H., Liu, X., Kornblum, H.I., and Wu, H. (2006). PTEN negatively regulates neural stem cell self-renewal by modulating G0-G1 cell cycle entry. *Proc. Natl. Acad. Sci. U. S. A.* 103, 111-116.

Gruber, R., Zhou, Z., Sukchev, M., Joerss, T., Frappart, P.O., and Wang, Z.Q. (2011). MCPH1 regulates

the neuroprogenitor division mode by coupling the centrosomal cycle with mitotic entry through the Chk1-Cdc25 pathway. *Nat. Cell Biol.* 13, 1325–1334

Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., Chen, H., Wang, J., Tang, H., Ge, W., et al. (2020). Construction of a human cell landscape at single-cell level. *Nature* 581, 303-309

Johnson, M.B., Wang, P.P., Atabay, K.D., Murphy, E.A., Doan, R.N., Hecht, J.L., and Walsh, C.A. (2015). Single-cell analysis reveals transcriptional heterogeneity of neural progenitors in human cortex. *Nat Neurosci* 18, 637–646.

Ladd, A.N., and Cooper, T.A. (2004). Multiple domains control the subcellular localization and activity of ETR-3, a regulator of nuclear and cytoplasmic RNA processing events. *J. Cell Sci.* 117, 3519-3529.

Lennox, A.L., Hoyer, M.L., Jiang, R., Johnson-Kerner, B.L., Suit, L.A., Venkataramanan, S., Sheehan, C.J., Alsina, F.C., Fregeau, B., Aldinger, K.A., et al. (2020). Pathogenic DDX3X Mutations Impair RNA Metabolism and Neurogenesis during Fetal Cortical Development. *Neuron* 106, 404-420

Mila, M., Alvarez-Mora, M.I., Madrigal, I., and Rodriguez-Revenga, L. (2018). Fragile X syndrome: An overview and update of the FMR1 gene. *Clin Genet* 93, 197–205.

Mitter, D., Pringsheim, M., Kaulisch, M., Plümacher, K.S., Schröder, S., Warthemann, R., Abou Jamra, R., Baethmann, M., Bast, T., Büttel, H.M., et al. (2018). FOXP1 syndrome: genotype-phenotype association in 83 patients with FOXP1 variants. *Genet Med* 20, 98–108.

Mohammad, L., Wiseman, J., Erickson, S., Yang, G., Mohammad, L., Wiseman, J., Erickson, S., and Yang, G. (2019). Protein Synthesis and Translational Control in Neural Stem Cell Development and Neurogenesis. In *The Oxford Handbook of Neuronal Protein Synthesis*, DOI: 10.1093/oxfordhb/9780190686307.013.21

O’Leary, N.A., Wright, M.W., Brister, J.R., Ciufu, S., Haddad, D., McVeigh, R., Rajput, B., Robbertse,

B., Smith-White, B., Ako-Adjei, D., et al. (2016). Reference sequence (RefSeq) database at NCBI: current status, taxonomic expansion, and functional annotation. *Nucleic Acids Res* *44*, D733-45.

Park, J., and Chung, K.C. (2013). New Perspectives of Dyrk1A Role in Neurogenesis and Neuropathologic Features of Down Syndrome. *Exp. Neurobiol.* *22*, 244–248

Parras, A., Anta, H., Santos-Galindo, M., Swarup, V., Elorza, A., Nieto-González, J.L., Picó, S., Hernández, I.H., Díaz-Hernández, J.I., Belloc, E., et al. (2018). Autism-like phenotype and risk gene mRNA deadenylation by CPEB4 mis-splicing. *Nature* *560*, 441–446.

Piazza, R., Magistroni, V., Redaelli, S., Mauri, M., Massimino, L., Sessa, A., Peronaci, M., Lalowski, M., Soliymani, R., Mezzatesta, C., et al. (2018). SETBP1 induces transcription of a network of development genes by acting as an epigenetic hub. *Nat. Commun.* *9*, 2192

Pimentel, H., Bray, N.L., Puente, S., Melsted, P., and Pachter, L. (2017). Differential analysis of RNA-seq incorporating quantification uncertainty. *Nat Methods* *14*, 687–690.

Pollen, A.A., Nowakowski, T.J., Chen, J., Retallack, H., Sandoval-Espinosa, C., Nicholas, C.R., Shuga, J., Liu, S.J., Oldham, M.C., Diaz, A., et al. (2015). Molecular identity of human outer radial glia during cortical development. *Cell* *163*, 55–67.

Raj, B., and Blencowe, B.J. (2015). Alternative Splicing in the Mammalian Nervous System: Recent Insights into Mechanisms and Functional Roles. *Neuron* *87*, 14–27.

Rakic, P. (2009). Evolution of the neocortex: a perspective from developmental biology. *Nat Rev Neurosci* *10*, 724–735.

Rodrigues, D.C., Harvey, E.M., Suraj, R., Erickson, S.L., Mohammad, L., Ren, M., Liu, H., He, G., Kaplan, D.R., Ellis, J., et al. (2020). Methylglyoxal couples metabolic and translational control of Notch signalling in mammalian neural stem cells. *Nat Commun* *11*, 2018.

Schultz, A.S., Preussner, M., Bunse, M., Karni, R., and Heyd, F. (2017). Activation-Dependent TRAF3 Exon 8 Alternative Splicing Is Controlled by CELF2 and hnRNP C Binding to an Upstream Intronic Element. *Mol Cell Biol* 37, e00488-16

Shitamukai, A., and Matsuzaki, F. (2012). Control of asymmetric cell division of mammalian neural progenitors. *Dev. Growth Differ.* 54, 277-286

Smoyer, C.J., and Jaspersen, S.L. (2014). Breaking down the wall: the nuclear envelope during mitosis. *Curr Opin Cell Biol* 26, 1–9.

Snijders Blok, L., Madsen, E., Juusola, J., Gilissen, C., Baralle, D., Reijnders, M.R., Venselaar, H., Helsmoortel, C., Cho, M.T., Hoischen, A., et al. (2015). Mutations in DDX3X Are a Common Cause of Unexplained Intellectual Disability with Gender-Specific Effects on Wnt Signaling. *Am J Hum Genet* 97, 343–352.

Sobreira, N., Schiettecatte, F., Valle, D., and Hamosh, A. (2015). GeneMatcher: a matching tool for connecting investigators with an interest in the same gene. *Hum Mutat* 36, 928–930.

Standart, N., and Weil, D. (2018). P-Bodies: Cytosolic Droplets for Coordinated mRNA Storage. *Trends Genet.* 34, 612-626

Subramaniam, D., Natarajan, G., Ramalingam, S., Ramachandran, I., May, R., Queimado, L., Houchen, C.W., and Anant, S. (2008). Translation inhibition during cell cycle arrest and apoptosis: Mcl-1 is a novel target for RNA binding protein CUGBP2. *Am J Physiol Gastrointest Liver Physiol* 294, G1025-32.

Subramanian, L., Calcagnotto, M.E., and Paredes, M.F. (2019). Cortical Malformations: Lessons in Human Brain Development. *Front Cell Neurosci* 13, 576.

Sun, T., and Hevner, R.F. (2014). Growth and folding of the mammalian cerebral cortex: from molecules to malformations. *Nat. Rev. Neurosci.* 15, 217-232

Taverna, E., Götz, M., and Huttner, W.B. (2014). The cell biology of neurogenesis: toward an understanding of the development and evolution of the neocortex. *Annu. Rev. Cell Dev. Biol.* 30, 465-502

Tervonen, T.A., Louhivuori, V., Sun, X., Hokkanen, M.E., Kratochwil, C.F., Zebryk, P., Castrén, E., and Castrén, M.L. (2009). Aberrant differentiation of glutamatergic cells in neocortex of mouse model for fragile X syndrome. *Neurobiol. Dis.* 33, 250-259

Veleva-Rotse, B.O., and Barnes, A.P. (2014). Brain patterning perturbations following PTEN loss. *Front. Mol. Neurosci.* 14;7:35

Wang, L., Wang, S., and Li, W. (2012). RSeQC: quality control of RNA-seq experiments. *Bioinformatics* 28, 2184–2185.

Yang, G., Smibert, C.A., Kaplan, D.R., and Miller, F.D. (2014). An eIF4E1/4E-T complex determines the genesis of neurons from precursors by translationally repressing a proneurogenic transcription program. *Neuron* 84, 723-739

Zahr, S.K., Yang, G., Kazan, H., Borrett, M.J., Yuzwa, S.A., Voronova, A., Kaplan, D.R., and Miller, F.D. (2018). A Translational Repression Complex in Developing Mammalian Neural Stem Cells that Regulates Neuronal Specification. *Neuron* 97, 520-537.e6.

Zhang, X., Chen, M.H., Wu, X., Kodani, A., Fan, J., Doan, R., Ozawa, M., Ma, J., Yoshida, N., Reiter, J.F., et al. (2016). Cell-Type-Specific Alternative Splicing Governs Cell Fate in the Developing Cerebral Cortex. *Cell* 166, 1147-1162.e15.

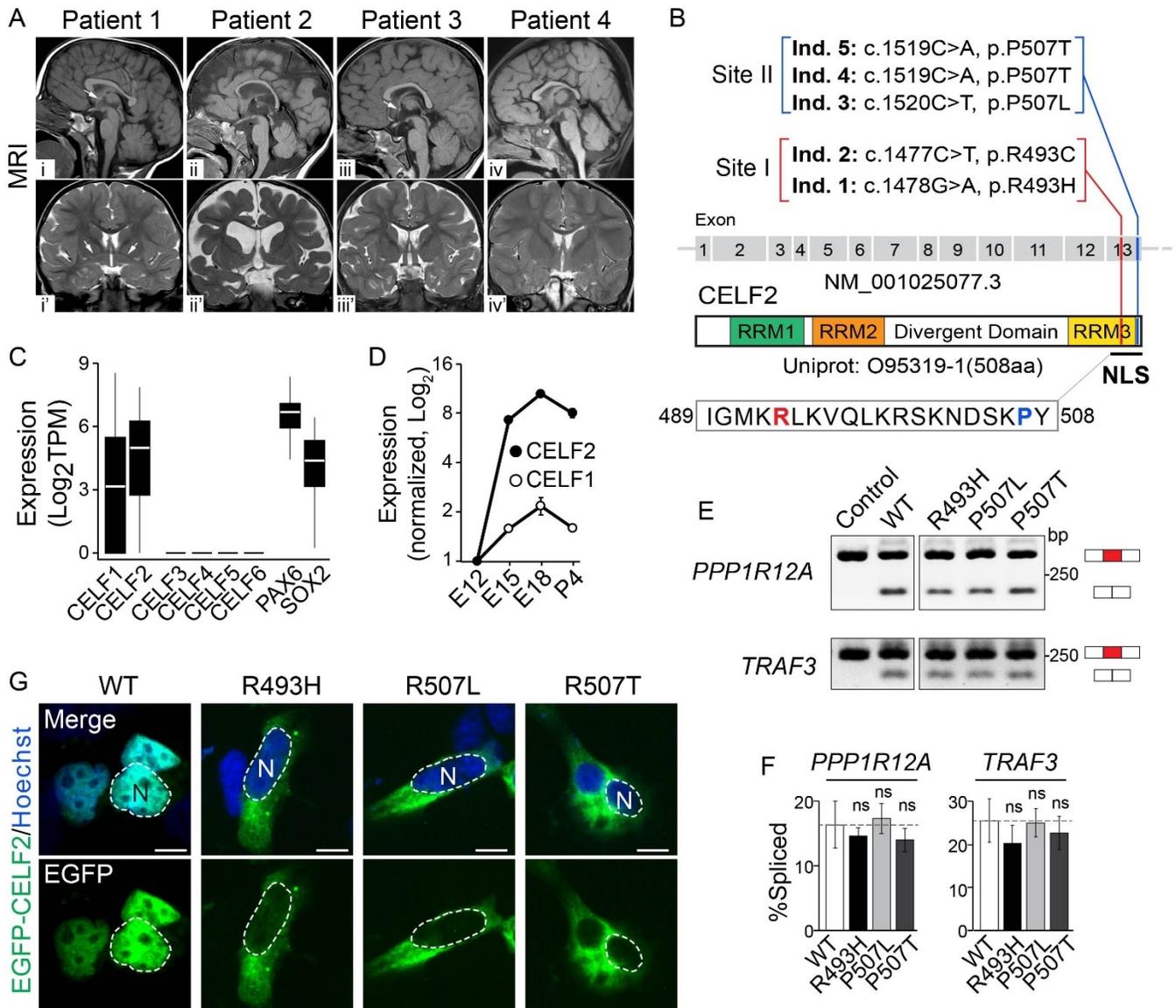


Figure 1. CELF2 pathogenic variants disrupt nucleocytoplasmic localization and cortical development. (A) Selected MR images. Sagittal T1-weighted (upper row) and coronal T2-weighted (lower row) images of patient 1 (i, i'), patient 2 (ii, ii'), patient 3 (iii, iii') and patient 4 (iv, iv') are shown. The images were obtained at age of 25 months for patient 1, 16 months for patient 2, 25 months for patient 3, and 41 months for patient 4. All patients have simplified gyral pattern, mild in patients 1, 3, and 4, and moderate in patient 2. The corpus callosum is normal in patient 1, normally developed but diffusely thinned in patient 2, normal in patient 3, and mildly hypoplastic in patient 4. Patients 1 and 3

have normally developed anterior commissure (arrows in i, ii', iii, iii'), but patients 2 and 4 have absence of anterior commissure. Patient 2 had moderate ventriculomegaly. **(B)** The schematic showing the location of the identified variants in the coding exons and the corresponding parts of the CELF2 protein with three RNA-recognition motifs (RRM1-3) (Ladd and Cooper, 2004). The affected residues are within a known nuclear-localization signal (NLS) in RRM3. **(C)** Box plots showing *CELFs* expression in human cortical NPCs (PAX6+, SOX2+) (single-cell RNA-seq dataset from Pollen et al., 2015). **(D)** Temporal qRT-PCR analysis of the expression of *CELF1* and *CELF2* in the developing mouse cortices at indicated timepoints. n = 3 samples each timepoints. **(E)** RT-PCR analysis of exon skipping events for PPP1R and TRAF3 in HEK293 cells transfected with empty vector (Control), wild-type (WT) or mutated CELF2. **(F)** Quantifications of the percentage of transcripts with the skipped exon. **(G)** Confocal images of HEK293 cells expressing EGFP-CELF2 (green) that carry different pathogenic variants and counterstained for Hoechst 33258 (blue). "N" denotes the nucleus. Dashed white lines denote nuclei (blue, N). Scale bar = 10 μ m. Data are presented as mean values +/- SEM.

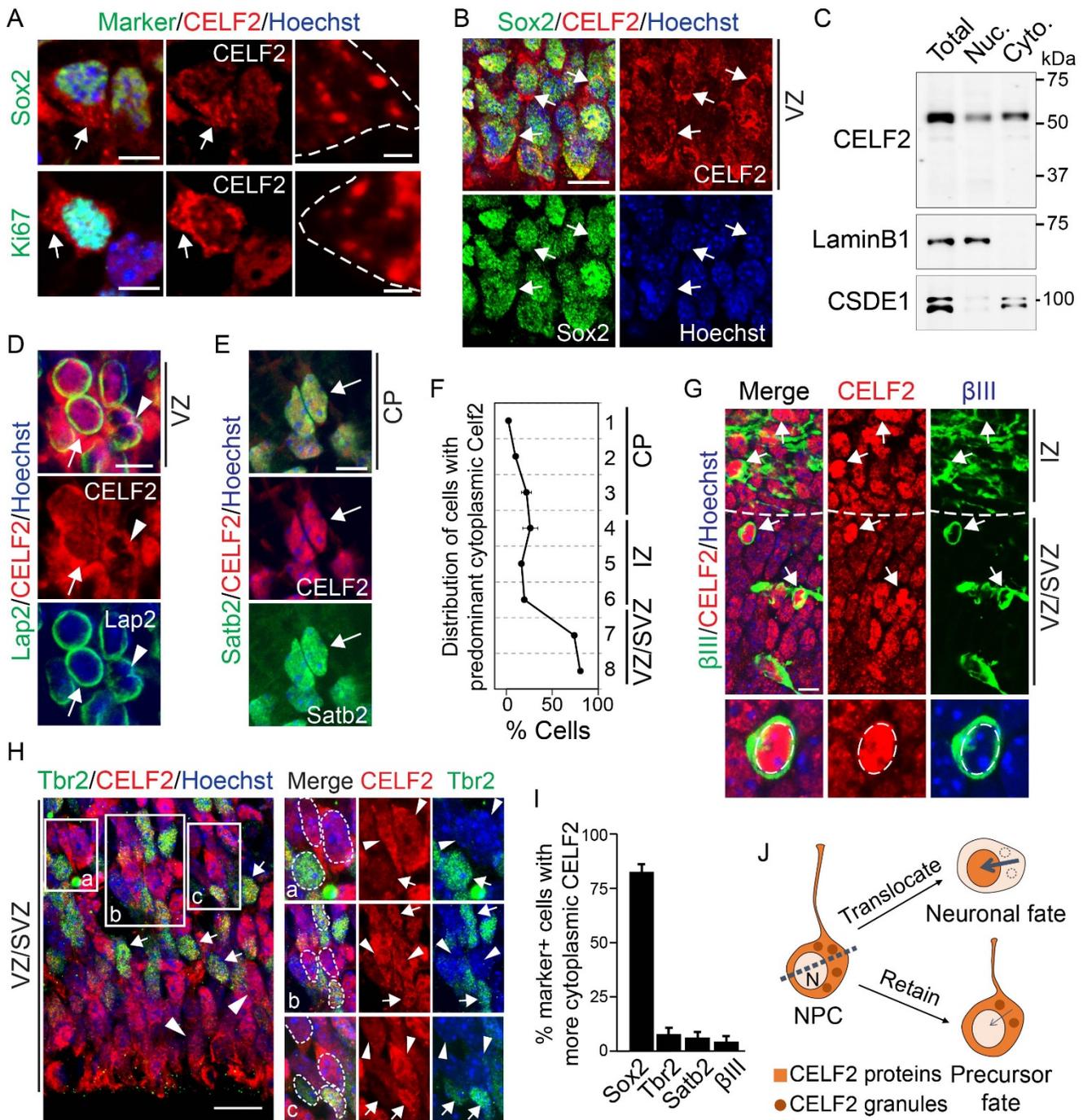


Figure 2. CELF2 nuclear translocation correlates with NPC differentiation. (A) Confocal images of cultured NPCs isolated from the E12.5 mouse cortex, immunostained for CELF2 (red) and Sox2+ or Ki67+ (both green). Arrow pointed regions are shown in higher magnification to the right with cell boundaries marked by dashed white lines. (B) Confocal images of the ventricular zone (VZ) of E12.5

cortical sections, immunostained for CELF2 (red) and Sox2 (green). Arrows denote NPCs with cytoplasmic CELF2. **(C)** Western blots of the lysate (Total), nuclear fraction (Nuc.) and cytoplasmic fraction (Cyto.) of the E12.5 cortex, probed for CELF2, LaminB1 and CSDE1. **(D)** Confocal images of E12.5 cortical sections, immunostained for CELF2 (red) and the nuclear envelop marker Lap2 (green). Arrows and arrowheads denote closed and open nuclear envelopes, respectively. **(E-I)** Analysis of E15.5 cortical sections, immunostained for CELF2 (red) and cell type markers (green) as indicated. **(E)** Confocal images of the cortical plate (CP) for the neuron-specific transcription factor Satb2. Arrows indicate cells with nuclear Satb2 and CELF2. **(F)** Quantifications of cells with predominant cytoplasmic CELF2 in eight equal-sized bins spanning the cortex as in **Figure S2F**. **(G)** Confocal images of the VZ and subventricular zone (SVZ) and intermediate zone (IZ) for β III-tubulin. Dashed white lines denote the boundary between the SVZ and IZ. An example of newborn neurons is shown in higher magnification to the bottom. **(H)** Images of the VZ/SVZ for the intermediate progenitor marker Tbr2. Regions with white boxes are shown in higher magnification to the right. Arrows and arrowheads denote Tbr2-positive and negative cells, respectively. Dashed white circles denote nuclei. **(I)** Quantifications of marker-positive cells with predominant cytoplasmic CELF2. **(J)** A schematic showing that CELF2 translocation into the nucleus marks NPCs committed to differentiation. Nuclei are stained by Hoechst 33258 dye (blue) in A, B, D, E, G, H. Scale bars = 10 μ m. Data are presented as mean values +/- SEM.

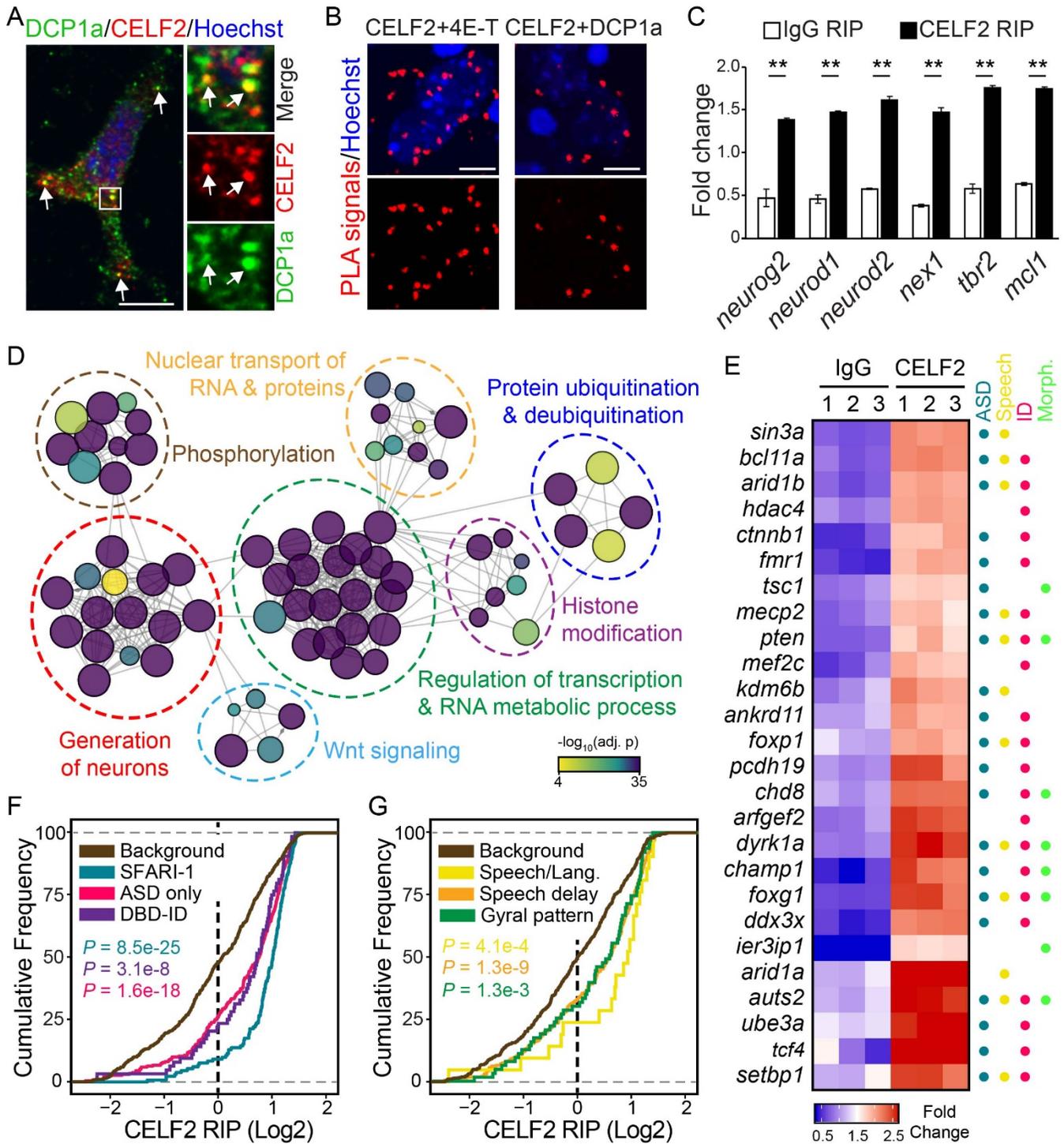


Figure 3. Cytoplasmic CELF2 orchestrates mRNAs encoding stem cell regulators and disease-

related factors. (A) Confocal images of E12.5 cultured cortical cells immunostained for CELF2 (red) and DCP1a (green). The white box is shown in higher magnification to the right. Arrows denote the

colocalization foci. **(B)** Confocal images of E12.5 cultured cortical cells showing PLA signals (red) for the interactions of CELF2 with 4E-T (left) or DCP1a (right). **(C)** RIP-seq analysis of proneurogenic mRNAs in the E12.5 cortex. n = 3 experiments. Data are presented as mean values +/- SEM. ** p<0.01. Two-sided unpaired student t-test. **(D)** Gene ontology analysis for CELF2 target mRNAs. The individual node represents different enriched pathways that are clustered based on the similarity of their functions. The sizes of the nodes are proportional to the number of genes included. Enrichment p-values are color-coded. **(E)** A heat map of select CELF2 targets from enriched pathways in **(D)**, showing their fold change (RIP/Input) in CELF2 and the IgG control RIPs. Disease conditions associated with each target are marked with colored dots. **(F, G)** Cumulative distribution of CELF2 fold change (RIP/Input, log₂ transformed) for 500 randomly chosen mRNAs from the transcriptome (dark brown) and mRNAs encoded by SFARI ASD genes and Geisinger database (DBD) ID genes and ASD-only genes **(F)** and genes related to speech and language and abnormal gyral pattern **(G)**. P-values were calculated using the two-sided Mann-Whitney U test. Nuclei are stained by Hoechst 33258 dye (blue) in A, B. Scale bars = 10 μm.

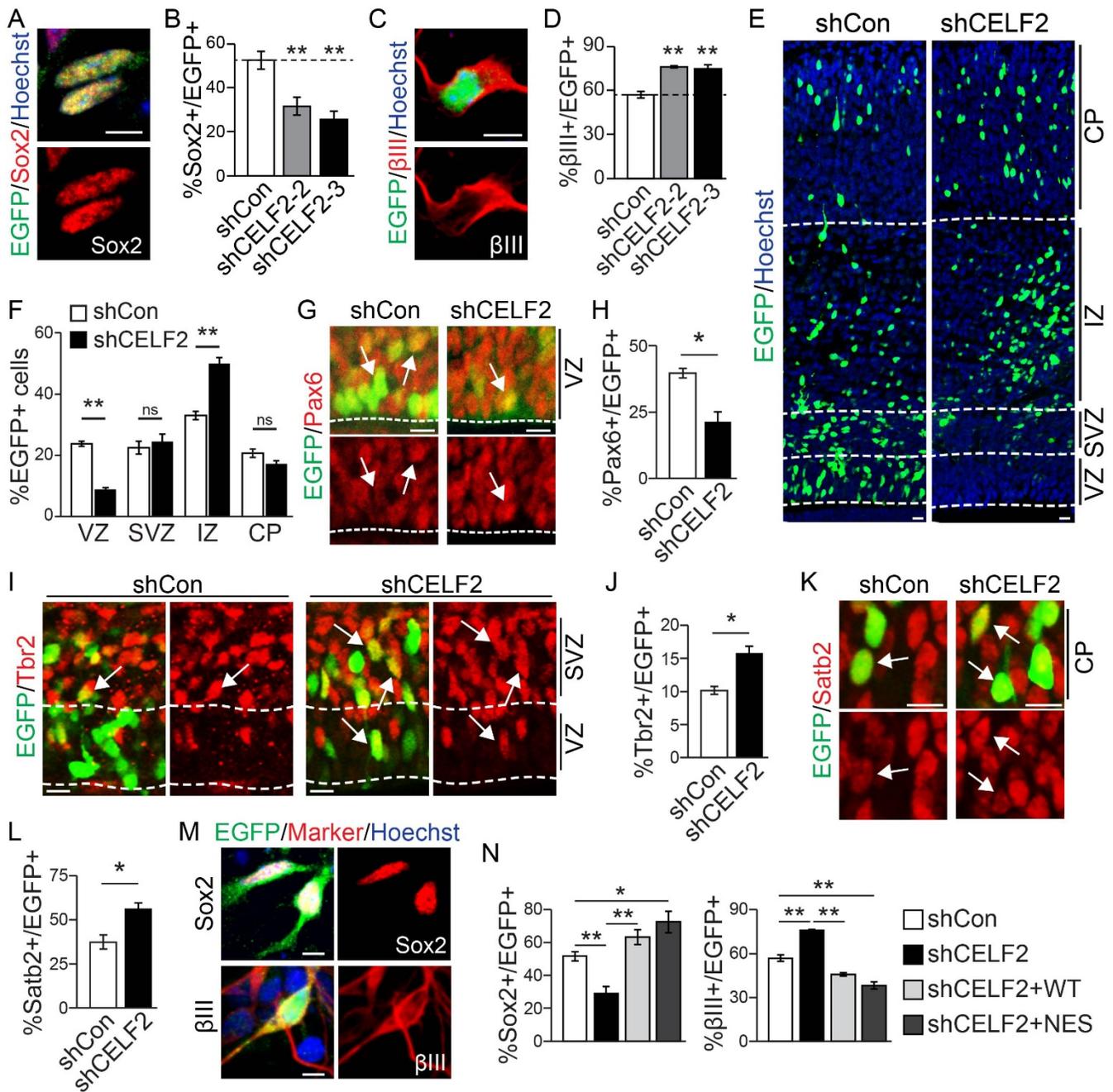


Figure 4. Cytoplasmic CELF2 regulates NPC fate decision. (A-D) Analysis of E12.5 cortical cultures transfected with control or CELF2 shRNAs and EGFP (green). (A) Confocal images and (B) quantifications of % EGFP+ cells that are also positive for Sox2 (red). n = 3 experiments. (C) Confocal images and (D) quantifications of % EGFP+ cells that are also positive for β III-tubulin (red). n = 3

experiments. **(E-L)** Analysis of E16.5 cortical sections *in utero* electroporated with control or CELF2 shRNAs and EGFP at E13.5 and immunostained for EGFP (green) and indicated markers (red). **(E)** Confocal images of the cortex. Dashed white lines denote boundaries of the VZ, SVZ, IZ and CP. **(F)** Distribution of EGFP+ cells across the cortex. n = 3 embryos each. **(G)** Images of the VZ for Pax6. Arrows denote double-positive cells. Dashed white lines denote the apical surfaces. **(H)** Quantifications of % of EGFP+ cells that are also positive for Pax6. n = 3 embryos each. **(I)** Images of the VZ and SVZ for Tbr2. Arrows denote double-positive cells. Dashed white lines denote the apical surfaces and the boundary between the VZ and SVZ. **(J)** Quantifications of % of EGFP+ cells that are also positive for Tbr2. n = 3 embryos each. **(K)** Images of the CP for Satb2. Arrows denote double-positive cells. **(L)** Quantifications of % of EGFP+ cells that are also positive for Satb2. n = 3 embryos each. **(M, N)** Analysis of E12.5 cortical cultures transfected with EGFP (green) and indicated plasmids for 3 days. **(M)** Images and **(N)** quantifications of % EGFP+ cells that are also positive for Sox2 or β III-tubulin (red). n = 3 experiments. Nuclei are stained by Hoechst 33258 dye (blue) in A, C, E, M. Scale bars = 10 μ m. Data are presented as mean values +/- SEM. * p<0.05, ** p<0.01, ns = p >0.05.

Figures

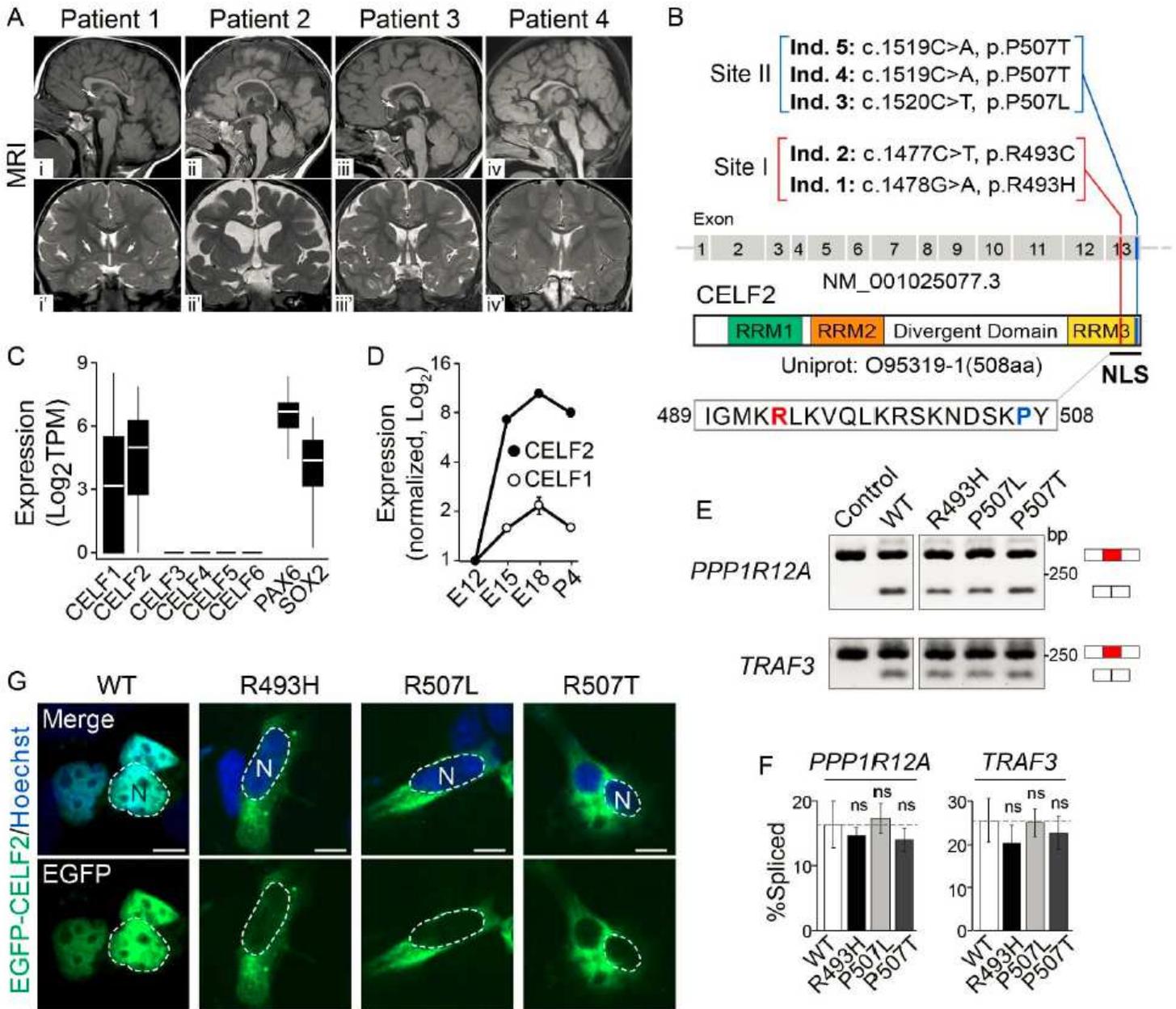


Figure 1

CELF2 pathogenic variants disrupt nucleocytoplasmic localization and cortical development. (A) Selected MR images. Sagittal T1-weighted (upper row) and coronal T2-weighted (lower row) images of patient 1 (i, i'), patient 2 (ii, ii'), patient 3 (iii, iii') and patient 4 (iv, iv') are shown. The images were obtained at age of 25 months for patient 1, 16 months for patient 2, 25 months for patient 3, and 41 months for patient 4. All patients have simplified gyral pattern, mild in patients 1, 3, and 4, and moderate in patient 2. The corpus callosum is normal in patient 1, normally developed but diffusely thinned in patient 2, normal in patient 3, and mildly hypoplastic in patient 4. Patients 1 and 3 have normally developed anterior commissure

(arrows in i, ii', iii, iii'), but patients 2 and 4 have absence of anterior commissure. Patient 2 had moderate ventriculomegaly. (B) The schematic showing the location of the identified variants in the coding exons and the corresponding parts of the CELF2 protein with three RNA-recognition motifs (RRM1-3) (Ladd and Cooper, 2004). The affected residues are within a known nuclear-localization signal (NLS) in RRM3. (C) Box plots showing CELFs expression in human cortical NPCs (PAX6+, SOX2+) (single-cell RNA-seq dataset from Pollen et al., 2015). (D) Temporal qRT-PCR analysis of the expression of CELF1 and CELF2 in the developing mouse cortices at indicated timepoints. n = 3 samples each timepoints. (E) RT-PCR analysis of exon skipping events for PPP1R and TRAF3 in HEK293 cells transfected with empty vector (Control), wild-type (WT) or mutated CELF2. (F) Quantifications of the percentage of transcripts with the skipped exon. (G) Confocal images of HEK293 cells expressing EGFP-CELF2 (green) that carry different pathogenic variants and counterstained for Hoechst 33258 (blue). "N" denotes the nucleus. Dashed white lines denote nuclei (blue, N). Scale bar = 10 μ m. Data are presented as mean values +/- SEM.

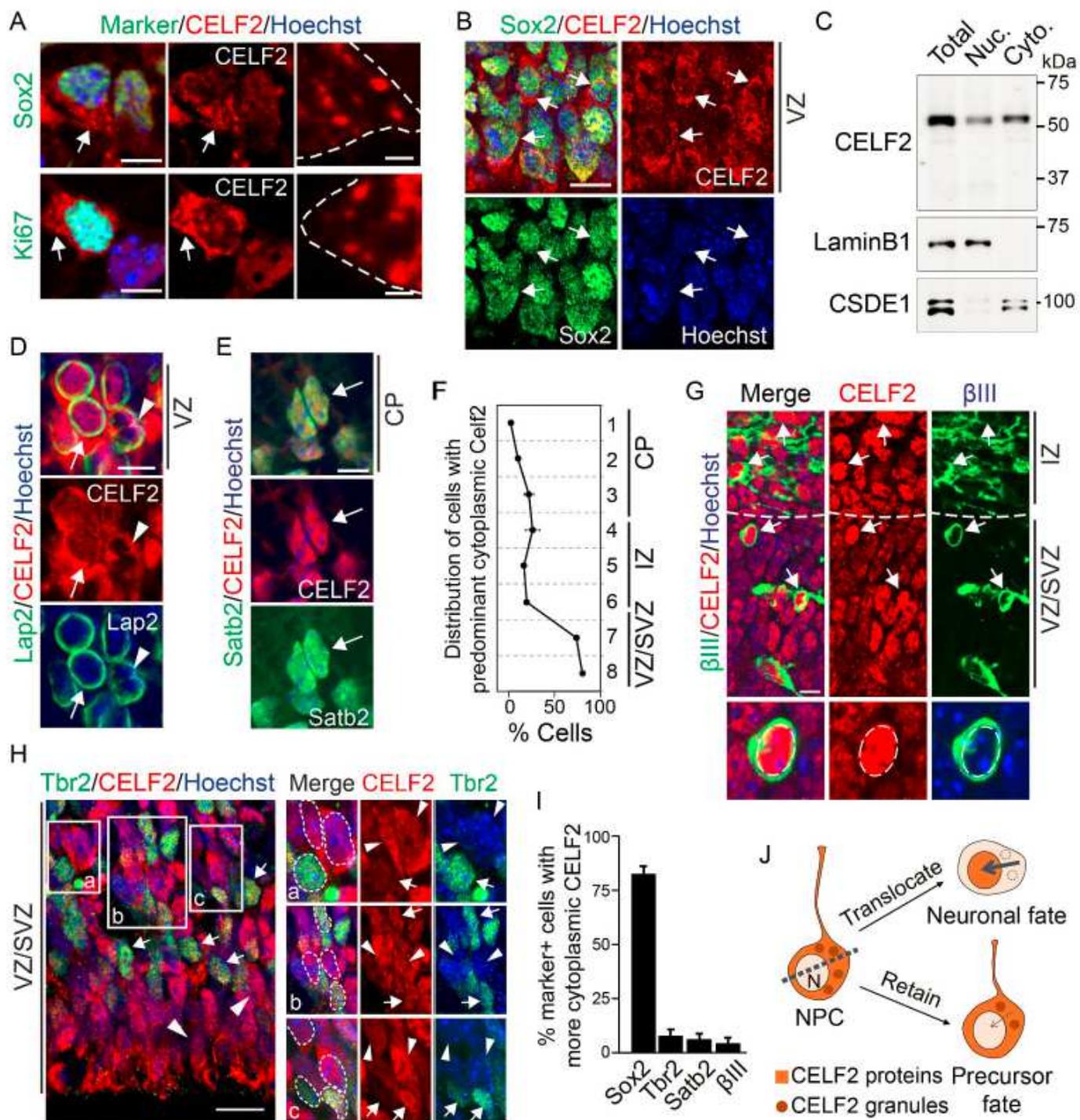


Figure 2

CELF2 nuclear translocation correlates with NPC differentiation. (A) Confocal images of cultured NPCs isolated from the E12.5 mouse cortex, immunostained for CELF2 (red) and Sox2+ or Ki67+ (both green). Arrow pointed regions are shown in higher magnification to the right with cell boundaries marked by dashed white lines. (B) Confocal images of the ventricular zone (VZ) of E12.5 cortical sections, immunostained for CELF2 (red) and Sox2 (green). Arrows denote NPCs with cytoplasmic CELF2. (C)

Western blots of the lysate (Total), nuclear fraction (Nuc.) and cytoplasmic fraction (Cyto.) of the E12.5 cortex, probed for CELF2, LaminB1 and CSDE1. (D) Confocal images of E12.5 cortical sections, immunostained for CELF2 (red) and the nuclear envelop marker Lap2 (green). Arrows and arrowheads denote closed and open nuclear envelopes, respectively. (E-I) Analysis of E15.5 cortical sections, immunostained for CELF2 (red) and cell type markers (green) as indicated. (E) Confocal images of the cortical plate (CP) for the neuron-specific transcription factor Satb2. Arrows indicate cells with nuclear Satb2 and CELF2. (F) Quantifications of cells with predominant cytoplasmic CELF2 in eight equal-sized bins spanning the cortex as in Figure S2F. (G) Confocal images of the VZ and subventricular zone (SVZ) and intermediate zone (IZ) for β III-tubulin. Dashed white lines denote the boundary between the SVZ and IZ. An example of newborn neurons is shown in higher magnification to the bottom. (H) Images of the VZ/SVZ for the intermediate progenitor marker Tbr2. Regions with white boxes are shown in higher magnification to the right. Arrows and arrowheads denote Tbr2-positive and negative cells, respectively. Dashed white circles denote nuclei. (I) Quantifications of marker-positive cells with predominant cytoplasmic CELF2. (J) A schematic showing that CELF2 translocation into the nucleus marks NPCs committed to differentiation. Nuclei are stained by Hoechst 33258 dye (blue) in A, B, D, E, G, H. Scale bars = 10 μ m. Data are presented as mean values +/- SEM.

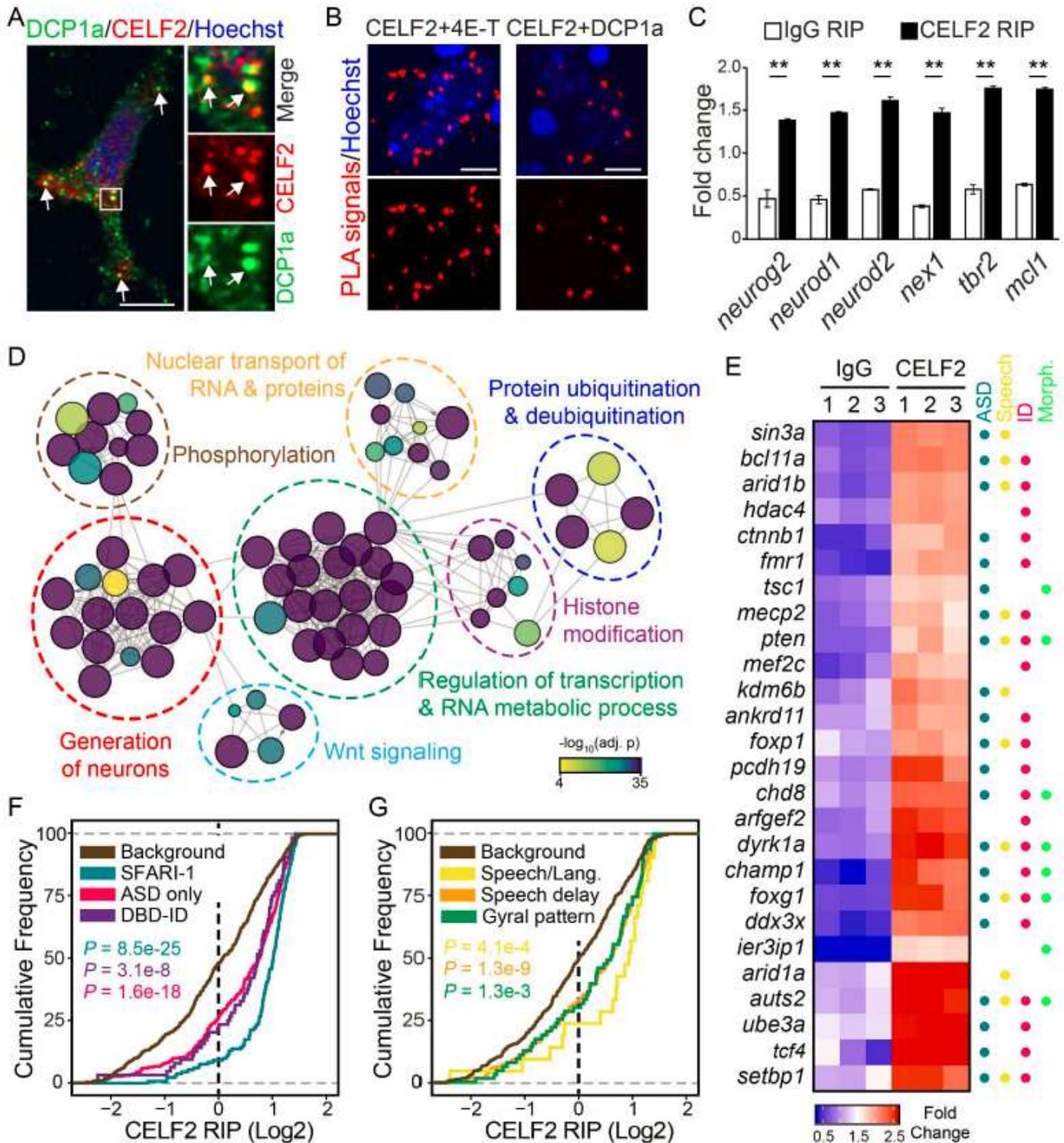


Figure 3

Cytoplasmic CELF2 orchestrates mRNAs encoding stem cell regulators and disease-related factors. (A) Confocal images of E12.5 cultured cortical cells immunostained for CELF2 (red) and DCP1a (green). The white box is shown in higher magnification to the right. Arrows denote the colocalization foci. (B) Confocal images of E12.5 cultured cortical cells showing PLA signals (red) for the interactions of CELF2 with 4E-T (left) or DCP1a (right). (C) RIP-seq analysis of proneurogenic mRNAs in the E12.5 cortex. $n = 3$

experiments. Data are presented as mean values \pm SEM. ** $p < 0.01$. Two-sided unpaired student t-test. (D) Gene ontology analysis for CELF2 target mRNAs. The individual node represents different enriched pathways that are clustered based on the similarity of their functions. The sizes of the nodes are proportional to the number of genes included. Enrichment p-values are color-coded. (E) A heat map of select CELF2 targets from enriched pathways in (D), showing their fold change (RIP/Input) in CELF2 and the IgG control RIPs. Disease conditions associated with each target are marked with colored dots. (F, G) Cumulative distribution of CELF2 fold change (RIP/Input, log₂ transformed) for 500 randomly chosen mRNAs from the transcriptome (dark brown) and mRNAs encoded by SFARI ASD genes and Geisinger database (DBD) ID genes and ASD-only genes (F) and genes related to speech and language and abnormal gyral pattern (G). P-values were calculated using the two-sided Mann-Whitney U test. Nuclei are stained by Hoechst 33258 dye (blue) in A, B. Scale bars = 10 μ m.

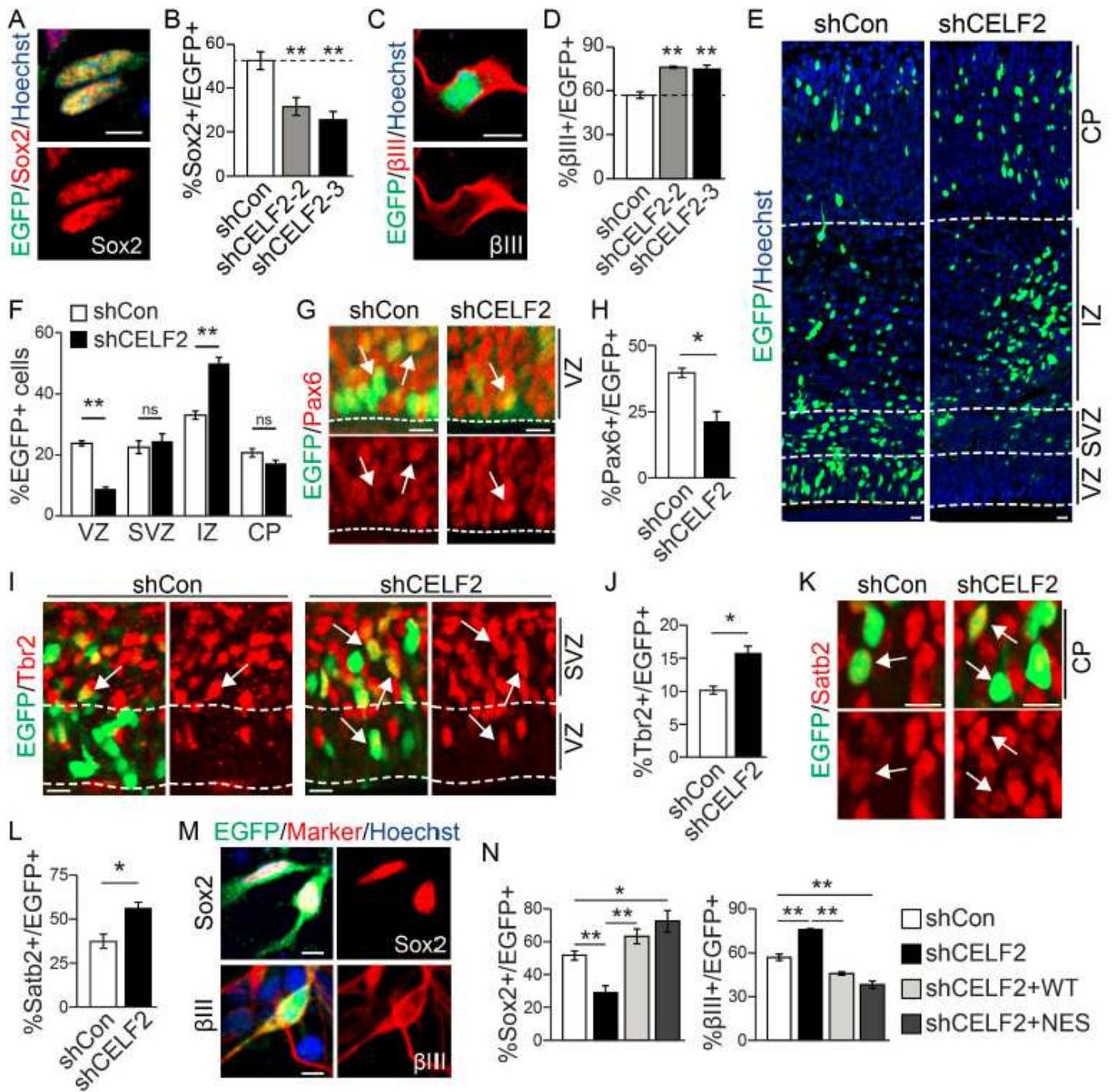


Figure 4

Cytoplasmic CELF2 regulates NPC fate decision. (A-D) Analysis of E12.5 cortical cultures transfected with control or CELF2 shRNAs and EGFP (green). (A) Confocal images and (B) quantifications of % EGFP+ cells that are also positive for Sox2 (red). n = 3 experiments. (C) Confocal images and (D) quantifications of % EGFP+ cells that are also positive for βIII-tubulin (red). n = 3 experiments. (E-L) Analysis of E16.5 cortical sections in utero electroporated with control or CELF2 shRNAs and EGFP at E13.5 and immunostained for EGFP (green) and indicated markers (red). (E) Confocal images of the cortex. Dashed

white lines denote boundaries of the VZ, SVZ, IZ and CP. (F) Distribution of EGFP+ cells across the cortex. n = 3 embryos each. (G) Images of the VZ for Pax6. Arrows denote double-positive cells. Dashed white lines denote the apical surfaces. (H) Quantifications of % of EGFP+ cells that are also positive for Pax6. n = 3 embryos each. (I) Images of the VZ and SVZ for Tbr2. Arrows denote double-positive cells. Dashed white lines denote the apical surfaces and the boundary between the VZ and SVZ. (J) Quantifications of % of EGFP+ cells that are also positive for Tbr2. n = 3 embryos each. (K) Images of the CP for Satb2. Arrows denote double-positive cells. (L) Quantifications of % of EGFP+ cells that are also positive for Satb2. n = 3 embryos each. (M, N) Analysis of E12.5 cortical cultures transfected with EGFP (green) and indicated plasmids for 3 days. (M) Images and (N) quantifications of % EGFP+ cells that are also positive for Sox2 or β III-tubulin (red). n = 3 experiments. Nuclei are stained by Hoechst 33258 dye (blue) in A, C, E, M. Scale bars = 10 μ m. Data are presented as mean values \pm SEM. * $p < 0.05$, ** $p < 0.01$, ns = $p > 0.05$.