Article

CRISPR/Cas9-mediated gene-knockout of *UPF3B* alters expression of cell cycle and neuron-specific genes

Pratap Chandra [1], Bhagyashree Deka [1], Sweta Kumari [1], Ayushi Rehman [1], Priyanka Yadav [1] and Kusum K. Singh [1,*]

[1] Department of Biosciences and Bioengineering, Indian Institute of Technology Guwahati, Guwaahti – 781039, Assam, India; singhpc90@iitg.ac.in (P.C.); bhagyashree@iitg.ac.in (B.D.); sweta176106105@iitg.ac.in (S.K.); ayushirehman@iitg.ac.in (A.R.); priyankayadav@iitg.ac.in (P.Y.)

* Correspondence: kusumsingh@iitg.ac.in (K.K.S.); Tel.: +91-(0)-361-2583206; Fax: +91-(0)-361-2582249

Abstract: UPF3B is a constituent of the classical nonsense-mediated mRNA decay (NMD) pathway that degrades both the aberrant transcripts and a set of physiological transcripts. In higher eukaryotes, UPF3B have significant biochemical functions in diverse cellular processes including NMD and translation. UPF3B plays a crucial role in neuronal development and differentiation. Next-generation sequencing technologies identified several loss-of-function mutations in the *UPF3B* gene that results in neuro-developmental disorders in humans. To uncover the mechanistic role of UPF3B in neuronal functions, we have generated the UPF3B-knockout mammalian cell line model system using CRISPR-Cas9 gene editing method. RNA-Sequencing Analysis of cellular transcriptome from *UPF3B*-KO cells identified specific genes involved in cell growth and neuronal functions. Altered expression of genes related to the axon guidance pathway delineated the UPF3B function to regulate the neuron-specific genes. Functional enrichment analysis identified the genes involved in the disorders related to mental health and intellectual disability. Our study has the potential to identify the direct players of intellectual disability and will have broader implications.

Keywords: Nonsense-mediated mRNA Decay, *UPF3B*-knockout, RNA-Sequencing, Intellectual disability, Neuro-developmental disorders.

1. Introduction

Nonsense-mediated mRNA decay (NMD), a conserved surveillance pathway in all eukaryotes, eradicates the aberrant transcripts and the accumulation of deleterious proteins in the cell [1]. The regulation of discrete sets of transcripts depends upon the assembly of different NMD factors; thus classifying NMD as a branched pathway [2, 3]. Up-Frameshift Suppressor 3 Homolog B (UPF3B) is a protein component of the classical NMD pathway that degrades aberrant and normal transcripts to maintain cellular homeostasis. However, the existence of UPF3B-independent alternative NMD pathways indicates that UPF3B-mediated NMD activity targets a specific subset of transcripts. Alternatives of the classical NMD pathway suggest that the UPF3B function is not limited to the mRNA turnover.

Several studies showed that UPF3B acts as a multifunctional protein, as described in a recent review [4]. UPF3B is involved in various protein-protein interactions (PPIs), which results in known biological functions. First, UPF3B is recruited to exon junction complex (EJC), an indicator for splicing, which is deposited over newly spliced mRNA [5]. Simultaneous interaction of UPF3B with nuclear export factors indicate that UPF3B-bound transcripts are assured to be exported to the cytoplasm. In the cytoplasm, UPF3B bridges EJC to NMD factor UPF2 [6], which further interacts with UPF1 to complete the NMD assembly [7]. Despite being part of the NMD complex, UPF3B is also involved in translation augmentation [8] and translation termination [9]. The UPF3B-eRF3a (eukaryotic release factor 3a) interaction is likely to delay the translation termination at premature

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

termination codon (PTC)-bearing transcripts by hampering the recognition of PTC via eRFs [9]. The role of UPF3B-eRF3a interactions in translation stimulation is yet to be explored. Furthermore, UPF3B expression is crucial for neuronal development and mature neuronal processes [10-12]. However, the underlying molecular mechanisms are poorly understood.

Over the last decade, growing evidence suggests that the *UPF3B* gene mutations are associated with various neuro-developmental disorders (NDDs). Initially, Tarpey et al. identified that protein truncation and loss of function of UPF3B due to mutations in the *UPF3B* gene led to the development of syndromic and non-syndromic X-linked intellectual disability (XLID) phenotype in patients [13]. Later studies have shown that UPF3B mutations are associated with non-syndromic XLID, with or without autism [11], childhood-onset schizophrenia (COS), and autism spectrum disorder (ASD) [14], schizophrenia [15], renal dysplasia, and variable developmental delays [16]. RNA-Seq analysis of the cellular transcriptome from UPF3B-deficient lymphoblastoid cells (LCLs) in patients with ID identified differentially expressed genes (DEGs) involved in neuronal processes, including axon guidance and axon growth [17]. Subsequently, Jolly et al. illustrated the upregulation of axon guidance genes in Upf3b-depleted mouse hippocampal neurons [10]. These studies identified a contrasting set of deregulated genes and suggested UPF3B-mediated regulation of a set of transcripts involved in neuronal development and function. Consequently, a suitable mammalian model system was required for an in-depth analysis of the mechanistic role of UPF3B in biological processes.

In a significant breakthrough in understanding the molecular mechanism of UPF3B-associated ID, Huang et al. developed a Upf3b-null mouse model showing behavioral and neurogenesis defects [18]. RNA-Seq analysis of the frontal cortex transcriptome from Upf3b-null mice identified several transcripts associated with neuronal processes, including dendritic arborization, promotion of axon growth, synaptic connections, and cellular excitability. However, the mouse model cannot represent human patient conditions. Researchers have developed *UPF3B*-knockout (*UPF3B*-KO) cell line models to study the UPF3B-mediated biological processes limited to the NMD pathway [19, 20]. In this study, we generated a *UPF3B*-KO HEK-293 cell line using CRISPR-Cas9 gene editing. We performed RNA-Seq analysis to identify DEGs involved in various biological processes. We also identified the distinct genes involved in the cell cycle and neuronal processes via the axon guidance pathway.

2. Results

2.1. Generation of UPF3B-knockout mammalian cell line

Previous studies have shown that *UPF3B* gene mutations result in NDDs with compromised mental stature and an underdeveloped central nervous system (CNS) in patients [11, 13-16]. Furthermore, neurogenesis defects were observed in a mouse model [18]. Mechanistically, UPF3B links EJC to the NMD complex and the eRF3a-GTP/eRF1 complex (Figure 1a), thereby determining the fate of transcripts via these interactions. To understand the mechanistic role of UPF3B in biological processes, we generated a mammalian cell line model system to mimic the patient conditions. We used CRISPR-Cas9 gene-editing method to perform *UPF3B*-KO in the HEK-293 cell line. First, we generated HEK-293 cells which stably expressed Cas9 protein along with GFP (Figure 1b). Furthermore, with the targeted delivery of a single guide-RNA corresponding to exon 1 of UPF3B (Figure S1a-b), we identified five potential *UPF3B*-KO clones. Western blot analysis of these clones showed no expression of UPF3B in clone 1; however, there was minimal expression of UPF3B in clone 2 and clone 3 (Figure 1c). Clone 4 and clone 5 showed no altered expression of UPF3B. Sequencing analysis of the *UPF3B* gene showed a 55 bp deletion in exon 1 downstream of the guide-RNA binding region (Figure 1d, Figure S1c). Using RNA-Seq, we also found evidence for the presence of the 55 bp deletion, thus validating the sequencing analysis (Supplementary Figure 1d). Furthermore, quantitative real-time polymerase chain reaction (qRT-PCR) analysis

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

showed the depletion of *UPF3B* mRNA levels in knockout cells (Figure 1e). These results indicate the successful generation of the UPF3B-deficient HEK-293 cell line, which we used as a model system for our study.

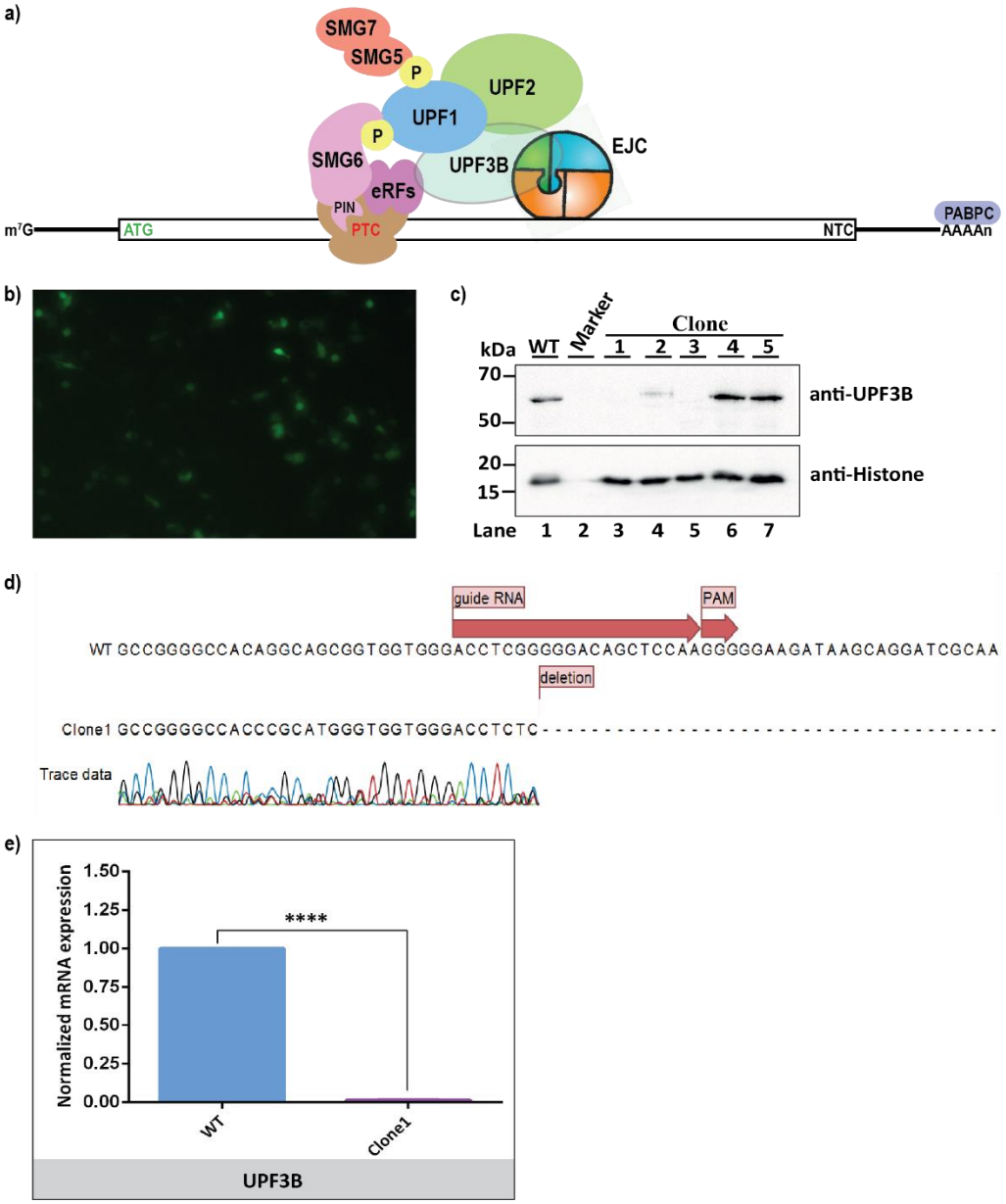

Figure 1. Generation of *UPF3B*-KO mammalian cell line model. a) Schematic representation of UPF3B-interacting protein complexes on spliced mRNA. b) Expression of GFP in Cas9-integrated HEK-293 cells after cumate induction. c) Western blot analysis of candidate clones for *UPF3B*-KO confirmation w.r.t. wild-type cells. d) Sanger sequencing of UPF3B exon1 sequences. e) qRT-PCR analysis of *UPF3B* mRNA expression.

2.2. UPF3B depletion and its affect NMD activity and mRNA levels of other NMD factors

Since UPF3B is a component of the classical NMD pathway, we were fascinated to discern the impact of UPF3B depletion on the NMD pathway. To check the alteration in NMD activity in *UPF3B*-KO cells, we quantified the mRNA expression levels of previously reported endogenous NMD substrates, *SRSF2, ZFAS1, GAS5*, and *ITGAE* [21]. SRSF2 is a pre-mRNA splicing factor, a member of serine/arginine (SR)-rich protein family, which auto-regulates its mRNA expression by generating the NMD-sensitive splice isoforms [22]. ZFAS1 represents a snoRNA host gene and encodes an mRNA with a short PTC-bearing ORF that undergoes degradation via NMD [23]. GAS5 and ITGAE transcripts possess intron retention that makes them NMD-sensitive. We observed no

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

significant fold change in mRNA expression of *SRSF2*, *ZFAS1*, *GAS5*, and *ITGAE* transcripts in *UPF3B*-KO cells compared to the wild-type (Figure 2a). Next, we examined whether the expression of other NMD factors is altered upon the loss of UPF3B, and whether alternative NMD pathways take over the NMD activity of UPF3B. To assess this possibility, we quantified the mRNA expression of the NMD factors UPF1, UPF2, and UPF3A (Figure 2b). We observed a significant decrease in mRNA levels of *UPF1* and *UPF2*, indicating the NMD activity mediated by UPF1 and UPF2 is probably altered. In accordance with previous studies, we also found that the mRNA level of *UPF3A* did not changed upon *UPF3B*-KO [24]. However, western blot analysis showed a significant upregulation in UPF3A protein expression (~1.9 fold) in the absence of UPF3B (Figure 2c).

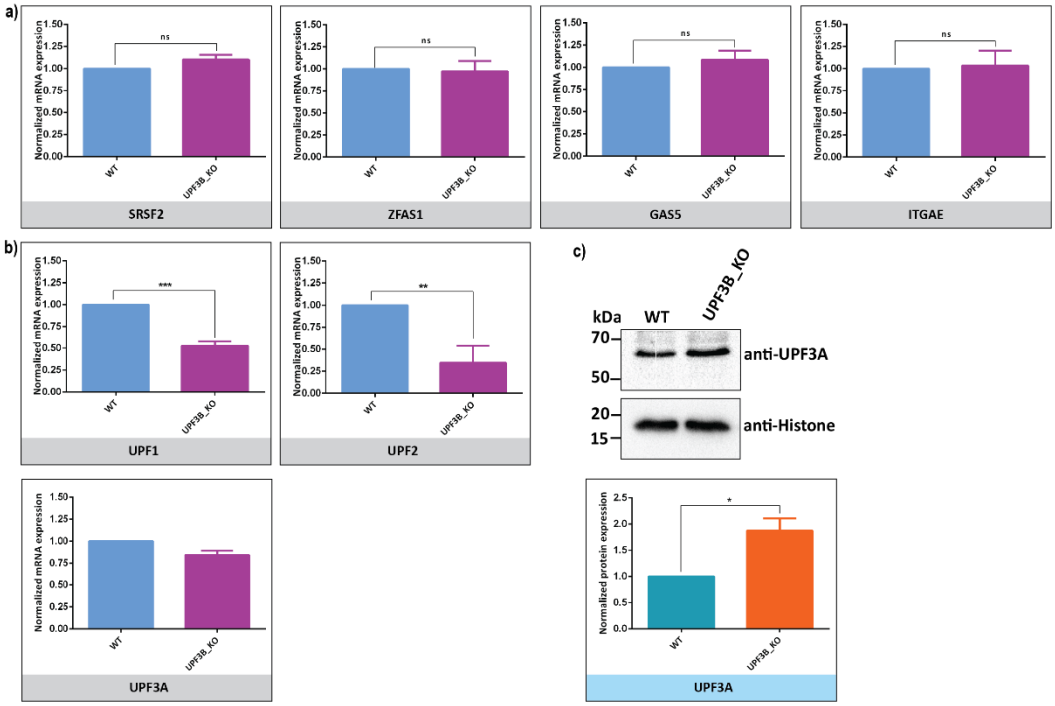

Figure 2. Impact of UPF3B-depletion on NMD activity. a) qRT-PCR analysis of NMD-sensitive transcripts, *SRSF2*, *ZFAS1*, *GAS5*, and *ITGAE*. b) qRT-PCR of other UPF genes, *UPF1*, *UPF2*, and *UPF3A*. c) Western blot analysis showing upregulated UPF3A protein expression (upper panel). Quantification of the normalized protein expression of UPF3A (lower panel).

Altogether, we found that *UPF3B*-KO does not alter the NMD activity, indicating NMD is a branched pathway. UPF3A is probably largely responsible for compensating UPF3B-mediated NMD activity in the complete absence of UPF3B.

2.3. Genome-wide identification of differentially regulated genes in UPF3B-KO cells

To gain insight into the global impact of *UPF3B*-KO on the cellular transcriptome, we performed RNA-Seq analysis on *UPF3B*-KO cell lines and parental wild-type cell lines (Figure 3). Principal component analysis and the Poisson distance among the samples demonstrated that the transcriptome signature of *UPF3B*-KO and wild-type cells clustered separately, whereas the sample replicates clustered together (Figure 3a-b). Furthermore, a heatmap of hierarchical clustering revealed the transcriptomic differences between *UPF3B*-KO and wild-type cells (Figure 3c). Differential gene expression analysis using DESeq2 revealed a total of 107 upregulated genes (6.5%) in *UPF3B*-KO cells compared to the wild-type (Adjusted p-value < 0.05; log2FC > 1.5), while, in contrast, 1526 genes (93.5%) were downregulated (Adjusted p-value < 0.05; log2FC < -1.5) (Figure 3d; Supplementary File S1). Among the significantly downregulated genes, we identified two genes, Midline 1 (*MID1*) and Anosmin 1 (*ANOS1*), which are directly associated with X-linked disorders displaying the DEGs. These genes might play a significant role in biological processes, including neuronal functions. Among 107 upregulated DEGs, PPI network analysis using the STRING database identified 94 annotated protein-coding genes (Figure S2a), where 14

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

genes were grouped to form a cluster (Figure 3e). Functional enrichment analysis showed that most of the genes were involved in cell differentiation.

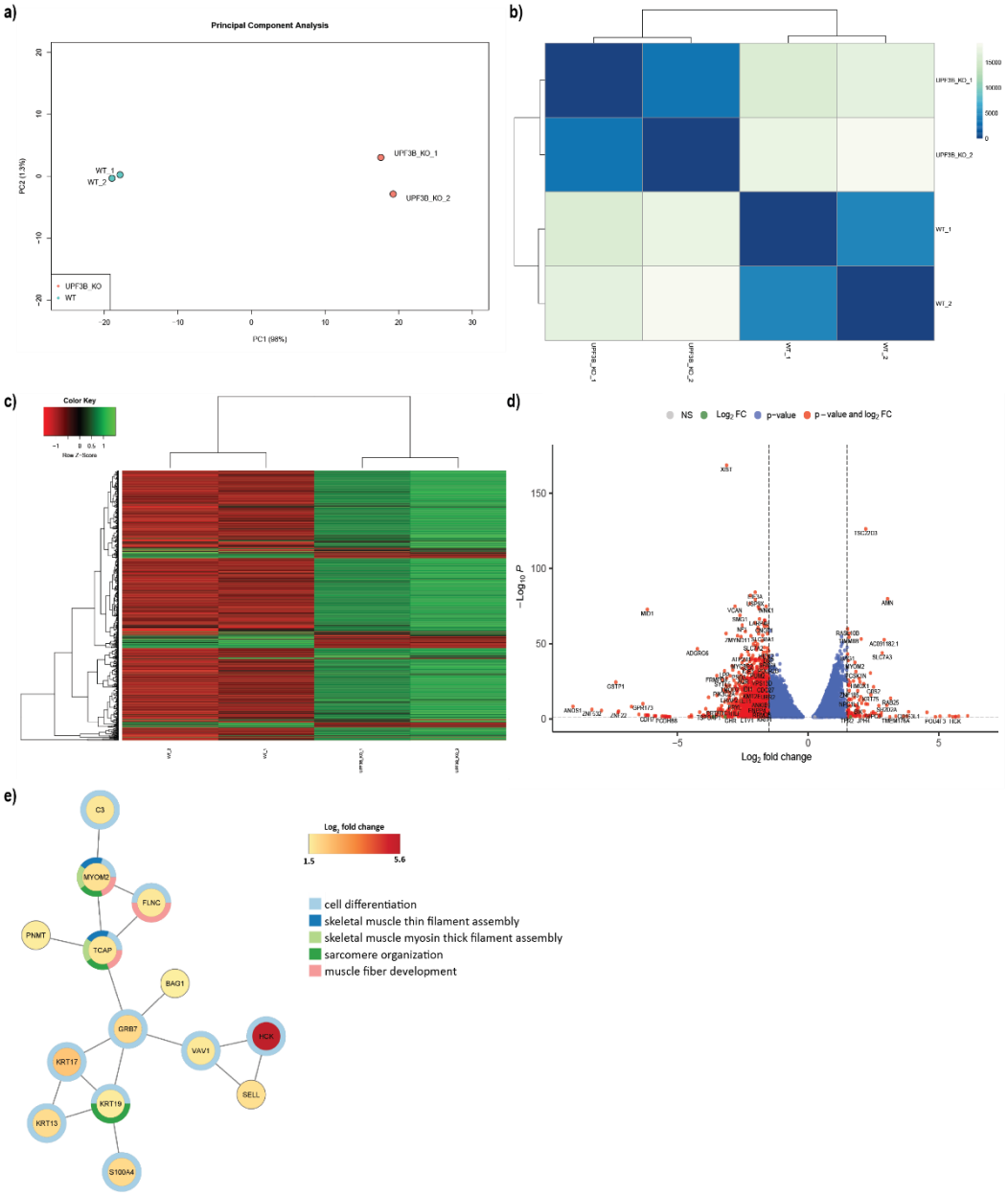

Figure 3. RNA-Seq analysis of *UPF3B*-KO cells. a) Principal component analysis (PCA) of *UPF3B*-KO and wild-type cells, with each point representing the biological replicates. Data was plotted along the first and second principal components. b) A sample heatmap showing the dissimilarity between counts calculated by the Poison distance. c) A heatmap of hierarchical clustering of top 1000 significant genes (Adjusted p-value < 0.05). d) Volcano plot showing the differential gene expression of RNA-Seq data (*UPF3B*-KO vs WT). The log2 fold change is plotted against the –log10 adjusted p-value. e) Protein-protein interaction shows a network of upregulated genes. The colour pattern of donut around the genes shows the biological processes associated with that gene.

To delineate the functional importance of DEGs, up/downregulated genes were classified according to the Gene Ontology (GO) classification of biological processes (Figure 4a), molecular functions (Figure S2b), and cellular components (Figure S2c). Based on GO, DEGs were observed to be significantly enriched in the regulation of transcription, response to DNA damage, cell division, and the cell cycle. In addition, the majority of DEGs were associated with metal ion binding and had ATP-binding features. These genes are localized in the nucleus and cytoplasm in equal proportions. In conclusion, UPF3B depletion significantly alters the expression of genes involved in the regulation of

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

transcription. Most of these genes require metal ions as cofactors and ATP as an energy source for their function and are confined to the nucleus.

2.4. Loss of UPF3B alters the overall cell growth with no apoptosis evident

GO analysis of DEGs revealed that genes are significantly enriched in cell division and cell cycle processes (Figure 4a). To understand how a lack of UPF3B alters cell growth and development, we identified cell cycle-related biological processes which are deregulated upon UPF3B depletion (Figure 4b). The circular plot of 10 deregulated cell cycle biological processes showed a vast number of genes were downregulated (blue dots); in contrast, few genes were upregulated (red dots). PPI analysis of 141 genes associated with cell cycle-related biological processes identified 140 proteins in the STRING database, of which 124 protein-coding genes formed the largest cluster (Figure 4c). These genes were associated with the cell cycle, cell division, and regulation of the cell cycle. In line with the GO data, we found that the proliferation of *UPF3B*-KO cells was compromised compared to that of the wild-type; however, there was no cell death (Figure 4d). The cell proliferation assay showed that the population doubling time of *UPF3B*-KO cells was almost twice the amount (48 h), compared to that of the wild-type cells (approximately 26 h) (Figure 4e). Next, we examined cell viability using the MTT assay and observed a decreased viability of *UPF3B*-KO cells (Figure 4f).

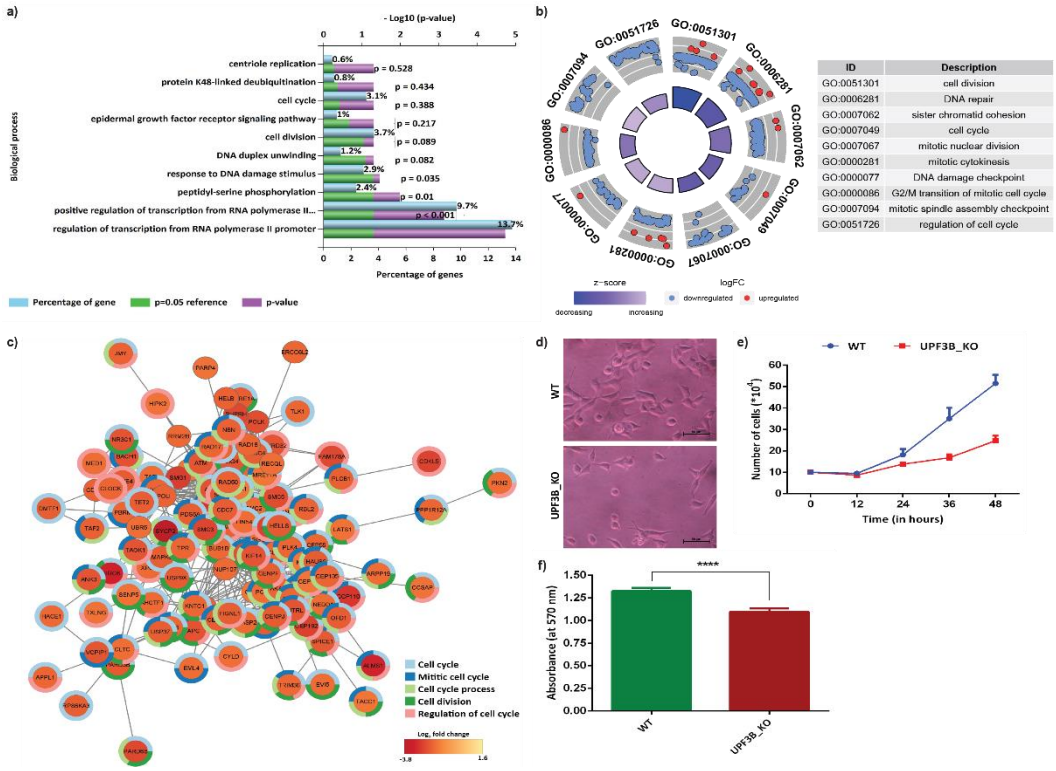

Figure 4. UPF3B-depletion alters the overall cell growth. a) Gene ontology classification of biological processes. b) A circular plot showing DEGs involved in cell cycle processes. Red dots represent upregulated genes whereas blue dots correspond the downregulated genes. z-score is calculated by the number of up- regulated genes minus the number of down- regulated genes divided by the square root of the count. c) PPI analysis of DEGs shows the largest cluster of 124 genes. The colour pattern of donut around the genes shows the cellular processes associated with that gene. d) Reduced cell growth of UPF3B-KO cells compared to wild-type cells. e) Cell proliferation assay of *UPF3B*-KO cells (red) compared to wild-type (blue). The number of cells are plotted with the time (in hours). d) MTT assay shows the decreased absorbance at 570 nm after 24 hours showing less number of cells in UPF3B-KO compared to WT.

To elucidate the molecular basis underlying the altered proliferation of *UPF3B*-KO cells, we evaluated the effect of *UPF3B*-KO on cell cycle progression using flow cytometry. Cell cycle analysis showed that cells were distributed more in the S-phase and less in the G1 phase after depletion of UPF3B (Figure 5a), suggesting an altered cell cycle progression

 doi:10.20944/preprints202111.0013.v1

upon *UPF3B*-KO. Furthermore, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis of DEGs showed that 12 genes were associated with the cell cycle pathway (hsa04110; Figure S3a). The cell cycle genes related to various phases of the cell cycle are shown in Figure 5c. The qRT-PCR analysis of 11 cell cycle genes showed a significant decrease in mRNA expression of *BUB1B*, *RBL1*, *RBL2*, *CDC7*, *CDK6*, and *SMC3* (Figure 5d). The mRNA expression of *CDC27* and *PRKDC* was downregulated, but not with a significant fold change. However, *GADD45G* mRNA levels were significantly upregulated. The DNA damage checkpoint genes, *ATM* and *ATR*, did not display an altered mRNA expression.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

Figure 5. *UPF3B*-KO induces cell cycle arrest without apoptosis of the cells. a) Cell cycle analysis performed using flow cytometry in *UPF3B*-KO cells. b) Apoptosis rate was tested using flow cytometry after UPF3B-depletion. c) Schematic representation of cell cycle genes associated with various cell cycle phases. d) qRT-PCR analysis of cell cycle genes.

 doi:10.20944/preprints202111.0013.v1

Flow cytometry analysis with Annexin V and propidium iodide (PI) staining revealed no difference in apoptotic rates of *UPF3B*-KO cells compared to the control cells (Figure 5b). Altogether, these results suggest that UPF3B is required for efficient cell growth and is probably an important regulator of the cell cycle.

2.5. UPF3B affects the expression of axon guidance and other neuron-specific genes

To analyze the impact of UPF3B-depletion on neuronal processes, we performed GO analysis of DEGs in biological processes and identified 12 neuronal processes associated with neuronal development and function (Figure 6a). We observed that several genes involved in neuronal processes were downregulated (blue dots), whereas only a few genes were upregulated (red dots). A total of 101 genes were related to 12 neuronal processes and PPI network, while 74 genes formed the largest cluster out of the rest of the identified genes (Figure 6b). The majority of the genes were involved in central nervous system development and brain development.

Next, the KEGG pathway analysis of DEGs revealed that 21 genes were associated with the axon guidance pathway (hsa04360; Figure S3b). A hierarchical clustering heatmap showed a difference in the transcript expression of genes involved in the axon guidance pathway (Figure 6c). The axon guidance genes, which are involved in various interactions and signaling pathways, are shown in Figure 6d. qRT-PCR analysis of 11 axon guidance genes showed significant decreases in mRNA levels of *SEMA3D*, *SEMA6D*, *PLXNC1*, *ARHEGF12*, *ITGB1*, *ROCK2*, *NTN4*, *PTPN11*, *TRPC4*, and *CAMK2D* (Figure 6e). The mRNA level of *ROBO1* was decreased but not significantly changed. SEMA3D, SEMA6D, ARHGEF12, PLXNC1, and ITGB1 were identified as protein components of the semaphorin interactions. Furthermore, NTN4, PTPN11, and TRPC4 were identified as protein components of the Netrin signaling pathway. The mRNA expression of other neuron-specific genes was also altered in *UPF3B*-KO cells. The mRNA level of the *ARC* gene was upregulated with significant fold change, whereas *ARHGAP21* mRNA expression decreased significantly (Figure 6f). However, there was no significant fold change in the mRNA levels of *FMR1* and *ATRX* genes.

Functional enrichment analysis of 1526 downregulated genes identified the genes associated with developmental disorders of mental health (75 genes), intellectual disability (64 genes), and specific developmental disorder (70 genes). PPI analysis of these disease-associated genes generated a cluster of 60 protein-coding genes shown in figure 7.

Altogether, UPF3B depletion deregulates the genes involved in axon guidance, thereby affecting the development of the central nervous system. Deregulation of neuron-specific genes results in developmental disorders.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

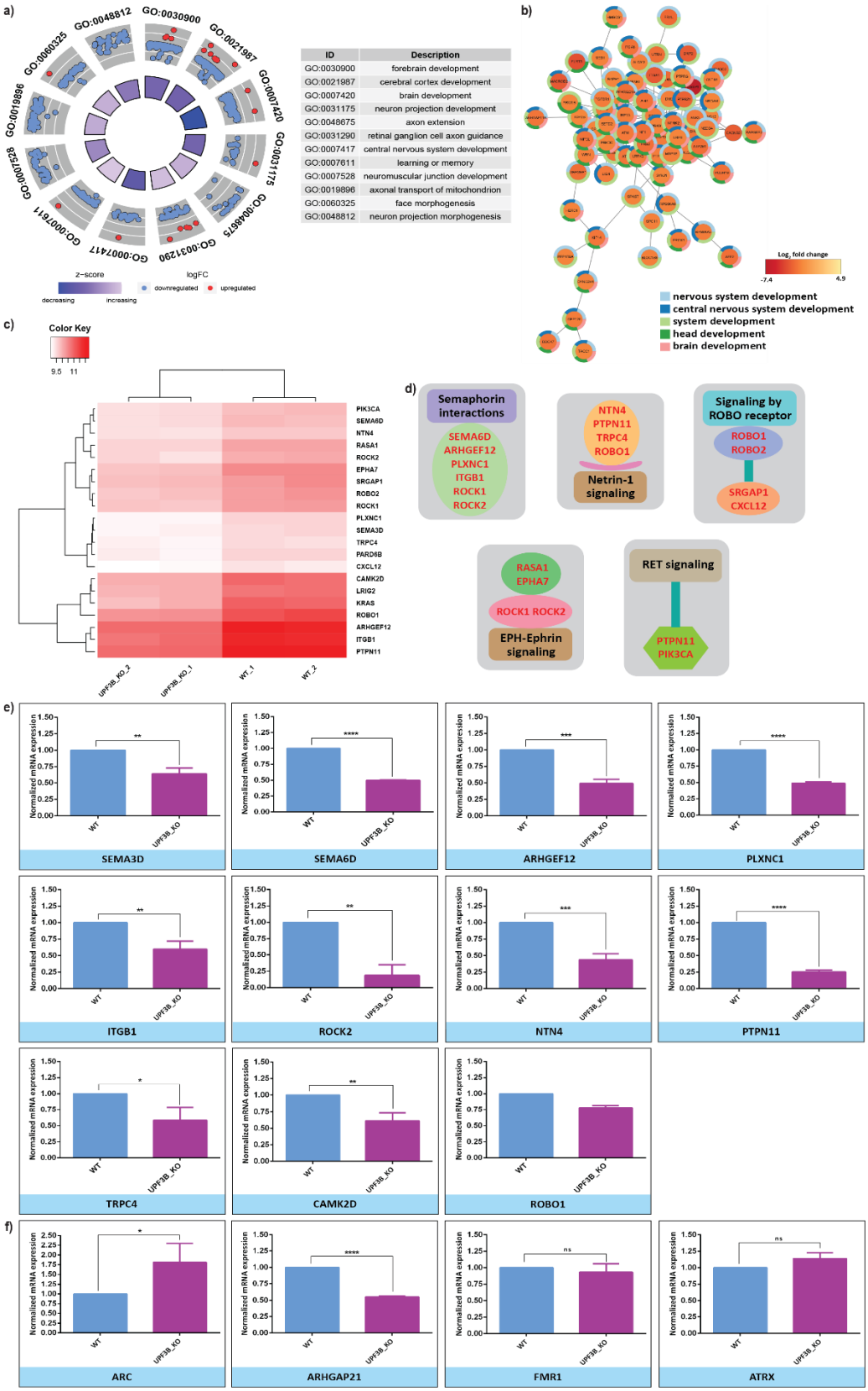

Figure 6. Loss of UPF3B alters the expression of genes associated with neuronal processes. a) Circular plot showing DEGs involved in the neuronal processes. Red dots represent upregulated genes whereas blue dots correspond the downregulated genes. z-score is calculated by the number of up- regulated genes minus the number of down- regulated genes divided by the square root of the count. b) PPI analysis of DEGs shows the largest cluster of 74 protein-coding genes. The colour pattern of donut around the genes shows the neuronal processes associated with that gene. c) A heatmap of hierarchical clustering of axon guidance genes in *UPF3B*-KO and WT samples. d)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

Clustering of axon guidance genes involved in various signaling cascades. e) qRT-PCR analysis of axon guidance genes. f) qRT-PCR of other neuron-specific genes.

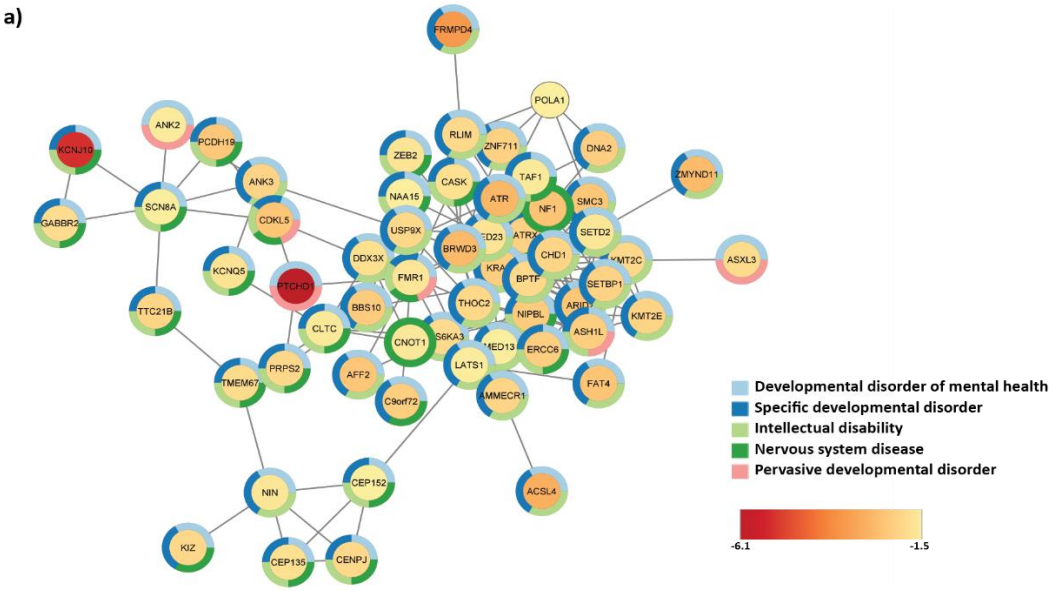

Figure 7. PPI analysis of downregulated genes shows the largest cluster of 60 genes. The colour pattern of donut around the genes shows the neuronal diseases associated with that gene.

3. Discussion

UPF3B mutations lead to compromised mental health and an underdeveloped central nervous system in patients with behavioral changes [11, 13-16]. Furthermore, neurogenesis defects were observed in a mouse model with *UPF3B* mutations [18]. However, the mechanistic role of UPF3B in neuronal development and functions is poorly understood. Researchers are endeavoring to generate a model system to study the functional aspects of UPF3B in CNS development. We developed a *UPF3B*-KO HEK-293 cell line that could be used as a model system to study the role of UPF3B in various biological processes in humans. HEK-293 cells are widely used in neurobiological studies, since they originate from the same precursor line (neural crest) as neurons; hence they have similar biological processes and regulatory mechanisms. Therefore, *UPF3B*-KO HEK-293 cells provide an approximation to address the questions relevant to the biology of neurons.

UPF3B is a component of the classical NMD pathway. However, loss of UPF3B does not affect NMD activity in HEK-293 cells, validating the existence of alternative NMD pathways [2, 3]. In our study, the downregulation of the mRNA levels of other NMD factors, UPF1 and UPF2, indicates the possibility that neither UPF1- nor UPF2-dependent NMD pathway performs the NMD activity. However, the upregulated UPF3A protein might be sufficient to compensate for the loss of UPF3B and restore NMD activity. In the future, it would be interesting to determine the subset of NMD-sensitive transcripts targeted by the UPF3B-mediated NMD pathway.

Subsequently, differential gene expression analysis of the cellular transcriptome was performed after *UPF3B*-KO identified many deregulated genes. Most DEGs were downregulated with a substantial fold change (1526 genes), whereas 107 genes were upregulated. *MID1* and *ANOS1* are two essential downregulated genes associated with X-linked disorders. Mutations in *MID1* and *ANOS1* lead to Opitz Gbbb syndrome [25] and Kallmann syndrome [26], respectively. GO analysis of DEGs classified most of the genes related to transcription regulation. Since UPF3B is involved in the homeostasis of the cellular transcriptome, it might regulate a set of transcripts involved in the regulation of transcription. Alternatively, there might be PPIs that claim that UPF3B functions as a mediator, if it is not directly involved.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

A subset of the DEGs is enriched in the biological processes involved in cell growth, which corresponds to our observation in *UPF3B*-KO cells with impaired cell growth without apoptosis. Cell cycle analysis of *UPF3B*-KO cells revealed altered cell cycle phases. The low distribution of *UPF3B*-KO cells in the G1-phase indicates that cells skip the preparatory phase for DNA replication. In contrast, *UPF3B*-KO cells are distributed more in the S-phase, indicating that the cells strive to perform DNA replication and transcription processes upon deregulation of the transcription factors. The genes involved in the cell cycle pathway, *BUB1B, RBL1, RBL2, CDC7, CDK6*, and *SMC3*, were significantly downregulated in *UPF3B*-KO cells. BUB1B, also known as BUBR1, is an essential component of the mitotic checkpoint, containing a spindle checkpoint function. The depletion of BUB1B in the G2 phase has been shown to accelerate the entry of cells to the M-phase [27]. BUB1B is highly upregulated in various cancer types, including lung cancer cells [28], gastric cancer [29], breast cancer cell lines and primary tumor samples [30]; thus, reflecting active cell division. Therefore, the decreased expression of BUB1B might be associated with the compromised growth of *UPF3B*-KO cells. Other cell cycle genes such as *RBL1, RBL2, CDC7, CDK6, SMC3* are involved in the regulation of cell cycle progression. The disease associated with CDK6 and SMC3 include primary autosomal recessive microcephaly [31] and cornelia de Lange syndrome with predominant mental retardation [32], respectively. The cell cycle genes might also affect neuronal processes leading to defects in brain function and development.

In addition, DEGs identified in *UPF3B*-KO cells were associated with neuronal processes, including CNS development. The axon guidance pathway was enriched with the DEGs involved in the neuronal processes, where all the relevant genes were significantly downregulated. These genes are involved in various signaling pathways implementing the axon guidance collectively. The genes *SEMA3D, SEMA6D, PLXNC1, ARHEGF12*, and *ITGB1* are involved in semaphorin interactions. SEMA3D is a member of class III semaphorins which are secreted signaling proteins, whereas SEMA6D, the class VI semaphorin, is a transmembrane signaling protein [33]. Semaphorins interact with the plexin receptors to induce semaphorin signaling. The plexin receptor, Plexin C1 (PLXNC1), interacts with semaphorins via its sema domain to execute the signaling cascade. Semaphorin-plexin signaling regulates neural circuit development, including axon growth, axon bundling, and pruning [34]. NTN4, PTPN11, TRPC4, and ROBO1 are components of the Netrin signaling pathway. Netrins are the secreted signaling molecules that have a conserved function of attracting the axons to the midline during brain commissure formation [35]. Semaphorins and netrins are two conserved families of axon guidance molecules. Changes in the expression of these molecules affect the neural circuit, leading to neurological disorders. Mutational screening in patients with Hirschsprung disease and genome-wide copy number variation analysis in patients with autism, identified deleterious variants of the SEMA3D gene [36, 37]. These findings suggest differential transcript expression and usage of neuron-specific genes in *UPF3B*-KO cells.

The neuron-specific gene *ARC* is a master regulator of synaptic plasticity in neurons. Depletion of UPF3B leads to the stabilization of *ARC* mRNA by UPF3A [38]. The upregulation of *ARC* mRNA in *UPF3B*-KO cells validates the role of UPF3B in regulating a subset of mRNAs involved in neuronal functions. PPI network analysis shows that the majority of downregulated genes are associated with neurodevelopmental disorders.

Altogether, *UPF3B*-KO in HEK-293 cells does not affect the NMD activity in the cells. However, loss of UPF3B alters the overall cellular growth by cell cycle arrest with no perceptible apoptosis. The downregulation of cell cycle genes is accompanied by the upregulation of genes involved in cell differentiation (Figure 3e), which suggests that UPF3B might promote the cell cycle progression and get depleted upon cellular differentiation. UPF3B depletion alters the expression of neuron-specific genes and genes associated with the axon guidance pathway. The deregulated genes are associated with neurodevelopmental disorders indicating the significance of UPF3B in neuronal function and develop-

 doi:10.20944/preprints202111.0013.v1

ment. Our study has identified the potential players of neuronal processes and CNS development and will provide more insight into the association of UPF3B with neuro-developmental disorders.

4. Materials and Methods

4.1. Mammalian Cell Culture

All mammalian cell culture experiments were performed in Flp-in T-Rex HEK-293 cell lines (R78007, Invitrogen). Cells were cultured in Dulbecco Modified Eagle Medium (DMEM), High glucose (AT007F, Himedia) supplemented with 10% (v/v) fetal bovine serum, FBS (RM10432, Himedia), and 1% Penicillin-streptomycin solution (A014, Himedia). Cells were maintained at 37°C in humidified condition with 5% CO2 supply.

4.2. Plasmids

The cumate-inducible PB-CuO-3xFLAG-Cas9-IRES-GFP-EF1α-CymR-Puro and pCI-FLAG-PB-Transposase plasmids were used for the stable genome integration of the Cas9 gene. 3xFLAG-Cas9 sequence was extracted from px330 plasmid and cloned into multiple cloning sites of PB-CuO-MCS-IRES-GFP-EF1α-CymR-Puro vector between NheI and NotI sites. PB-CuO-3xFLAG-Cas9-IRES-GFP-EF1α-CymR-Puro plasmid was a generous gift from Prof. Neils Gehring.

4.3. Cell Transfection

PB-CuO-3xFLAG-Cas9-IRES-GFP-EF1α-CymR-Puro and pCI-FLAG-PB-Transposase plasmids were co-transfected to 106 Flp-in T-Rex 293 cells using PEI transfection reagent (1:3 ratio). 72-hour post-transfection, cells were trypsinized and replenished with DMEM medium containing 2µg/mL puromycin (ant-pr, InvivoGen). The cells were selected in puromycin for 14 days; afterward, the final selection was made using DMEM medium with 3µg/mL puromycin for 5 more days. The cumate inducer (268402, Sigma) was used to induce the GFP expression and visualized under the microscope (Axiovert A1, Zeiss). Genomic DNA was isolated from the cells using HiPurA Mammalian Genomic DNA Purification Kit (MB506, Himedia) following the manufacturer's protocol. PCR amplification was performed using Taq 2X master mix (M0270L, NEB) and specific primers (Supplementary File S2) for confirmation.

4.4. CRISPR/Cas9-mediated Gene-knockout using Single-guide RNA

Predesigned UPF3B CRISPR RNA (crRNA) and trans-crRNA (1075927, IDT) were synthesized from Integrated DNA Technologies (IDT). UPF3B guide RNA was prepared by adding equimolar concentrations of crRNA and tracrRNA. The equal volume (1 µL each) of crRNA and tracrRNA mixed with 98 µL Nuclease-free duplex buffer (11-01-03-01, IDT) was heated at 95°C for 5 minutes and lowered down to the room temperature for the formation of 1µM UPF3B guide RNA. Approximately 60,000 cells (stable HEK-293 cells generated in previous section) were supplemented with the growth medium containing cumate inducer. After 24 hours of induction, UPF3B guide RNA was transfected to the cells using Lipofectamine RNAiMAX (13778030, Invitrogen). After confluency, cells were serially diluted in 96-well plate, the wells with single-cell were identified and marked. Cells were grown and stored using freezing medium (90% FBS and 10% Dimethyl sulfoxide, DMSO) in a deep freezer (−80°C).

4.5. SDS-PAGE and Western Blotting

106 cells from individual clones were seeded in 35 mm dishes. After 24 hours of seeding, the growth medium was removed and were washed twice with Dulbecco's phosphate-buffered saline, DPBS (TS1006, Himedia). Cells were replenished with RIPA buffer and were scrapped using a cell scraper (960052, Tarsons). Lysed cells were transferred to the tubes on ice. Cells were entirely lysed by sonication using Ultrasonic Homogenizer (Fisher Scientific). The homogenate was centrifuged and the supernatant was transferred to a new tube and stored in a deep freezer (−80°C). A 12% SDS-PAGE gel was prepared with BIO-RAD mini gel cast system using combs with 10 well. The protein concentration of whole-cell lysates was measured using PierceTM Detergent Compatible Bradford Assay reagent (PI23246, Thermo Scientific) at 595 nm in a microplate reader (Biorad). The

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 1 November 2021 doi:10.20944/preprints202111.0013.v1

cell lysate samples (20 µg each) were resolved on 12% SDS-PAGE gel using SDS running buffer in a vertical electrophoresis tank at 220 V for 45 min. After resolving the protein samples in electrophoresis, gel and nitrocellulose membrane (S020A330RI, Axiva) were equilibrated in 1x transfer buffer (25 mM Tris base, 52 mM glycine and 20% methanol) for 5 min. The blot sandwich was prepared as follows: 4 filter papers, nitrocellulose membrane, gel, and 4 filter papers (from anode to cathode). The transfer was completed using the PierceTM Power blotter (Thermo Scientific) semi-dry system at 19 V and 0.5 A for 25 min. The membrane was stained with Ponceau S (0.25g Ponceau S and 0.5 mL acetic acid were added to prepare 50 mL of staining solution) for 2-3 minutes to visualize the bands. The membrane was washed twice with Tris-buffered saline-Tween (TBST; 20 mM Tris base, 150 mM, pH 7.5; 0.1% w/v Tween-20) for 5 min. The membrane was blocked with 5%(w/v) skim milk in TBST. After blocking, the membrane was washed six times with TBST for 5 minutes each. The following primary antibodies were used to probe the membrane: Histone (1:10,000 dilution; BioBharati) and UPF3B (1:2000 dilution). All primary antibodies were diluted in TBST with 1% skim milk. The membrane was incubated with primary antibodies overnight at 4°C. The membranes were washed thrice with TBST, and anti-rabbit secondary antibody (1:5,000 dilution) was used to probe the membranes for 1 hour at room temperature. The membranes were developed using ClarityTM Western ECL Substrate (1705061, Biorad) and visualized in ChemiDoc XRS+ system (Biorad).

4.6. Polymerase Chain Reaction and Sequencing

Genomic DNA from the individual clones was isolated using HiPurA Mammalian Genomic DNA Purification Kit (MB506, Himedia) following the manufacturer's protocol. PCR amplification was performed using gene specific primers (Supplementary File S2). PCR amplicons were resolved on 1% agarose gel. UPF3B PCR products were purified using PCR purification kit (28104, QIAGEN) and Sanger sequencing was performed.

4.7. Quantitative real-time PCR

106 cells were seeded in 35 mm dish in triplicates for wild-type and *UPF3B*-KO cells each. Total RNA was isolated using TRIzol (15596018, Invitrogen). 1 µg of RNA from each replicate was used for cDNA synthesis using Super Reverse Transcriptase (MuLV) kit (BB-E0043, Biobharati). For qRT-PCR, the reaction mixture contained, 5 µL of PowerUp SYBR Green Master Mix (A25742, Applied Biosystems); 1 µL cDNA template (1:3 dilution); 0.5 µL specific primers (0.25 µL reverse primer and 0.25 µL forward primer), and 3.5 µL Nuclease-FreeWater. Actin gene was used as an endogenous control gene for normalizing the mRNA expression.

4.8. RNA-Seq analysis

Total RNA was isolated from *UPF3B*-KO and wild-type cell lines in duplicates using TRIzol method (15596018, Invitrogen), quality evaluated by TapeStation (Agilent), and sequenced with an Illumina NovaSeq 6000 High-Throughput Sequencing System (Illumina). Adaptor contamination was removed from raw reads using cutadapt (v3.3) [39]. Trimmed reads were mapped to the reference human genome (GRCh38.p13) using the RNA-seq aligner STAR (v2.7.8a) [40]. Counts for each gene were quantified using htseq-count (v0.13.5) [41] and annotated using the Genecode GRCh38 human genome. Reads were filtered, such that genes without at least one sample with at least 10 raw reads and one RPKM reads were removed from the analysis. The count data were normalized and differential expression was performed using the R (v.3.1.1) package DESeq2 (v.1.4.5) (Bioconductor) [42]. In brief, DESeq2 uses negative binomial generalized linear models and shrinkage estimation for dispersions and fold changes to improve stability and interpretability of the estimates. It reports a P-value and an adjusted P-value using the Benjamini–Hochberg procedure. Genes with an adjusted P-value of 0.05 were considered differentially expressed unless otherwise noted. Other plots were constructed using the R (v.3.1.1) package ggplots. All functional enrichment analyses were generated using Funrich. Protein-protein interaction maps were generated using Cytoscape (v3.8.2) [43].

4.9. Cell Proliferation Assay

105 cells from wild-type and knockout lines were seeded in 24-well plates in triplicates. Relative cell proliferation was determined by cell counting at the time intervals of

 doi:10.20944/preprints202111.0013.v1

12, 24, 36 and 48 hours using the Trypan blue (302643, Sigma) staining method in hemocytometer.

4.10. MTT Assay

Cells were harvested and seeded onto 96-well plates (100 µL, 5 × 104/well) and placed in an incubator in humidified condition (37°C, 5% CO2) for 24 hours. Then, cells treated with different factors were added (an equal volume of PBS for the blank control group) with 6 replicates for each group. After 24 h culture, with the addition of 50 µL MTT (5 g/L) (TC191, Himedia), the cells were cultured for another 4 h (37 ◦C). Finally, we aspirated the supernatant, added 150 µL DMSO to each well and shook it well with a plate shaker. After the crystals were dissolved, a microplate reader was taken to examine each well's OD value at a wavelength of 570 nm.

4.11. Cell Cycle and Apoptosis Assays

106 cells from wild-type and knockout lines were seeded in 60 mm dish. After 24 hours, cells were harvested and fixed in ice-cold 70% ethanol at -20°C overnight. After removal of ethanol, cells were washed with PBS and treated with RNase I. RNase treated cells were stained with propidium iodide (PI) (V13241, Life technologies) and measured by the flow cytometry. The cells were stained with Annexin V and PI and subsequently the ratio of apoptotic cells was tested by flow cytometry.

4.12. Validation of Gene Expression

The validation of differentially expressed genes with significant fold changes were performed using the qRT-PCR. The method consists of two major steps: (1) Total RNA was isolated from wild-type and knockout samples in triplicates. 2 µg of RNA was reverse transcribed using Super Reverse Transcriptase (MuLV) kit (BB-E0043, Biobharati). (2) main qRT-PCR reaction. The reaction mixture of each sample consisted of: 5 µL of PowerUp SYBR Green Master Mix (A25742, Applied Biosystems); 1 µL cDNA template (1:3 and 1:5 dilutions); 0.5 µL specific primers (0.25 µL reverse primer and 0.25 µL forward primer), and 3.5 µL Nuclease-FreeWater. Actin gene was used as an endogenous control gene for normalizing the mRNA expression. The relative gene expression was quantified using the comparative Ct method (ΔΔCt, where ΔCt = [Ct of target genes] - [Ct of endogenous control gene]).

4.13. Statistical Methods

All statistics were calculated using Prism Version 6.0 (GraphPad) and Microsoft Excel.

5. Conclusions

UPF3B depletion in mammalian cell lines showed reduced cell growth caused by the cell cycle arrest. The genes associated with the cell cycle and neuronal processes were significantly downregulated in *UPF3B*-KO cells. These downregulated genes were associated with the previously reported neuro-developmental disorders. UPF3B seems to play a critical role in regulating the expression of neuron-specific genes either by regulating a subset of neuronal transcripts via NMD or by mediating the interaction with the transcriptional regulators. Further, there is a scope to identify the neuron-specific transcripts and other NMD-sensitive transcripts with an altered expression upon UPF3B depletion that might play a significant role in neuronal development.

Supplementary Materials: Supplementary Information; Supplementary File S1; Supplementary File S2.

Author Contributions: Conceptualization, K.K.S.; methodology, P.C., B.D., S.K., A.R. and P.Y.; software, P.C.; validation, P.C., B.D. and A.R.; formal analysis, P.C., K.K.S. and B.D.; investigation, P.C., B.D. and K.K.S.; resources, K.K.S.; IITG; data curation, P.C.; writing—original draft preparation, P.C. and K.K.S.; writing—review and editing, P.C., B.D. and K.K.S.; visualization, P.C.; supervision, K.K.S.; project administration, K.K.S.; funding acquisition, K.K.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by ICMR (2020-0476/CMB/ADOC-BMS) and DST-SERB (CRG/2019/001352). P.C., B.D., S.W., A.R. and P.Y. are funded under the scholarships provided by the Ministry of Human Resource Development (MHRD), Govt. of India.

Data Availability Statement: The link will be provided.

Acknowledgments: We acknowledge the facilities provided by the Indian Institute of Technology, Guwahati, Assam, India. We thank Prof. Neils Gehring for the cell lines and plasmids. We acknowledge Ms. Renu Sharma and Mr. Adhiraj Nath for their technical assistance.

Conflicts of Interest: The authors declare no competing interests.

References

1. Cheng, J. and L.E. Maquat, Nonsense codons can reduce the abundance of nuclear mRNA without affecting the abundance of pre-mRNA or the half-life of cytoplasmic mRNA. Mol Cell Biol, 1993. 13(3): p. 1892-902.
2. Gehring, N.H., et al., Exon-Junction Complex Components Specify Distinct Routes of Nonsense-Mediated mRNA Decay with Differential Cofactor Requirements. Molecular Cell, 2005. 20(1): p. 65-75.
3. Chan, W.-K., et al., An alternative branch of the nonsense-mediated decay pathway. The EMBO journal, 2007. 26(7): p. 1820-1830.
4. Deka, B., P. Chandra, and K.K. Singh, Functional roles of human Up-frameshift suppressor 3 (UPF3) proteins: From non-sense-mediated mRNA decay to neurodevelopmental disorders. Biochimie, 2021. 180: p. 10-22.
5. Kim, V.N., N. Kataoka, and G. Dreyfuss, Role of the Nonsense-Mediated Decay Factor hUpf3 in the Splicing-Dependent Exon-Exon Junction Complex. Science, 2001. 293(5536): p. 1832.
6. Kadlec, J., E. Izaurralde, and S. Cusack, The structural basis for the interaction between nonsense-mediated mRNA decay factors UPF2 and UPF3. Nature Structural &Amp; Molecular Biology, 2004. 11: p. 330.
7. Kadlec, J., et al., Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1. 2006. 12(10): p. 1817-1824.
8. Kunz, J.B., et al., Functions of hUpf3a and hUpf3b in nonsense-mediated mRNA decay and translation. RNA (New York, N.Y.), 2006. 12(6): p. 1015-1022.
9. Neu-Yilik, G., et al., Dual function of UPF3B in early and late translation termination. The EMBO journal, 2017. 36(20): p. 2968-2986.
10. Jolly, L.A., et al., The UPF3B gene, implicated in intellectual disability, autism, ADHD and childhood onset schizophrenia regulates neural progenitor cell behaviour and neuronal outgrowth. Human Molecular Genetics, 2013. 22(23): p. 4673-4687.
11. Laumonnier, F., et al., Mutations of the UPF3B gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism. Molecular Psychiatry, 2009. 15: p. 767.
12. Notaras, M., et al., UPF2 leads to degradation of dendritically targeted mRNAs to regulate synaptic plasticity and cognitive function. Molecular Psychiatry, 2020. 25(12): p. 3360-3379.
13. Tarpey, P.S., et al., Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. Nature Genetics, 2007. 39: p. 1127.
14. Addington, A.M., et al., A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. Molecular Psychiatry, 2010. 16: p. 238.
15. Szyszka, P., et al., A nonconservative amino acid change in the UPF3B gene in a patient with schizophrenia. 2012. 22(3): p. 150-151.
16. Lynch, S.A., et al., Broadening the phenotype associated with mutations in UPF3B: Two further cases with renal dysplasia and variable developmental delay. European Journal of Medical Genetics, 2012. 55(8): p. 476-479.
17. Nguyen, L.S., et al., Transcriptome profiling of UPF3B/NMD-deficient lymphoblastoid cells from patients with various forms of intellectual disability. Molecular Psychiatry, 2011. 17: p. 1103.
18. Huang, L., et al., A Upf3b-mutant mouse model with behavioral and neurogenesis defects. Molecular Psychiatry, 2018. 23(8): p. 1773-1786.
19. Wallmeroth, D., et al., UPF3A and UPF3B are redundant and modular activators of nonsense-mediated mRNA decay in human cells. 2021.
20. Yi, Z., et al., Mammalian UPF3A and UPF3B activate NMD independently of their EJC binding. 2021.
21. Boehm, V., et al., SMG5-SMG7 authorize nonsense-mediated mRNA decay by enabling SMG6 endonucleolytic activity. Nature Communications, 2021. 12(1): p. 3965.
22. Sureau, A., et al., SC35 autoregulates its expression by promoting splicing events that destabilize its mRNAs. Embo j, 2001. 20(7): p. 1785-96.
23. Lykke-Andersen, S., et al., Human nonsense-mediated RNA decay initiates widely by endonucleolysis and targets snoRNA host genes. 2014. 28(22): p. 2498-2517.
24. Chan, W.-K., et al., A UPF3-mediated regulatory switch that maintains RNA surveillance. Nature Structural & Molecular Biology, 2009. 16: p. 747.
25. Fontanella, B., G. Russolillo, and G. Meroni, MID1 mutations in patients with X-linked Opitz G/BBB syndrome. Hum Mutat, 2008. 29(5): p. 584-94.

26. Albuisson, J., et al., Kallmann syndrome: 14 novel mutations in KAL1 and FGFR1 (KAL2). Hum Mutat, 2005. 25(1): p. 98-9.
27. Park, S.Y., et al., Depletion of BubR1 promotes premature centrosomal localization of cyclin B1 and accelerates mitotic entry. Cell Cycle, 2009. 8(11): p. 1754-64.
28. Haruki, N., et al., Molecular analysis of the mitotic checkpoint genes BUB1, BUBR1 and BUB3 in human lung cancers. Cancer Letters, 2001. 162(2): p. 201-205.
29. Grabsch, H., et al., Overexpression of the mitotic checkpoint genes BUB1, BUBR1, and BUB3 in gastric cancer--association with tumour cell proliferation. J Pathol, 2003. 200(1): p. 16-22.
30. Yuan, B., et al., Increased Expression of Mitotic Checkpoint Genes in Breast Cancer Cells with Chromosomal Instability. 2006. 12(2): p. 405-410.
31. Hussain, M.S., et al., CDK6 associates with the centrosome during mitosis and is mutated in a large Pakistani family with primary microcephaly. Hum Mol Genet, 2013. 22(25): p. 5199-214.
32. Deardorff, M.A., et al., Mutations in cohesin complex members SMC3 and SMC1A cause a mild variant of cornelia de Lange syndrome with predominant mental retardation. Am J Hum Genet, 2007. 80(3): p. 485-94.
33. Alto, L.T. and J.R. Terman, Semaphorins and their Signaling Mechanisms. Methods Mol Biol, 2017. 1493: p. 1-25.
34. Pasterkamp, R.J., Getting neural circuits into shape with semaphorins. Nat Rev Neurosci, 2012. 13(9): p. 605-18.
35. Lai Wing Sun, K., J.P. Correia, and T.E. Kennedy, Netrins: versatile extracellular cues with diverse functions. Development, 2011. 138(11): p. 2153-69.
36. Luzón-Toro, B., et al., Mutational Spectrum of Semaphorin 3A and Semaphorin 3D Genes in Spanish Hirschsprung patients. PLOS ONE, 2013. 8(1): p. e54800.
37. Sbacchi, S., et al., Functional annotation of genes overlapping copy number variants in autistic patients: focus on axon pathfinding. Current genomics, 2010. 11(2): p. 136-145.
38. Shum, E.Y., et al., The Antagonistic Gene Paralogs Upf3a and Upf3b Govern Nonsense-Mediated RNA Decay. Cell, 2016. 165(2): p. 382-395.
39. Martin, M., Cutadapt removes adapter sequences from high-throughput sequencing reads. 2011, 2011. 17(1): p. 3 %J EMBnet.journal.
40. Dobin, A., et al., STAR: ultrafast universal RNA-seq aligner. Bioinformatics (Oxford, England), 2013. 29(1): p. 15-21.
41. Anders, S., P.T. Pyl, and W. Huber, HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics, 2014. 31(2): p. 166-169.
42. Love, M.I., W. Huber, and S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology, 2014. 15(12): p. 550.
43. Shannon, P., et al., Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res, 2003. 13(11): p. 2498-504.Author 1, A.B.; Author 2, C.D. Title of the article. *Abbreviated Journal Name* **Year**, *Volume*, page range.