Oncogene
https://doi.org/10.1038/s41388-019-0714-9

ARTICLE

Nudt21 regulates the alternative polyadenylation of *Pak1* and is predictive in the prognosis of glioblastoma patients

Yuan Chu[1,2] · Nathan Elrod[1] · Chaojie Wang[3] · Lei Li[4] · Tao Chen[2] · Andrew Routh[1,5] · Zheng Xia[3] · Wei Li[4] · Eric J. Wagner[1,5] · Ping Ji[1]

Received: 17 May 2018 / Revised: 9 December 2018 / Accepted: 18 January 2019

Abstract

Alternative polyadenylation (APA) has emerged as a prevalent feature associated with cancer development and progression. The advantage of APA to tumor progression is to induce oncogenes through 3′-UTR shortening, and to inactivate tumor suppressor genes via the re-routing of microRNA competition. We previously identified the Mammalian Cleavage Factor I-25 (CFIm25) (encoded by *Nudt21* gene*)* as a master APA regulator whose expression levels directly impact the tumorigenicity of glioblastoma (GBM) in vitro and in vivo. Despite its importance, the role of *Nudt21* in GBM development is not known and the genes subject to *Nudt21* APA regulation that contribute to GBM progression have not been identified. Here, we find that *Nudt21* is reduced in low grade glioma (LGG) and all four subtypes of high grade glioma (GBM). Reduced expression of *Nudt21* associates with worse survival in TCGA LGG cohorts and two TCGA GBM cohorts. Moreover, although CFIm25 was initially identified as biochemically associated with both CFIm59 and CFIm68, we observed three CFIm distinct subcomplexes exist and CFIm59 protein level is dependent on *Nudt21* expression in GBM cells, but CFIm68 is not, and that only *CFIm59* predicts prognosis of GBM patients similar to *Nudt21*. Through the use of Poly(A)-Click-Seq to characterize APA, we define the mRNAs subject to 3′-UTR shortening upon *Nudt21* depletion in GBM cells and observed enrichment in genes important in the *Ras* signaling pathway, including *Pak1*. Remarkably, we find that *Pak1* expression is regulated by *Nudt21* through its 3′-UTR APA, and the combination of *Pak1* and *Nudt21* expression generates an even stronger prognostic indicator of GBM survival versus either value used alone. Collectively, our data uncover *Nudt21* and its downstream target *Pak1* as a potential "combination biomarker" for predicting prognosis of GBM patients.

Supplementary information The online version of this article (https://doi.org/10.1038/s41388-019-0714-9) contains supplementary material, which is available to authorized users.

✉ Eric J. Wagner
ejwagner@utmb.edu

✉ Ping Ji
piji@utmb.edu

1 Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX, USA

2 Endoscopy Center, Zhongshan Hospital and Endoscopy Research Institute, Fudan University, Shanghai, China

3 Department of Molecular Microbiology and Immunology, Computational Biology Program, OHSU, Portland, OR 97273, USA

4 Daniel Duncan Cancer Center, Baylor College of Medicine, Houston, TX, USA

5 Sealy Centre for Structural Biology and Molecular Biophysics, University of Texas Medical Branch, Galveston, TX, USA

Introduction

Glioblastoma multiforme (GBM) is the most common primary brain malignant tumor, which is hallmarked by an aggressive tumor behavior [1–3]. Despite new biological insights and advances in treatment, GBM still remains an incurable disease, with an average survival of less than 15 months due to its resistance to therapy [4]. Developing novel strategies to explore early biomarkers to predict prognosis in GBM patients is therefore of utmost importance. The upregulation of oncogenes and downregulation of tumor suppressor genes, similar to other cancer types, play causative roles in GBM formation and progression. Research into the mechanisms related to the regulation of oncogenes and tumor suppresser genes has primarily focused on genomic mutations, copy number variation, and transcription dysregulation. By copy number alteration alone, 66%, 70%, and 59% of the GBM patient samples

Published online: 31 January 2019

contain changes in core components of *RB*, *TP53*, and *RTK* pathways, respectively. Moreover, it has been found that 88% of the GBM samples harbor at least one genetically anomalous event in the core *RTK/RAS/PI(3)K* signaling pathway [5, 6]. In comparison, less is understood about the oncogenic importance of so-called "dark matter" of the genome [7], which includes epigenetic modifications of DNA or histones as well as changes in 3′-UTR length through alternative polyadenylation (APA) [8–10].

The initial evidence implicating APA as important for cell division was attained from the observation that when T-cells transition from a quiescent to proliferative state, there is a global shift to express mRNA with shortened 3′-UTRs [11, 12]. Further, it was reported that cancer cell lines are significantly enriched in mRNA containing shortened 3′-UTRs relative to non-transformed cells [13]. These provocative observations generated a model where cancer cells can selectively shorten the 3′-UTR of oncogenes through the process of APA in order to evade the repressive effects of microRNA and RNA binding proteins [14–17]. As about half of human genes have more than one polyadenylation site [18–20], APA is potentially a general and pervasive form of post-transcriptional regulation, which can activate oncogenes in *cis* by 3′-UTR shortening in different cancer types, including GBM [13, 21, 22]. Complicating this relatively straightforward model are observations from multiple groups demonstrating that 3′-UTR shortening does not obligatorily result in increased protein levels [11] and that tumors, as opposed to cell lines, are not always prone to shorten 3′-UTRs with a bias toward oncogenes [23]. This has led to an alternative mechanism where 3′-UTR shortening observed in tumors occurs in order to redirect microRNA to more efficiently downregulate tumor suppressor mRNA. Indeed, we have recently demonstrated that the tumor suppressor *PTEN* is downregulated in breast tumors by microRNA re-routing after 3′-UTR shortening of transcripts that share microRNA target sites [24]. This *trans* model, which is predicted to disrupt competing endogenous RNA (ceRNA) networks, may exist in addition to the upregulation of oncogenes and underscores the complexity of interplay between mRNA and microRNA that exists in tumors.

The process of cleavage and polyadenylation relies on a molecular machinery that contains four distinct protein subcomplexes: the cleavage and polyadenylation specificity factor (CPSF), cleavage stimulation factor (CstF), mammalian cleavage factor I (CFIm) and cleavage factor II (CFIIm) [8, 25]. Among these complexes, members of the CFIm complex play an important regulatory role in poly(A) site (PAS) selection [26, 27]. The CFIm complex contains three peptides: CFIm25 (also known as CPSF5 or NUDT21), CFIm59 (also known as CPSF7), and CFIm68 (also known as CPSF6) [28, 29]. While initially purified as a complex [28], structural studies demonstrated that CFIm likely exists as two distinct heterotetramers where two CFIm25 subunits associate as a dimer that then further binds with either two additional CFIm59 or CFIm68 subunits [30]. It is not clear what the relative abundance of these heterotetramers are in cells nor if their existence contributes differentially to APA regulation. The specificity of either tetrameric complex is derived from the RNA binding domain of CFIm25, which has high affinity for UGUA sequences [25]. The RNA binding domains of CFIm59 or CFIm68 are thought to be non-specific and may only serve to increase the overall binding strength of the heterotetramer to RNA. Recently, it has been shown that CFIm25 stimulates the use of PASs that are enriched in UGUA elements, which are most prevalent at distal PASs (dPAS) [31]. This model is supported by previous data showing that knockdown of *Nudt21* results in the use of proximal PASs (pPAS) resulting in the global shortening of 3′-UTRs [9, 19, 32]. In our previous work, we demonstrated that depletion of *Nudt21* not only causes increased pPAS usage but also increases cell proliferation and enhances GBM cell tumorigenicity [9, 33]. However, the importance of *Nudt21* to human GBM, the prognostic value of *Nudt21*-regulated APA targets to GBM, and the role of the other two CFIm subunits to GBM remain unclear.

In this study, we show that *Nudt21* expression is reduced in TCGA LGG and GBM patients and is associated with worse survival in TCGA LGG grade II cohort and two-independent TCGA GBM (grade IV) cohorts. Among the three CFIm subunits, CFIm25 plays a critical role in CFIm complex formation and CFIm59 stability in GBM cells. Using Poly(A)-Click-Seq, we conducted a genome-wide APA analysis in GBM cells with *Nudt21* knockdown (KD) in order to identify novel downstream target genes of *Nudt21* as potential biomarkers for predicting prognosis of GBM patients. These efforts uncovered *Pak1*, a p21- activated kinase, as one of target mRNAs whose APA is governed by *Nudt21* that plays a significant role in promoting tumorigenicity in *Nudt21* depleted cells. Using this information, we find that the combination of *Nudt21* and *Pak1* expression is a strong prognostic indicator of GBM patient survival. Collectively, these results extend our understanding of *Nudt21* importance in GBM and the role that APA plays in GBM tumor progression.

Results

Reduced expression of *Nudt21* predicts worse survival in low grade glioma and GBM patients

In our previous study, we showed that *Nudt21* is a master regulator of 3′-UTR alternative polyadenylation, and

functions as a tumor suppressor in GBM [9]. To gain insight into the importance of *Nudt21* in glioma development and progression in humans, we took advantage of the comprehensive TCGA LGG and GBM datasets and compared *Nudt21* expression in different grades of glioma and found *Nudt21* expression significantly reduced in grade III and grade IV compare to grade II glioma (Fig. 1a). We also performed overall patient survival analysis to examine the prognostic role of *Nudt21*. Kaplan–Meier plot analysis showed that reduced expression of *Nudt21* was associated with worse overall survival in the grade II cohort (Fig. 1b). We also compared *Nudt21* expression in TCGA GBM cohort from Project Betastasis and found *Nudt21* expression significantly reduced in all four subtypes of GBM patients compared to normal control samples (Fig. 1c). Kaplan–Meier plot analysis showed that reduced expression of *Nudt21* was associated with worse overall survival in the TCGA GBM patient cohort (log-rank test $p = 0.015$, $n = 348$) (Fig. 1d). To further confirm the prognostic value of *Nudt21*, we obtained the *Nudt21* expression data from GDC TCGA GBM patients based on RNA-seq data, and then stratified the patients into two groups based on top and bottom 30% of *Nudt21* gene expression. Similar to the exon array data in Betastasis, Kaplan–Meier plot analysis showed that reduced *Nudt21* expression is associated with worse overall survival (log-rank test $p = 0.0073$, $n = 173$) (Fig. 1e). It is noted that expression of *CFIm59* is also significantly decreased in all four subtypes of GBM patient samples and associated with worse survival in the TCGA GBM patient cohort using Project Betastasis online representation tool (Fig. S1A and S1B, $p = 0.026$, $n = 348$). However, the expression of *CFIm68* is significantly increased in the proneural subtype of GBM patients (Fig. S1C) and expression of *CFIm68* had no correlation with overall survival (Fig. S1D, $p = 0.29$, $n = 348$). These results agree with our previous work that reduction in *Nudt21* increases tumorigenicity and suggest that CFIm25 and CFIm59 but not CFIm68 may be an important component of its APA regulation in GBM.

CFIm25 is a key component for CFIm complex formation and modulates CFIm59 protein stability, but not CFIm68

The three CFIm subunits, CFIm25/CFIm59/CFIm68 were originally isolated from HeLa nuclear extract as a fraction required for pre-mRNA 3′-end processing with the implication that these three subunits interacted as a complex [28]. However, subsequent structural studies indicated that the CFIm complex likely exists as two heterotetramers both containing two copies of CFIm25 with one containing two copies of CFIm59 or the other containing two copies of CFIm68 [19, 30]. Given our observation that both CFIm25 and CFIm59 possess prognostic value in GBM but not CFIm68, we next asked the question if we could detect distinct CFIm heterotetramers in GBM cells. To achieve this, we subjected two GBM cell lysates to immunoprecipitation with an individual antibody against CFIm25, CFIm59, or CFIm68. We tested both the immunoprecipitates and the depleted supernatants for the presence of each of the subunits. Consistent with structural studies, we found that CFIm25 pulls down both CFIm59 and CFIm68 equally in both cells and that depletion of CFIm25 in cell lysates removed almost all CFIm59 and CFIm68 (Fig. 2a, lane 1 vs. 2 and 5 vs. 6). This result further supports the model that CFIm25 is common to both CFIm subunits. In contrast, immunoprecipitation of CFIm59 completely depleted its levels from supernatants, but reduced the level of CFIm25 by ~50% (Fig. 2a, lane 1 vs. 3 and 5 vs. 7). The levels of CFIm68 detected in the CFIm59-depleted supernatants were marginally reduced relative to control-treated cells suggesting that CFIm59 and CFIm68 do not interact significantly in GBM cells. Analysis of GBM lysates immunoprecipitated with CFIm68 antibodies yielded a similar result in that levels of CFIm68 were undetected while CFIm25 was depleted ~50% and CFIm59 levels were only marginally effected (Fig. 2a, lane 1 vs. lane 4 and lane 5 vs. 8). Altogether, these results indicate that CFIm25 likely forms complexes with CFIm59 and/ or CFIm68 in GBM cells and those three complexes (CFIm25/CFIm59/CFIm68, CFIm25/CFIm59, and CFIm25/CFIm68) are co-existing in GBM cells (Fig. 2a).

Given the critical role of CFIm25 to forming those three CFIm subcomplexes, we next asked whether CFIm25 could interact with either subunit when tested in isolation. To achieve this, we performed pairwise two yeast hybrid assays using each subunit of the CFIm complex fused to the Gal4-activation domain tested against *Nudt21* fused to the Gal4 DNA binding domain. Importantly, *Saccharomyces cerevisiae* lack CFIm complex subunits in their genome reducing the potential for bridging interactions with endogenous proteins. We observed that *Nudt21* could support growth on selective media when paired with itself or either of the *CFIm* subunits, but not the empty vector suggesting that direct interaction of CFIm25 is responsible for heterotetramer formation (Fig. 2b). Finally, given the existence of three distinct CFIm complexes in GBM cells, we asked whether these subunits exhibited interdependent protein stability. To that end, we knocked down each one of CFIm subunits with siRNA in both LN229 and U251 cells and determined protein level of each CFIm subunit by western blotting. We found that downregulation of CFIm25 reduced CFIm59 protein level in both cell lines, but did not affect CFIm68 protein level. Also, depletion of CFIm59 and CFIm68 did not affect the expression of any other non-targeted subunit (Fig. 2c). These data suggest that CFIm25

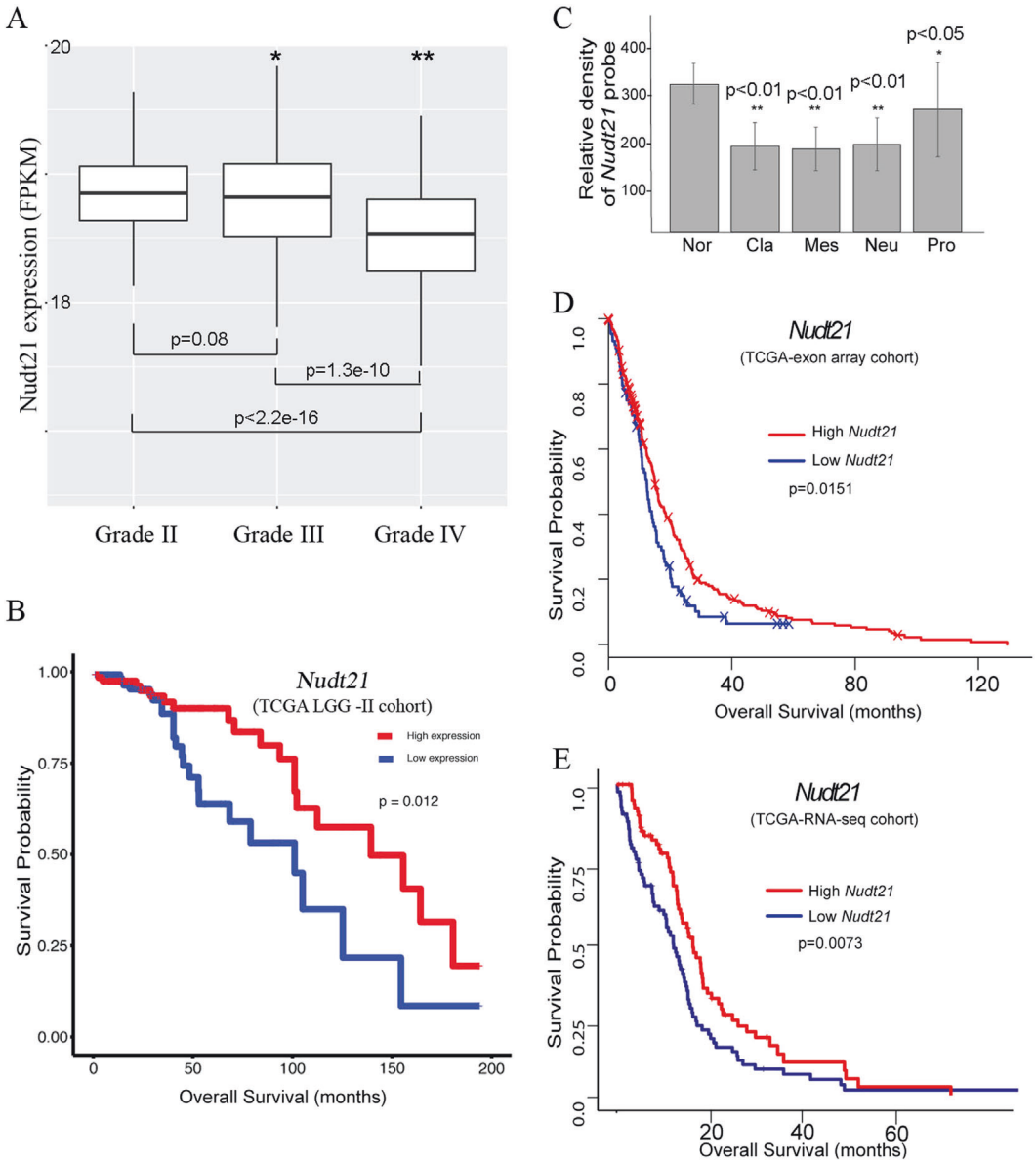

Fig. 1 Reduced expression of *Nudt21* in TCGA LGG and GBM cohorts and *Nudt21* correlation with reduced survival in TCGA LGG cohort and two TCGA GBM patient cohorts. **a** Reduced expression of *Nudt21* in TCGA low grade glioma patients and TCGA GBM patients based on RNAseq data. $P = 0.08$ grade II ($n = 259$) versus grade III ($n = 266$); $p < 2.2e{-}16$ grade II versus grade IV ($n = 173$) and $p = 1.3e{-}10$ grade III versus grade IV. **b** Kaplan–meier plot analysis showing that reduced expression of *Nudt21* is associated with worse survival in the TCGA LGG grade II patients ($n = 259$, $p = 0.012$). **c** Downregulation of *Nudt21* in four subtypes of GBM compared to normal brain samples, based on exon array data with Affymetrix human exon 1.0 ST platform. (Nor: Normal control ($n = 11$); Cla: Classical ($n = 52$); Mes: Mesenchymal ($n = 56$); Neu: Neural ($n = 31$); Pro: Proneural ($n = 54$), $*p < 0.05$, $**p < 0.01$). **d**, **e** Kaplan–meier plot analysis showing that reduced expression of *Nudt21* is associated with worse survival in the TCGA GBM exon array cohort ($p = 0.015$, $n = 348$), and GDC TCGA GBM RNAseq cohort ($p = 0.0073$, $n = 173$)

is a key component for CFIm complex formation and is required to maintain CFIm59 stability in GBM cells.

Nudt21 regulates broad 3′-UTR APA genes enriched in *Ras* signaling pathway

Given our previously report that downregulation of *Nudt21* increases tumorigenicity of GBM cells [9] and the results presented here, we next wanted to define all of the APA targets of *Nudt21* in GBM cells. Therefore, we first transfected either *Nudt21* targeting siRNA or control siRNA into LN229 cells and assessed knockdown using western blotting. In each biological replicate we observed specific reduction in the level of CFIm25 protein (Fig. 3a). From each of the replicates, we isolated total RNA and subjected it to Poly(A)-Click-Seq (PAC-seq) analysis. The PAC-seq

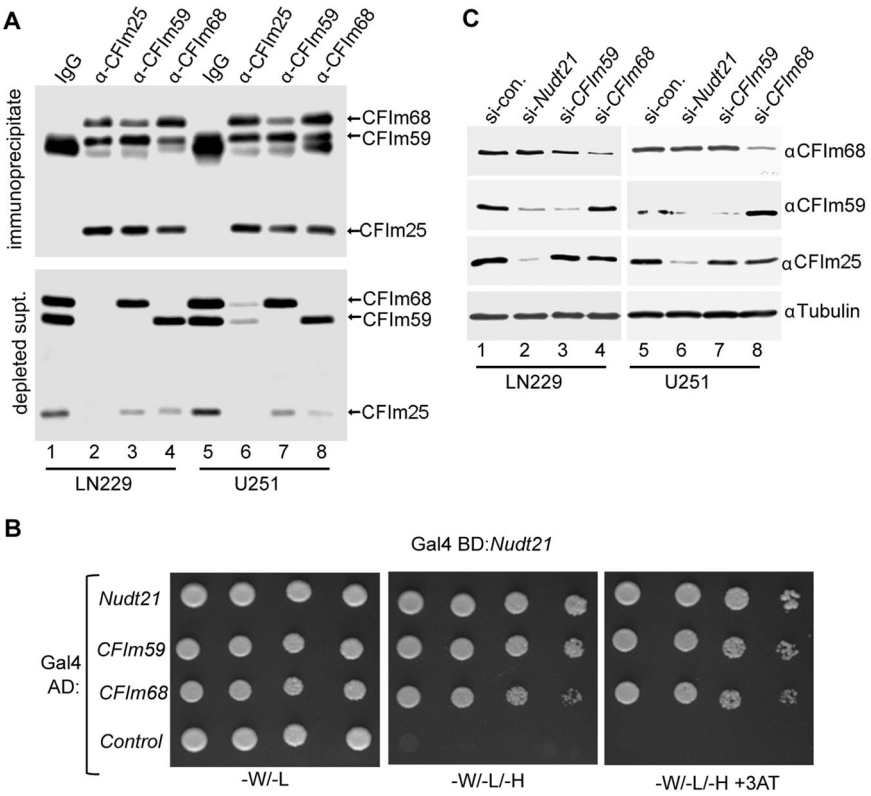

Fig. 2 Three CFIm subunits form three distinct subcomplexes and CFIm25 stabilizes CFIm59 but not CFIm68 in GBM cell lines. **a** Co-immunoprecipitation (Co-IP) of CFIm complex members with an individual anti-CFIm subunit antibody. Lysates from LN229 and U251 cell lines were incubated with protein A/G agarose beads conjugated with anti-CFIm25, or anti-CFIm59, or anti-CFIm68 antibody. Anti-IgG serves as a control. Co-IP pull-down complex and the depletion supernatant were subjected to analysis using a 10% SDS-PAGE and the composition of CFIm subunits were detected by western blotting with indicated antibodies. A representative image is shown from three-independent experiments, CFIm subunits in Co-IP complex (upper panel) and in depletion supernatant (lower panel). **b** Interaction of CFIm subunits was determined by two yeast hybrid screen. Data show CFim25 interacts with itself and the two other individual CFIm subunits and empty vector serves as a negative control. **c** *Nudt21* knockdown decreases CFIm59 expression, but has no effect on CFIm68 expression in GBM cell lines. A representative data are shown from three-independent experiments

approach was specifically developed to sequence the 3′-end regions of poly-adenylated RNAs genome-wide to map APA events [32]. This is accomplished in part by spiking in azido-dVTPs (azido-dNTPs lacking azido-dTTP to prevent termination within poly(A) tails) into the reverse transcription reaction to generate azido-blocked cDNAs and then click- ligating a 5' alkyne-modified adaptor using copper-catalyzed cycloaddition [34]. From all six RNA isolates, PAC-seq libraries were efficiently and equally amplified (Fig. 3b), then subjected to gel purification and sequenced using the Illumina platform.

Among all the detected mRNA 3′-UTRs, we found 18% mRNAs ($n = 1226$) that had two or more poly(A) sites and could therefore undergo APA. Over half of those mRNA (10% of the total mRNA population ($n = 695$)) were found to exhibit significant (>20% change) APA upon *Nudt21* KD (Fig. 3c). When we filtered those mRNA that had more than 1 poly(A) site to include only those that showed >twofold change in poly(A) site usage and a change of at least 20% in the percent distal poly(A) site usage (% dPAS), 93.6% exhibited 3′-UTR shortening ($n = 350$) while only 6.4% exhibited 3′-UTR lengthening ($n = 24$) (Fig. 3d). To evaluate the quality of our PAC-seq data and the specificity of *Nudt21*-regulated APA, we created gene tracks and visualized the results using the UCSC genome browser. We first inspected PAC-seq data for *Vma21*, which is a previously established target of *Nudt21*-regulated APA, and found that the *Vma21* 3′-UTR was shortened in all three *Nudt21* KD cells compared to the three control cells (Fig. 3e). Importantly, by comparing PAC-seq results with defined poly(A) sites determined by both the polyA database (PDB) and poly(A) seq analysis we found complete agreement in the placement of the two alternative poly(A) sites. Moreover, we inspected the relative poly(A) site usage for the *Mxra7* gene, which contains a complex pattern of PASs and two distinct 3′-UTR regions. We observed that PAC-seq could accurately identify 5/6 annotated PASs and more importantly, we observed that *Nudt21* KD does not change the

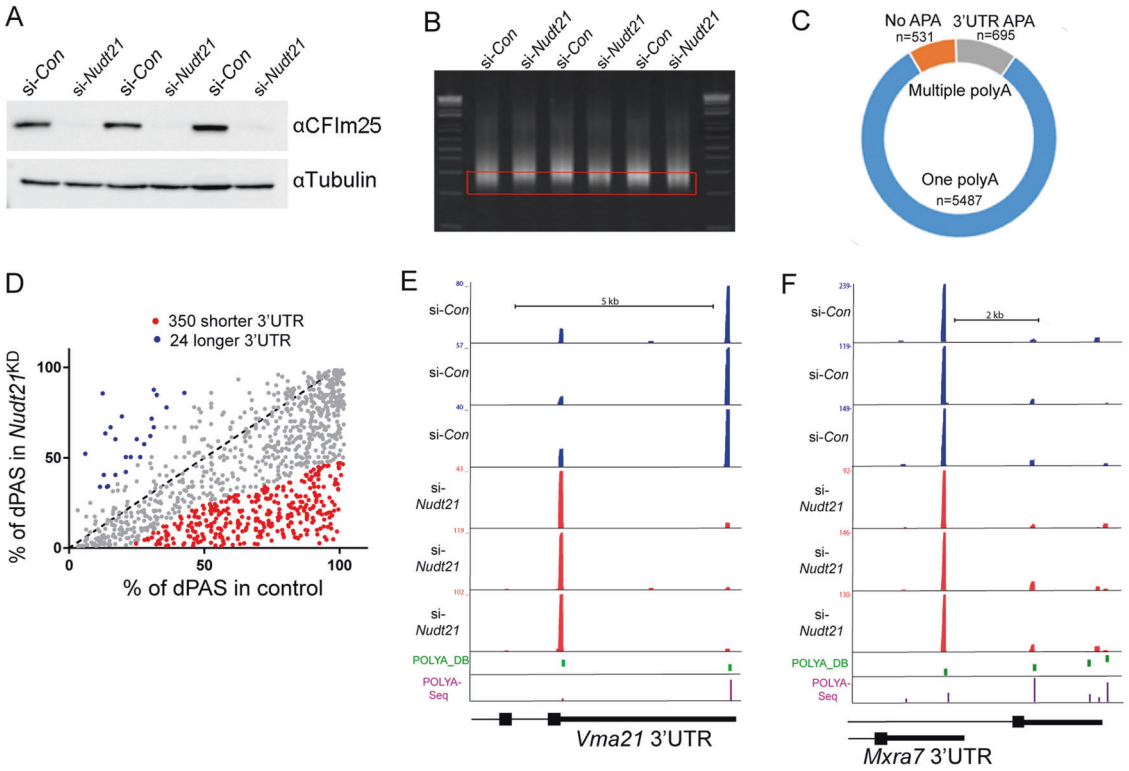

Fig. 3 Identification of key signaling pathways whose APA is regulated by *Nudt21* in GBM cells. **a** Knockdown of *Nudt21* in LN229 was validated by western blotting for three-independent biological replicates. **b** Preparation of PAC-seq library by poly azido-nucleotides. The pooled PAC-seq libraries from 200 bp to 300 bp were isolated from a 2% agarose gene and sequenced using Miseq. **c** Detection of 3′-UTR APA events in *Nudt21* KD cells. The total six bar coded libraries were loaded on a flowcell (illumine V3 kit) and analyzed with a Miseq and total sequence reads were trimmed and aligned to USCS human genome browser. 82% genes ($n = 5487$) with one poly(A) site; 8% genes ($n = 531$) with two more poly(A) sites no APA and 10% genes ($n = 695$) with 3′-UTR APA. **d** Identification of 3′-UTR shortening genes. Filtering of mRNA that had more than 1 poly(A) site that showed a greater than twofold change in site usage and a change of at least 20% in the percent distal poly(A) site usage (% dPAS) upon *Nudt21* KD

overall usage of those PASs relative to control (Fig. 3f). Taken together, these results indicate that the PAC-seq data generated from knockdown is highly congruent with previously defined PAS databases and that the *Nudt21* KD does not cause a "general" phenotype of 3′-UTR shortening but rather that there is a specific subset of genes regulated by *Nudt21*. By using KEGG map-signaling pathway analysis program, we found that enrichment of 3′-UTR shortening genes are related to several tumor-driving signaling pathways, including *Ras* signaling pathway with 12 3′-UTR shortening genes identified (Table 1).

The Pak1/2 3′-UTR undergoes shortening in response to *Nudt21* inactivation

We decided to further investigate the 12 members of the Ras signaling pathway because of their high enrichment among genes whose APA is regulated by *Nudt21* and because of their established importance in GBM [5]. All 12 of these genes exhibited 3′-UTR shortening to various degrees in *Nudt21* KD cells. *Pak2* was the top candidate gene with 83% 3′-UTR shortening, which was followed by *Gng12* with 80% shortening, *Ets1* with 57% shortening, and *Pak1* with 56% shortening (Table 1). *Pak1* and *Pak2* are isoforms encoded by distinct genes and produce the p21-activated kinase, which is a key node in oncogenic signaling pathways that controls cell growth, survival and motility of cancer cells [35]. The expression and activity of *Pak1* is increased in many human cancers and is associated with poor prognosis [36, 37]. Therefore, we focused on *Pak1/2* as the target genes of *Nudt21* for further study.

The genome coverage tracks of PAC-seq loaded into the UCSC genome browser were used to visualize the 3′-UTR of both *Pak1* and *Pak2* (Fig. 4a, b). In the case of *Pak1*, we observed three poly(A) sites in the polyA_DB [38] and all three could be observed as utilized in LN229 PAC-seq data. Importantly, *Nudt21* KD resulted in a significant shift from the preferred usage of the most distal PAS in control-treated cells towards the most proximal PAS. In the case of *Pak2*, a similar result was also observed except that the shift occurred in the middle of the 3′-UTR upon *Nudt21* KD and most miRNA binding sites were remained.

To validate these results, we utilized the RNase H alternative polyadenylation assay (RHAPA), which is a

Table 1 Enrichment of 3′-UTR shortening genes in *Ras* signaling pathway

	Gene	Poly(A)	% of 3′-US	Function	Ref
1	*Pak2*	3	83%	*Pak2* is a p21-activated kinase, a close homolog of *Pak1*. Key nodes in oncogenic signaling pathways controlling growth, survival and motility of cancer cells.	[50]
2	*Gng12*	3	80%	Guanine nucleotide-binding proteins (G proteins) are involved as a modulator or transducer in various transmembrane signaling systems.	[51]
3	*Ets1*	2	57%	Transcription factor. Directly controls the expression of cytokine and chemokine genes in a wide variety of different cellular contexts.	[52]
4	*Pak1*	2	56%	*Pak1* is involved in a variety of pathways in biological processes, such as proliferation. Its activity increases in many cancers and is associated with poor prognosis.	[37]
5	*NRas*	3	56%	Ras proteins comprise a fimily of low-molecular weight GTPases, function as a conduit for signals received from receptor tyrosine kinase on the cell surface.	[53]
6	*Gnb4*	3	55%	Belong to Guanine nucleotide-binding proteins (G proteins) and is a key player in the signal transduction from membrane receptors to intracellular effectors.	[54]
7	*Tgfa*	3	54%	*TGF*α is a mitogenic polypeptide that is able to bind to bind to the EGFR and to act synergistically with TGFβ to promote anchorage-independent cell proliferation in soft agar.	[55]
8	*Gnb1*	3	35%	Belong to Guanine nucleotide-binding proteins (G proteins) and is involved as a modulator or transducer in various transmembrane signaling systems.	[56]
9	*Rac1*	2	32%	A small G protein regulates a wide variety of processes in the cell, including growth, differentiation, movement and lipid vesicle transport.	[57]
10	*Gng2*	2	31%	This gene encodes one of the gamma subunits of a guanine nucleotide-binding protein, involved in signaling mechanisms across membrane.	[58]
11	*Grb2*	2	27%	Adapter protein that provides a critical link between cell surface growth factor receptors and the Ras signaling pathway.	[59]
12	*Epha2*	2	19%	Membrane-bound ephrin-A family ligand 2 residing on adjacent cells, leading to contact-dependent bidirectional signaling into neighboring cells.	[60]

qRT-PCR-based approach specifically designed to quantify APA events [39]. In this approach, total RNA is incubated with a DNA primer complementary to a region within the 3′-UTR in between the two differentially utilized PASs (Fig. 4c, d). Then, RNase H is added in order to cleave the RNA present within the DNA/RNA hybrid thereby separating the distal and proximal region of the 3′-UTR. Following oligo-dT primed reverse transcription, primers specific to either the distal or proximal region are used to quantify the relative use of the dPAS versus the pPAS. Using RHAPA, we observed that after treatment of either LN229 or U251 cells with *Nudt21* siRNA, there was a clear and preferred use of proximal poly(A) sites compared to distal poly(A) site for both *Pak1* and *Pak2* (Fig. 4e, f). These results confirmed that *Pak1* and *Pak2* 3′-UTR APA are regulated by *Nudt21* in GBM cells.

Depletion of *Nudt21* induced *Pak1* expression and promotes cell proliferation and migration

Next, we evaluated the consequence of *Pak1/2* 3′-UTR APA regulated by *Nudt21* in terms of both mRNA and protein expression. First, we determined the expression of *Pak1/2* mRNA by quantifying the read coverage in PAC-seq in both control and *Nudt21* KD cells. *Pak1* mRNA was slightly increased in *Nudt21* KD cells ($n = 3$, $p = 0.37$, *t*-test) whereas *Pak2* mRNA was not changed ($n = 3$, $p = 0.97$, *t*-test) (Fig. S2). Second, we measured the protein level of PAK1/2 in both control and *Nudt21* KD cells by western blot. We observed that PAK2 is barely detectable in both LN229 and U251 cells and did not increase appreciably upon *Nudt21* KD. We reasoned that the low level of expression and the lack of a translational increase upon *Nudt21* KD is because *Pak2* is subject to 3′-UTR repression that is not alleviated after 3′-UTR shortening induced by *Nudt21* KD. Given that the APA of *Pak2* only results in removal of ~50% of its 3′-UTR, it is likely that inhibition by microRNA is not attenuated. In contrast, we observed that PAK1 protein levels present in LN229 cells were significantly increased upon *Nudt21* KD (Fig. 5a), which is consistent with its increase in mRNA level and the fact that APA in response to *Nudt21* KD shortens nearly all of the *Pak1* 3′-UTR. Finally, to demonstrate that the increase in *Pak1* mRNA and protein is mediated by its 3′-UTR, we created a heterologous reporter containing a Renilla luciferase open reading frame followed by the *Pak1* 3′-UTR. This reporter, along with a control Firefly luciferase reporter was transfected into either control or *Nudt21* siRNA transfected cells and then luciferase expression as measured. In *Nudt21* KD cells, the levels of luciferase expression was

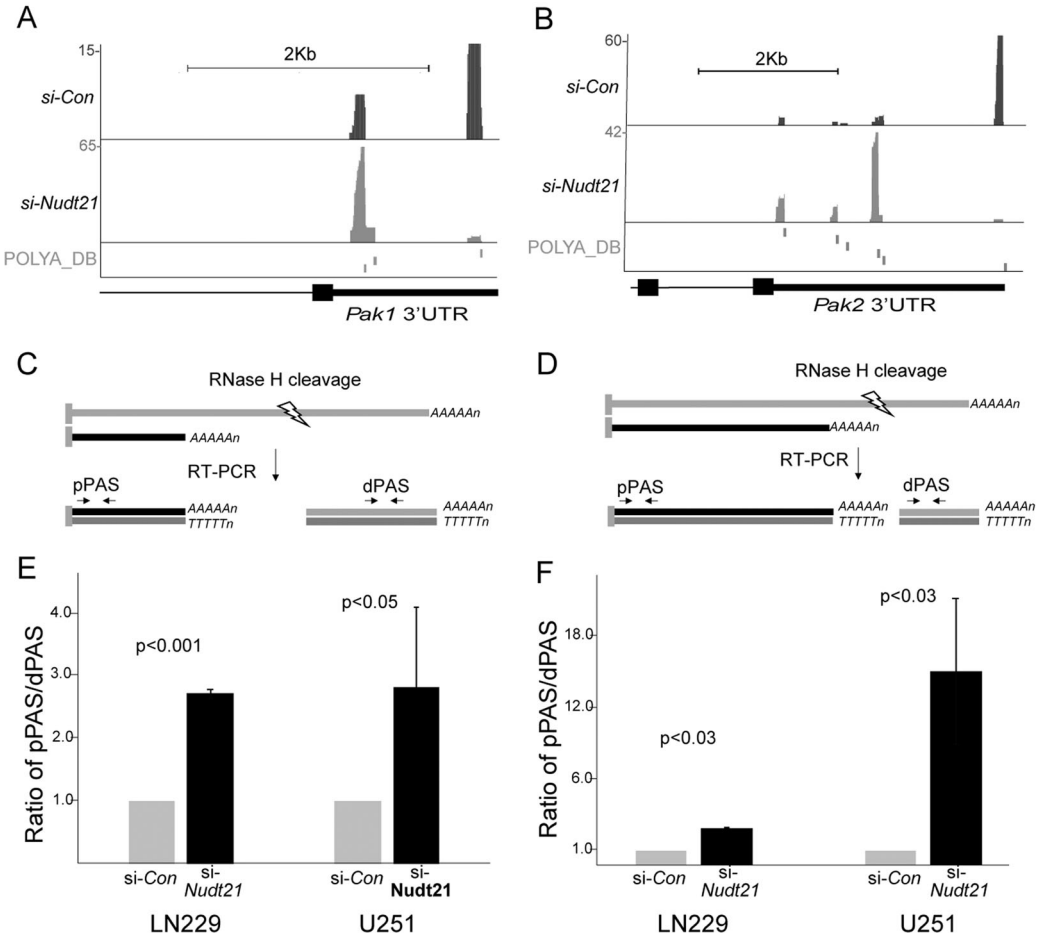

Fig. 4 3′-UTR shortening of *Pak1* and *Pak2* regulated by *Nudt21*. **a** Two poly(A) sites of *Pak1* were detected in 3′-UTR and PAC-seq reads density in proximal poly(A) site (pPAS) was increased in *Nudt21* KD cells compare to control cells. A representative 3′-UTR APA profile of *Pak1* was shown. **b** Three poly(A) sites were found in 3′-UTR of *Pak2* and the sequence reads in distal poly(A) site (dPAS) was significantly decreased in *Nudt21* KD cells compare to control cells. A representative 3′-UTR APA profile of *Pak2* is shown. Numbers on *y*-axis indicated PAC-Seq read coverage. **c** Diagram of RNase H dependent APA (RHAPA) for validation of *Pak1* 3′-UTR APA. **d** Usage of 3′-UTR PPAS of *Pak1* was increased in both in LN229 and U251 with *Nudt21* KD. Data from real time PCR are shown as fold change of usage in pPAS/dPAS. Data are shown as mean ± standard deviation. ($p < 0.001$ in LN229 and $p = 0.041$ in U251, two side *T*-test). **e** Diagram of RHAPA for validation of *Pak2* 3′-UTR APA. **f** Usage of pPAS of *Pak2* was increased in both in LN229 and U251 with *Nudt21* KD. Data are shown as fold change of usage in pPAS/dPAS. Experiments were performed in triplicate with the data shown as mean ± standard deviation. ($p = 0.022$ in LN229 and $p = 0.025$ in U251, two side *T*-test)

increased indicating that the *Pak1* 3′-UTR is sufficient to mediate upregulation of a heterologous mRNA. Altogether, these results show that *Pak1* is a bona fide target of *Nudt21*-regulated APA and that its mRNA and protein levels increase upon shortening of its 3′-UTR. Further, given the lack of any increase in PAK2 levels in response to *Nudt21* KD, we chose to focus our analysis on *Pak1*.

To explore the consequence of *Pak1* upregulation regulated by *Nudt21* in GBM, we focused on two oncogenic properties including anchorage-independent colony formation and cell migration. Using soft agar assay, we found that depletion of *Nudt21* significantly promoted LN229 and U251 anchorage-independent growth ability, as shown by increased colony numbers, while depletion of *Pak1* reduces colony formation. Interestingly, co-depletion of both *Nudt21* and *Pak1* eliminated the benefit of *Nudt21* on anchorage-independent growth suggesting that *Pak1* is indeed an important APA target of *Nudt21* for this process (Fig. 5c, d). Similarly, we observed increased cell migration in LN229 and U251 cells when *Nudt21* is knocked down whereas depletion of *Pak1* reduces the number of migration cells. Similar to the anchorage-independent growth assays, co-depletion of both *Nudt21* and *Pak1* reduced the increased cell migration observed in *Nudt21* depleted cells to a level similar to control-treated cells (Fig. 5e, f). Our results suggest that *Pak1* is an important 3′-UTR APA target regulated by *Nudt21* and its expression is vital for *Nudt21*-dependent increases in cell migration and anchorage-independent colony formation.

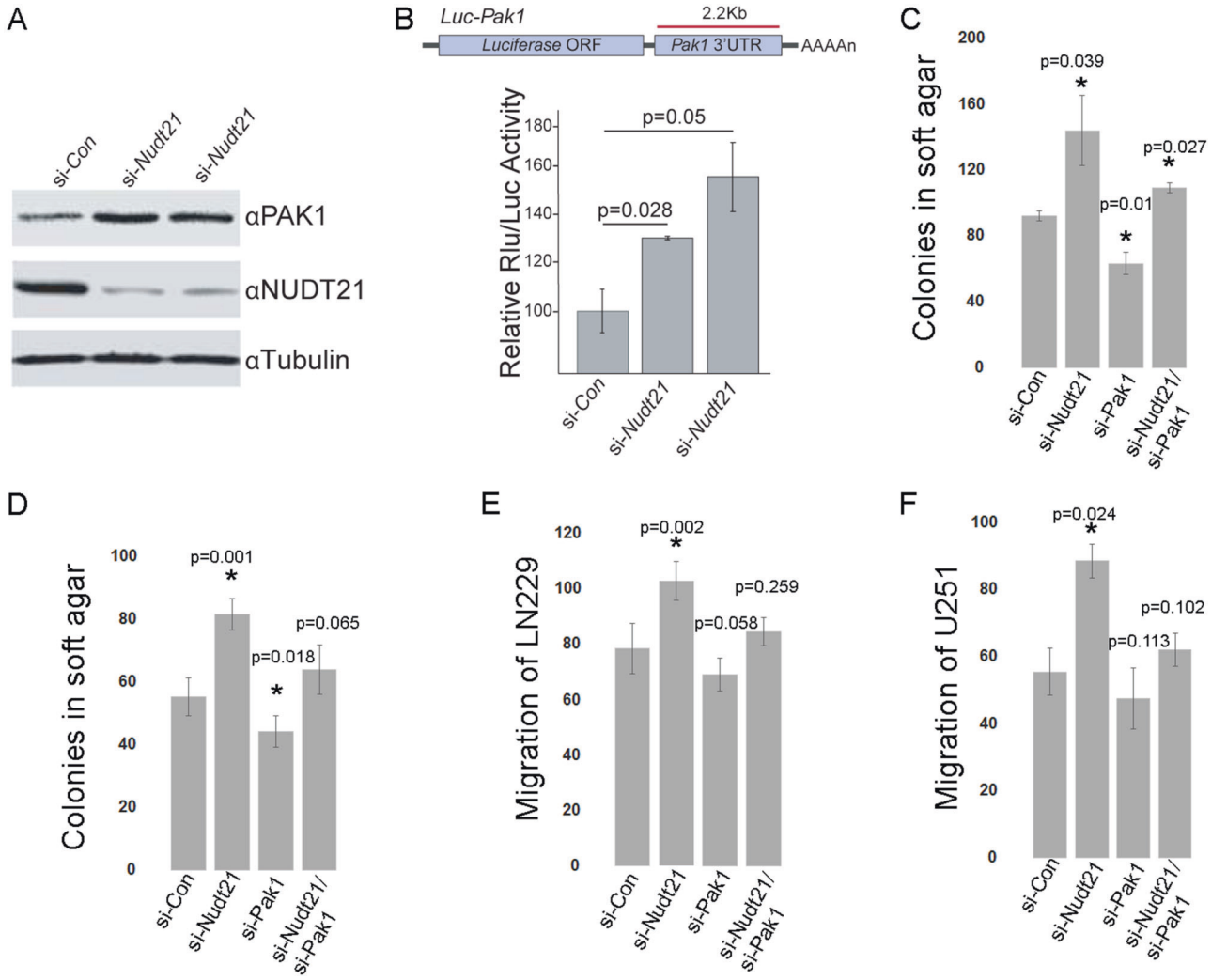

Fig. 5 Upregulation of *Pak1* regulated by *Nudt21* and promoting GBM cell proliferation. **a** *Pak1* expression increased in LN229 with *Nudt21* KD. Cell lysate from siRNA-control and siRNA-*Nudt21* was separated in 10% SDS-PAGE and protein expression was detected by western blot with indicated antibodies. A representative profile of *Pak1* was shown from three-independent experiments. **b** *Pak1* 3′-UTR-mediated luciferase activity was increased in LN229 cells with *Nudt21* KD. Data are shown as average luciferase activity ± standard deviation from three-independent experiments in triplicate ($p = 0.028$ and $p = 0.05$ two side *T*-test). **c**, **d** Effects of knockdown of *Nudt21* and/or *Pak1* on colony formation in LN229 cells and U251 cells. The ability of colony formation was evaluated by soft agar assays. Comparison of the difference of this ability between groups of negative control (NC), knockdown of *Nudt21*, knockdown of *Pak1* or combination of knockdown of *Nudt21* and *Pak1* was performed. $*p < 0.05$ was assumed significant difference between the siRNA knockdown group cells versus control siRNA transfected cells. **e**, **f** Effects of knockdown of *Nudt21* and /or *Pak1* on cell migration in LN229 and U251 cells. LN229 and U251 cells were transfected with siRNA for *Nudt21* and/or *Pak1* for 24 h and the transfected cells were seeded into the transwell inserts of a CytoSelect 24-well cell migration plate. Qualification data showed the mean of migrating cells ± standard deviations in different groups of NC, knockdown of *Nudt21*, knockdown of *Pak1* or combination of knockdown of *Nudt21* and *Pak1*. $*p < 0.05$ was assumed significant difference between the knockdown group cells versus control siRNA transfected cells

Pak1 expression provides an additional biomarker in predicting prognosis of GBM patients

Given that *Pak1* APA regulation by *Nudt21* is important for the increased tumorigenicity of GBM cells in response to *Nudt21* downregulation, we decided to evaluate whether expression of *Pak1* serves as a biomarker for prognosis of GBM patients. In the TCGA exon array GBM cohort, we found that high expression of *Pak1* was correlated with poor overall survival among GBM patients ($n = 348$, $p = 0.0233$) (Fig. 6a), while high expression of *Pak2* didn't show significant correlation with overall survival ($n = 348$, $P = 0.7554$) (Fig S3). Interestingly, combination of high level of *Pak1* expression and low level of *Nudt21* expression has greater predictive power than stratifying for *Pak1* expression or *Nudt21* expression alone ($p = 0.0001$, $n = 348$) (Fig. 6b). To further confirm the prognostic value of *Pak1* expression, we evaluated the association of clinical

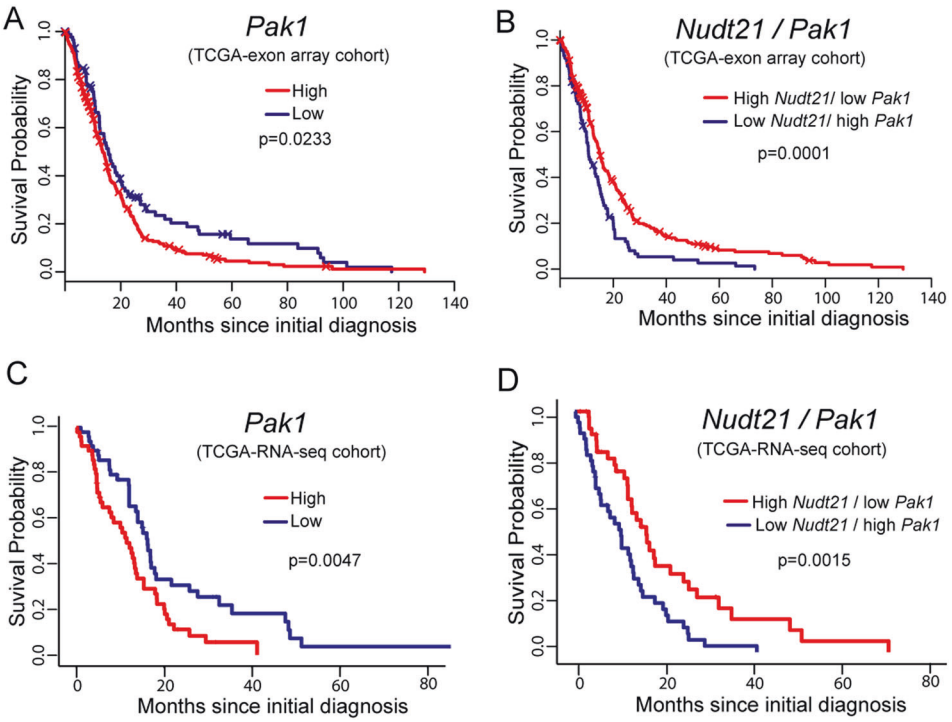

Fig. 6 Induced expression of *Pak1* predicts worse survival in two TCGA GBM cohorts. **a**, **c** Kaplan–meier plots analyses indicated that increased expression of *Pak1* predicts worse survival in the TCGA GBM exon array cohort ($p = 0.023$, $n = 348$), and GDC TCGA GBM RNAseq cohort ($p = 0.0047$, $n = 173$). **b**, **d** Kaplan–meier plot analyses showed that combination of high level of *Pak1* expression and low level of *Nudt21* expression stratified prognosis of GBM patients in TCGA GBM exon array cohort ($p = 0.00017$, $n = 348$) and GDC TCGA GBM RNAseq cohort ($p = 0.0015$, $n = 173$)

survival with *Pak1* expression data on a second and independent dataset, the GDC TCGA GBM patients based on RNA-seq data. Consistently, we found that increased *Pak1* expression was associated with worse overall survival in the GBM patient cohort ($p = 0.0047$, $n = 173$) (Fig. 6c). Furthermore, combination of high level of *Pak1* expression and low level of *Nudt21* expression also demonstrated more power to predict prognosis of GBM patients ($p = 0.0015$, $n = 173$) (Fig. 6d). These data suggest that both *Pak1* and its upstream regulator, *Nudt21*, serve as a biomarker for predicting prognosis of GBM patients and combination of *Nudt21* and *Pak1* expression improves its prediction value in prognosis of GBM patients.

Discussion

Despite advances in GBM diagnosis and treatments, the overall outcome of patients with GBM remains poor due to the extremely aggressive and highly infiltrative nature of GBM [2, 40]. To date, our understanding of GBM tumorigenesis has focused largely on genomic mutations, copy number alteration, and transcription dysregulation, however, post-transcriptional and translational mechanisms of GBM progression remain underappreciated. Our results suggest that post-transcriptional regulation plays a critical role in regulating APA events and, in turn gene expression related to cancer signaling pathways. Further, our results demonstrate the importance of APA regulators and their targets to the prognosis of GBM patients. Specifically, we have identified that *Nudt21* regulates the APA of a broad spectrum of mRNA in GBM, with target genes enriched in the *Ras* signaling pathway. We find the activated oncogenic function of *Pak1* is potentiated by *Nudt21*-regulated 3′-UTR shortening and therefore can contribute to GBM development and progression. Further, our results reveal that both *Nudt21* and *Pak1* may serve as a biomarker for predicting prognosis of GBM patients and imply an important role in GBM development and progression.

Recently, it has become increasingly clear that APA plays a critical role in regulation of gene expression, oncogenic signaling pathway, and tumorigenesis [18]. Our previous work demonstrated that *Nudt21* is a critical APA regulator in GBM that possess properties similar to classic tumor suppressors. Reduction in *Nudt21* expression enhances growth of GBM tumors in mice while its overexpression decreases tumor size [9]. Here, find that *Nudt21* expression in human LGG and GBM patients within two distinct TCGA cohorts also behaves consistently with that of a tumor suppressor in that decreased expression correlates with reduced survival. Moreover, *Nudt21* expression is reduced in LGG grade II and grade III, and all four GBM subtypes relative to normal brain tissue. Collectively, these results suggest that reduction in *Nudt21* is an important component of GBM tumor progression. Recently, a genome-wide CRISPR screen revealed that *Nudt21* expression is a major barrier for creation of iPS cells from fibroblasts [31]. In this context, downregulation of *Nudt21* will induce broad shortening of 3′-UTRs enhancing de-differentiation of cells leading to a stem-like state [31] and

this may represent the reason why GBM tumors preferentially reduce *Nudt21* as they progress towards further de-differentiation.

The initial identification of the CFIm complex showed that it consisted of three proteins: CFIm25 and two closely related proteins, CFIm59 and CFIm68. Depletion of these subunits from nuclear extracts inhibited in vitro cleavage and polyadenylation and the addition of only CFIm25 and CFIm68 is required to reconstitute activity [27]. Structural studies show that CFIm25 homodimerizes and then interacts with two copies of either of the other two subunits [41]. Our results indicate that indeed three distinct CFIm complexes are co-existing in GBM cells. Others have shown that knockdown of *Nudt21* or *CFIm68*, but not *CFIm59*, leads to significant 3′-UTR shortening suggesting that the two heterotetramers are not functionally equivalent [30]. In human GBM patients, *CFIm59* expression correlates with patient survival similar to *Nudt21* but *CFIm68* does not present any correlation. In fact, CFIm68 expression appears to be increased in GBM proneural subtype. The explanation for these observations is not known, but the lack of CFIm68 correlation could derive from reports of CFIm68 functioning in RNA processing events that are independent of its association with CFIm25. For example, CFIm68 is involved in histone pre-mRNA processing and stimulates mRNA export [42], but CFIm25 does not. Reduction of CFIm68 expression in GBM tumors may provide the benefit of broad 3′-UTR shortening but may also reduce the efficiency of histone mRNA biogenesis, which could be deleterious to tumor survival. The precise function of CFIm59 is not yet known but its downregulation in GBM tumors could disrupt the balance of CFIm25 containing tetramers leading to selective 3′-UTR shortening. Further studies will be necessary to fully understand CFIm59 contribution to APA.

Over the past decade, the importance of APA to human physiology has become more evident. APA has been found to play a role in various biological or pathological conditions, such as cell fate, development, cancer, and heart disease [23, 43–45]. This research has spurred innovation to develop new technologies in order both detect and quantify APA in these settings. Initial studies of APA involved the use of microarrays, which have several well-known limitations, such as lack of sensitivity, limited dynamic range, and possible competitive cross-hybridization. This approach was replaced by multiple partitioned RNA-seq techniques custom designed to enrich for sequences near poly(A) sites. This advantage of these approaches is that far less reads are required to identify APA events throughout the genome. In this study, we used a recently developed tool, called PAC-seq [32], to detect and quantify APA events that change in response to *Nudt21* KD in GBM cells that may reflect the targets of *Nudt21* that change APA in response to its downregulation in GBM patients. PAC-seq utilizes click-chemistry to specifically enrich for the junction of the 3′-UTR and poly(A) tail. Using this approach, we demonstrated that *Nudt21* regulated a broad the APA of a broad spectrum of mRNA in GBM cells, especially those enriched in cancer-related signaling pathways, including the *Ras* signaling pathway.

Among the genes whose APA is regulated by *Nudt21*, we identified that *Pak1* and *Pak2*, which are components of the *Ras* signaling pathway and been found to be important for tumor progression, undergo 3′-UTR shortening in response to *Nudt21* downregulation. It has been reported that *Pak1* expression correlates with poor prognosis in solid tumors [46]. Moreover, phosphorylation of PAK1 can promote migration/invasion in GBM and is associated with shorter survival [47]. Here, we find that depletion of *Nudt21* in GBM cells changes the usage of dPAS to pPAS of both *Pak1/2* but only *Pak1* mRNA levels were observed to increase. This is likely because the amount of the 3′-UTR removed upon *Pak2* shortening is relatively small and may not remove enough-negative elements from the 3′-UTR to enhance its stability. The increased expression of both endogenous *Pak1* and the *Pak1* 3′-UTR luciferase reporter after *Nudt21* KD demonstrates that *Nudt21* is a key regulator of *Pak1* levels. Moreover, our observations that co-knockdown of *Nudt21* and *Pak1* abrogates the proliferative and migration benefit imparted by *Nudt21* KD suggests that *Pak1* may be one of the important downstream targets of *Nudt21*. Consistent with this hypothesis, we discovered that high expression of *Pak1* was associated with worse survival of GBM patients based on the data from two-independent TCGA datasets: The data of GBM exon array are from Project Betastasis, which includes cohort from REMBRANDT and TCGA ($n = 349$) and it is quite different from DGC TCGA GBM RNA-seq cohort ($n = 173$), there are only 11.3% of all patients overlap. Notably, the combination of *Nudt21* expression and *Pak1* expression has more power to predict prognosis of GBM patients compared to *Pak1* expression only. All of the above implies that *Nudt21* and its downstream target, *Pak1* play an important role in GBM development and progression. Targeting *Pak1* while devising methods to increase expression of *Nudt21* (possibly through microRNA inhibition) represents potential therapy targets for GBM treatment.

Materials and methods

Cell lines and reagents

GBM cell lines (LN229, U251) were obtained from ATCC (Manassas, VA, USA) and maintained in DMEM medium (Sigma-Aldrich, St. Louis, MO) supplemented with 10% fetal calf serum (Sigma), penicillin (100 U/ml)/streptomycin

(100 µg/ml). The cells are mycoplasma free by incubating Plasmocin (InvivoGen, Cat# ant-MPT) for 2 weeks before transfection. Luciferase reporter vector (pLightSwitch *Pak1* 3′-UTR) was purchased from SWITCHGEAR genomics (Menlo Park, CA, USA). Lipofectamine 2000 and Lipofectamine RNAimax were purchased from (Thermo Fisher scientific, Grand Island, NY, USA). The siRNAs for *Nudt21* (ID: SASI_Hs01_00146875~77) and *Pak1* (ID: SASI_Hs01_00087968 and SASI_Hs02_00334074) were synthesized from Sigma-Genosys (Woodland, TX). RNAi experiments were performed as previously described [48]. Antibodies: anti-CFIm25, anti-CFIm59 and anti-CFIm68 antibodies (Santa Cruz, Dallas, TX, Cat# sc-271767); anti-GAPDH (Sigma, St. Loius, MO, G9545); anti-PAK1 and PAK2 antibodies (Cell Signaling, Danvers, MA).

Gene expression and survival association acquisition and analysis

The Cancer Genome Atlas (TCGA) gene expression data (Affymetrix Human Exon 1.0 ST) and clinical information for GBM samples were obtained from Project Betastasis (www.betastasis.com). The dataset included 349 GBM patient samples with clinic information and microarray data are available. The expression value of *Nudt21*, *CFIm59*, *CFIm68*, *Pak1* and *Pak2* was collected for each case, and presents as mean ± standard deviation. *p*-value is calculated with two side *T*-test for statistical comparison. Kaplan–Meier plots are used to measure the survival association with gene expression at 25% thresholds and the days from the date of diagnosis to the time of death using Betastasis online representation tool with a log-rank test. The RNA-seq expression data for LGG and GBM samples was obtained from UCSC Xena browser (https://xena.ucsc.edu/), which have 259 grade II LGG samples, 266 grade III LGG samples and 173 GBM (grade IV). The patients were stratified into two main groups based on top and bottom 30% of gene expression profiles, and survival plot was calculated using the standard cox proportional hazards model implemented in "survival" R package.

PAC-seq processing and 3′-UTR APA analysis

Poly(A)-ClickSeq (PAC-seq) was used to investigate 3′-UTR APA usage in a genome-wide fashion [32]. Briefly, 1 µg of total RNA from three control cells and three *Nudt21* KD cells were reverse transcribed with partial p7 adaptor (Illumina_4N_21T: GTGACTGGAGTTCAGACGTGTGC TCTTCCGATCTNNNN TTTTTTTTTT TTTTTTTTTT) and dNTPs with the addition of spiked-in azido-nucleotides (AzVTPs) at 5:1. p5 adaptor (5'Hexynyl-NNNNA-GATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAG ATCT CGGTGGTCGCCGTATCATT, IDT) was clicked to 5′ end of cDNA by CuAAC. The libraries were amplified for 21 cycles with Universal primer (AATGATACGG CGACCACCGAG) and 3′ indexing primer (for example Index 1: CAAGCAGAAGACGGCATACGAGATCGTGA TGTG ACTGGAGTTCAGACGTGT) and purified on a 2% agarose gel at molecular size from 200 bp to 300 bp. Then the libraries were pooled and sequenced by Miseq with illumina Miseq Reagent Kit V3 (Cat# MS-102-3001). The reads were processed and quality filtered as previously described [32] mapped to the human genome (hg19) using Hisat2 [49] using the default mapping parameters, with the exception of disallowing softpads at the 3′-end of the mapped read in order to prevent mis-annotation of the poly (A) site. To analyze the alternative polyadenylation, sites found within the terminal exon of genes annotated in the UCSC genome browser were extracted and compared between control and *Nudt21* KD cells. If multiple poly(A) sites were found within the terminal exon and if the relative usage of these was altered by >20% between the control and *Nudt21* KD cells, then these poly(A) sites were deemed to be alternatively poly-adenylated.

Co-immunoprecipitation assay and western blotting

Protein A/G plus-agarose beads were used in the co-immunoprecipitation assay according to the manufacturer's instructions (Santa Cruz Biotechnology, Texas, USA). Briefly, LN229 and U251 cells were grown in 4 × 10 cm dishes with DMEM medium with 10% fetal bovine serum in a humidified incubator containing 5% CO_2 at 37 °C. Cells were collected with 1 ml of RIPA lysate buffer with 1X protein inhibitor cocktail for each dish for a single co-IP assay. The cell lysate was centrifuged at 10,000 × *g* for 15 min at 4 °C after sonicated and incubated with magnetic beads conjugated with anti-NUDT21 antibody, anti-CFIm59 antibody, anti-CFIm68 antibody, mouse IgG as a negative control. The pull-down complex was analyzed by western blotting with indicated antibodies. For western blotting, 30–50 µg of total protein was loaded onto an 8–12% SDS-PAGE gel and transferred onto PVDF membrane, followed by blocking and probing with indicated primary antibodies, and HRP-conjugated secondary antibodies, and visualized by an ECL plus chemiluminescence. GAPDH was used as an internal control.

Protein–protein interaction with two yeast hybrid assays

Yeast two-hybrid assays were carried out in PJ69-4a and PJ49-4α. Human *Nudt21*, *CFIm59* and *CFIm68* genes were cloned into pOBD vector as a bait, and human *Nudt21* gene was cloned into pOAD as a prey using conventional cloning strategy. All clones were sequenced to confirm their

identity. pOBD plasmids were transformed into PJ69-4a yeast and were selected on tryptophan-dropout medium; pOAD plasmids were transformed into PJ49-4 α yeast and were selected on leucine-dropout medium. Double transformants were created by mating the yeast strains followed by selection on medium lacking both tryptophan and leucine. Interactions were tested through serial dilution of diploid yeast followed by plating on medium lacking tryptophan and leucine or on medium lacking tryptophan, leucine, and histidine that also were supplemented with 1 mM 3-amino-1,2,4-triazole.

KEGG signaling pathway analysis

The KEGG (Kyoto encyclopedia of genes and genomes) PATHWAY database records networks of molecular interactions in the cells and dealing with biological pathways. In order to explore the signaling pathways associated with 3′-UTR APA events, we collected all the 686 3′-UTR shortening genes in *Nudt21* KD cells and performed KEGG analysis by using KEGG map-pathway tool. We identified 12 *Nudt21*-mediated APA genes enriched in *RAS* signaling pathway.

Quantification of 3′-UTR alternative polyadenylation by RHAPA

TRIzol reagent (400 µl/well, Life Technologies) was added to the transfected cells in a 6-well plate and left on the cells for 5 min at room temperature and transferred to 1.5 ml microfuge tubes. Then, 80 µl of chloroform was added to the tubes, and vigorously shaken by hand for 15 s, followed by centrifugation (~10,000 × *g*) for 15 min after incubation at room temperature for 10 min. The RNA-containing aqueous phase was removed and transferred to a fresh microfuge tube and equal volume of isopropanol was added to precipitate RNA. To quantify alternative polyadenylation, gene-specific antisense DNA oligonucleotides whose sequences were between proximal polyadenylation site (pPAS) and distal polyadenylation site (dPAS) in 3′-UTR region were used to form an RNA:DNA hybrid which was cut by RNase H. Then oligo dT primers were used to convert poly(A) contained transcripts to cDNA. Both usage of pPAS and dPAS was quantified by real time PCR with specific primer pairs. The distal primer pair was designed to amplify the 3′-UTR region just before the dPAS to detect long transcripts that use the dPAS. The proximal primer pair was designed to amplify 3′-UTR region before the pPAS and detect short transcripts that use the pPAS. Fold change of pPAS/dPAS was calculated in *Nudt21* KD cells over control cells. *GAPDH* was used as an internal control. Oligonucleotides used for RHAPA: *Pak1* 3′-UTR distal forward, 5′-TCTCCCACTATGGTAGGACCCCT-3′; reverse, 5′-TGTGCTGCAGAGGCAGT-3′ AGT; proximal forward, 5′-ATTGTGCCAAGCCTT CTGTG-3′; reverse, 5′-GGAAATGGGAGAAGCAAGGC-3′; for RNase H cleavage 5′-CAACAC CCAGTGTAAGCATT-3′. *Pak2* 3′-UTR distal forward, 5′-ACCTGTGCCTCTAACAAGCG-3′; reverse, 5′-CAGCGGAGACAGGAAGACAAT-3′; proximal forward, 5′-TCAGGCTTGGCTCTAGGAAC-3′; reverse, 5′-GAGGGAAGAAGGTAGTGGCA-3′; for RNase H cleavage 5′-TCAGATTTAAATGGGAATTTTCATTCTAAAGAGATGA-3′.

Pak1 3′-UTR-mediated luciferase reporter assays

Luciferase reporter constructs with *Pak1* 3′-UTR (pLightSwitch *Pak1* 3′-UTR) and pGL4 Luciferase reporter vector were purchased from SWITCHGEAR genomics (Menlo Park, CA, USA) and Promega (Madison, WI, USA). Briefly, LN229 cells were transfected with the indicated siRNAs for 24 h and then were co-transfected with *Pak1* 3′-UTR luciferase reporter and pGL4 Luciferase reporter vectors using lipofectamine 2000 Transfection Reagent. Then luciferase activity was assayed 24 h after transfection using Dual-Luciferase Reporter Assay (Promega) according to manufacturer's protocol.

Soft agar colony formation assay

LN229 and U251 cells were transfected with the indicated siRNAs for 24 h and were used to determine anchorage-dependent growth. For the base layer, 1.5% of UltraPure low melting point agarose was mixed 1:2 with 1X DMEM media and plated in 6-well plates giving a 1 ml bottom layer of 0.5% agar. Then LN229 and U251 transfected cells were added to 1X DMEM, and mixed with 1.5% agar to give a 0.375% agar cell suspension at 2×10^3 cells/ml. One milliliter was dispensed into each well as a second layer. The agar was covered with 1 ml of 1X DMEM and incubated in a humidified incubator at 37 °C (5% CO_2) for 2 weeks, formed colonies were photographed and counted. Three-independent experiments were performed in duplicate.

Cell migration assay

LN229 and U251 cells were seeded into a 6-well plate overnight and transfected with the indicated siRNAs mixed with LipofectamineRNAiMax for 24 h. The transfected cells were suspended in serum free 1X DMEM medium at 5×10^5 cells/ml, and seeded 300 µl suspended cells into the transwell inserts of a CytoSelect 24-well cell migration assay plate (Cell Biolabs Inc). The cells in the transwell inserts were incubated in 500 µl 10% FBS 1X DMEM medium for 6 h in a humidified tissue culture incubator at

37 °C 5% CO_2 atmosphere. Remove the non-migrating cells in the tranwell inserts with a cotton tipped swab and fixed the cells in the staining buffer for 5 min. Wash the inserts twice with 1X PBS and count the migrating cells under the microscope in three fields of triplicate inserts each cell line. The data represented the mean ± standard deviations.

Statistical analysis

The two-tailed Student's *t*-test was used for determining the probability of difference between the test group and the control group. The log-rank test compares the survival times of two or more groups. Statistical significance was assumed at $p <= 0.05$.

Acknowledgements This work was supported by the US National Institutes of Health (NIH) grants R01CA193466-01 to W.L. and E.J. W., the Cancer Prevention Research Institute of Texas grant CPRIT RP140800 to E.J.W., and NIH R03 CA223893-01 to P.J.

Author contributions E.J.W and P.J. conceived the project, designed the experiments and Y.C. and N.E. performed the experiments and data analysis. C.W. and Z.X. performed survival analysis. N.E. and A. R. performed the PAC-Seq APA analysis. L.L. and W.L. performed TCGA GBM RNA-seq data analysis, T.C. helped with data analysis. Y.C., P.J., and E.J.W. wrote the manuscript with input from N.E., L.L., T.C., and A.R.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Wen PY, Kesari S. Malignant gliomas in adults. N Engl J Med. 2008;359:492–507.
2. Chen J, Li Y, Yu TS, McKay RM, Burns DK, Kernie SG, et al. A restricted cell population propagates glioblastoma growth after chemotherapy. Nature. 2012;488:522–6.
3. Wang R, Chadalavada K, Wilshire J, Kowalik U, Hovinga KE, Geber A, et al. Glioblastoma stem-like cells give rise to tumour endothelium. Nature. 2010;468:829–33.
4. Kang TW, Choi SW, Yang SR, Shin TH, Kim HS, Yu KR, et al. Growth arrest and forced differentiation of human primary glioblastoma multiforme by a novel small molecule. Sci Rep. 2014;4:5546.
5. Cancer Genome Atlas Research N. Comprehensive genomic characterization defines human glioblastoma genes and core pathways. Nature. 2008;455:1061–8.
6. Brennan CW, Verhaak RG, McKenna A, Campos B, Noushmehr H, Salama SR, et al. The somatic genomic landscape of glioblastoma. Cell. 2013;155:462–77.
7. Diederichs S, Bartsch L, Berkmann JC, Frose K, Heitmann J, Hoppe C, et al. The dark matter of the cancer genome: aberrations in regulatory elements, untranslated regions, splice sites, non-coding RNA and synonymous mutations. EMBO Mol Med. 2016;8:442–57.
8. Tian B, Manley JL. Alternative polyadenylation of mRNA precursors. Nat Rev Mol Cell Biol. 2017;18:18–30.
9. Masamha CP, Xia Z, Yang J, Albrecht TR, Li M, Shyu AB, et al. CFIm25 links alternative polyadenylation to glioblastoma tumour suppression. Nature. 2014;510:412–6.
10. Singh P, Alley TL, Wright SM, Kamdar S, Schott W, Wilpan RY, et al. Global changes in processing of mRNA 3' untranslated regions characterize clinically distinct cancer subtypes. Cancer Res. 2009;69:9422–30.
11. Gruber AR, Martin G, Muller P, Schmidt A, Gruber AJ, Gumienny R, et al. Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. Nat Commun. 2014;5:5465.
12. Elkon R, Drost J, van Haaften G, Jenal M, Schrier M, Oude Vrielink JA, et al. E2F mediates enhanced alternative polyadenylation in proliferation. Genome Biol. 2012;13:R59.
13. Mayr C, Bartel DP. Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell. 2009;138:673–84.
14. Muller S, Rycak L, Afonso-Grunz F, Winter P, Zawada AM, Damrath E, et al. APADB: a database for alternative polyadenylation and microRNA regulation events. Database. 2014;2014. https://doi.org/10.1093/database/bau076.
15. Sun M, Ding J, Li D, Yang G, Cheng Z, Zhu Q. NUDT21 regulates 3'-UTR length and microRNA-mediated gene silencing in hepatocellular carcinoma. Cancer Lett. 2017;410:158–68.
16. Masamha CP, Xia Z, Peart N, Collum S, Li W, Wagner EJ, et al. CFIm25 regulates glutaminase alternative terminal exon definition to modulate miR-23 function. RNA. 2016;22:830–8.
17. Akman BH, Can T, Erson-Bensan AE. Estrogen-induced upregulation and 3'-UTR shortening of CDC6. Nucleic Acids Res. 2012;40:10679–88.
18. Masamha CP, Wagner EJ. The contribution of alternative polyadenylation to the cancer phenotype. Carcinogenesis. 2018;39:2–10.
19. Hardy JG, Norbury CJ. Cleavage factor Im (CFIm) as a regulator of alternative polyadenylation. Biochem Soc Trans. 2016;44:1051–7.
20. Ozsolak F, Kapranov P, Foissac S, Kim SW, Fishilevich E, Monaghan AP, et al. Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. Cell. 2010;143:1018–29.
21. Fu Y, Sun Y, Li Y, Li J, Rao X, Chen C, et al. Differential genome-wide profiling of tandem 3' UTRs among human breast cancer and normal cells by high-throughput sequencing. Genome Res. 2011;21:741–7.
22. Shao J, Zhang J, Zhang Z, Jiang H, Lou X, Huang B, et al. Alternative polyadenylation in glioblastoma multiforme and changes in predicted RNA binding protein profiles. OMICS. 2013;17:136–49.
23. Xia Z, Donehower LA, Cooper TA, Neilson JR, Wheeler DA, Wagner EJ, et al. Dynamic analyses of alternative polyadenylation from RNA-seq reveal a 3'-UTR landscape across seven tumour types. Nat Commun. 2014;5:5274.
24. Park HJ, Ji P, Kim S, Xia Z, Rodriguez B, Li L, et al. 3' UTR shortening represses tumor-suppressor genes in trans by disrupting ceRNA crosstalk. Nat Genet. 2018;50:783–89.
25. Gruber AR, Martin G, Keller W, Zavolan M. Cleavage factor Im is a key regulator of 3' UTR length. RNA Biol. 2012;9:1405–12.
26. Zhu Y, Wang X, Forouzmand E, Jeong J, Qiao F, Sowd GA, et al. Molecular mechanisms for CFIm-mediated regulation of mRNA alternative polyadenylation. Mol Cell. 2018;69:62–74e4.

27. Brown KM, Gilmartin GM. A mechanism for the regulation of pre-mRNA 3' processing by human cleavage factor Im. Mol Cell. 2003;12:1467–76.
28. Ruegsegger U, Beyer K, Keller W. Purification and characterization of human cleavage factor Im involved in the 3' end processing of messenger RNA precursors. J Biol Chem. 1996;271:6107–13.
29. Yang Q, Gilmartin GM, Doublie S. Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3' processing. Proc Natl Acad Sci USA. 2010;107:10062–7.
30. Yang Q, Coseno M, Gilmartin GM, Doublie S. Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. Structure. 2011;19:368–77.
31. Brumbaugh J, Di Stefano B, Wang X, Borkent M, Forouzmand E, Clowers KJ, et al. Nudt21 controls cell fate by connecting alternative polyadenylation to chromatin signaling. Cell. 2017;172:629–31.
32. Routh A, Ji P, Jaworski E, Xia Z, Li W, Wagner EJ. Poly(A)-ClickSeq: click-chemistry for next-generation 3-end sequencing without RNA enrichment or fragmentation. Nucleic Acids Res. 2017;45:e112.
33. Han T, Kim JK. Driving glioblastoma growth by alternative polyadenylation. Cell Res. 2014;24:1023–4.
34. Routh A, Head SR, Ordoukhanian P, Johnson JE. ClickSeq: fragmentation-free next-generation sequencing via click ligation of adaptors to stochastically terminated 3'-azido cDNAs. J Mol Biol. 2015;427:2610–6.
35. Kumar R, Gururaj AE, Barnes CJ. p21-activated kinases in cancer. Nat Rev Cancer. 2006;6:459–71.
36. Bagheri-Yarmand R, Mandal M, Taludker AH, Wang RA, Vadlamudi RK, Kung HJ, et al. Etk/Bmx tyrosine kinase activates Pak1 and regulates tumorigenicity of breast cancer cells. J Biol Chem. 2001;276:29403–9.
37. Holm C, Rayala S, Jirstrom K, Stal O, Kumar R, Landberg G. Association between Pak1 expression and subcellular localization and tamoxifen resistance in breast cancer patients. J Natl Cancer Inst. 2006;98:671–80.
38. Tian B, Hu J, Zhang H, Lutz CS. A large-scale analysis of mRNA polyadenylation of human and mouse genes. Nucleic Acids Res. 2005;33:201–12.
39. Cornett AL, Lutz CS. RHAPA: a new method to quantify alternative polyadenylation. Methods Mol Biol. 2014;1125:157–67.
40. Denysenko T, Gennero L, Roos MA, Melcarne A, Juenemann C, Faccani G, et al. Glioblastoma cancer stem cells: heterogeneity, microenvironment and related therapeutic strategies. Cell Biochem Funct. 2010;28:343–51.
41. Kim S, Yamamoto J, Chen Y, Aida M, Wada T, Handa H, et al. Evidence that cleavage factor Im is a heterotetrameric protein complex controlling alternative polyadenylation. Genes Cells. 2010;15:1003–13.
42. Sowd GA, Serrao E, Wang H, Wang W, Fadel HJ, Poeschla EM, et al. A critical role for alternative polyadenylation factor CPSF6 in targeting HIV-1 integration to transcriptionally active chromatin. Proc Natl Acad Sci USA. 2016;113:E1054–63.
43. Ji Z, Lee JY, Pan Z, Jiang B, Tian B. Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci USA. 2009;106:7028–33.
44. Brumbaugh J, Di Stefano B, Wang X, Borkent M, Forouzmand E, Clowers KJ, et al. Nudt21 controls cell fate by connecting alternative polyadenylation to chromatin signaling. Cell. 2018;172:106–20e21.
45. Creemers EE, Bawazeer A, Ugalde AP, van Deutekom HW, van der Made I, de Groot NE, et al. Genome-wide polyadenylation maps reveal dynamic mrna 3'-end formation in the failing human heart. Circ Res. 2016;118:433–8.
46. Chen MJ, Wu DW, Wang YC, Chen CY, Lee H. PAK1 confers chemoresistance and poor outcome in non-small cell lung cancer via beta-catenin-mediated stemness. Sci Rep. 2016;6:34933.
47. Aoki H, Yokoyama T, Fujiwara K, Tari AM, Sawaya R, Suki D, et al. Phosphorylated Pak1 level in the cytoplasm correlates with shorter survival time in patients with glioblastoma. Clin Cancer Res. 2007;13(22 Pt 1):6603–9.
48. Wagner EJ, Garcia-Blanco MA. RNAi-mediated PTB depletion leads to enhanced exon definition. Mol Cell. 2002;10:943–9.
49. Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods. 2015;12:357–60.
50. Edlinger L, Berger-Becvar A, Menzl I, Hoermann G, Greiner G, Grundschober E, et al. Expansion of BCR/ABL1+cells requires PAK2 but not PAK1. Br J Haematol. 2017;179:229–41.
51. Larson KC, Lipko M, Dabrowski M, Draper MP. Gng12 is a novel negative regulator of LPS-induced inflammation in the microglial cell line BV-2. Inflamm Res. 2010;59:15–22.
52. Nagarajan P, Chin SS, Wang D, Liu S, Sinha S, Garrett-Sinha LA. Ets1 blocks terminal differentiation of keratinocytes and induces expression of matrix metalloproteases and innate immune mediators. J Cell Sci. 2010;123(Pt 20):3566–75.
53. Johnson DB, Smalley KS, Sosman JA. Molecular pathways: targeting NRAS in melanoma and acute myelogenous leukemia. Clin Cancer Res. 2014;20:4186–92.
54. Riemann K, Struwe H, Alakus H, Obermaier B, Schmitz KJ, Schmid KW, et al. Association of GNB4 intron-1 haplotypes with survival in patients with UICC stage III and IV colorectal carcinoma. Anticancer Res. 2009;29:1271–4.
55. Jasonni VM, Amadori A, Santini D, Ceccarelli C, Naldi S, Flamigni C. Epidermal growth factor receptor (EGF-R) and transforming growth factor alpha (TGFA) expression in different endometrial cancers. Anticancer Res. 1995;15:1327–32.
56. Zimmermannova O, Doktorova E, Stuchly J, Kanderova V, Kuzilkova D, Strnad H, et al. An activating mutation of GNB1 is associated with resistance to tyrosine kinase inhibitors in ETV6-ABL1-positive leukemia. Oncogene. 2017;36:5985–94.
57. Li TF, Qin SH, Ruan XZ, Wang X. p120-catenin participates in the progress of gastric cancer through regulating the Rac1 and Pak1 signaling pathway. Oncol Rep. 2015;34:2357–64.
58. Yajima I, Kumasaka MY, Yamanoshita O, Zou C, Li X, Ohgami N, et al. GNG2 inhibits invasion of human malignant melanoma cells with decreased FAK activity. Am J Cancer Res. 2014;4:182–8.
59. Giubellino A, Burke TR Jr., Bottaro DP. Grb2 signaling in cell motility and cancer. Expert Opin Ther Targets. 2008;12:1021–33.
60. Walker-Daniels J, Hess AR, Hendrix MJ, Kinch MS. Differential regulation of EphA2 in normal and malignant cells. Am J Pathol. 2003;162:1037–42.