circHIPK3 nucleates IGF2BP2 and functions as a competing endogenous RNA

Trine Line Hauge Okholm[1-3,*], Andreas Bjerregaard Kamstrup[4], Morten Muhlig Nielsen[1,3], Anne Kruse Hollensen[4], Mette Laugesen Graversgaard[4], Lasse Sommer Kristensen[5], Søren Vang[1], Samuel S. Park[6], Gene W. Yeo[6], Lars Dyrskjøt[1,3], Jørgen Kjems[4,7], Jakob Skou Pedersen[1,3,8,9,*], Christian Kroun Damgaard[4,9,*]

[1]Department of Molecular Medicine (MOMA), Aarhus University Hospital, 8200 Aarhus N, Denmark

[2]Departments of Otolaryngology-Head and Neck Surgery and Microbiology & Immunology, University of California, San Francisco, San Francisco, CA, USA

[3]Department of Clinical Medicine, Aarhus University, 8000 Aarhus C, Denmark

[4]Department of Molecular Biology and Genetics (MBG), Aarhus University, 8000 Aarhus C, Denmark

[5]Department of Biomedicine, Aarhus University, 8000 Aarhus C, Denmark

[6]Department of Cellular and Molecular Medicine, University of California, San Diego, La Jolla, CA 92093, USA

[7]Interdisciplinary Nanoscience Center (iNANO), Aarhus University, 8000 Aarhus C, Denmark

[8]Bioinformatics Research Center (BiRC), Aarhus University, 8000 Aarhus C, Denmark

[9]These authors contributed equally

[*]Co-corresponding authors

Co-corresponding authors:

Trine Line Hauge Okholm, PhD, Departments of Otolaryngology-Head and Neck Surgery and Microbiology & Immunology, University of California, San Francisco, San Francisco, CA, USA, email: trineline.okholm@ucsf.edu

Jakob Skou Pedersen, Professor, Department of Molecular Medicine, Aarhus University Hospital, Palle Juul-Jensens Boulevard 99, 8200 Aarhus N, Denmark, email: jakob.skou@clin.au.dk.

Christian Kroun Damgaard, Associate Professor, Department of Molecular Biology and Genetics, Universitetsbyen 81, 8000 Aarhus C, email: ckd@mbg.au.dk

Abstract

Circular RNAs (circRNAs) represent a class of widespread endogenous RNAs that regulate gene expression and thereby influence cell biological decisions with implications for the pathogenesis of several diseases. Here, we disclose a novel gene-regulatory role of circHIPK3 by combining analyses of large genomics datasets and mechanistic cell biological follow-up experiments. Specifically, we use temporal depletion of circHIPK3 or specific RNA binding proteins (RBPs) and identify several perturbed genes by RNA sequencing analyses. Using expression-coupled motif analyses of mRNA expression data from various knockdown experiments, we identify an 11-mer motif within circHIPK3, which is also enriched in genes that become downregulated upon circHIPK3 depletion. By mining eCLIP datasets, we find that the 11-mer motif constitutes a strong binding site for IGF2BP2 and validate this circHIPK3-IGF2BP2 interaction experimentally using RNA-immunoprecipitation and competition assays in bladder cancer cell lines. Our results suggest that circHIPK3 and IGF2BP2 mRNA targets compete for binding. Since the identified 11-mer motif found in circHIPK3 is enriched in upregulated genes following IGF2BP2 knockdown, and since IGF2BP2 depletion conversely globally antagonizes the effect of circHIPK3 knockdown on target genes, our results suggest that circHIPK3 can sequester IGF2BP2 as a competing endogenous RNA (ceRNA), leading to target mRNA stabilization. As an example of a circHIPK3-regulated gene, we focus on the *STAT3* mRNA as a specific substrate of IGF2BP2 and validate that manipulation of circHIPK3 regulates IGF2BP2-*STAT3* mRNA binding and thereby *STAT3* mRNA levels. However, absolute copy number quantifications demonstrate that IGF2BP2 outnumbers circHIPK3 by orders of magnitude, which is inconsistent with a simple 1:1 ceRNA hypothesis. Instead, we show that circHIPK3 can nucleate multiple copies of IGF2BP2, potentially via phase separation, to produce IGF2BP2 condensates. Finally, we show that circHIPK3 expression correlates with overall survival of patients with bladder cancer. Our results are consistent with a model where relatively few cellular circHIPK3 molecules function as inducers of IGF2BP2 condensation thereby regulating STAT3 and other key factors for cell proliferation and potentially cancer progression.

Introduction

Circular RNAs (circRNAs) are endogenous predominantly non-coding RNAs, formed from precursor mRNA through backsplicing events (Starke et al., 2015). Although first thought to be by-products or a result of erroneous splicing (Cocquerelle et al., 1993), recent improvements in next-generation sequencing of non-polyadenylated transcriptomes and bioinformatics technology have revealed the existence of thousands of human circRNAs (Jeck et al., 2013; Liu and Chen, 2022; Memczak et al., 2013). Now, accumulating evidence suggests that circRNAs regulate numerous cell biological processes (reviewed in (Liu and Chen, 2022)) and have been implicated in tumorigenesis (S. Chen et al., 2019; Vo et al., 2019). Some circRNAs possess biomarker potential (reviewed in (Kristensen et al., 2022; Zhang et al., 2018)) owing to expression specificity (Memczak et al., 2013), structural stability (Jeck et al., 2013), and abundance in exosomes (Y. Li et al., 2015), blood (Memczak et al., 2015), plasma (Bahn et al., 2015), and urine (Vo et al., 2019).

The functional impact of circRNAs in cancer is still a major area of research. Many circRNAs can interact with RNA binding proteins (RBPs) (Ashwal-Fluss et al., 2014; Du et al., 2016) and several circRNAs are suggested to function as miRNA sponges (F. Li et al., 2015; Li et al., 2017a; Zheng et al., 2016). However, besides ciRS-7 and circSRY, which can bind a large number of miRNAs (Hansen et al., 2013; Memczak et al., 2013), circRNAs are not enriched for more miRNA binding sites than would be expected by chance (Guo et al., 2014). Recently, it was discovered that some circRNAs can be translated into proteins under certain conditions (Chen et al., 2021; Legnini et al., 2017; Pamudurti et al., 2017) but the majority of circRNAs likely do not encode proteins (Stagsted et al., 2019). Only a few circRNAs have been intensely studied and also functionally described *in vivo*, e.g., ciRS-7 (Piwecka et al., 2017). However, functional implications are usually restricted to specific conditions, cells, and tissues, with small data sets and few experimental validation studies being general caveats that provide inconsistent conclusions. Additionally, there is a general lack of knowledge on functional mechanisms, beyond potential miRNA sponging, which is often based on studies of overexpression of circRNAs and/or miRNA mimics and reporter assays.

CircHIPK3 (chr11:33307958-33309057) (hg19) is the most prevalent alpha circRNA, i.e., it is the circRNA most often found to be highest expressed in human samples (Stagsted et al., 2019). It originates from the second exon of the gene *HIPK3* and possesses highly conserved regions across species (Zheng et al., 2016). Accumulating evidence indicates that circHIPK3 possesses biomarker potential and could be a potential target in cancer therapy (reviewed in (Wen et al., 2020; Xie et al., 2020)). While circHIPK3 is often deregulated in cancer, studies indicate that circHIPK3 play dual roles in tumorigenesis; some studies report that circHIPK3 is downregulated in cancers and possess tumor suppressor functions, e.g., in bladder cancer (Li et al., 2017a; Okholm et al., 2017), ovarian cancer (Teng et al., 2019), and gastric cancer

(Ghasemi et al., 2019), while other studies describe the opposite effect of circHIPK3 in the same and other cancer types, e.g., glioma cancer (Jin et al., 2018), colorectal cancer (Zeng et al., 2018), prostate cancer (Cai et al., 2019), ovarian cancer (Liu et al., 2018), lung cancer (X. Chen et al., 2019), gastric cancer (Cheng et al., 2018), and hepatocellular carcinoma (Chen et al., 2018; X. Chen et al., 2019; Zheng et al., 2016). Several of these studies conclude that circHIPK3 regulates a wide range of genes by sponging different miRNAs in various cancer tissues, e.g., miR-124-3p (Chen et al., 2018; X. Chen et al., 2019; Zheng et al., 2016), miR-338-3p (Cai et al., 2019), miR-193a-3p (D. Chen et al., 2019), miR-7 (Zeng et al., 2018), miR-654 (Jin et al., 2018), and miR-558 (Li et al., 2017a).

In a previous study, we found that circHIPK3 is expressed in patients with non-muscle invasive bladder cancer (NMIBC) and correlates positively with bladder cancer progression independently of the parent gene (Okholm et al., 2017). Additionally, we found that circHIPK3 is higher expressed than the corresponding linear transcript and possesses key biological features, e.g., enrichment of evolutionary conserved regions and miRNA target sites, indicating a regulatory function of its own.

Here, we extensively evaluate the functional impact of circHIPK3 in bladder cancer cells by combining large-scale computational analyses with comprehensive validation experiments. By conducting time-series knockdown (KD) experiments, we identify thousands of genes that are deregulated upon circHIPK3 KD across different bladder cancer cell lines. We find no evidence for a miRNA sponge function of circHIPK3, since predicted miRNA-targeted mRNAs are not significantly affected by efficient circHIPK3 depletion. Instead, using a computational method, Regmex, that evaluates the enrichment of nucleotide motifs in deregulated genes (Nielsen et al., 2018), we identify an 11-mer circHIPK3 motif that is also enriched in mRNAs that become downregulated upon circHIPK3 knockdown. We examine a large eCLIP data set of RNA binding protein (RBP) targets and discover that the motif is part of a binding site for IGF2BP2. By performing RNA immunoprecipitation (RIP), we demonstrate that IGF2BP2 associates with circHIPK3, primarily via the 11-mer motif. We show that IGF2BP2 KD leads to upregulation of genes containing the 11-mer motif and that IGF2BP2 KD counteracts the effect of circHIPK3 KD on target genes and proliferation pathways. One of these targets, *STAT3*, is downregulated upon circHIPK3 KD, while TP53, a known target of STAT3 inhibition, becomes upregulated. Consistent with a circHIPK3 competing endogenous RNA (ceRNA) model, we confirm that *STAT3* mRNA is bound more efficiently by IGF2BP2 upon circHIPK3 depletion. Conversely, this apparent competition is dependent on the 11-mer motif, since its presence within circHIPK3 reduces the *STAT3* mRNA-IGF2BP2 interaction. However, careful quantifications of copy numbers of circHIPK3 and IGF2BP2 suggest that a simple ceRNA model is unlikely as IGF2BP2 outnumbers circHIPK3 by orders of magnitude. Since ectopic expression of circHIPK3 induces IGF2BP2 condensation within cells, we propose that even a few circHIPK3 molecules can function as a

ceRNAs by nucleating numerous IGF2BP2 molecules per circRNA. Finally, by analyzing large patient cohorts with rich clinical annotations, we show that the expression of circHIPK3 is positively associated with overall survival in bladder cancer patients.

Results

Thousands of genes are deregulated upon circHIPK3 knockdown

To comprehensively study the role of circHIPK3 in the regulation of gene expression in bladder cancer, we performed time-series knockdown (KD) experiments in bladder cancer (BC) cell lines followed by RNA sequencing (RNA-seq). First, we evaluated the expression of circHIPK3 (the second exon of *HIPK3*, position chr11:33,307,958-33,309,057 (hg19), Figure 1A) and the corresponding linear transcript in non-malignant (n = 4) and metastatic (n = 7) BC cell lines (Hedegaard et al., 2016; Okholm et al., 2017) using the CIRI2-pipeline (Gao et al., 2018). In all cell lines except for one, circHIPK3 is higher expressed than the linear transcript (Figure 1B). Based on circHIPK3 expression and cell line stability, we chose J82 and UMUC3 for circHIPK3 KD experiments. We designed an siRNA against the backsplicing junction of circHIPK3 to specifically target the circle but not the corresponding linear transcript (Figure 1A, Methods). We transfected cells with the circHIPK3-specific siRNA or a scrambled (scr) siRNA as negative control and measured the KD efficiency of circHIPK3 using divergent primers against the unique backsplicing junction (Figure 1C and Supplementary Table S1). Next, we harvested RNA at five time points (0, 12, 24, 48, and 72 hours post siRNA transfection) in three biological replicates and quantified mRNA expression using QuantSeq (Moll et al., 2014). Additionally, we included control samples without siRNA transfection to ensure that gene perturbation is not a consequence of off-target effects or transfection *per se* (Supplementary Figure S1A). Based on mRNA expression, we assessed overall sample similarity. As expected, samples cluster according to time (Figure 1D).

Next, we performed differential expression analyses and identified thousands of genes that are differentially expressed (DE) upon circHIPK3 KD across time in UMUC3 (n = 3,072) and J82 (n = 2,389) (FDR < 0.1, Benjamini-Hochberg correction), with a large overlap (n = 1,104) between cell lines (P < 2.2e-16, Fisher's Exact Test, Figure 1E). While few genes are DE at 12 hours, the number of DE genes increases dramatically at 24 hours (Figure 1F and Supplementary Figure S1B), where 1,241 (UMUC3) and 642 (J82) genes are significantly down- or upregulated (Wald Test, Benjamini-Hochberg correction with FDR < 0.1, Figure 1G). Convincingly, 92% of the shared DE genes between UMUC3 and J82 cells show the same perturbation profile in both cell lines (Supplementary Figure S1C), providing reliability of DE genes and indicating shared functions of circHIPK3 across BC cell lines.

Lack of evidence for miRNA-sponge functions

The most widespread proposed function of circRNAs is as miRNA sponges although this is in many cases based on circRNA/miRNA mimic overexpression studies or reporter assays, which may not provide concluding evidence for an endogenous function (Jarlstad Olesen and S Kristensen, 2021). Nevertheless, several studies have suggested that circHIPK3 sponges multiple miRNAs in various tissues (Chen et al., 2018; Jin et al., 2018; Li et al., 2017a; Shan et al., 2017; Stoll et al., 2018; Zeng et al., 2018; Zheng et al., 2016). Based on AGO-CLIP data of experimentally defined miRNA binding sites, we previously identified 11 distinctly conserved miRNA target sites in circHIPK3 (Hamilton et al., 2016, 2013; Okholm et al., 2017) (Figure 2A and Supplementary Table S2) constituting binding sites for 23 miRNAs (circHIPK3-miRNAs). We profiled the expression of ~790 miRNAs in J82 (Supplementary Figure S2A and Supplementary Table S3) and evaluated if the 3'UTRs of downregulated genes upon circHIPK3 KD are enriched with target sites for expressed circHIPK3-miRNAs in J82. Neither the target sites for the highest expressed miRNA (miR-148-3p; CAGUGCA) or any of the other circHIPK3-miRNAs are enriched among downregulated genes ($P > 0.05$, Chi-square test, Figure 2B and Supplementary Figure S2B, C), indicating that circHIPK3 do not sponge these miRNAs.

Abundance and regulatory roles of miRNAs are highly cell type and tissue dependent (Landgraf et al., 2007; Ludwig et al., 2016). Therefore, circHIPK3 could bind different miRNAs in bladder cancer cell lines than the ones proposed by the AGO-CLIP data. To evaluate this, we enumerated all possible unique 7-mers (n = 1,032) in the sequence of circHIPK3 (circHIPK3 7-mers). Of these, 148 correspond to target sites for miRNAs (miRNA 7-mers). No circHIPK3 7-mers occur more than three times.

If circHIPK3 functions as a miRNA sponge, we predict that target genes of the miRNA become downregulated upon circHIPK3 KD. We used a computational method, Regmex (Nielsen et al., 2018), to evaluate the enrichment of circHIPK3 7-mers in the 3'UTRs of deregulated genes upon circHIPK3 KD (Figure 2C and Supplementary Table S4). Regmex assigns a positive score to motifs enriched in 3'UTRs of upregulated genes and, correspondingly, a negative score to motifs enriched in 3'UTRs of downregulated genes. In line with perturbed gene expression profiles, the most extreme motif scores were seen 24 hours post-transfection in UMUC3, while motif perturbation in J82 was seen across multiple time points (Supplementary Figure S2D). Although motif scores were slightly lower for miRNA 7-mers than for 7-mers that do not constitute a miRNA target site, the median motif score was above 0 and the effect size small in both cell lines (effect size < 0.25, $P < 0.05$, Wilcoxon Rank Sum Test, Supplementary Figure S2E), indicating that 7-mers constituting miRNA target sites are not overall enriched in downregulated genes. Specifically, the miRNA-7-mer with the most negative motif score in both cell lines constitute a target site for miR-1207-5p, however, miR-1207-5p is not conserved and has no reported functional roles. Taken together,

we found no clear evidence for circHIPK3 to possess miRNA-sponge functions in our experiments, although this has previously been suggested for circHIPK3 in numerous studies (Cai et al., 2019; D. Chen et al., 2019; Chen et al., 2018; X. Chen et al., 2019; Jin et al., 2018; Li et al., 2017b; Zeng et al., 2018; Zheng et al., 2016; Zhou et al., 2021)(Cai et al., 2019; D. Chen et al., 2019; Chen et al., 2018; X. Chen et al., 2019; Jin et al., 2018; Li et al., 2017b; Zeng et al., 2018; Zheng et al., 2016).

A long motif in circHIPK3 is enriched in downregulated genes upon circHIPK3 knockdown

Some circRNAs have been shown to interact with other regulatory factors like RNA binding proteins (RBPs) (Ashwal-Fluss et al., 2014; Du et al., 2016; Hollensen et al., 2020). While miRNAs typically rely on perfect pairing of the miRNA seed region to the target site (Bartel, 2009), RBP binding sites are generally less well-defined as RBP binding often depends on contextual properties, including secondary structure and nucleotide compositions (Dominguez et al., 2018).

We assessed the ability of circHIPK3 to regulate target genes through other regulatory factors than miRNAs by examining the expression correlation of all circHIPK3 7-mers (Figure 2D). Interestingly, the sequence composition of nine of the top ten 7-mers with the most negative motif scores in UMUC3 are shifted by one base or only differ by a single or two bases. These 7-mers are also among the most negatively scoring motifs in J82 cells (Supplementary Figure S2F), suggesting downregulation of motif-containing mRNAs upon circHIPK3 knockdown. Aligning these nine 7-mers to circHIPK3 reveals a larger 11-mer within circHIPK3, AGGCCCCC(A/T)GC (Figure 2E). Allowing up to two base substitutions, we evaluated the presence of the 11-mer motif in the 3'UTR of genes. Interestingly, the motif is enriched in downregulated genes compared to upregulated genes ($P < 1e-39$) and genes that are not differentially expressed (not DE) ($P < 1e-28$) in both UMUC3 (Chi-square Test, Figure 2F) and J82 cells ($P < 0.0001$ for both comparisons, Chi-square Test, Supplementary Figure S2G). This finding was further validated in a third bladder cancer cell line FL3 ($P < 2e-39$ for both comparisons, Chi-square Test, Supplementary Figure S2H).

Our results show that the 7-mers with the most negative motif scores comprise a larger circHIPK3 motif with some variability and suggest that circHIPK3 regulate target genes containing this motif. Thus, we hypothesized that the 11-mer motif in circHIPK3 comprise a binding site for RBPs that may be crucial for its function, potentially by sponging or otherwise regulating specific RBPs.

The 11-mer motif in circHIPK3 constitutes a binding site for IGF2BP2

Next, we evaluated if the 11-mer motif represents a binding site for RNA binding proteins. Since RBP targets are generally poorly defined and several factors influence RBP binding, e.g., contextual features and

expression of cofactors, sequence-based analyses are likely not sufficient to capture true RNA-RBP interactions. Therefore, to identify RBPs that target the 11-mer in circHIPK3, we analyzed a comprehensive eCLIP data set with experimentally defined binding sites of 150 RBPs in HepG2 and K562 cells (Van Nostrand et al., 2016). This revealed high confidence binding sites of several RBPs in the circHIPK3 loci (circHIPK3-RBPs) (Figure 3A). Specifically, binding sites for GRWD1 (HepG2) and IGF2BP2 (K562) overlap the 11-mer.

To interact, both circHIPK3 and RBPs should be expressed in the same subcellular compartments of the cells. We analyzed circHIPK3 localization in total RNA-Seq data sets of fractionated cells (ENCODE Project Consortium, 2012) and found that circHIPK3 is highly expressed in the cytoplasm of HepG2 and K562 cells (Supplementary Figure S3A), which corroborates our previous findings in the bladder cancer cell lines T24, HCV29, and FL3 (Okholm et al., 2017). From immunofluorescence imaging of RBP localization in HepG2 (Van Nostrand et al., 2018), we found that all circHIPK3-RBPs occupy the cytoplasmic fraction of the cell, except for U2AF1 and U2AF2, which span the intron-exon boundary at 3' splice sites and are instrumental for mRNA splicing (Supplementary Table S5). To confirm circHIPK3-RBP interactions, we performed RNA immunoprecipitation (RIP) of identified circHIPK3-RBPs as well as an unrelated control RBP, UCHL5. Using divergent primers targeting the unique backsplice junction of circHIPK3 (Supplementary Table S1), we specifically verified circHIPK3 pull-down in IGF2BP1 and IGF2BP2 RIP experiments by RT-PCR in both cell lines (Figure 3B). Faint bands indicate that circHIPK3 also binds other RBPs albeit with lower affinity. Notably, circHIPK3 interacts weakly with GRWD1 in HepG2 but not in K562, as suggested by the eCLIP data.

Studies have shown that circRNAs that serve as miRNA sponges compete for binding of miRNAs and thereby influence abundance of the corresponding target mRNAs (Hansen et al., 2013). If circHIPK3 competes for RBP binding in a similar manner, we would expect deregulated genes upon circHIPK3 KD to contain RBP binding sites as well. Based on the eCLIP data, we identified all transcriptome-wide binding sites of the RBPs predicted to interact with circHIPK3 (Figure 3A). Focusing on K562-RBPs, we found that while genes that become upregulated upon circHIPK3 knockdown generally contain more circHIPK3-RBP binding sites (Supplementary Figure S3B), IGF2BP2 binding sites are enriched in genes that become downregulated upon circHIPK3 knockdown in both UMUC3, J82, and FL3 (all $P < 0.05$, Chi-square tests, Figure 3C and Supplementary Figure S3C). This suggests that IGF2BP2 may become sequestered by circHIPK3 via the 11-mer motif and allow for stabilization of target mRNAs.

We evaluated how often binding sites of circHIPK3-RBPs overlap the 11-mer motif and found that this is more often the case for IGF2BP2 binding sites than binding sites of the other circHIPK3-RBPs (Supplementary Figure S3D). Furthermore, IGF2BP2-11-mer overlap is enriched in downregulated genes

(P < 0.01, Chi-square test, Supplementary Figure S3E). We found no enrichment of HepG2-RBPs in downregulated genes (Supplementary Figure S3F).

IGF2BP2 knockdown counteracts the effect of circHIPK3 knockdown

If circHIPK3 functions as a ceRNA for IGF2BP2, downregulation of IGF2BP2 would be expected to have the opposite effect of circHIPK3 KD on target genes. Therefore, we evaluated if genes affected by circHIPK3 KD in bladder cancer cell lines are also affected in ENCODE experiments of IGF2BP2 KD and found a significant overlap of altered genes (P < 2.2e-16, Fisher's Exact Tests, Figure 3D and Supplementary Figure S3G). The overlap was largest for downregulated genes in both UMUC3 (P < 2e-15, Chi-square test, Figure 3E) and J82 cells (P < 0.001, Chi-square test, Supplementary Figure S3H). Interestingly, the genes that are downregulated by circHIPK3 KD are primarily upregulated upon IGF2BP2 KD suggesting an mRNA destabilization function of IGF2BP2 (Figure 3F and Supplementary Figure S3I). RegMex analysis of all possible 7-mers showed that the 7-mers that constitute the 11-mer motif in circHIPK3 are among the highest scoring motifs upon IGF2BP2 KD in K562 (Figure 3G). Accordingly, the 11-mer motif and IGF2BP2 binding sites are highly enriched in upregulated genes (P < 6e-19, Chi-square test, Figure 3H, I). While there were no IGF2BP2 eCLIP data available in HepG2, analogous enrichments were observed in upregulated genes upon IGF2BP2 KD in HepG2, suggesting a similar mechanism in HepG2 (Supplementary Figure S3J, K).

Taken together, our results suggest that the 11-mer motif in circHIPK3 constitutes a binding site for IGF2BP2 and that IGF2BP2 binding sites and the 11-mer motif are enriched in downregulated mRNA transcripts upon circHIPK3 KD, while the opposite is observed upon IGF2BP2 KD. Furthermore, we find that IGF2BP2 KD counteracts the effect of circHIPK3 KD on target mRNAs, indicating that the circHIPK3-IGF2BP2 complex positively regulates mRNAs targeted by IGF2BP2.

circHIPK3 interacts with IGF2BP2 and affects genes controlling cell cycle progression

To verify circHIPK3-IGF2BP2 interactions in bladder cancer we performed RIP experiments in the bladder cancer cell line FL3 for the two RBPs that have binding sites across the 11-mer: IGF2BP2 and GRWD1. Additionally, we included IGF2BP1, which has binding sites across one of the downregulated 7-mers constituting the 11-mer motif and several binding sites in the 3'UTR of the parent gene. Consistent with our findings above, circHIPK3 interacts strongly with IGF2BP2 in bladder cancer cells but not with GRWD1 (Figure 4A and Supplementary Figure S4A). CircHIPK3 also interacts with IGF2BP1, albeit with lower binding affinity than observed for IGF2BP2. RIP experiments of IGF2BP2 upon circHIPK3 overexpression (OE), using Laccase2-driven circRNA expression (Kramer et al., 2015), confirmed tight and specific binding between circHIPK3 and IGF2BP2 (P < 0.002, T-test, Figure 4B).

To confirm regulatory implications of circHIPK3-IGF2BP2 binding in bladder cancer, we evaluated the presence of the 11-mer motif in deregulated genes in previously performed experiments of IGF2BP2 KD in bladder cancer cells (Okholm et al., 2020). Consistent with our observations in K562 and HepG2, we found that the 11-mer motif is enriched in upregulated genes compared to downregulated genes upon IGF2BP2 KD in UMUC3 ($P < 7e-04$, Chi-square Test, Figure 4C) and J82 cells ($P < 0.002$, Chi-square Test, Supplementary Figure S4B).

IGF2BP2 is a post-transcriptional regulator involved in many mRNA processes, including mRNA stability, transportation, localization, and translation (reviewed in (Cao et al., 2018)). Using a peptide spanning the third and fourth K-homology domain of IGF2BP2 (KH34 domains), a recent SELEX study identified a bipartite recognition element (RE), which confers high affinity RNA binding: CUCAC-(N10-15)-(A/U)-GG-(A/U) (Biswas et al., 2019). Here, we found a significant overlap between genes containing the 11-mer motif and the IGF2BP2 KH34 RE in their 3'UTRs ($P < 2.2e-16$, Fisher's Exact Test, Supplementary Figure S4C). Additionally, the IGF2BP2 KH34 RE is enriched in downregulated genes upon circHIPK3 KD in UMUC3 cells ($P < 0.002$, Chi-square Test, Supplementary Figure S4D).

Studies have shown that IGF2BP2 is overexpressed in many cancers, correlates with progression, and promotes cancer cell proliferation and migration (reviewed in (Cao et al., 2018)), while we and others have reported that circHIPK3 is downregulated in bladder cancer (Okholm et al., 2017) and possess tumor suppressor functions (Teng et al., 2019).

To assess the functional impact of circHIPK3 and IGF2BP2 in bladder carcinogenesis, we performed gene set enrichment analysis of 50 hallmarks of cancer gene sets (Liberzon et al., 2015) and found that several proliferation pathways, e.g., E2F targets and G2/M checkpoint, are upregulated upon circHIPK3 KD in both UMUC3 ($P < 0.05$, T-test, Figure 4D) and J82 cells ($P < 0.05$, T-test, Supplementary Figure S4E). In contrast, these pathways are downregulated upon IGF2BP2 knockdown ($P < 0.05$, T-test, Figure 4E, Supplementary Figure S4F), suggesting that the circHIPK3-IGF2BP2-target axis may affect cell cycle progression. However, using an EdU-flowcytometry assay (Supplementary Figure S4G), we found only subtle differences in the distribution between G1-, S- or G2/M-phases of the cell-cycle, when measuring these parameters upon knockdown of either circHIPK3 or IGF2BP2 in FL3 cells.

We observed no difference in *IGF2BP2* gene expression between circHIPK3 KD and scr samples in UMUC3 ($P = 0.46$, Likelihood Ratio Test, Supplementary Figure S4H) and only a slight, though significant, difference

at later time points in J82 cells (P = 0.004, Likelihood Ratio Test, Supplementary Figure S4H), consistent with no direct effect of circHIPK3 on IGF2BP2 levels.

Taken together, our results suggest that circHIPK3 possesses important regulatory roles in bladder cancer by interacting with the oncogenic RBP IGF2BP2, which potentially affects proliferative pathways.

circHIPK3 functions as a ceRNA for IGF2BP2

To further elucidate the functional role of circHIPK3-IGF2BP2 interactions in bladder cancer, we identified specific targets of circHIPK3 and IGF2BP2. An interesting example is the *STAT3* gene, which contains the 11-mer motif in its 3'UTR (allowing two mismatches), becomes downregulated upon circHIPK3 KD in UMUC3 (P_24h = 3e-04, Wald Test, Figure 5A) and J82 cells (P_48h = 0.01, Wald Test, Supplementary Figure S5A), while it is upregulated upon IGF2BP2 depletion in both cell lines (P < 0.001, Wald Test, Figure 5B and Supplementary Figure S5B). IGF2BP2 RIP experiments confirmed a *STAT3*-IGF2BP2 interaction in bladder cancer cells (P = 0.04, T-test, Figure 5C). Interestingly, upon immunoprecipitation of IGF2BP2, we observed significantly more *STAT3* mRNA binding after circHIPK3 depletion (48h), supporting a regulatory role of circHIPK3 as a ceRNA (P < 0.007, T-test, Figure 5D and Supplementary Figure S5C). We analyzed the IGF2BP2 binding affinity of three other transcripts (*NEU*, *SLC7A5*, and *TRAPPC9*) containing the 11-mer motifs in their 3'UTRs, which are also downregulated upon circHIPK3 KD, and conversely upregulated upon IGF2BP2 KD, and found a similar significant enrichment upon circHIPK3 KD (P < 0.03, T-test, Supplementary Figure S5D). The relative IGF2BP2 binding of circHIPK3 (IP/INPUT ratio) was not affected by circHIPK3 KD, suggesting that excess IGF2BP2 does not affect the amount bound per circHIPK3 molecule (P = 0.35, T-test, Supplementary Figure S5E).

To assess the significance of the 11-mer motif in circHIPK3-IGF2BP2 interactions, we initially transfected FL3 cells with antisense oligonucleotides (ASO) specifically targeting the 11-mer motif in circHIPK3 but not IGF2BP2 target genes, where up to two mismatches were allowed. Although not statistically significant, FLAG-IGF2BP2 binding affinity to *STAT3* is 2.9 times higher in the presence of the 11-mer ASO compared to control samples (P = 0.078, T-test, Figure 5E). Additionally, the relative circHIPK3-IGF2BP2 binding affinity is 30% lower in 11-mer ASO samples (P = 0.27, T-test, Supplementary Figure S5F). Next, to directly evaluate the importance of the 11-mer motif in IGF2BP2 binding of circHIPK3, we mutated the sequence in a pcDNA3.1+-Laccase circRNA expression construct and transfected HEK293 cells with either wildtype or mutant circHIPK3. Interestingly, we found that binding of IGF2BP2 was reduced significantly (~2-fold, P = 0.028) when the 11-mer motif had been scrambled, while binding of *STAT3* mRNA to IGF2BP2 was increased (~7-fold, P = 0.036) (Figure 5F).

It is well established that STAT3 regulates *TP53* expression by repressing its transcription (Niu et al., 2005). If the observed downregulation of *STAT3*, upon knockdown of circHIPK3 is functionally relevant, we predict that *TP53* mRNA and protein levels would rise. Consistent with our predictions, *TP53* mRNA is upregulated upon circHIPK3 KD at later time points in UMUC3 cells (P = 1.7e-05, Wald Test, Figure 5G). Additionally, western blot analysis confirmed that circHIPK3 KD indeed leads to downregulation of STAT3 protein, while p53 protein levels rise (P < 0.05, T-test, Figure 5H). Finally, immunofluorescence staining showed that nuclear and overall STAT3 protein levels are significantly lower upon circHIPK3 KD (P < 0.003, Wilcoxon Rank Sum Test, Figure 5I), while nuclear p53 protein levels are higher (P = 1.7e-30, Wilcoxon Rank Sum, Test, Figure 5J).

Our experiments suggest that circHIPK3 functions as a ceRNA for IGF2BP2 most likely by specifically interacting with IGF2BP2 at the 11-mer binding site (and potentially secondary sites), which affects IGF2BP2 binding affinity to other target genes. This in turn represses STAT3 expression while upregulating p53.

The ratio between circHIPK3 and IGF2BP2 cannot explain a simple competing endogenous RNA mode of action

To function as a ceRNA through direct binding, circHIPK3 levels must at least roughly match those of IGF2BP2. To directly assess the number of circHIPK3 and IGF2BP2 protein molecules per cell, we first designed a backsplicing junction DNA template for generation of qPCR standard curves (Figure 6A and Supplementary Figure S6A-B). Isolating RNA from a known number of cells then allowed us to estimate the number of circHIPK3 molecules to ~45 (FL3), ~51 (J82), and ~40 (UMUC3) per cell (Figure 6B and Supplementary Figure S6C). Assuming similar primer efficiencies in qPCR reactions, these results correlate well with obtained circHIPK3/mRNA ratios estimated from RT-qPCR reactions amplifying mRNAs/cDNAs from housekeeping genes (e.g. GAPDH or ß-Actin) (data not shown). As IGF2BP2 is generally an abundant ubiquitous protein with an estimated ~10.000 copies per cell in HeLa cells (Bekker-Jensen et al., 2017), we next wanted to quantify IGF2BP2 content in FL3, J82 and UMUC3 cells. To this end, we expressed and purified Twin-streptagged-IGF2BP2 from HEK293 Flp-in T-Rex cells and quantified the yield by Coomassie staining by comparing to known amounts of bovine serum albumin (Figure 6A and Supplementary Figure S6D). From known amounts of purified IGF2BP2 in western blots with lysates from a known number of bladder cancer cells, we estimate that each cell contains ~216.000 (FL3), ~127.500 (J82), and ~261.000 (UMUC3) IGF2BP2 molecules (Figure 6B and Supplementary Figure S6E-F).

These results suggest that IGF2BP2 significantly outnumbers circHIPK3 by several orders of magnitude and hence that a competing endogenous RNA function for circHIPK3, in its simplest form, is unlikely. How can circHIPK3 then affect IGF2BP2 function? One possibility is that circHIPK3 nucleates numerous IGF2BP2 proteins on each circRNA molecule potentially starting at the 11-mer motif. To test this possibility, we performed immunofluorescence targeting endogenous IGF2BP2 in UMUC3 cells with circHIPK3 overexpression and hypothesized that if circHIPK3 nucleates IGF2BP2 molecules, it would be visible as cellular condensates. Indeed, we observed a significantly higher prevalence of large cytoplasmic condensates in cells expressing high levels of circHIPK3 compared to controls (80% vs 29%) (Figure 6C-D), which may contribute to IGF2BP2 accumulation by much fewer circHIPK3 molecules. Our results suggest that while circHIPK3 is unable to sequester IGF2BP2 in a 1:1 fashion, due to the large excess of IGF2BP2, the circRNA may efficiently nucleate numerous IGF2BP2 molecules, which are then growing into visible phase-separated condensates upon circHIPK3 overexpression.

circHIPK3 expression is positively correlated with overall survival

Studies have shown that *HIPK3* possesses oncogenic features (Curtin and Cotter, 2004), while the derived circHIPK3 is downregulated in cancer and possesses tumor suppressor functions (Teng et al., 2019).
We extended a previously described cohort (Hedegaard et al., 2016; Okholm et al., 2017; Strandgaard et al., 2022; Taber et al., 2020), and analyzed the expression of circHIPK3 in more than 500 samples from patients with bladder cancer. The expression of circHIPK3 is significantly upregulated in patients with non-muscle invasive bladder cancer (NMIBC) compared to patients with muscle invasive bladder cancer (MIBC) ($P < 3.3e-09$, Wilcoxon Rank Sum Test, Figure 6E).

We previously reported that circHIPK3 expression correlates positively with bladder cancer progression-free survival independently of the parent gene (Okholm et al., 2017). In a newly generated and hence independent cohort of patients with NMIBC ($n = 56$) (Strandgaard et al., 2022), we found that circHIPK3 levels are positively associated with overall survival ($P < 0.009$, Log-Rank Test, Figure 6F). The association was less pronounced for *HIPK3* levels ($P = 0.052$, Log-Rank Test, Figure 6F). Our clinical analyses confirm that high circHIPK3 expression levels are correlated with good clinical outcomes.

In conclusion, our study provides evidence that the most prominent alpha circRNA, circHIPK3, interacts with IGF2BP2 to regulate the expression of hundreds of genes containing the 11-mer motif in its 3'-UTR, including *STAT3*. Since circHIPK3 expression induces IGF2BP2 condensates, we suggest that such protein-RNA nucleation events may, also at the submicroscopic level, enhance the potential for circHIPK3 to titrate out multiple IGF2BP2 molecules even though it is found in low numbers within cells (Figure 6G). Alternatively, circHIPK3 may recruit unknown enzymatic modifiers of IGF2BP2, which in turn could affect

its activity. Both scenarios could explain the functional output from a small number of circHIPK3 molecules in cells and await further experimentation.

Discussion

In this study, we characterized the functional role of circHIPK3 in bladder cancer cells by perturbation studies, comprehensive analyses of large genomics data sets, and experimental hypothesis evaluation. In contrast to previous studies suggesting a role for circHIPK3 as a miRNA sponge (reviewed in (Zhou et al. 2021)), our results indicate that manipulation of circHIPK3 levels does not affect genes targeted by miRNAs with seed region binding sites also present within circHIPK3. Specifically, based on AGO-CLIP data profiling miRNA target sites from 32 independent experiments, we found no indication that circHIPK3 functions as a miRNA sponge. In addition, since most reports claiming circRNA-specific miRNA-sponging are based on inflated circRNA expression, use of miRNA mimics and reporter assays, a conclusion that circRNAs are generally functioning as miRNA sponges should be taken with some caution. Indeed, our estimation of less than a hundred circHIPK3 molecules per cell, questions the miRNA sponge hypothesis, as this entails specific sequence-specific binding of miRNA-guided Argonaute proteins. It should be noted however, that miRNA expression and target sites are often tissue specific and the lack of bladder cancer tissue in the AGO-CLIP data could influence our findings. Nonetheless, several studies argue against a general miRNA-sponge function for most circRNAs (Guo et al., 2014; Militello et al., 2017).

Here, we identified an 11-mer motif present in the sequence of circHIPK3 that constitutes a binding site for IGF2BP2 in RIP assays. IGF2BP2 is a multifunctional RBP that affects hundreds of mRNAs and various RNA processes (Cao et al., 2018). It belongs to a family of three conserved RBPs (IGF2BP1, IGF2BP2, and IGF2BP3), that are important modulators of gene expression, especially during development, and contain highly similar RNA binding domains; two N-terminal RNA-recognition motifs (RRM) and four C-terminal hnRNPK Homology (KH) domains (reviewed in (Bell et al., 2013)). Recently, Biswas et al. identified a consensus binding motif for IGF2BP2, CUCAC-(N10-15)-(A/U)-GG-(A/U), containing the core elements CA and GG required for specific binding of the IGF2BP2 KH3 and KH4 domain, respectively (Biswas et al., 2019). Although in reverse order, both core REs are present in the 11-mer motif, and we observed a significant overlap between genes containing the consensus sequence and the 11-mer motif (allowing two base substitutions).

While some RBPs contain well-established binding domains and bind to well-defined motifs, most RBPs rely on sequence context and structural features (Dominguez et al., 2018; Doyle and Kiebler, 2012), suggesting that RNA structure could explain the discrepancies observed here. Additionally, IGF2BP2 might

recognize the swapped orientation of the REs, as observed for IGF2BP1 (Patel et al., 2012). Although earlier work have shown that the KH34 domains are the major RNA binding domains of all IGF2BPs (Chao et al., 2010; Patel et al., 2012), more recent studies demonstrate that additional RNA binding domains are involved in RNA recognition and distinction between targets (Schneider et al., 2019; Wächter et al., 2013), complicating the establishment of a consensus motif and binding sites predictions. Furthermore, other external factors likely influence binding affinity and preference, and cooperativity both intra- and intermolecularly might be restricted to certain cells, tissues, and environment, as observed for circFoxo3, which interacts with different RBPs depending on the biological setting (Du et al., 2017a, 2017b, 2016). Since nucleotide sequence is thus insufficient to fully predict RBP binding sites, we used large-scale motif-enrichment analysis of RNA-Seq data and experimentally validated RBP target sites from eCLIP data to identify circHIPK3-IGF2BP2 mRNA targets and regulatory tendencies that are consistent across ENCODE and BC cell lines. While we do not attempt to define a new recognition motif for IGF2BP2, our results are in line with previous work describing the necessity of two appropriately spaced REs that are crucial for binding by all IGF2BPs (Biswas et al., 2019; Patel et al., 2012; Schneider et al., 2019).

While circHIPK3 is frequently downregulated in cancer and positively correlated to good clinical outcomes (Li et al., 2017a; Okholm et al., 2017), IGF2BP2 enhances genomic instability, is upregulated in cancer, and involved in cancer progression (Barghash et al., 2016; Kessler et al., 2015; Ye et al., 2016). Our results show that IGF2BP2 knockdown counteracts the effect of circHIPK3 knockdown on target genes and proliferation pathways, indicating that circHIPK3 might function as a reservoir for IGF2BP2. By focusing on circHIPK3 and IGF2BP2 regulated targets, we provide evidence that circHIPK3-IGF2BP2 interactions affect the expression of *STAT3*. *STAT3* is persistently activated in a wide diversity of cancers and is positively associated with cell proliferation, cell survival, invasion, and metastasis (reviewed in (Kamran et al., 2013)). By conducting RIP experiments after circHIPK3 knockdown, we observed a stronger IGF2BP2 association with *STAT3* mRNA (~7-fold enrichment), along with a concomitant *STAT3* mRNA and protein downregulation. Consistent with this finding, a study recently used an mRNA-RBP tethering reporter assay and found a general mRNA destabilizing effect of an IGF2BP2-mRNA association (Luo et al., 2020). This strong association is dependent on the 11-mer sequence, since a circHIPK3 variant with a mutated motif decreased binding significantly in RIP assays. These results point to a regulatory role of circHIPK3 as a ceRNA for IGF2BP2 binding.

Surprisingly, our stoichiometric analysis of copy-numbers of circHIPK3 and IGF2BP2 in BC cell lines are not readily compatible with a simple ceRNA hypothesis, since IGF2BP2 outnumbers circHIPK3 by orders of magnitude. However, we demonstrated that exogenous expression of circHIPK3 produces IGF2BP2-containing condensates, suggesting that circHIPK3 can initiate a nucleation process that will likely

condensate/phase separate multiple proteins, which could at least partly explain how a few cellular circHIPK3 molecules can sequester many IGF2BP2 proteins. Another possibility is that the cell only has a limited number of available IGF2BP2 molecules that can dynamically interact with either circHIPK3 or other mRNA targets, leaving enough circles for a proper ceRNA function. More speculatively, circHIPK3 may scaffold an enzymatic complex that can modify IGF2BP2 post-translationally and change its function (e.g. RNA-binding, localization or interaction with other factors) - a scenario where a few circRNAs can change the function of numerous RBPs. Still, the regulatory interplay could be more complex and circHIPK3 could act as a scaffold that facilitates interaction between several RBPs or as part of a larger functional unit. While we provide evidence that circHIPK3 interacts with IGF2BP2 to regulate IGF2BP2 target genes, the detailed molecular mechanism remains to be fully elucidated.

Taken together, we identified an 11-mer motif within circHIPK3 that is important for the regulation of multiple cancer-related genes via IGF2BP2 binding. We propose a ceRNA model where a few circHIPK3 molecules can sequester multiple IGF2BP2 proteins by inducing their condensation.

Materials and methods

Plasmids

To generate plasmids for exogenous expression of Twin-Strep-tagged IGF2BP1 (transcript variant 1, NM_006546.4), GRWD1 (transcript variant 1, NM_031485.4) or GFP, cDNA libraries from HEK293 and HeLa cells were used as templates for standard PCR reactions using primers containing BamHI-NotI restrictions sites that amplify the CDS. XhoI-NotI restriction sites were used for cloning IGF2BP2 (transript variant 1, NM_006548.6). The amplicons were inserted into pcDNA5-Twin-Streptag. pcDNA5-Twin-Streptag was made by inserting a PCR product (containing a Twin-Streptag followed by a TEV protease site) using a pDSG-IBA-Twin-Strep-Tag vector as template (IBA-lifesciences) into the HindIII-ApaI site of pcDNA5-FRT-TO (Invitrogen).

To generate plasmids for exogenous expression of circHIPK3, the second exon encoding circHIPK3 was PCR-amplified from cDNA prepared from RNA isolated from HEK293/HeLa cells and inserted into PacI/SacII-digested pcDNA3.1(+)-Laccase2-MCS-exon-vector (Kramer et al., 2015) to generate the pcDNA3.1(+)-Laccase2-circHIPK3 vector. pcDNA3.1(+) Laccase2-MCS-Exon-Vector was a gift from Jeremy Wilusz (Addgene plasmid # 69893; http://n2t.net/addgene:69893 ; RRID:Addgene_69893). The 11-mer motif mutation within pcDNA3.1(+)-Laccase2-circHIPK3 vector was generated by overlap-extension PCR including overlapping primers containing the mutated 11-mer sequence.

HIPK3 Scr 11-mer F: TTGCATTTCCTAGAAGACGGCTACAGATGTGGATTGAAGCGCAAGAGTG
HIPK3 Scr 11-mer R: CTTCAATCCACATCTGTAGCCGTCTTCTAGGAAATGCAATCTGTTTCGC
Outer amplification primers: Laccase F: CAACGAATCTAGTATACCTT. Laccase R: ACAATAGTATAGAAACTGCC. All constructs were verified by sanger sequencing.

circHIPK3 expression in bladder cancer cell lines

Whole transcriptome RNA-Seq data from eleven bladder cancer cell lines (HCV29, HT1197, Nhutert, RT4, FL3, HT-1376, J82, LUL2, SLT4, T24, and UMUC3) were generated and mapped in a previous study (Hedegaard et al., 2016). We used the CIRI2 pipeline (v2.0.6)(Gao et al., 2018) to profile the expression of circHIPK3. Before running the CIRI2 pipeline, we trimmed reads with trim_galore (v0.4.1). We aligned the reads to the human genome (hg19) using bwa (v0.7.15) and samtools (v1.3). The CIRI2 pipeline was run with a GTF file (hg19) to annotate the overlapping gene of the circRNAs.

Time-series siRNA-mediated knockdown

A total of $3x10^5$ cells (UMUC3 or J82) were seeded per well in 6-well plates (2 ml per well). Twenty-four hours later, cells were transfected with an siRNA targeting the unique backsplice junction of circHIPK3 (Sense: GGUACUACAG|GUAUGGCCU[dTdT] and antisense: AGGCCAUAC|CUGUAGUACC[dTdT], "|" denote backsplice junction) using SilentFect (Bio-Rad) in biological triplicates: 100 µl OptiMEM (Gibco) + 2 µl SilentFect (Bio-Rad) was incubated at room temperature (RT) for 5 min before mixing with 2 µl siRNA (20 µM stock) pre-diluted in 100 µl OptiMEM. After gentle mixing by pipetting, the solution was incubated for 20 mins at RT prior to dropwise addition to cells (final concentration of 20 nM). Cells were incubated for 12, 24, 48 or 72 hours before harvest by addition of 1 ml Trizol (Thermo Fisher Scientific) to each well followed by mixing. As negative control, we used MISSION siRNA Universal Negative Control #2 (Sigma-Aldrich). For knockdown of mRNAs encoding RNA-binding proteins, UMUC3 or J82 cells were transfected as indicated above and incubated for 56 hrs. Harvested RNA was purified according to manufacturer's protocol with an additional 400 µl chloroform extraction to increase RNA quality. A parallel set of identically treated samples were used to harvest cells for protein lysates by addition of 300 µl 2 X SDS sample buffer [4% SDS, 20% glycerol, 10% 2-mercaptoethanol, 0.004% bromophenol blue and 0.125 M Tris-HCl, pH 6.8] per well followed by incubation at 90°C for 8-10 min until all cell material is dissolved.

Knockdown efficiency was assessed by qRT-PCR. RNA samples were treated with DNase I (Thermo Fisher Scientific) according to the manufacturer's protocol. First-strand cDNA synthesis was carried out using the Maxima First Strand cDNA synthesis Kit for qPCR (Thermo Fisher Scientific) according to the manufacturer's protocol. qPCR reactions were prepared using gene-specific primers and Platinum SYBR Green qPCR Supermix-UDG (Thermo Fisher Scientific) according to the manufacturer's protocol. An

AriaMx Real-time PCR System (Agilent Technologies) was used for quantification of RNA levels and the X_0 method was used for calculations of relative RNA levels (Thomsen et al., 2010) normalized to GAPDH or beta-actin mRNA as indicated.

We quantified mRNA expression using QuantSeq (Moll et al., 2014). Initial quality control was performed using Dropsense96 (Trinean) and 2100 Bioanalyzer (Agilent). Sequencing libraries were generated using the QuantSeq 3'mRNA Library Prep Kit Protocol (Lexogen). We used 1,000 ng RNA as input and 11 PCR cycles were applied. Library sizes were measured on LabChip GX (PerkinElmer) or Bioanalyzer (Agilent) while concentrations were measured on Qubit 3.0 (invitrogen). Finally, we applied the KAPA Library Quantification Kit (Kapa Biosystems) to ensure library quality for sequencing.

The raw reads were converted to fastq format and demultiplexed using Illumina's bcl2fastq v2.20.0.422 and library adapters were removed from the read pairs (trim_galore v0.4.1). Reads were mapped to the human genome (hg19) using TopHat2 (version 2.1.1) (Kim et al., 2013) and Bowtie2 (version 2.1.0.0) (Langmead and Salzberg, 2012), and Cufflinks (v2.1.1) (Trapnell et al., 2010) and HTSeq (v0.6.1p1) (Anders et al., 2015) were used to estimate the transcripts abundance using transcript information from GENCODE v19. Samtools (v1.3) (Li et al., 2009) and Picard (v2.0.1) were used for quality control and statistics.

Differential expression analysis of time-series experiments

For circHIPK3 knockdown experiments, each cell line was analyzed separately. For the differential expression analysis, we excluded protein coding genes with a maximum expression < 1 RPM and the 25% lowest-variance genes across samples. We used DESeq2 on raw counts to identify genes that are differentially expressed between conditions. Briefly, DESeq2 uses a likelihood ratio test to identify genes that show a different expression profile over time across conditions as well as the genes that show any difference across conditions (Love et al., 2014). This approach will not identify genes that move up or down over time in the same way in both conditions. Thus, we used the built-in Wald Test of DESeq2 to evaluate gene expression differences between conditions at each time point.

miRNA expression profiling by NanoString nCounter

The nCounter Human v3 miRNA panel (NanoString Technologies), which targets 799 miRNAs, was used for miRNA profiling in the cell line J82 according to the manufacturer's instructions. In brief, 100 ng of total RNA input and a hybridization time of 20 hours was used. The products were analyzed on the nCounter™ SPRINT (NanoString Technologies) platform. The raw data were processed using the nSOLVER 4.0 software (NanoString Technologies); first, a background subtraction was performed using the max of

negative controls, and then positive control normalization was performed using the geometric mean of all positive controls. Finally, a second normalization was performed using the expression levels of *GAPDH*.

AGO-CLIP data of miRNA target sites

Generation of AGO-CLIP data was described previously (Hamilton et al., 2016, 2013; Okholm et al., 2017). To identify miRNA targets in the sequence of circHIPK3, we restricted our analysis to conserved miRNAs as this reduces false positives as described in (Okholm et al., 2017). We evaluated the presence of miR-1207-5p in the entire AGO-CLIP data with no restriction on conservation.

Identification of motifs in deregulated genes

We used Regmex (Nielsen et al., 2018) to identify motifs enriched in 3'UTRs of down- and upregulated genes. For circHIPK3 knockdown experiments, we ranked genes according to a differential expression score calculated as the -log10(P-value) that is signed according to whether the genes were upregulated (positive) or downregulated (negative). The score was calculated on gene regulation 24 hours post-transfection. Regmex was run with all 1,032 unique 7-mers found in circHIPK3.
For ENCODE IGF2BP2 KD experiments, genes were ranked by the fold change expression between KD and control samples, and Regmex was run with all possible 7-mers.

circHIPK3 expression in HepG2 and K562 cells

We downloaded all total and fractionated samples for the cell lines HepG2 (n = 5) and K562 (n = 11) from ENCODE (https://www.encodeproject.org/) as described in (Okholm et al., 2020). CircHIPK3 expression was profiled using the CIRI2 pipeline as described above for the bladder cancer cell lines and in (Okholm et al., 2020).

eCLIP of HepG2 and K562

The eCLIP data of RNA binding protein (RBP) targets from HepG2 and K562 was obtained from (Van Nostrand et al., 2016) and processed as described in a previous study (Okholm et al., 2020). Briefly, we only considered RBP binding sites that are at least 8-fold enriched in the RBP immunoprecipitation under consideration compared to a size-matched input control. We restricted our analysis to RBPs with binding sites in exonic regions from known and novel protein-coding transcripts (hg19). RBP binding sites < 4 bp or an exon-RBP binding site overlap < 4 bp were disregarded. When analyzing enrichment of RBP binding sites in genes affected by circHIPK3 KD, only the genes expressed in the ENCODE cell line and the bladder cancer cell line under consideration were evaluated.

RNA immunoprecipitation

The RNA immunoprecipitation experiments were performed as described in (Okholm et al., 2020). The following antibodies were obtained for RNA immunoprecipitation in HepG2 and K562 cells: GRDW1 (Bethyl A301-576A Lot 1), UCHL5 (Bethyl A304-099A Lot 1), YBX3 (A303-070A Lot 1), IGF2BP1 (MBL RN007P Lot 004), and IGF2BP2 (MBL RN008P Lot 005). Rabbit IgG Isotype Control (Invitrogen Cat# 02-6102) was used as negative control. K562 cells (American Type Culture Collection (ATCC®), CCL-243™) were grown in RPIM 1640 media (Gibco™, Life Technologies, 11875119) with 10% FBS (Gibco™ Life Technologies, 26140079). HepG2 cells (American Type Culture Collection (ATCC®), HB-8065) were grown in DMEM (Gibco™, Thermo Fisher Scientific) with 10% FBS (Gibco™ Life Technologies, 26140079). 20 x 10^6 snap frozen cells were lysed in 1ml of iCLIP Lysis Buffer (50mM Tris-HCl pH 7.4, 100mM NaCl, 1% NP-40 (Igepal CA630), 0.1% SDS, 0.5% Sodium Deoxycholate) with 5.5µl Protease Inhibitor Cocktail Set III EDTA Free (EMD Millipore Corp.539134-1ML) and 11µl Murine RNase Inhibitor (New England BioLabs Inc® M0314L) for 15 mins and were then centrifuged at 20,000 g for 20 mins at 4°C. The supernatant was placed in a solution containing specific primary(10µg)/secondary(1.25g) (anti-Rabbit magnetic DynaBeads, Invitrogen, 11204) antibody-antibody (incubated on a rotator at 25°C for 45 minutes) to immunoprecipitate overnight on a rotator at 4°C. The RNA-RBP pull down was then purified by stringently washing with NET-2 wash buffer (5mM Tris-HCl pH 7.5, 150mM NaCl, 0.1% Triton X 100). Isolation of RNA from the RNA-RBP complexes was accomplished with the addition of TRIzol™ Reagent (Invitrogen™, 15596018) followed by the Direct-zol™ RNA MiniPrep (Zymo Research Cat No. R2052). Isolated RNA was reverse transcribed with SuperScript™ III First Strand Synthesis System (Invitrogen™ 18080051) using Random Hexamers (ThermoFisher Scientific N8080127) and circRNAs were amplified using GoTaq® DNA Polymerase (Promega, M3005); 4µl of 1:2.5 diluted cDNA, 1µl of each primer for 34 cycles at the following conditions: strand separation at 95°C for 30s, primer hybridization at 55°C for 30s, and elongation at 72°C for 20s followed by a final elongation step at 72°C for 5 min. Amplicons were run on a 3% Agarose Gel at 135V for 35 minutes at 4°C alongside a 50bp ladder marker. CircHIPK3 bound by RBPs was identified on the gel based on amplicon size.

For RNA immunoprecipitation in bladder cancer, 1.65x10^6 FL3 cells were seeded in a P10 dish and transfected 24 hrs later with 12 µg total DNA (2 µg Twin-Strep-Tag-RBP expression vector and 10 µg pcDNA3 PL). Forty-eight hours later cells were washed with PBS and placed on ice. One milliliter cold lysis buffer (50 mM TRIS-HCl pH 7.5, 10 mM NaCl, 5 mM $MgCl_2$, 0.5% Triton-X 100, 1% Hexane-1,6-diol, 1 pill Complete - protease inhibitor EDTA-free (Roche) pr. 10 mL) was added per plate and cells were scraped off and transferred to an Eppendorf tube. Samples were mixed and spun (13,000 RPM, 15 min at 4ºC) and 50 µL supernatant was transferred to 50 µL SDS load buffer while 100 µL supernatant was transferred to 0.5 mL Trizol (INPUT). 800 µL supernatant was incubated with pre-equilibrated MagStrep type 3 XT beads (IBA Life sciences) and rotated for ≥2 hrs at 4ºC. 500 µL supernatant was collected (FT) and samples were

washed 4x with 1.5 mL WASH1 buffer (10 mM TRIS-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 0.1% Triton-X 100) and 2x with WASH2 buffer (10 mM TRIS-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$). During the last wash sample beads were divided 1:10 (RNA:protein) and added 40 µL SDS load buffer or 0.5 mL Trizol (IP). The incubation time for siRNA KD was 48 hrs. INPUT, FT, and IP samples were analyzed with western blotting and RT-qPCR.

IGF2BP2 knockdown

RBP knockdown samples in K562 and HepG2 were produced for the ENCODE project by Brenton Graveley's Lab, UConn (ENCODE Project Consortium, 2012). RNA-Seq data of polyadenylated transcripts were obtained from ENCODE (https://www.encodeproject.org/). When comparing genes affected by IGF2BP2 KD and circHIPK3 KD, only the genes expressed in the ENCODE cell line knockdown experiment and the bladder cancer cell line under consideration were evaluated. IGF2BP2 knockdown in the bladder cancer cell lines UMUC3 and J82 were described in a previous study(Okholm et al., 2020).

Pathway analysis

We used the R package gage (Luo et al., 2009) for pathway analyses of 50 hallmarks of cancer gene sets from The Molecular Signatures Database (MSigDB) (Liberzon et al., 2015). For exploratory pathway analyses upon circHIPK3 KD in UMUC3, we used a significance level of FDR < 0.1. To verify specific observations upon circHIPK3 KD in J82 and IGF2BP2 KD in UMUC3 and J82, we used a significance level of $P < 0.05$.

EdU Click-IT proliferation assay

The cell cycle assay was performed as described in detail elsewhere (Hollensen et al., 2020). Briefly, labeling of newly synthesized DNA was carried out using Click-iT Plus EdU Alexa Flour Flow Cytometry Assay Kit (Thermo Scientific) according to manufacturer's protocol. Notably, the cell culture medium of FL3 cells cultured in 6-cm dishes was supplemented with 10 µM EdU for 90 minutes. To stain total DNA, cells with already detected EdU were resuspended in 300 µL 1x ClickiT saponin-based permeabilization and wash reagent from the assay kit. Subsequently, RNase A was added to a final concentration of 0.2 mg/mL. After 5 minutes of incubation at RT, propidium iodide was added to a final concentration of 5 µg/mL and the cells were incubated for 30 minutes at RT. Incorporated EdU and total DNA levels were analyzed on a NovoCyte Quanteon Flow Cytometer System (Agilent). The incubation time for siRNA KD was 48 hrs. Data analysis was carried out in the FLOWJO software (BD Biosciences).

Immunofluorescence microscopy

Autoclaved 18 mm coverslips were placed in 12-well plates and covered with 0.01% poly-L-lysine for 2-5 minutes and allowed to dry. The dry poly-L-lysine coated coverslips were then placed under UV light for 10 minutes before sealing with Parafilm. The coated coverslips were stored at RT until use. siRNA-transfected cells grown on poly-L-lysine coated coverslips for 48 hrs and washed twice in PBS and fixed using 4% paraformaldehyde in phosphate-buffered saline (PBS) at RT for 15 minutes. The fixed cells were then washed 3 X in 1 ml PBS. Fixed cells were permeabilized using 1% horse serum/0.5% Triton X-100 in PBS at RT for 10-15 minutes and then washed twice in PBS containing 1% horse serum. Cells were incubated with primary antibodies diluted (1:500-1:1000) in PBS/1% horse serum for 1 hr, prior to 3 x 1 ml washes in PBS/1% horse serum. Secondary antibodies AlexaFlour 488 (Thermo Scientific) was used at a dilution of 1:1000 in PBS/1% horse serum and incubated 1 hr at RT prior to 3 X washes in PBS/1% horse serum. Cells were then counterstained with 1 μg/mL DAPI in PBS/1% horse serum for 15-30 seconds before washing with PBS. Coverslips were rinsed in ddH2O and mounted onto a microscope slide using ProLong Gold Antifade Mountant. Antibodies: HIPK3 (Rabbit anti-HIPK3 from Abcam, ab72538), TP53 (Rabbit anti-p53 from Cell Signaling Technologies (CST), 2527S) and STAT3 (Mouse anti-STAT, Cell Signaling Technologies (CST), 124H6, 9139).

Antisense oligonucleotides (ASO) experiment

Antisense oligos were co-transfected into FL3 cells with pcDNA3-Twin-Strep-Tag-IGF2BP2 expression plasmid using Lipofectamine 2000 transfection reagent (Thermo) according to manufacturer's protocol for combined DNA/siRNA transfection. Briefly, 1.65×10^6 FL3 cells were seeded in a P10 dish and transfected 24 hrs later with 6 μg total DNA (2 μg pcDNA3-Twin-Strep-Tag-RBP expression vector and 6 μg pcDNA3 PL) and 20 nM final concentration of ASO. Cells were incubated for another 48 hrs prior to Twin-Strep-tag-RBP RIP as described above.

ASOs used in the study: 11-mer ASO, 5'- mC*mA*T*mC*G*C* T*G*mG*G*mG*G*mC*C*T* T*mC*T*mA*mG -3' Control ASO, 5'-mC*mG*T*A*mG*G*T*mG*G*C*A*T*mC*mG*C*mC*C*T*mC*mG -3'. m= 2'-O-Methyl RNA base *Phosphorothioate bond.

circHIPK3 knockdown and overexpression

For circHIPK3 KD (time-series excluded, see above), human cancer cell lines FL3, HCT-116, UMUC3 and J82 were grown in DMEM high glucose supplemented with 10% fetal bovine serum (FBS) and 1% penicillin/streptomycin (P/S). Cells (HCT-116, J82, UMUC3, FL3) were seeded in a 12-well plate (1 mL per well), 6-well plate (2 mL per well) or P6 dish (4 mL per well) - approximately 20% confluent at the day of transfection. The volumes and amounts mentioned below cover transfections made in 12-well plates. Double volumes were used for transfection in 6-well plates, 4x volumes in P6 dishes and 10x volumes in

P10 dishes. Cells were transfected with siRNA 24 hrs later using SilentFect (Bio-Rad): 37.5 µL OptiMEM (Gibco) + 1.25 µL SilentFect (Bio-Rad) was incubated at RT for 5 min before mixing with 1.1 µL siRNA (20 µM stock) pre-diluted in 37.5 µL OptiMEM. After gentle mixing by pipetting, the solution was incubated for 20 mins at RT prior to dropwise addition to cells (final concentration of 20 nM). Cells were incubated for 48 hrs before harvest. For overexpression studies 2.5 µL lipofectamine 2000 (Thermo Fisher Scientific) and 0.5-1.0 µg total DNA (pcDNA3.1(+)-Laccase2-circHIPK3 or pcDNA3.1(+)-Laccase2-MCS was used.

For RNA purification, cells were harvested by addition of 1 ml Trizol (Thermo Fisher Scientific) followed by mixing. RNA was purified according to the manufacturer's protocol with an additional chloroform extraction to increase quality.

For cDNA synthesis and qPCR, ≥ 1 µg of total RNA was DNase I (EN0525, Thermo Fisher Scientific) treated according to the manufacturer's description. Between 500-800 ng total RNA was used for first strand cDNA synthesis following the description of the Maxima First Strand cDNA Synthesis Kit for RT-qPCR (K1671, Thermo Fisher Scientific). The cDNA concentration for each sample was diluted to 5.0-2.0 ng/µL. Per well of an Agilent skirted plate, 7.5 µL Platinum SYBR Green qPCR SuperMix-UDG (Thermo Fisher Scientific) was mixed with 1 µL cDNA (5.0-2.0 ng/µL), 6.25 µL nuclease-free H_2O and 0.125 µL of each primer (10 µM) to a total reaction volume of 15 µL. Technical triplicates were made for each target gene. Samples were run on an AriaMx Real-Time PCR System and data was analyzed according to the X(0) method described elsewhere(Thomsen et al., 2010).

Quantification of IGF2BP2 copy number

IGF2BP2 copy numbers were measured by comparing signal intensity in western blotting experiments using known amounts of purified IGF2BP2 protein along with cell lysates from a known number of FL3, UMUC3 and J82 cells. Twin-Streptagged IGF2BP2 was expressed in HEK293 Flp-In T-Rex cells (Invitrogen) by tetracycline induction (200 ng/ml) for 36 hrs prior to cell lysis using hypotonic lysis buffer: 50 mM TRIS-HCl pH 7.5, 10 mM NaCl, 5 mM $MgCl_2$, 0.5% Triton-X 100, 1% Hexane-1,6-diol, 1 pill Complete - protease inhibitor EDTA-free (Roche) pr. 10 mL. Cells were scraped off and transferred to an Eppendorf tube. The lysate was added NaCl (final concentration of 150 mM), mixed and spun (18,000 x g, 15 min at 4°C). The supernatant was incubated with pre-equilibrated MagStrep type 3 XT beads (IBA Life sciences) and rotated for ≥2 hrs at 4ºC. The magnetic beads were washed 8 x with 1.5 mL Wash buffer (10 mM TRIS-HCl pH 7.5, 150 mM NaCl, 5 mM $MgCl_2$, 0.1% Triton-X 100). Twin-Streptagged IGF2BP2 was eluted by boiling beads 90°C, 5 mins in 100 µl 2 x SDS load buffer. Purified IGF2BP2 concentration was quantified and diluted to 100 ng/µl sample based on comparison with Coomassie staining (GelCode blue (Thermo scientific)) of known amounts of Bovine Serum Albumin (BSA).

Quantification of circHIPK3 copy number

$0.8 \cdot 10^6$ FL3, J82, or UMUC3 cells were harvested in TRIzol™ Reagent (Invitrogen™, 15596018) and RNA was isolated according to manufacturer's protocol. DNase treatment and cDNA synthesis were carried out using the Maxima™ H Minus cDNA Synthesis Master Mix, with dsDNase (Thermo Fisher Scientific, M1682) according to manufacturer's protocol. qPCR was performed with cDNA corresponding to 0.033% of the total amount of isolated RNA, 2 pmol of each of the primers 5'-GGTCGGCCAGTCATGTATCA-3' and 5'-ACACAACTGCTTGGCTCTACT-3', and Platinum SYBR Green qPCR SuperMix-UDG (Thermo Fisher Scientific, 11733038) according to manufacturer's protocol. circHIPK3 copy number per cell was calculated based on a standard curve made from serial dilutions of a DNA fragment (Twist Bioscience) carrying the sequence amplified by the circHIPK3 primers used for qPCR and starting at $2.5 \cdot 10^7$ molecules per qPCR-reaction. circHIPK3 qPCR primers: C73: GGTCGGCCAGTCATGTATCA and C74: ACACAACTGCTTGGCTCTACT.

Bladder cancer patient cohorts

Previously analyzed bladder cancer cohorts were mapped as described in (Hedegaard et al., 2016). For new cohorts (Strandgaard et al., 2022; Taber et al., 2020), RNA was paired-end sequenced (2 x 151 bp) using an Illumina NextSeq 500 instrument. Reads were demultiplexed using bcl2fastq v2.20.0.422 trimmed for traces of adapters using Trim Galore v0.4.1, and mapped to the hg19 genome build using tophat v2.1.1. Gene expression was estimated using cufflinks v2.1.1 and HTseq v0.6.1. CircRNA expression was quantified using the CIRI2 pipeline as described above for HepG2 and K562.

Statistical analyses

All statistical tests were performed in R (R Core Team, 2019; RStudio Team, 2016). *DESeq2* was used for differential expression analyses of time-series circHIPK3 KD data (see above) and for RBP knockdown data. Fisher's exact test was used to evaluate the significance of overlap between two gene sets. The chi-square test was used to evaluate different enrichments in downregulated, upregulated, and non-differentially expressed genes. The non-parametric Wilcoxon Rank Sum Test was used to evaluate differences between groups in large bladder cancer patient cohorts. The T-test was used to evaluate differences between conditions in wet-lab experiments. For multiple testing correction, we used Benjamini-Hochberg correction and statistical differences were declared significant at FDR < 0.1. When multiple testing was not applied, statistical differences were declared significant at $P < 0.05$. Most plots were produced with the R package *ggplot2 (Wickham, 2016)*. The R packages *ggfortify (Tang et al., 2016)* and *ggalt* were used to produce PCA plots and *GGally* were used to correlate samples. The R package

VennDiagram was used to produce Venn diagrams. *Survival (Therneau, 2015; Therneau and Grambsch, 2000)* and *Survminer (Kassambara et al., 2019)* were used to produce Kaplan-Meier plots and curves were compared statistically by the log-rank test. Additionally, the packages *dplyr* and *reshape* were used to analyze data.

Ethics approval and consent to participate

Informed written consent to take part in future research projects was obtained from all patients, and the specific project was approved by the National Committee on Health Research Ethics (#1706291; #1708266).

Data availability

All generated sequencing data have been deposited in NCBI's Gene Expression Omnibus. The circHIPK3 time-series knockdown sequencing data in UMUC3 and J82 cells are accessible through accession number GSE148824, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE148824. Sensitive personal data (raw data from RNA-Seq and clinical information) cannot be shared publicly because of Danish legislation regarding sharing and processing of sensitive personal data. The publicly available datasets supporting the conclusions of this article, e.g., eCLIP data, IGF2BP2 knockdown data and total RNA-Seq data from HepG2 and K562 etc., are available in the ENCODE repository, https://www.encodeproject.org/, as described in the Methods sections.

Competing interests

G.W.Y. is a co-founder, member of the Board of Directors, on the SAB, equity holder, and paid consultant for Locanabio and Eclipse BioInnovations. G.W.Y. is a visiting professor at the National University of Singapore. G.W.Y.'s interests have been reviewed and approved by the University of California, San Diego in accordance with its conflict-of-interest policies. The authors declare that they have no other competing interests.

Funding

This work was supported by generous grants from several funds. TLHO, MMN, and JSP were supported by the Lundbeck Foundation (R191-2015-1515), the Danish Cancer Society (R124-A7869), the Danish Council for Independent Research | Medical Sciences (FSS; DFF - 7016-00379), the Harboe Foundation (19110), and Aage and Johanne Louis-Hansen Foundation (19-2B-503). TLHO was also supported by a Carlsberg Foundation Internationalisation Fellowship (CF19-0493). ABK and CKD were supported by the Danish Cancer Society (R167-A11105) and the Lundbeck foundation (R370-2021-858). CKD was also

supported by the Carlsberg Foundation (CF20-0236 and CF18-0212). AKH was supported by Dagmar Marshalls Fond. GWY is supported by NIH R01 HG004659, U24 HG009889, and an Allen Distinguished Investigator Award, a Paul G. Allen Frontiers Group advised grant of the Paul G. Allen Foundation. We thank the NEYE Foundation for supporting the experimental follow-up studies.

Author's contributions

TLHO, JSP and CKD conceived and developed the project. TLHO and CKD performed the data analyses and generated all figures. CKD and ABK conducted perturbation experiments of circHIPK3 and RBPs in UMUC3, J82, FL3, and HT-116 cells, ASO and 11-mer mutation experiments and proliferation assays including data analyses. AKH performed IGF2BP2 and circHIPK3 copy-number determinations, including data analyses. MLG and CKD performed Western blotting and immunofluorescence experiments including data analyses. MMN performed Regmex analyses. SSP and GWY performed RNA immunoprecipitation in HepG2 and K562 cells, while CKD and AK performed RNA immunoprecipitation in FL3 cells. SV profiled circRNA and gene expression in RNA-Seq data sets. LSK and JK profiled miRNA expression. LD contributed clinical cohorts. JSP and CKD supervised the project. TLHO, CKD, and JSP wrote the manuscript with input from the other authors.

Acknowledgements

We thank Jeetayu Biswas and Robert Singer for the list of genes with IGF2BP2 KH34 RE (Biswas et al., 2019). Flow cytometry was performed at the FACS Core Facility, Aarhus University, Denmark. We thank Micki Grünzig and Karina Hjorth for technical assistance.

References

Abdelmohsen K, Panda AC, Munk R, Grammatikakis I, Dudekula DB, De S, Kim J, Noh JH, Kim KM, Martindale JL, Gorospe M. 2017. Identification of HuR target circular RNAs uncovers suppression of PABPN1 translation by CircPABPN1. *RNA Biol* **14**:361–369.

Anders S, Pyl PT, Huber W. 2015. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**:166–169.

Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S. 2014. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* **56**:55–66.

Bahn JH, Zhang Q, Li F, Chan T-M, Lin X, Kim Y, Wong DTW, Xiao X. 2015. The landscape of microRNA, Piwi-interacting RNA, and circular RNA in human saliva. *Clin Chem* **61**:221–230.

Barghash A, Golob-Schwarzl N, Helms V, Haybaeck J, Kessler SM. 2016. Elevated expression of the IGF2 mRNA binding protein 2 (IGF2BP2/IMP2) is linked to short survival and metastasis in esophageal adenocarcinoma. *Oncotarget* **7**:49743–49750.

Bartel DP. 2009. MicroRNA Target Recognition and Regulatory Functions. *Cell* **136**:215.

Begley DA, Berkenpas MB, Sampson KE, Abraham I. 1997. Identification and sequence of human PKY, a putative kinase with increased expression in multidrug-resistant cells, with homology to yeast protein kinase Yak1. *Gene* **200**:35–43.

Bekker-Jensen DB, Kelstrup CD, Batth TS, Larsen SC, Haldrup C, Bramsen JB, Sørensen KD, Høyer S, Ørntoft TF, Andersen CL, Nielsen ML, Olsen JV. 2017. An Optimized Shotgun Strategy for the Rapid Generation of Comprehensive Human Proteomes. *Cell Syst* **4**:587–599.e4.

Bell JL, Wächter K, Mühleck B, Pazaitis N, Köhn M, Lederer M, Hüttelmaier S. 2013. Insulin-like growth factor 2 mRNA-binding proteins (IGF2BPs): post-transcriptional drivers of cancer progression? *Cell Mol Life Sci* **70**:2657–2675.

Biswas J, Patel VL, Bhaskar V, Chao JA, Singer RH, Eliscovich C. 2019. The structural basis for RNA selectivity by the IMP family of RNA-binding proteins. *Nat Commun* **10**:4440.

Cai C, Zhi Y, Wang K, Zhang P, Ji Z, Xie C, Sun F. 2019. CircHIPK3 overexpression accelerates the proliferation and invasion of prostate cancer cells through regulating miRNA-338-3p. *Onco Targets Ther* **12**:3363–3372.

Cao J, Mu Q, Huang H. 2018. The Roles of Insulin-Like Growth Factor 2 mRNA-Binding Protein 2 in Cancer and Cancer Stem Cells. *Stem Cells Int* **2018**:4217259.

Chao JA, Patskovsky Y, Patel V, Levy M, Almo SC, Singer RH. 2010. ZBP1 recognition of -actin zipcode induces RNA looping. *Genes & Development*. doi:10.1101/gad.1862910

Chen C-K, Cheng R, Demeter J, Chen J, Weingarten-Gabbay S, Jiang L, Snyder MP, Weissman JS, Segal E, Jackson PK, Chang HY. 2021. Structured elements drive extensive circular RNA translation. *Mol Cell* **81**:4300–4318.e13.

Chen D, Lu X, Yang F, Xing N. 2019. Circular RNA circHIPK3 promotes cell proliferation and invasion of prostate cancer by sponging miR-193a-3p and regulating MCL1 expression. *Cancer Manag Res* **11**:1415–1423.

Cheng J, Zhuo H, Xu M, Wang L, Xu H, Peng J, Hou J, Lin L, Cai J. 2018. Regulatory network of circRNA–miRNA–mRNA contributes to the histological classification and disease progression in gastric cancer. *Journal of Translational Medicine*. doi:10.1186/s12967-018-1582-8

Chen G, Shi Y, Liu M, Sun J. 2018. circHIPK3 regulates cell proliferation and migration by sponging miR-124 and regulating AQP3 expression in hepatocellular carcinoma. *Cell Death Dis* **9**:175.

Chen S, Huang V, Xu X, Livingstone J, Soares F, Jeon J, Zeng Y, Hua JT, Petricca J, Guo H, Wang M, Yousif F, Zhang Y, Donmez N, Ahmed M, Volik S, Lapuk A, Chua MLK, Heisler LE, Foucal A, Fox NS, Fraser M, Bhandari V, Shiah Y-J, Guan J, Li J, Orain M, Picard V, Hovington H, Bergeron A, Lacombe L, Fradet Y, Têtu B, Liu S, Feng F, Wu X, Shao YW, Komor MA, Sahinalp C, Collins C, Hoogstrate Y, de Jong M, Fijneman RJA, Fei T, Jenster G, van der Kwast T, Bristow RG, Boutros PC, He HH. 2019. Widespread and Functional RNA Circularization in Localized Prostate Cancer. *Cell* **176**:831–843.e22.

Chen X, Mao R, Su W, Yang X, Geng Q, Guo C, Wang Z, Wang J, Kresty LA, Beer DG, Chang AC, Chen G. 2019. Circular RNA circHIPK3 modulates autophagy via MIR124-3p-STAT3-PRKAA/AMPKα signaling in STK11 mutant lung cancer. *Autophagy*. doi:10.1080/15548627.2019.1634945

Cocquerelle C, Mascrez B, Hétuin D, Bailleul B. 1993. Mis-splicing yields circular RNA molecules. *FASEB J* **7**:155–160.

Curtin JF, Cotter TG. 2004. JNK regulates HIPK3 expression and promotes resistance to Fas-mediated apoptosis in DU 145 prostate carcinoma cells. *J Biol Chem* **279**:17090–17100.

Dominguez D, Freese P, Alexis MS, Su A, Hochman M, Palden T, Bazile C, Lambert NJ, Van Nostrand EL, Pratt GA, Yeo GW, Graveley BR, Burge CB. 2018. Sequence, Structure, and Context Preferences of Human RNA Binding Proteins. *Mol Cell* **70**:854–867.e9.

Doyle M, Kiebler MA. 2012. A zipcode unzipped. *Genes Dev*.

Du WW, Fang L, Yang W, Wu N, Awan FM, Yang Z, Yang BB. 2017a. Induction of tumor apoptosis through a circular RNA enhancing Foxo3 activity. *Cell Death Differ* **24**:357–370.

Du WW, Yang W, Chen Y, Wu Z-K, Foster FS, Yang Z, Li X, Yang BB. 2017b. Foxo3 circular RNA promotes cardiac senescence by modulating multiple factors associated with stress and senescence responses. *Eur Heart J* **38**:1402–1412.

Du WW, Yang W, Liu E, Yang Z, Dhaliwal P, Yang BB. 2016. Foxo3 circular RNA retards cell cycle progression via forming ternary complexes with p21 and CDK2. *Nucleic Acids Res* **44**:2846–2858.

ENCODE Project Consortium. 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**:57–74.

Gao Y, Zhang J, Zhao F. 2018. Circular RNA identification based on multiple seed matching. *Brief Bioinform* **19**:803–810.

Ghasemi S, Emadi-Baygi M, Nikpour P. 2019. Down-regulation of circular RNA ITCH and circHIPK3 in gastric cancer tissues. *Turk J Med Sci* **49**:687–695.

Guo JU, Agarwal V, Guo H, Bartel DP. 2014. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol* **15**:409.

Hamilton MP, Rajapakshe K, Hartig SM, Reva B, McLellan MD, Kandoth C, Ding L, Zack TI, Gunaratne PH, Wheeler DA, Coarfa C, McGuire SE. 2013. Identification of a pan-cancer oncogenic microRNA superfamily anchored by a central core seed motif. *Nat Commun* **4**:2730.

Hamilton MP, Rajapakshe KI, Bader DA, Cerne JZ, Smith EA, Coarfa C, Hartig SM, McGuire SE. 2016. The Landscape of microRNA Targeting in Prostate Cancer Defined by AGO-PAR-CLIP. *Neoplasia* **18**:356–370.

Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J. 2013. Natural RNA circles function as efficient microRNA sponges. *Nature* **495**:384–388.

Hedegaard J, Lamy P, Nordentoft I, Algaba F, Høyer S, Ulhøi BP, Vang S, Reinert T, Hermann GG, Mogensen K, Thomsen MBH, Nielsen MM, Marquez M, Segersten U, Aine M, Höglund M, Birkenkamp-Demtröder K, Fristrup N, Borre M, Hartmann A, Stöhr R, Wach S, Keck B, Seitz AK, Nawroth R, Maurer T, Tulic C, Simic T, Junker K, Horstmann M, Harving N, Petersen AC, Calle ML, Steyerberg EW, Beukers W, van Kessel KEM, Jensen JB, Pedersen JS, Malmström P-U, Malats N, Real FX, Zwarthoff EC, Ørntoft TF, Dyrskjøt L. 2016. Comprehensive Transcriptional Analysis of Early-Stage Urothelial Carcinoma. *Cancer Cell* **30**:27–42.

Hollensen AK, Thomsen HS, Lloret-Llinares M, Kamstrup AB, Jensen JM, Luckmann M, Birkmose N, Palmfeldt J, Jensen TH, Hansen TB, Damgaard CK. 2020. nucleates a transcription inhibitory complex to balance neuronal differentiation. *Elife* **9**. doi:10.7554/eLife.58478

Jarlstad Olesen MT, S Kristensen L. 2021. Circular RNAs as microRNA sponges: evidence and controversies. *Essays*

Biochem **65**:685–696.

Jeck WR, Sorrentino JA, Wang K, Slevin MK, Burd CE, Liu J, Marzluff WF, Sharpless NE. 2013. Circular RNAs are abundant, conserved, and associated with ALU repeats. *RNA* **19**:141–157.

Jin P, Huang Y, Zhu P, Zou Y, Shao T, Wang O. 2018. CircRNA circHIPK3 serves as a prognostic marker to promote glioma progression by regulating miR-654/IGF2BP3 signaling. *Biochem Biophys Res Commun* **503**:1570–1574.

Kamran MZ, Patil P, Gude RP. 2013. Role of STAT3 in cancer metastasis and translational advances. *Biomed Res Int* **2013**:421821.

Kassambara A, Kosinski M, Biecek P. 2019. survminer: Drawing Survival Curves using "ggplot2". R package version 0.4.6.

Kelly S, Greenman C, Cook PR, Papantonis A. 2015. Exon Skipping Is Correlated with Exon Circularization. *J Mol Biol* **427**:2414–2417.

Kessler SM, Laggai S, Barghash A, Schultheiss CS, Lederer E, Artl M, Helms V, Haybaeck J, Kiemer AK. 2015. IMP2/p62 induces genomic instability and an aggressive hepatocellular carcinoma phenotype. *Cell Death Dis* **6**:e1894.

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol* **14**:R36.

Kofoed RH, Betzer C, Lykke-Andersen S, Molska E, Jensen PH. 2018. Investigation of RNA Synthesis Using 5-Bromouridine Labelling and Immunoprecipitation. *J Vis Exp*. doi:10.3791/57056

Kramer MC, Liang D, Tatomer DC, Gold B, March ZM, Cherry S, Wilusz JE. 2015. Combinatorial control of Drosophila circular RNA expression by intronic repeats, hnRNPs, and SR proteins. *Genes Dev* **29**:2168–2182.

Kristensen LS, Jakobsen T, Hager H, Kjems J. 2022. The emerging roles of circRNAs in cancer and oncology. *Nat Rev Clin Oncol* **19**:188–206.

Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, Lin C, Socci ND, Hermida L, Fulci V, Chiaretti S, Foà R, Schliwka J, Fuchs U, Novosel A, Müller R-U, Schermer B, Bissels U, Inman J, Phan Q, Chien M, Weir DB, Choksi R, De Vita G, Frezzetti D, Trompeter H-I, Hornung V, Teng G, Hartmann G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler CE, Nagle JW, Ju J, Papavasiliou FN, Benzing T, Lichter P, Tam W, Brownstein MJ, Bosio A, Borkhardt A, Russo JJ, Sander C, Zavolan M, Tuschl T. 2007. A mammalian microRNA expression atlas based on small RNA library sequencing. *Cell* **129**:1401–1414.

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* **9**:357–359.

Legnini I, Di Timoteo G, Rossi F, Morlando M, Briganti F, Sthandier O, Fatica A, Santini T, Andronache A, Wade M, Laneve P, Rajewsky N, Bozzoni I. 2017. Circ-ZNF609 Is a Circular RNA that Can Be Translated and Functions in Myogenesis. *Mol Cell* **66**:22–37.e9.

Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. 2015. The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Systems*. doi:10.1016/j.cels.2015.12.004

Li F, Zhang L, Li W, Deng J, Zheng J, An M, Lu J, Zhou Y. 2015. Circular RNA ITCH has inhibitory effect on ESCC by suppressing the Wnt/β-catenin pathway. *Oncotarget* **6**:6001–6013.

Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*

25:2078–2079.

Liu C-X, Chen L-L. 2022. Circular RNAs: Characterization, cellular roles, and applications. *Cell* **185**:2016–2034.

Liu N, Zhang J, Zhang L-Y, Wang L. 2018. CircHIPK3 is upregulated and predicts a poor prognosis in epithelial ovarian cancer. *Eur Rev Med Pharmacol Sci* **22**:3713–3718.

Li Y, Zheng F, Xiao X, Xie F, Tao D, Huang C, Liu D, Wang M, Wang L, Zeng F, Jiang G. 2017a. CircHIPK3 sponges miR-558 to suppress heparanase expression in bladder cancer cells. *EMBO reports*. doi:10.15252/embr.201643581

Li Y, Zheng F, Xiao X, Xie F, Tao D, Huang C, Liu D, Wang M, Wang L, Zeng F, Jiang G. 2017b. CircHIPK3 sponges miR-558 to suppress heparanase expression in bladder cancer cells. *EMBO Rep*. doi:10.15252/embr.201643581

Li Y, Zheng Q, Bao C, Li S, Guo W, Zhao J, Chen D, Gu J, He X, Huang S. 2015. Circular RNA is enriched and stable in exosomes: a promising biomarker for cancer diagnosis. *Cell Res* **25**:981–984.

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**:550.

Ludwig N, Leidinger P, Becker K, Backes C, Fehlmann T, Pallasch C, Rheinheimer S, Meder B, Stähler C, Meese E, Keller A. 2016. Distribution of miRNA expression across human tissues. *Nucleic Acids Res* **44**:3865–3877.

Luo E-C, Nathanson JL, Tan FE, Schwartz JL, Schmok JC, Shankar A, Markmiller S, Yee BA, Sathe S, Pratt GA, Scaletta DB, Ha Y, Hill DE, Aigner S, Yeo GW. 2020. Large-scale tethered function assays identify factors that regulate mRNA stability and translation. *Nat Struct Mol Biol* **27**:989–1000.

Luo W, Friedman MS, Shedden K, Hankenson KD, Woolf PJ. 2009. GAGE: generally applicable gene set enrichment for pathway analysis. *BMC Bioinformatics* **10**:161.

Memczak S, Jens M, Elefsinioti A, Torti F, Krueger J, Rybak A, Maier L, Mackowiak SD, Gregersen LH, Munschauer M, Loewer A, Ziebold U, Landthaler M, Kocks C, le Noble F, Rajewsky N. 2013. Circular RNAs are a large class of animal RNAs with regulatory potency. *Nature* **495**:333–338.

Memczak S, Papavasileiou P, Peters O, Rajewsky N. 2015. Identification and Characterization of Circular RNAs As a New Class of Putative Biomarkers in Human Blood. *PLoS One* **10**:e0141214.

Militello G, Weirick T, John D, Döring C, Dimmeler S, Uchida S. 2017. Screening and validation of lncRNAs and circRNAs as miRNA sponges. *Brief Bioinform* **18**:780–788.

Moll P, Ante M, Seitz A, Reda T. 2014. QuantSeq 3′ mRNA sequencing for RNA quantification. *Nature Methods*. doi:10.1038/nmeth.f.376

Nielsen MM, Tataru P, Madsen T, Hobolth A, Pedersen JS. 2018. Regmex: a statistical tool for exploring motifs in ranked sequence lists from genomics experiments. *Algorithms Mol Biol* **13**:17.

Niu G, Wright KL, Ma Y, Wright GM, Huang M, Irby R, Briggs J, Karras J, Douglas Cress W, Pardoll D, Jove R, Chen J, Yu H. 2005. Role of Stat3 in Regulating p53 Expression and Function. *Molecular and Cellular Biology*. doi:10.1128/mcb.25.17.7432-7440.2005

Okholm TLH, Nielsen MM, Hamilton MP, Christensen L-L, Vang S, Hedegaard J, Hansen TB, Kjems J, Dyrskjøt L, Pedersen JS. 2017. Circular RNA expression is abundant and correlated to aggressiveness in early-stage bladder cancer. *NPJ Genom Med* **2**:36.

Okholm TLH, Sathe S, Park SS, Kamstrup AB, Rasmussen AM, Shankar A, Chua ZM, Fristrup N, Nielsen MM, Vang S,

Dyrskjøt L, Aigner S, Damgaard CK, Yeo GW, Pedersen JS. 2020. Transcriptome-wide profiles of circular RNA and RNA-binding protein interactions reveal effects on circular RNA biogenesis and cancer pathway expression. *Genome Med* **12**:112.

Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, Shenzis S, Samson M, Dittmar G, Landthaler M, Chekulaeva M, Rajewsky N, Kadener S. 2017. Translation of CircRNAs. *Mol Cell* **66**:9–21.e7.

Patel VL, Mitra S, Harris R, Buxbaum AR, Lionnet T, Brenowitz M, Girvin M, Levy M, Almo SC, Singer RH, Chao JA. 2012. Spatial arrangement of an RNA zipcode identifies mRNAs under post-transcriptional control. *Genes Dev* **26**:43–53.

Piwecka M, Glažar P, Hernandez-Miranda LR, Memczak S, Wolf SA, Rybak-Wolf A, Filipchyk A, Klironomos F, Cerda Jara CA, Fenske P, Trimbuch T, Zywitza V, Plass M, Schreyer L, Ayoub S, Kocks C, Kühn R, Rosenmund C, Birchmeier C, Rajewsky N. 2017. Loss of a mammalian circular RNA locus causes miRNA deregulation and affects brain function. *Science* **357**. doi:10.1126/science.aam8526

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rochat-Steiner V, Becker K, Micheau O, Schneider P, Burns K, Tschopp J. 2000. FIST/HIPK3: a Fas/FADD-interacting serine/threonine kinase that induces FADD phosphorylation and inhibits fas-mediated Jun NH(2)-terminal kinase activation. *J Exp Med* **192**:1165–1174.

RStudio Team. 2016. RStudio: Integrated Development for R. RStudio, Inc., Boston, MA.

Sampson KE, McCroskey MC, Abraham I. 1993. Identification of a 170 kDa membrane kinase with increased activity in KB-V1 multidrug resistant cells. *J Cell Biochem* **52**:384–395.

Schneider T, Hung L-H, Aziz M, Wilmen A, Thaum S, Wagner J, Janowski R, Müller S, Schreiner S, Friedhoff P, Hüttelmaier S, Niessing D, Sattler M, Schlundt A, Bindereif A. 2019. Combinatorial recognition of clustered RNA elements by the multidomain RNA-binding protein IMP3. *Nat Commun* **10**:2266.

Shan K, Liu C, Liu B-H, Chen X, Dong R, Liu X, Zhang Y-Y, Liu B, Zhang S-J, Wang J-J, Zhang S-H, Wu J-H, Zhao C, Yan B. 2017. Circular Noncoding RNA HIPK3 Mediates Retinal Vascular Dysfunction in Diabetes Mellitus. *Circulation* **136**:1629–1642.

Stagsted LV, Nielsen KM, Daugaard I, Hansen TB. 2019. Noncoding AUG circRNAs constitute an abundant and conserved subclass of circles. *Life Sci Alliance* **2**. doi:10.26508/lsa.201900398

Starke S, Jost I, Rossbach O, Schneider T, Schreiner S, Hung L-H, Bindereif A. 2015. Exon circularization requires canonical splice signals. *Cell Rep* **10**:103–111.

Stoll L, Sobel J, Rodriguez-Trejo A, Guay C, Lee K, Venø MT, Kjems J, Laybutt DR, Regazzi R. 2018. Circular RNAs as novel regulators of β-cell functions in normal and disease conditions. *Mol Metab* **9**:69–83.

Strandgaard T, Lindskrog SV, Nordentoft I, Christensen E, Birkenkamp-Demtröder K, Andreasen TG, Lamy P, Kjær A, Ranti D, Wang YA, Bieber C, Prip F, Rasmussen J, Steiniche T, Birkbak N, Sfakianos J, Horowitz A, Jensen JB, Dyrskjøt L. 2022. Elevated T-cell Exhaustion and Urinary Tumor DNA Levels Are Associated with Bacillus Calmette-Guérin Failure in Patients with Non-muscle-invasive Bladder Cancer. *Eur Urol* **82**:646–656.

Taber A, Christensen E, Lamy P, Nordentoft I, Prip F, Lindskrog SV, Birkenkamp-Demtröder K, Okholm TLH,

Knudsen M, Pedersen JS, Steiniche T, Agerbæk M, Jensen JB, Dyrskjøt L. 2020. Molecular correlates of cisplatin-based chemotherapy response in muscle invasive bladder cancer by integrated multi-omics analysis. *Nat Commun* **11**:1–15.

Tang Y, Horikoshi M, Li W. 2016. ggfortify: Unified Interface to Visualize Statistical Results of Popular R Packages. *The R Journal*. doi:10.32614/rj-2016-060

Teng F, Xu J, Zhang M, Liu S, Gu Y, Zhang M, Wang X, Ni J, Qian B, Shen R, Jia X. 2019. Comprehensive circular RNA expression profiles and the tumor-suppressive function of circHIPK3 in ovarian cancer. *Int J Biochem Cell Biol* **112**:8–17.

Therneau T. 2015. A Package for Survival Analysis in S. Version 2.38.

Therneau TM, Grambsch PM. 2000. Modeling Survival Data: Extending the Cox Model. *Springer, New York*.

Thomsen R, Sølvsten CAE, Linnet TE, Blechingberg J, Nielsen AL. 2010. Analysis of qPCR data by converting exponentially related Ct values into linearly related X0 values. *J Bioinform Comput Biol* **8**:885–900.

Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* **28**:511–515.

Van Nostrand EL, Freese P, Pratt GA, Wang X, Wei X, Xiao R, Blue SM, Chen J-Y, Cody NAL, Dominguez D, Olson S, Sundararaman B, Zhan L, Bazile C, Bouvrette LPB, Bergalet J, Duff MO, Garcia KE, Gelboin-Burkhart C, Hochman M, Lambert NJ, Li H, Nguyen TB, Palden T, Rabano I, Sathe S, Stanton R, Su A, Wang R, Yee BA, Zhou B, Louie AL, Aigner S, Fu X-D, Lécuyer E, Burge CB, Graveley BR, Yeo GW. 2018. A Large-Scale Binding and Functional Map of Human RNA Binding Proteins. *bioRxiv*. doi:10.1101/179648

Van Nostrand EL, Pratt GA, Shishkin AA, Gelboin-Burkhart C, Fang MY, Sundararaman B, Blue SM, Nguyen TB, Surka C, Elkins K, Stanton R, Rigo F, Guttman M, Yeo GW. 2016. Robust transcriptome-wide discovery of RNA-binding protein binding sites with enhanced CLIP (eCLIP). *Nat Methods* **13**:508–514.

Vo JN, Cieslik M, Zhang Y, Shukla S, Xiao L, Zhang Y, Wu Y-M, Dhanasekaran SM, Engelke CG, Cao X, Robinson DR, Nesvizhskii AI, Chinnaiyan AM. 2019. The Landscape of Circular RNA in Cancer. *Cell* **176**:869–881.e13.

Wächter K, Köhn M, Stöhr N, Hüttelmaier S. 2013. Subcellular localization and RNP formation of IGF2BPs (IGF2 mRNA-binding proteins) is modulated by distinct RNA-binding domains. *Biol Chem* **394**:1077–1090.

Wen J, Liao J, Liang J, Chen X-P, Zhang B, Chu L. 2020. Circular RNA HIPK3: A Key Circular RNA in a Variety of Human Cancers. *Frontiers in Oncology*. doi:10.3389/fonc.2020.00773

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis. Springer.

Xie Y, Yuan X, Zhou W, Kosiba AA, Shi H, Gu J, Qin Z. 2020. The circular RNA HIPK3 (circHIPK3) and its regulation in cancer progression: Review. *Life Sci* **254**:117252.

Ye S, Song W, Xu X, Zhao X, Yang L. 2016. IGF2BP2 promotes colorectal cancer cell proliferation and survival through interfering with RAF-1 degradation by miR-195. *FEBS Lett* **590**:1641–1650.

Zeng K, Chen X, Xu M, Liu X, Hu X, Xu T, Sun H, Pan Y, He B, Wang S. 2018. CircHIPK3 promotes colorectal cancer growth and metastasis by sponging miR-7. *Cell Death Dis* **9**:417.

Zhang X-O, Wang H-B, Zhang Y, Lu X, Chen L-L, Yang L. 2014. Complementary sequence-mediated exon circularization. *Cell* **159**:134–147.

Zhang Z, Yang T, Xiao J. 2018. Circular RNAs: Promising Biomarkers for Human Diseases. *EBioMedicine* **34**:267–274.

Zheng Q, Bao C, Guo W, Li S, Chen J, Chen B, Luo Y, Lyu D, Li Y, Shi G, Liang L, Gu J, He X, Huang S. 2016. Circular RNA profiling reveals an abundant circHIPK3 that regulates cell growth by sponging multiple miRNAs. *Nat Commun* **7**:11215.

Zhou J, Wang B, Bin X, Xie C, Li B, Liu O, Tang Z. 2021. CircHIPK3: Key Player in Pathophysiology and Potential Diagnostic and Therapeutic Tool. *Front Med* **8**:615417.

Figure legends

Figure 1: Thousands of genes are deregulated upon circHIPK3 knockdown

(A) circHIPK3 is produced from the second exon of the *HIPK3* gene. We designed a circHIPK3-specific siRNA targeting the backsplice junction of circHIPK3 to specifically knockdown the expression of circHIPK3. **(B)** Expression of circHIPK3 and the corresponding linear transcript in eleven bladder cancer cell lines. UMUC3 and J82 were chosen for knockdown experiments based on circHIPK3 expression levels and cell line stability. CPM = counts per million. **(C)** Knockdown efficiency of circHIPK3 in UMUC3 and J82 at different time points post circHIPK3 (yellow) or scramble (grey) siRNA transfection. Expression is normalized to GAPDH levels. **(D)** PCA plot of gene expression in UMUC3 and J82. PCA plots are based on the genes with the 50% most variance across all cell line samples. Gene expression is log transformed (natural logarithm) and added a pseudocount of 1. **(E)** Overlap of differentially expressed genes across time and/or conditions between UMUC3 (n = 3,072) and J82 (n = 2,389). P-value obtained by Fisher's Exact Test. **(F**) Number of differentially expressed genes at each time point (Wald test, Benjamini-Hochberg correction with FDR < 0.1). Only genes with perturbed expression profiles across time and/or conditions in UMUC3 (n = 3,072) or J82 (n = 2,389) are considered. **(G)** Differential expression analysis between circHIPK3 KD and scr siRNA samples 24 hour post-transfection (Wald test). The log2 fold changes (circHIPK3 KD vs scr) are plotted against the negative log10(p-values). Colors indicate if genes are significantly down- (red) or upregulated (blue) or not differentially expressed (Not DE, green) after Benjamini-Hochberg correction, FDR < 0.1. Vertical lines indicate a log2FC > 1 or < -1.

Figure 2: A long motif in circHIPK3 is enriched in downregulated genes upon circHIPK3 knockdown

(A) Illustration of conserved miRNA binding sites in circHIPK3 based on AGO-CLIP data. **(B)** Percentage of genes in each group with miR-148ab-3p/152 target sites (CAGUGCA) in their 3'UTRs. Gene regulation is based on circHIPK3 KD vs scr siRNA transfection 24 hours (24h) post-transfection in J82. P-values obtained by Chi-square Test. **(C)** Procedure for motif enrichment analysis using Regmex. First, we extract all 7-mers in the sequence of circHIPK3. For each 7-mer, e.g., CCCTGCC, we identify their presence in the 3'UTR of genes. We order all genes according to a differential expression score calculated as the -log10(P-value) multiplied by the fold change direction, e.g., 1 for upregulated genes and -1 for downregulated genes. Then we calculate a motif score for each 7-mer based on their occurrences in either down- or upregulated genes. If a 7-mer has a positive motif score it means that it is enriched in the 3'UTR of genes that are upregulated upon circHIPK3 KD. Conversely, 7-mers with a negative motif score are primarily found in genes that are downregulated upon circHIPK3 KD. The 7-mer, CCCTGCC, is used for illustration purposes. **(D)** Regmex motif scores for circHIPK3 7-mers (UMUC3, 24h). Alignment of the ten 7-mers with the most negative motif scores are shown. 7-mers are ranked from most negative to most positive motif

scores. Numbers correspond to rank. **(E)** Illustration of circHIPK3 and position of the ten motifs with the lowest motif scores. The 7-mers with the most negative motif scores found in circHIPK3 comprise a larger 11-mer, AGGCCCCCAGC, present in the sequence of circHIPK3. **(F)** Percentage of genes in each group containing the 11-mer motif upon circHIPK3 KD in UMUC3 cells (24h). P-values obtained by Chi-square Test.

Figure 3: The 11-mer motif in circHIPK3 constitutes a binding site for IGF2BP2

(A) Illustration of RBP binding sites in circHIPK3 based on eCLIP data in the ENCODE cell lines HepG2 and K562. **(B)** RNA immunoprecipitation of circHIPK3-RBPs and others in HepG2 and K562 cells. Bands confirm that circHIPK3 interacts with IGF2BP1 and IGF2BP2. Marker indicates 50 bp. **(C)** Percentage of genes in each group containing IGF2BP2 binding sites (K562) in UMUC3 (24h). P-values obtained by Chi-square Test. **(D)** Overlap between genes that are affected by circHIPK3 KD in UMUC3 and IGF2BP2 KD in K562. P-value obtained by Fisher's Exact Test. **(E)** Percentage of genes in each group affected by IGF2BP2 KD in K562. P-values obtained by Chi-square Test. **(F)** Regulation of genes affected by both circHIPK3 KD (UMUC3, 24h) and IGF2BP2 KD (K562). The x-axis indicates gene regulation upon circHIPK3 KD in UMUC3. Percentage on y-axis and colors show how these genes are regulated upon IGF2BP2 KD. Downregulated genes upon circHIPK3 KD are mainly upregulated upon IGF2BP2 KD and vice versa. **(G)** Regmex motif enrichment analysis upon IGF2BP2 KD in K562. 7-mers are ranked from most positive to most negative motif scores. The ten 7-mers with the most negative motif scores in UMUC3 are shown. Numbers indicate motif rank. All possible 7-mers are evaluated (n = 16,384). **(H-I)** Percentage of genes in each group containing the 11-mer motif (H) and IGF2BP2 binding sites (I) upon IGF2BP2 KD in K562. P-values obtained by Chi-square Test.

Figure 4: circHIPK3 interacts with IGF2BP2 and affects genes controlling cell cycle progression

(A) Relative enrichment of circHIPK3 levels between IP and input for the three RBPs GRWD1, IGF2BP1, and IGF2BP2 in the bladder cancer cell line FL3. GFP was used as a negative control. **(B)** Relative enrichment of circHIPK3 levels between IP and input for IGF2BP2 upon circHIPK3 overexpression (OE). GFP was used as a negative control. Error bars reflect standard deviation of biological triplicates. **(C)** Percentage of genes in each group containing the 11-mer motif upon IGF2BP2 KD in UMUC3. P-values obtained by Chi-square Test. **(D)** Gene set enrichment analysis of 50 hallmarks of cancer upon circHIPK3 KD (24h) in UMUC3 cells (FDR < 0.1 for all shown pathways). **(E)** Downregulated hallmarks of cancer upon IGF2BP2 KD in UMUC3 (P-value < 0.05 for all shown pathways).

Figure 5: circHIPK3 functions as a competing endogenous RNA for IGF2BP2

(A) *STAT3* mRNA expression upon circHIPK3 KD in the time-course perturbation experiment in UMUC3 cells. Expression represents DESeq2 normalized counts. P-value (24h) obtained by Wald Test. **(B)** Expression of *STAT3* upon IGF2BP2 knockdown in UMUC3 cells. P-value obtained by Wald Test. **(C)** Relative enrichment of *STAT3* and β-actin levels between IGF2BP2 IP and input in FL3 cells. GFP was used as a negative control. P-value obtained by T-test. **(D)** Relative enrichment of *STAT3* levels between IGF2BP2 IP and input upon circHIPK3 KD in FL3 cells. P-value obtained by T-test. Control sample (scr) has been normalized to 1. **(E)** Relative enrichment of *STAT3* and *GAPDH* levels between IGF2BP2 IP and input after 11-mer ASO transfection in FL3 cells. P-values obtained by T-test. **(F)** Relative enrichment of circHIPK3, *STAT3*, and *GAPDH* levels between IGF2BP2 IP and input after transfection of wildtype and mutated circHIPK3 in HEK293 cells. P-values obtained by T-test. **(G)** *TP53* mRNA expression upon circHIPK3 KD in the time-course perturbation experiment in UMUC3 cells. Expression represents DESeq2 normalized counts. P-value (72h) obtained by Wald Test. **(H)** Western blot of STAT3 and p53 protein upon circHIPK3 KD in UMUC cells. Quantifications on the left are internally normalized to HuR and then scr is normalized to 1. P-values obtained by T-test. **(I-J)** Immunofluorescence staining for STAT3 (I) or p53 (J) in fixed UMUC3 cells (as indicated - green) subjected to control (scr) or circHIPK3 KD. Nuclei were counterstained using DAPI (as indicated - blue). Mean signal intensities are quantified within nuclei or entire cells to the right. P-values obtained by Wilcoxon Rank Sum Test. **(B-F,H)** Error bars reflect standard deviation of biological triplicates.

Figure 6: circHIPK3 levels are low in BC cells compared to IGF2BP2 and high circHIPK3 levels correlate with increased survival of bladder cancer patients

(A) Flow chart used to perform absolute quantification of circHIPK3 and IGF2BP2 protein levels in BC cells. **(B)** Estimated levels of circHIPK3 and IGF2BP2 (copies per cell). **(C)** Immunofluorescence using IGF2BP2 antibody (red channel) with cells transfected with circHIPK3 expression plasmid (upper panel) or control plasmid (lower panel). Both plasmids encode enhanced green fluorescent protein (GFP) from a separate expression cassette to identify transfected cells (green channel). White arrows indicate IGF2BP2 condensates **(D)** Quantification of transfected cells containing large IGF2BP2 condensates - % positive cells (28 and 34 cells were counted). Error bars reflect standard error of mean (SE). **(E)** Expression of circHIPK3 in samples from patients with non-muscle invasive bladder cancer (NMIBC) and muscle invasive bladder cancer (MIBC). CPM = counts per million. P-value obtained by Wilcoxon Rank Sum Test. **(F)** Kaplan–Meier overall survival plots of circHIPK3 (left) and *HIPK3* (right). Median expression of circHIPK3 (0.192 CPM) and *HIPK3* (7.96 FPKM) used as cutoff. P-values obtained by Log-Rank Test. **(G)** Working model that illustrates potential mechanism for circHIPK3. Our studies show that circHIPK3 contains an 11-mer motif (illustrated by red bases) that comprise a binding site for IGF2BP2 (in green). At normal levels,

circHIPK3 interacts strongly with IGF2BP2 (middle panel), potentially involving nucleation of many IGF2BP2 proteins on single circHIPK3 molecules - exacerbated at high circHIPK3 levels, which induce condensates (left panel). This allows for only a few accessible IGF2BP2 molecules to bind target mRNAs including *STAT3*. Upon circHIPK3 knockdown and low circHIPK3 levels (right panel), IGF2BP2 is more accessible to interact more strongly with IGF2BP2 target genes, which contain the 11-mer motif in their 3'UTRs, incl. *STAT3*, and which are subsequently downregulated. Conversely, upon IGF2BP2 knockdown, IGF2BP2 target genes containing 11-mer are upregulated (not illustrated).

Supplementary Tables

Supplementary Table S1. qPCR primers for detecting circHIPK3 (Table_S1_Primers_circHIPK3.xlsx)

Supplementary Table S2. Conserved miRNA binding sites in circHIPK3 based on AgoClip data (Table_S2_circHIPK3_miRNAs.xlsx)

Supplementary Table S3. miRNA expression in J82 (Table_S3_miRNA_profiling_J82.xlsx)

Supplementary Table S4. 7-mer motifs in circHIPK3 and motif scores from RegMex analysis based on gene regulations 24 hours post-transfection in UMUC3 and 48 hours post-transfection in J82 (Table_S4_circHIPK3_7mers.xlsx)

Supplementary Table S5. Subcellular localization of circHIPK3-RBPs in HepG2 (Table_S5_HepG2_RBP_subcellular_location.txt)

Figure 1

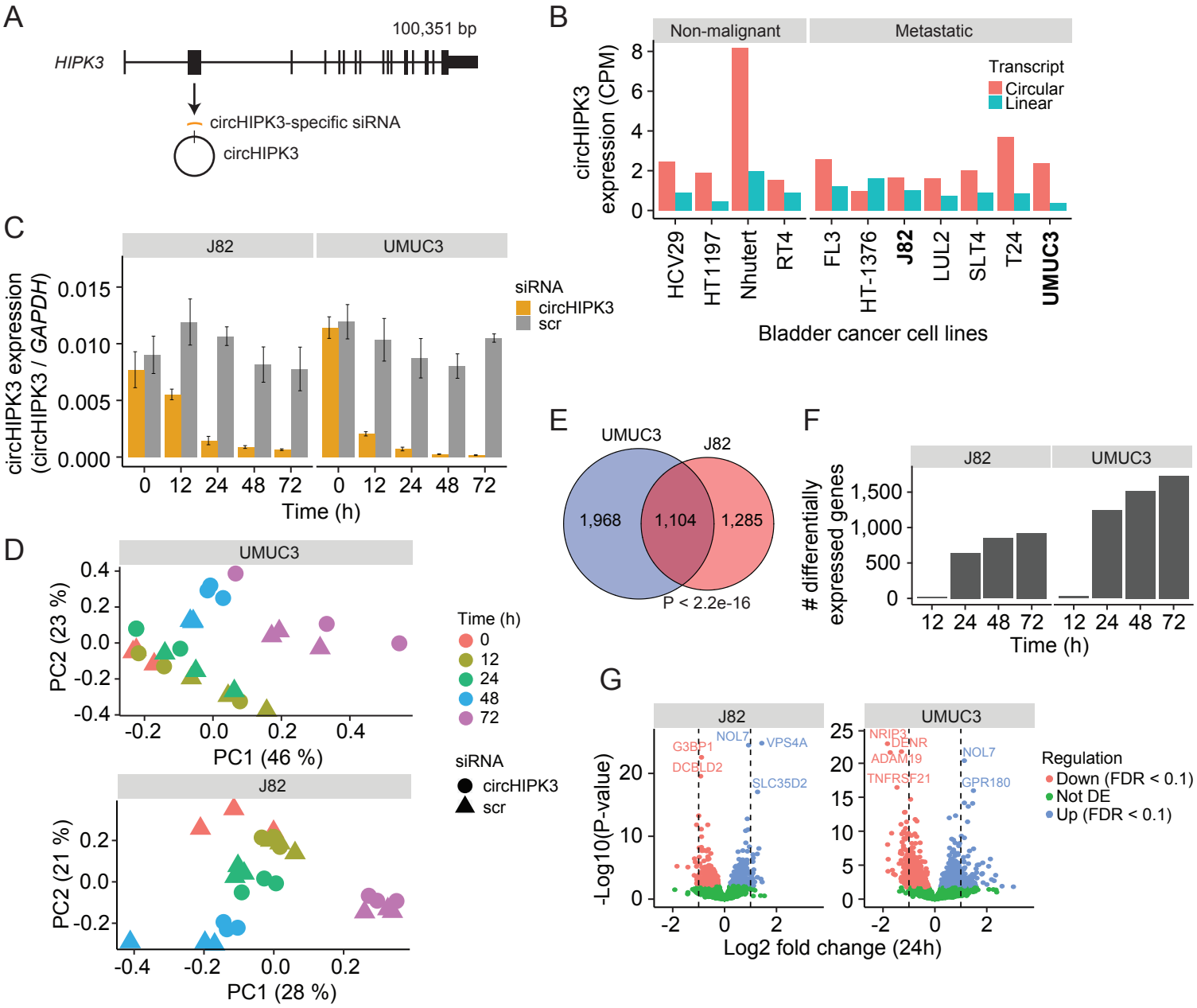

A
100,351 bp
HIPK3
circHIPK3-specific siRNA
circHIPK3
B
Non-malignant
Metastatic
circHIPK3 expression (CPM)
Transcript
Circular
Linear
HCV29
HT1197
Nhutert
RT4
FL3
HT-1376
J82
LUL2
SLT4
T24
UMUC3
Bladder cancer cell lines
C
J82
UMUC3
circHIPK3 expression (circHIPK3 / GAPDH)
siRNA
circHIPK3
scr
Time (h)
D
UMUC3
PC2 (23 %)
PC1 (46 %)
J82
PC2 (21 %)
PC1 (28 %)
Time (h)
siRNA
circHIPK3
scr
E
UMUC3
J82
1,968
1,104
1,285
P < 2.2e-16
F
J82
UMUC3
differentially expressed genes
Time (h)
G
J82
UMUC3
-Log10(P-value)
Log2 fold change (24h)
G3BP1
DCBLD2
NOL7
VPS4A
SLC35D2
NRIP3
DENR
ADAM19
TNFRSF21
NOL7
GPR180
Regulation
Down (FDR < 0.1)
Not DE
Up (FDR < 0.1)

Figure 2

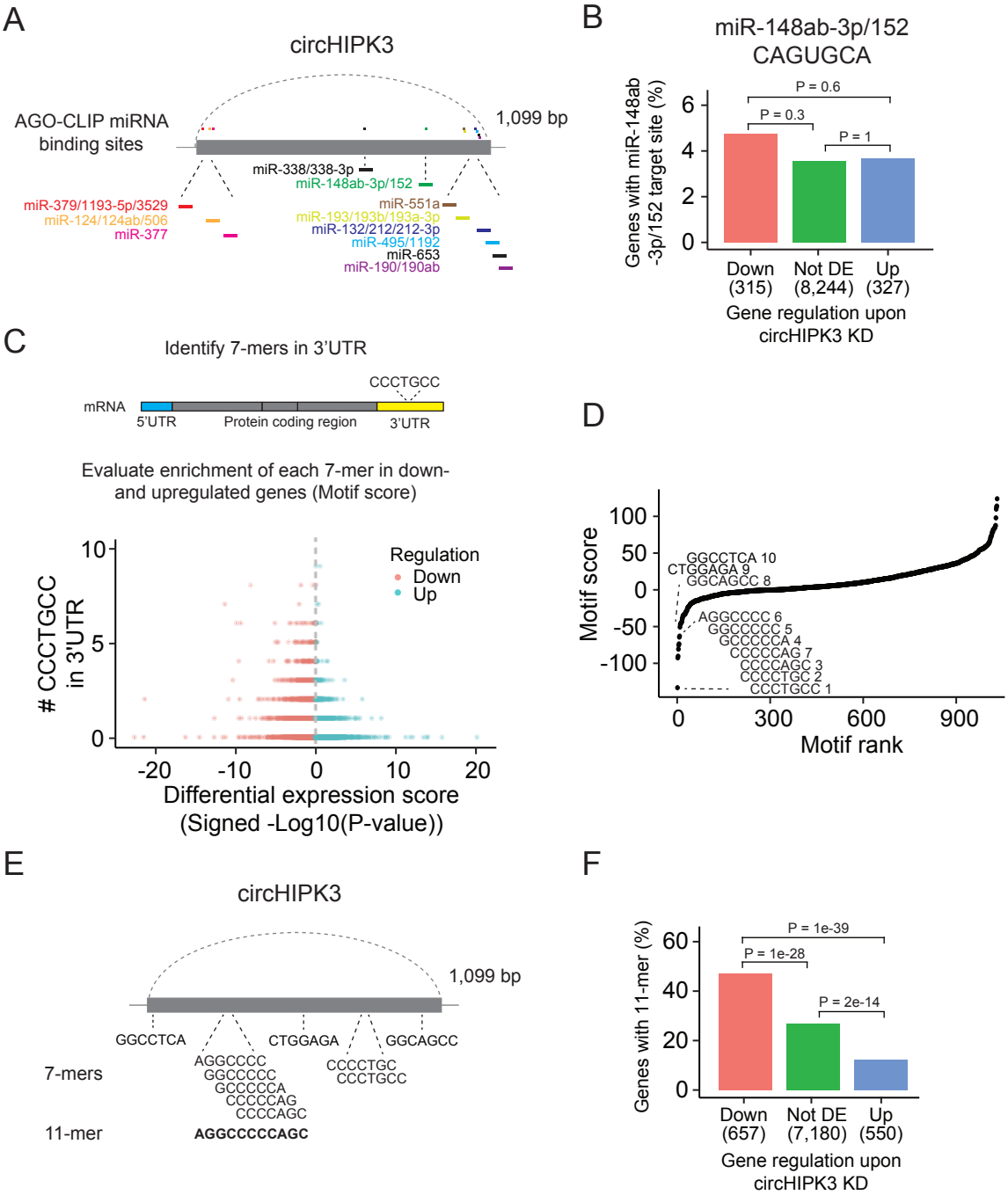

A
circHIPK3
AGO-CLIP miRNA binding sites
1,099 bp
miR-338/338-3p
miR-148ab-3p/152
miR-551a
miR-379/1193-5p/3529
miR-124/124ab/506
miR-377
miR-193/193b/193a-3p
miR-132/212/212-3p
miR-495/1192
miR-653
miR-190/190ab
B
miR-148ab-3p/152
CAGUGCA
Genes with miR-148ab -3p/152 target site (%)
P = 0.6
P = 0.3
P = 1
Down (315)
Not DE (8,244)
Up (327)
Gene regulation upon circHIPK3 KD
C
Identify 7-mers in 3'UTR
mRNA
CCCTGCC
5'UTR
Protein coding region
3'UTR
Evaluate enrichment of each 7-mer in down- and upregulated genes (Motif score)
CCCTGCC in 3'UTR
Regulation
Down
Up
Differential expression score (Signed -Log10(P-value))
D
Motif score
GGCCTCA 10
CTGGAGA 9
GGCAGCC 8
AGGCCCC 6
GGCCCCC 5
GCCCCCA 4
CCCCCAG 7
CCCCAGC 3
CCCCTGC 2
CCCTGCC 1
Motif rank
E
circHIPK3
1,099 bp
GGCCTCA
CTGGAGA
GGCAGCC
7-mers
AGGCCCC
GGCCCCC
GCCCCCA
CCCCCAG
CCCCAGC
CCCCTGC
CCCTGCC
11-mer
AGGCCCCCAGC
F
Genes with 11-mer (%)
P = 1e-39
P = 1e-28
P = 2e-14
Down (657)
Not DE (7,180)
Up (550)
Gene regulation upon circHIPK3 KD

Figure 3

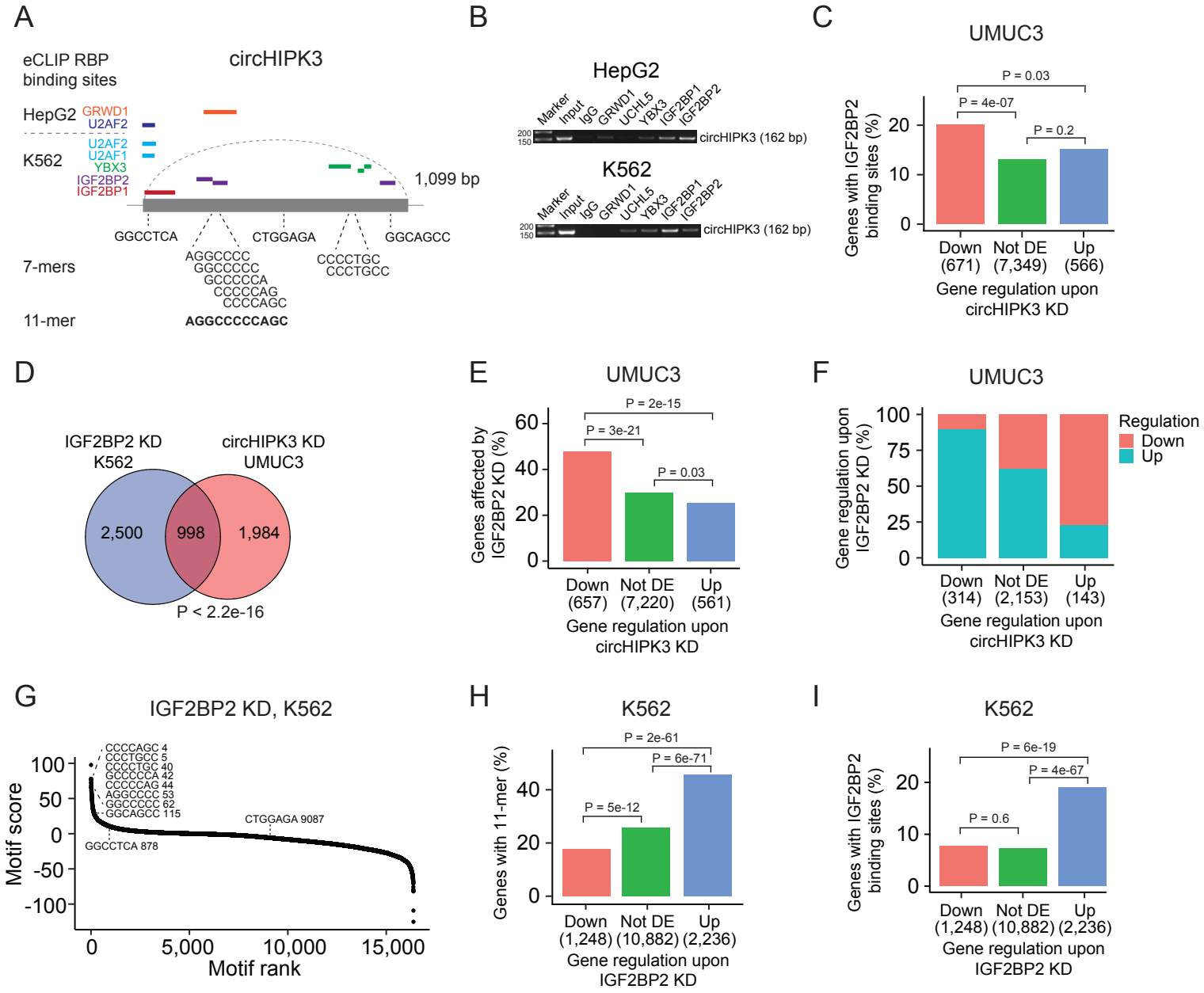

A
eCLIP RBP binding sites
circHIPK3
HepG2
GRWD1
U2AF2
K562
U2AF2
U2AF1
YBX3
IGF2BP2
IGF2BP1
1,099 bp
GGCCTCA
CTGGAGA
GGCAGCC
7-mers
AGGCCCC
GGCCCCC
GCCCCCA
CCCCCAG
CCCCAGC
CCCCTGC
CCCTGCC
11-mer
AGGCCCCCAGC
B
HepG2
Marker
Input
IgG
GRWD1
UCHL5
YBX3
IGF2BP1
IGF2BP2
200
150
circHIPK3 (162 bp)
K562
circHIPK3 (162 bp)
C
UMUC3
P = 0.03
P = 4e-07
P = 0.2
Genes with IGF2BP2 binding sites (%)
Down (671)
Not DE (7,349)
Up (566)
Gene regulation upon circHIPK3 KD
D
IGF2BP2 KD K562
circHIPK3 KD UMUC3
2,500
998
1,984
P < 2.2e-16
E
UMUC3
P = 2e-15
P = 3e-21
P = 0.03
Genes affected by IGF2BP2 KD (%)
Down (657)
Not DE (7,220)
Up (561)
Gene regulation upon circHIPK3 KD
F
UMUC3
Regulation
Down
Up
Gene regulation upon IGF2BP2 KD (%)
Down (314)
Not DE (2,153)
Up (143)
Gene regulation upon circHIPK3 KD
G
IGF2BP2 KD, K562
CCCCAGC 4
CCCTGCC 5
CCCCTGC 40
GCCCCCA 42
CCCCCAG 44
AGGCCCC 53
GGCCCCC 62
GGCAGCC 115
CTGGAGA 9087
GGCCTCA 878
Motif score
Motif rank
H
K562
P = 2e-61
P = 6e-71
P = 5e-12
Genes with 11-mer (%)
Down (1,248)
Not DE (10,882)
Up (2,236)
Gene regulation upon IGF2BP2 KD
I
K562
P = 6e-19
P = 4e-67
P = 0.6
Genes with IGF2BP2 binding sites (%)
Down (1,248)
Not DE (10,882)
Up (2,236)
Gene regulation upon IGF2BP2 KD

Figure 4

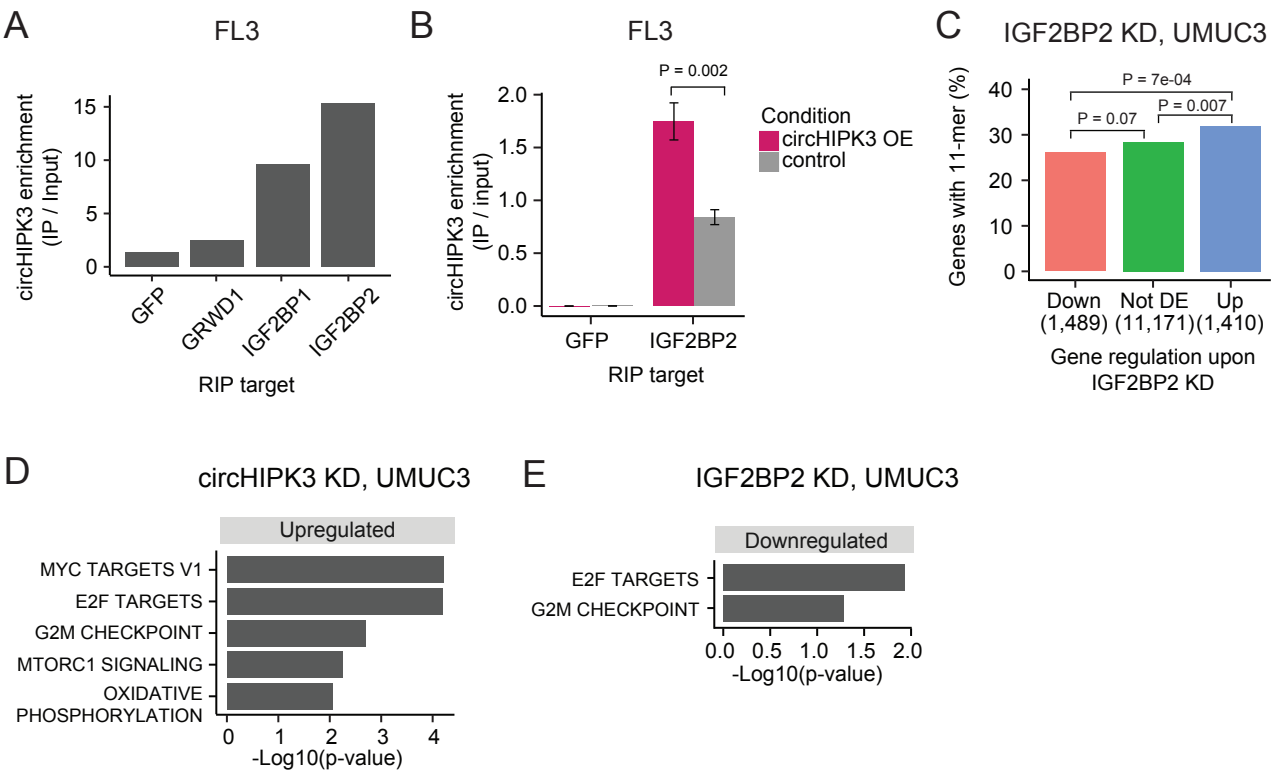

A
FL3
circHIPK3 enrichment (IP / Input)
15
10
5
0
GFP
GRWD1
IGF2BP1
IGF2BP2
RIP target
B
FL3
P = 0.002
circHIPK3 enrichment (IP / input)
2.0
1.5
1.0
0.5
0.0
GFP
IGF2BP2
RIP target
Condition
circHIPK3 OE
control
C
IGF2BP2 KD, UMUC3
P = 7e-04
P = 0.07
P = 0.007
Genes with 11-mer (%)
40
30
20
10
0
Down (1,489)
Not DE (11,171)
Up (1,410)
Gene regulation upon IGF2BP2 KD
D
circHIPK3 KD, UMUC3
Upregulated
MYC TARGETS V1
E2F TARGETS
G2M CHECKPOINT
MTORC1 SIGNALING
OXIDATIVE PHOSPHORYLATION
0
1
2
3
4
-Log10(p-value)
E
IGF2BP2 KD, UMUC3
Downregulated
E2F TARGETS
G2M CHECKPOINT
0.0
0.5
1.0
1.5
2.0
-Log10(p-value)

Figure 5

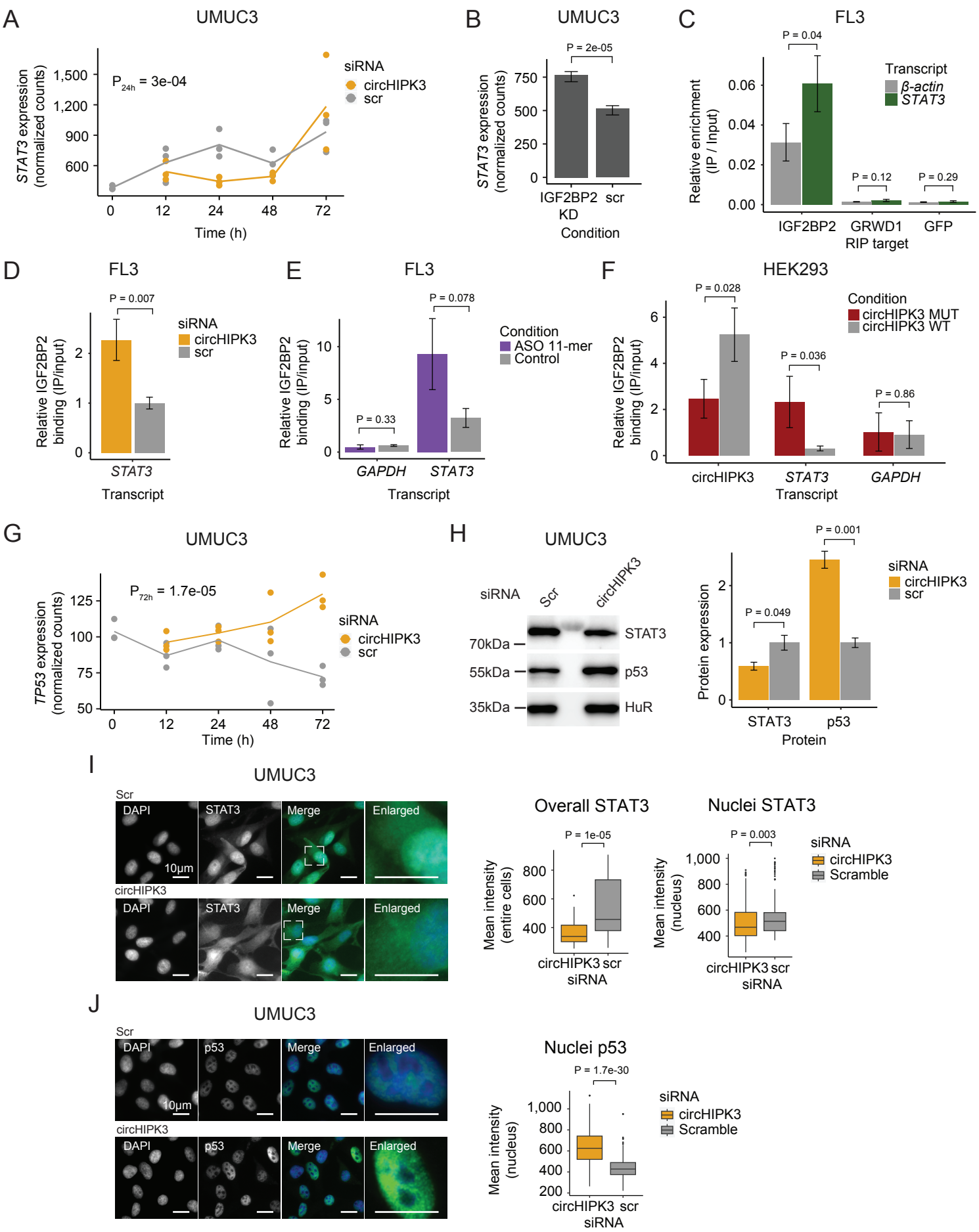
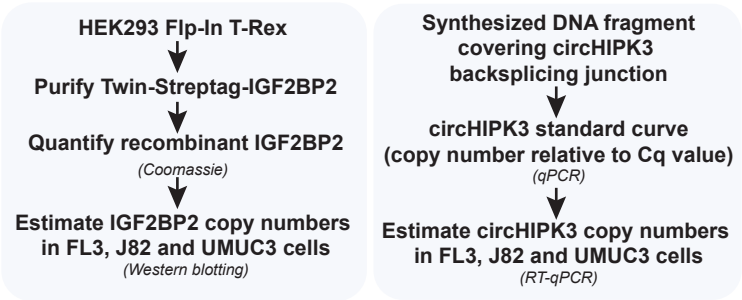

A
UMUC3
STAT3 expression (normalized counts)
P24h = 3e-04
siRNA
circHIPK3
scr
Time (h)
B
UMUC3
P = 2e-05
STAT3 expression (normalized counts)
IGF2BP2 KD
scr
Condition
C
FL3
P = 0.04
Relative enrichment (IP / Input)
Transcript
β-actin
STAT3
P = 0.12
P = 0.29
IGF2BP2
GRWD1
GFP
RIP target
D
FL3
P = 0.007
Relative IGF2BP2 binding (IP/input)
siRNA
circHIPK3
scr
STAT3
Transcript
E
FL3
P = 0.078
P = 0.33
Condition
ASO 11-mer
Control
GAPDH
STAT3
Transcript
F
HEK293
P = 0.028
P = 0.036
P = 0.86
Condition
circHIPK3 MUT
circHIPK3 WT
circHIPK3
STAT3
GAPDH
Transcript
G
UMUC3
P72h = 1.7e-05
TP53 expression (normalized counts)
siRNA
circHIPK3
scr
Time (h)
H
UMUC3
siRNA
Scr
circHIPK3
70kDa
55kDa
35kDa
STAT3
p53
HuR
P = 0.001
P = 0.049
Protein expression
STAT3
p53
Protein
I
UMUC3
Scr
DAPI
STAT3
Merge
Enlarged
10μm
circHIPK3
Overall STAT3
P = 1e-05
Mean intensity (entire cells)
Nuclei STAT3
P = 0.003
Mean intensity (nucleus)
siRNA
circHIPK3
Scramble
circHIPK3 scr
siRNA
J
UMUC3
Scr
DAPI
p53
Merge
Enlarged
10μm
circHIPK3
Nuclei p53
P = 1.7e-30
Mean intensity (nucleus)
siRNA
circHIPK3
Scramble
circHIPK3 scr
siRNA

Figure 6

A

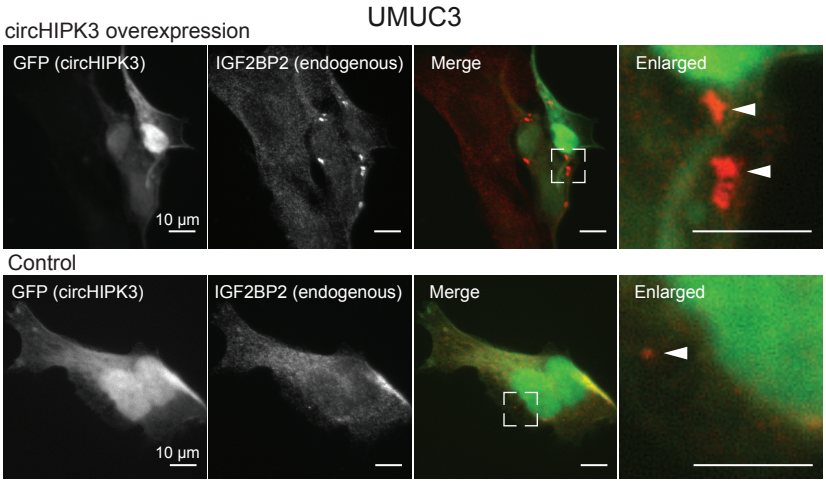


B

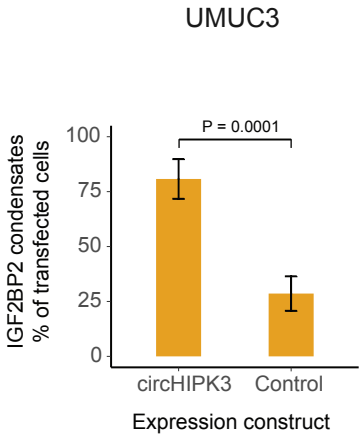
Cell line	circHIPK3*	IGF2BP2*
FL3	40	215892
J82	51	127483
UMUC3	45	260903

*Copies per cell

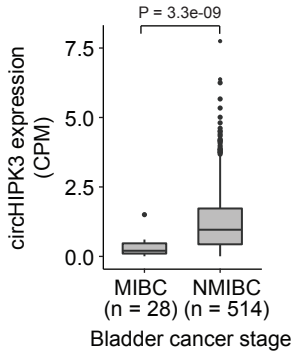
C



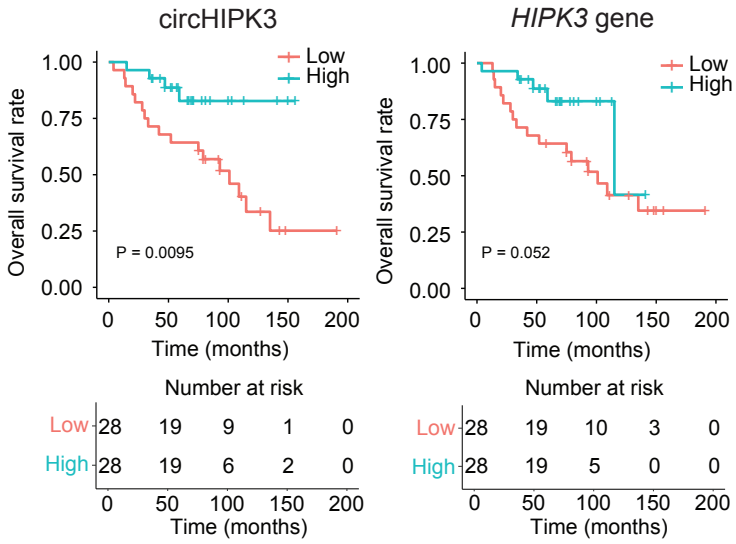
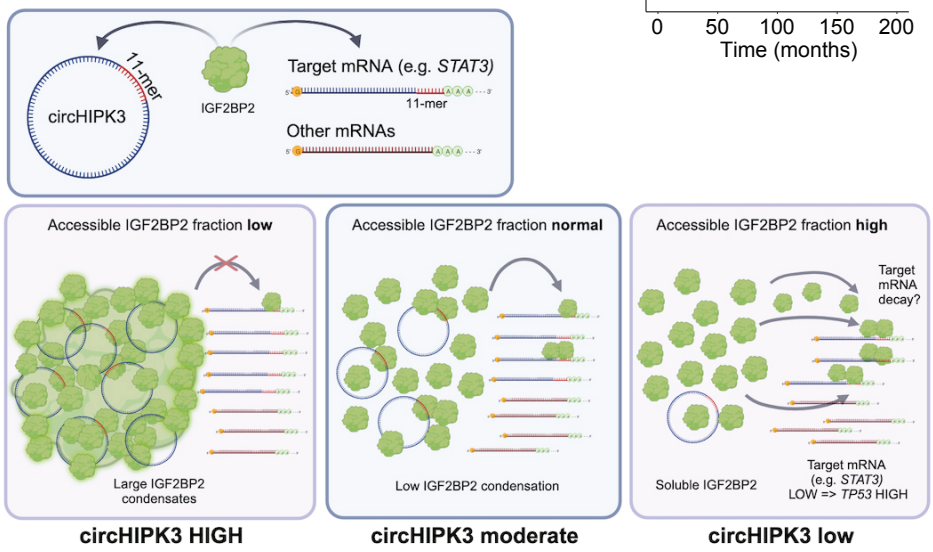
D



E



F



G