Functional landscape of circular RNAs in human cancer cells

Peter Hyunwuk Her[1,2,*], Tiantian Li[1,*], Ziwei Huang[3,*], Xin Xu[1], Weining Yang[1], Moliang Chen[1], Mona Teng[1], Sujun Chen[1], Yong Zeng[1], Stanley Liu[1,4,5], Benjamin Haibe-Kains[1,2], Fraser Soares[1,#], Jie Ming[3,#], Housheng Hansen He[1,2,#]

[1] Princess Margaret Cancer Center, University Health Network, Toronto, Canada
[2] Department of Medical Biophysics, University of Toronto, Toronto, Canada
[3] Department of Breast and Thyroid Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China
[4] Sunnybrook Research Institute, Sunnybrook Health Sciences Centre, Toronto, Ontario, Canada
[5] Department of Radiation Oncology, Sunnybrook Odette Cancer Centre, Toronto, Ontario, Canada
[6] Ontario Institute for Cancer Research, Toronto, Canada

[*] These authors contributed equally
[#]Correspondence: fraser.soares@utoronto.ca (F.S), mingjiewh@126.com (J.M) and hansen.he@uhn.ca (H.H.H)

Abstract

Circular RNAs (circRNAs) constitute a novel class of noncoding RNAs showcasing distinct tissue- and cell-specific expression patterns. Despite the extensive profiling of circRNAs, their individual functions remain poorly understood. To fill this gap, we designed a genome-wide library of 65,300 shRNAs, targeting 9,688 clinically relevant circRNAs and 3,981 of their linear parental genes, and conducted functional screening in seven types of human cancer. We identified a total of 1,342 essential circRNAs (13.8% of screened) that impact cell proliferation in at least one cell line, and in 96.5% of the cases, the linear counterparts are not essential. While the majority of circRNAs are functional in a cell type-specific manner, a shared common subset emerges as functional regulators across all examined cell lines. We further discerned the functional relevance of two cancer type-specific circRNAs, circMPP6(2,3,4,5) in prostate cancer and circACVR2A(2,3,4) in breast cancer. For a comprehensive presentation of the functional circRNA landscape in cancer, we introduce FunCirc, an online database

encompassing functional oncogenic circRNAs across cancer cell lines and several studies, coupled with circRNA expression profiles from diverse cancer and tissue types. Our work enhances the understanding of circRNA functions in cancer and provides the scientific community with a resource to further investigate their intricate roles.

Introduction

Research into cancer biology has traditionally centered around protein-coding genes; however, a recent surge in transcriptomic investigations has turned the spotlight onto other classes of RNA molecules in cancer. Remarkably, more than 90% of the RNAs originating from the human genome fall within the realm of noncoding RNAs (ncRNAs), many of which remain ambiguously defined[1,2]. With recent advances in RNA therapeutics, clinical trials involving ncRNAs are ongoing with the vast majority of ncRNAs being tested as diagnostic or prognostic biomarkers[3]. Classified primarily by size, ncRNAs encompass a diverse array of molecules, including small ncRNAs such as microRNAs, transfer RNAs (tRNAs), PIWI-interacting RNAs (piRNAs), small nucleolar RNAs (snoRNAs) and small nuclear RNAs (snRNAs), and long non-coding RNAs (lncRNAs), which typically exceed 200 nucleotides (nts) in length[4,5].

Among the myriad of ncRNA classes, circular RNAs (circRNAs) have emerged as a distinct and intriguing group. Generated through a unique back-splicing mechanism, circRNAs exhibit a covalently closed loop structure devoid of polyadenylated tails[4,5]. Although a minority may harbor protein coding potential, the majority function as ncRNAs, modulating gene expression and cellular processes through various mechanisms including interactions with transcription factors, sequestration of microRNAs, and scaffolding proteins[5]. Their circular structure confers enhanced stability and resistance against degradation by exoribonucleases, making them attractive candidates as diagnostic and prognostic biomarkers in cancer[6,7].

Transcriptomic profiling studies across various cancer cohorts have unveiled dysregulated expression patterns of circRNAs, hinting at their potential significance in tumorigenesis and disease progression[8–11]. In order to assess the functional significance of particular circRNAs, we conducted a pilot loss-of-function analysis using shRNA screens targeting 1,500 most abundantly expressed circRNAs in prostate cancer cell lines[9]. Out of the screened circRNAs, a collective of 171 (11.3% of the total screened) exhibited essential roles within at least one cell line, which in most cases, autonomously modulate cell survival, independent of their linear counterparts. Additional studies have leveraged RNA-targeting CRISPR loss of function screens to identify circRNAs with potentially important roles in modulating cellular function[11–14]. However, despite the wealth of data identifying circRNAs, the functional characterization of these molecules remains relatively poor, with studies being limited by small library size or within individual cancer type[9,11,12].

While numerous studies have documented circRNA expression across cancers, their functional roles remain largely uncharacterized due to a lack of systematic, large-scale functional resources. To bridge this gap, we developed FunCirc, a curated online

database integrating genome-scale functional screening of clinically relevant circRNAs. Our approach identified a core set of circRNAs essential for cell survival across multiple cancer cell lines and highlighted cancer type-specific dependencies. By centralizing functionally validated circRNA data, FunCirc serves as a valuable resource for RNA therapeutics and biomarker discovery, paving the way for more precise cancer treatment strategies.

Results

Curation and functional screening of clinically relevant circular RNAs

To identify circRNA for genome-scale functional screening, we compiled a comprehensive collection of clinically relevant circRNAs sourced from published cohort studies. Notably, oncogenic circRNA profiling data from Vo *et al.*[8] and Chen *et al.*[9] contributed mostly to this endeavor. Vo *et al.* uncovered and characterized 232 655 circRNAs in over 2,000 samples across various cancer types through exosome capture RNA sequencing. Chen *et al.* profiled the transcriptomic landscape of circRNAs in localized prostate samples using ultra-deep total RNA-seq on 144 tumors and characterized 76 311 circRNAs. Augmenting these datasets were in-house patient samples from breast (48 samples, 47 759 circRNAs), lung (63 samples, 15 253 circRNAs) and brain cancer (3 samples, 9 899 circRNAs) cohorts, contributing further to the curation of circRNAs for screening **(Figure 1A)**.

Integration of these datasets resulted in 250 380 circRNAs that were filtered based on their expression levels. First, circRNAs were filtered by expression in at least 10% of all samples of each cancer type. Then, they were further filtered by expression in the top 40% of each cancer type. This resulted in an initial list of 18 845 circRNAs for shRNA design, 1643 of which were overlapping in all cancer types.

We employed a multi-step scoring and filtering process to identify high-confidence shRNA candidates for these circRNAs while minimizing off-target effects (**Figure 1B, Methods**). First, we generated candidate sequences by extracting 21-mers across each transcript, skipping the first 25 nt at the 5′ end and the last 150 nt near the 3′ end to avoid poor or highly variable knockdown regions. For circRNAs specifically, we scanned a 34-nt region around the back-splice junction (BJS), capturing 17 nt on each side, and selected 21-mer windows overlapping at least 4 nt of the splice site to ensure circRNA-specific targeting. Each candidate then received an intrinsic score based on a set of rules that penalize or reward features predictive of successful knockdown or undesirable characteristics (e.g., extreme GC content).

Next, we computed a miRNA seed factor to account for potential microRNA-like off-target matches. Candidates with higher predicted seed complementarity to known miRNAs or off-target sites were penalized accordingly, while those showing minimal seed overlap were rewarded. Using the product of the intrinsic score and miRNA seed factor as an intermediate ranking metric, we then assessed sequence specificity via BLASTN searches against the targeted transcriptome. Perfect or near-perfect matches to unintended targets lowered a "specificity factor," while partial mismatches incurred smaller penalties.

Candidates above an acceptance threshold were further refined by computing a final adjusted score, integrating all intrinsic, seed, and specificity measures. Finally, to avoid overlapping shRNAs that might compete for the same BJS or regulatory regions, we enforced minimal spacing among selected candidates. Only those with acceptable overlap metrics and high adjusted scores advanced to the final library design phase, ensuring each circRNA target was represented by four highly specific shRNAs.

To assess the robustness of our scoring pipeline, we compared the final adjusted scores among the top four shRNA designs (Top 1–Top 4) for each circRNA. The distribution of Top 1 scores was shifted toward higher values compared to Top 2, Top 3, and Top 4, reflecting our rank-based selection criteria (**Figure 1C**). Pairwise correlation analysis among top1–top4 revealed moderate to strong positive correlations (Pearson's r ranging from 0.60 to 0.80; $p < 0.0001$), indicating consistent identification of high-scoring shRNAs across multiple ranks. We used a threshold of Adjusted Score > 3 and Adjusted Score > 5 to retain shRNAs with sufficient on-target potency and minimal predicted off-target effects, further refining our candidate list to 9,688 circRNAs. Collectively, these analyses demonstrate that our multi-step pipeline consistently prioritizes the most promising shRNA designs for each circRNA target.

Positive controls such as core essential linear genes and cell line specific essential genes were selected based on previous screening studies[15] and the DepMap database. Core essential circRNAs, and cell line specific essential circRNAs were also included from previous studies[8–10]. Negative controls encompassing non-targeting shRNA sequences for GFP, RFP, LacZ, and circRNAs from viral pathogens, as well as non-essential linear genes. The final shRNA library consisted of 65,300 shRNAs, with 4 shRNAs encompassing each target **(Figure 1D, E)**.

To assess the functional significance of circRNAs across diverse cancer types on a genome-wide scale, shRNA dropout screens were conducted employing the lentiviral library within 7 cell lines. The cell lines were carefully chosen to represent a broad spectrum of cancer types—including lung (A549), prostate (22Rv1), colorectal (HCT116), melanoma (A375), thyroid (FTC133), breast (MCF7), and cervical (HeLa) cancers, to ensure that our genome-wide shRNA dropout screens capture the functional relevance of circRNAs across diverse tumor contexts. Following puromycin selection,

genomic DNA was sampled on Days 0, 8, and 16, and shRNA inserts were PCR amplified for next-generation sequencing **(Figure 1A)**. All experiments were performed in triplicates to ensure robustness of the screening.

The shRNA sequencing data were processed using MaGeCK to identify depleted or enriched transcripts. To check for any positional bias in our shRNA designs, we compared the average knockdown across the four shRNA positions for each target. In A549 (**Figure 1F**) and across all cell lines (**Supplemental Figure 1A-F**), there was little difference among positions, indicating minimal positional effects.

We used one-way ANOVA followed by Tukey's HSD to compare the mean LFC values among the *circ*, *neg*, and *pos* groups across all cell lines. The negative control group exhibited a mean LFC that was 0.019 units higher than that of the circRNA group (diff = +0.01887, 95% CI: [0.00876, 0.02898], padj = 3.61×10^{-5}), indicating that circRNAs are modestly more depleted than negative controls. In contrast, the positive control group demonstrated a markedly more negative LFC compared to the circRNA group (diff = −1.78130, 95% CI: [−1.88595, −1.67665], padj < 1×10^{-16}) and the negative controls (diff = −1.80017, 95% CI: [−1.90481, −1.69553], padj < 1×10^{-16}). These results suggest that while circRNAs exhibit a moderate knockdown effect relative to negative controls, their depletion is substantially less pronounced than that observed in positive controls, aligning with the concept that circRNAs may modulate cellular processes in a more subtle manner than essential genes.

To assess additional potential off-target knockdowns, we calculated Z-scores for each shRNA relative to the distribution of negative controls. In A549 (**Figure 1G**), and across all cell lines (**Supplemental Figure 1G-L**) the Z-scores for circRNA and linear gene targets cluster around zero, similar to negative controls, suggesting minimal off-target effects (**Figure 1G**). By contrast, the positive-control set displays a clear left-shift (more negative Z-scores), reflecting strong knockdown as expected. Quantifying how many shRNAs exceed $|Z| > 2$ in each group in A549 (**Figure 1H**) and across all cell lines (**Supplemental Figure 1M-R**) confirms that positive controls have the highest fraction of "significant" hits (>30%), while circ, pos circ, and linear each show only a small fraction of off-target–level depletions. This demonstrates that the bulk of circRNAs do not exhibit large, unintended knockdown effects and supports the specificity of our shRNA library design.

To assess how consistently each target's four shRNAs produced similar knockdown effects, we calculated the standard deviation (SD) of the $\log_2$ fold change (LFC) for all targets in each group. In A549 **(Figure 1I)** and across all cell lines **(Supplemental Figure 1S-X)** circRNAs (*circ*) exhibit the lowest median SD, whereas positive controls (*pos*) display the highest variability, with the other groups (*pos circ*, *linear*, and *neg*) falling in between. In A549, Kruskal–Wallis test confirmed significant overall differences

among the five groups (Kruskal-Wallis chi-squared = 434.53, $p < 2.2 \times 10^{-16}$). Pairwise Dunn tests (Benjamini–Hochberg–adjusted) further indicated that *circ* had significantly lower SD than *linear*, *neg*, and *pos*, while *pos* was significantly more variable than all other groups ($p < 0.05$). This suggests that circRNAs are generally knocked down more consistently by their four shRNAs.

To further ensure that our shRNA design pipeline does not inadvertently introduce biases that could affect the screening outcomes, we evaluated the relationship between the final adjusted score and the observed biological effects. Specifically, we computed Pearson's correlations between the final adjusted score and both the MaGeCK RRA score and the $\log_2$ fold change for circRNAs. It is important to note that each shRNA has a score associated with it, leading to a large number of observations which can cause negligible effects to become statistically significant. In the case of A549, this produced very weak correlations, which were statistically significant ($r \approx -0.074$ for the RRA score and $r \approx -0.080$ for the LFC), with highly significant p-values ($p < 3.5\times10^{-13}$ and $p < 3.4\times10^{-15}$, respectively), a pattern which was consistent across most cell lines (**Supplementary Table 1 and 2**) . The statistical significance and the minimal correlation coefficients indicate that our sequence-based scoring metric does not substantially influence the observed knockdown effects. This lack of meaningful correlation supports the notion that our final adjusted score is not confounded by intrinsic sequence biases and that the functional readouts from the dropout screens reflect genuine biological effects rather than artifacts of shRNA design.

While our multi-step scoring pipeline effectively minimizes off-target effects and maximizes on-target knockdown, several limitations remain. First, in silico specificity checks cannot entirely eliminate the possibility of unanticipated off-targets, particularly in highly repetitive or under-annotated genomic regions. Second, circRNA annotations are still evolving, and our filtering criteria may exclude certain novel or lowly expressed circRNAs. Finally, we focused on 21-mer windows, but alternative shRNA lengths or designs might further optimize knockdown. Therefore, following the screen we sought to validate top circRNA hits with orthogonal methods to confirm both specificity and functional relevance.

Functional screening identifies known linear dependencies

Following processing using MaGeCK, as shown in 22Rv1 cells, the dropout or enriched levels, as measured by fold change of normalized shRNA abundances, was highly correlated between time points (Day 8 vs. Day 0 and Day 16. vs Day 0) for all the screens **(Figure 2A, Supplementary Table 3)**. Furthermore, positive controls showed a significantly higher dropout score (Mann-Whitney U test = $2.56\text{x}10^{-32}$, CLES = 0.21), and negative controls showed a significantly lower dropout score (Mann-Whitney U test =

8.61×10^{-48}, CLES = 0.55) **(Figure 2B, Supplementary Figure 2A-F)**, compared to all the linear genes, respectively ("linear" group in **Figure 1B**). The depleted linear transcripts from our screens demonstrated significant positive correlation with publicly available screening data in corresponding cell lines obtained from DepMap[16] **(Figure 2C, Supplementary Table 4)**.

We next investigated the cell line-specific and common dropout and enriched linear transcripts identified from our screens. We discovered a total of 491 dropout genes and 363 enriched genes that show significant dependency in at least one cell line. Among the seven cell lines analyzed, HCT116 displayed the highest number of essential dependencies, with 257 identified, while FTC133 showed the lowest, with 173 dependencies **(Figure 2D)**. In total, 580 linear genes demonstrated unique dependencies in individual cell lines across cancer types, with the largest subset consisting of 95 dependencies specific to MCF7. Notably, several of these genes, such as MED1, ESR1, HHAT and FOXA1, have been previously reported to be involved in breast cancer biology[17,18,19,20]. Upon analyzing the combined score, which accounts for both the fold change and the statistical significance, MCF7 cell line showed the most substantial changes across these genes, further confirming the prominent role for these genes in breast cancer **(Figure 2E)**. A total of 274 genes exhibited dependencies in two or more cell lines, with a set of six genes, including COPA, CDC42, LONP1, SUPT6H, WDR3 and RPS24, depleted across all cell lines (**Figure 2F, Supplementary Figure 3G**). These common essential genes have been previously identified and characterized as oncogenes linked to various cancers [21,22,23,24,25,26].

Taken together, our screens identified known cell-type specific and common linear gene dependencies, validating the fidelity of our screens.

Functional screening identifies core and cell-type specific essential circRNAs

Building on the insights gained from linear transcripts analyses, we utilized MAGeCK to determine the depletion or enrichment of circRNAs in our screens. Similar as observed for linear transcripts, the changes in normalized abundance of shRNAs targeting circRNA back-splicing sites, measured as fold change, showed a strong positive correlation between time points (Day 8 vs Day 0 and Day 16 vs Day 0) in 22Rv1 (correlation = 0.90, p-value = < 0.001) **(Figure 3A)**, a pattern consistent across all screens in other cell lines **(Supplementary Table 5)**. Furthermore, positive control circRNAs in 22Rv1 showed a significantly higher dropout compared to all circRNAs (4.95×10^{-54}, CLES = 0.34), while negative controls showed a significantly lower dropout score (Mann-Whitney U test = 2.28×10^{-36}, CLES = 0.54) **(Figure 3B)**, a finding replicated in majority of the screens **(Supplementary Figure 2G-L)**.

Our differential shRNA abundance analysis identified a total of 3,573 circRNAs showing differential dependencies in at least one cell line. The number of differential

dependencies ranged broadly, from 571 in FTC133 to 1,198 in 22Rv1 **(Figure 3C)**, reflecting the diverse dependence of circRNAs across the cell lines. Importantly, in 96.5% of the cases, circRNA dependencies were distinct from their linear parental genes in at least one cell line **(Figure 3D, Supplementary Figure 3A-F).** Among these circRNAs, 1,955 demonstrated unique dependencies in individual cell lines, with prevalence varying from 12-16% across different cell lines **(Figure 3C)**. The largest subset, consisting of 331 dependencies, was specific to 22Rv1 cells. Notably, many of these, including cricINO80, circMPP6(2,3,4,5) and circUSP34, with functions unknown in prostate cancer **(Figure 3E)**.

While the prevailing trend in circRNA dependencies predominantly displayed functional uniqueness associated with specific cancer cell types, we identified 1,618 circRNAs that showed dependencies in at least two cell lines. Additionally, a core group of circRNAs demonstrated shared functionality across all the seven cancer cell lines screened. This group includes six essential or positive regulators of cell growth (circABHD17B(4,3,2), circGOLIM4(8,7,6), circMTR(3-28), circPSEN1(2,3), circPTK2(7,6,5,4), and circPTK2(7,6,5,4)) and four negative regulators of cell growth (circADGRB3(18,19,20), circARFGEF1(4,3), circCUL1(2-9), and circRacGAP1(17,16,15)) **(Figure 3F)**. Importantly, isoforms of circPSEN1, circARFGEF1, circPTK2 and circRACGAP1 align with previously reported functional roles, reinforcing the robustness and reliability of our screening approach.[27–34].

To evaluate the scope and uniqueness of our circRNA screening library, we compared the circRNAs included in our study with those from previous functional screens. Our library encompassed 9,853 circRNAs and 190 positive control circRNAs, which is significantly more than Li et al.[12] (762 circRNAs), Liu et al.[35] (3,354 circRNAs), and Chen et al.[9] (1,507 circRNAs). After accounting for shared circRNAs, 139 circRNAs were unique to Li et al., 199 to Chen et al., and 1087 to Liu et al., highlighting that our study covers a broader and more comprehensive set of circRNAs (**Supplementary Figure 4A**). This substantial increase in the number of circRNAs screened demonstrates the comprehensiveness of our library, providing a significantly expanded resource for the identification and functional characterization of circRNAs in cancer.

To further assess the utility of our screening, we compared the essential circRNAs identified in our study to those reported by Liu et al., who identified 1,756 circRNAs essential in at least one cell line. Our study identified 1,342 essential circRNAs (plus 78 essential positive control circRNAs) resulting in 1,420 unique circRNAs essential in at least one cell line (**Supplementary Figure 4B**). Of these, 1,220 were novel and did not overlap with those identified by Liu et al. Notably, 200 circRNAs were shared between the two studies. Furthermore, we observed positive correlations in the $\log_2$ fold changes of overlapping circRNAs screened in our A375 cell line with the skin cell lines from Liu et al. ($r = 0.23$, $p = 2.05 \times 10^{-3}$) (**Supplementary Figure 4C**). This overlap, along with the

identification of novel candidates, underscores the utility of our screen in expanding the understanding of circRNA functionality across diverse cellular contexts.

We also compared our results with those from Li et al., who utilized CRISPR-Cas13 screen to identify essential circRNAs in three different cell lines: HT29 (67 circRNAs), 293FT (62 circRNAs), and HeLa (63 circRNAs). The overlaps from this comparison were minimal, with most circRNAs being unique to our screening study (**Supplementary Figure 4D**). Specifically, only 10 essential circRNAs were shared with HeLa, two with 293FT, and one across all studies, highlighting the expanded coverage of our methodology and the unique contributions in identifying additional circRNAs that may have been missed by previous targeted screens.

When compared to our previous work in prostate cancer, 76 of the previous 171 essential circRNAs identified by Chen et al. overlapped with our essential circRNAs (**Supplementary Figure 4E)**. In addition, in the 22Rv1 prostate cancer cell line, our findings correlated with those from Chen et al. ($r = 0.25$, $p = 1.19 \times 10^{-3}$) (**Supplementary Figure 4F**). Collectively, these comparisons with previous studies not only validate our screening approach but also emphasize the novelty and utility of our work in expanding the known landscape of functional circRNAs in cancer.

Functional validation of core dropout and enriched circRNAs

To confirm the functionality of the core circRNA dependencies, we first designed divergent primers targeting back-splice junctions. PCR amplification of cDNA from two representative cell lines 22Rv1 and A549, followed by sanger sequencing, verified the back-splicing sites for nine out of the ten core circRNAs. circSCFD1(7,8,9,10), due to unsuccessful amplification, was excluded from subsequent analyses. Subsequent RNase R treatment showed that these circRNAs are highly resistant to exonuclease mediated degradation, contrasting sharply with the linear gene GAPDH, which demonstrated high sensitivity to the treatment **(Figure 4A, B)**.

To assess the functional role of the core circRNAs, we designed three siRNAs targeting the back splice junction sites for the five core dropout circRNAs and the four enriched circRNAs **(Figure 4C)**. We performed cell proliferation assays following the knockdown (KD) of each circRNA in 22Rv1 and A549 cells. Cell proliferation was monitored using the Incucyte live cell imaging system, and fold changes and statistical analyses were performed between the knockdown and non-targeting siRNA control groups when the former reached a plateau. Upon siRNA-mediated knockdown of these circRNAs, we observed a significant reduction in cell proliferation for the five dropout circRNAs posited to enhance cell survival, consistent with their anticipated biological roles **(Figure 4D)**. In contrast, siRNA silencing of the four enriched circRNAs implicated in growth inhibition resulted in a marked increase in proliferation rates in both cell lines **(Figure 4E)**. These

findings suggest that the core circRNAs play crucial roles in modulating cell growth across different cancer types. This series of experiments not only highlights the functional importance of these core circRNAs, but also validates the reliability of our screening.

Context specific dependence of circRNA in prostate cancer

In an extension of our validation efforts, we further investigated the function of circRNAs specific to individual cell lines. Notably, circMPP6(2,3,4,5), a previously unreported circRNA, demonstrated dropout exclusively in the 22Rv1 cells. The circular transcript, encompassing 653 nucleotides, is composed of four exons produced by back splicing of exons 2 through 5 **(Figure 5A)**. To assess the presence of circMPP6(2,3,4,5) across our panel of screened cell lines, we designed divergent primers targeting this specific back-splice junction. Our analysis revealed that while parental linear MPP6 transcript demonstrated comparable abundance in all the seven cell lines, circMPP6(2,3,4,5) is preferentially highly expressed in 22Rv1 cells, underscoring its potential relevance in this cancer type **(Figure 5B)**.

The circular nature of circMPP6(2,3,4,5) was further confirmed through RNAse-R treatment, which significantly reduced the level of linear MPP6, while the circular form demonstrated a high level of resistance to the treatment. Sanger sequencing of the PCR amplicons verified the sequence of the back-splice junction site **(Figure 5C, D)**.

To explore the functional significance of circMPP6(2,3,4,5), we designed two siRNAs specifically targeting the back-splicing site. RT-qPCR analysis following siRNA treatment demonstrated an approximate 60% reduction of circMPP6(2,3,4,5) abundance in 22Rv1 cells, with no obvious effect on the level of the linear parental gene (**Figure 5E**), affirming the specificity of these siRNAs. Unlike in other cell lines, where circMPP6(2,3,4,5) knockdown did not yield noticeable impact, the proliferation of 22Rv1 cells was significantly reduced, highlighting a critical, cell line-specific role of circMPP6(2,3,4,5) **(Figure 5F)**.

Next, we sought to further characterize the biological functions of circMPP6(2,3,4,5) by performing invasion and migration assays in 22Rv1 cells **(Figure 5G)**. Knockdown of circMPP6(2,3,4,5) led to a considerable decrease in these metastatic capabilities, suggesting a role in controlling prostate cancer cell progression.

Further insights into the function of circMPP6(2,3,4,5) were gained through RNA sequencing analysis following circMPP6(2,3,4,5) knockdown. This analysis identified 36 downregulated and 37 upregulated genes **(Figure 5H)**. The downregulated genes were significantly enriched in terms associated with genome stability processes such as double-strand break repair via homologous recombination and chromosome organization. In contrast, genes activated by circMPP6(2,3,4,5) knockdown were

predominantly linked to regulation of blood pressure, inflammatory response, regulation of cytokine production, and the promotion of programmed cell death and apoptosis **(Figure 5I)**.

Context specific dependence of circRNA in breast cancer

Having successfully validated the function of circMPP6(2,3,4,5) in prostate cancer, we extended our investigation to explore additional context specific essential circRNAs across different cancer types. circACVR2A(2,3,4) emerged as a top ranked cell-line specific circRNA, displaying unique shRNA dropout in MCF7 breast cancer cells. This circRNA comprises three exons, generated through back splicing of exons 2,3 and 4 with a total length of 437 nucleotides **(Figure 6A)**. The back-splice junction sequence was confirmed by sanger sequencing and gel electrophoresis analysis of the PCR amplicons **(Supplementary Figure 5A)**. To assess the presence of circACVR2A(2,3,4) across our panel of screened cell lines, we designed divergent primers targeting this specific back-splice junction. Our analysis revealed that similar to the parental linear ACVR2A transcript, circACRV2A demonstrated comparable abundance in all the seven cell lines, underscoring that circRNA expression alone might not be indicative of its function (**Figure 6 B**).

Expression analysis revealed that circACVR2A(2,3,4) levels were significantly higher in breast cancer cell lines, including luminal subtype MCF7 and T47D, and triple-negative subtype MDA-MB-231, BT549 and HCC1937, compared to the breast epithelial cell line MCF10A **(Figure 6C)**. Subcellular fractionation followed by RT-qPCR analysis revealed that circACVR2A(2,3,4) predominantly localizes to the cytoplasm in T47D and MDA-MB-231 cells **(Supplementary Figure 5B)**. This cytoplasmic localization was further confirmed by RNA fluorescence in situ hybridization (FISH) assays, which visually depicted the circACVR2A(2,3,4)'s distribution in cytoplasm **(Figure 6D)**. Treatment with Actinomycin D demonstrated that circACVR2A(2,3,4) is highly stable over time in both MDA-MB-231 and T47D cell lines, compared to the linear counterpart **(Figure 6E),** suggesting long half life of the circular transcript.

To investigate the specific function of circACVR2A(2,3,4), we designed two siRNAs specifically targeting the back-splice site. RT-qPCR analysis showed that both siRNAs effectively reduced the expression level of circACVR2A(2,3,4), with minimal impact on the expression of parental linear transcript **(Supplementary Figure 5C)**, indicating the specificity of the siRNAs. We subsequently assessed the effect of circACVR2 knockdown on cell proliferation across a panel of cell lines. While circACVR2A(2,3,4) knockdown did not significantly affect proliferation in 22Rv1, A549, HCT116, FTC133, HeLa and A375, a notable decrease in proliferation was observed in MCF7 cells **(Figure 6F)**, consistent with our shRNA screening results showing specific shRNA dropout in MCF7 cells. Significant reduction in cell proliferation was consistently observed in additional breast cancer cell line model MDA-MB-231 and T47D cell

(Figure 6G), underscoring the specific impact of circACVR2 on breast cancer cell proliferation. Furthermore, colony formation assays demonstrated a substantial reduction in colony numbers following circACVR2A(2,3,4) silencing **(Figure 6H)**. This was corroborated by EdU assays, which showed a decrease in DNA synthesis rates in cells with circACVR2A(2,3,4) knockdown **(Supplementary Figure 5D)**. Together, these results highlight the critical role of circACVR2 in regulating cell proliferation, specifically in breast cancer cells.

To elucidate the mechanism underlying the function of circACVR2A(2,3,4), RNA sequencing analysis was performed in MDA-MB-231 cells with and without siRNA-mediated silencing of this circRNA. A total of 89 downregulated and 49 upregulated genes were identified upon silencing of circACVR2A(2,3,4) **(Figure 6I)**. The downregulated genes were significantly enriched in terms associated with genome stability processes such as centrosome duplication, centrosome cycle and microtubule organizing center organization. In contrast, genes activated by circACVR2A(2,3,4) knockdown were predominantly linked to regulation of chemotaxis, sensory perception, bone morphogenesis, extracellular matrix organization, extracellular structure organization and external encapsulating structure organization **(Figure 6J)**.

To further assess the function of circACVR2A(2,3,4) *in vivo*, we utilized murine breast cancer 4T1 cells treated with either control shRNA (shNC) or shRNA targeting circACVR2A(2,3,4) (shcircACVR2A). These cells were subcutaneously injected into BALB/c mice to establish xenografts. The resulting tumors from the shcircACVR2A-infected cells exhibited significantly reduced growth rate and volume compared to those from the control group **(Figure 6K-L)**, indicating that circACVR2A(2,3,4) promotes tumor growth. Live imaging of these subcutaneous tumors provided additional confirmation of the attenuated tumor growth in the shcircACVR2A group, visually documenting the reduced tumor progression **(Figure 6M)**.

These results collectively underscore the critical role of circACVR2A(2,3,4) in regulating breast cancer growth both *in vitro* and *in vivo*.

FunCirc: a circRNA Function and Clinical Annotation Database

Leveraging data generated from our transcriptome-wide and previously published targeted functional circRNA screens, we developed FunCirc, a comprehensive and open compendium tailored for the scientific community. FunCirc provides clinical annotations and essentiality information for circRNAs across seven major cancer cell types **(Figure 7)**. Designed with the needs of the research community in mind, FunCirc offers an extensive collection of over 10,000 circRNAs across multiple studies, enabling detailed exploration of their functions and relationships with their linear counterparts through a user-centric interface (https://funcirc-shiny-153878979048.northamerica-northeast2.run.app/).

The platform allows for users to investigate the functional roles of circRNAs invarious cell lines and timepoints, integrating circRNA expression profiles in clinical datasets of over 2,200 patient tumors and cell lines, collected from public studies as well as in-house cohorts [8,9]. FunCirc allows for the selection of splice isoforms of interest, generating bar plots that depict the normalized count per million reads. This feature provides a nuanced view of circRNA prevalence across different cancer types, aiding in the identification of potential therapeutic targets and biomarkers. Moreover, the FunCirc dataset is quantitatively robust, comprising uniformly processed and normalized data ready for immediate meta-analysis and in-depth downstream examination. This standardized processing contrasts with other databases that amalgamate data from diverse sequencing protocols, introducing platform-specific biases that can complicate comparative analysis.

Consequently, FunCirc stands out as a singularly curated, quantitative portal, hosting an unparalleled range of cancer-specific circRNAs derived from clinical samples. This platform provides a reliable and comprehensive platform for advancing our understanding of circRNAs in the context of cancer biology and therapy, enables researchers to unlock new insights into the molecular underpinnings of cancer and beyond.

Discussion

RNA sequencing technologies have significantly expanded our knowledge of the human transcriptome, revealing the presence of hundreds and thousands of circRNAs. Despite their dynamic expression pattern in various tissues and diseases, including cancer, the functions of these transcripts remain largely unexplored. Majority of circRNAs function as noncoding RNA, and share transcription start sites with their parental linear transcripts, features render them not suitable for recently developed CRISPR screening techniques that target protein-coding gene exon DNA or the transcriptional start sites of lncRNAs[36]. Three recent studies[9,11,12], including one from our group, have utilized pooled shRNA and Cas13 sgRNA designed to specifically target the unique back-splice junctions of circRNAs. However, these efforts have been limited by small library sizes and the number of cell lines used, resulting in an incomplete picture of circRNA functionality.

In this study, we designed over 60,000 shRNAs to target approximately 10,000 cancer-associated circRNAs and their parental linear transcripts, and conducted comprehensive functional screens in cell lines from seven major cancer types. These extensive screens identified a common group of essential circRNAs across multiple cancer cell lines, akin to 'fitness genes' in protein-coding gene screens. Some of these circRNAs have already been documented to play functional roles in specific cancer

types. The revelation of their consistent and shared roles across various cancers is of significance, highlighting that circRNAs may have distinct regulatory capabilities, serving as key determinants of cell survival and cancer progression. Ongoing and future research aimed at elucidating how these shared functions are regulated to execute specific biological actions will deepen our understanding of circRNAs in cancer biology[13,37].

Our comparisons with previous studies highlight not only the robustness of our screening approach but also its significant contribution to expanding the landscape of functional circRNAs in cancer. By screening a larger number of circRNAs across multiple cancer types, we identified both common and unique essential circRNAs. The substantial overlap with previous studies validates our findings, while the unique circRNAs identified exclusively in our study suggest new avenues for research and potential therapeutic targets.

Our investigation uncovered hundreds of circRNAs exhibiting specific functions unique to certain cell lines, reinforcing the concept of context-dependent expression of circRNAs[38]. Notably, we identified circMPP6(2,3,4,5), previously unreported and now recognized as essential in prostate cancer cells. Membrane palmitoylated proteins (MPPs), including MPP6, are involved in crucial biological processes such as cell adhesion and polarity and are often dysregulated in cancers[39,40]. Interestingly, the linear counterpart of MPP6 was associated with prostate cancer progression and was significantly overexpressed in hepatocellular carcinoma[39]. Specifically, the genes exhibiting differential expression were primarily concentrated in processes related to the synthesis of genetic materials, alongside the WNT signaling pathway. Despite some pathway overlaps between the circular and linear target genes, our screening results suggest a unique role for circMPP6(2,3,4,5), distinct from its linear counterpart, particularly since MPP6 itself was not essential for cell survival.

Another circRNA, circACVR2A(2,3,4), proved essential in breast cancer but not in other studied cell lines, despite known roles in suppressing bladder cancer proliferation and metastasis via the miR-626/EYA4 pathway[41]. This highlights the versatile and complex regulatory roles of circRNAs, which can exhibit diverse functions across different cancer types. Our findings suggest that circRNAs like circACVR2A(2,3,4) can adapt to distinct cellular environments, underscoring their potential as pivotal elements in cancer progression and therapy resistance. However, the limited range of cancer cell lines tested poses a constraint, and the full spectrum of circRNA functionality might extend beyond our current observations.

Although shRNA screening offers several advantages, its application to circRNAs presents unique challenges. The design of shRNAs targeting the back splice junctions of circRNAs can lead to overlaps, potentially resulting in skewed or biased screening results. This overlap is a significant confounder that could obscure the accurate

assessment of circRNA function through shRNA-based methods. In our study, several circRNAs were excluded because an overlap-free shRNA design was unfeasible, necessitating the removal of these candidates to avoid off-target effects from their linear parental genes. A promising alternative that we explored is the use of RNA-targeting type VI CRISPR effectors, known as Cas13 RNases[11,12]. These emergent tools offer a distinct approach for studying circRNAs and could serve as valuable assets for cross-validating experimental outcomes. Considering their reported precision and unique guide design, these tools could provide a new avenue to explore circRNAs for which shRNA design is challenging. However, it's important to recognize that, like any evolving technology, the efficacy and reliability of Cas 13 RNases and similar approaches require further development and validation to address their varying extent of efficiency and off-target effects[42,43].

The increasing interest in circRNA research has led to the development of several circRNA databases such as MiOncoCirc[8], CircBase[44], CIRCpedia[45], circRNAdb[46], and CSCD[47]. FunCirc stands out as the first extensive, cancer-focused resource that integrates functional and clinical data on circRNAs. FunCirc provides fold change values in cellular proliferation across various cancer cell lines post-knockdown of circRNAs and their corresponding linear parental genes. This dual assessment not only highlights the unique contributions of circRNAs but also explores the potential interactions between circRNAs and their linear counterparts in cancer progression. Moreover, FunCirc bridges laboratory findings with clinical data, enriching circRNA research with a necessary clinical dimension.

In summary, we report comprehensive transcriptome-wide circRNA shRNA screens, offering functional annotations across cell lines of diverse cancer types. Our extensive analysis uncovered a group of core circRNAs that positively or negatively regulate cell survival across all cell lines examined as well as hundreds of circRNAs that are crucial for cell survival in a cell-type-specific manner. By validating the functional relevance of circRNAs across various cancer subtypes, our findings underscore their broad impact on cancer biology and the potential impact of circRNAs across varied cancer types. Combining functional and clinical data, our work establishes a robust foundation for future research endeavors, aiming to further elucidate the multifaceted roles of circRNAs in cancer and potentially other diseases.

Methods

shRNA library design

Clinically relevant circRNAs for functional validation in cell lines were curated from a list of oncogenic circRNAs compiled from the following published cohort studies: Vo JN et al. 2019, Chen et al. 2019, Wang et al. X, and from additional in-house patient samples [46 Breast cancer cohort (10 normal: 36 tumor) and 3 brain cancer samples). CircRNAs were prioritized based on expression across samples and cancer type. In brief, circRNAs expressed in at least 10% of the samples per cancer type were selected and the top 30-40% expressed in each cancer were selected for library design (9,688 circRNAs). shRNAs were designed for each circRNAs as previously described (Chen et al. 2019). In brief, a 34-nucleotide long sequence around the back splice junction site, with 17 nucleotides on each side, was extracted. A custom script was developed to perform batch shRNA designing with the GPP web portal (https://portals.broadinstitute.org/gpp/public/seq/search). To minimize off-target effects for all candidates, we chose one representative RefSeq transcript and scanned from 25 nt downstream of the coding sequence (CDS) start to 150 nt upstream of the transcript end. We extracted every 21 nt window within this region as a potential short hairpin RNA (shRNA) "candidate." This exclusion of the extreme 5′ and 3′ regions aims to avoid regulatory elements and regions prone to poor knockdown. Each 21 nt candidate received an intrinsic score derived from rules that either reward or penalize features predictive of efficient knockdown. Factors considered include GC content, presence or absence of specific motifs, and positional attributes (e.g., avoiding extremely close proximity to the CDS boundaries). The sum of these rewards and penalties yields a single numeric score (the "intrinsic factor") for each candidate. To minimize off-target effects, each candidate was aligned against the entire transcriptome. Candidates exhibiting near-perfect identity (≥ 19 out of 21 nt) to non-target transcripts were heavily penalized or removed altogether. Partial matches (≥ 17 nt) incurred smaller penalties. In certain cases, if a candidate aligned uniquely to a functionally important target (e.g., a known essential gene), a small positive "bonus" factor was applied. The result of this process is a "specificity factor," which, if it fell below a threshold (e.g., 0.7), caused the candidate to be excluded. The final adjusted score for each candidate was computed by combining (often via multiplication) the intrinsic factor, the specificity factor, and additional penalty or reward terms—such as a miRNA seed factor—to account for potential microRNA-like off-target effects. Candidates with very low specificity or excessively high off-target penalties scored close to zero and were not considered further. In typical linear gene knockdown designs, only the CDS and possibly the 3′ UTR were targeted. Candidates overlapping the CDS–3′ UTR boundary were included if they met length and scoring requirements. For circRNA targets, we additionally scanned across the back-splice junction, ensuring at least 4 nt overlap on each side of the junction to achieve circRNA-specific targeting. Finally, all potential targets were sorted

by the adjusted score. For circRNA parental transcripts, the shRNA candidates were scanned from the TRC shRNA library.

The shRNA oligo array was synthesized as 92-mer oligonucleotides (Twist Biosciences), GGAAAGGACGAAACACCGGNNNNNNNNNNNNNNNNNNNNCTCGAGNNNNNNNNNNNNNNNNNNNNNNNNTTTTTGAATTCTCGACCTCGAGACA (N's denote shRNA 21nt-target sequence, sense and antisense). Library construction has been previously described (Chen et al., 2019). In brief, the array was amplified by PCR (KAPA HiFi HotStart ReadyMix, KK2601) as a pool using the following primers: TAACTTGAAAGTATTTCGATTTCTTGGCTTTATATATCTTGTGGAAAGGACGAAACACCGG (Forward) and CCCCCTTTTCTTTTAAAATTGTGGATGAATACTGCCATTTGTCTCGAGGTCGAGAATTC (Reverse). The PCR product was purified and then cloned in the pLKO.1 vector using AgeI/EcoRI (NEB). Ligation was performed using the NEBuilder HiFi DNA Assembly Cloning Kit (NEB, Cat. #E5520S) and transformed into an electrocompetent strain (Endura™ Electrocompetent Cells, 24 rxns (DUOs), Cat. 60242-2). 10-15 Gibson ligation reactions were pooled together to achieve > 300X coverage. Colonies were scraped off plates using LB and plasmid DNA was extracted (NA0310 Sigma GenElute HP Plasmid Maxiprep Kit). The library was submitted for next generation sequencing to confirm adequate library representation of each shRNA.

shRNA library screening and analysis

Library virus was generated as previously described [9] and each cell line was titrated with library virus to achieve a low MOI (~0.3-0.4). Specifically, MOI was determined by infecting 5-10 million cells with varying amounts of library virus for 24 hours, which were then passed into media with or without 1ug/mL puromycin (ThermoFisher) for 72 hours. A ratio between these two populations was calculated to determine the infection efficiency to achieve a MOI of 0.3. The amount of library virus was scaled up along with the number of cells to ensure that on average every shRNA was represented in 300 cells. For each screen, cells were split into 3 independent batches, passed every 3-4 days, and maintained at 300x coverage for the duration of the screen. A sample was taken from each independent batch on Day 0 (T0), Day 8 (T8) and Day 16 (T16) post-selection for genomic DNA analysis and ~40-50 million cells were collected per sample. Genomic DNA was extracted (QIAamp Blood Maxi kit, Cat. # 51194) and shRNA inserts were amplified by PCR using standard Illumina next-generation sequencing library preparation kits. The input amount of genomic DNA was calculated to achieve 250x coverage of the library, which was then sequenced on an Illumina HiSeq 2500. Prior to performing screens, all cell lines were authenticated by STR (Geneprint10 panel, Sickkids TCGA) and tested for mycoplasma contamination using the EZ-PCR mycoplasma Test Kit (20-700-20, Biological Industries). All screens were performed with low passage cell lines starting between passages 3-6. Prior to library virus infection, 5 million cells of each cell type were collected for downstream RNA-Seq.

Essential circRNAs were identified from shRNA screen sequencing results with MAGeCK (v0.5.4) software[48]. Sequencing FASTQ files were first trimmed so that the shRNA sequence starts at the first base. The counts of each shRNA in different time points were then counted from the trimmed files with MAGeCK 'count' function with default parameters. The Robust Rank Aggregation (RRA) algorithm was applied on the count table. A p value of 0.01 was used as threshold for essentiality.

Cell culture

MCF7, HeLa, A549, 22Rv1, HCT116, and A375 cell lines were obtained from the American Type Culture Collection (ATCC® CCL-185, CRL-2505, CCL-247, CRL-1619), FTC133 from Sigma (94060901-1VL). 22Rv1 and A549 cells were cultured in RPMI1640 media, HCT116 in McCoy's 5A media, MCF7, HeLa and FTC133 in DMEM media. All cell culture media was supplemented with 10% FBS (Wisent) and 1% Penicillin and Streptomycin (450-201-EL, Wisent). 293FT cells were cultured in DMEM medium containing 10% FBS (080150, Wisent), L-glutamine (25030-081, ThermoFisher), and non- essential amino acids (11140–050, ThermoFisher) supplemented with 500 µg/mL Geneticin (4727894001, Sigma-Aldrich). All cells were cultured at 37° in 5% CO2.

Transfection

Control siRNA and all siRNAs targeting linear and circular candidate genes were purchased from MilliporeSigma. Lipofectamine RNAiMax reagent (ThermoFisher; Cat. #13778150) was used to transfect siRNAs. Cells were plated and transfected with the siRNA-lipid complex diluted in Opti-MEM I Reduced Serum Medium (ThermoFisher; Cat. #31985070).

Cell Proliferation Assays

Cells were seeded in a SARSTEDT 96-well cell culture plate (Cat. # 83.3924) to be 60-80% confluent at transfection. Cells were imaged for 7 days using IncuCyte S3 live cell imaging system (Sartorius, Göttingen, Germany) and the surface area occupied by the cells was calculated, which was then expressed as percent cell confluence. MISSION® siRNA Universal Negative Control #2 (MilliporeSigma) was used as a negative control.

Invasion Migration Assays

For the migration assay, Corning BioCoat Control Inserts with 8.0µm Pore Polyester (PET) Membrane (Cat # 08-774-162) were used. Cells were pre-treated with siRNAs specific for the circRNA of interest or a non-targeting siRNA for 48 hours. Cells were then trypsinized and resuspended to a concentration of 1*10^5 cells/ml for A549 cells and 2*10^5 cells/ml for 22Rv1. in serum-free RPMI-1640. A total of 500 µl cell suspension was added into the upper chamber, whilst the lower chamber was treated

with 750 µl of RPMI-1640 with 10% FBS. Following incubation at 37°C for 20 h, the medium was discarded. Migrated cells were fixed with 100% ethanol for 30 seconds then stained with 0.5% crystal violet at room temperature, washed three times followed by removal of any additional dye with kimwipes and cotton swabs. Representative images were taken under an inverted microscope equipped with a camera.

For the invasion assay, Corning BioCoat Growth Factor Reduced Matrigel Invasion Chamber with 8.0µm Pore Polyester (PET) Membrane (Cat # 08-774-193) were used. Cells were pretreated and trypsinized as discussed above. Prior to adding cell suspension, 500 µl of serum-free RPMI-1640 was added to the transwells and incubated at 37°C for 2 h. The following steps including cell plating, incubation, fixing and staining were conducted as discussed above for the migration assay. Representative images were taken under an inverted microscope equipped with a camera.

Real-time PCR

Total RNA was purified with the RNeasy Mini Kit (QIAGEN, Cat. # 74106) and DNA was removed by performing on-column DNAse treatment (QIAGEN, Cat. # 79254). cDNA was reverse transcribed using the High Capacity cDNA Reverse Kit (4368814, Applied Biosystems).

For RNase R treatment, one µg total RNA was incubated for 15 min at 37 °C with or without 10 U RNase R (Epicenter, Cat. #RNR07250,) followed by phenol-chloroform (ThermoFisher, Cat. #15593-049) purification. cDNA was reverse transcribed using the High-Capacity cDNA Reverse Transcription Kit (Applied Biosystems, Cat. #4368814).

Divergent primers to detect circRNAs were designed for PCR products of sizes ranging from 100-150 bp and spanning the junction site. Convergent primers that targeted the parental linear mRNA counterparts were exclusive of the circRNA generating region for the same size range. Expression levels of circRNAs and their parental linear RNAs were quantified using the primers listed in (Table S4) and PowerUp SYBR Green Master Mix (Applied Biosystems, Cat. #A25742,) in StepOnePlus Real-Time PCR System (Applied Biosystems) or CFX96 Touch Real-Time PCR Detection System (Bio-Rad). RPS28 was used as the endogenous control gene. The results of RT-qPCR were analyzed by the 2-ΔΔCT method.

RNA-seq Mapping and Data processing

RNA was extracted with TRIzol (ThermoFisher, 15596026) and processed with a RNA-seq library preparation kit (Illumina, RS-122-2101) to produce libraries for deep sequencing on Illumina Novaseq X Plus. Library preparation and sequencing were performed according to the manufacturer's protocol. RNA-seq raw reads were first filtered by trim_galore (v 0.5.0) then mapped to the human genome (hg38) by using

STAR (V 2.4.2a) [49] software with default parameters. The hg38 GENCODE gene list was used for all transcription level analysis. RNA-seq reads strands were determined by RSeQC (V 2.6.1)[50,51]. HTSeq (0.11.0)[52] was used to get gene level read counts from STAR mapped bam files. The resultant gene read count table was subjected to DESeq2 (V 1.22.2) [53] for differential gene analysis and a cutoff of 0.01 for FDR was chosen to identify significant differential genes. $\log_2$FoldChange < (-1) and $\log_2$FoldChange > 1 were chosen as up regulated genes and down regulated genes respectively.

Data Analysis & Visualization

Data analysis was conducted using R (v4.1.0) and Python (v3.8.18). Data visualization was conducted using ggplot2 (v3.36) [54], MAGeCKFlute (v1.6.5) [48,55], forestplot (v1.10), VennDiagram (v1.7.3) [56], ComplexUpset (v1.3.3)[57], ComplexHeatmap (v.2.12.0)[58], BoutrosLab.plotting.general (v. 7.1.0)[59], Guitar (v2.16.0)[60], ClusterProfiler (v4.8.3)[61], DESeq2 (v1.40.2)[53], AnnotationDbi (v.1.62.2)[62], DOSE (v.3.26.2)[63], org.Hs.eg.db (v.3.17.0)[64], EnhancedVolcano (v1.18.0)[65], Seaborn (v.0.12.2)[66], Scipy (v1.10.1)[67], Matplotlib (v3.7.3)[68]. Figure 1A was created using BioRender.com. Database was created using RShiny[69].

Data Availability

The DepMap DEMETER2 Data v6 data was downloaded from the DepMap portal (https://depmap.org/portal/download/all/)[16]. CircRNA data from cancer clinical samples were obtained from the MiOncoCirc database (https://mioncocirc.github.io/)[8]. The Raw sequencing data of shRNA screening and RNA-seq generated in this study can be accessed at NCBI under the accessions GSE269567 (reviewer token: gpadomagbhcdfst).

Code Availability

All R packages used are available online as described in the method section. Code used to generate the FunCirc database can be found here: https://github.com/HansenHeLab/FunCirc

Author Contributions

Designed studies: P.H., T.L., F.S., M.J. and H.H.H.

Performed experiments: P.H., F.S., Z.H., X.X., W.Y., M.T.,

Data Analysis: P.H., T.L., M.C., S.C., Y.Z.

Manuscript First Draft: P.H., F.S., Z.H., T.L., M.J., and H.H.H.

Revised & approved manuscript: All authors

Conflicts of Interest

The authors disclose no conflicts.

Acknowledgement

This work was supported by the Princess Margaret Cancer Foundation (886012001223 to H.H.H.), Canadian Cancer Society (TAG2018-2061), CIHR operating grants (142246, 152863, 152864 and 159567 to H.H.H.), Terry Fox New Frontiers Program Project Grant (PPG19-1090 and PPG23-1124 to H.H.H.). H.H.H. holds Tier 1 Canada Research Chair in RNA Medicine. X.X. was supported by the Prostate Cancer Foundation Young Investigator Award (21YOUN06). M.T. was supported by a CIHR Doctoral Award. P.H.H was supported by a MOHCCN Health Informatics Award and a CIHR Doctoral Award. J.M. was supported by the Science Fund of the National Natural Science Foundation of China (No. 82270830) and Hubei Provincial Science & Technology Innovation Team Grant (No. 2022CFB072).

Inclusion & Ethics Statement

We actively took measures to prevent stigmatization, incrimination, or discrimination. Safety measures were implemented to protect participants and researchers, with risk management plans established for potential health, safety, and security risks. We also ensured that local and regional research relevant to this study was appropriately cited, with efforts made to promote diverse perspectives in our reference list.

References

1. Slack, F. J. & Chinnaiyan, A. M. The Role of Non-coding RNAs in Oncology. *Cell* vol. 179 1033–1055 Preprint at https://doi.org/10.1016/j.cell.2019.10.017 (2019).
2. Esposito, R. *et al.* Hacking the Cancer Genome: Profiling Therapeutically Actionable Long Non-coding RNAs Using CRISPR-Cas9 Screening. *Cancer Cell* **35**, 545–557 (2019).
3. Anfossi, S., Babayan, A., Pantel, K. & Calin, G. A. Clinical utility of circulating non-coding RNAs - an update. *Nat. Rev. Clin. Oncol.* **15**, 541–563 (2018).
4. St Laurent, G., Wahlestedt, C. & Kapranov, P. The Landscape of long noncoding RNA classification. *Trends Genet.* **31**, 239–251 (2015).
5. Chen, L.-L. The expanding regulatory mechanisms and cellular functions of circular RNAs. *Nat. Rev. Mol. Cell Biol.* **21**, 475–490 (2020).
6. Liu, J., Zhang, X., Yan, M. & Li, H. Emerging Role of Circular RNAs in Cancer. *Frontiers in Oncology* vol. 10 Preprint at https://doi.org/10.3389/fonc.2020.00663 (2020).
7. Rajappa, A., Banerjee, S., Sharma, V. & Khandelia, P. Circular RNAs: Emerging Role in Cancer Diagnostics and Therapeutics. *Front Mol Biosci* **7**, 577938 (2020).
8. Vo, J. N. *et al.* The Landscape of Circular RNA in Cancer. *Cell* **176**, 869–881.e13 (2019).
9. Chen, S. *et al.* Widespread and Functional RNA Circularization in Localized Prostate Cancer. *Cell* **176**, 831–843.e22 (2019).
10. Smid, M. *et al.* The circular RNome of primary breast cancer. *Genome Res.* **29**, 356–366 (2019).
11. Zhang, Y. *et al.* Optimized RNA-targeting CRISPR/Cas13d technology outperforms shRNA in identifying functional circRNAs. *Genome Biol.* **22**, 41 (2021).

12. Li, S. *et al.* Screening for functional circular RNAs using the CRISPR-Cas13 system. *Nat. Methods* **18**, 51–59 (2021).
13. He, A. T., Liu, J., Li, F. & Yang, B. B. Targeting circular RNAs as a therapeutic approach: current strategies and challenges. *Signal Transduct Target Ther* **6**, 185 (2021).
14. Chen, S., Zhang, J. & Zhao, F. Screening Linear and Circular RNA Transcripts from Stress Granules. *Genomics Proteomics Bioinformatics* (2022) doi:10.1016/j.gpb.2022.01.003.
15. Hart, T. *et al.* High-Resolution CRISPR Screens Reveal Fitness Genes and Genotype-Specific Cancer Liabilities. *Cell* **163**, 1515–1526 (2015).
16. Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564 (2017).
17. Nagpal, N. *et al.* Essential role of MED1 in the transcriptional regulation of ER-dependent oncogenic miRNAs in breast cancer. *Sci. Rep.* **8**, (2018).
18. Seachrist, D. D., Anstine, L. J. & Keri, R. A. FOXA1: A Pioneer of Nuclear Receptor Action in Breast Cancer. *Cancers* **13**, (2021).
19. Matevossian, A. & Resh, M. D. Hedgehog Acyltransferase as a target in estrogen receptor positive, HER2 amplified, and tamoxifen resistant breast cancer cells. *Mol. Cancer* **14**, 1–15 (2015).
20. Brett, J. O., Spring, L. M., Bardia, A. & Wander, S. A. ESR1 mutation as an emerging clinical biomarker in metastatic hormone receptor-positive breast cancer. *Breast Cancer Res.* **23**, 1–15 (2021).
21. del Mar Maldonado, M. & Dharmawardhane, S. Targeting Rac and Cdc42 GTPases in Cancer. *Cancer Res.* **78**, 3101 (2018).
22. Xiao, X.-H. *et al.* Regulating Cdc42 and Its Signaling Pathways in Cancer: Small Molecules and MicroRNA as New Treatment Candidates. *Molecules : A Journal of Synthetic*

Chemistry and Natural Product Chemistry **23**, (2018).

23. Lee, Y. G. *et al.* LONP1 and ClpP cooperatively regulate mitochondrial proteostasis for cancer cell survival. *Oncogenesis* **10**, 1–14 (2021).
24. Bedi, U. *et al.* SUPT6H controls estrogen receptor activity and cellular differentiation by multiple epigenomic mechanisms. *Oncogene* **34**, (2015).
25. Su, W. *et al.* Overexpressed WDR3 induces the activation of Hippo pathway by interacting with GATA4 in pancreatic cancer. *J. Exp. Clin. Cancer Res.* **40**, 1–15 (2021).
26. Wang, Y. *et al.* RPS24 knockdown inhibits colorectal cancer cell migration and proliferation in vitro. *Gene* **571**, (2015).
27. Chen, H.-H. *et al.* Circular RNA detection identifies circPSEN1 alterations in brain specific to autosomal dominant Alzheimer's disease. *Acta Neuropathol Commun* **10**, 29 (2022).
28. Yao, S. *et al.* CircRNA ARFGEF1 functions as a ceRNA to promote oncogenic KSHV-encoded viral interferon regulatory factor induction of cell invasion and angiogenesis by upregulating glutaredoxin 3. *PLoS Pathog.* **17**, e1009294 (2021).
29. Yu, D. & Zhang, C. Circular RNA PTK2 Accelerates Cell Proliferation and Inhibits Cell Apoptosis in Gastric Carcinoma via miR-139-3p. *Dig. Dis. Sci.* **66**, 1499–1509 (2021).
30. Zhou, F. *et al.* Circular RNA Protein Tyrosine Kinase 2 Promotes Cell Proliferation, Migration and Suppresses Apoptosis via Activating MicroRNA-638 Mediated MEK/ERK, WNT/β-Catenin Signaling Pathways in Multiple Myeloma. *Front. Oncol.* **11**, 648189 (2021).
31. Ma, L. *et al.* Silencing of circRACGAP1 sensitizes gastric cancer cells to apatinib via modulating autophagy by targeting miR-3657 and ATG7. *Cell Death Dis.* **11**, 169 (2020).
32. Zhou, Z. *et al.* Circular RNAs act as regulators of autophagy in cancer. *Mol Ther Oncolytics* **21**, 242–254 (2021).

33. Yang, Q. *et al.* Circular RNA expression profiles during the differentiation of mouse neural stem cells. *BMC Syst. Biol.* **12**, 128 (2018).

34. Jiang, Z. *et al.* Circular RNA protein tyrosine kinase 2 (circPTK2) promotes colorectal cancer proliferation, migration, invasion and chemoresistance. *Bioengineered* **13**, 810–823 (2022).

35. Liu, L. *et al.* Systematic loss-of-function screens identify pathway-specific functional circular RNAs. *Nat. Cell Biol.* **26**, 1359–1372 (2024).

36. Liu, S. J. *et al.* CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* **355**, (2017).

37. Liu, X. *et al.* Circular RNA: An emerging frontier in RNA therapeutic targets, RNA therapeutics, and mRNA vaccines. *J. Control. Release* **348**, 84–94 (2022).

38. Misir, S., Wu, N. & Yang, B. B. Specific expression and functions of circular RNAs. *Cell Death Differ.* **29**, 481–491 (2022).

39. Cheng, Q. *et al.* Elevated MPP6 expression correlates with an unfavorable prognosis, angiogenesis and immune evasion in hepatocellular carcinoma. *Front. Immunol.* **14**, 1173848 (2023).

40. Xu, F. *et al.* Downregulating SynCAM and MPP6 expression is associated with ovarian cancer progression. *Oncol. Lett.* **18**, 2477–2483 (2019).

41. Dong, W. *et al.* Circular RNA ACVR2A suppresses bladder cancer cells proliferation and metastasis through miR-626/EYA4 axis. *Mol. Cancer* **18**, 95 (2019).

42. Ai, Y., Liang, D. & Wilusz, J. E. CRISPR/Cas13 effectors have differing extents of off-target effects that limit their utility in eukaryotic cells. *Nucleic Acids Res.* **50**, e65 (2022).

43. Bayoumi, M. & Munir, M. Potential Use of CRISPR/Cas13 Machinery in Understanding

Virus-Host Interaction. *Front. Microbiol.* **12**, 743580 (2021).

44. Glažar, P., Papavasileiou, P. & Rajewsky, N. circBase: a database for circular RNAs. *RNA* **20**, 1666–1670 (2014).
45. Zhang, X.-O. *et al.* Diverse alternative back-splicing and alternative splicing landscape of circular RNAs. *Genome Res.* **26**, 1277–1287 (2016).
46. Chen, X. *et al.* circRNADb: A comprehensive database for human circular RNAs with protein-coding annotations. *Sci. Rep.* **6**, 34985 (2016).
47. Xia, S. *et al.* CSCD: a database for cancer-specific circular RNAs. *Nucleic Acids Res.* **46**, D925–D929 (2018).
48. Li, W. *et al.* MAGeCK enables robust identification of essential genes from genome-scale CRISPR/Cas9 knockout screens. *Genome Biol.* **15**, 554 (2014).
49. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
50. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184–2185 (2012).
51. Wang, L. *et al.* Measure transcript integrity using RNA-seq data. *BMC Bioinformatics* **17**, 58 (2016).
52. Anders, S., Pyl, P. T. & Huber, W. HTSeq--a Python framework to work with high-throughput sequencing data. *Bioinformatics* **31**, 166–169 (2015).
53. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014).
54. Wickham, H. *ggplot2: Elegant Graphics for Data Analysis*. (Springer, 2016).
55. Wang, B. *et al.* Integrative analysis of pooled CRISPR genetic screens using MAGeCKFlute. *Nat. Protoc.* **14**, 756–780 (2019).

56. Chen, H. & Boutros, P. C. VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**, 35 (2011).
57. Lex, A., Gehlenborg, N., Strobelt, H., Vuillemot, R. & Pfister, H. UpSet: Visualization of Intersecting Sets. *IEEE Trans. Vis. Comput. Graph.* **20**, 1983–1992 (2014).
58. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849 (2016).
59. P'ng, C. *et al.* BPG: Seamless, automated and interactive visualization of scientific data. *BMC Bioinformatics* **20**, 42 (2019).
60. Cui, X. *et al.* Guitar: An R/Bioconductor Package for Gene Annotation Guided Transcriptomic Analysis of RNA-Related Genomic Features. *Biomed Res. Int.* **2016**, 8367534 (2016).
61. Wu, T. *et al.* clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* **2**, 100141 (2021).
62. AnnotationDbi. *Bioconductor* https://bioconductor.org/packages/AnnotationDbi.
63. Yu, G., Wang, L.-G., Yan, G.-R. & He, Q.-Y. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* **31**, 608–609 (2015).
64. Carlson, M. *org.Hs.eg.db*. (Bioconductor, 2017). doi:10.18129/B9.BIOC.ORG.HS.EG.DB.
65. Blighe, K. *EnhancedVolcano*. (Bioconductor, 2018). doi:10.18129/B9.BIOC.ENHANCEDVOLCANO.
66. Waskom, M. seaborn: statistical data visualization. *J. Open Source Softw.* **6**, 3021 (2021).
67. Virtanen, P. *et al.* SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
68. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **9**, 90–95

(May-June 2007).

69. https://cran.r-project.org/web/packages/shiny/index.html.

Figures and Figure Legends

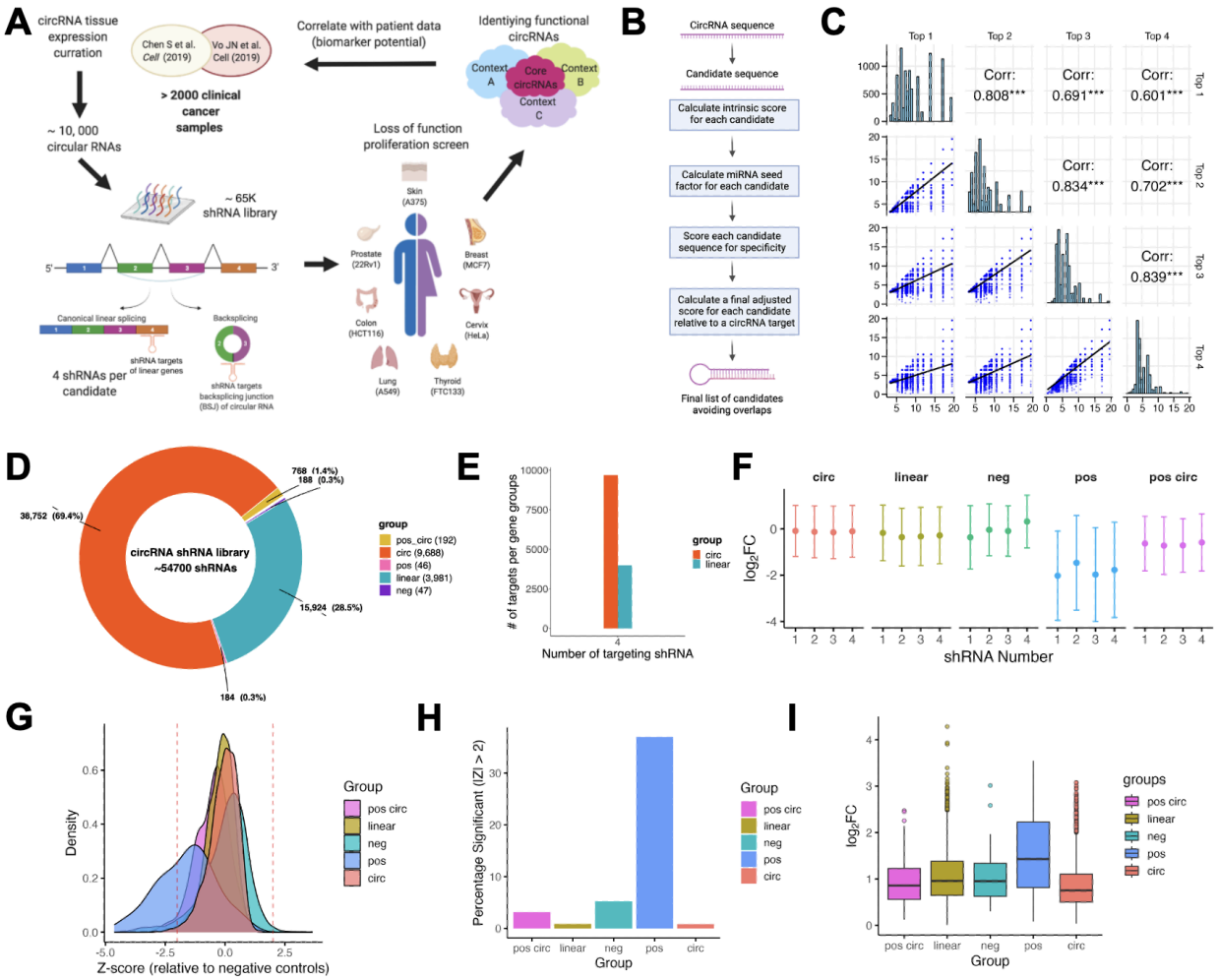


Figure 1: Design and quality control of shRNA screening library targeting cancer associated circRNAs and parental linear transcripts.

A) Schematic of pan-cancer shRNA screening. Approximately 10,000 clinically associated circRNAs were filtered from public and internal clinical cohorts, and shRNAs were designed to specifically target the backsplice site of each circRNAs as well as exons of parental linear transcripts. Functional screens were performed in seven major cancer types, including Prostate, Breast, Colon, Lung, Tyroid, Cervix and Skin cancers.
B) Schematic of the multi-step shRNA design pipeline for circRNAs
C) Pairwise correlation of final adjusted scores among the top four shRNA designs for each circRNA. Scatter plots in the lower off-diagonal cells show the relationship between Top 1–Top 4 ranks, with Pearson's correlation coefficients (r) and significance

(**p < 0.0001) reported in the upper off-diagonal cells. Histograms along the diagonal depict the score distributions for each rank.
D) Overview of categories of genes included in the circRNA shRNA screening library.
E) Summary of number of shRNAs per linear or circRNA target in the screening library.
F) For each group (circ, linear, negative control, positive control, and positive circ control), the mean $\log_2$ fold change (± SD) is shown for each of the four shRNA positions targeting the same gene or circRNA. Circles represent the average knockdown across all targets in that group for a given position, with vertical bars indicating the standard deviation.
G) Each point represents the average $\log_2$ fold change (LFC) for a given target plotted against its standardized Z-score, where Z = (LFC − mean neg)/SD neg. The dashed horizontal lines mark Z = ±2, commonly used as a significance cutoff.
H) Bar chart showing the fraction of shRNAs in each group with |Z| > 2, indicating significant depletion or enrichment relative to negative controls. Positive controls (green) have the highest fraction of "hits" (~30%).
I) Boxplot comparing within-target variability (SD of the $\log_2$ fold change) for each of the four shRNAs targeting a given gene or circRNA. Dots represent individual genes/circRNAs, boxes denote the interquartile range with the median line, and whiskers extend up to 1.5 × IQR.

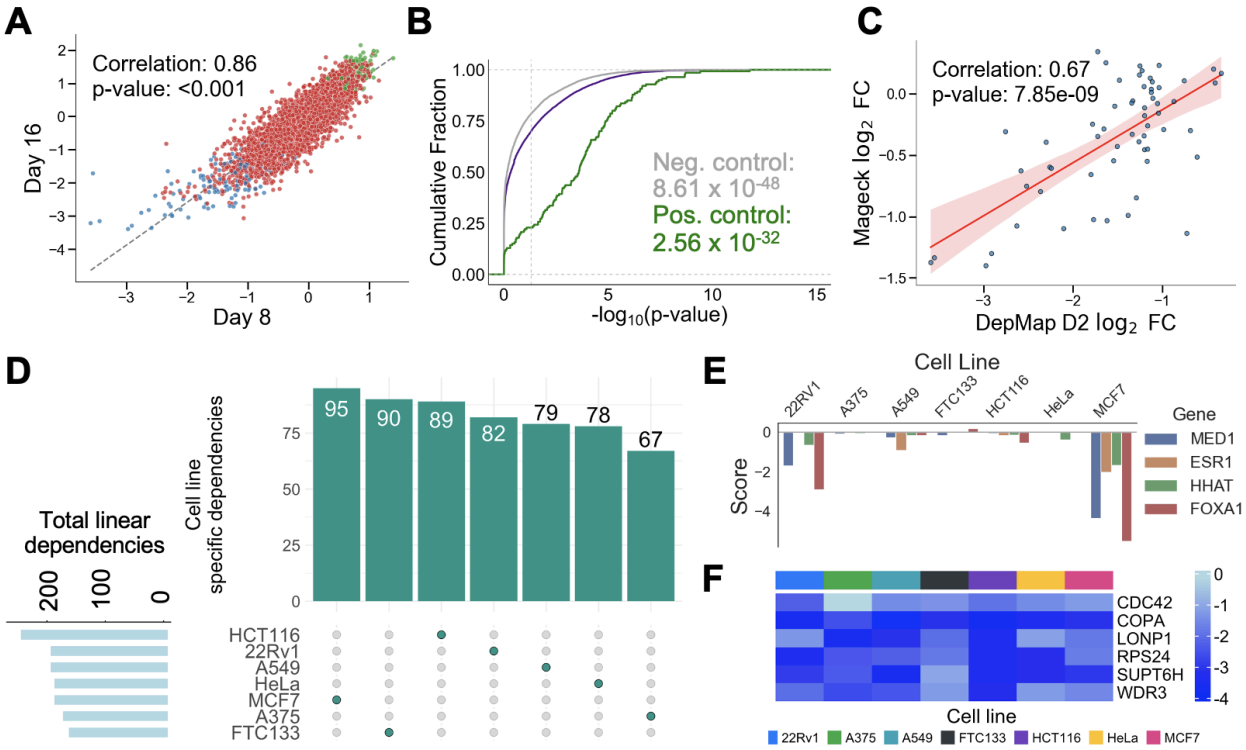


Figure 2: shRNA screens identify common and cancer type specific dependencies of linear transcripts.

A) Dropout score measured as log_2-transformed fold change (log_2(FC)) in the 22Rv1 cell line of linear transcripts (red), positive controls (green), and negative controls (blue). x and y axes represent comparisons for T8 and T16 samples relative to T0 separately.
B) Cumulative distribution of the negative log_{10}(p value) for all linear transcripts. Linear transcripts, and positive and negative controls comparing T8 versus T0 samples in the 22Rv1 cell line. p values (Mann–Whitney U test) are relative to the all linear group (purple line).
C) Comparison of dropout score in 22Rv1 cells for the union of essential linear transcripts identified from the screens and the DepMap shRNA screening project.
D) UpSet plot of cell-line specific and common linear dependencies. Linear transcripts bar graph on the left corner represents the total number of linear dependencies present for each cell line. Columns represent each cell line, with the number of cell-line specific dependencies listed.
E) Comparison of dropout score of previously identified linear transcripts with known functional relevance in MCF7.
F) Heatmap showing dropout score of common essential linear transcripts identified in our shRNA screens.

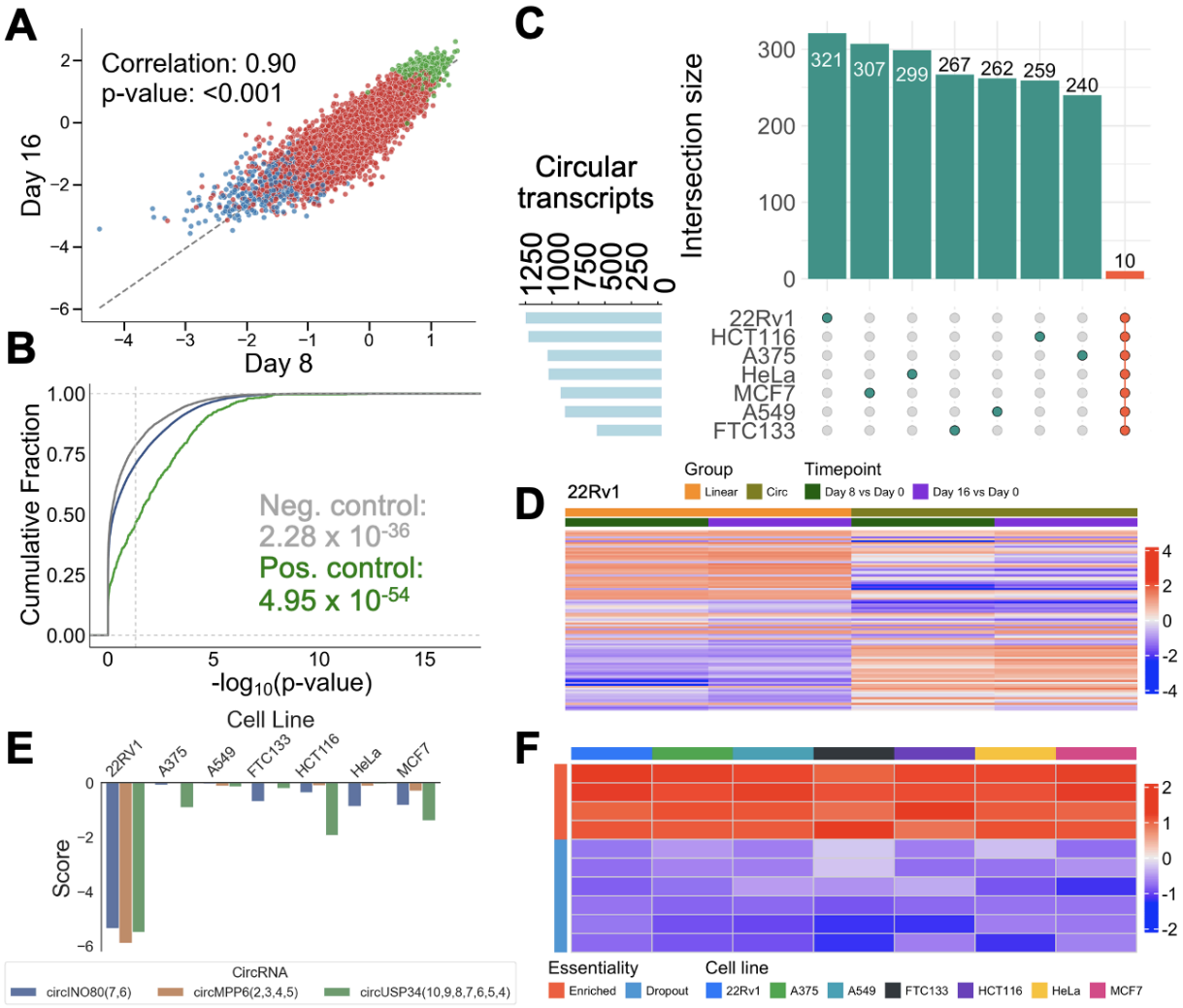


Figure 3: shRNA screens identify common and cancer type specific dependencies of circRNA transcripts.

A) Dropout score measured as $\log_2$(FC) in the 22Rv1 cell line of circular transcripts (red), positive controls (green), and negative controls (blue). x and y axes represent comparisons for T8 and T16 samples relative to T0 separately.

B) Cumulative distribution of the negative $\log_{10}$(p value) for all circular transcripts. Circular transcripts, and positive and negative controls comparing T8 versus T0 samples in the 22Rv1 cell line. p values (Mann–Whitney U test) are relative to the all circular group (blue line).

C) UpSet plot of cell-line specific and common circular dependencies. Circular transcripts bar graph on the left corner represents the total number of circular dependencies present for each cell line. First seven columns (teal) represent each cell line, with the number of functional circular transcripts specific to each cell line listed. The final column (orange) represents the functional circular transcripts shared across all the

seven cell lines.
D) Depletion of linear and circRNAs following shRNA screening in 22Rv1 cell line. Heatmap shows essential circRNAs along with their linear parental gene. Color represents log_2(FC) of T8 and T16 compared to T0, with red and blue shades representing negative (depleted) and positive (enriched) values separately.
E) Comparison of dropout score of representative functional unknown circRNAs specific to 22Rv1.
F) Heatmap showing dropout score of common dropout and enriched circular transcripts identified in our shRNA screens.

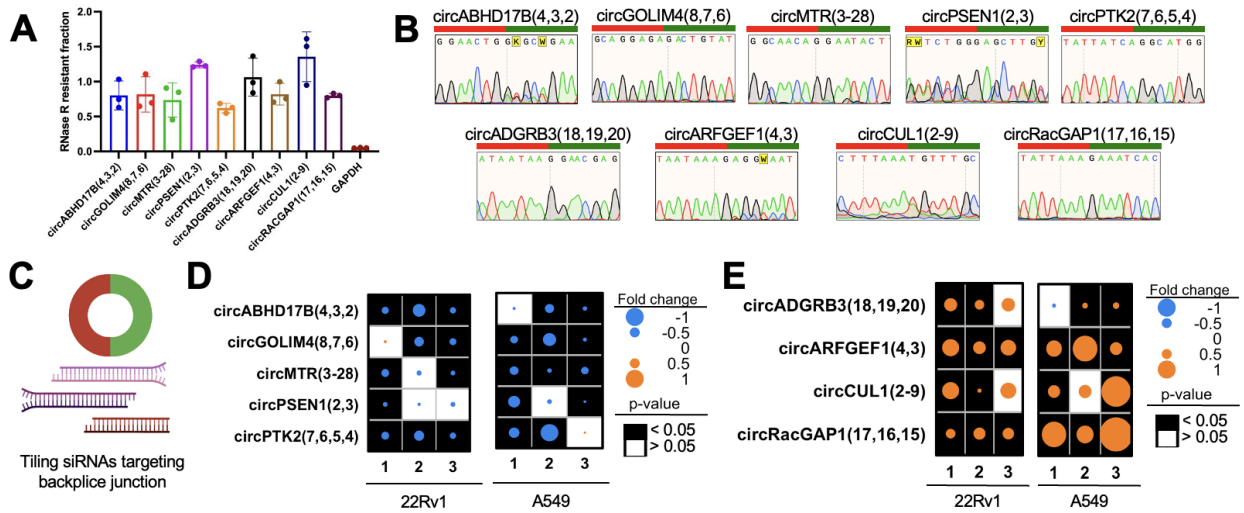


Figure 4 :Functional validation of core dropout and enriched circRNAs.
A) Core common circRNAs detected by the divergent primers are resistant to RNase R. Data are normalized to their corresponding values in the untreated samples. GAPDH is used as a linear transcript control to demonstrate degradation following RNase R treatment.
B) Sanger sequencing of the PCR product by the divergent primers targeting the junction sites of the core common circRNAs to confirm back-splice sequences. The junction between the orange region and the blue region indicates the 'head-to-tail' splicing site.
C) Schematic of tiling siRNA design targeting the backspliced junction of circRNAs.
D-E) Illustration of proliferation assays following circRNA KD in 22Rv1 and A549 cells for the five dropout (D) or four enriched (E) circRNAs. Background shading indicates p value; size and color of dot shows log_2(FC) of treatment over control. The numbers indicate different siRNA.

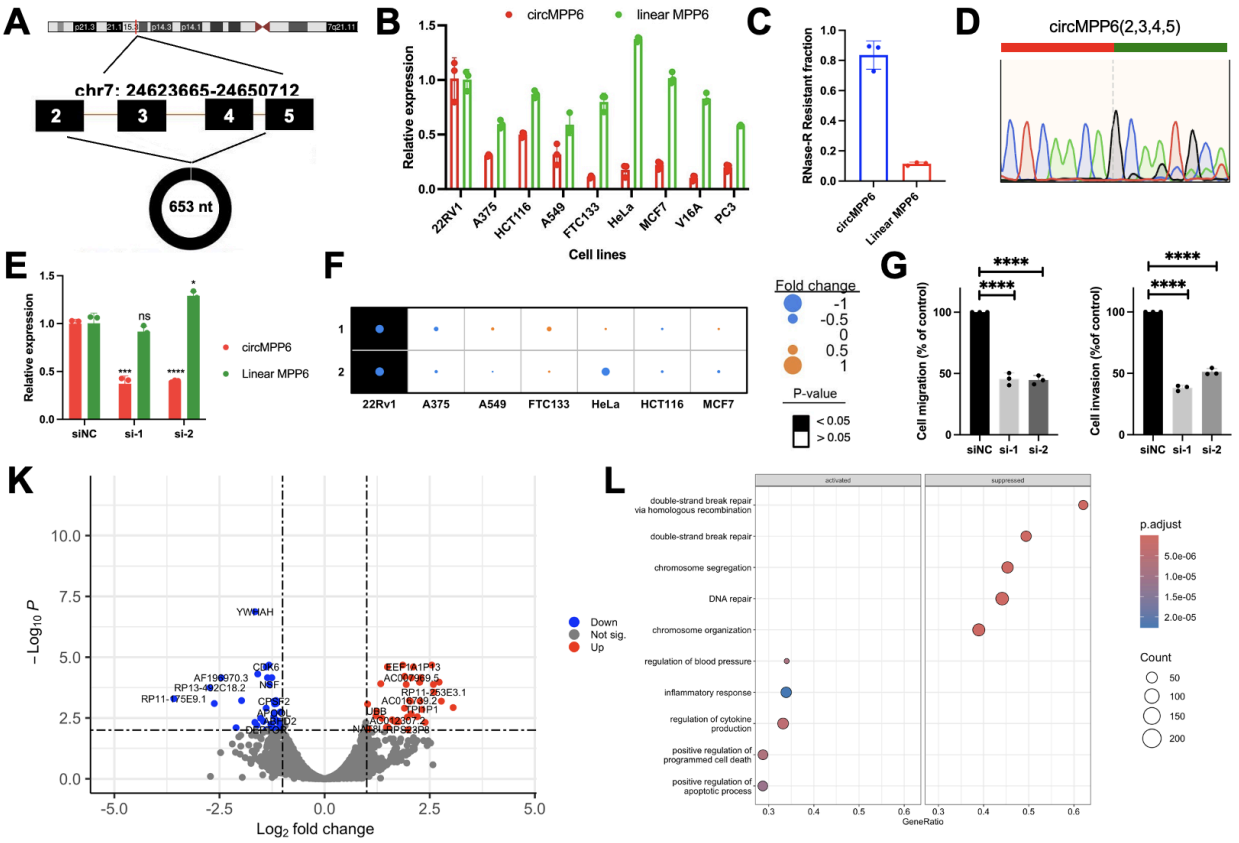


Figure 5: Functional validation of context specific dependence circRNA in prostate cancer.

A) Schematic diagram illustrating the formation of circMPP6(2,3,4,5).

B) RT-qPCR for circMPP6(2,3,4,5) and its linear parental gene across all cell lines screened. Data are normalized to the 22Rv1 cell line.

C) circMPP6(2,3,4,5) detected by the divergent primers is resistant to RNase R. Data are normalized to their corresponding values in the untreated samples.

D) Sanger sequencing of the PCR product by the divergent primers targeting the junction sites of circMPP6(2,3,4,5) to confirm back-splice sequences. The junction between the orange region and the blue region indicates the 'head-to-tail' splicing site.

E) siRNA KD efficiency of circMPP6(2,3,4,5) as determined by RT-qPCR. siNC: Non-targeting siRNA control, si-1, si-2: two tiled siRNAs targeting circMPP6(2,3,4,5). ns, no significance; *, $P < 0.05$; ***, $P < 0.001$; ****, $P < 0.0001$.

F) Summary of proliferation assays following circRNA KD across all cell lines for circMPP6(2,3,4,5).

G) Migration-invasion assay in 22Rv1 cells following KD of circMPP6(2,3,4,5). NT si: Non-targeting siRNA control, si-1 and si-2: two tiled siRNAs targeting circMPP6(2,3,4,5). ****, $P < 0.0001$.

(H) Volcano plot showing RNA-seq data in 22Rv1 cells with and without KD of circMPP6(2,3,4,5). The $\log_2$(FC) indicates the mean differential expression level for

each gene contrasting KD and control samples. Grey dots represent non significant (Not sig.) differentially expressed genes between circMPP6(2,3,4,5) KD and control groups, the blue dots represent down-regulated (Down) genes and red dots represent up-regulated genes (Up).
(I) GO classifications of the differentially expressed genes ($log_2(FC)>1$, $p<0.05$) between circMPP6(2,3,4,5) KD and control groups.

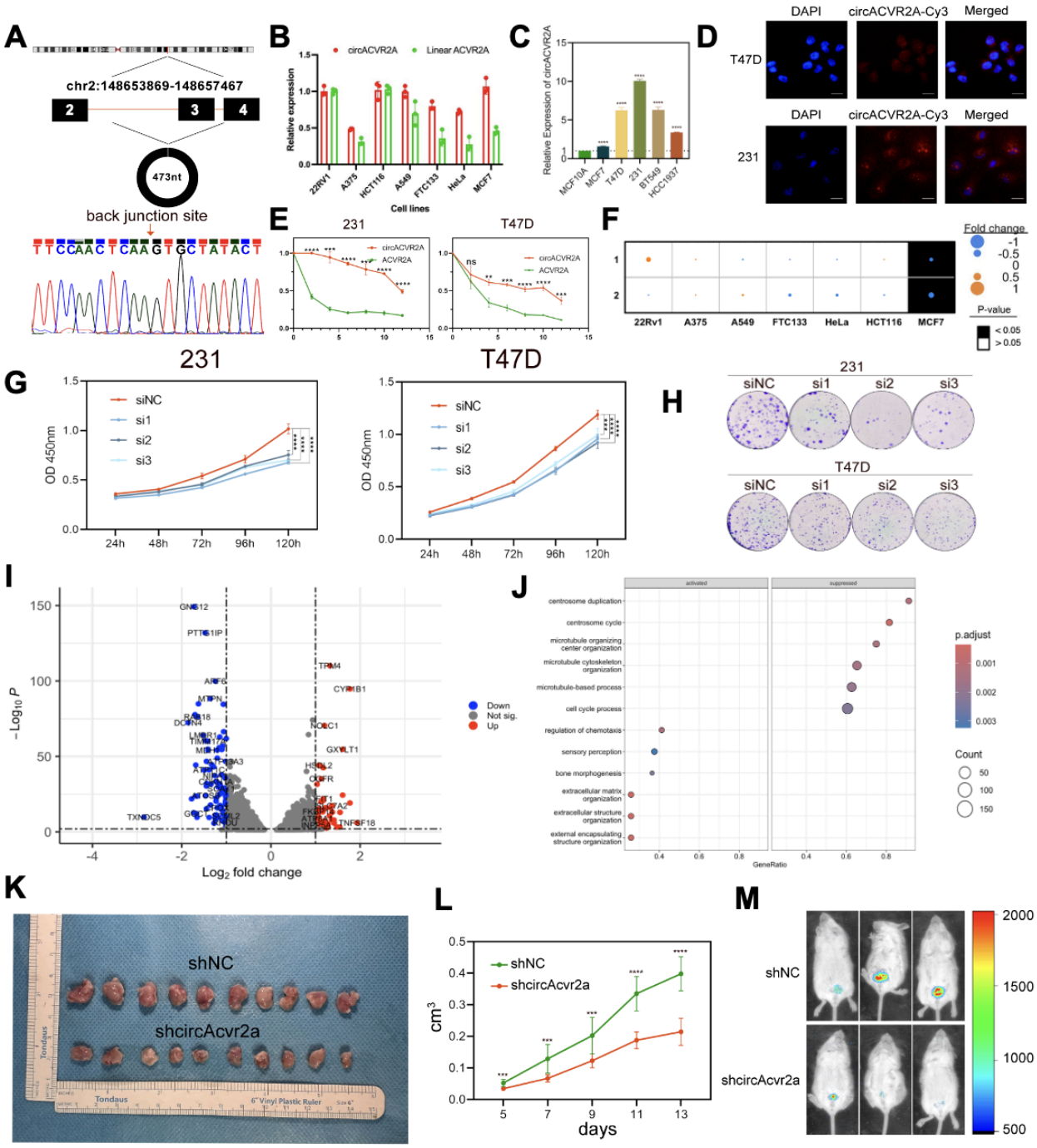


Figure 6: Functional validation of context specific dependence circRNA in breast cancer.

A) Schematic diagram illustrating the formation and sanger sequencing of circACVR2A(2,3,4).

B) RT-qPCR for circACVR2A(2,3,4) and its linear parental gene across all cell lines screened. Data are normalized to the MCF7 cell line.

C) The expression level of circACVR2A(2,3,4) in breast epithelial and cancer cell lines.
D) Representative RNA FISH assay showed the cellular localization Scale bar, 10 μm.
E) Relative expression of linear and circular ACVR2A detected by RT-qPCR treated with Actinomycin.
F) Summary of proliferation assays following circACVR2A(2,3,4) KD across all seven screened cell lines.
(G) The proliferative abilities of transfected T47D and 231 measured by CCK8 and (H) colony formation assays.
(I) Volcano plot showing RNA-seq data in T47D cells with and without KD of circACVR2A(2,3,4). The log2(FC) indicates the mean differential expression level for each gene contrasting circACVR2A(2,3,4) KD and control cells. Each dot represents one gene. After siRNA KD, grey dots represent non significant (Not sig.) differentially expressed genes between circACVR2A(2,3,4) KD and control groups, the blue dots represent down-regulated (Down) genes and red dots represent up-regulated genes (Up).
(J) GO classifications of the differentially expressed genes (log2(FC)>1, p<0.05) between circACVR2A(2,3,4) KD and control groups.
K) shNC/shcircACVR2A infected 4T1 were subcutaneously injected in the BALB/c mouse to establish a xenograft model.
L) Tumor growth curve for circACVR2A(2,3,4) knockdown and control groups.
M) Live imaging of the subcutaneous tumors.
ns, no significance; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$.

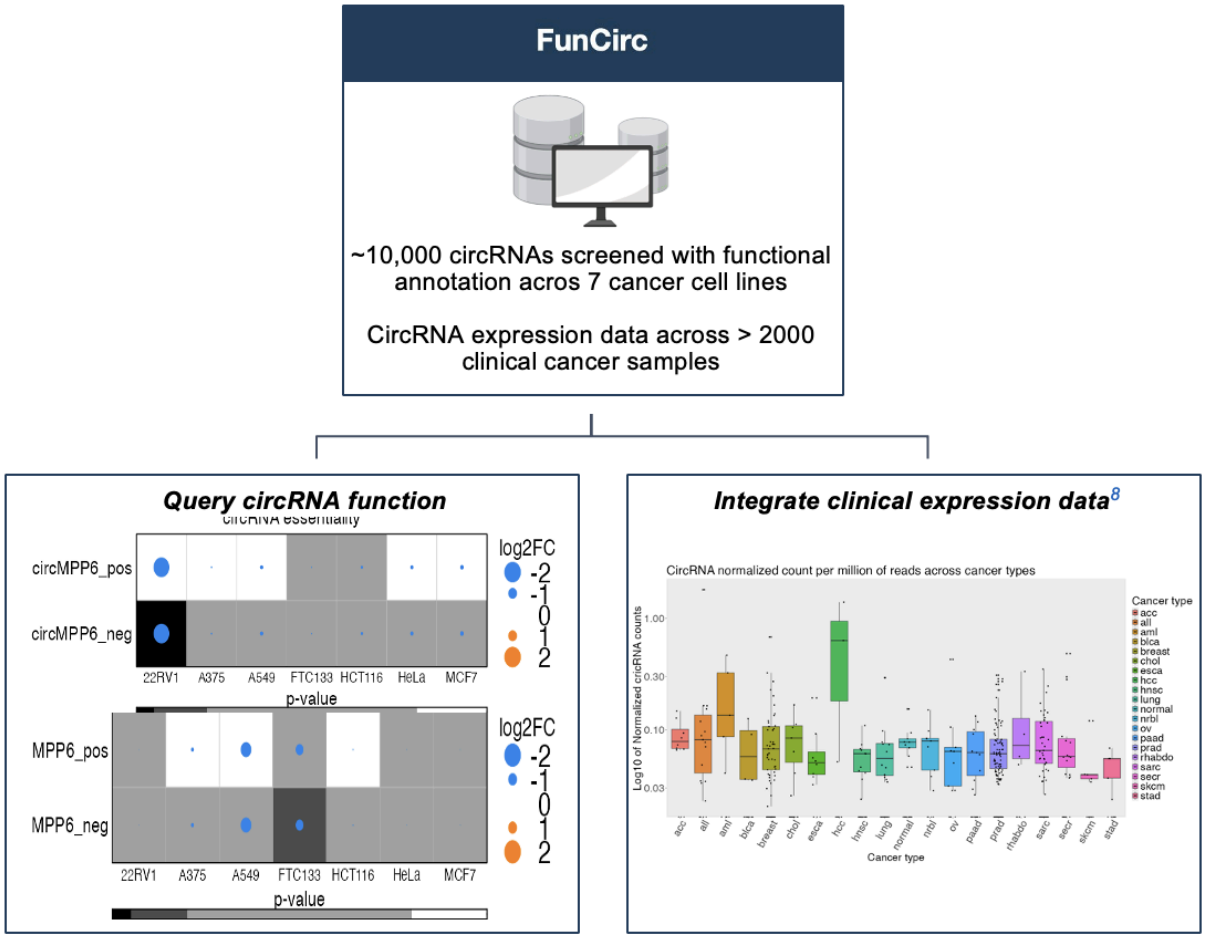
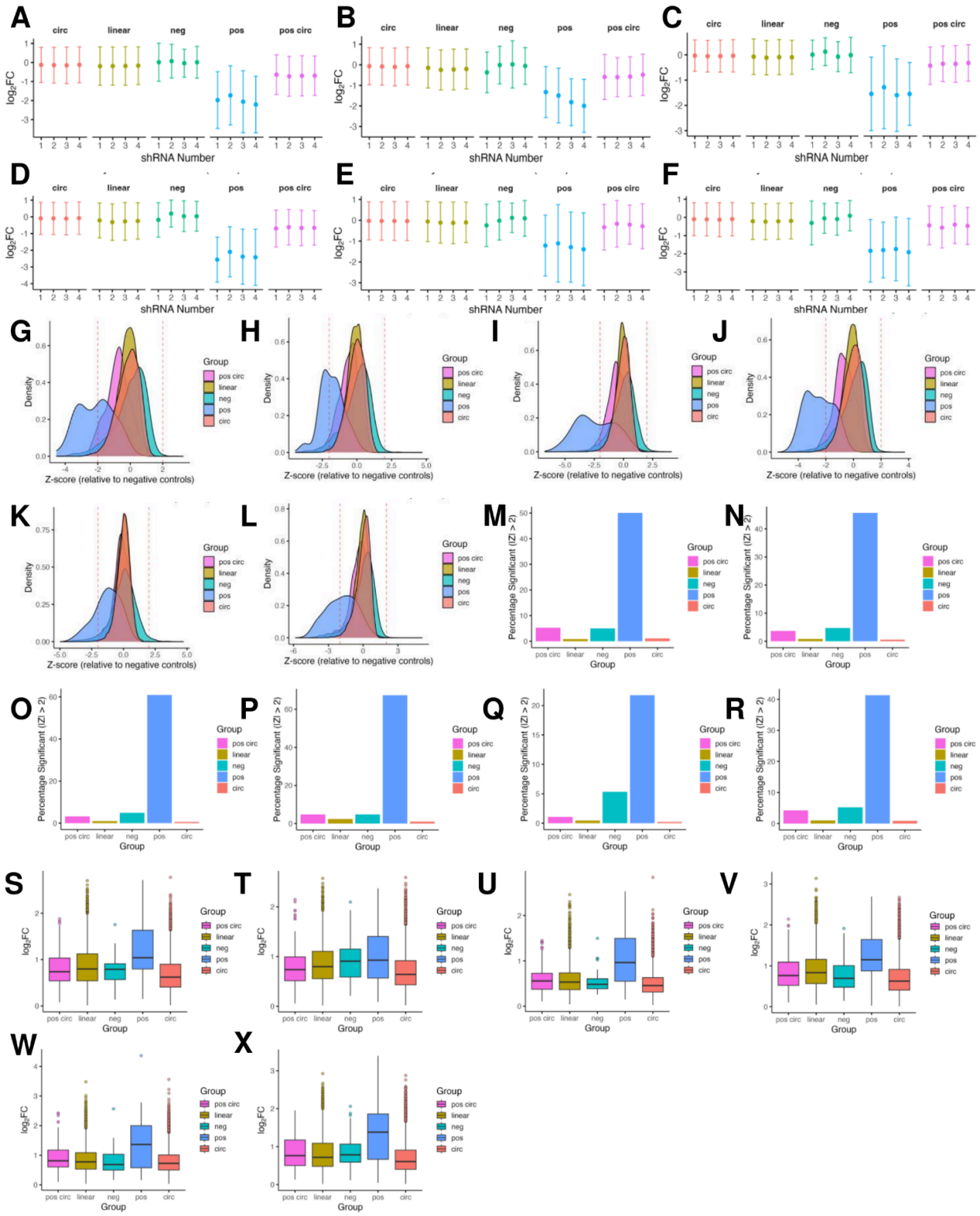


Figure 7: Overview of the FunCirc portal.

A schematic overview of the construction of FunCirc (upper panel) and the essentiality and expression outputs (bottom panels).

Supplementary Material

Supplementary Figure S1

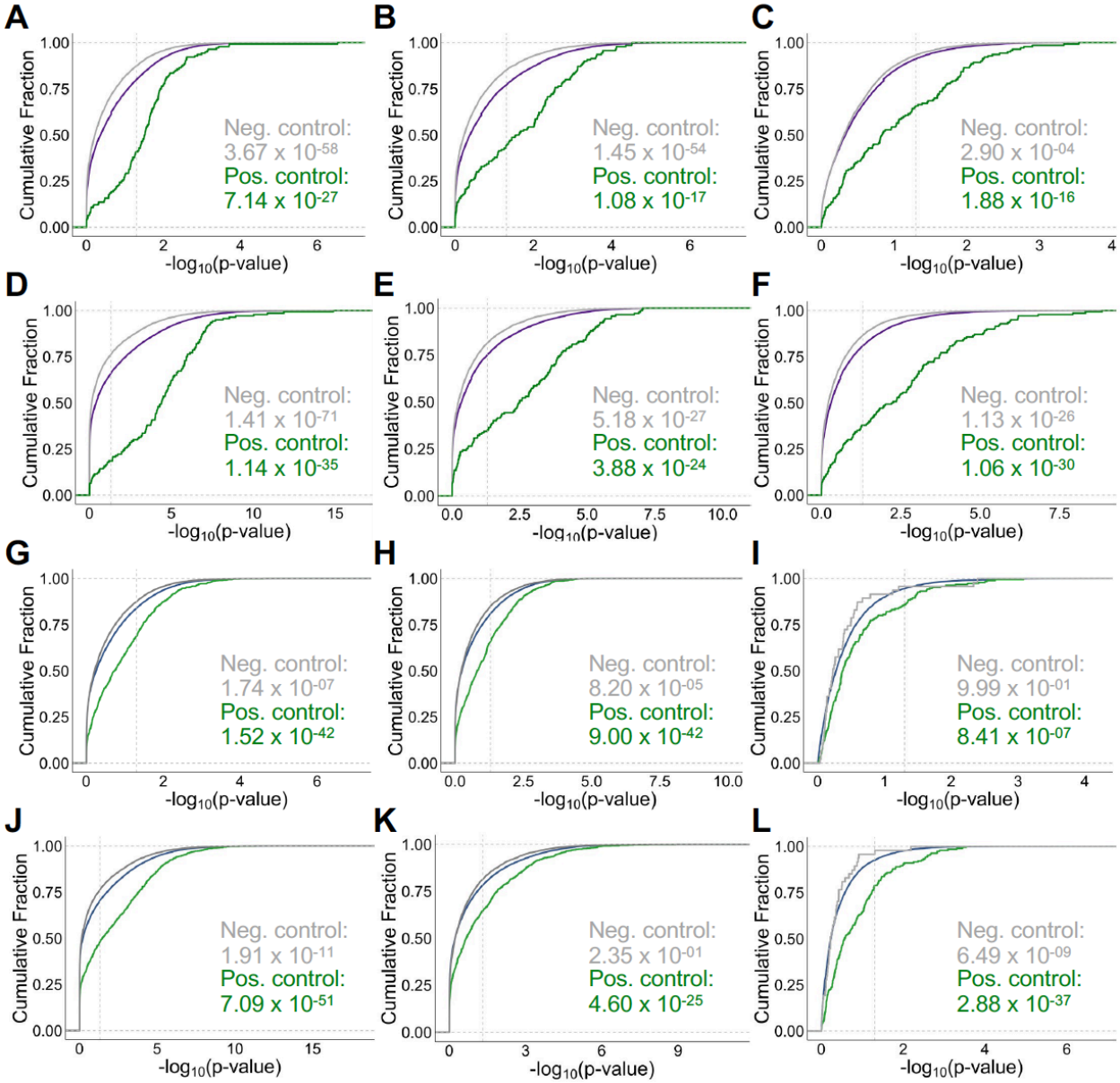


Supplementary Figure 1: Design and quality control of shRNA screening library

targeting cancer associated circRNAs and parental linear transcripts across all cell lines.

A-F) (A) 22Rv1, (B) A375, (C) MCF7, (D) HCT116, (E) FTC133, (F) HeLa. For each group (circ, linear, negative control, positive control, and positive circ control), the mean $\log_2$ fold change (± SD) is shown for each of the four shRNA positions targeting the same gene or circRNA. Circles represent the average knockdown across all targets in that group for a given position, with vertical bars indicating the standard deviation.
G-L) (G) 22Rv1, (H) A375, (I) MCF7, (J) HCT116, (K) FTC133, (L) HeLa. Each point represents the average $\log_2$ fold change (LFC) for a given target plotted against its standardized Z-score, where Z = (LFC − mean neg)/SD neg. The dashed horizontal lines mark Z = ±2, commonly used as a significance cutoff.
M-R) (M) 22Rv1, (N) A375, (O) MCF7, (P) HCT116, (Q) FTC133, (R) HeLa. Bar chart showing the fraction of shRNAs in each group with |Z| > 2, indicating significant depletion or enrichment relative to negative controls. Positive controls (green) have the highest fraction of “hits” (~30%).
S-X) (S) 22Rv1 (Kruskal-Wallis chi-squared = 563.63, p-value < 0.001), (T) A375 (Kruskal-Wallis chi-squared = 440.48, p-value <0.001), (U) MCF7 (Kruskal-Wallis chi-squared = 277.96, p-value <0.001), (V) HCT116 (Kruskal-Wallis chi-squared = 680.35, p-value <0.001), (W) FTC133 (Kruskal-Wallis chi-squared = 73.24, p-value <0.001), (X) HeLa (Kruskal-Wallis chi-squared = 258.94, p-value <0.001). Boxplot comparing within-target variability (SD of the $\log_2$ fold change) for each of the four shRNAs targeting a given gene or circRNA. Dots represent individual genes/circRNAs, boxes denote the interquartile range with the median line, and whiskers extend up to 1.5 × IQR.

Supplementary Figure S2



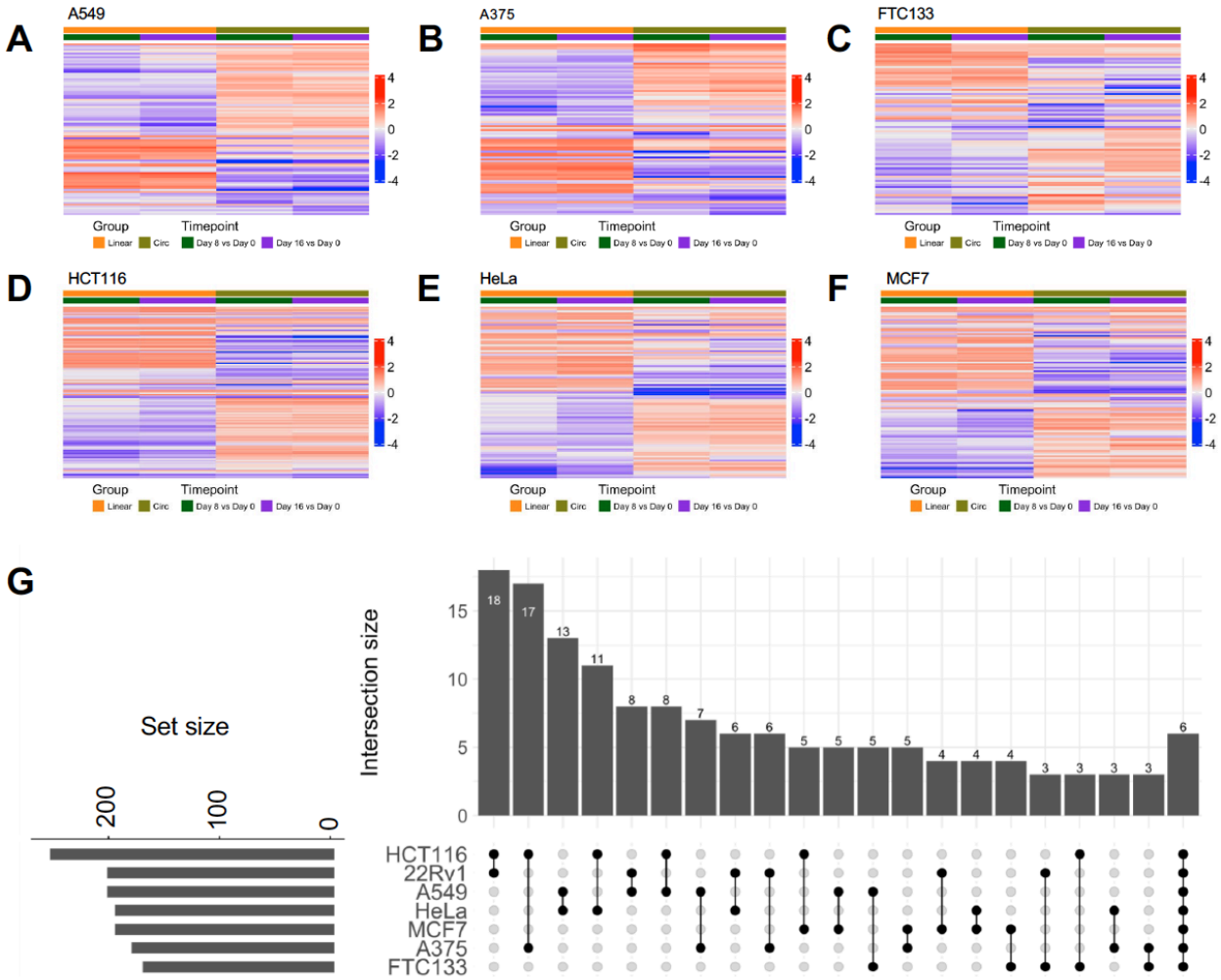
Supplementary Figure 2: Quality measurement of shRNA screens of circular and linear transcripts.

A-F) Cumulative distribution of the negative $\log_{10}$ transformed p-value for all linear transcripts, positive and negative controls comparing Day 8 versus Day 0 samples in A375, A549, FTC133, HCT116, HeLa, MCF7 cell lines. One-sided Mann-Whitney U test was used to compare groups to the linear group (purple line). CLES between the linear and positive control group are 0.24, 0.29, 0.30, 0.20, 0.25 and 0.22 respectively. CLES between the linear and negative control group are 0.56, 0.56, 0.51, 0.56, 0.54 and 0.54 respectively.

G-L) Cumulative distribution of the negative log10 transformed P value for all circular

transcripts, positive and negative controls comparing Day 8 versus Day 0 samples in A375, A549, FTC133, HCT116, HeLa, MCF7 cell lines. One-sided Mann-Whitney U test was used to compare groups to the linear group (purple line). CLES between the linear and positive control group are 0.36, 0.36, 0.44, 0.34, 0.39 and 0.37 respectively. CLES between the linear and negative control group are 0.52, 0.51, 0.49, 0.52, 0.50 and 0.52 respectively.

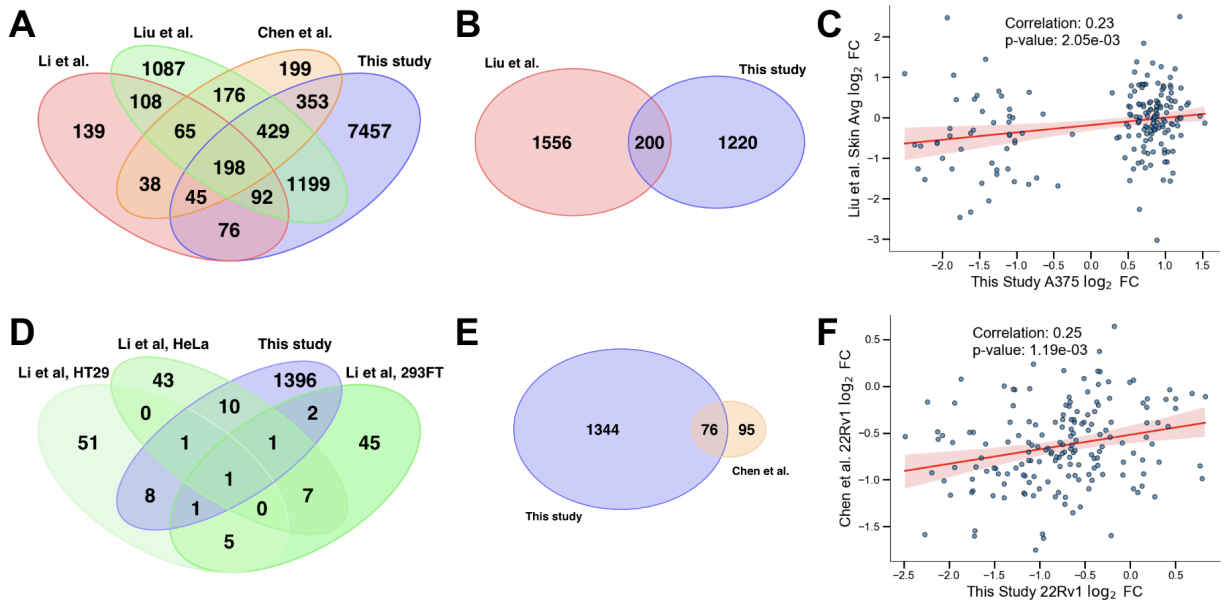
Supplementary Figure S3



Supplementary Figure 3: circular and linear dependencies identified in the shRNA screens.
A-F) Depletion of linear and circular RNAs following shRNA screening in A549 (A), A375 (B), FTC133 (C), HCT116 (D), HeLa (E) and MCF7 (F) cell lines. Heatmap shows essential circRNAs along with their linear parental gene. Color represents $\log_2$(FC) of T8 and T16 compared to T0, with red and blue shades representing negative (depleted) and positive hits.

G) UpSet plot of linear dependencies common for two sets of cell lines along with the set common across cell lines. Linear transcripts bar graph on the left corner represents the total number of linear dependencies present for each cell line. Columns represent each set of cell lines, with the number of transcripts listed in or above the column.

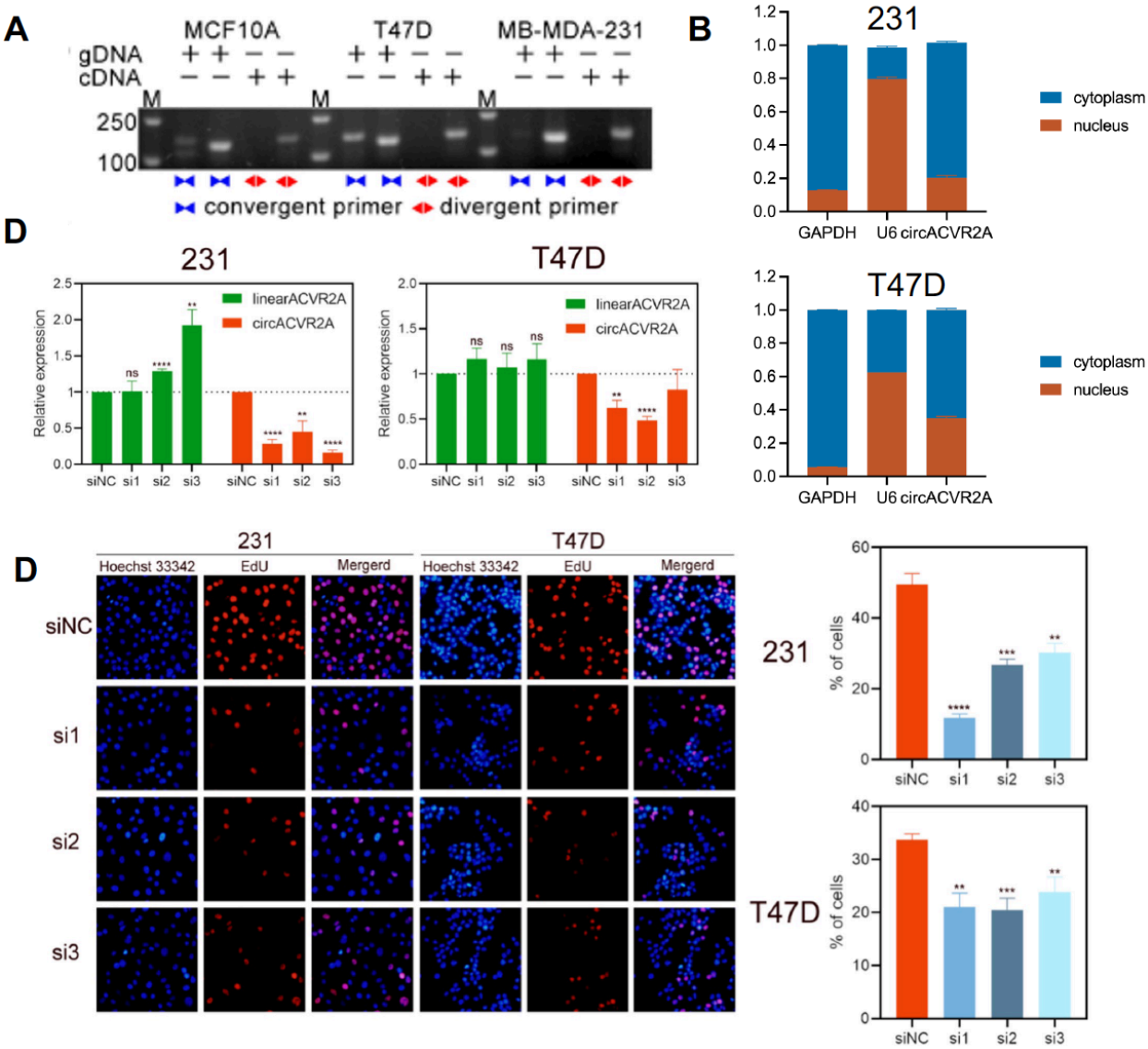
Supplementary Figure S4



Supplementary Figure 4: overlap of circRNAs screened in our study compared with previous studies

A) Venn diagram illustrating the overlap of circRNAs screened across four studies: Li et al. (red, 762 circRNAs), Liu et al. (green, 3,354 circRNAs), Chen et al. (orange, 1,507 circRNAs), and our study (blue, 9,853 circRNAs). The diagram shows that 429 circRNAs are shared among all four studies, while a significant number of circRNAs are unique to our study.
B) Venn diagram showing the overlap of essential circRNAs identified in at least one cell line between Liu et al. (1,556 essential circRNAs) and our study (1,344 essential circRNAs), with 200 shared essential circRNAs.
C) Scatter plot comparing circRNA $\log_2$ fold changes from the A375 melanoma cell line in our study (x-axis) and the average skin $\log_2$ fold change from Liu et al. (y-axis). ($r = 0.23$, $p = 2.05 \times 10^{-3}$)
F) Scatter plot comparing circRNA $\log_2$ fold changes from the 22Rv1 prostate cancer cell line in our study (x-axis) and Chen et al.'s results (y-axis). ($r = 0.25$, $p = 1.19 \times 10^{-3}$)

Supplementary Figure S5



Supplementary Figure 5: Functional validation of context specific dependence circRNA in breast cancer.

A) Identification of circACVR2A(2,3,4) back-splice junction sites.

B) The circACVR2A(2,3,4) abundance detected by cytoplasm/nucleus separation assay.

C) siRNA knockdown efficiency in T47D and 231 cell lines.

D) The proliferative abilities of transfected T47D and 231 cells measured by EdU assay.

Supplementary Table 1:

Comparisons final adjusted circRNA shRNA score and corresponding average log2 fold

change from the screen

Supplementary Table 2:

Comparisons final adjusted circRNA shRNA score and corresponding T8vsT0 robust rank agreement score from the screen

Supplementary Table 3:

Comparison of $\log_2$ fold change across cell lines for linear transcripts, positive controls, and negative controls across T8 and T16 samples relative to T0.

Supplementary Table 4:

Comparison of dropout score across the screened cell lines for the union of essential linear transcripts and the DepMap shRNA screening project.

Supplementary Table 5:

Comparison of $\log_2$ fold change across cell lines for circular transcripts, positive controls, and negative controls across T8 and T16 samples relative to T0.

Supplementary Table 6:

Primer, oligo sequences and siRNAs used in this study.