



OPEN ACCESS

Original research

# CCT3-*LINC00326* axis regulates hepatocarcinogenic lipid metabolism

Jonas Nørskov Søndergaard ,<sup>1</sup> Christian Sommerauer ,<sup>1</sup> Ionut Atanasoai ,<sup>1</sup> Laura C Hinte ,<sup>1</sup> Keyi Geng ,<sup>1</sup> Giulia Guiducci ,<sup>2</sup> Lars Bräutigam ,<sup>3</sup> Myriam Aouadi ,<sup>4</sup> Lovorka Stojic ,<sup>2</sup> Isabel Barragan ,<sup>5</sup> Claudia Kutter <sup>1</sup>

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2021-325109>).

<sup>1</sup>Department of Microbiology, Tumor, and Cell Biology, Science for Life Laboratory, Karolinska Institute, Stockholm, Sweden

<sup>2</sup>Barts Cancer Institute, Centre for Cancer Cell and Molecular Biology, John Vane Science Centre, Queen Mary University of London, London, UK

<sup>3</sup>Comparative Medicine, Karolinska Institute, Stockholm, Sweden

<sup>4</sup>Department of Medicine, Karolinska Institute, Stockholm, Sweden

<sup>5</sup>Department of Physiology and Pharmacology, Karolinska Institute, Stockholm, Sweden

## Correspondence to

Dr Claudia Kutter, Department of Microbiology, Tumor, and Cell Biology, Science for Life Laboratory, Karolinska Institute, Stockholm, Sweden; [claudia.kutter@ki.se](mailto:claudia.kutter@ki.se)

Received 9 May 2021

Accepted 9 December 2021

Published Online First  
12 January 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY. Published by BMJ.

**To cite:** Søndergaard JN, Sommerauer C, Atanasoai I, et al. *Gut* 2022;**71**:2081–2092.

## ABSTRACT

**Objective** To better comprehend transcriptional phenotypes of cancer cells, we globally characterised RNA-binding proteins (RBPs) to identify altered RNAs, including long non-coding RNAs (lncRNAs).

**Design** To unravel RBP-lncRNA interactions in cancer, we curated a list of ~2300 highly expressed RBPs in human cells, tested effects of RBPs and lncRNAs on patient survival in multiple cohorts, altered expression levels, integrated various sequencing, molecular and cell-based data.

**Results** High expression of RBPs negatively affected patient survival in 21 cancer types, especially hepatocellular carcinoma (HCC). After knockdown of the top 10 upregulated RBPs and subsequent transcriptome analysis, we identified 88 differentially expressed lncRNAs, including 34 novel transcripts. CRISPRa-mediated overexpression of four lncRNAs had major effects on the HCC cell phenotype and transcriptome. Further investigation of four RBP-lncRNA pairs revealed involvement in distinct regulatory processes. The most noticeable RBP-lncRNA connection affected lipid metabolism, whereby the non-canonical RBP CCT3 regulated *LINC00326* in a chaperonin-independent manner. Perturbation of the CCT3-*LINC00326* regulatory network led to decreased lipid accumulation and increased lipid degradation *in cellulo* as well as diminished tumour growth *in vivo*.

**Conclusions** We revealed that RBP gene expression is perturbed in HCC and identified that RBPs exerted additional functions beyond their tasks under normal physiological conditions, which can be stimulated or intensified via lncRNAs and affected tumour growth.

## INTRODUCTION

Liver cancer encompasses a collection of clinically diverse tumour subtypes that arise from malignant liver cells. Hepatocellular carcinoma (HCC) is the most common form affecting ~80% of all patients.<sup>1</sup> The cause of HCC is often attributed to intrinsic, extrinsic and unknown idiopathic factors.<sup>2</sup> Perturbation of cellular homeostasis leading to uncontrolled cell growth and proliferation is characteristic for HCC but the underlying molecular consequences are only partly understood. In recent years, more focus has been placed on studying deregulated long non-coding RNAs (lncRNAs) in cancer.<sup>3,4</sup> However, functionality of many lncRNAs still remains to be

## Significance of this study

### What is already known on this subject?

- ⇒ RNA-binding proteins (RBPs) play crucial roles in cancer.
- ⇒ RBPs can function through long non-coding RNAs (lncRNAs).
- ⇒ Cell type-specific expression patterns of lncRNAs have strong diagnostic, prognostic and therapeutic value.

### What are the new findings?

- ⇒ Integrative multiomics approach, including a new liver cancer cohort dataset, describes the transcriptional landscape in hepatocellular carcinoma (HCC).
- ⇒ RBPs are significantly deregulated in HCC mounting in reduced patient survival.
- ⇒ Pathological RBPs control metabolic activity and apoptosis in HCC cells through lncRNAs.
- ⇒ Chaperonin complex subunit CCT3 moonlights as a novel RBP.
- ⇒ CCT3 functions in HCC lipid metabolism via the long non-coding RNA *LINC00326* *in cellulo* and *in vivo*.

### How might it impact on clinical practice in the foreseeable future?

- ⇒ Diagnostic and prognostic potential of HCC stage-specific response of CCT3 and *LINC00326*.
- ⇒ Beneficial therapeutic effects of lncRNA overexpression at a specific HCC disease stage.

explored. lncRNAs are transcripts longer than 200 nucleotides that either originate from intergenic regions (lincRNAs) or coincide within transcriptional units of different genes.<sup>5</sup> Given the unique cell type- and disease-specific expression patterns of lncRNAs,<sup>6,7</sup> they are emerging targets for biomarker and therapeutic developments since their presence affects primarily the diseased cell. lncRNAs regulate various processes, such as cell cycle, proliferation, apoptosis and cell death.<sup>5</sup> These actions are likely mediated through interaction with RNA-binding proteins (RBPs).<sup>8</sup>

As vital enzymes, RBPs control RNA regulatory pathways.<sup>9</sup> RBP activity is adjusted to the cellular demand of RNA transcripts. Altering RBP

gene expression levels has profound implications on cellular physiology and contributes to the phenotypic abnormalities commonly observed in atypical and cancer cells.<sup>10</sup> RBPs contain known or predicted RNA-binding domains (RBD). For example, the canonical RBP Insulin Like Growth Factor 2 mRNA Binding Protein 1 (IGF2BP1) contains six RBDs through which IGF2BP1 regulates mRNA stability, such as, by impeding access of miRNAs to their targets.<sup>11</sup> Over 500 proteins with classical RBDs have been identified in human cells.<sup>12</sup> Moreover, additional proteins with RNA-binding capacity were found through newer technologies (online supplemental table S1). While these non-canonical RBPs have well-established biological functions, they can also moonlight as RBPs.<sup>13</sup> For instance, Tripartite Motif 25 (TRIM25) ubiquitinates proteins for degradation<sup>14</sup> and as a non-canonical RBP binds RNA to regulate innate immune response pathways.<sup>14</sup> Likewise, Alpha Enolase (ENO1) is indispensable in glycolysis, and its enzymatic activity is abolished through the interaction with a lncRNA.<sup>15</sup>

Due to the frequent dependency between lncRNAs and RBPs,<sup>8</sup> we here used a novel RBP-centric approach to identify functional lncRNAs. We curated a list of 2282 RBPs reported in RNA interactome capture experiments across different cell types (online supplemental table S1) and found aberrant RBP gene expression profiles in two HCC patient cohorts. Perturbation of selected RBPs in *in cellulo* and *in vivo* settings revealed the underlying regulatory networks through which RBPs can act. Our approach led to the identification of new functional lncRNAs and revealed their regulatory roles in HCC. Specifically, we found that *LINC00326* regulates lipid metabolism through its interaction with the non-canonical RBP *CCT3*.

## RESULTS

### Differentially expressed RNA-binding proteins impact survival of patients with hepatocellular carcinoma

Deregulation of genes is a key feature of cancer. Due to their regulatory capacity, genes encoding for RBPs have gained more attention.<sup>10</sup> We therefore built a comprehensive catalogue of genes with RNA-binding capacity in human by inspecting RNA interactome capture experiments and gene ontology (GO) databases (online supplemental table S1). The resulting list comprised 1321 canonical (containing known RBDs) and 959 non-canonical (without any characterised RBDs) RBPs (online supplemental table S1). We compared gene expression levels and survival probabilities of RBP *versus* all other protein-coding genes across 21 different human cancer types (The Cancer Genome Atlas, TCGA).<sup>16</sup> RBP gene expression was higher irrespective of the cancer type (two-tailed unpaired student's t-test,  $p < 0.001$ ) (figure 1A), which is in accordance to previous estimates.<sup>17</sup> Division of RBPs into canonical and non-canonical RBPs yielded the same results (online supplemental figure S1a). To investigate the impact of RBP deregulation on patient survival, we compared the cox proportional hazard coefficient (coxph) of RBP and non-RBP genes for each cancer type.<sup>16</sup> We found the largest fold change (FC) in HCC (LIHC, 6.3 FC) followed by sarcoma (SARC, 5.1 FC) (figure 1B). Kidney cancers (KIRP and KIRC) had relatively high absolute coxph values for RBP and non-RBP genes and thus lower FCs (4.2 and 1.6, respectively) than the other top-ranked cancer types. Therefore, RBP gene expression had a greater prognostic value in liver cancer than any other investigated cancer type, which might be explained due to high proliferation rate<sup>17</sup> and global activation of RBP gene copies.<sup>18</sup> By comparing paired data of tumour and peritumour tissue from 50 LIHC cohort patients (online supplemental figure S1b),

we found 92 upregulated and 68 downregulated RBP genes (figure 1C, online supplemental table S2).

To validate our findings in a TCGA-independent HCC cohort, we profiled matched pairs of primary tumour and peritumour liver samples of 24 patients with HCC from the Australian Victorian Cancer Biobank (Australia HCC) by RNA-sequencing (RNA-seq) (figure 1D). Similar to the TCGA LIHC cohort, the expression pattern separated the tumour and peritumour tissue samples in the Australia HCC cohort (online supplemental figure S1c). Our differential gene expression analysis revealed 111 upregulated and 64 downregulated RBP genes (figure 1D–F, online supplemental tables S3 and S4). In both cohorts, we identified a common set of 63 upregulated and 47 downregulated RBP genes (figure 1E,F). We complemented our analysis by including gene expression data of the human HCC cell lines HepG2 and Huh7.<sup>19</sup> A total number of 26 upregulated and 23 downregulated RBP genes showed similar gene expression patterns between all HCC datasets (figure 1E,F, online supplemental figure S1e).

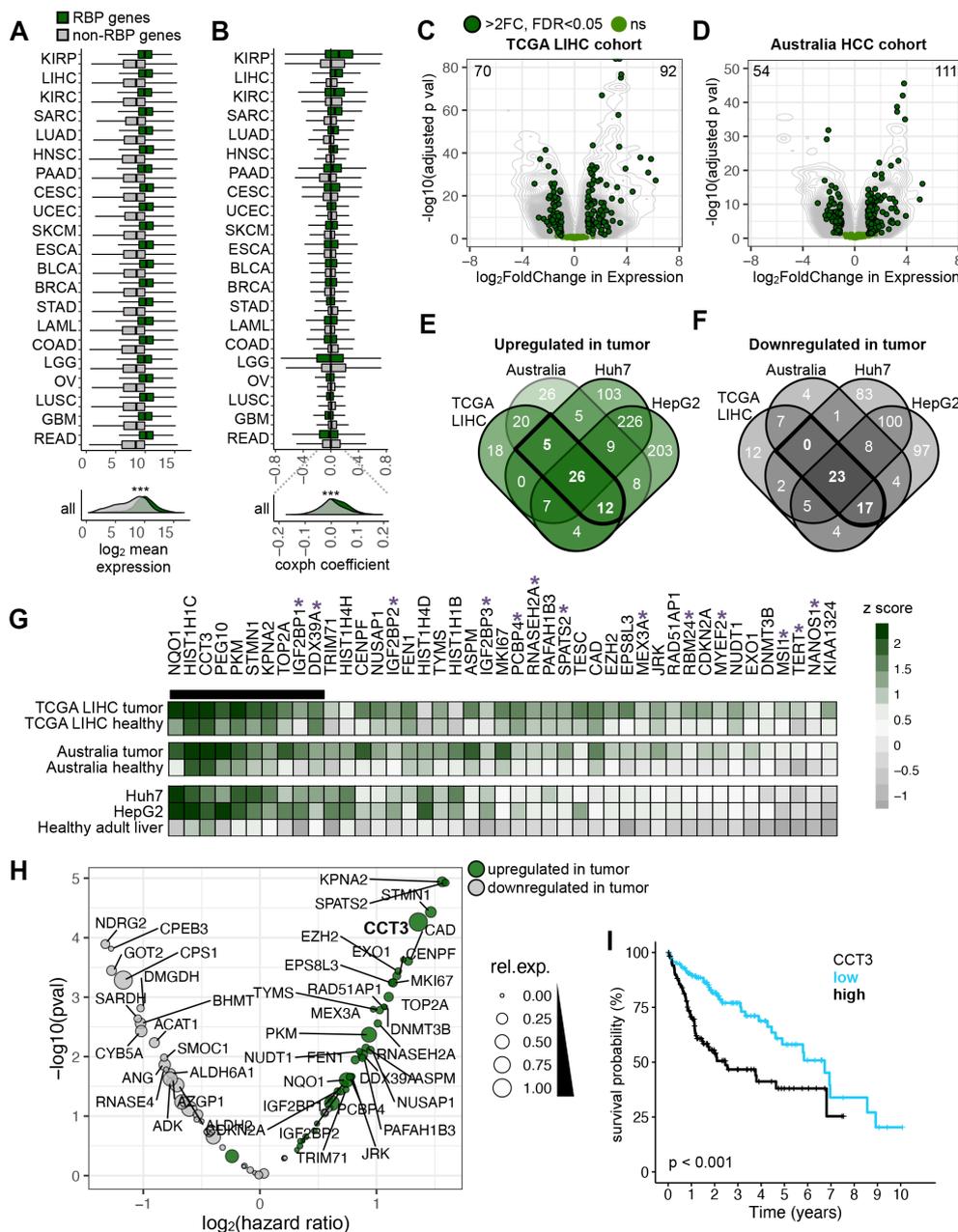
We ranked z-score gene expression values across all datasets and selected the top 10 highest expressed canonical and non-canonical RBPs that were upregulated in HCC (*NQO1*, *HIST1H1C*, *CCT3*, *PEG10*, *PKM*, *STMN1*, *KPNA2*, *TOP2A*, *IGF2BP1*, *DDX39A*) (figure 1G). By analysing Cap Analysis of Gene Expression transcriptome data,<sup>20</sup> we confirmed upregulation of these RBP genes in HCC (online supplemental figure S1d). The selected RBPs have paralogous genes and belong to different protein families that are composed of diverse predicted protein domains (online supplemental figure S1f). Protein expression of these RBPs was detectable in multiple subcellular compartments (online supplemental figure S1g).

These upregulated RBP genes carried a spectrum of mutations and copy number variations in HCC. In particular, Chaperonin Containing TCP1 Subunit 3 (*CCT3*) was frequently amplified (online supplemental figure S1h). We inspected the hazard ratio (HR) and found that high RBP gene expression levels were associated with poor patient survival (figure 1H). The majority (7/10) of RBPs had a significant prognostic value (corrected for tumour stage, age and gender) (figure 1H,I and online supplemental figure S2). In comparison to all RBPs, high expression of Karyopherin Subunit Alpha 2 (*KPNA2*) had the largest HR closely followed by *CCT3* and Stathmin 1 (*STMN1*) (figure 1H). Within their respective gene families, *CCT3*, *IGF2BP1*, *KPNA2*, NAD(P)H Quinone Dehydrogenase 1 (*NQO1*), and *STMN1* had the highest HR, while DExD-Box Helicase 39A (*DDX39A*), Histone Cluster 1 H1 Family Member C (*HIST1H1C*), Paternally Expressed 10 (*PEG10*) and DNA Topoisomerase II Alpha (*TOP2A*) had gene family members with a higher HR (online supplemental figure S3a). Elevated gene expression levels for most family members was associated with decreased patient survival (figure 1I, online supplemental figure S3b,c). Similar results were previously obtained when dividing *CCT3* gene expression levels into halves<sup>21</sup> instead of terciles.

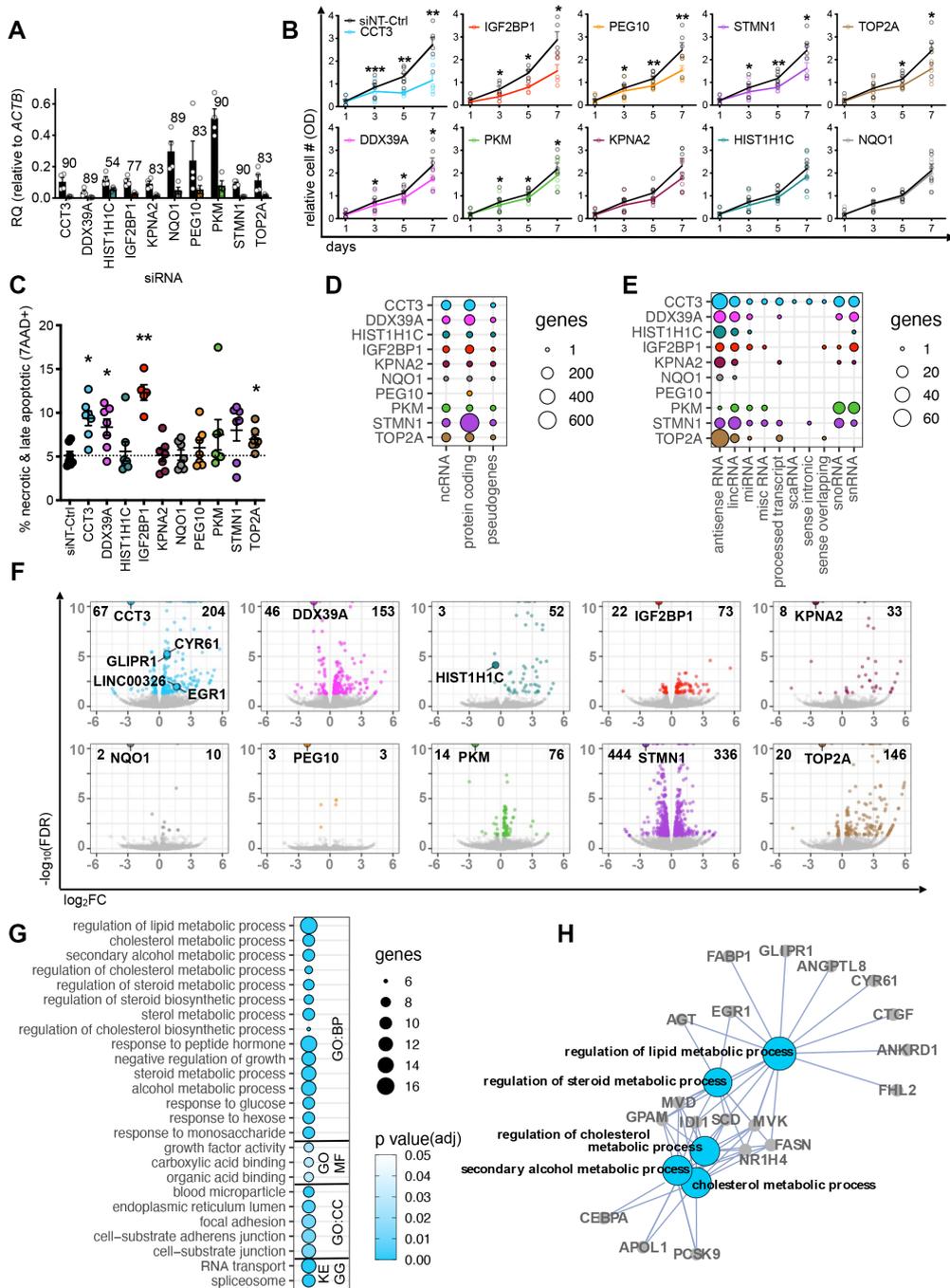
In summary, we identified a set of RBPs belonging to diverse gene families that were deregulated in HCC. RBP gene expression levels could therefore be used as potential prognostic markers in HCC patients.

### RBP knockdown reduces cancer growth and changes non-coding RNA expression in HCC cell lines

In order to dissect the functional roles of upregulated RBP genes in the HCC cohorts and cell lines, we performed siRNA-mediated gene knockdown (KD) in Huh7 and HepG2 (figure 2,



**Figure 1** RBPs are deregulated in cancer and affect patient survival. (A,B) Boxplots of (A) gene expression level and (B) cox proportional hazard (coxph) coefficient of protein-coding genes grouped into RBP (green) and non-RBP (grey) genes. Each row represents a different cancer type defined by TCGA. Each cancer type consisted of 144 to 1006 patients. Hinges correspond to the first and third quartiles, and whiskers correspond to the 1.5-times interquartile range. (C,D) Volcano plots demonstrate differentially expressed (DE) genes in the two HCC cohorts (C) TCGA-LIHC and (D) Australia HCC. Data points represent DE RBP genes (dark green), not significantly (ns)-DE RBP genes (FDR>0.05, light green) and all other genes (grey). (E,F) Four-way Venn diagrams intersect the number of (E) upregulated and (F) downregulated RBP genes in the TCGA and Australia HCC cohorts as well as in liver cancer cell lines Huh7 and HepG2. Intersections highlighted (bold) show the number of RBP genes commonly deregulated in HCC cohorts and cell lines. (G) Heatmap displays changes in expression levels for commonly upregulated RBP genes in HCC as highlighted in (E). Expression level is sorted by average tumour z-score from left to right. Black bar marks the top 10 highest expressed RBP genes, and purple asterisk marks RBPs with a canonical RBD. Colour gradient indicates z-score differences (green: high; grey: low). (H) Volcano plot displays comparison of the top and bottom tercile in RBP gene expression levels and hazard ratios within the TCGA-LIHC cohort (377 patients) (grey: downregulated and green: upregulated in tumour). The size of the circle represents the gene expression level of each RBP relative to each other (broad: high, narrow: low). (I) Kaplan-Meier plot shows the association of *CCT3* gene expression level and 10-year survival within the top and bottom tercile within the TCGA-LIHC cohort (377 patients) (black: high and blue: low *CCT3* gene expression levels). Statistics: log-rank (Mantel-Cox). BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical squamous cell carcinoma and endocervical adenocarcinoma; COAD, colon adenocarcinoma; DE, differentially expressed; ESCA, oesophageal carcinoma; GBM, glioblastoma multiforme; HCC, hepatocellular carcinoma; HNSC, head and neck squamous cell carcinoma; KIRC, kidney renal clear cell carcinoma; KIRP, kidney renal papillary cell carcinoma; LAML, acute myeloid leukaemia; LGG, brain lower grade glioma; LIHC, liver hepatocellular carcinoma; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; RBP, RNA-binding proteins; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; TCGA, The Cancer Genome Atlas; UCEC, uterine corpus endometrial carcinoma.



**Figure 2** Reduced expression of RBP genes impact various cellular and molecular responses in HCC cell lines. (A) Bar graph displays RBP gene expression levels before (black bars) and after siRNA-mediated RBP-KD (coloured bar) relative to *ACTB* in HepG2 and Huh7 cells determined by qPCR (n=4 mean,  $\pm$ SEM). Number above each bar shows the average KD efficiency in percent. Colour code: *CCT3* (blue), *DDX39A* (magenta), *HIST1H1C* (turquoise), *IGF2BP1* (red), *KPNA2* (plum), *NQO1* (grey), *PEG10* (yellow), *PKM* (green), *STMN1* (purple) and *TOP2A* (brown). Individual replicates are displayed by white circles. (B) Line graphs show the relative number of metabolically active cells (measured by optical density) over 7 days after siRNA-mediated RBP-KDs assayed by the MTT assay (n=5). Black line: non-targeting siRNA control (siNT-Ctrl), coloured line: RBP-specific siRNA KD. (C) Dot plot represents the percentage of dead cells after siRNA-mediated RBP-KDs (colour-coded) after 5 days (n=7, mean,  $\pm$ SEM). (B,C) Statistics: paired two-tailed t-test, \*p<0.05, \*\*p<0.01, \*\*\*p<0.001. (D,E) Circle plots display number of genes per RNA biotype affected by RBP-KDs. The diameter of the circles corresponds to the number of genes in each category. Deregulated genes falling into different ncRNA subcategories are shown in figure 2E. (F) Volcano plots demonstrate DE genes after RBP-KDs. Data points represent significantly DE genes (colour-coded by RBP-KD, FDR<0.05) and not significantly DE genes (grey, FDR>0.05). Bolded numbers on the top of each graph indicate total number of DE genes (FDR<0.05). The wider circle in each plot highlights the downregulated RBP. Genes with an FDR value smaller than  $1 \times 10^{-10}$  were collapsed at  $1 \times 10^{-10}$ . (G) Circle plot shows GO term and KEGG pathway enrichment analysis of deregulated genes after the *CCT3*-KD (FDR<0.05). The diameter of the circles corresponds to the number of genes in each GO or KEGG term and the colour code represents varying degrees of significance (white: high and blue: low p value). (H) Interaction network displays connections of the five most significant GO BP terms in figure 2G. GO term is bolded and gene names are highlighted. BP, biological process; CC, cellular compartment; DE, differentially expressed; KD, knockdown; MF, molecular function; OD, optical density; RQ, relative quantity.

online supplemental figure S4a–f). We achieved high RBP-KD efficiencies and confirmed a mean reduction of 80% (range: 54%–90%) by RT-qPCR (figure 2A). For the majority (7/10) of RBPs, the KDs resulted in a significant decrease in the number of metabolically active cells 3 days post-transfection (MTT assay) (figure 2B) due to an increase of cell death (FACS for late apoptosis/necrosis) (figure 2B,C), especially after KD of *CCT3* and *IGF2BP1*.

To investigate the molecular mechanisms underlying the observed cellular phenotypes of the RBP-KDs, we profiled gene expression by RNA-seq. Each RBP-KD affected different annotated RNA biotypes (12 different categories) (figure 2D–G, online supplemental tables S5–S7). The majority of differentially expressed (DE) genes were protein-coding (mean: 120, range: 6–697) followed by non-coding (mean: 44, range: 0–105) and pseudogenes (mean: 6, range: 0–18) (figure 2D). Closer inspection of the non-coding RNA biotype revealed that genes encoding for antisense and lincRNAs were frequently deregulated after the RBP-KD (figure 2E). The *CCT3*-KD showed the highest diversity in RNA biotypes and affected the highest number of non-coding genes ( $n=105$ ) among all RBPs tested. The *STMN1*-KD led to the highest ( $n=780$ ), while the *PEG10*-KD led to the lowest ( $n=6$ ) number of DE genes (figure 2F). The number of DE genes observed in our RBP-KD experiments is in accordance to previous shRBP-KD experiments ( $n=235$ ) in K562 cells (online supplemental figure S4g).<sup>22,23</sup> To understand whether the altered expression of RBP genes impacts specific regulatory processes, we performed a GO term and KEGG pathway analysis of the DE genes for each RBP-KD (figure 2G,H, online supplemental figure S5, online supplemental table S8). Deregulated genes belonged to GO terms distinct for each RBP-KD, e.g., lipid metabolism (*CCT3*-KD), angiogenesis (*DDX39A*-KD), response to oxygen levels (*IGF2BP1*-KD), RNA transport (*PKM*-KD), lipid localisation and transport (*STMN1*-KD) and DNA packaging and conformation change (*TOP2A*-KD) (online supplemental figure S5, online supplemental table S8).

In sum, deregulation of some RBPs in HCC caused diverse alteration of gene regulatory programmes linked to metabolism and lipogenesis.

### Each RBP-KD is accompanied by gene expression changes of a specific set of annotated and novel lincRNA genes

Since GO terms are largely curated based on information obtained from protein-coding genes, the functional association of deregulated non-coding RNAs remains uncertain. Beyond protein-coding genes, we found 54 deregulated annotated lincRNAs (figure 3A) of which 69% (37/54) and 31% (17/54) were deregulated in one or more RBP-KDs, respectively. LincRNAs were more up- (76%, 41/54) than downregulated (24%, 13/54) after RBP-KDs (figure 3B). Among the upregulated lincRNAs were Lung Cancer Associated Transcript 1 (*LUCAT1*) and Promoter of CDKN1A Antisense DNA Damage Activated RNA (*PANDAR*), which have been linked to various cancer types.<sup>24,25</sup> Hierarchical clustering of the annotated deregulated lincRNAs revealed three distinct groups of commonly up- and downregulated lincRNAs as well as lincRNAs that were specifically affected in a single RBP-KD experiment (figure 3B). We expected that lincRNAs in the latter group have a specific rather than a general role in cancer pathways (figure 3B).

At the point of the analysis, 7307 lincRNA genes had been annotated in the human genome (Gencode GRCh38 v.27). To identify novel lincRNAs that were only detectable after RBP-KDs, we *de novo* annotated lincRNA in our RNA-seq data

and found 34 novel lincRNAs that were upregulated in at least one RBP-KD (figure 3C–E, online supplemental table S9). It is possible that those novel lincRNAs have been undetected previously due to rapid turnover in the presence of highly abundant RBPs.

Overall, our data-driven approach revealed that the majority of novel lincRNAs were specific for one RBP-KD, in particular after the *CCT3*-KD (figure 3D).

### Overexpression of lincRNA genes alters cancer cell phenotypes towards early apoptosis

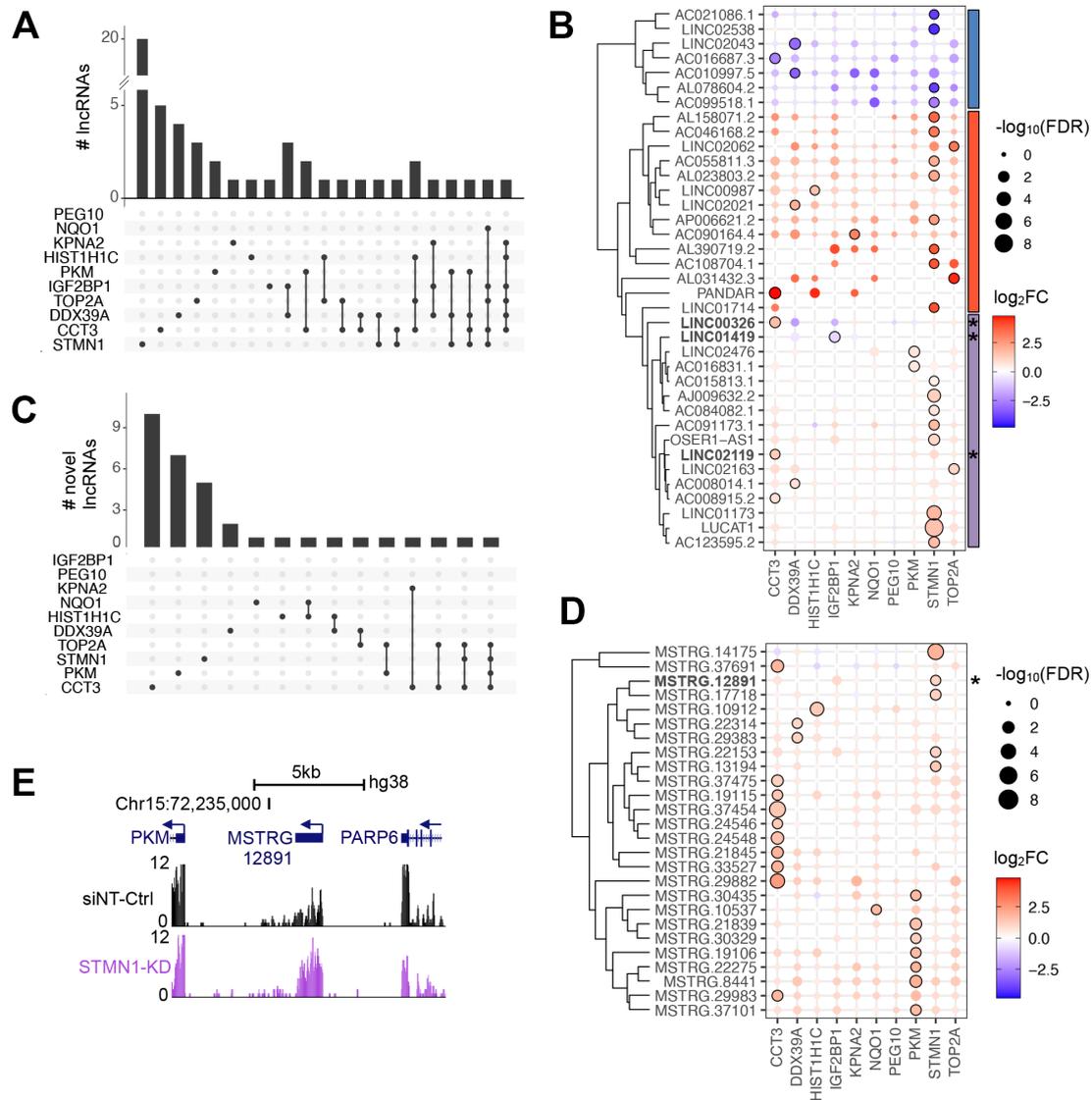
To further characterise the roles of lincRNAs in HCC, we applied stringent selection criteria including (1) FC in lincRNA gene expression, (2) RBP-specific dependency, (3) lincRNA abundance and (4) visual inspection in the genome browser. This led to a subsequent analysis of four lincRNAs, *LINC00326*, *LINC01419*, *LINC02119* and *MSTRG.12891*. Since lower RBP gene expression levels resulted in increased expression levels of lincRNA genes (figure 3B,D), we hypothesised that an increase in lincRNA gene expression reveals a liver cancer cell-specific phenotype. We used CRISPR-VPR activation (CRISPRa) with three target-specific guide RNAs (gRNAs) to overexpress selected lincRNA genes (lincRNA-OE) and achieved a 2-fold to 20-fold increase of lincRNA gene expression that lasted for several days (figure 4A,B).

In all four lincRNA-OE experiments, we detected a reduced number of metabolically active Huh7 and HepG2 cells and early cell apoptosis, which was significant for *LINC00326* and *MSTRG.12891* (figure 4C,D, online supplemental figure S6a–e). To explain the phenotype, we performed RNA-seq. Overall, lincRNA-OEs resulted in more upregulation than downregulation of genes (figure 4E, online supplemental figure S6f, online supplemental table S10–12) and many of the deregulated genes were specific for the respective lincRNA-OE (figure 4F). In contrast to the RBP-KDs, lincRNA-OEs resulted in frequent alteration of protein-coding genes and only few changes in ncRNA and pseudogene expression (figure 4G, online supplemental figure S6g–h). We performed GO term and KEGG pathway analyses (online supplemental figure S7, online supplemental table S13) and found an enrichment in lipid transporter activity after *LINC00326*-OE and *MSTRG.12891*-OE and growth factor binding after *LINC01419*-OE and *LINC02119*-OE.

In conclusion, lincRNA-OEs resulted in increased early apoptosis and reduced metabolic activity as well as transcriptional alteration of specific biological pathways. In particular, we found the strongest perturbation when increasing gene expression of *CCT3*-dependent *LINC00326* in lipid metabolism.

### A *CCT3*-*LINC00326* network regulates lipid metabolism

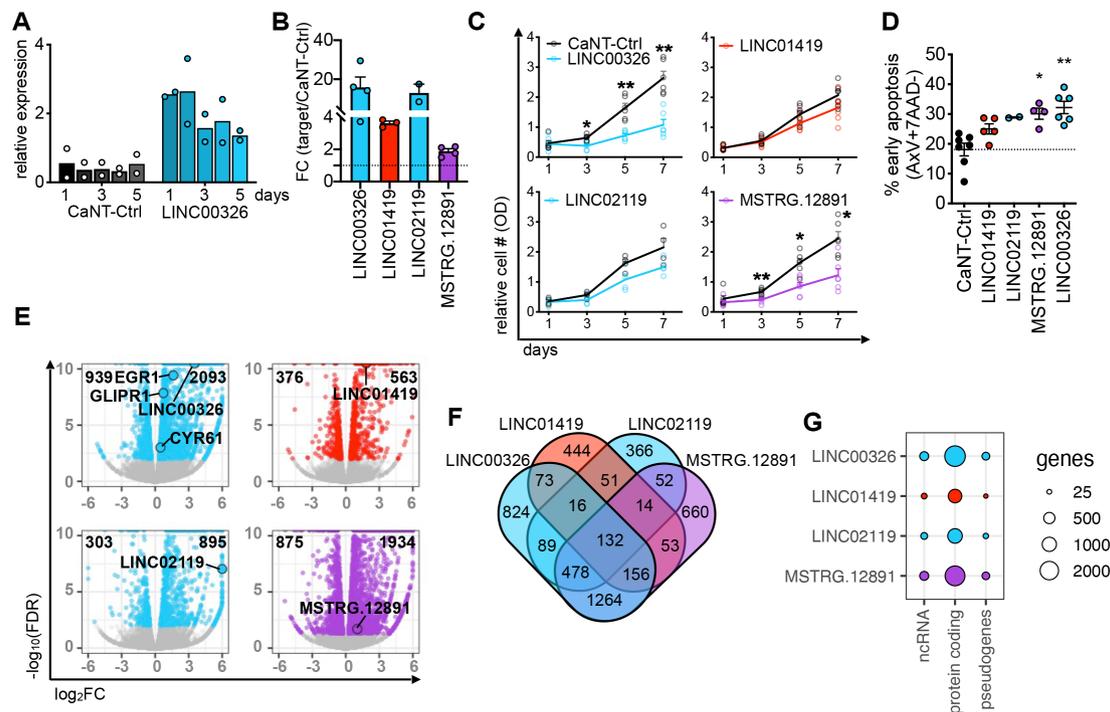
Due to the strongest phenotypical severity, we further investigated the interaction of *CCT3* and *LINC00326*. After overexpression, we found that *LINC00326* but not the liver-specific control lincRNA *HULC* was coimmunoprecipitated with *CCT3* in liver cancer cells (figure 5A). Other components of the chaperonin-complex did not coimmunoprecipitate with *LINC00326*, thus indicating a chaperonin-independent function. Likewise, KD of these chaperonin-components did not lead to *LINC00326* upregulation (online supplemental figure S8a,b). In addition, *CCT3* protein and *LINC00326* were located in the same cellular compartment enabling their interaction (figure 5B,C, online supplemental figures S1g, S8c–e,<sup>26</sup>). *CCT3* RNA stoichiometry data further suggested an enrichment of a chaperonin-independent function in HCC when compared



**Figure 3** RBP-KD affects lincRNA gene expression levels. (A,C) Bar graph show the number of (A) annotated and (C) novel DE lincRNA genes detectable after RBP-KDs. The frequency of lincRNA genes in one (black dot) or multiple (black dots connected by a line) RBP-KD experiments is shown. (B,D) Circle plots display the occurrence of (B) annotated and (D) novel lincRNA genes per RBP-KD. The diameter of the circles corresponds to varying degrees of significance (large: high, and narrow: low FDR value, black line:  $\text{FDR} < 0.05$ ). The colour code represents fold change (red: upregulated and blue: downregulated). Vertical bars specify the three most common clusters defining lincRNAs as either consistently downregulated (blue) or upregulated (red), or with varying pattern deregulation across the ten RBP-KD (purple). A star (\*) marks lincRNAs used for further investigation. (E) The UCSC genome browser view demonstrates the genomic location of the novel lincRNA *MSTRG.12891* in between genes encoding for *PKM* and *PARP6*. Arrows indicate direction of gene transcription. Gene expression patterns in Huh7 cells transfected with siNT-Ctrl (black) or siRNA-mediated KD of *STMN1* (purple) are shown. The y-axis of each track specifies normalised RNA-seq read intensity.

with healthy tissue and 14 other cancer types (online supplemental figure S8f). Remarkably, the double-KD of *CCT3* and *LINC00326* rescued the *CCT3*-KD phenotype, further verifying their functional dependence (figure 5D,E). Functional inspection of the 70 commonly deregulated genes of the *CCT3*-KD and *LINC00326*-OE (figure 5F) suggested an involvement in the regulation of lipid metabolic processes, response to decreased oxygen levels and angiogenesis (figure 5G–H). Lipids are degraded via peroxidation.<sup>27</sup> Accordingly, a balanced cellular oxygen supply is interrelated with processes regulating vascularisation, such as angiogenesis.<sup>28</sup> None of the other RBP-lincRNA interactions assayed in this study acted through these biological pathways (online supplemental figure S9a–e). Most genes (8/10) were regulated in similar directions in the *CCT3*-KD and the *LINC00326*-OE (figure 5I). When including our HCC cohort

and HCC cell line datasets (figure 1E,F), we noticed that a cluster consisting of genes encoding for Early growth response protein 1 (*EGR1*), Glioma pathogenesis-related protein 1 (*GLIPR1*) and Cysteine-rich angiogenic inducer 61 (*CYR61*) were frequently lower expressed than their corresponding non-carcinogenic controls (figure 5I). This suggests that increased expression of these genes may contribute to a more physiologically normal cellular phenotype. To examine how the *CCT3*-*LINC00326* core genes were regulated, we determined DNA-binding motifs enriched in the promoter regions of the 70 commonly deregulated genes over 1000 random sequences (figure 5F,J, online supplemental figure S9f). The most significant motifs were recognised by CREM/CREB/ATF transcription factors (TFs) (figure 5J), which were entwined with regulation of lipid metabolism and general Pol II transcription (figure 5K). Moreover, ChIP-seq data



**Figure 4** Overexpression of lincRNA genes in HCC cell lines causes molecular and cellular alterations. (A) Bar graphs exemplify the increase in relative gene expression of *LINC00326* over 5 days after transfection of a CRISPRa vector with non-targeting gRNA-CRISPRa controls (CaNT-Ctrl, black gradient) or three lincRNA-specific gRNAs (*LINC00326*, blue gradient) determined by RT-qPCR. (B) Bar graph demonstrate the fold change in lincRNA gene expression 2 days after CRISPRa transfection of lincRNA-specific versus CaNT-Ctrl determined by RT-qPCR ( $n=2-4$ , mean,  $\pm$ SEM). The colour-code links the lincRNA to the respective RBP-KD experiment in which the lincRNA was identified (blue: *CCT3*, red: *IGF2BP1*, purple: *STMN1*). (C) Line graphs show relative increase in metabolically active HCC cell number over 7 days after CRISPRa transfection with CaNT-Ctrl (black) or lincRNA-specific gRNAs (coloured) determined by MTT assay. (D) Dot graph shows the percentage of early apoptotic cells 5 days after CRISPRa transfection with CaNT-Ctrl (black) and lincRNA-specific gRNAs (coloured) determined by FACS ( $n=2-8$ , mean,  $\pm$ SEM). (A–D) Each biological replicate of Huh7 and HepG2 is displayed by circles. Graphs are coloured according to the colour-code selected for the RBP partner through which the lincRNA was identified (blue: *CCT3*, red: *IGF2BP1*, purple: *STMN1*). Statistics: paired two-tailed t-test, \* $p<0.05$ , \*\* $p<0.01$ . (E) Volcano plots demonstrate DE genes 2 days after CRISPRa transfection of lincRNA-specific versus CaNT-Ctrl determined by RNA-seq in Huh7. Data points represent significantly DE genes (coloured,  $FDR<0.01$ ) and not significantly DE genes (grey,  $FDR>0.01$ ). Bolded numbers on the top of each graph indicate total number of DE genes ( $FDR<0.01$ ). Circle highlights lincRNA and lincRNA-interacting genes investigated. (F) Four-way Venn diagram intersects the number of DE genes after each lincRNA-OE experiment ( $FDR<0.01$ ). (G) Circle plot displays the number of genes per RNA biotype affected by the lincRNA-OE. The diameter of the circles corresponds to the number of genes in each category.

in liver and liver cancer cell lines<sup>23</sup> revealed direct binding of the CREM/CREB/ATF TFs to the promoters of the lipid metabolism related genes *EGR1*, *CYR61* and *GLIPR1* (figure 5L).

Thus, coordinated regulation of CREM/CREB/ATF TFs could lead to the observed expression changes of genes controlling lipid metabolism, hypoxia and angiogenesis.

### The *CCT3-LINC00326* network reduces tumour burden *in cellulo* and *in vivo*

Because our molecular data revealed that *CCT3* and *LINC00326* affected lipid metabolism, we examined the cellular impact of this RBP-lincRNA interaction. By comparing *CCT3*-KD and *LINC00326*-OE to their respective controls, we measured a significant increase in lipid degradation (1.9 and 1.8 FC, respectively) (figure 6A), a significant decrease in lipid accumulation ( $-2.1$  and  $-1.5$  FC, respectively) (figure 6B) and elevated levels of reactive oxygen species (ROS) (both 1.3 FC) although not statistically significant (figure 6C). Since these assays confirmed that alteration of *CCT3* and *LINC00326* gene expression levels modulate regulation of lipid metabolism, we inspected publicly available data from patient liver biopsies with lipid metabolism disorders (GSE126848 and TCGA). We found an increase in

*CCT3* gene expression, which correlated with the severity of metabolic-associated fatty liver disease and the pathological stage of HCC (figure 6D,E). This indicated that *CCT3* functionality in cell lines can be recapitulated in the human body and is associated with one of the HCC aetiologies. *LINC00326* was not profiled in this study.

Because *LINC00326* had not been assessed *in vivo*, we performed cell line-derived xenograft experiments. We used human Huh7-GFP cells with reduced *CCT3* or elevated *LINC00326* gene expression, injected them into zebrafish embryos and monitored their cell growth over time. *CCT3*-KD and *LINC00326*-OE resulted in a significant suppression in tumour growth in comparison to the respective controls (figure 6F–J) confirming that low *CCT3* and high *LINC00326* gene expression reduced tumour burden.

In summary, our study demonstrated that functional lincRNAs can be identified by using an RBP-centric approach, through which we uncovered that the *CCT3-LINC00326* interaction regulates lipid metabolism in cancer cells. We show that modulation of lincRNA biogenesis via RBPs can alter cancer cell-specific activities, such as cancer cell survival and tumour growth (figure 6K).

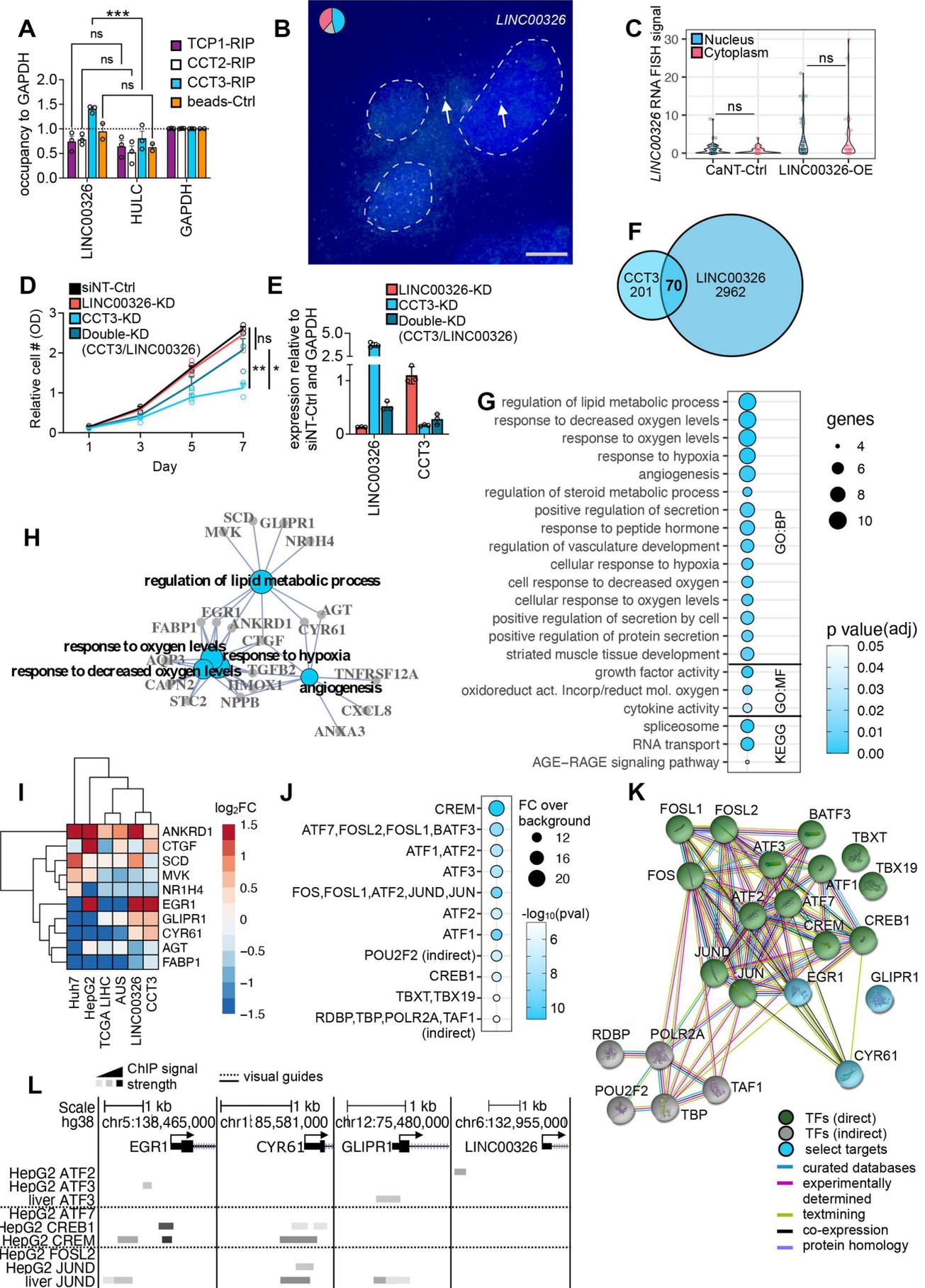


Figure 5 Continued

**Figure 5** The CCT3-*LINC00326* interactome regulates lipid metabolism. (A) Bar graph shows enrichment of *LINC00326* compared with GAPDH and HULC (negative controls) over input control after RNA immunoprecipitation with a TCP1 (CCT1)- (purple), CCT2- (white) or CCT3-specific (blue) antibody versus the beads-only (no antibody) control (orange) followed by RT-qPCR with gene- and strand-specific primers (n=3, mean, +SEM). Each biological replicate is displayed by circles. Statistics: ANOVA with a Bonferroni's multiple comparison test, \*\*\*p<0.001, ns: non-significant. (B) Microscopic image of single-molecule RNA FISH using exonic probes for *LINC00326* (white dots and arrows) in *LINC00326*-OE Huh7 HCC cells. DAPI (blue) marks the nucleus. Pie chart represents the fraction of signals in the nucleus (blue) or cytoplasm (pink) in cells, or cells without any signal (grey). Scale bar: 5  $\mu$ m. (C) Violin plots show quantification of *LINC00326* RNA FISH signal localisation in Huh7 HCC cells (n=40–53, statistics: paired two-tailed t-test, ns: non-significant). (D) Line graphs shows the relative number of metabolically active cells (measured by optical density, OD) over 7 days after the siRNA-mediated CCT3-KD and/or *LINC00326*-KD assayed by MTT assay (n=3, mean, +SEM, statistics: ANOVA with a Bonferroni's multiple comparison test, \*p<0.05, \*\*p<0.01). (E) Bar charts show siRNA-KD-efficiencies of targeted RBP genes (figure 5D) (n=3, mean, +SEM). (F) Two-way Venn diagram intersects the number of deregulated genes after CCT3-KD and *LINC00326*-OE. (G) Circle plot shows GO term and KEGG pathway enrichment analysis of the 70 commonly deregulated genes after CCT3-KD and *LINC00326*-OE. The diameter of the circles corresponds to the number of genes in each GO or KEGG term and the colour code represents varying degrees of significance (white: high and blue: low p value). (H) Interaction network displays connections of the five most significant GO BP terms shown in figure 5G. GO term is bolded and gene names are highlighted. (I) Heatmap (unsupervised clustering) displays the fold change in expression levels for lipid metabolic process genes (figure 5H) when comparing HCC cohorts and cell lines, CCT3-KD and *LINC00326*-OE over non-cancerous or NT controls, respectively. Colour gradient indicates log<sub>2</sub>FC differences (red: high; blue: low). (J) Circle plot demonstrates enrichment of TF-binding motifs in the promoter regions of the 70 commonly deregulated genes of the CCT3-KD and *LINC00326*-OE. The diameter of the circles corresponds to the fold change over background controls and the colour code represents varying degrees of significance (white: high and blue: low p value). Identified motifs for each TF are shown in online supplemental figure S9f. (K) StringDB interaction network shows the links of the TFs identified in figure 5F and known interaction partners (direct: green, indirect: grey). Two direct connections to lipid metabolism genes are highlighted (blue). (L) The UCSC genome browser view demonstrates genomic location of three lipid metabolism-associated genes and the *LINC00326* gene. Arrows indicate direction of gene transcription. Horizontal bars indicate ChIP-seq signals (black: strong; grey: weak) for available TF-binding events in HepG2 or liver cells (ENCODE).

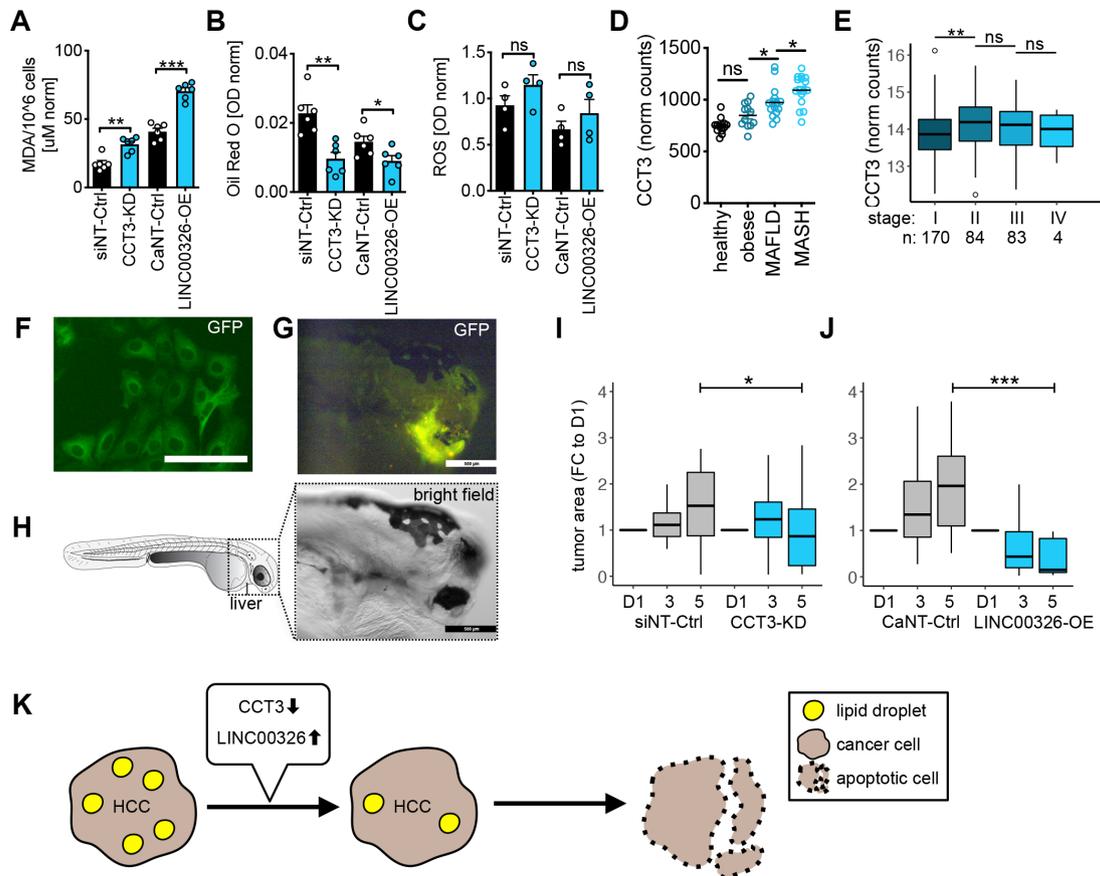
## DISCUSSION

Through advanced transcriptomic RBP-RNA profiling, RBP-RNA complex purification and functional screening (online supplemental table S1), the number of proteins with RNA-binding capacity has risen from a few hundreds to ~2300. *Bona fide* functions of canonical RBPs have been well characterised, and the role of non-canonical RBPs in RNA metabolism are beginning to be unravelled.<sup>15</sup> By using an unbiased RBP-centric approach, we found that both canonical and non-canonical RBPs perturb liver cancer pathology by acting through an entangled network that required the involvement of lincRNAs. Phenotypic alterations were particularly severe after reducing gene expression of *CCT3* and *IGF2BP1*. Interestingly, the *CCT3*-KD influenced a large number of ncRNAs, especially lincRNAs and antisense RNAs. In contrast, the *IGF2BP1*-KD affected largely protein-coding transcripts, which has been observed previously.<sup>29</sup> The preference towards specific RNA types could be explained through differences in RNA binding modalities. *CCT3* is a non-canonical RBP without apparent RBDs, whereas *IGF2BP1* represents a canonical RBP exerting its RNA binding activity through six RBDs.<sup>11</sup> *CCT3* together with seven other CCT subunits is known for forming a stoichiometrically even cytosolic chaperonin complex that ensures proper protein-folding. Chaperonin-independent functions of each member have been speculated because of transcriptional and phenotypic differences after altering expression levels of individual CCT genes<sup>30,31</sup> and disparities in protein abundance across intracellular compartments.<sup>32</sup> Despite the high amino acid identity, each member evolved differences in protein regions that are essential for substrate specificity.<sup>33</sup> Since the underlying genetic sequences of CCT genes are under purifying selection, new paralog-specific functions have been developing,<sup>34</sup> perhaps even as a consequence of emerging new non-coding RNA substrates. It is therefore plausible that *CCT3* functions as a non-canonical RBP independent of its role in the chaperonin complex. To gain a better understanding whether RBPs act through regulation of lincRNAs in HCC, we assayed four lincRNAs for which our data indicated a strong functional connection with RBPs. Three of the four lincRNAs (*LINC00326*, *LINC01419* and *LINC02119*)

were previously annotated and one represented a novel lincRNA (*MSTRG.12891*). When altering the expression of these lincRNAs, we observed major transcriptional and phenotypic changes in HCC cells.

Across all assays performed, the *CCT3*-*LINC00326* interaction caused the most severe molecular and cellular effects. We detected elevated *CCT3* gene expression in HCC, which is in accordance to previous reports and underscores its prognostic value in HCC.<sup>35,36</sup> Overall, *CCT3* was highly expressed in malignant cells when compared with various tissue types (figure 1, online supplemental figure S10a). In contrast, *LINC00326* abundance was low in tumour tissues but increased after the *CCT3*-KD in liver cancer cell lines (figure 3). Under normal physiological conditions, *LINC00326* was only detectable in testis (online supplemental figure S10b). Based on *LINC00326* gene expression patterns during spermatogenesis,<sup>37</sup> *LINC00326* may function in cell proliferation and controlled apoptosis to eliminate irreparable damaged germ cells during development but the exact regulatory mechanisms remain to be determined. Interestingly, *LINC00326* gene expression is almost completely diminished in testicular cancers (online supplemental figure S10b) and gradually decreased with increased testicular cancer severity (online supplemental figure S10c). This implied that antiproliferative and proapoptotic properties of *LINC00326* (figure 4) peaks at the early stages of carcinogenesis. Besides, single-cell RNA-seq data from healthy testis showed high *CCT3* gene expression at early stages of spermatogenesis. Reduced *CCT3* gene expression at a later developmental stage increased *LINC00326* levels (online supplemental figure S10d–f). This temporal expression pattern in testis is in accordance with our *CCT3*-KD and suggests that high levels of *CCT3* suppress *LINC00326*.

Reducing gene expression from a high to a moderate level has smaller effects than increasing gene expression from a low to a high level.<sup>38</sup> Accordingly, KD of the highly expressed *CCT3* caused a small increase of *LINC00326* transcript abundance, while CRISPRa-mediated OE substantially stimulated *LINC00326* gene expression and could explain why more genes were affected by *LINC00326*-OE than *CCT3*-KD in liver cancer



**Figure 6** The CCT3-*LINC00326* network affects lipid metabolism and tumour growth *in vitro* and *in vivo*. (A–C) Bar graphs show comparison of (A) malondialdehyde (MDA) production (lipid degradation), (B) Oil Red O staining (lipid accumulation) and (C) ROS production of CCT3-KD or *LINC00326*-OE (blue) to the respective NT controls (black) 48 hours after transfection (n=4–6, mean, ±SEM). Each biological replicate is displayed by circles. Statistics: paired two-tailed t-test, (D) Dot graph displays library size-normalised *CCT3* mRNA expression level in adult human individuals with normal and obese weight, MAFLD and MASH (n=2–8, mean). *LINC00326* was not assayed. Statistics: one-way ANOVA. (E) Boxplot of normalised *CCT3*-expression in the TCGA-LIHC cohort divided by main pathological cancer stage. Statistics: one-way ANOVA with Tukey Honest Significant Differences test. (F–H) Microscopy images of TUBULIN-GFP expressing Huh7 cells (F) *in vitro* (scale bar: 100µM) and (G–H) *in vivo* in zebrafish xenografts (scale bar: 500 µM). (I, J) Box plots show changes in tumour area in zebrafish xenografts after (I) CCT3-KD (n=20–21, mean, ±SEM) or (J) *LINC00326*-OE (n=21–28, mean, ±SEM). Individual zebrafish were followed over 5 days and tumour area is given relative fold change to day 1 (D1) after injection. Statistics: one-way ANOVA, \*p<0.05, \*\*\*p<0.001. (K) Schematic model for CCT3-*LINC00326* regulation of lipid metabolism. Reducing *CCT3* or increasing *LINC00326* gene expression in liver cancer cells inhibits lipid accumulation and promotes lipid degradation (peroxidation). Due to the strong dependency of cancer cells towards high lipogenesis, this in turn slows down cancer growth and promotes cell death. \*p<0.05, \*\*p<0.01, \*\*\*p<0.001, MAFLD, metabolic-associated fatty liver disease; MASH, metabolic-associated steatohepatitis; NS, non-significant; ROS, reactive oxygen species.

cells. Inspection of genes overlapping both perturbations revealed an involvement in controlling lipid metabolism pathways. Our comprehensive cell-based assays confirmed that *LINC00326*-OE and *CCT3*-KD led to a decrease in lipid accumulation and increase in peroxidation. Our findings support the increasing recognition of lipid metabolism as cancer confounder.<sup>39–40</sup> Response to hypoxia and oxygen levels as well as angiogenesis were additional commonly induced pathways, which are consistent with prior reports ascribing atypical hypoxic and angiogenic conditions to tumours.<sup>41</sup> These pathways are linked mechanistically whereby a hypoxic condition leads to an excess of NADPH, which is used for lipogenesis, and has been proposed to maintain a balanced redox environment.<sup>42</sup> Accordingly, we also observed a trend towards increased ROS-production after *CCT3*-KD or *LINC00326*-OE. Taken together, our data supported that the CCT3-*LINC00326* network plays a vital role in liver cancer pathology via perturbing lipid metabolism.

Inspection of the promoter regions of deregulated genes emerging after altering *CCT3* and *LINC00326* gene expression revealed a common set of enriched transcription factors (TFs) that acted in a coordinated manner (figure 5K). Our network analysis showed that the TFs CREM, CREB and ATF were not only linked to the Pol II machinery but also to genes that regulate lipid metabolism. Furthermore, we found experimental evidence of TF-binding to promoters of lipid metabolic genes (figure 5L). For instance, we confirmed strong CREM/CREB1-binding to the *EGR1* promoter. *EGR1* is in itself a TF, and accumulating evidence substantiates its tumour suppressing role in HCC.<sup>43</sup> Loss of *EGR1* and tumour development are connected through oncogenic RAS-PI3K signalling<sup>44</sup>, which is also a top GO term that we identified after *LINC00326*-OE (online supplemental figure S7). *CCT3*-KD and *LINC00326*-OE did not lead to upregulation of these TF genes but TF activity can still be altered. Previous studies described that ATF2<sup>45</sup> and JUN<sup>46</sup> proteins interact with

CCT3. It is therefore plausible that after upregulation of CCT3 in HCC, CCT3 binds to TFs (including ATF2), sequesters and thereby prevents TFs from binding to promoter regions of genes involved in lipid metabolism and *LINC00326* (online supplemental figure S11a). Our RIP-qPCR assay confirmed the interaction of CCT3 and *LINC00326* (figure 5A). When *LINC00326* binds to CCT3, it may impede CCT3's confinement of ATF2, thereby releasing ATF2 from its inactive state and thus allowing transcription of lipid metabolism genes and *LINC00326* itself (online supplemental figure S11b). Further studies will shed more light on the regulatory intricacies of the proposed CCT3-*LINC00326* network.

Most RBPs were identified in more than one recent study (average: 6.5, median: 5) (online supplemental table S1) and many have not yet been validated. With 2282 detectable RBPs in HCC, including 959 non-canonical, the future challenge is to establish the exact functional links between RBP and RNA. Nevertheless, by using an RBP-centric approach, we have identified prognostically relevant RBPs and lincRNAs with major functional molecular and cellular roles in HCC. The combination of loss-of-function and gain-of-function *in cellulo* and *in vivo* experiments allowed us to construct networks that regulate oncogenic lipid metabolism, impair cellular energy consumption and increase intracellular oxidative stress. Given that the lincRNAs investigated in this study are barely expressed under normal physiological conditions, we speculate that they act as an inhibitor to prevent cellular transformation of healthy hepatocytes into malignant cells. As such, they could represent novel markers for liver pathologies and molecular targets for future HCC treatment approaches.

## MATERIALS AND METHODS

Cell-based and molecular assays as well as xenograft experiments are described in the online supplemental materials and methods.

Supplementary tables and microscopic imaging files are accessible via Figshare: <https://figshare.com/s/2c05765158269b3b4ff2>, <https://figshare.com/s/a83dbee52555e922ca8d> and <https://figshare.com/s/08b0f84f2ea241b03c8d>.

Datasets generated in this study are deposited under ArrayExpress accessions E-MTAB-8915, E-MTAB-9587 and E-MTAB-9586.

Scripts used for bioinformatics analyses are available on Github: <https://github.com/jonasns/LiveRNome>.

### Patient material

Patients (75% men and 25% women) taking part in this study had HCC from Hepatitis B virus infection, non-alcoholic fatty liver disease, alcoholic steatohepatitis, hereditary haemochromatosis and other HCC-triggering conditions. See patient information on ArrayExpress E-MTAB-8915. It was neither possible nor appropriate to involve patients or the public in the design, conduct, reporting or dissemination plans of our research.

**Twitter** Jonas Nørskov Søndergaard @jonasns, Christian Sommerauer @ChrisSommerauer, Ionut Atanasoai @IonutAtanasoai, Laura C Hinte @laurahinte, Keyi Geng @keyi\_G, Myriam Aouadi @AouadiLab, Lovorka Stojic @stojic\_lovorka and Claudia Kutter @ClaudiaKutter

**Acknowledgements** We are grateful for receiving the constitutively TUBULIN-GFP expressing Huh7 cells from Ralf Bartenschlager and the CRISPR-dCas9-VPR overexpression vector from George Church. We thank the laboratories of Marc Friedländer and Vicent Pelechano as well as Hassan Foroughi Asl and Volker Lauschke for helpful feedback concerning experimental procedures, data analyses and data presentation. We appreciated critical comments on the manuscript by Carlos Gallardo Dodd. We thank the TCGA Research Network ([www.cancer.gov/](http://www.cancer.gov/)

tcga) for their publicly available data and the zebrafish core facility at the Karolinska Institute for service and husbandry of animals.

**Contributors** JNS and CK conceptualised the project. JNS, CS, IA, LCH, KG, LB, GG, LS and IB performed the laboratory experiments. IB was responsible for obtaining the Australia HCC patient samples. JNS did the analysis and visualised the data. JNS, LS, IB and CK acquired funding. JNS and CK wrote the original draft. CK acted as guarantor. All authors contributed to the review and editing process.

**Funding** This work was supported by the Knut & Alice Wallenberg foundation (KAW 2016.0174, CK), Ruth & Richard Julin foundation (2017-00358, 2018-00328, 2020-00294, CK); SFO-SciLifeLab fellowship (SFO\_004, CK), Swedish Research Council (2019-05165, CK), Lillian Sagen & Curt Ericsson research foundation (2021-00427, CK), Gösta Miltons research foundation (2021-00527, CK), Chinese Scholarship Council (KG, CK), KI-KID funding (2016-00189 and 2018-00904, CK), SNIC projects (20177-154 and sllstore2017022), Nilsson-Ehle Endowments (JNS), Barts and London Charity (MGU0404, LS) Cancer Research UK (RCCFEL100007, LS), AIRC Fellowship for Abroad (GG), Tornspiran Foundation (IB) and Svenska Läkaresällskapet (SLS-693561 and SLS-694791) (IB).

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants. Tissue specimens were obtained under informed consent from 24 patients undergoing liver resection of HCC from the Australian Victorian Biobank, according to the approval of local ethics committee (application number: 2010/541-31/1 and 2017/719-31/2). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available in a public, open access repository. All data relevant to the study are included in the article or uploaded as supplementary information. All data generated or analysed during this study are included in the article. Supplementary tables and microscopic imaging files are accessible via Figshare: <https://figshare.com/s/2c05765158269b3b4ff2>, <https://figshare.com/s/a83dbee52555e922ca8d> and <https://figshare.com/s/08b0f84f2ea241b03c8d>. Original sequencing data are available from ArrayExpress under accession numbers: E-MTAB-8915, E-MTAB-9587 and E-MTAB-9586. Scripts used for bioinformatics analyses are available on Github: <https://github.com/jonasns/LiveRNome>. Raw data not available in the online supplemental tables can be found in supplemental data 1. Additionally, all relevant data are available from the authors on request.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

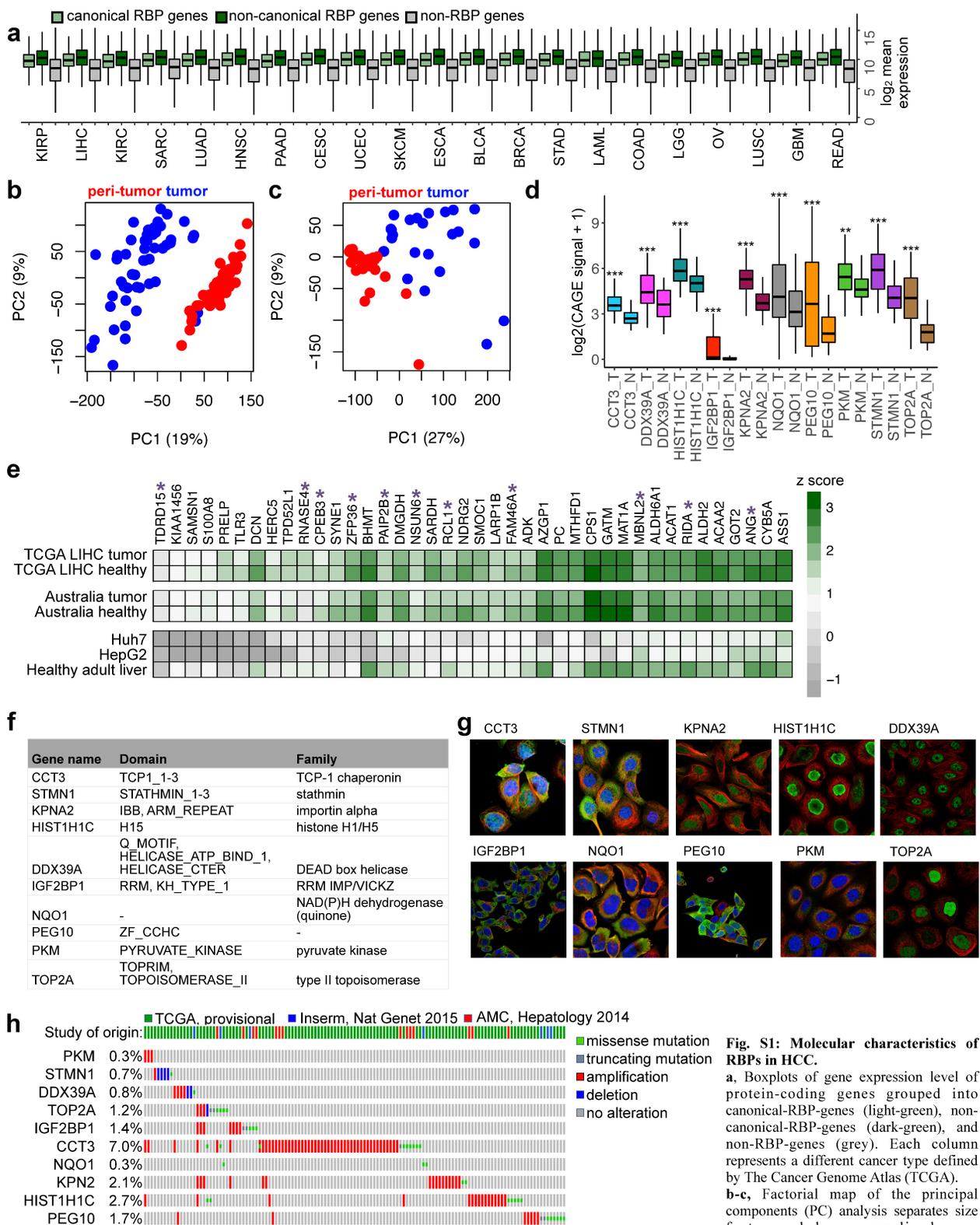
### ORCID iDs

Jonas Nørskov Søndergaard <http://orcid.org/0000-0002-4438-6756>  
 Christian Sommerauer <http://orcid.org/0000-0001-7132-7172>  
 Ionut Atanasoai <http://orcid.org/0000-0002-0757-6681>  
 Laura C Hinte <http://orcid.org/0000-0002-4226-0009>  
 Keyi Geng <http://orcid.org/0000-0003-0892-7460>  
 Giulia Guiducci <http://orcid.org/0000-0002-0632-6808>  
 Lars Bräutigam <http://orcid.org/0000-0003-2834-1512>  
 Myriam Aouadi <http://orcid.org/0000-0001-6256-7107>  
 Lovorka Stojic <http://orcid.org/0000-0001-6691-3396>  
 Isabel Barragan <http://orcid.org/0000-0002-8586-6502>  
 Claudia Kutter <http://orcid.org/0000-0002-8047-0058>

## REFERENCES

- Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424.
- Sia D, Villanueva A, Friedman SL, et al. Liver Cancer Cell of Origin, Molecular Class, and Effects on Patient Prognosis. *Gastroenterology* 2017;152:745–61.

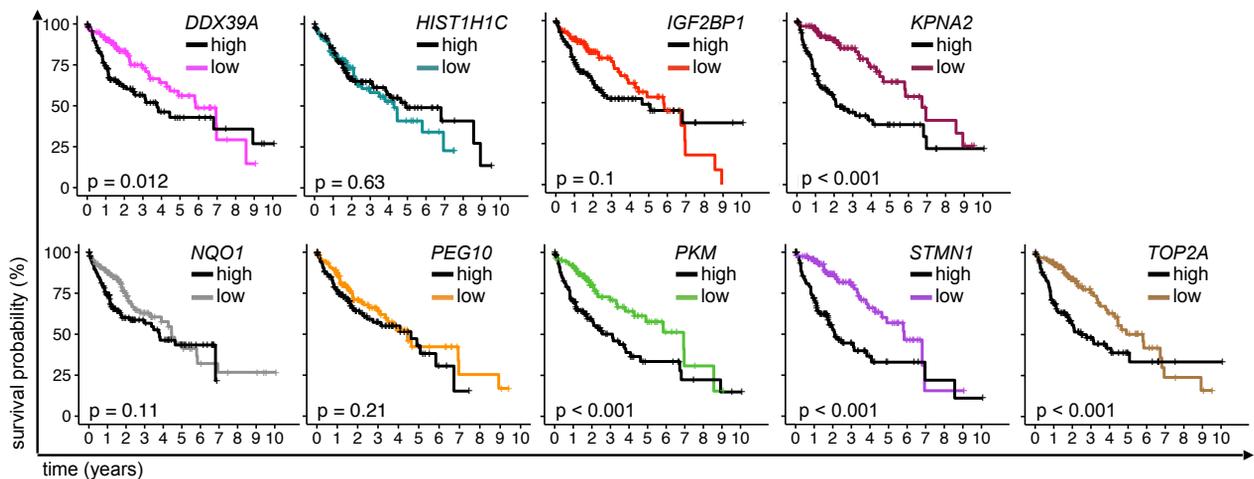
- 3 Iyer MK, Niknafs YS, Malik R, *et al.* The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet* 2015;47:199–208.
- 4 Klingsberg M, Matsuda A, Diederichs S, *et al.* Non-Coding RNA in hepatocellular carcinoma: mechanisms, biomarkers and therapeutic targets. *J Hepatol* 2017;67:603–18.
- 5 Gutschner T, Diederichs S, Gutschner T. The hallmarks of cancer: a long non-coding RNA point of view. *RNA Biol* 2012;9:703–19.
- 6 Arun G, Diermeier SD, Spector DL. Therapeutic targeting of long non-coding RNAs in cancer. *Trends Mol Med* 2018;24:257–77.
- 7 Kutter C, Watt S, Stefflova K, *et al.* Rapid turnover of long noncoding RNAs and the evolution of gene expression. *PLoS Genet* 2012;8:e1002841.10.1371/journal.pgen.1002841
- 8 Wong C-M, Tsang FH-C, Ng IO-L. Non-Coding RNAs in hepatocellular carcinoma: molecular functions and pathological implications. *Nat Rev Gastroenterol Hepatol* 2018;15:137–51.
- 9 Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet* 2014;15:829–45.
- 10 Pereira B, Billaud M, Almeida R. Rna-Binding proteins in cancer: old players and new actors. *Trends Cancer* 2017;3:506–28.
- 11 Müller S, Bley N, Glaß M, *et al.* Igf2Bp1 enhances an aggressive tumor cell phenotype by impairing miRNA-directed downregulation of oncogenic factors. *Nucleic Acids Res* 2018;46:6285–303.
- 12 Castello A, Fischer B, Frese CK, *et al.* Comprehensive identification of RNA-binding domains in human cells. *Mol Cell* 2016;63:696–710.
- 13 Beckmann BM, Horos R, Fischer B, *et al.* The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun* 2015;6:10127.
- 14 Choudhury NR, Nowak JS, Zuo J, *et al.* Trim25 is an RNA-specific activator of Lin28a/TuT4-mediated uridylation. *Cell Rep* 2014;9:1265–72.
- 15 Yu S, Li N, Huang Z, *et al.* A novel lncRNA, TCONS\_00006195, represses hepatocellular carcinoma progression by inhibiting enzymatic activity of ENO1. *Cell Death Dis* 2018;9:1–13.
- 16 Anaya J. OncoLnc: linking TCGA survival data to mRNAs, miRNAs, and lncRNAs. *PeerJ Comput Sci* 2016;2:e67.
- 17 Zhang B, Babu KR, Lim CY, *et al.* A comprehensive expression landscape of RNA-binding proteins (RBPs) across 16 human cancer types. *RNA Biol* 2020;17:211–26.
- 18 Dang H, Takai A, Forgues M, *et al.* Oncogenic activation of the RNA binding protein NELFE and Myc signaling in hepatocellular carcinoma. *Cancer Cell* 2017;32:101–14.
- 19 Rudolph KLM, Schmitt BM, Villar D, *et al.* Codon-driven translational efficiency is stable across diverse mammalian cell states. *PLoS Genet* 2016;12:e1006024–23.
- 20 Hashimoto K, Suzuki AM, Dos Santos A, *et al.* Cage profiling of ncRNAs in hepatocellular carcinoma reveals widespread activation of retroviral LTR promoters in virus-induced tumors. *Genome Res* 2015;25:1812–24.
- 21 Liu Y, Zhang X, Lin J, *et al.* CCT3 acts upstream of YAP and TFCP2 as a potential target and tumour biomarker in liver cancer. *Cell Death Dis* 2019;10:1–15.
- 22 Van Nostrand EL, Freese P, Pratt GA, *et al.* A large-scale binding and functional map of human RNA-binding proteins. *Nature* 2020;583:711–9.
- 23 Davis CA, Hitz BC, Sloan CA, *et al.* The encyclopedia of DNA elements (encode): data portal update. *Nucleic Acids Res* 2018;46:D794–801.
- 24 Carlevaro-Fita J, Lanzas A, Feuerbach L, *et al.* Cancer lncRNA census reveals evidence for deep functional conservation of long noncoding RNAs in tumorigenesis. *Commun Biol* 2020;3:1–16.
- 25 Das Mahapatra K, Pasquali L, Søndergaard JN, *et al.* A comprehensive analysis of coding and non-coding transcriptomic changes in cutaneous squamous cell carcinoma. *Sci Rep* 2020;10:1–12.
- 26 Cui X, Hu Z-P, Li Z, *et al.* Overexpression of chaperonin containing TCP1, subunit 3 predicts poor prognosis in hepatocellular carcinoma. *World J Gastroenterol* 2015;21:8588–604.
- 27 Behn C, Araneda OF, Llanos AJ, *et al.* Hypoxia-Related lipid peroxidation: evidences, implications and approaches. *Respir Physiol Neurobiol* 2007;158:143–50.
- 28 Medina J, Arroyo AG, Sánchez-Madrid F, *et al.* Angiogenesis in chronic inflammatory liver disease. *Hepatology* 2004;39:1185–95.
- 29 Huang H, Weng H, Sun W, *et al.* Recognition of RNA N<sup>6</sup>-methyladenosine by IGF2BP proteins enhances mRNA stability and translation. *Nat Cell Biol* 2018;20:285–95.
- 30 Spiess M, Echbarshi M, Svanström A, *et al.* Over-Expression analysis of all eight subunits of the molecular chaperone CCT in mammalian cells reveals a novel function for CCTdelta. *J Mol Biol* 2015;427:2757–64.
- 31 Amit M, Weisberg SJ, Nadler-Holly M, *et al.* Equivalent mutations in the eight subunits of the chaperonin CCT produce dramatically different cellular and gene expression phenotypes. *J Mol Biol* 2010;401:532–43.
- 32 Queiroz RML, Smith T, Villanueva E, *et al.* Comprehensive identification of RNA-protein interactions in any organism using orthogonal organic phase separation (OOPS). *Nat Biotechnol* 2019;37:169–78.
- 33 Chagoyen M, Carrascosa JL, Pazos F, *et al.* Molecular determinants of the ATP hydrolysis asymmetry of the CCT chaperonin complex. *Proteins* 2014;82:703–7.
- 34 Archibald JM, Logsdon Jr. JM, Doolittle WF. Origin and evolution of eukaryotic chaperonins: phylogenetic evidence for ancient duplications in CCT genes. *Mol Biol Evol* 2000;17:1456–66.
- 35 Yao L, Zou X, Liu L. The TCP1 ring complex is associated with malignancy and poor prognosis in hepatocellular carcinoma. *Int J Clin Exp Pathol* 2019;12:3329–43.
- 36 Hou J-Y, Wu H-Y, He R-Q, *et al.* Clinical and prognostic value of chaperonin containing t-complex 1 subunit 3 in hepatocellular carcinoma: a study based on microarray and RNA-sequencing with 4272 cases. *Pathol Res Pract* 2019;215:177–94.
- 37 Zhu Z, Li C, Yang S, *et al.* Dynamics of the transcriptome during human spermatogenesis: predicting the potential key genes regulating male gametes generation. *Sci Rep* 2016;6:19069.
- 38 Chavez A, Tuttle M, Pruitt BW, *et al.* Comparison of Cas9 activators in multiple species. *Nat Methods* 2016;13:563–7.
- 39 Beloribi-Djefafalia S, Vasseur S, Guillaumond F. Lipid metabolic reprogramming in cancer cells. *Oncogenesis* 2016;5:e189.
- 40 Hausser J, Szekely P, Bar N, *et al.* Tumor diversity and the trade-off between universal cancer tasks. *Nat Commun* 2019;10.
- 41 Hanahan D, Weinberg RA. The hallmarks of cancer. *Cell* 2000;100:57–70.
- 42 Menendez JA, Lupu R. Fatty acid synthase and the lipogenic phenotype in cancer pathogenesis. *Nat Rev Cancer* 2007;7:763–77.
- 43 Magee N, Zhang Y. Role of early growth response 1 in liver metabolism and liver cancer. *Hepatoma Res* 2017;3:268–77.
- 44 Shin SY, Bahk YY, Ko J, *et al.* Suppression of Egr-1 transcription through targeting of the serum response factor by oncogenic H-ras. *Embo J* 2006;25:1093–103.
- 45 Lau E, Kluger H, Varsano T, *et al.* Pkc $\epsilon$  promotes oncogenic functions of ATF2 in the nucleus while blocking its apoptotic function at mitochondria. *Cell* 2012;148:543–55.
- 46 Li X, Wang W, Wang J, *et al.* Proteomic analyses reveal distinct chromatin-associated and soluble transcription factor complexes. *Mol Syst Biol* 2015;11:775.



**Fig. S1: Molecular characteristics of RBPs in HCC.**

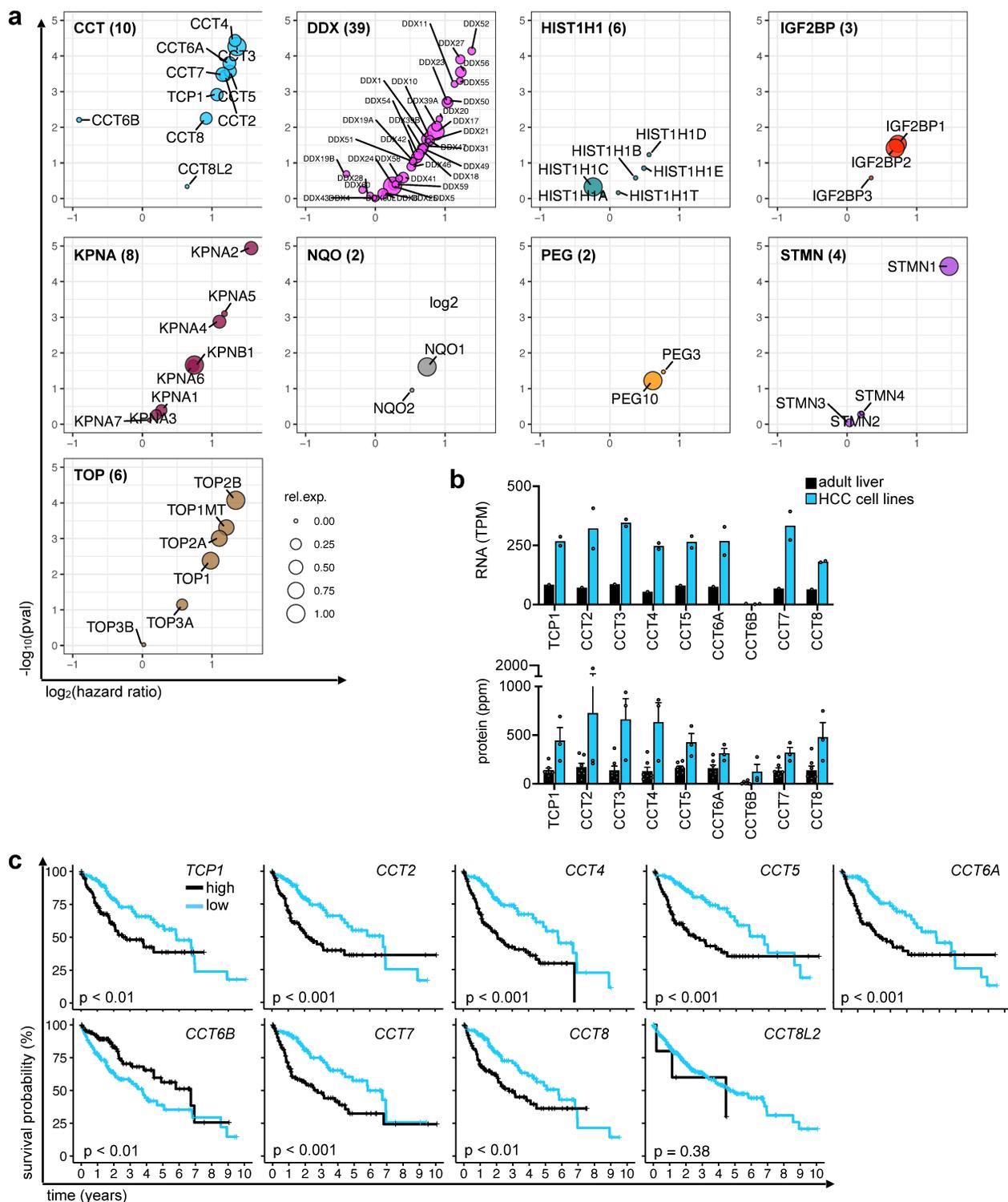
**a**, Boxplots of gene expression level of protein-coding genes grouped into canonical-RBP-genes (light-green), non-canonical-RBP-genes (dark-green), and non-RBP-genes (grey). Each column represents a different cancer type defined by The Cancer Genome Atlas (TCGA). **b-c**, Factorial map of the principal components (PC) analysis separates size factor and  $\log_2$  normalized gene expression levels of the (b) TCGA LIHC

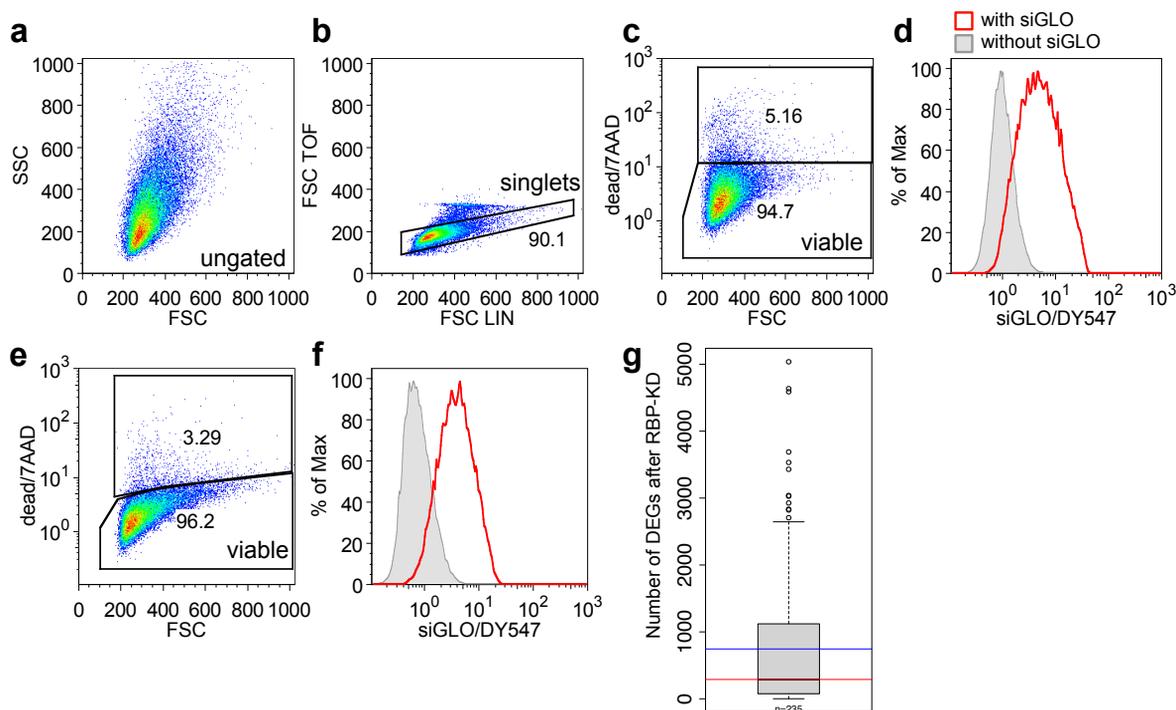
(50 patients) and (c) Australia HCC (24 patients) cohort into tumor (blue) and peritumor (red) patient samples. The proportion of variance explained by each principal component is indicated in parenthesis. **d**, Box plots represent the expression levels of the ten selected RBP genes in the FANTOM HCC cohort (50 patients) (Hashimoto *et al.*, Genome Res., 2015). Hinges correspond to the first and third quartiles, and whiskers correspond to the 1.5-times interquartile range. T: tumor; N: peritumor tissue. Statistics: paired two-tailed t-test. \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . **e**, Heatmap displays changes in expression level for commonly downregulated RBP genes in HCC as highlighted in (Fig. 1e). Expression level is sorted by average tumor z-score from right to left. Color gradient indicates z-score differences (green: high; grey: low). **f**, Table displays protein domains and family classification of the ten selected RBPs according to ExPASy (<https://prosite.expasy.org>). **g**, Microscopy images showing cellular location of immunofluorescently labeled RBPs (blue: DAPI staining of the nucleus, red: microtubules, green: RBP) in human cells (HepG2, A-431, U-2 OS, U-251) retrieved from the Human Protein Atlas (<https://www.proteinatlas.org>). **h**, Mutational status of the 10 selected RBP genes from three independent cohort studies (top panel, green: CGA, blue: INSERM, green: AMC) deposited in cBioPortal (<http://www.cbioportal.org>) shows proportional frequency of missense mutations (green), truncations (purple), amplifications (red), deletions (blue) and no alterations (grey). The percentage indicates the total number of alterations in all patients ( $n=865$ ).



**Fig. S2: Survival analysis of the TCGA-LIHC cohort.**

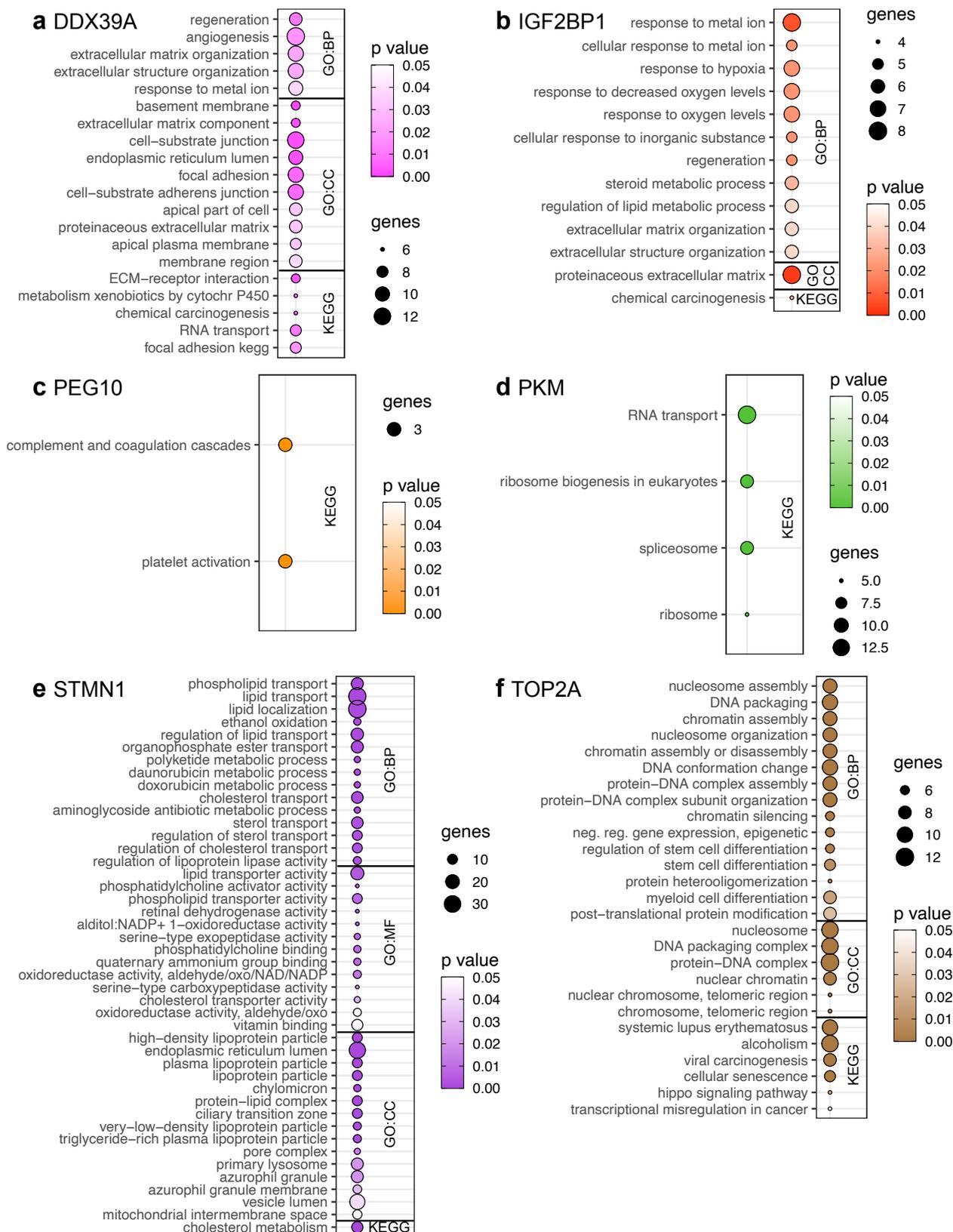
Kaplan-Meier plots show the association of RBP gene expression level and ten-year survival within the top and bottom tercile of 377 TCGA-LIHC (black: high and colored: low RBP gene expression levels, color code: DDX39A (magenta), HIST1H1C (turquoise), IGF2BP2 (red), *KPNA2* (plum), NQO1 (grey), PEG10 (yellow), PKM (green), STMN1 (purple) and TOP2A (brown)). Statistics: log-rank (Mantel-Cox).





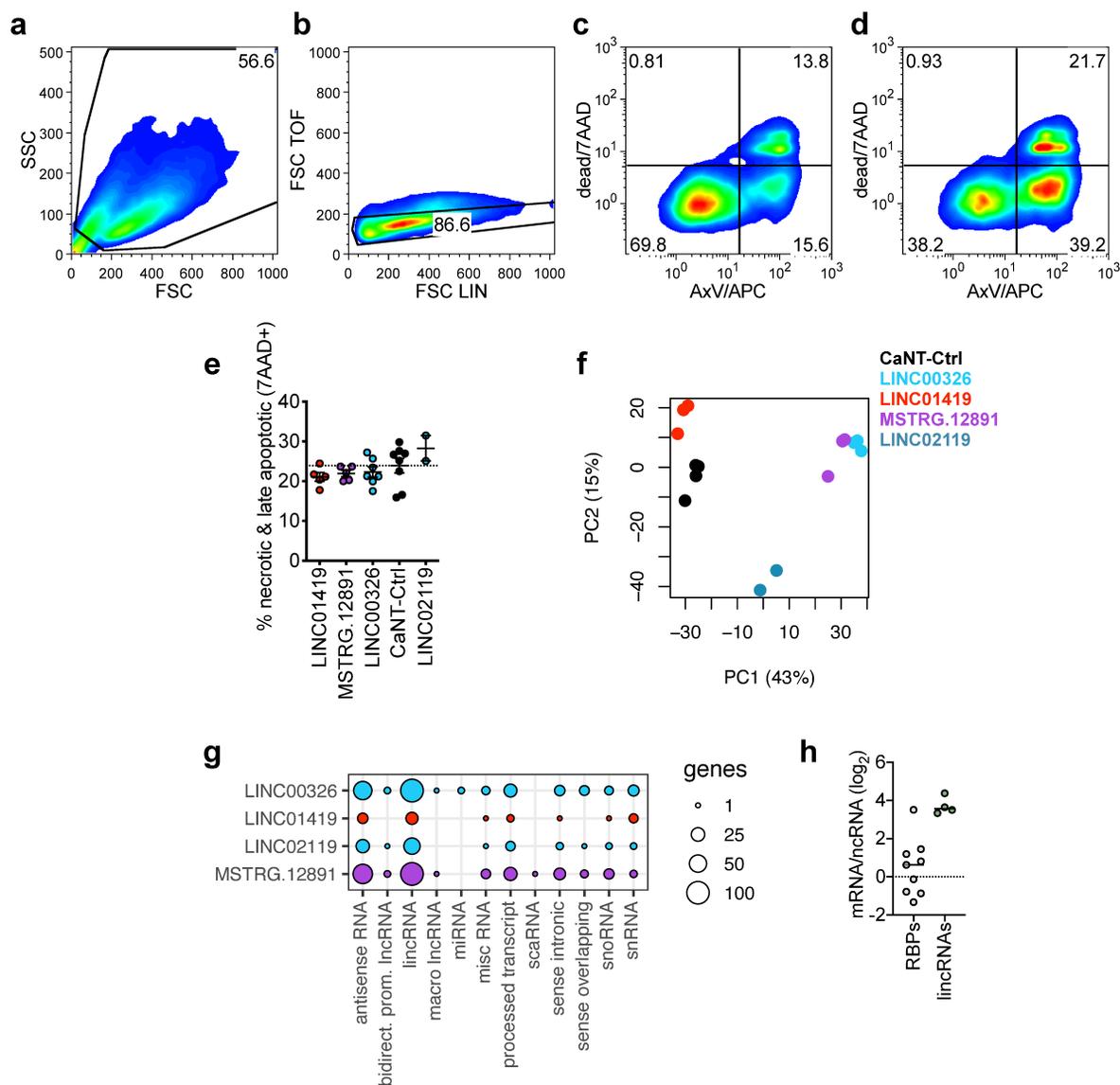
**Fig. S4: siRNA transfection efficiency and RBP-KD perturbation effects in human HCC cells**

**a-f,** Flow cytometry plots displaying gating strategy and transfection efficiency of siRNA (here siGLO) in **a-d** Huh7 and **e-f** HepG2 cells after electroporation. The completely shifted flow cytometry histogram indicates 100% transfection efficiency. Cell dead marker 7AAD was used to count dead and viable cells (shown as percentage). **g,** Boxplot showing the number of significantly differentially expressed genes (DEGs) upon RBP-KD in the ENCODE project database. Y axis represents number of DEGs; n represents the number of KD experiments; the box encloses all the data points between the upper and lower quartile; whiskers are drawn to the furthest data point falling within 1.5-times the interquartile range. Points represent outliers. The red line is median of the distribution and the blue line the mean (both median and mean are annotated with their respective values). Data is from (Van Nostrand *et al.*, Nature, 2020; Davis *et al.*, Nucleic Acids Res., 2018).



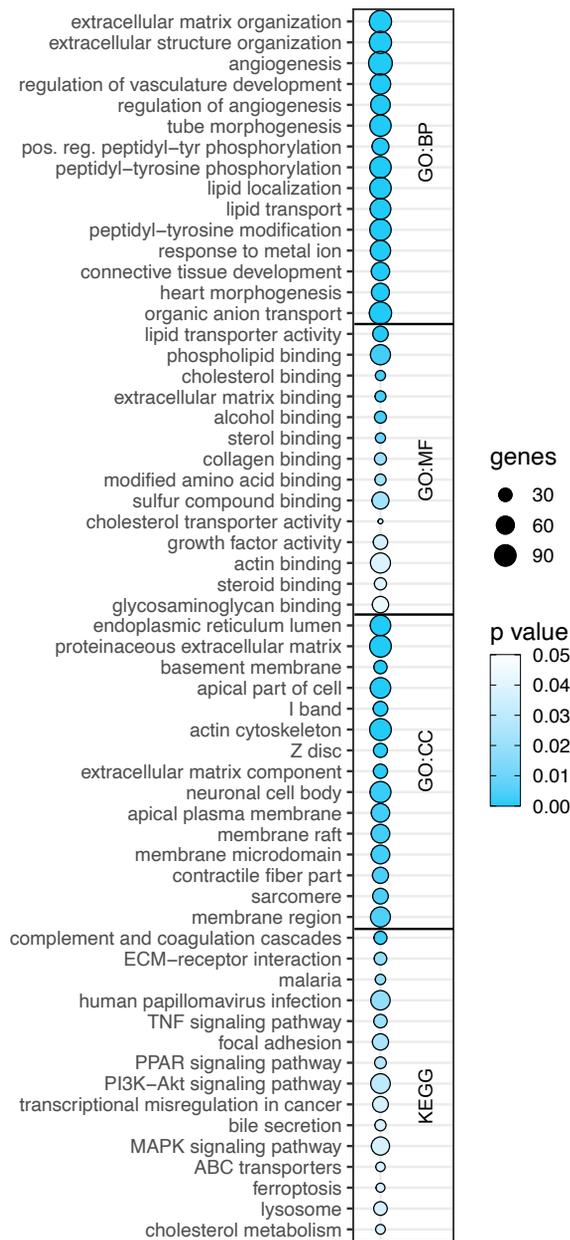
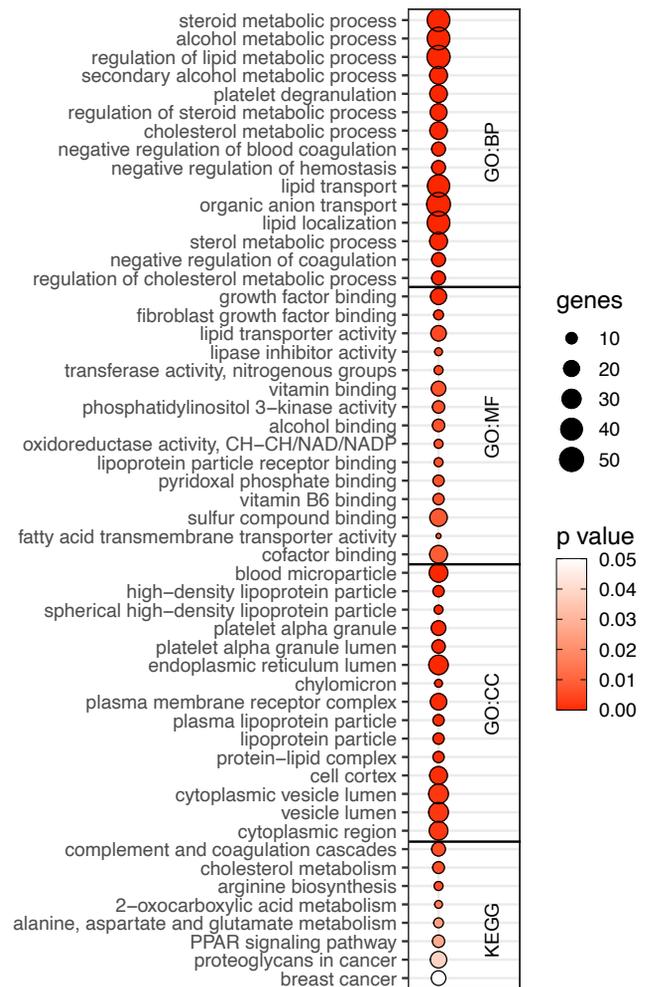
**Fig. S5: Gene Ontology (GO) and KEGG analyses of deregulated genes upon RBP-KD**

**a-f**, Circle plot shows GO term and KEGG pathway enrichment analysis of deregulated genes (FDR<0.05) after the siRNA-mediated KD of RBP **a**, DDX39A; **b**, IGF2BP1; **c**, PEG10; **d**, PKM; **e**, STMN1 and **f**, TOP2A. No significant pathways were found for genes deregulated upon KD of HIST1H1C, KPNA2, and NQO1. The diameter of the circles corresponds to the number of genes in each GO or KEGG term and the color code represents varying degrees of significance for the respective RBP-KD (white: high and colored: low adjusted p-value). BP, biological process; MF, molecular function; CC, cellular compartment.



**Fig. S6: Cellular and molecular phenotype assessment upon lincRNA-OE**

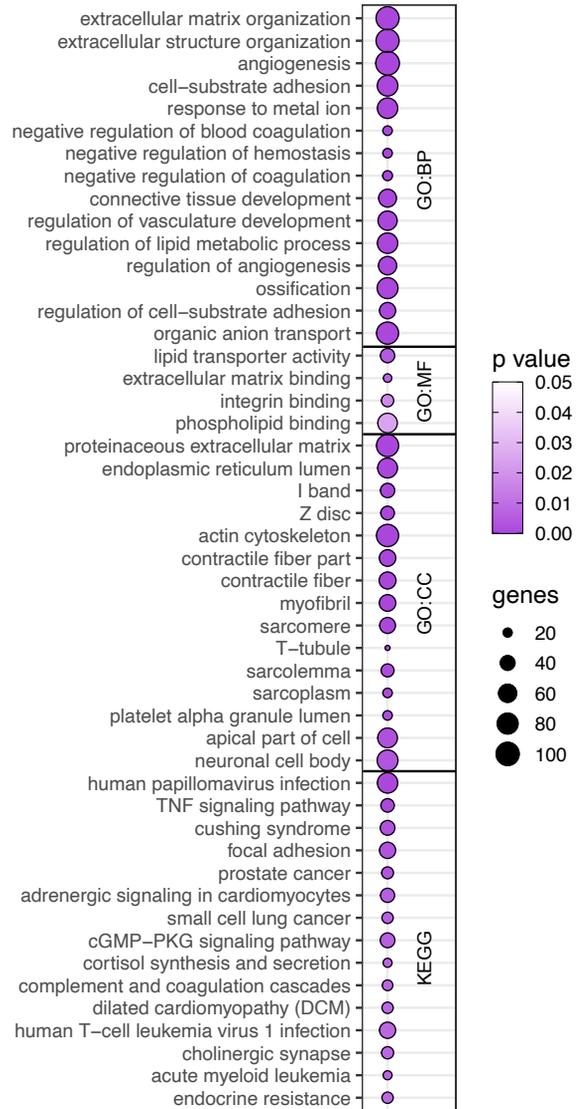
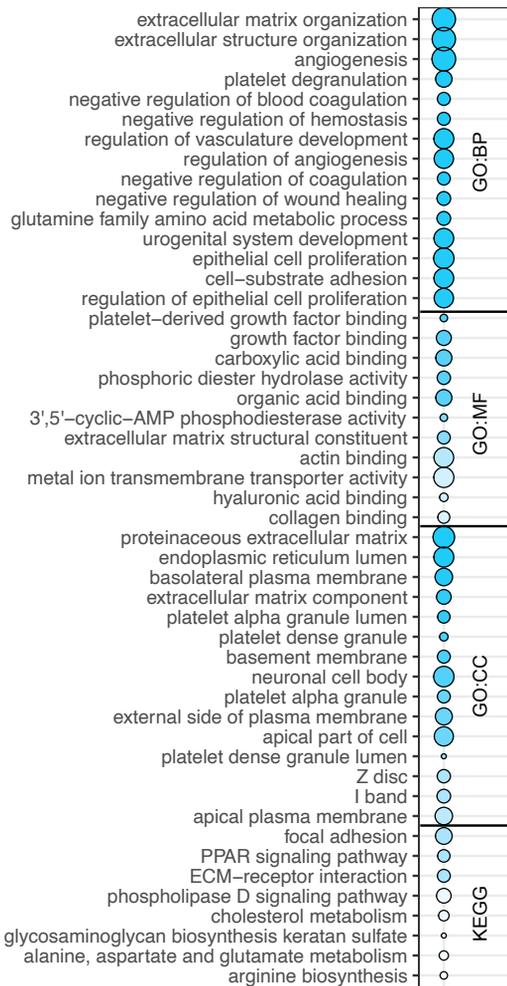
**a-d**, Flow cytometry gating strategy for assaying early apoptosis (Annexin V, AxV) and late apoptosis/necrosis (7AAD). **e**, dot plot displays the percentage of necrotic and late apoptotic (7AAD<sup>+</sup> cells assayed by flow cytometry) in Huh7 and HepG2 cells after CRISPRa-mediated lincRNA-OE after five days (n=2-8, mean, +/-SEM). The color-code links the lincRNA to the respective RBP-KD experiment through which it was identified (blue: CCT3, red: IGF2BP1, purple: STMN1). **f**, Factorial map of the principal components (PC) analysis separates size factor and log<sub>2</sub> normalized gene expression levels of the CRISPRa-mediated lincRNA-OE. The proportion of variance explained by each principal component is indicated in parenthesis. **g**, Circle plots display number of genes per non-coding RNA biotype affected by the CRISPRa-mediated lincRNA-OE. The diameter of the circles corresponds to the number of genes in each category. **h**, Dot plot of mRNA/lincRNA ratio after RBP-KD (n=10) or lincRNA-OE (n=4). The mean ratio is indicated by a black line.

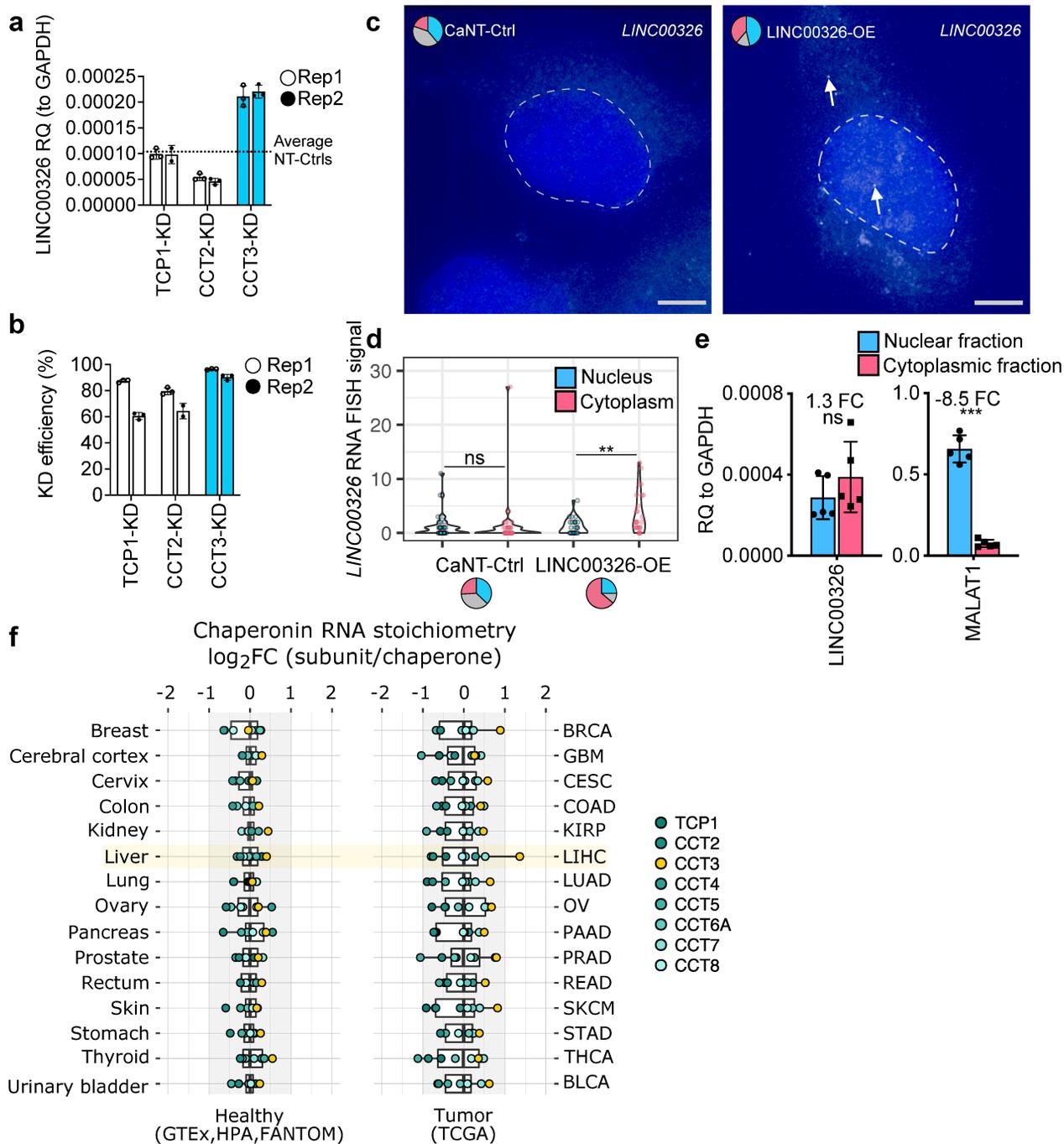
**a LINC00326****b LINC01419****Fig. S7: Gene Ontology (GO) and KEGG analyses of deregulated genes upon OE**

**a-d**, Circle plot shows GO term and KEGG pathway enrichment analysis of deregulated genes (FDR<0.01) after the CRISPRa-mediated lincRNA-OE **a**, *LINC00326*; **b**, *LINC01419*; **c**, *LINC02119* and **d**, *MSTR.12891*. The size of the circles indicates number of genes in each GO or KEGG term and the color represents varying degrees of significance for the respective lincRNA-OE experiment (white: high and colored: low adjusted p-value). The color-code links the lincRNA to the respective RBP-KD experiment through which it was identified (blue: CCT3, red: IGF2BP1, purple: STMN1). A maximum of 15 terms are shown in each GO or KEGG term (full list is available in Table S13). BP, biological process; MF, molecular function; CC, cellular compartment. Figure continues on the next page.

**c** LINC02119

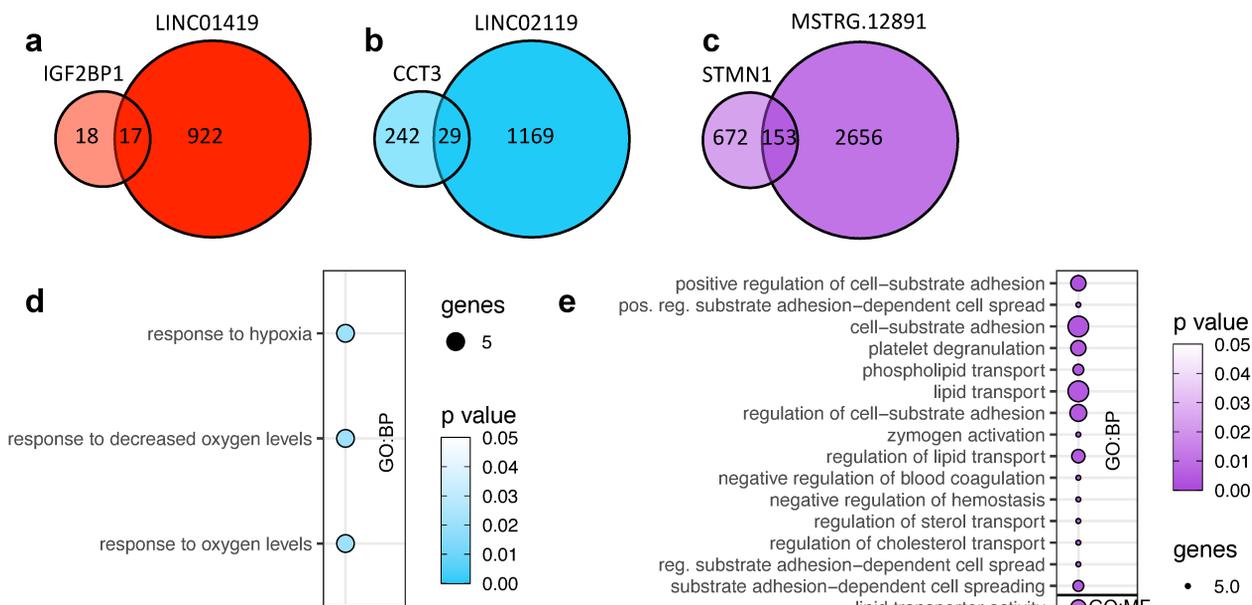
**d** MSTRG.12891





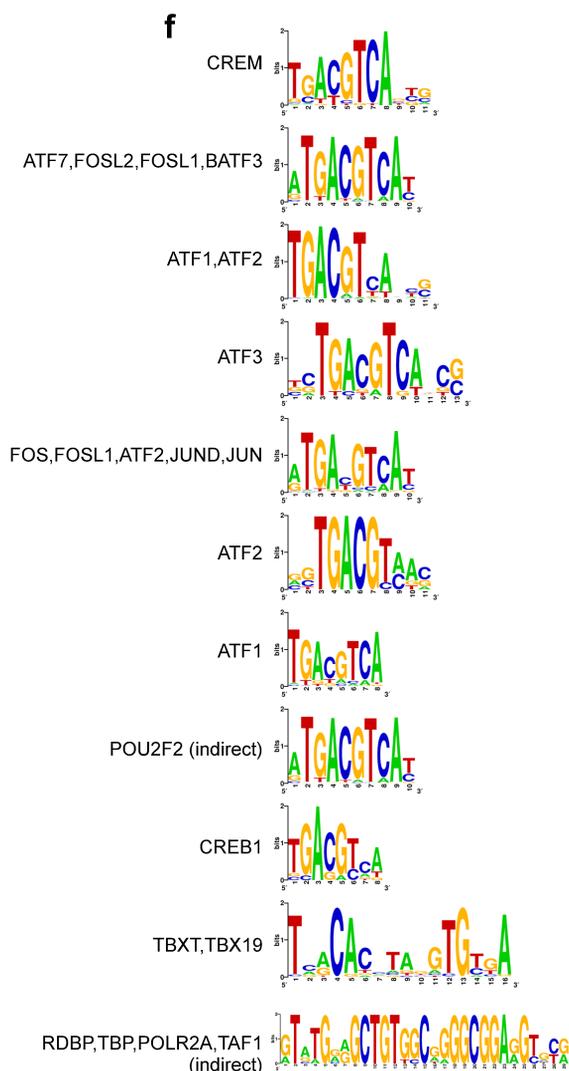
**Fig. S8: CCT3 works in a chaperoning-independent manner to interact with LINC00326.**

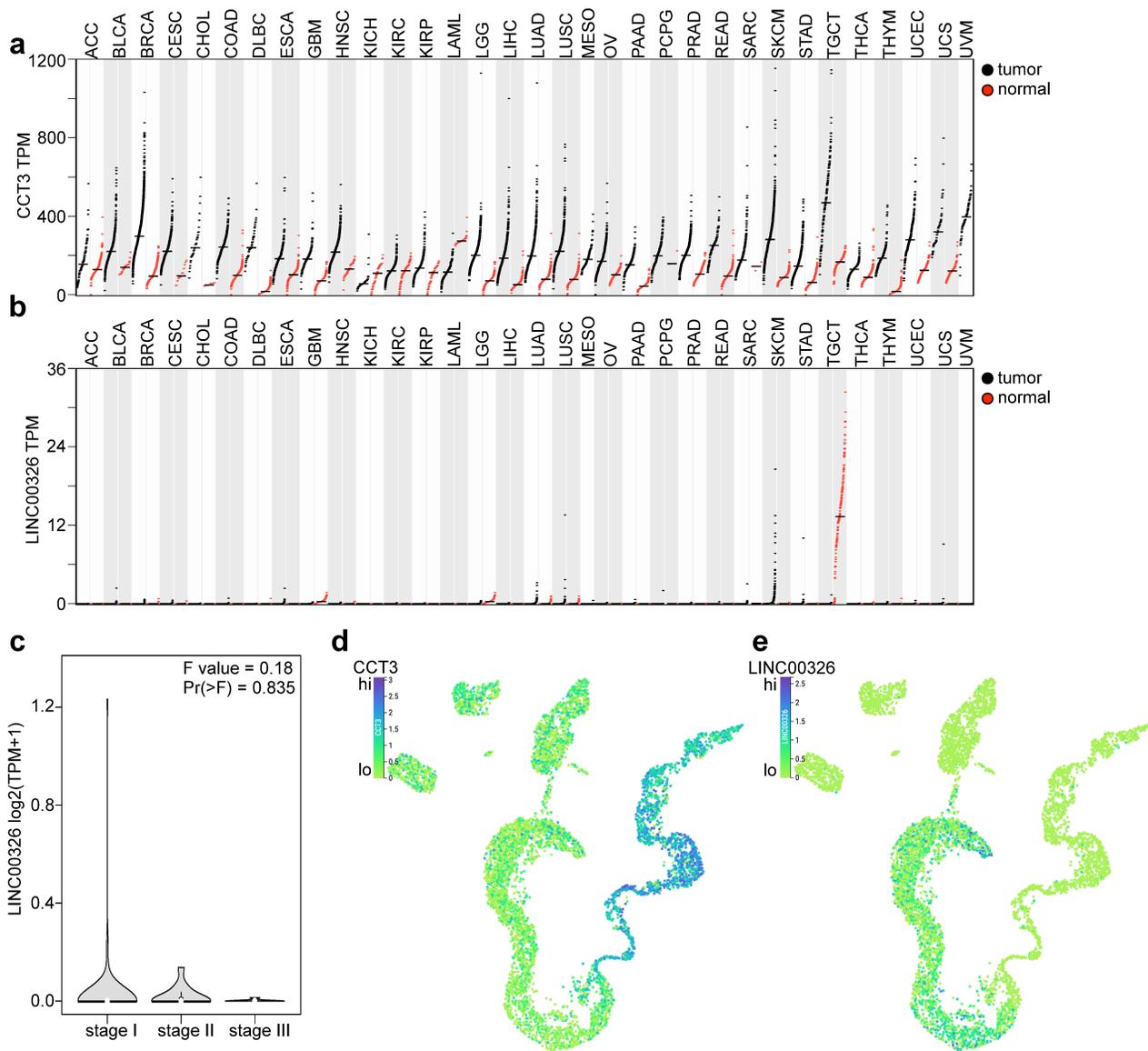
**a**, Bar graph shows the gene expression levels of *LINC00326* after CCT-family member KD. **b**, Bar graph shows KD-efficiency of samples in **a**. **c**, Microscopic images of single-molecule RNA FISH using exonic probes for *LINC00326* (white dots and arrows) in CaNT-Ctrl (left) and *LINC00326*-OE (right) Huh7 HCC cells. DAPI (blue) marks the nucleus. Pie chart represents the fraction of signals in the nucleus (blue) or cytoplasm (pink) in cells, or cells without any signal (grey). Scale bar: 5  $\mu$ m. **d**, Violin plots show quantification of *LINC00326* RNA FISH signal localization in HepG2 cells (n = 27-35). The pie chart represents the fraction of signal in the nucleus (blue) or cytoplasm (pink). Statistics: paired two-tailed t-test, \*\* p<0.01, ns: not significant. **e**, Bar graphs show gene expression patterns of *LINC00326* and *MALAT1* in separate cellular fractions of *LINC00326*-OE HCC cells. Quantification was performed by qPCR and gene expression values were normalized to *GAPDH*. Statistical significance and absolute fold change (FC) are displayed above the graphs. Statistics: paired two-tailed t-test, \*\*\* p<0.001. (n = 5). **f**, Box plots shows RNA stoichiometry data for subunits of the CCT chaperonin complex. The x axis value is computed as the fold change between the median gene expression level of the subunit and the median gene expression level of all subunits. The left boxes show healthy tissues and the right boxes show corresponding tissue tumors. Data were retrieved from the Human Protein Atlas (v. 20.1, <https://www.proteinatlas.org>).



**Fig. S9: Gene Ontology (GO) and KEGG analyses of deregulated genes shared upon RBP-KD and lincRNA-OE**

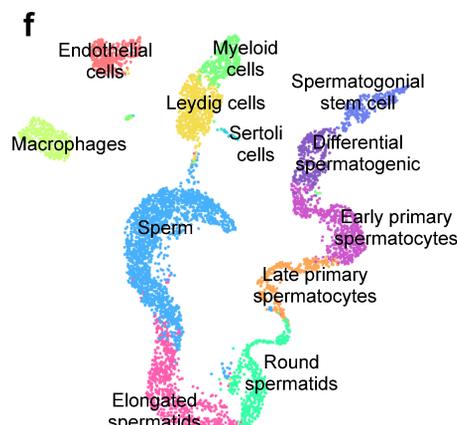
**a-c**, Two-way Venn diagrams intersect the number of deregulated genes in the RBP-KD and lincRNA-OE. **a**, *IGF2BP1*-KD and *LINC01419*-OE, **b**, *CCT3*-KD and *LINC2119*-OE, **c**, *STMN1*-KD and *MSTR.12891*-OE. **d-e**, GO and KEGG pathway analysis of overlapping genes from (b) *CCT3*-KD and *LINC2119*-OE and (c) *STMN1*-KD and *MSTR.12891*-OE. No significant pathways were found for genes deregulated upon *IGF2BP1*-KD and *LINC01419*-OE. The diameter of the circles corresponds to the number of genes in each GO or KEGG term and the color code represents varying degrees of significance for the respective RBP-KD and lincRNA-OE (white: high and colored: low adjusted p-value). BP, biological process; MF, molecular function; CC, cellular compartment. A maximum of 15 terms are shown in each category (full list is available in Table S13). **f**, Identified motifs for each TF in **Fig. 5f**.

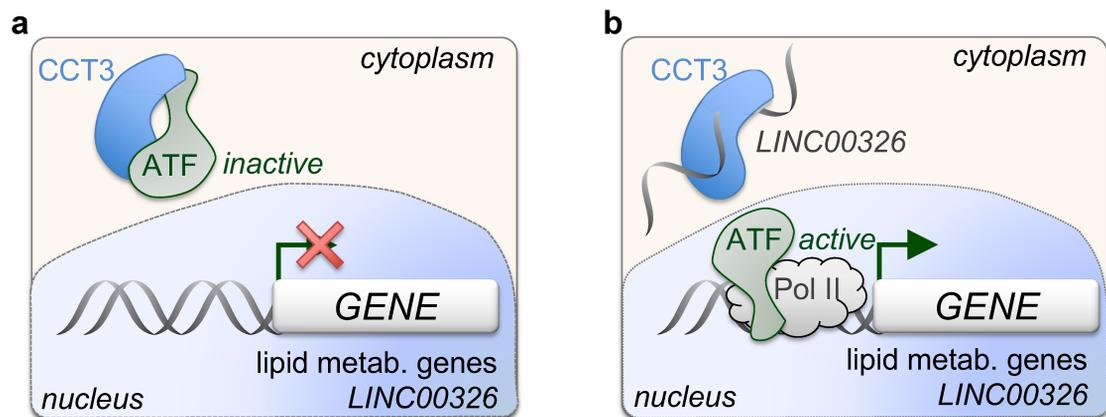




**Fig. S10: LINC00326 is highly expressed in healthy testis and diminished upon cancer progression**

**a-b**, Dot plot demonstrating the expression level in TPM of (a) CCT3 and (b) LINC00326 across different tissues from the TCGA and GTEx dataset. ACC (n=77-128 (tumor-normal)), BLCA (n=404-28), BRCA (n=1085-291), CESC (n=306-13), CHOL (n=36-9), COAD (n=275-349), DLBC (n=47-337), ESCA (n=182-286), GBM (n=163-207), HNSC (n=519-44), KICH (n=66-53), KIRC (n=523-100), KIRP (n=286-60), LAML (n=173-70), LGG (n=518-207), LIHC (n=369-160), LUAD (n=483-347), LUSC (n=486-338), MESO (n=87-0), OV (n=426-88), PAAD (n=179-171), PCPG (n=182-3), PRAD (n=492-152), READ (n=92-318), SARC (n=262-2), SKCM (n=460-558), STAD (n=408-211), TGCT (n=137-165), THCA (n=512-337), THYM (n=118-339), UCEC (n=174-91), UCS (n=57-78), UVM (n=79-0). **c**, LINC00326 expression in tumor biopsies from testis cancer patients divided by cancer stage. **d-e**, UMAP of single cell RNA expression of (d) CCT3 (e) LINC00326 in (f) human testis cells. a-c data are from: GEPIA (Tang *et al.* Nucleic Acids Res., 2017), and d-e are from Guo *et al.* Cell Res., 2018, visualised at covid19cellatlas.org (accessed on 21.3.21). ACC: adrenocortical carcinoma, BLCA: bladder urothelial carcinoma, BRCA: breast invasive carcinoma, CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma, CHOL: cholangiocarcinoma, COAD: colon adenocarcinoma, DLBC: lymphoid neoplasm diffuse large B-cell lymphoma, ESCA: esophageal carcinoma, GBM: glioblastoma multiforme, HNSC: head and neck squamous cell carcinoma, KICH: kidney chromophobe, KIRC: kidney renal clear cell carcinoma, KIRP: kidney renal papillary cell carcinoma, LAML: acute myeloid leukemia, LGG: brain lower grade glioma, LIHC: liver hepatocellular carcinoma, LUAD: lung adenocarcinoma, LUSC: lung squamous cell carcinoma, MESO: mesothelioma, OV: ovarian serous cystadenocarcinoma, PAAD: pancreatic adenocarcinoma, PCPG: pheochromocytoma and paraganglioma, PRAD: prostate adenocarcinoma, READ: rectum adenocarcinoma, SARC: sarcoma, SKCM: skin cutaneous melanoma, STAD: stomach adenocarcinoma, TGCT: testicular germ cell tumors, THCA: thyroid carcinoma, THYM: thymoma, UCEC: uterine corpus endometrial carcinoma, UCS: uterine carcinosarcoma, UVM: uveal melanoma.





**Fig. S11: Molecular model for LINC00326 regulating lipid metabolism-associated gene expression**

**a**, In HCC, CCT3 moonlights as an RBP, perhaps due to increased gene expression of CCT3. CCT3 binds to transcription factor ATF mRNA or forms a heteromeric protein complex. This causes that CCT3 sequester away ATF from initiating transcription of genes associated with lipid metabolism or *LINC00326*. **b**, when expressed, *LINC00326* RNA interacts with CCT3. This consequently limits the binding of CCT3 and ATF protein, thus releasing ATF and allowing transcription of lipid metabolism genes and *LINC00326* itself. A similar molecular mechanism occurs when CCT3 is reduced, thereby enabling expression of *LINC00326*.

## Supplementary Materials & Methods

### “The CCT3-LINC00326 axis regulates hepatocarcinogenic lipid metabolism”

Søndergaard *et al.*

---

#### ***Data and code availability***

Supplementary tables can be found here: <https://figshare.com/s/2c05765158269b3b4ff2>

The datasets generated in this study are available in the ArrayExpress repository, under accession numbers:

E-MTAB-8915

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8915>,

username: Reviewer\_E-MTAB-8915, password: xpizwqkc),

E-MTAB-9587

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9587>,

username: Reviewer\_E-MTAB-9587, password: 25ddT0yi), and

E-MTAB-9586

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9586>,

username: Reviewer\_E-MTAB-9586, password: M2tY9UPH). All scripts used for

bioinformatics analysis are available on Github: <https://github.com/jonasns/LiverNome>.

Single molecule RNA FISH imaging files generated in this study are available here:

Treated cells: <https://figshare.com/s/a83dbec52555e922ca8d>

Untreated cells: <https://figshare.com/s/08b0f84f2ea241b03c8d>

#### ***Identification of RBPs from prior studies***

The list of 2,282 proteins that can bind RNA in liver was generated by combining information from published datasets [1,2,11–19,3–10]. The list includes both canonical and non-canonical RBPs defined by whether a known RBD is present.

#### ***Cell culture***

All assay plates, dishes and culture flasks were purchased from Sarstedt. HepG2 and Huh7 human HCC cell lines were obtained from American Type Culture Collection (ATCC, Rockville, MD) with certified genotype. Huh7 cells constitutively expressing tubulin-GFP [20] were kindly provided by Prof. Bartenschlager (University of Heidelberg). All cell lines were mycoplasma free when periodically tested with MycoplasmaCheck (Eurofins Genomics). To ensure authenticity, cell lines were initially genotyped by short-tandem repeat genetic profiling (STR) using the Power-Plex\_16HS\_Cell Line panel and analyzed using Applied Biosystems Gene Mapper ID v3.2.1 software by the external provider Genetica DNA Laboratories (LabCorp Specialty Testing Group) and continuously assessed phenotypically. Huh7 and HepG2 cells were cultured in T75 flasks at 37°C and 5% CO<sub>2</sub> atmosphere using Dulbecco's Modified Eagle Medium (DMEM) supplemented with 1/100 Penicillin/Streptomycin (P/S, Sigma) and 10% fetal bovine serum (Hyclone, GE healthcare).

Huh7 cells were maintained by splitting 1/6 three times a week, HepG2 by splitting 1/4 three times a week. This was done by aspirating the medium, gently washing the cells with phosphate buffered saline (PBS) without  $Mg^{2+}$  (Sigma) and detaching them with 2mL of a trypsin-EDTA solution (Sigma) for 3-5min. Trypsin was inactivated with a minimum of 10-fold surplus of culture medium before a cell fraction was passaged. To ensure authenticity, cell lines were initially genotyped by short-tandem repeat genetic profiling (STR) using the PowerPlex\_16HS\_Cell Line panel and analyzed using Applied Biosystems Gene Mapper ID v3.2.1 software by the external provider Genetica DNA Laboratories (LabCorp Specialty Testing Group) and continuously assessed phenotypically.

### ***Patient material***

Tissue specimens were obtained under informed consent from 24 patients undergoing liver resection of HCC from the Australian Victorian Biobank, according to the approval of local ethics committee (application number: 2010/541-31/1 and 2017/719-31/2). The patients taking part in this study were 75% males and 25% females, and presented with HCC from Hepatitis B virus infection, non-alcoholic fatty liver disease, alcoholic steatohepatitis, hereditary haemochromatosis and other HCC triggering conditions. See patient information on ArrayExpress E-MTAB-8915 (reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8915>, username: Reviewer\_E-MTAB-8915, password: xpizwqkc). It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

### ***Survival analysis of TCGA data***

Survival analysis was done in R (v. 3.4.3) using survival (v. 3.1-8) and survminer (v. 0.4.6). The script is available at: <https://github.com/jonasns/LiveRNome>. A total of 377 patients were used for the analysis, and divided into terciles based on normalized htseq-counts of mRNA expression. Metadata and raw htseq-counts were acquired from <https://www.cancer.gov/tcga> (accessed on 2017.02.17).

### ***siRNA-mediated KD of RBPs***

At 70-80% confluency, cells were harvested by trypsination and subjected to electroporation. Small interfering RNA (siRNA) POOLS targeting *PEG10*, *HIST1H1C*, *KPNA2*, *PKM*, *NQO1*, *TOP2A*, *CCT3*, *STMN1*, *DDX39A*, and *IGF2BP1* was used (Thermo Scientific). siRNA ON-TARGETplus Non-Targeting siRNA#1 (Thermo Scientific) was used as control. Cells were washed twice in PBS and once in OptiMEM without phenol red (Invitrogen). A total of 5µg (18.8µL of a 20µM stock) siRNA was transferred to a 4mm cuvette (Bio-Rad), and 5 million cells were added in 200µl OptiMEM and incubated for 3min before being pulsed with an exponential decay pulse at 300V, 250µF, in a Genepulser II (Bio-Rad). Immediately after electroporation, the cells were transferred to pre-heated (37°C) phenol red-free DMEM culture medium supplemented with 10% (v/v) FBS and without the addition of antibiotics. Electroporation conditions were optimized using DY547-siGLO RISC-free control siRNA (Dharmacon) and flow cytometry (FACSNavios, Beckman Coulter, Navios Cytometry List Mode Data Acquisition and Analysis Software version 1.3). Optimization results were analyzed using FlowJo (v. 8.2). Gating strategy can be found in Fig. S4.

### ***RNA extraction and DNase-treatment***

All reagents and devices were purchased from ThermoFisher Scientific, if not stated otherwise. Upon cell harvest, 700µL Qiazol (Qiagen) was directly added onto the cells on ice and mixed. At this point, the cell extract was either stored at -80°C or directly added 140µL chloroform. This Qiazol/chloroform mixture was shaken for 30 sec and incubated at RT for 2.5min, before centrifugation at 9,000g for 5min at 4 °C. The mixture separated into a lower chloroform-phenol phase, an interphase and an upper, aqueous phase. The latter was carefully transferred to a new reaction tube before adding 1 volume isopropanol, inverting the tube 5 times followed by 10min incubation at RT. The mixture was centrifuged at 9,000g and 4°C for 10min and the supernatant discarded. The pellet was washed carefully using 700µL cold 70% ethanol, flicking the tube and centrifuged at 15,000g and 4°C for 5min. The supernatant was discarded entirely, and the pellet air-dried for 5min, before resuspension in 20–50µL nuclease-free water (NF-H<sub>2</sub>O, Ambion). RNA concentration was determined by nanodrop (Nanodrop 2000c). When possible, 10µg RNA was mixed with NF-H<sub>2</sub>O, 5µL TurboDNase buffer, 1µL TurboDNase and 1µL RNase Inhibitor (RiboLock, Invitrogen) in a total volume of 50µL and incubated at 37°C for 30min. Afterwards, 5µL Turbo DNase Inactivation reagent was added to the tube and incubated at RT for 5min. During incubation, the tube was occasionally flicked. The tube was centrifuged at 10,000g and 4°C for 2min and the supernatant carefully transferred to a new reaction tube. In order to achieve higher purity, the RNA was precipitated overnight. This was done by adding 19µL 3M sodium acetate (pH 5.2), 1µL GlycoBlue (Ambion), 135µL NF-H<sub>2</sub>O and 600µL ice-cold 99.8% ethanol, homogenizing and storing the tube at -80°C. The next day, the sample was centrifuged at 12,000g for 30min, the supernatant discarded, and the pellet washed twice with 700µL 70% ethanol. The pellet was air-dried, resuspended in 25µL NF-H<sub>2</sub>O and the RNA concentration determined by nanodrop.

For the Australian HCC cohort, total RNA from the liver biopsies was extracted using the RNeasy mini kit (Qiagen) and treated with DNase RNase free kit (Qiagen).

### ***cDNA synthesis and qPCR***

All reagents and devices were purchased from ThermoFisher Scientific, if not stated otherwise. 1-2µg of pure DNase-treated RNA was added 1µL random primers (250ng/µL) and 1µL dNTP Mix (10m each) to a final volume of 13.5µL. The sample was incubated in a thermocycler (ProFlex PCR) at 65°C for 5min. Afterwards, 7µL of a mastermix containing 1µL RNase-Inhibitor, 4µL First-Strand Buffer, 1µL 0.1M DTT and 0.5µL SuperScript II reverse transcriptase were added. Samples were incubated in the thermocycler at 25°C for 10min, 42°C for 50min and 70°C for 15min and a final hold at 4°C.

The cDNA used for qPCR was diluted 1/10 to 1/25 depending on the experiment. 3.6X Master mixes were prepared, containing 3.6µL of primer mix (forward and reverse primer, 2.5µM each), 10.8µL NF-H<sub>2</sub>O and 18µL SYBR-Green Mix (PowerUp) per sample. 3.6µL of diluted cDNA was added to a 96-well PCR plate and 32.4µL of the prepared master mix added. After mixing, three times 10µL were transferred to a 384-well plate to obtain three technical replicates per sample. After sealing the plate and a short centrifugation, it was run on Real-Time-PCR machine (QuantStudio 5 or CFX384 (Bio-Rad)). The following

parameters were set: initial heating steps of 3min at 50°C followed by 3min at 95°C. 40 cycles of 15sec at 95°C and 30sec at 60°C. Melt curve: 60-95°C using 0.5°C increments every 5sec. Quantitative PCR data was checked for correct amplification and dissociation of the products. mRNA levels of the genes of interest were normalized to mRNA levels of the housekeeping genes *GAPDH* or *ACTB* and were calculated according to the ddCT method.

### **Primer design**

Exon sequences of desired genes were obtained from NCBI Reference Sequence Database (RefSeq) and used as input on [www.primer3.ut.ee](http://www.primer3.ut.ee) (version 4.1.0). The following parameters were set: T<sub>m</sub> 60-65°C with optimum of 60°C, GC-content 40-60% with optimum of 50% and product size range from 100-300 bp. Primer results were validated using the in-silico PCR tool in UCSC Genome Browser (Human, Assembly: Dec. 2013 – GRCh38/hg38, target: genome assembly). Whenever possible, the primers were designed to be intron-spanning. Primers used can be found in Table S14.

### **MTT proliferation assay**

Directly after electroporation, 5,000 Huh7 or 7,000 HepG2 cells were seeded into 96 well plates in triplicates. Cells were assayed between day 1 and day 7 every 48h. For this purpose, the medium was aspirated and a mixture of 10µL MTT (Methylthiazolyldiphenyl-tetrazolium bromide, Sigma, 4mg/mL in 1X PBS) and 60µL culture medium added to the cells, before incubating at 37°C and 5% CO<sub>2</sub> atmosphere for 1 h. After carefully removing the medium, 100µL of MTT-lysis buffer per well (for 100mL MTT lysis buffer: 90mL isopropanol, 2.5mL 10% SDS, 4mL HCl, 3.5mL H<sub>2</sub>O) was added onto the cells and incubated on an orbital shaker for 15-30min. The lysate was resuspended, the bottom of the assay plate wiped with ethanol and the absorption at 595nm measured on a plate reader (Spectramax i3x, Molecular Devices). Background values were subtracted from the obtained values before normalization to the average value of the plate in order to compare between different experiments.

### **Viability measures**

Supernatants containing non-adherent cells and adherent cells (harvested by trypsination) were harvested directly into FACS tubes. For 7-AAD single staining, cells were washed 1X in PBS with 1% BSA (Sigma), followed by staining in 100µL buffer with 5µL 7-AAD Viability Staining Solution (eBioscience) on ice in the dark for 15min. Cells were acquired directly on a flow cytometer (FACSNavios, Beckman Coulter) without washing away the staining buffer. For detection of early apoptosis, cells were washed twice in PBS (without BSA), and resuspended in 100µL Annexin V binding buffer (10mM HEPES (pH 7.4), 150mM NaCl, 2.5mM CaCl<sub>2</sub> in H<sub>2</sub>O) containing 5µL 7AAD and 5µL Annexin V-APC (eBioscience). Cells were stained on ice in the dark for 30min and acquired as for 7-AAD single staining. Data were acquired by flow cytometry (FACSNavios, Beckman Coulter, Navios Cytometry List Mode Data Acquisition and Analysis Software version 1.3) and analyzed using FlowJo (v. 8.2). Gating strategy can be found in Fig. S6.

### **Sequencing of RNA libraries**

The library preparation was carried out using the Illumina TruSeq® Stranded RNA Library Prep Kit v2 (dual index) according to the manufacturer's instruction and ribosomal depletion was achieved with Ribo-Zero Gold. The quality of every cDNA library was determined on high-sensitivity DNA chips (Agilent Bioanalyzer) according to the manufacturer's protocol. Libraries were quantified with the KAPA-SYBR FAST qPCR kit (Roche). The sequencing run was performed either with the NextSeq 500/550 High Output v2 kit (Illumina) for 150 cycles, paired end, on a Illumina NextSeq 500 or HiSeq Rapid SBS Kit v2 for 50 cycles, single end on a Illumina HiSeq2500. All raw data (fastq files) are accessible under

ArrayExpress:

E-MTAB-9587

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9587>, username: Reviewer\_E-MTAB-9587, password: 25ddT0yi) and

E-MTAB-9586

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9586>, username: Reviewer\_E-MTAB-9586, password: M2tY9UPH)

### ***RNA-seq data QC and processing***

Next generation sequencing read quality were assessed with FastQC (v. 0.11.5) and RseQC (v. 2.6.4). Adaptor sequences were trimmed and low-quality reads removed using Trimmomatic (v. 0.36). All sequencing reads aligning (HiSAT2, v. 2.1.0) to annotated ribosomal RNA genes were discarded. High-quality and ribosomal RNA depleted sequencing reads were aligned to the genome using HiSAT2. Novel RNA transcripts were identified using StringTie (v. 1.3.3). Using sorted bam files (Samtools v. 1.5), the number of aligned reads were counted (featurecount in subread package v. 1.5.2) for both annotated transcripts and novel references generated with StringTie. After normalization (TMM: trimmed mean of M-values), a differential gene expression analysis (edgeR v. 3.20.7) was performed in R (v. 3.4.3). For the two patient cohorts (TCGA and Australia), DESeq (v. 1.14.1 in bioconductor v. 3.4) was used for differential gene expression analysis. In general, we use edgeR for  $n < 6$  and DESeq for  $n > 6$  replicates. Significant DE genes was distinguished by a false discovery rate (FDR) under 0.05 or 0.01. Gene ontology analysis was performed in R (v. 3.4.3) with clusterProfiler (v. 3.6.0) and org.Hs.eg.db (v. 3.5.0), with p values corrected by the Benjamini-Hochberg FDR procedure. Additionally, the following dependent package versions were installed: DOSE (v. 3.4.0), AnnotationDbi (v. 1.40.0), IRanges (v. 2.12.0), S4Vectors (v. 0.16.0), BiocGenerics (v. 0.24.0), and Biobase (v. 2.38.0). All scripts are available on Github: <https://github.com/jonasns/LiveRNome>.

### ***CRISPRa vector generation and cell electroporation***

SP-dCas9-VPR [21] were acquired from Addgene plasmid # 63798 (deposited by George Church). Three gRNAs targeting within the 300 bp upstream region of the transcriptional start site were designed for each lincRNA using the Broad Institute design tool (<https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design>, accessed on 2018.01.08). Each gRNA was checked for off-target effects using both the Broad Institute tool's internal checkup and by BLATing them in UCSC genome browser. gRNAs were assembled into SP-dCas9-VPR using Gibson assembly (NEB). Non-targeting gRNAs were

from [22] and is not found anywhere in the human genome. The gRNA sequences can be found in Table S14. HepG2 and Huh7 cells were electroporated using the NEON electroporation system (Invitrogen) as previously described [20]. Transfection efficiency was checked by qPCR of target genes.

### ***Lipid experiments***

Assays for lipid peroxidation (MDA) (Abcam, ab233471), DCFDA cellular ROS detection (Abcam, ab113851) and hepatic lipid accumulation/steatosis (ORO, Abcam, ab133131) were conducted according to the manufacturers' protocols. For the MDA assay, cells were lysed directly in the culture dish with a modified 4°C cold RIPA buffer (10mM Tris-HCl pH 8.0, Sigma), 150mM NaCl (Sigma), 1mM Na<sub>2</sub>EDTA (Fisher), 1mM EGTA (VWR), 1% NP-40 (Sigma), 1% sodium deoxycholate (VWR), 0.1% SDS (Sigma). Lysates were kept on ice unless otherwise stated. The MDA assay was performed 48h after transfection, while for the ROS and ORO assays, cells were collected, counted and re-plated in equal numbers 48h after transfection, followed by readout the day after. The ROS assay was performed in 96wp, while ORO was performed in 24wp, until after dye extraction, which was transferred to 96wp. All experiments were read out on a Spectramax i3x.

### ***Zebrafish xenograft experiments***

Zebrafish experimentation was performed at the zebrafish core facility at Comparative Medicine, Karolinska Institute, Sweden. Husbandry and breeding were in accordance to the ethical permit Dnr 14049-2019 approved by the Stockholm North Ethical Board. Zebrafish experimentation in embryos younger than five days are excluded from the normative on animal testing by the EU directive 2010/63/EU.

Huh7 cells constitutively expressing TUBULIN-GFP [23] were cultured identical to standard Huh7 cells, except for adding 10ug/mL Blasticidin S HCl (Thermo Fisher) to the growth medium. TUBULIN-GFP Huh7 cells were grown to 80% confluency in a T75 flask. Medium was removed and cells were washed twice with pre-warmed 1xPBS. Five mL of pre-warmed 1xPBS was added to the cell culture followed by 20µL of Vybrant DiICell-labeling solution (Thermo Fisher). After incubating for 20min at 37°C excessive dye was washed away with 1xPBS and cells were harvested using Accutase (Sigma). Accutase was inactivated with 10-fold excess culture medium and filtered through a pre-washed 40µM cell strainer (Sigma). After pelleting (500g, 4min) cells were resuspended in 1mL of 2% (w/v) polyvinylpyrrolidone (Sigma) in PBS and transported on ice to the zebrafish core facility. Cells were re-pelleted, resuspended by flicking the tube, and transplanted into the perivitelline space of 48h old zebrafish embryos (AB genotype). Tumor growth was followed by daily by epifluorescence microscopy (Leica) from day 1 to 5 after injection. Tumor area was quantified using Fiji in ImageJ (v. 1.52).

### ***Promoter analysis***

Promoter analysis was done using the gimmeroc command in GimmeMotifs v0.13.1 with fasta sequences 400bp upstream of DEG transcription start site as input and 1000 random sequences with equal length and GC content from hg38 as background.

***RNA immunoprecipitation followed by quantitative PCR (RIP-qPCR)***

RIP-qPCR was done using the Magna RIP RBP IP kit (Sigma) according to the manufacturer's instructions. Cell lysates were prepared from Huh7 or HepG2 HCC cell lines with *LINC00326*-OE, and immunoprecipitated using CCT3 Ab 10571-1-AP (Proteintech, lot #00021192) rotating overnight at 4°C. HULC was used as a negative control.

***Subcellular fractionation***

Subcellular fractions were prepared as described in [17]. Cultured cells at an approximate confluence of 90 % were harvested, counted, washed with cold PBS and  $\sim 0.5 \times 10^7$  cells were transferred to a reaction tube. Cells were centrifuged at  $500 \times g$  and 4°C for 5 min, the supernatant discarded, and the cells very gently resuspended in 380  $\mu$ L hypotonic lysis buffer (HLB) containing 100 U RNase Inhibitor. After 10 min incubation on ice, the lysate was centrifuged at  $1000 \times g$  and 4°C for 3 min. The supernatant containing the cytoplasmic fraction was carefully transferred to a new reaction tube, while leaving the pellet on ice. 1 mL RNA precipitation solution was directly added to the supernatant, and stored at -20°C for at least 1 h. The pellet containing the nuclear fraction was washed three times by addition of 1 mL HLB and centrifugation at  $200 \times g$  at 4°C for 2 min. Subsequently, the pellet was dissolved in 1 mL Qiazol and kept on ice for short- or at -80°C for long-term storage. After 1 h of incubation at -20°C, the cytosolic fraction was vortexed for 30 sec and centrifuged at  $18,000 \times g$  and 4°C for 15 min. After discarding the supernatant, the pellet was washed with ice-cold 70% ethanol, centrifuged at  $18,000 \times g$  and 4°C for 5 min and air dried. Subsequently, the pellet was resuspended in 1 mL Qiazol and mixed until the pellet was fully dissolved. RNA from the cytosolic and nuclear fractions was extracted and the reverse transcription followed by quantitative PCR was performed.

***Single-molecule RNA FISH***

Cells were grown on coverslips in 12-well plates and were left untreated or were transfected with *LINC00326* or appropriate negative control guides with CRISPRa. The cells were briefly washed with 1xPBS (diluted in RNase-free H<sub>2</sub>O) and fixed with 75% methanol + 25% acetic acid at RT for 10 min. Following fixation, cells were washed twice with 1xPBS. The coverslips were briefly rehydrated with Wash Buffer A (SMF-WA1-60, Biosearch) (Formamide, Thermo Fischer Scientific, AM9342) before FISH. The Stellaris FISH Probes (*LINC00326* exonic probes, Q570) were designed using the Stellaris online RNA FISH probe designer (Stellaris Probe Designer, version 4.2), which was produced by LGC Biosearch Technologies. The *LINC00326* probes were added to the hybridization buffer (SMF-HB1-10, LGC Biosearch Technologies) at a final concentration of 250 nM, and hybridization was carried out in a humidified chamber at 37 °C overnight. The following day, the cells were washed twice with Wash Buffer A (SMF-WA1-60, LGC Biosearch Technologies) at 37 °C for 30 min. The second wash contained 4,6-diamidino-2-phenylindole for nuclear staining (5  $\mu$ g/ml, Merck-Sigma, D9542). The coverslips were then briefly washed with wash buffer B (SMF-WB1-20, LGC Biosearch Technologies) for 5 min at RT, equilibrated 5 min in base glucose buffer (2x SSC, 0.4% glucose solution 49163, Merck-Sigma, in RNase-free H<sub>2</sub>O), and then incubated 5 min in Base Glucose buffer supplemented with 1:100 dilution of glucose oxidase (stock 3.7 mg/mL, G2133, Merck-Sigma) and catalase (stock 4 mg/mL,

219261-100KU, Merck-Sigma). Finally, the coverslips were mounted with ProlongGold Antifade Mountant (P36934, Thermo Fischer Scientific). Images were captured with wide-field DeltaVision microscope (Applied Precision, LLC) equipped with a Coolsnap HQ2 12 bit camera with 1x1 binning and 896x896 frame size. We acquired 20-30 optical slices depending on the thickness of the cell lines at 0.3- $\mu$ m intervals using Olympus 100x (1.4 numerical aperture) oil immersion objectives. Deconvolution of three-dimensional stacks was performed with built-in DeltaVision SoftWoRx Imaging software (Applied Precision, LLC). Maximum intensity projections were generated in Fiji and subjected for manual quantification using Fiji. The sequences of LINC00326 RNA FISH probe can be found in Table S14.

### Supplementary Materials & Methods References

- 1 Castello A, Fischer B, Eichelbaum K, *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 2012;**149**:1393–406. doi:10.1016/j.cell.2012.04.031
- 2 Castello A, Fischer B, Frese CK, *et al.* Comprehensive Identification of RNA-Binding domains in human cells. *Mol Cell* 2016;**63**:696–710. doi:10.1016/j.molcel.2016.06.029
- 3 Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet* 2014;**15**:829–45. doi:10.1038/nrg3813
- 4 Baltz AG, Munschauer M, Schwanhäusser B, *et al.* The mRNA-bound proteome and Its global occupancy profile on protein-coding transcripts. *Mol Cell* 2012;**46**:674–90. doi:10.1016/j.molcel.2012.05.021
- 5 Caudron-Herger M, Rusin SF, Adamo ME, *et al.* R-DeeP: proteome-wide and quantitative identification of RNA-dependent proteins by density gradient ultracentrifugation. *Mol Cell* 2019;**75**:184-199.e10. doi:10.1016/j.molcel.2019.04.018
- 6 Bao X, Guo X, Yin M, *et al.* Capturing the interactome of newly transcribed RNA. *Nat Methods* 2018;**15**:213–20. doi:10.1038/nmeth.4595
- 7 Perez-Perri JI, Rogell B, Schwarzl T, *et al.* Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat Commun* 2018;**9**. doi:10.1038/s41467-018-06557-8
- 8 Mullari M, Lyon D, Jensen LJ, *et al.* Specifying RNA-Binding Regions in Proteins by Peptide Cross-Linking and Affinity Purification. *J Proteome Res* 2017;**16**:2762–72. doi:10.1021/acs.jproteome.7b00042
- 9 Trendel J, Schwarzl T, Horos R, *et al.* The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell* 2019;**176**:391-403.e19. doi:10.1016/j.cell.2018.11.004
- 10 Brannan KW, Jin W, Huelga SC, *et al.* SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell* 2016;**64**:282–93. doi:10.1016/j.molcel.2016.09.003
- 11 Treiber T, Treiber N, Plessmann U, *et al.* A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Mol Cell* 2017;**66**:270-284.e13. doi:10.1016/j.molcel.2017.03.014
- 12 He C, Sidoli S, Warneford-Thomson R, *et al.* High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Mol Cell* 2016;**64**:416–30. doi:10.1016/j.molcel.2016.09.034
- 13 Conrad T, Albrecht AS, De Melo Costa VR, *et al.* Serial interactome capture of the human cell nucleus. *Nat Commun* 2016;**7**. doi:10.1038/ncomms11212

- 14 Beckmann BM, Horos R, Fischer B, *et al.* The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun* 2015;**6**. doi:10.1038/ncomms10127
- 15 Sundararaman B, Zhan L, Blue SM, *et al.* Resources for the Comprehensive Discovery of Functional RNA Elements. *Mol Cell* 2016;**61**:903–13. doi:10.1016/j.molcel.2016.02.012
- 16 Hendrickson D, Kelley DR, Tenen D, *et al.* Widespread RNA binding by chromatin-associated proteins. *Genome Biol* 2016;**17**. doi:10.1186/s13059-016-0878-3
- 17 Cook KB, Kazan H, Zuberi K, *et al.* RBPDB: A database of RNA-binding specificities. *Nucleic Acids Res* 2011;**39**. doi:10.1093/nar/gkq1069
- 18 Ray D, Kazan H, Cook KB, *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 2013;**499**:172–7. doi:10.1038/nature12311
- 19 Kwon SC, Yi H, Eichelbaum K, *et al.* The RNA-binding protein repertoire of embryonic stem cells. *Nat Struct Mol Biol* 2013;**20**:1122–30. doi:10.1038/nsmb.2638
- 20 Søndergaard JN, Geng K, Sommerauer C, *et al.* Successful delivery of large-size CRISPR/Cas9 vectors in hard-to-transfect human cells using small plasmids. *Commun Biol* 2020;**3**:1–6. doi:10.1038/s42003-020-1045-7
- 21 Chavez A, Scheiman J, Vora S, *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* 2015;**12**:326–8. doi:10.1038/nmeth.3312
- 22 Liu SJ, Horlbeck MA, Cho SW, *et al.* CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* 2017;**355**:eaah7111. doi:10.1126/science.aah7111
- 23 Cortese M, Goellner S, Acosta EG, *et al.* Ultrastructural Characterization of Zika Virus Replication Factories. *Cell Rep* 2017;**18**. doi:10.1016/j.celrep.2017.02.014

## Supplemental Material

### “The CCT3-*LINC00326* axis regulates hepatocarcinogenic lipid metabolism”

Søndergaard *et al.*

---

#### SUPPLEMENTARY FIGURES

Figure S1. Molecular characteristics of RBPs in HCC

Figure S2. Survival analysis of the TCGA-LIHC cohort

Figure S3. Estimation of hazard ratio and HCC patient survival based on RBP gene expression

Figure S4. siRNA transfection efficiency and RBP-KD perturbation effects in human HCC cells

Figure S5. Gene Ontology (GO) and KEGG analyses of deregulated genes upon RBP-KD

Figure S6. Cellular and molecular phenotype assessment upon lincRNA-OE

Figure S7. GO and KEGG analyses of deregulated genes upon OE

Figure S8. CCT3 works in a chaperoning-independent manner to interact with *LINC00326*

Figure S9. GO and KEGG analyses of deregulated genes shared upon RBP-KD and lincRNA OE

Figure S10. *LINC00326* is highly expressed in healthy testis and diminished upon cancer progression

Figure S11. Molecular model for *LINC00326* regulating lipid metabolism-associated gene expression

#### SUPPLEMENTARY TABLES

Table S1. hRBP list

Table S2. TCGA LIHC DEG

Table S3. AUS counts

Table S4. AUS DEG

Table S5. RBP KD RNAseq raw counts

Table S6. RBP KD RNAseq TPM

Table S7. RBP KD RNAseq DEG

Table S8. RBP KD RNAseq GO KEGG

Table S9. novel genes info

Table S10. lincRNA OE RNAseq raw counts

Table S11. lincRNA OE RNAseq TPM

Table S12. lincRNA OE RNAseq DEG

Table S13. lincRNA OE and RBP overlap GO KEGG

Table S14. Primer list

Supplementary tables and microscopic imaging files are accessible via Figshare:

<https://figshare.com/s/2c05765158269b3b4ff2>

<https://figshare.com/s/a83dbec52555e922ca8d>

<https://figshare.com/s/08b0f84f2ea241b03c8d>

#### SUPPLEMENTARY MATERIALS & METHODS

Cell-based and molecular assays as well as xenograft experiments are described in the supplementary materials and methods.

#### ARRAYEXPRESS ACCESSION

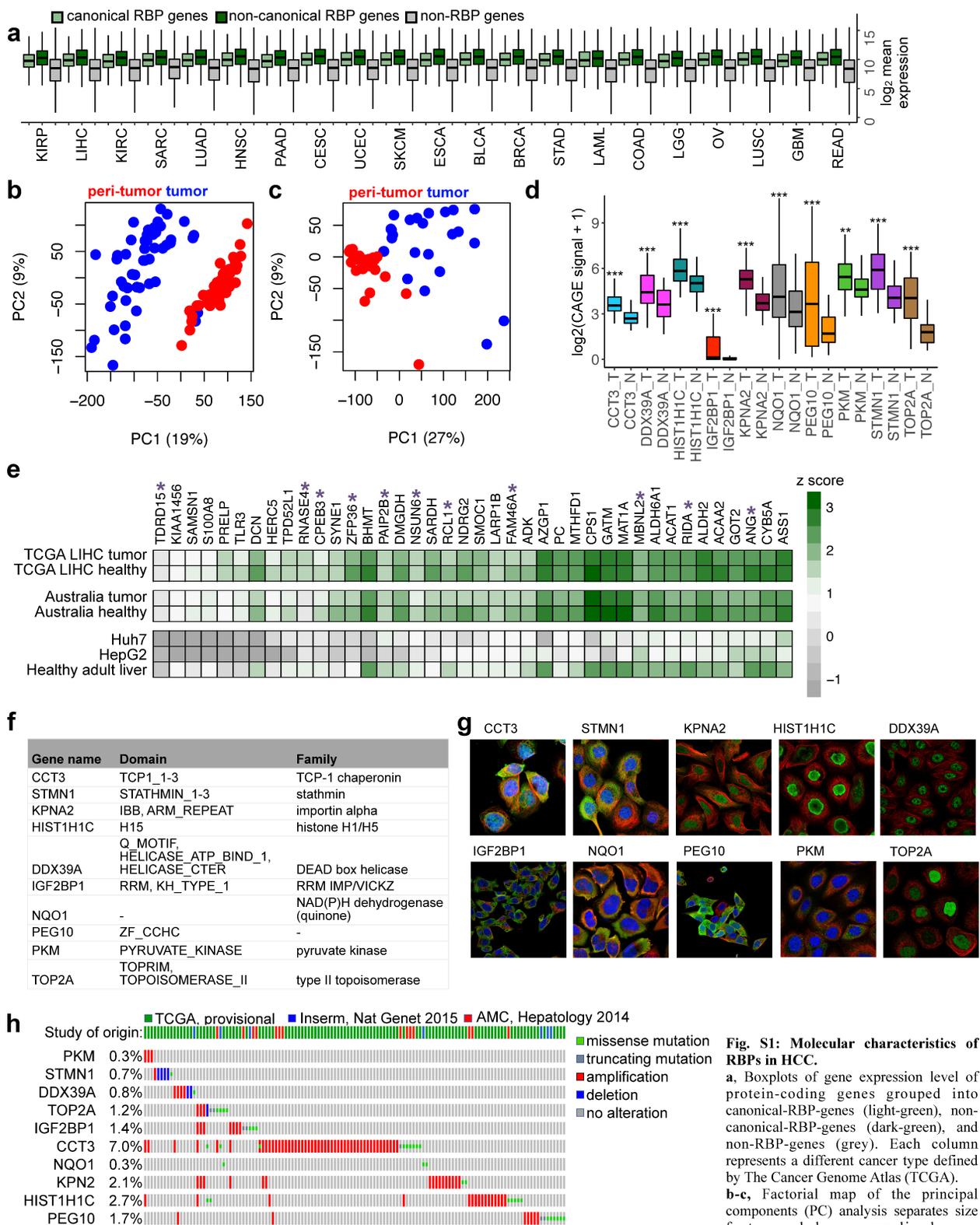
Australian HCC patient cohort RNA-seq: E-MTAB-8915

RBP-KD RNA-seq: E-MTAB-9587

lincRNA-OE RNA-seq: E-MTAB-9586

#### CODE ACCESSIBILITY

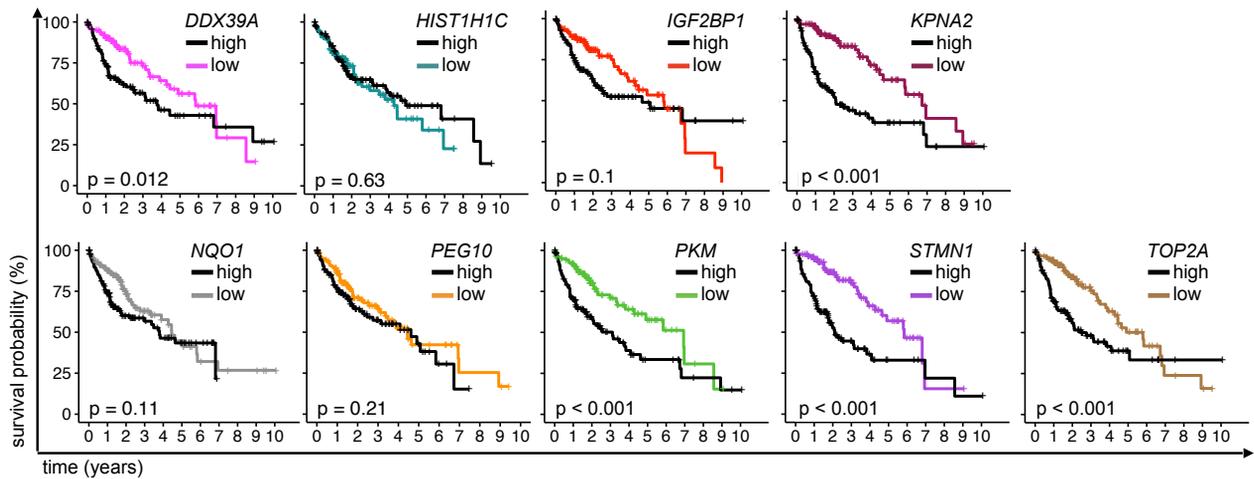
[github.com/jonasns/LiveRNome](https://github.com/jonasns/LiveRNome)



**Fig. S1: Molecular characteristics of RBPs in HCC.**

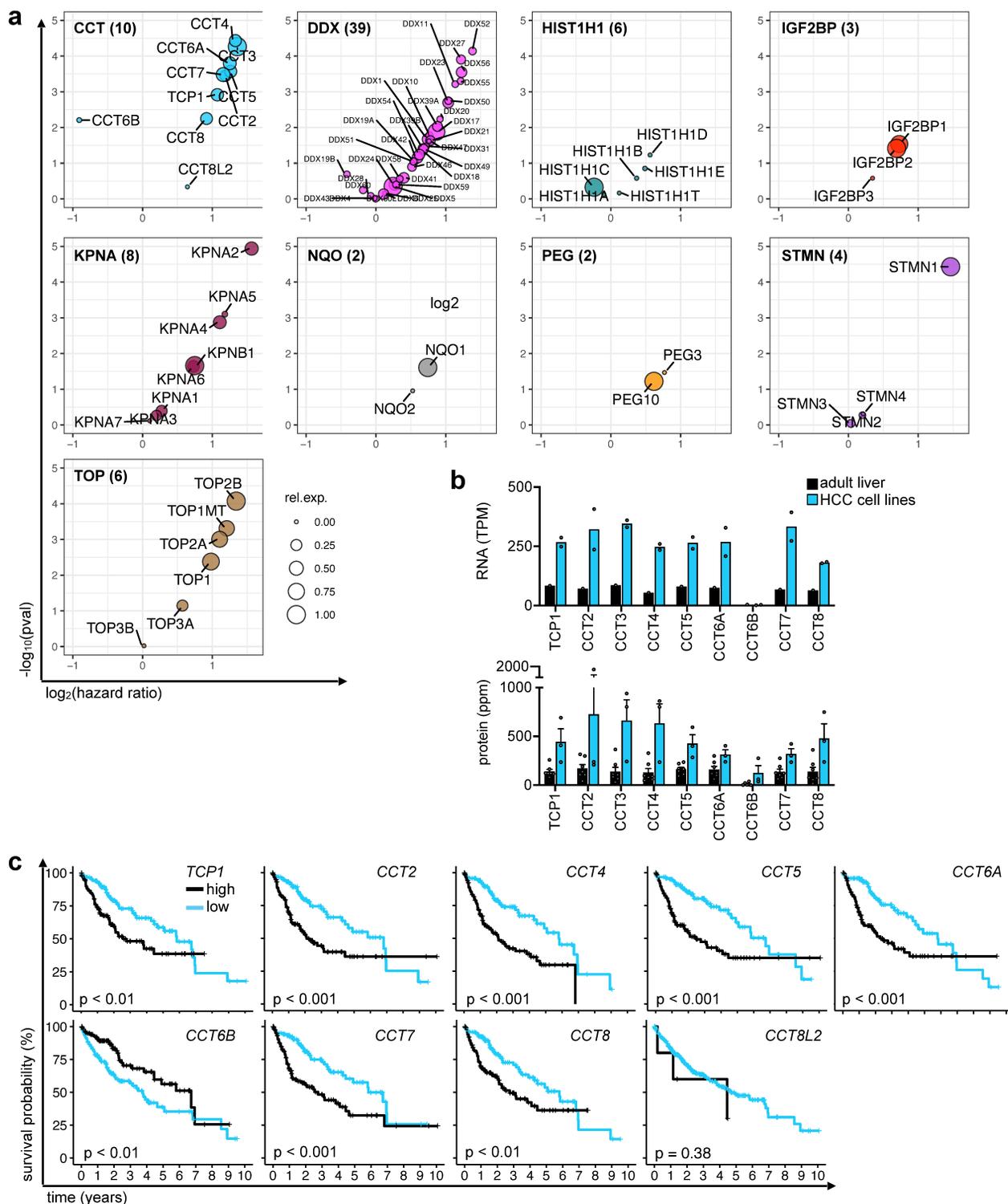
**a**, Boxplots of gene expression level of protein-coding genes grouped into canonical-RBP-genes (light-green), non-canonical-RBP-genes (dark-green), and non-RBP-genes (grey). Each column represents a different cancer type defined by The Cancer Genome Atlas (TCGA). **b-c**, Factorial map of the principal components (PC) analysis separates size factor and  $\log_2$  normalized gene expression levels of the (b) TCGA LIHC

(50 patients) and (c) Australia HCC (24 patients) cohort into tumor (blue) and peritumor (red) patient samples. The proportion of variance explained by each principal component is indicated in parenthesis. **d**, Box plots represent the expression levels of the ten selected RBP genes in the FANTOM HCC cohort (50 patients) (Hashimoto *et al.*, Genome Res., 2015). Hinges correspond to the first and third quartiles, and whiskers correspond to the 1.5-times interquartile range. T: tumor; N: peritumor tissue. Statistics: paired two-tailed t-test. \*\* $p < 0.01$ , \*\*\* $p < 0.001$ . **e**, Heatmap displays changes in expression level for commonly downregulated RBP genes in HCC as highlighted in (Fig. 1e). Expression level is sorted by average tumor z-score from right to left. Color gradient indicates z-score differences (green: high; grey: low). **f**, Table displays protein domains and family classification of the ten selected RBPs according to ExPASy (<https://prosite.expasy.org>). **g**, Microscopy images showing cellular location of immunofluorescently labeled RBPs (blue: DAPI staining of the nucleus, red: microtubules, green: RBP) in human cells (HepG2, A-431, U-2 OS, U-251) retrieved from the Human Protein Atlas (<https://www.proteinatlas.org>). **h**, Mutational status of the 10 selected RBP genes from three independent cohort studies (top panel, green: CGA, blue: INSERM, green: AMC) deposited in cBioPortal (<http://www.cbioportal.org>) shows proportional frequency of missense mutations (green), truncations (purple), amplifications (red), deletions (blue) and no alterations (grey). The percentage indicates the total number of alterations in all patients ( $n=865$ ).



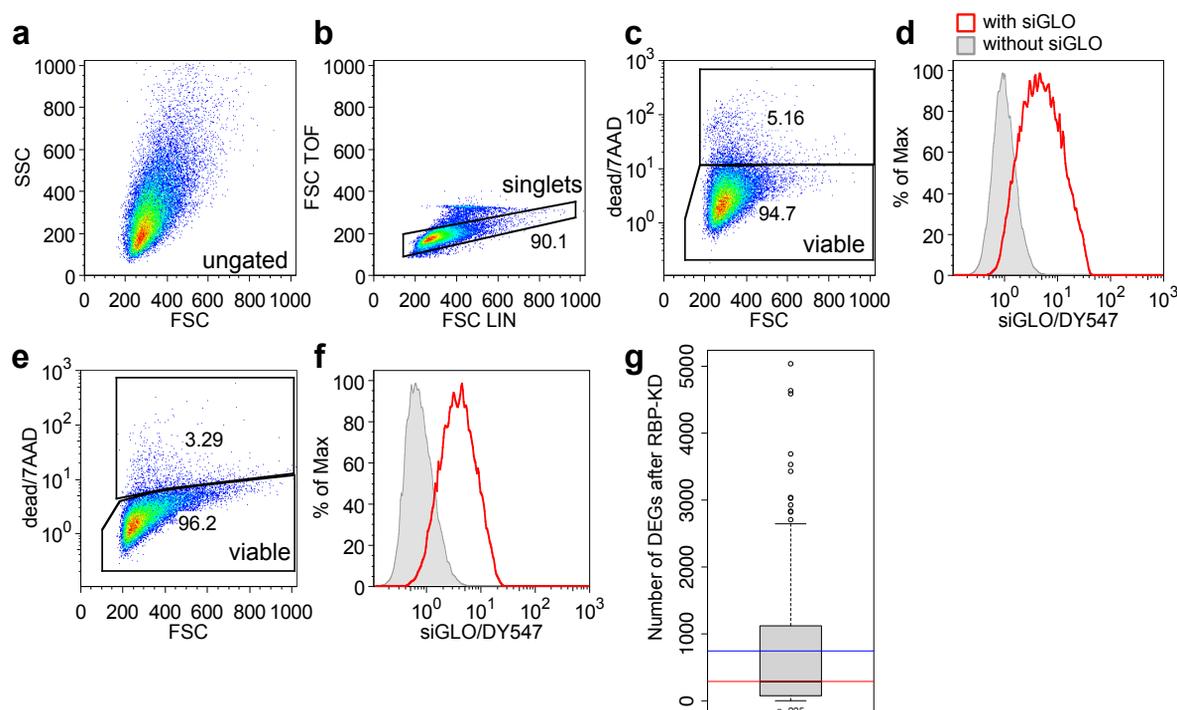
**Fig. S2: Survival analysis of the TCGA-LIHC cohort.**

Kaplan-Meier plots show the association of RBP gene expression level and ten-year survival within the top and bottom tercile of 377 TCGA-LIHC (black: high and colored: low RBP gene expression levels, color code: DDX39A (magenta), HIST1H1C (turquoise), IGF2BP2 (red), *KPNA2* (plum), NQO1 (grey), PEG10 (yellow), PKM (green), STMN1 (purple) and TOP2A (brown)). Statistics: log-rank (Mantel-Cox).



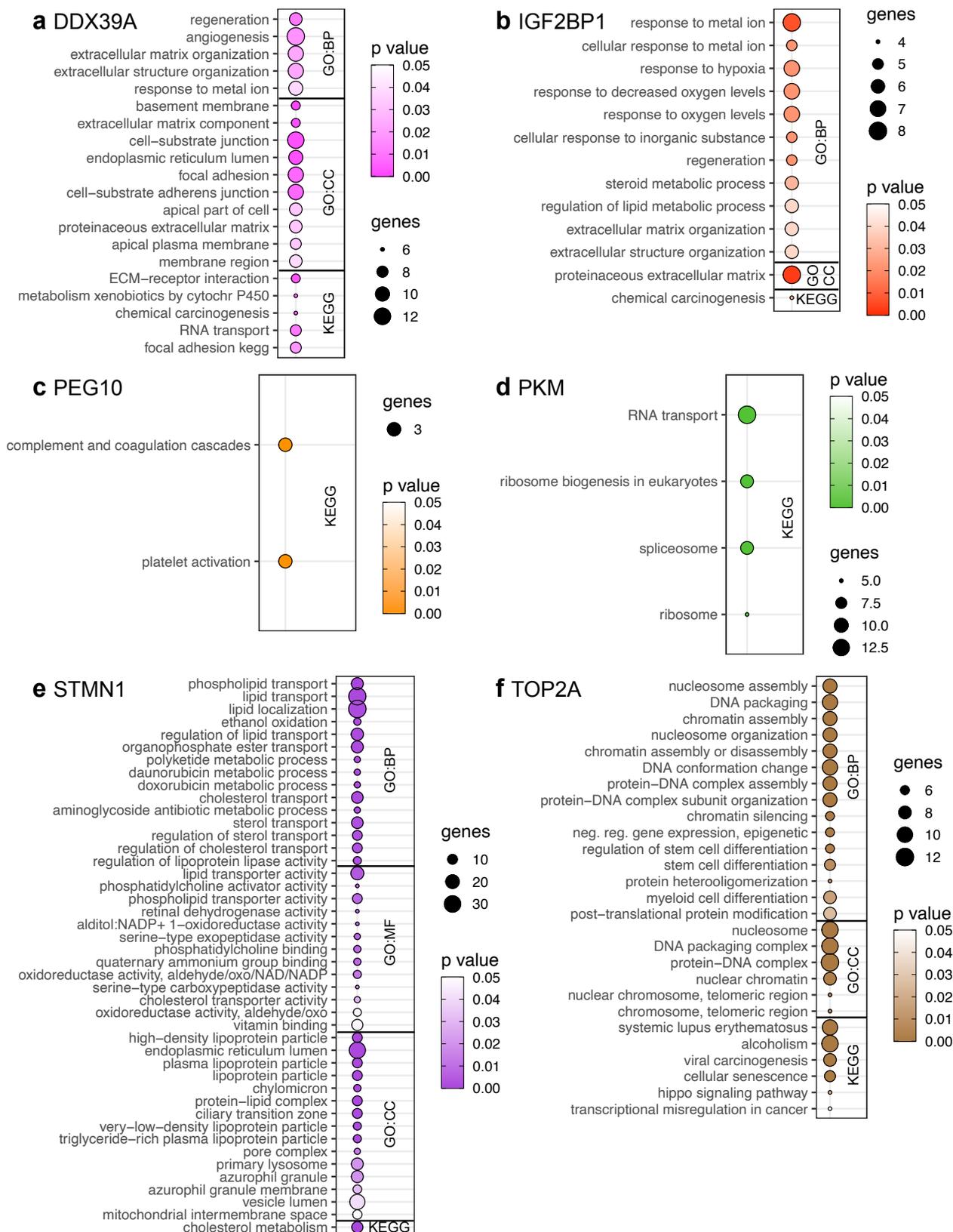
**Fig. S3: Estimation of hazard ratio and HCC patient survival based on RBP gene expression.**

**a**, Volcano plots display statistical significance ( $FDR < 0.05$ , likelihood ratio test) and hazard ratio comparing the top and bottom tertiles of RBP genes grouped by family. Numbers in brackets after the name of the gene family represent the total number of family members. PKM does not have any family members and is hence not displayed. The size of the circle represents the relative gene expression level of each gene per family (broad: high, narrow: low). **b**, Barplots demonstrate RNA (top) and protein (bottom) expression of the CCT3 family members. Data were obtained from Rudolph *et al.*, PLoS Genet, 2016; PAXdb v. 4.1 pax-db.org, accessed on 2021-03-24). **c**, Kaplan-Meier plots show the effect of CCT3-family members expression on ten-year survival within the top and bottom tertile of the TCGA-LIHC cohort (377 patients). Statistics: log-rank (Mantel-Cox).



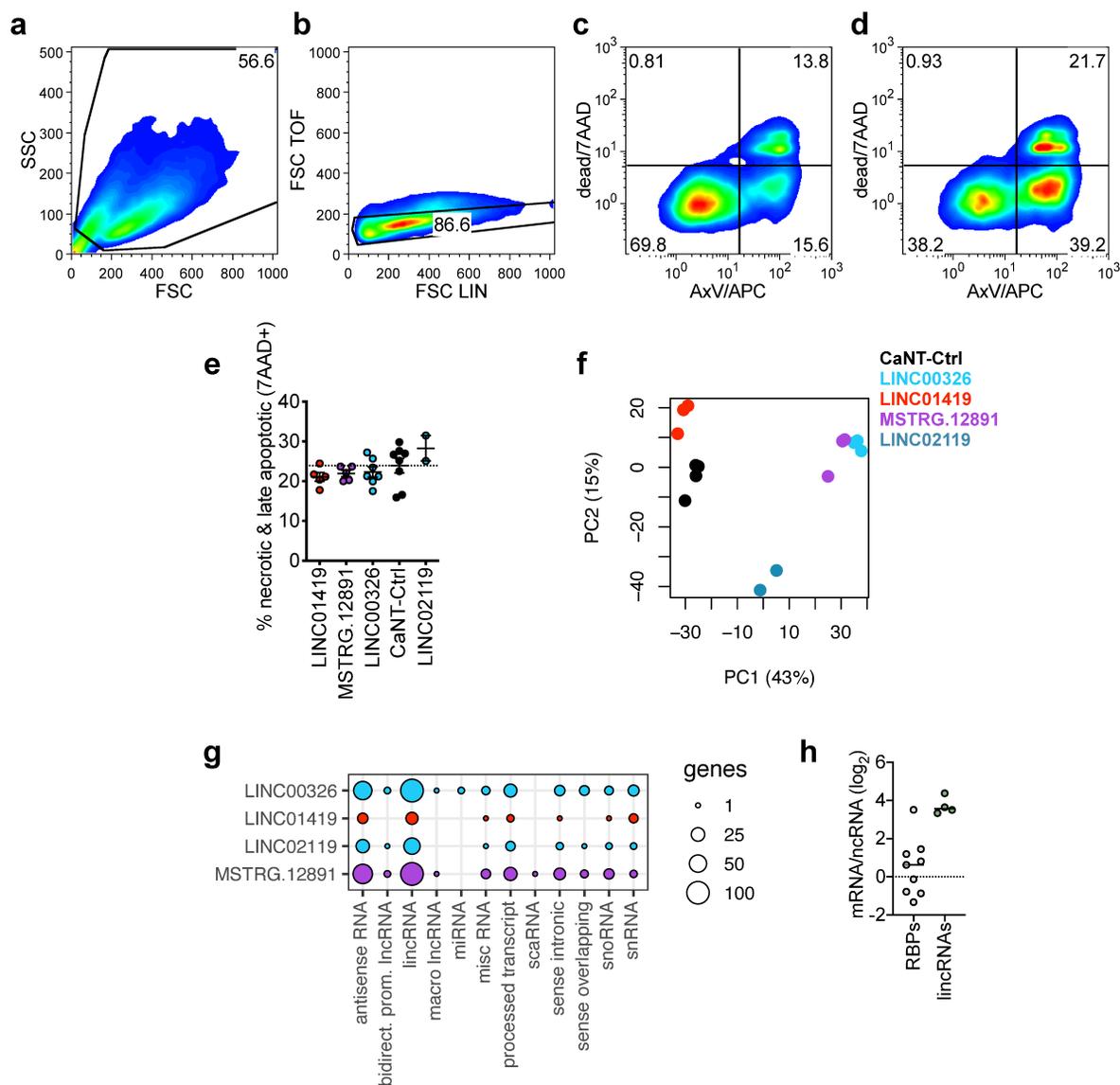
**Fig. S4: siRNA transfection efficiency and RBP-KD perturbation effects in human HCC cells**

**a-f**, Flow cytometry plots displaying gating strategy and transfection efficiency of siRNA (here siGLO) in **a-d** Huh7 and **e-f** HepG2 cells after electroporation. The completely shifted flow cytometry histogram indicates 100% transfection efficiency. Cell dead marker 7AAD was used to count dead and viable cells (shown as percentage). **g**, Boxplot showing the number of significantly differentially expressed genes (DEGs) upon RBP-KD in the ENCODE project database. Y axis represents number of DEGs; n represents the number of KD experiments; the box encloses all the data points between the upper and lower quartile; whiskers are drawn to the furthest data point falling within 1.5-times the interquartile range. Points represent outliers. The red line is median of the distribution and the blue line the mean (both median and mean are annotated with their respective values). Data is from (Van Nostrand *et al.*, Nature, 2020; Davis *et al.*, Nucleic Acids Res., 2018).



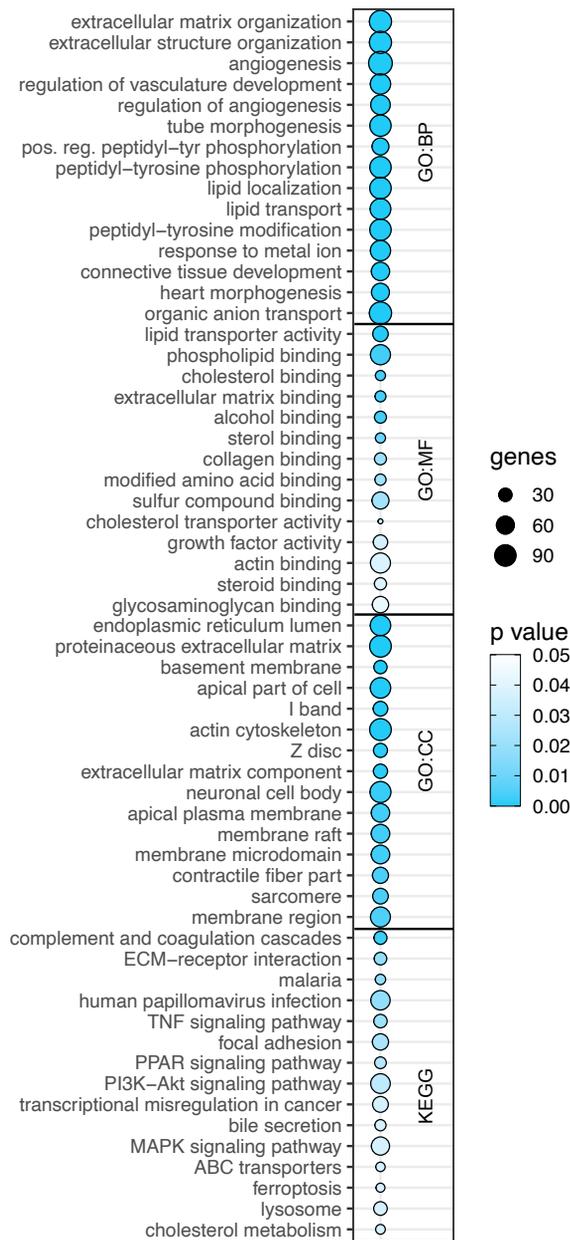
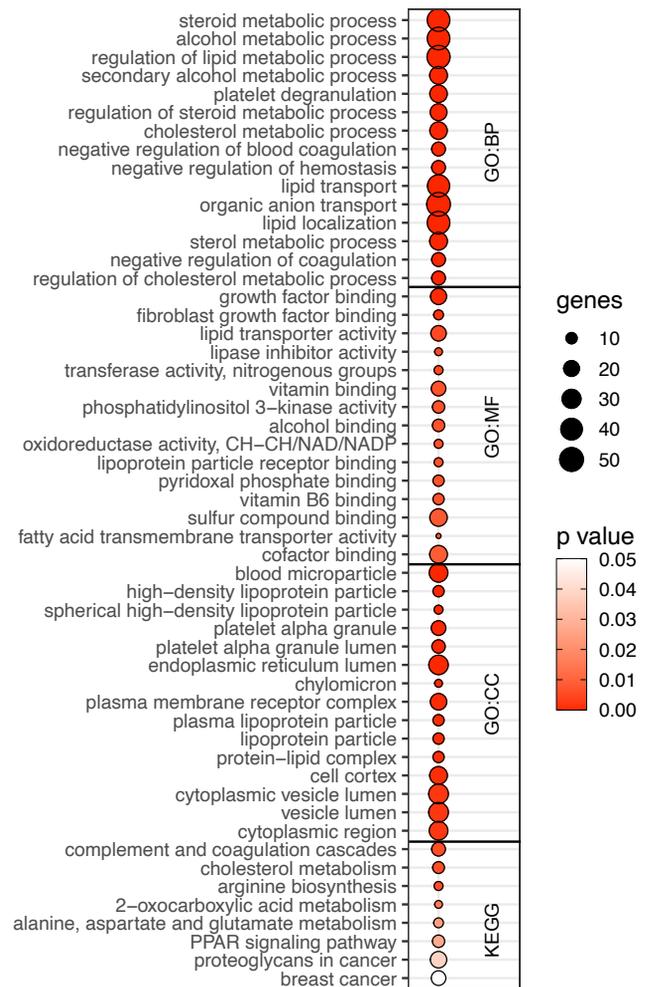
**Fig. S5: Gene Ontology (GO) and KEGG analyses of deregulated genes upon RBP-KD**

**a-f**, Circle plot shows GO term and KEGG pathway enrichment analysis of deregulated genes (FDR<0.05) after the siRNA-mediated KD of RBP **a**, DDX39A; **b**, IGF2BP1; **c**, PEG10; **d**, PKM; **e**, STMN1 and **f**, TOP2A. No significant pathways were found for genes deregulated upon KD of HIST1H1C, KPNA2, and NQO1. The diameter of the circles corresponds to the number of genes in each GO or KEGG term and the color code represents varying degrees of significance for the respective RBP-KD (white: high and colored: low adjusted p-value). BP, biological process; MF, molecular function; CC, cellular compartment.



**Fig. S6: Cellular and molecular phenotype assessment upon lincRNA-OE**

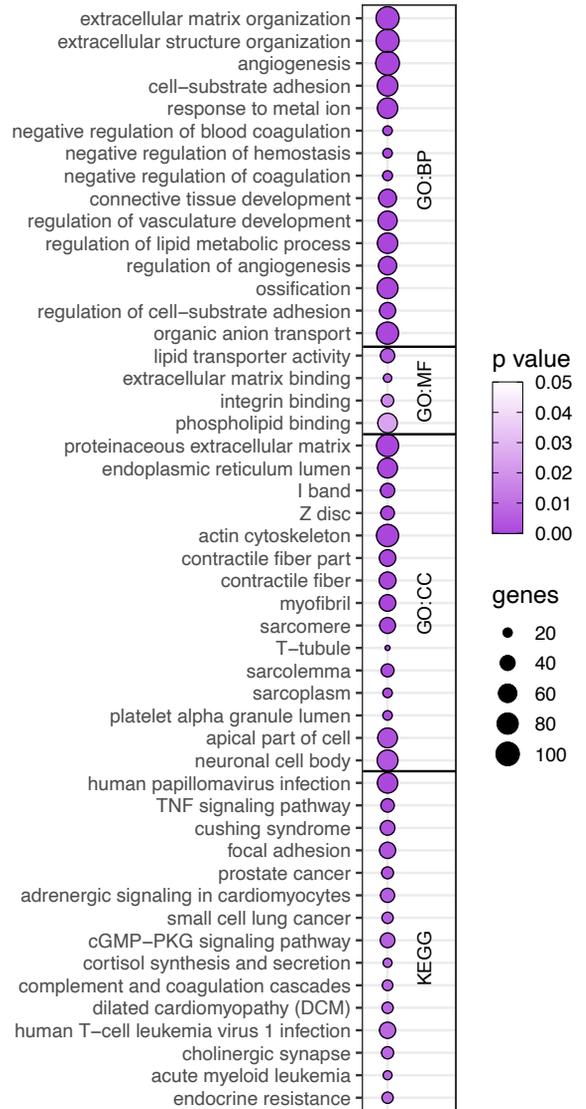
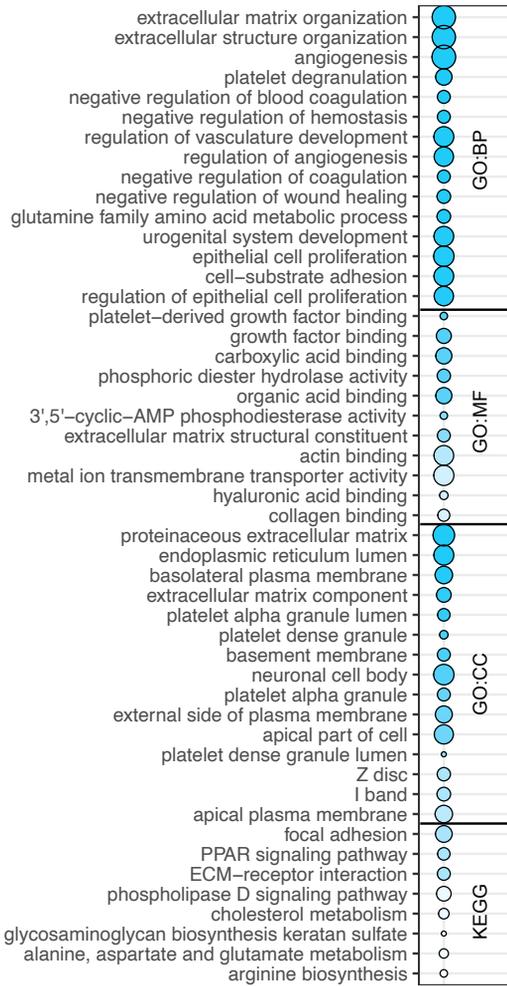
**a-d**, Flow cytometry gating strategy for assaying early apoptosis (Annexin V, AxV) and late apoptosis/necrosis (7AAD). **e**, dot plot displays the percentage of necrotic and late apoptotic (7AAD<sup>+</sup> cells assayed by flow cytometry) in Huh7 and HepG2 cells after CRISPRa-mediated lincRNA-OE after five days (n=2-8, mean, +/-SEM). The color-code links the lincRNA to the respective RBP-KD experiment through which it was identified (blue: CCT3, red: IGF2BP1, purple: STMN1). **f**, Factorial map of the principal components (PC) analysis separates size factor and log<sub>2</sub> normalized gene expression levels of the CRISPRa-mediated lincRNA-OE. The proportion of variance explained by each principal component is indicated in parenthesis. **g**, Circle plots display number of genes per non-coding RNA biotype affected by the CRISPRa-mediated lincRNA-OE. The diameter of the circles corresponds to the number of genes in each category. **h**, Dot plot of mRNA/lincRNA ratio after RBP-KD (n=10) or lincRNA-OE (n=4). The mean ratio is indicated by a black line.

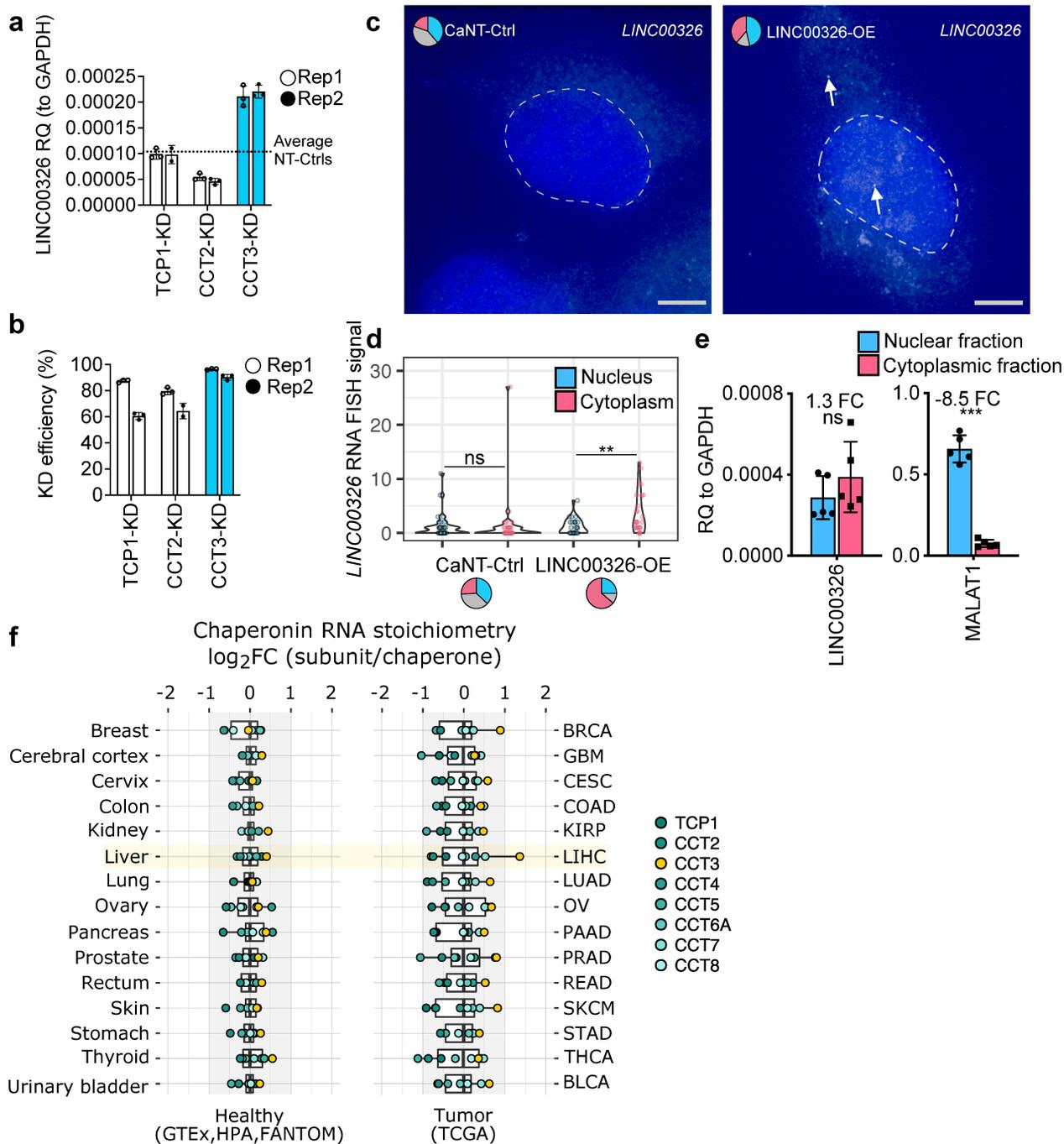
**a** LINC00326**b** LINC01419**Fig. S7: Gene Ontology (GO) and KEGG analyses of deregulated genes upon OE**

**a-d**, Circle plot shows GO term and KEGG pathway enrichment analysis of deregulated genes (FDR<0.01) after the CRISPRa-mediated lincRNA-OE **a**, *LINC00326*; **b**, *LINC01419*; **c**, *LINC02119* and **d**, *MSTR.12891*. The size of the circles indicates number of genes in each GO or KEGG term and the color represents varying degrees of significance for the respective lincRNA-OE experiment (white: high and colored: low adjusted p-value). The color-code links the lincRNA to the respective RBP-KD experiment through which it was identified (blue: CCT3, red: IGF2BP1, purple: STMN1). A maximum of 15 terms are shown in each GO or KEGG term (full list is available in Table S13). BP, biological process; MF, molecular function; CC, cellular compartment. Figure continues on the next page.

**c** LINC02119

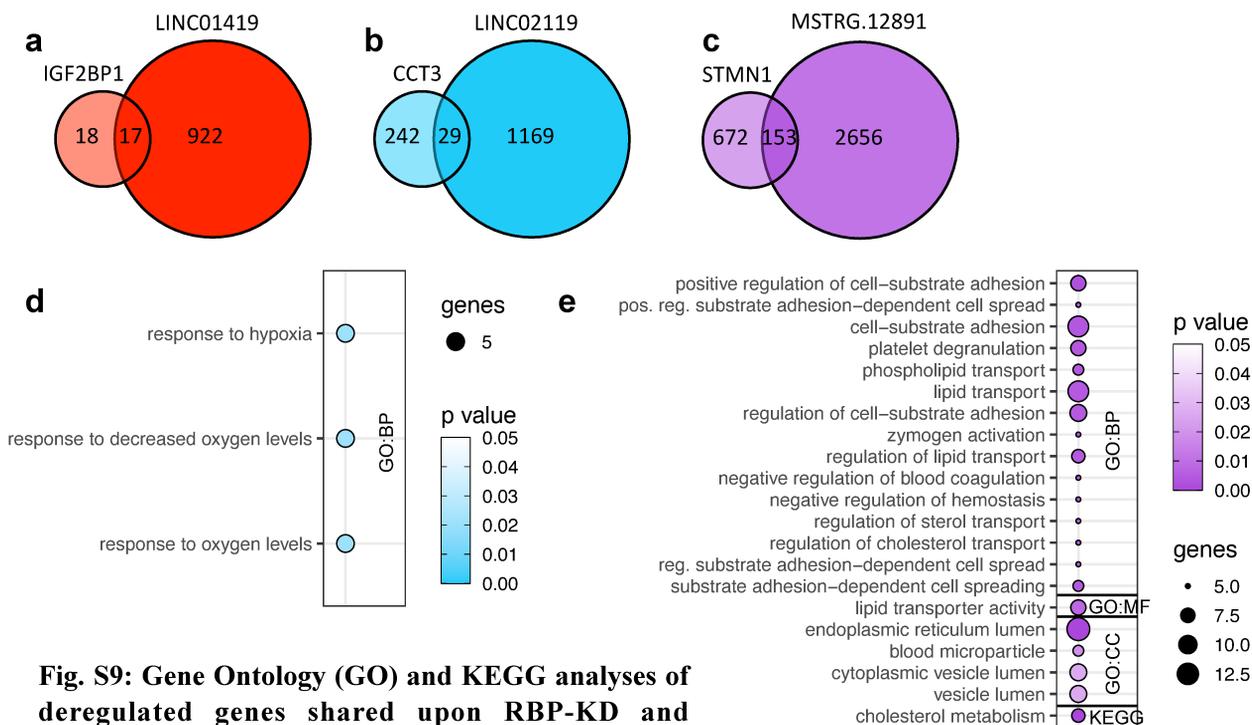
**d** MSTRG.12891





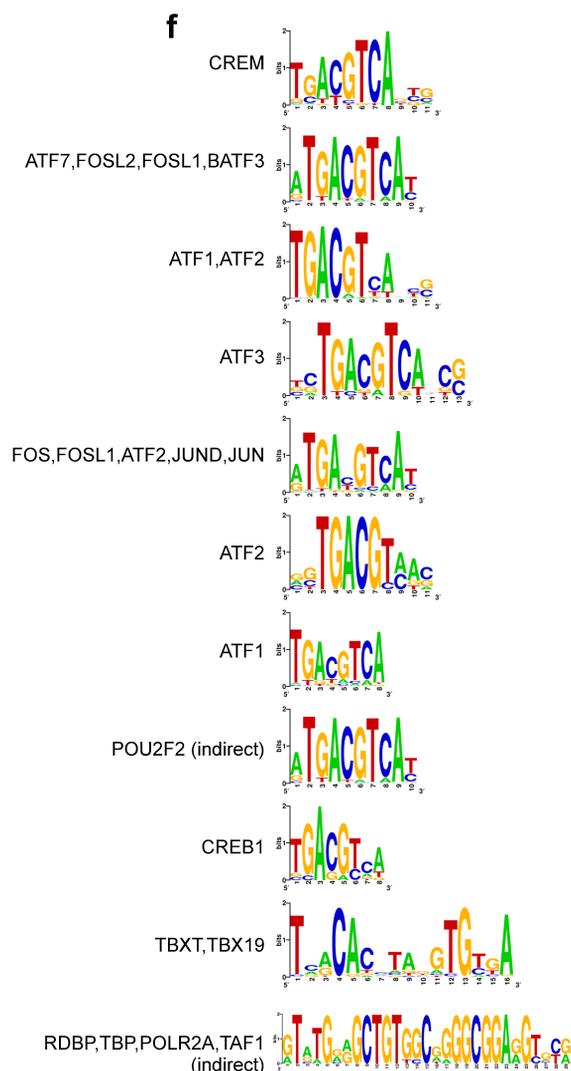
**Fig. S8: CCT3 works in a chaperoning-independent manner to interact with LINC00326.**

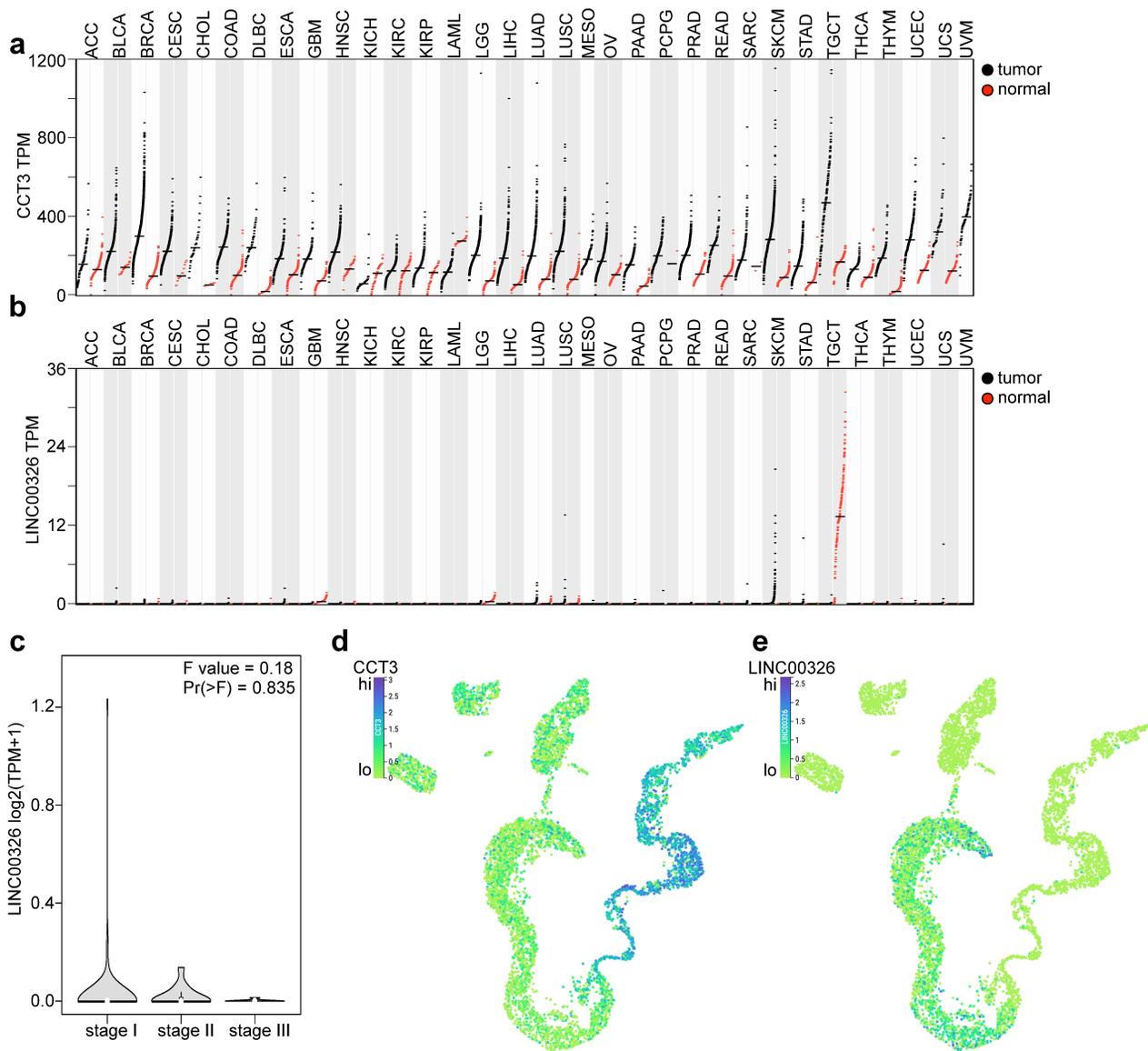
**a**, Bar graph shows the gene expression levels of *LINC00326* after CCT-family member KD. **b**, Bar graph shows KD-efficiency of samples in **a**. **c**, Microscopic images of single-molecule RNA FISH using exonic probes for *LINC00326* (white dots and arrows) in CaNT-Ctrl (left) and *LINC00326*-OE (right) Huh7 HCC cells. DAPI (blue) marks the nucleus. Pie chart represents the fraction of signals in the nucleus (blue) or cytoplasm (pink) in cells, or cells without any signal (grey). Scale bar: 5  $\mu$ m. **d**, Violin plots show quantification of *LINC00326* RNA FISH signal localization in HepG2 cells (n = 27-35). The pie chart represents the fraction of signal in the nucleus (blue) or cytoplasm (pink). Statistics: paired two-tailed t-test, \*\* p<0.01, ns: not significant. **e**, Bar graphs show gene expression patterns of *LINC00326* and *MALAT1* in separate cellular fractions of *LINC00326*-OE HCC cells. Quantification was performed by qPCR and gene expression values were normalized to *GAPDH*. Statistical significance and absolute fold change (FC) are displayed above the graphs. Statistics: paired two-tailed t-test, \*\*\* p<0.001. (n = 5). **f**, Box plots shows RNA stoichiometry data for subunits of the CCT chaperonin complex. The x axis value is computed as the fold change between the median gene expression level of the subunit and the median gene expression level of all subunits. The left boxes show healthy tissues and the right boxes show corresponding tissue tumors. Data were retrieved from the Human Protein Atlas (v. 20.1, <https://www.proteinatlas.org>).



**Fig. S9: Gene Ontology (GO) and KEGG analyses of deregulated genes shared upon RBP-KD and lincRNA-OE**

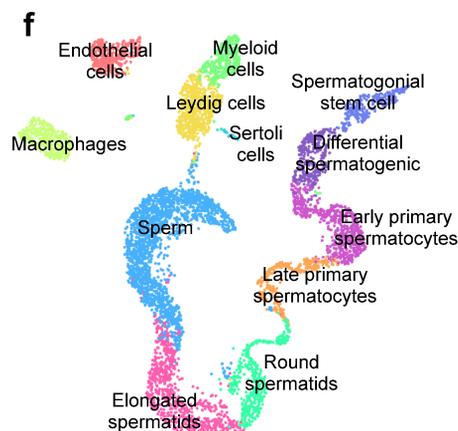
**a-c**, Two-way Venn diagrams intersect the number of deregulated genes in the RBP-KD and lincRNA-OE. **a**, *IGF2BP1*-KD and *LINC01419*-OE, **b**, *CCT3*-KD and *LINC2119*-OE, **c**, *STMN1*-KD and *MSTR.12891*-OE. **d-e**, GO and KEGG pathway analysis of overlapping genes from (b) *CCT3*-KD and *LINC2119*-OE and (c) *STMN1*-KD and *MSTR.12891*-OE. No significant pathways were found for genes deregulated upon *IGF2BP1*-KD and *LINC01419*-OE. The diameter of the circles corresponds to the number of genes in each GO or KEGG term and the color code represents varying degrees of significance for the respective RBP-KD and lincRNA-OE (white: high and colored: low adjusted p-value). BP, biological process; MF, molecular function; CC, cellular compartment. A maximum of 15 terms are shown in each category (full list is available in Table S13). **f**, Identified motifs for each TF in **Fig. 5f**.

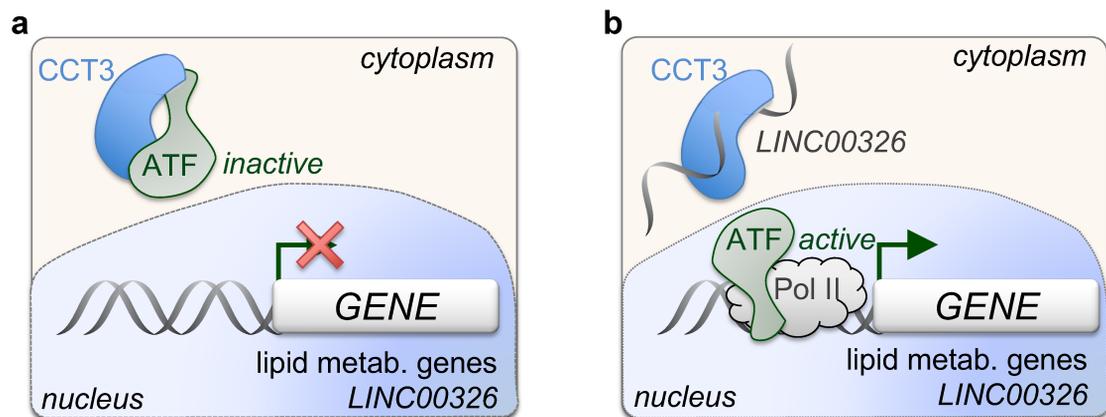




**Fig. S10: LINC00326 is highly expressed in healthy testis and diminished upon cancer progression**

**a-b**, Dot plot demonstrating the expression level in TPM of (a) CCT3 and (b) LINC00326 across different tissues from the TCGA and GTEx dataset. ACC (n=77-128 (tumor-normal)), BLCA (n=404-28), BRCA (n=1085-291), CESC (n=306-13), CHOL (n=36-9), COAD (n=275-349), DLBC (n=47-337), ESCA (n=182-286), GBM (n=163-207), HNSC (n=519-44), KICH (n=66-53), KIRC (n=523-100), KIRP (n=286-60), LAML (n=173-70), LGG (n=518-207), LIHC (n=369-160), LUAD (n=483-347), LUSC (n=486-338), MESO (n=87-0), OV (n=426-88), PAAD (n=179-171), PCPG (n=182-3), PRAD (n=492-152), READ (n=92-318), SARC (n=262-2), SKCM (n=460-558), STAD (n=408-211), TGCT (n=137-165), THCA (n=512-337), THYM (n=118-339), UCEC (n=174-91), UCS (n=57-78), UVM (n=79-0). **c**, LINC00326 expression in tumor biopsies from testis cancer patients divided by cancer stage. **d-e**, UMAP of single cell RNA expression of (d) CCT3 (e) LINC00326 in (f) human testis cells. a-c data are from: GEPIA (Tang *et al.* Nucleic Acids Res., 2017), and d-e are from Guo *et al.* Cell Res., 2018, visualised at covid19cellatlas.org (accessed on 21.3.21). ACC: adrenocortical carcinoma, BLCA: bladder urothelial carcinoma, BRCA: breast invasive carcinoma, CESC: cervical squamous cell carcinoma and endocervical adenocarcinoma, CHOL: cholangiocarcinoma, COAD: colon adenocarcinoma, DLBC: lymphoid neoplasm diffuse large B-cell lymphoma, ESCA: esophageal carcinoma, GBM: glioblastoma multiforme, HNSC: head and neck squamous cell carcinoma, KICH: kidney chromophobe, KIRC: kidney renal clear cell carcinoma, KIRP: kidney renal papillary cell carcinoma, LAML: acute myeloid leukemia, LGG: brain lower grade glioma, LIHC: liver hepatocellular carcinoma, LUAD: lung adenocarcinoma, LUSC: lung squamous cell carcinoma, MESO: mesothelioma, OV: ovarian serous cystadenocarcinoma, PAAD: pancreatic adenocarcinoma, PCPG: pheochromocytoma and paraganglioma, PRAD: prostate adenocarcinoma, READ: rectum adenocarcinoma, SARC: sarcoma, SKCM: skin cutaneous melanoma, STAD: stomach adenocarcinoma, TGCT: testicular germ cell tumors, THCA: thyroid carcinoma, THYM: thymoma, UCEC: uterine corpus endometrial carcinoma, UCS: uterine carcinosarcoma, UVM: uveal melanoma.





**Fig. S11: Molecular model for LINC00326 regulating lipid metabolism-associated gene expression**

**a**, In HCC, CCT3 moonlights as an RBP, perhaps due to increased gene expression of CCT3. CCT3 binds to transcription factor ATF mRNA or forms a heteromeric protein complex. This causes that CCT3 sequester away ATF from initiating transcription of genes associated with lipid metabolism or *LINC00326*. **b**, when expressed, *LINC00326* RNA interacts with CCT3. This consequently limits the binding of CCT3 and ATF protein, thus releasing ATF and allowing transcription of lipid metabolism genes and *LINC00326* itself. A similar molecular mechanism occurs when CCT3 is reduced, thereby enabling expression of *LINC00326*.

## Supplementary Materials & Methods

### “The CCT3-LINC00326 axis regulates hepatocarcinogenic lipid metabolism”

Søndergaard *et al.*

---

#### **Data and code availability**

Supplementary tables can be found here: <https://figshare.com/s/2c05765158269b3b4ff2>

The datasets generated in this study are available in the ArrayExpress repository, under accession numbers:

E-MTAB-8915

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8915>,

username: Reviewer\_E-MTAB-8915, password: xpizwqkc),

E-MTAB-9587

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9587>,

username: Reviewer\_E-MTAB-9587, password: 25ddT0yi), and

E-MTAB-9586

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9586>,

username: Reviewer\_E-MTAB-9586, password: M2tY9UPH). All scripts used for

bioinformatics analysis are available on Github: <https://github.com/jonasns/LiverNome>.

Single molecule RNA FISH imaging files generated in this study are available here:

Treated cells: <https://figshare.com/s/a83dbec52555e922ca8d>

Untreated cells: <https://figshare.com/s/08b0f84f2ea241b03c8d>

#### **Identification of RBPs from prior studies**

The list of 2,282 proteins that can bind RNA in liver was generated by combining information from published datasets [1,2,11–19,3–10]. The list includes both canonical and non-canonical RBPs defined by whether a known RBD is present.

#### **Cell culture**

All assay plates, dishes and culture flasks were purchased from Sarstedt. HepG2 and Huh7 human HCC cell lines were obtained from American Type Culture Collection (ATCC, Rockville, MD) with certified genotype. Huh7 cells constitutively expressing tubulin-GFP [20] were kindly provided by Prof. Bartenschlager (University of Heidelberg). All cell lines were mycoplasma free when periodically tested with MycoplasmaCheck (Eurofins Genomics). To ensure authenticity, cell lines were initially genotyped by short-tandem repeat genetic profiling (STR) using the Power-Plex\_16HS\_Cell Line panel and analyzed using Applied Biosystems Gene Mapper ID v3.2.1 software by the external provider Genetica DNA Laboratories (LabCorp Specialty Testing Group) and continuously assessed phenotypically. Huh7 and HepG2 cells were cultured in T75 flasks at 37°C and 5% CO<sub>2</sub> atmosphere using Dulbecco's Modified Eagle Medium (DMEM) supplemented with 1/100 Penicillin/Streptomycin (P/S, Sigma) and 10% fetal bovine serum (Hyclone, GE healthcare).

Huh7 cells were maintained by splitting 1/6 three times a week, HepG2 by splitting 1/4 three times a week. This was done by aspirating the medium, gently washing the cells with phosphate buffered saline (PBS) without  $Mg^{2+}$  (Sigma) and detaching them with 2mL of a trypsin-EDTA solution (Sigma) for 3-5min. Trypsin was inactivated with a minimum of 10-fold surplus of culture medium before a cell fraction was passaged. To ensure authenticity, cell lines were initially genotyped by short-tandem repeat genetic profiling (STR) using the PowerPlex\_16HS\_Cell Line panel and analyzed using Applied Biosystems Gene Mapper ID v3.2.1 software by the external provider Genetica DNA Laboratories (LabCorp Specialty Testing Group) and continuously assessed phenotypically.

### ***Patient material***

Tissue specimens were obtained under informed consent from 24 patients undergoing liver resection of HCC from the Australian Victorian Biobank, according to the approval of local ethics committee (application number: 2010/541-31/1 and 2017/719-31/2). The patients taking part in this study were 75% males and 25% females, and presented with HCC from Hepatitis B virus infection, non-alcoholic fatty liver disease, alcoholic steatohepatitis, hereditary haemochromatosis and other HCC triggering conditions. See patient information on ArrayExpress E-MTAB-8915 (reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-8915>, username: Reviewer\_E-MTAB-8915, password: xpizwqkc). It was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans of our research.

### ***Survival analysis of TCGA data***

Survival analysis was done in R (v. 3.4.3) using survival (v. 3.1-8) and survminer (v. 0.4.6). The script is available at: <https://github.com/jonasns/LiveRNome>. A total of 377 patients were used for the analysis, and divided into terciles based on normalized htseq-counts of mRNA expression. Metadata and raw htseq-counts were acquired from <https://www.cancer.gov/tcga> (accessed on 2017.02.17).

### ***siRNA-mediated KD of RBPs***

At 70-80% confluency, cells were harvested by trypsination and subjected to electroporation. Small interfering RNA (siRNA) POOLS targeting *PEG10*, *HIST1H1C*, *KPNA2*, *PKM*, *NQO1*, *TOP2A*, *CCT3*, *STMN1*, *DDX39A*, and *IGF2BP1* was used (Thermo Scientific). siRNA ON-TARGETplus Non-Targeting siRNA#1 (Thermo Scientific) was used as control. Cells were washed twice in PBS and once in OptiMEM without phenol red (Invitrogen). A total of 5µg (18.8µL of a 20µM stock) siRNA was transferred to a 4mm cuvette (Bio-Rad), and 5 million cells were added in 200µl OptiMEM and incubated for 3min before being pulsed with an exponential decay pulse at 300V, 250µF, in a Genepulser II (Bio-Rad). Immediately after electroporation, the cells were transferred to pre-heated (37°C) phenol red-free DMEM culture medium supplemented with 10% (v/v) FBS and without the addition of antibiotics. Electroporation conditions were optimized using DY547-siGLO RISC-free control siRNA (Dharmacon) and flow cytometry (FACSNavios, Beckman Coulter, Navios Cytometry List Mode Data Acquisition and Analysis Software version 1.3). Optimization results were analyzed using FlowJo (v. 8.2). Gating strategy can be found in Fig. S4.

### ***RNA extraction and DNase-treatment***

All reagents and devices were purchased from ThermoFisher Scientific, if not stated otherwise. Upon cell harvest, 700µL Qiazol (Qiagen) was directly added onto the cells on ice and mixed. At this point, the cell extract was either stored at -80°C or directly added 140µL chloroform. This Qiazol/chloroform mixture was shaken for 30 sec and incubated at RT for 2.5min, before centrifugation at 9,000g for 5min at 4 °C. The mixture separated into a lower chloroform-phenol phase, an interphase and an upper, aqueous phase. The latter was carefully transferred to a new reaction tube before adding 1 volume isopropanol, inverting the tube 5 times followed by 10min incubation at RT. The mixture was centrifuged at 9,000g and 4°C for 10min and the supernatant discarded. The pellet was washed carefully using 700µL cold 70% ethanol, flicking the tube and centrifuged at 15,000g and 4°C for 5min. The supernatant was discarded entirely, and the pellet air-dried for 5min, before resuspension in 20–50µL nuclease-free water (NF-H<sub>2</sub>O, Ambion). RNA concentration was determined by nanodrop (Nanodrop 2000c). When possible, 10µg RNA was mixed with NF-H<sub>2</sub>O, 5µL TurboDNase buffer, 1µL TurboDNase and 1µL RNase Inhibitor (RiboLock, Invitrogen) in a total volume of 50µL and incubated at 37°C for 30min. Afterwards, 5µL Turbo DNase Inactivation reagent was added to the tube and incubated at RT for 5min. During incubation, the tube was occasionally flicked. The tube was centrifuged at 10,000g and 4°C for 2min and the supernatant carefully transferred to a new reaction tube. In order to achieve higher purity, the RNA was precipitated overnight. This was done by adding 19µL 3M sodium acetate (pH 5.2), 1µL GlycoBlue (Ambion), 135µL NF-H<sub>2</sub>O and 600µL ice-cold 99.8% ethanol, homogenizing and storing the tube at -80°C. The next day, the sample was centrifuged at 12,000g for 30min, the supernatant discarded, and the pellet washed twice with 700µL 70% ethanol. The pellet was air-dried, resuspended in 25µL NF-H<sub>2</sub>O and the RNA concentration determined by nanodrop.

For the Australian HCC cohort, total RNA from the liver biopsies was extracted using the RNeasy mini kit (Qiagen) and treated with DNase RNase free kit (Qiagen).

### ***cDNA synthesis and qPCR***

All reagents and devices were purchased from ThermoFisher Scientific, if not stated otherwise. 1-2µg of pure DNase-treated RNA was added 1µL random primers (250ng/µL) and 1µL dNTP Mix (10m each) to a final volume of 13.5µL. The sample was incubated in a thermocycler (ProFlex PCR) at 65°C for 5min. Afterwards, 7µL of a mastermix containing 1µL RNase-Inhibitor, 4µL First-Strand Buffer, 1µL 0.1M DTT and 0.5µL SuperScript II reverse transcriptase were added. Samples were incubated in the thermocycler at 25°C for 10min, 42°C for 50min and 70°C for 15min and a final hold at 4°C.

The cDNA used for qPCR was diluted 1/10 to 1/25 depending on the experiment. 3.6X Master mixes were prepared, containing 3.6µL of primer mix (forward and reverse primer, 2.5µM each), 10.8µL NF-H<sub>2</sub>O and 18µL SYBR-Green Mix (PowerUp) per sample. 3.6µL of diluted cDNA was added to a 96-well PCR plate and 32.4µL of the prepared master mix added. After mixing, three times 10µL were transferred to a 384-well plate to obtain three technical replicates per sample. After sealing the plate and a short centrifugation, it was run on Real-Time-PCR machine (QuantStudio 5 or CFX384 (Bio-Rad)). The following

parameters were set: initial heating steps of 3min at 50°C followed by 3min at 95°C. 40 cycles of 15sec at 95°C and 30sec at 60°C. Melt curve: 60-95°C using 0.5°C increments every 5sec. Quantitative PCR data was checked for correct amplification and dissociation of the products. mRNA levels of the genes of interest were normalized to mRNA levels of the housekeeping genes *GAPDH* or *ACTB* and were calculated according to the ddCT method.

### **Primer design**

Exon sequences of desired genes were obtained from NCBI Reference Sequence Database (RefSeq) and used as input on [www.primer3.ut.ee](http://www.primer3.ut.ee) (version 4.1.0). The following parameters were set: T<sub>m</sub> 60-65°C with optimum of 60°C, GC-content 40-60% with optimum of 50% and product size range from 100-300 bp. Primer results were validated using the in-silico PCR tool in UCSC Genome Browser (Human, Assembly: Dec. 2013 – GRCh38/hg38, target: genome assembly). Whenever possible, the primers were designed to be intron-spanning. Primers used can be found in Table S14.

### **MTT proliferation assay**

Directly after electroporation, 5,000 Huh7 or 7,000 HepG2 cells were seeded into 96 well plates in triplicates. Cells were assayed between day 1 and day 7 every 48h. For this purpose, the medium was aspirated and a mixture of 10µL MTT (Methylthiazolyldiphenyl-tetrazolium bromide, Sigma, 4mg/mL in 1X PBS) and 60µL culture medium added to the cells, before incubating at 37°C and 5% CO<sub>2</sub> atmosphere for 1 h. After carefully removing the medium, 100µL of MTT-lysis buffer per well (for 100mL MTT lysis buffer: 90mL isopropanol, 2.5mL 10% SDS, 4mL HCl, 3.5mL H<sub>2</sub>O) was added onto the cells and incubated on an orbital shaker for 15-30min. The lysate was resuspended, the bottom of the assay plate wiped with ethanol and the absorption at 595nm measured on a plate reader (Spectramax i3x, Molecular Devices). Background values were subtracted from the obtained values before normalization to the average value of the plate in order to compare between different experiments.

### **Viability measures**

Supernatants containing non-adherent cells and adherent cells (harvested by trypsination) were harvested directly into FACS tubes. For 7-AAD single staining, cells were washed 1X in PBS with 1% BSA (Sigma), followed by staining in 100µL buffer with 5µL 7-AAD Viability Staining Solution (eBioscience) on ice in the dark for 15min. Cells were acquired directly on a flow cytometer (FACSNavios, Beckman Coulter) without washing away the staining buffer. For detection of early apoptosis, cells were washed twice in PBS (without BSA), and resuspended in 100µL Annexin V binding buffer (10mM HEPES (pH 7.4), 150mM NaCl, 2.5mM CaCl<sub>2</sub> in H<sub>2</sub>O) containing 5µL 7AAD and 5µL Annexin V-APC (eBioscience). Cells were stained on ice in the dark for 30min and acquired as for 7-AAD single staining. Data were acquired by flow cytometry (FACSNavios, Beckman Coulter, Navios Cytometry List Mode Data Acquisition and Analysis Software version 1.3) and analyzed using FlowJo (v. 8.2). Gating strategy can be found in Fig. S6.

### **Sequencing of RNA libraries**

The library preparation was carried out using the Illumina TruSeq® Stranded RNA Library Prep Kit v2 (dual index) according to the manufacturer's instruction and ribosomal depletion was achieved with Ribo-Zero Gold. The quality of every cDNA library was determined on high-sensitivity DNA chips (Agilent Bioanalyzer) according to the manufacturer's protocol. Libraries were quantified with the KAPA-SYBR FAST qPCR kit (Roche). The sequencing run was performed either with the NextSeq 500/550 High Output v2 kit (Illumina) for 150 cycles, paired end, on a Illumina NextSeq 500 or HiSeq Rapid SBS Kit v2 for 50 cycles, single end on a Illumina HiSeq2500. All raw data (fastq files) are accessible under

ArrayExpress:

E-MTAB-9587

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9587>,  
username: Reviewer\_E-MTAB-9587, password: 25ddT0yi) and

E-MTAB-9586

(reviewer link: <http://www.ebi.ac.uk/arrayexpress/experiments/E-MTAB-9586>,  
username: Reviewer\_E-MTAB-9586, password: M2tY9UPH)

### ***RNA-seq data QC and processing***

Next generation sequencing read quality were assessed with FastQC (v. 0.11.5) and RseQC (v. 2.6.4). Adaptor sequences were trimmed and low-quality reads removed using Trimmomatic (v. 0.36). All sequencing reads aligning (HiSAT2, v. 2.1.0) to annotated ribosomal RNA genes were discarded. High-quality and ribosomal RNA depleted sequencing reads were aligned to the genome using HiSAT2. Novel RNA transcripts were identified using StringTie (v. 1.3.3). Using sorted bam files (Samtools v. 1.5), the number of aligned reads were counted (featurecount in subread package v. 1.5.2) for both annotated transcripts and novel references generated with StringTie. After normalization (TMM: trimmed mean of M-values), a differential gene expression analysis (edgeR v. 3.20.7) was performed in R (v. 3.4.3). For the two patient cohorts (TCGA and Australia), DESeq (v. 1.14.1 in bioconductor v. 3.4) was used for differential gene expression analysis. In general, we use edgeR for  $n < 6$  and DESeq for  $n > 6$  replicates. Significant DE genes was distinguished by a false discovery rate (FDR) under 0.05 or 0.01. Gene ontology analysis was performed in R (v. 3.4.3) with clusterProfiler (v. 3.6.0) and org.Hs.eg.db (v. 3.5.0), with p values corrected by the Benjamini-Hochberg FDR procedure. Additionally, the following dependent package versions were installed: DOSE (v. 3.4.0), AnnotationDbi (v. 1.40.0), IRanges (v. 2.12.0), S4Vectors (v. 0.16.0), BiocGenerics (v. 0.24.0), and Biobase (v. 2.38.0). All scripts are available on Github: <https://github.com/jonasns/LiveRNome>.

### ***CRISPRa vector generation and cell electroporation***

SP-dCas9-VPR [21] were acquired from Addgene plasmid # 63798 (deposited by George Church). Three gRNAs targeting within the 300 bp upstream region of the transcriptional start site were designed for each lincRNA using the Broad Institute design tool (<https://portals.broadinstitute.org/gpp/public/analysis-tools/sgrna-design>, accessed on 2018.01.08). Each gRNA was checked for off-target effects using both the Broad Institute tool's internal checkup and by BLATing them in UCSC genome browser. gRNAs were assembled into SP-dCas9-VPR using Gibson assembly (NEB). Non-targeting gRNAs were

from [22] and is not found anywhere in the human genome. The gRNA sequences can be found in Table S14. HepG2 and Huh7 cells were electroporated using the NEON electroporation system (Invitrogen) as previously described [20]. Transfection efficiency was checked by qPCR of target genes.

### ***Lipid experiments***

Assays for lipid peroxidation (MDA) (Abcam, ab233471), DCFDA cellular ROS detection (Abcam, ab113851) and hepatic lipid accumulation/steatosis (ORO, Abcam, ab133131) were conducted according to the manufacturers' protocols. For the MDA assay, cells were lysed directly in the culture dish with a modified 4°C cold RIPA buffer (10mM Tris-HCl pH 8.0, Sigma), 150mM NaCl (Sigma), 1mM Na<sub>2</sub>EDTA (Fisher), 1mM EGTA (VWR), 1% NP-40 (Sigma), 1% sodium deoxycholate (VWR), 0.1% SDS (Sigma). Lysates were kept on ice unless otherwise stated. The MDA assay was performed 48h after transfection, while for the ROS and ORO assays, cells were collected, counted and re-plated in equal numbers 48h after transfection, followed by readout the day after. The ROS assay was performed in 96wp, while ORO was performed in 24wp, until after dye extraction, which was transferred to 96wp. All experiments were read out on a Spectramax i3x.

### ***Zebrafish xenograft experiments***

Zebrafish experimentation was performed at the zebrafish core facility at Comparative Medicine, Karolinska Institute, Sweden. Husbandry and breeding were in accordance to the ethical permit Dnr 14049-2019 approved by the Stockholm North Ethical Board. Zebrafish experimentation in embryos younger than five days are excluded from the normative on animal testing by the EU directive 2010/63/EU.

Huh7 cells constitutively expressing TUBULIN-GFP [23] were cultured identical to standard Huh7 cells, except for adding 10ug/mL Blasticidin S HCl (Thermo Fisher) to the growth medium. TUBULIN-GFP Huh7 cells were grown to 80% confluency in a T75 flask. Medium was removed and cells were washed twice with pre-warmed 1xPBS. Five mL of pre-warmed 1xPBS was added to the cell culture followed by 20µL of Vybrant DiICell-labeling solution (Thermo Fisher). After incubating for 20min at 37°C excessive dye was washed away with 1xPBS and cells were harvested using Accutase (Sigma). Accutase was inactivated with 10-fold excess culture medium and filtered through a pre-washed 40µM cell strainer (Sigma). After pelleting (500g, 4min) cells were resuspended in 1mL of 2% (w/v) polyvinylpyrrolidone (Sigma) in PBS and transported on ice to the zebrafish core facility. Cells were re-pelleted, resuspended by flicking the tube, and transplanted into the perivitelline space of 48h old zebrafish embryos (AB genotype). Tumor growth was followed by daily by epifluorescence microscopy (Leica) from day 1 to 5 after injection. Tumor area was quantified using Fiji in ImageJ (v. 1.52).

### ***Promoter analysis***

Promoter analysis was done using the gimmeroc command in GimmeMotifs v0.13.1 with fasta sequences 400bp upstream of DEG transcription start site as input and 1000 random sequences with equal length and GC content from hg38 as background.

***RNA immunoprecipitation followed by quantitative PCR (RIP-qPCR)***

RIP-qPCR was done using the Magna RIP RBP IP kit (Sigma) according to the manufacturer's instructions. Cell lysates were prepared from Huh7 or HepG2 HCC cell lines with *LINC00326*-OE, and immunoprecipitated using CCT3 Ab 10571-1-AP (Proteintech, lot #00021192) rotating overnight at 4°C. HULC was used as a negative control.

***Subcellular fractionation***

Subcellular fractions were prepared as described in [17]. Cultured cells at an approximate confluence of 90 % were harvested, counted, washed with cold PBS and  $\sim 0.5 \times 10^7$  cells were transferred to a reaction tube. Cells were centrifuged at  $500 \times g$  and 4°C for 5 min, the supernatant discarded, and the cells very gently resuspended in 380  $\mu$ L hypotonic lysis buffer (HLB) containing 100 U RNase Inhibitor. After 10 min incubation on ice, the lysate was centrifuged at  $1000 \times g$  and 4°C for 3 min. The supernatant containing the cytoplasmic fraction was carefully transferred to a new reaction tube, while leaving the pellet on ice. 1 mL RNA precipitation solution was directly added to the supernatant, and stored at -20°C for at least 1 h. The pellet containing the nuclear fraction was washed three times by addition of 1 mL HLB and centrifugation at  $200 \times g$  at 4°C for 2 min. Subsequently, the pellet was dissolved in 1 mL Qiazol and kept on ice for short- or at -80°C for long-term storage. After 1 h of incubation at -20°C, the cytosolic fraction was vortexed for 30 sec and centrifuged at  $18,000 \times g$  and 4°C for 15 min. After discarding the supernatant, the pellet was washed with ice-cold 70% ethanol, centrifuged at  $18,000 \times g$  and 4°C for 5 min and air dried. Subsequently, the pellet was resuspended in 1 mL Qiazol and mixed until the pellet was fully dissolved. RNA from the cytosolic and nuclear fractions was extracted and the reverse transcription followed by quantitative PCR was performed.

***Single-molecule RNA FISH***

Cells were grown on coverslips in 12-well plates and were left untreated or were transfected with *LINC00326* or appropriate negative control guides with CRISPRa. The cells were briefly washed with 1xPBS (diluted in RNase-free H<sub>2</sub>O) and fixed with 75% methanol + 25% acetic acid at RT for 10 min. Following fixation, cells were washed twice with 1xPBS. The coverslips were briefly rehydrated with Wash Buffer A (SMF-WA1-60, Biosearch) (Formamide, Thermo Fischer Scientific, AM9342) before FISH. The Stellaris FISH Probes (*LINC00326* exonic probes, Q570) were designed using the Stellaris online RNA FISH probe designer (Stellaris Probe Designer, version 4.2), which was produced by LGC Biosearch Technologies. The *LINC00326* probes were added to the hybridization buffer (SMF-HB1-10, LGC Biosearch Technologies) at a final concentration of 250 nM, and hybridization was carried out in a humidified chamber at 37 °C overnight. The following day, the cells were washed twice with Wash Buffer A (SMF-WA1-60, LGC Biosearch Technologies) at 37 °C for 30 min. The second wash contained 4,6-diamidino-2-phenylindole for nuclear staining (5  $\mu$ g/ml, Merck-Sigma, D9542). The coverslips were then briefly washed with wash buffer B (SMF-WB1-20, LGC Biosearch Technologies) for 5 min at RT, equilibrated 5 min in base glucose buffer (2x SSC, 0.4% glucose solution 49163, Merck-Sigma, in RNase-free H<sub>2</sub>O), and then incubated 5 min in Base Glucose buffer supplemented with 1:100 dilution of glucose oxidase (stock 3.7 mg/mL, G2133, Merck-Sigma) and catalase (stock 4 mg/mL,

219261-100KU, Merck-Sigma). Finally, the coverslips were mounted with ProlongGold Antifade Mountant (P36934, Thermo Fischer Scientific). Images were captured with wide-field DeltaVision microscope (Applied Precision, LLC) equipped with a Coolsnap HQ2 12 bit camera with 1x1 binning and 896x896 frame size. We acquired 20-30 optical slices depending on the thickness of the cell lines at 0.3- $\mu$ m intervals using Olympus 100x (1.4 numerical aperture) oil immersion objectives. Deconvolution of three-dimensional stacks was performed with built-in DeltaVision SoftWoRx Imaging software (Applied Precision, LLC). Maximum intensity projections were generated in Fiji and subjected for manual quantification using Fiji. The sequences of LINC00326 RNA FISH probe can be found in Table S14.

### Supplementary Materials & Methods References

- 1 Castello A, Fischer B, Eichelbaum K, *et al.* Insights into RNA biology from an atlas of mammalian mRNA-binding proteins. *Cell* 2012;**149**:1393–406. doi:10.1016/j.cell.2012.04.031
- 2 Castello A, Fischer B, Frese CK, *et al.* Comprehensive Identification of RNA-Binding domains in human cells. *Mol Cell* 2016;**63**:696–710. doi:10.1016/j.molcel.2016.06.029
- 3 Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet* 2014;**15**:829–45. doi:10.1038/nrg3813
- 4 Baltz AG, Munschauer M, Schwanhäusser B, *et al.* The mRNA-bound proteome and Its global occupancy profile on protein-coding transcripts. *Mol Cell* 2012;**46**:674–90. doi:10.1016/j.molcel.2012.05.021
- 5 Caudron-Herger M, Rusin SF, Adamo ME, *et al.* R-DeeP: proteome-wide and quantitative identification of RNA-dependent proteins by density gradient ultracentrifugation. *Mol Cell* 2019;**75**:184-199.e10. doi:10.1016/j.molcel.2019.04.018
- 6 Bao X, Guo X, Yin M, *et al.* Capturing the interactome of newly transcribed RNA. *Nat Methods* 2018;**15**:213–20. doi:10.1038/nmeth.4595
- 7 Perez-Perri JI, Rogell B, Schwarzl T, *et al.* Discovery of RNA-binding proteins and characterization of their dynamic responses by enhanced RNA interactome capture. *Nat Commun* 2018;**9**. doi:10.1038/s41467-018-06557-8
- 8 Mullari M, Lyon D, Jensen LJ, *et al.* Specifying RNA-Binding Regions in Proteins by Peptide Cross-Linking and Affinity Purification. *J Proteome Res* 2017;**16**:2762–72. doi:10.1021/acs.jproteome.7b00042
- 9 Trendel J, Schwarzl T, Horos R, *et al.* The Human RNA-Binding Proteome and Its Dynamics during Translational Arrest. *Cell* 2019;**176**:391-403.e19. doi:10.1016/j.cell.2018.11.004
- 10 Brannan KW, Jin W, Huelga SC, *et al.* SONAR Discovers RNA-Binding Proteins from Analysis of Large-Scale Protein-Protein Interactomes. *Mol Cell* 2016;**64**:282–93. doi:10.1016/j.molcel.2016.09.003
- 11 Treiber T, Treiber N, Plessmann U, *et al.* A Compendium of RNA-Binding Proteins that Regulate MicroRNA Biogenesis. *Mol Cell* 2017;**66**:270-284.e13. doi:10.1016/j.molcel.2017.03.014
- 12 He C, Sidoli S, Warneford-Thomson R, *et al.* High-Resolution Mapping of RNA-Binding Regions in the Nuclear Proteome of Embryonic Stem Cells. *Mol Cell* 2016;**64**:416–30. doi:10.1016/j.molcel.2016.09.034
- 13 Conrad T, Albrecht AS, De Melo Costa VR, *et al.* Serial interactome capture of the human cell nucleus. *Nat Commun* 2016;**7**. doi:10.1038/ncomms11212

- 14 Beckmann BM, Horos R, Fischer B, *et al.* The RNA-binding proteomes from yeast to man harbour conserved enigmRBPs. *Nat Commun* 2015;**6**. doi:10.1038/ncomms10127
- 15 Sundararaman B, Zhan L, Blue SM, *et al.* Resources for the Comprehensive Discovery of Functional RNA Elements. *Mol Cell* 2016;**61**:903–13. doi:10.1016/j.molcel.2016.02.012
- 16 Hendrickson D, Kelley DR, Tenen D, *et al.* Widespread RNA binding by chromatin-associated proteins. *Genome Biol* 2016;**17**. doi:10.1186/s13059-016-0878-3
- 17 Cook KB, Kazan H, Zuberi K, *et al.* RBPDB: A database of RNA-binding specificities. *Nucleic Acids Res* 2011;**39**. doi:10.1093/nar/gkq1069
- 18 Ray D, Kazan H, Cook KB, *et al.* A compendium of RNA-binding motifs for decoding gene regulation. *Nature* 2013;**499**:172–7. doi:10.1038/nature12311
- 19 Kwon SC, Yi H, Eichelbaum K, *et al.* The RNA-binding protein repertoire of embryonic stem cells. *Nat Struct Mol Biol* 2013;**20**:1122–30. doi:10.1038/nsmb.2638
- 20 Søndergaard JN, Geng K, Sommerauer C, *et al.* Successful delivery of large-size CRISPR/Cas9 vectors in hard-to-transfect human cells using small plasmids. *Commun Biol* 2020;**3**:1–6. doi:10.1038/s42003-020-1045-7
- 21 Chavez A, Scheiman J, Vora S, *et al.* Highly efficient Cas9-mediated transcriptional programming. *Nat Methods* 2015;**12**:326–8. doi:10.1038/nmeth.3312
- 22 Liu SJ, Horlbeck MA, Cho SW, *et al.* CRISPRi-based genome-scale identification of functional long noncoding RNA loci in human cells. *Science* 2017;**355**:eaah7111. doi:10.1126/science.aah7111
- 23 Cortese M, Goellner S, Acosta EG, *et al.* Ultrastructural Characterization of Zika Virus Replication Factories. *Cell Rep* 2017;**18**. doi:10.1016/j.celrep.2017.02.014

## Supplemental Material

### “The CCT3-*LINC00326* axis regulates hepatocarcinogenic lipid metabolism”

Søndergaard *et al.*

---

#### SUPPLEMENTARY FIGURES

Figure S1. Molecular characteristics of RBPs in HCC

Figure S2. Survival analysis of the TCGA-LIHC cohort

Figure S3. Estimation of hazard ratio and HCC patient survival based on RBP gene expression

Figure S4. siRNA transfection efficiency and RBP-KD perturbation effects in human HCC cells

Figure S5. Gene Ontology (GO) and KEGG analyses of deregulated genes upon RBP-KD

Figure S6. Cellular and molecular phenotype assessment upon lincRNA-OE

Figure S7. GO and KEGG analyses of deregulated genes upon OE

Figure S8. CCT3 works in a chaperoning-independent manner to interact with *LINC00326*

Figure S9. GO and KEGG analyses of deregulated genes shared upon RBP-KD and lincRNA OE

Figure S10. *LINC00326* is highly expressed in healthy testis and diminished upon cancer progression

Figure S11. Molecular model for *LINC00326* regulating lipid metabolism-associated gene expression

#### SUPPLEMENTARY TABLES

Table S1. hRBP list

Table S2. TCGA LIHC DEG

Table S3. AUS counts

Table S4. AUS DEG

Table S5. RBP KD RNAseq raw counts

Table S6. RBP KD RNAseq TPM

Table S7. RBP KD RNAseq DEG

Table S8. RBP KD RNAseq GO KEGG

Table S9. novel genes info

Table S10. lincRNA OE RNAseq raw counts

Table S11. lincRNA OE RNAseq TPM

Table S12. lincRNA OE RNAseq DEG

Table S13. lincRNA OE and RBP overlap GO KEGG

Table S14. Primer list

Supplementary tables and microscopic imaging files are accessible via Figshare:

<https://figshare.com/s/2c05765158269b3b4ff2>

<https://figshare.com/s/a83dbec52555e922ca8d>

<https://figshare.com/s/08b0f84f2ea241b03c8d>

#### SUPPLEMENTARY MATERIALS & METHODS

Cell-based and molecular assays as well as xenograft experiments are described in the supplementary materials and methods.

#### ARRAYEXPRESS ACCESSION

Australian HCC patient cohort RNA-seq: E-MTAB-8915

RBP-KD RNA-seq: E-MTAB-9587

lincRNA-OE RNA-seq: E-MTAB-9586

#### CODE ACCESSIBILITY

[github.com/jonasns/LiveRNome](https://github.com/jonasns/LiveRNome)