ZNF204P is a stemness-associated oncogenic long non-coding RNA in hepatocellular carcinoma

Ji-Hyun Hwang1, Jungwoo Lee1, Won-Young Choi1, Min-Jung Kim1, Jiyeon Lee2, Khanh Hoang Bao Chu3, Lark Kyun Kim2,* & Young-Joon Kim1,3,*

INTRODUCTION

Hepatocellular carcinoma (HCC) is the main form (90%) of primary liver cancer (1), which is the fourth most common cause of cancer-related death worldwide. Cancer stem cell (CSC) model has been attributed to create tumor heterogeneity, in which a subset of tumor cells capable of self-renewal and extensive proliferation result in high recurrence rates and chemoresistance acquisition (2). CSC-induced tumor heterogeneity is associated with poor prognosis in several cancers (3) and thus CSC-specific markers (extracellular and intracellular) have been reported as crucial therapeutic targets (4) to selectively eradicate cells with tumorogenic capabilities. As CSCs share their key properties such as self-renewal and proliferation with normal stem cells, there is an unsurprising large overlap between CSC and normal stem cell surface markers such as CD133, CD90 and EpCAM (5). However, studies indicate that detailed curation is still required to fully distinguish CSCs (4). Transcription factors OCT4 and SOX2, best defined as essential players in maintaining pluripotency and self-renewal in embryonic stem cells/induced pluripotent stem cells (6), can be reactivated in CSCs showing similar properties (7), albeit under an aberrant light. Identification of such stemness-related factors would be useful in addressing CSCs-related malignant phenotypes (4).

ROR, HOTAIR and UCA1 are some well-established CSC-regulating long non-coding RNAs (lncRNAs) through various molecular mechanisms involving proliferation, self-renewal and metastasis promotion (8). lncRNAs are ideal target factors as their regulation spans epigenetic, post-transcriptional and transcriptional levels (9) with reported function in various cancers (10) and cancer-type specific expression (11). Pseudogenes which were first defined as ‘junk’ genomic loci, have recently undergone resuscitation as a novel IncRNA class, namely pseudogene-derived IncRNAs with high relevance in cancer initiation and progression (12). Pseudogenes which were first defined as ‘junk’ genomic loci, have recently undergone resuscitation as a novel IncRNA class, namely pseudogene-derived IncRNAs with high relevance in cancer initiation and progression (12). Pseudogene-encoding IncRNAs such as PTENP1 (13) and DUXAP10 (14) have demonstrated their critical role as competing endogenous RNAs by binding competitively to shared microRNAs (miRNAs) that target either their cognate gene or unrelated genes (14). Thus, identifying stemness-related factors in form of IncRNAs would...
be highly relevant in deepening understanding of gene circuits regulating cancer stemness.

In our study, we propose a user-friendly method of using public microarray data to identify IncRNA stemness-related factors in liver cancer, followed by preliminary experimental validation. Our candidate ZNF204P is a pseudogene-derived IncRNA and its knockdown inhibits cell proliferation, migration/invasive and colony formation properties. The cytosolic ZNF204P shares a miRNA-binding site (miRNA-145-5p) with OCT4 and SOX2, implying its role as a molecular decoy for tumor-suppressive miRNA-145-5p by interfering with its targeted degradation of OCT4 and SOX2. To the best of our knowledge, this study is the first to publish experimental validation of ZNF204P, and our data collectively indicates that ZNF204P is a stemness-associated oncogenic IncRNA in HCC with possible function as a miRNA-decoy, providing another avenue to identify prognostic and/or therapeutic targets.

RESULTS

Stemness-related candidate selection

Microarray datasets (GSE50206, GSE25097) were utilized for stemness-related candidate selection. HCC samples from GSE25097 were selected based on elevated expression of various stemness markers (EpCAM, Sall4, BMI-1 and SOX12) (15-17). Up-regulated genes in embryonic stem cells (ESCs) relative to human dermal fibroblasts (HDFs) (Fig. 1A, left) and those in selected HCC samples compared to healthy samples (Fig. 1A, right, Supplementary Table 1) were shortlisted through GEO2R analysis. log2FC-filtered significantly dysregulated genes were overlapped to find common gene candidates (Fig. 1B, Supplementary Fig. 1A, B) that regulate stemness in both normal stem cells and tumor samples. Gene ontology (GO) analysis of resulting 139 genes presented association with stemness regulation and maintenance (Fig. 1C). We focused on ZNF204P, the sole pseudogene-derived IncRNA given its importance as a growing novel class of IncRNA in cancer stemness regulation (Supplementary Table 1). We found that ZNF204P is up-regulated in tumor samples (Fig. 1D) with poor prognosis (Fig. 1E). Similar analysis patterns are reflected in well-known stemness-driven tumorigenic factors (18, 19) MYCN and MYB (Supplementary Fig. 2). Furthermore, gene set enrichment analysis depicted enrichment for increased cell proliferation and maintenance of stem cells in ZNF204P-high group (Fig. 1F), consolidating our stand that ZNF204P is a stemness-associated oncogenic IncRNA.

ZNF204P knockdown impairs cell proliferation, cell migration and invasion in HepG2 cell line

FANTOM CAT browser was accessed to define transcription start site (TSS) of ZNF204P (Fig. 2A) for CRISPR interference (CRISPR) single-guide RNA (sgRNA) design. We found that most of the FANTOM-annotated ZNF204P transcripts corresponded to the cytosolic ZNF204P and most were weakly expressed in HepG2 cell line (Fig. 2B). Combined use of 2 sgRNAs showed the best KD efficacy. (C) ZNF204P KD led to reduced cell proliferation over time. (D) Reduced number of colonies were observed post-ZNF204P KD in colony formation assay (left) and quantification of colonies in an average of 3 fields per group (right). All data were acquired from three independent experiments (n = 2). *P < 0.05; **P < 0.01; ***P < 0.001 (two-tailed unpaired t-test analysis); ns indicates no significance.
to the same CAGE (Cap Analysis of Gene Expression) cluster (Fig. 2A) and combination of 2 custom-designed sgRNAs produced the highest knockdown (KD) efficacy relative to the control group (Fig. 2B) and was thus used in all subsequent KD experiments.

In vitro cell survival assays in form of Cell-counting Kit-8 showed that ZNF204P KD significantly inhibits cell survival (Fig. 2C) and number of colonies were drastically reduced in KD groups (Fig. 2D). Next, we examined the effect of ZNF204P KD on tumor cell migration through Transwell and wound healing assays. Following knockdown of ZNF204P, both migratory (Fig. 3A) and invasive (Fig. 3B) abilities of HepG2 were significantly weakened compared to control groups. ZNF204P regulation over cell migration was further substantiated when inhibited wound closure presented in the knockdown group relative to the control group (Fig. 3C). Altogether, our data presents ZNF204P as an oncogenic IncRNA that plays a critical role in cell proliferation, self-renewal and migration/invasion in HepG2 cells.

ZNF204P as a predicted miRNA-145 sponge

Fractionated RT-qPCR with MALAT1 and DANCN as positive controls established ZNF204P as a cytosolic IncRNA (Fig. 4A). Since the main mechanism of action of pseudogene- IncRNAs is through competing endogenous RNAs (ceRNAs), whereby IncRNAs compete for miRNAs that bind to cognate or unre-

lating genes, we first examined the relationship between ZNF204P and its parental gene ZNF79. As no shared miRNA sites were detected (Supplementary Fig. 3A, Supplementary Table 4), we investigated the mode of function of ZNF204P in relation to well-known regulators of stemness in both normal and cancer stem cells – OCT4, SOX2 and Nanog. miRNA-145-5p stood out within the predicted 7 shared miRNAs between ZNF204P and different combinations of OCT4/SOX2/Nanog (Fig. 4B, Supplementary Table 4) as its tumor suppressor role by degrading OCT4 and SOX2 has been covered previously (20, 21). Based on confirmation of down-regulated expression of miRNA-145-5p in TCGA-LIHC tumor samples (Fig. 4C) and corresponding decrease in OCT4 and SOX2 in ZNF204P-inhibited groups (Fig. 4D), we postulated that ZNF204P could function as a pluripotency regulator within cancer stem cells by decoying tumor suppressive miRNAs such as miRNA-145-5p. Collectively, our results show that ZNF204P is tumor-specific and exerts its oncogenic function through cell proliferation and migration, with possible function as competing endogenous RNA for tumor-suppressor miRNA-145-5p.

DISCUSSION

Outstanding issues despite the establishment of CSC model are i.) refining isolation and identification of CSCs through extracellular/intracellular factors and ii.) unraveling the complex mechanism of how such factors dysregulate self-renewal and
ZNF204P as an oncogenic lncRNA in HCC
Ji-Hyun Hwang, et al.

extensive proliferation in CSCs. Extracellular markers such as EpCAM, CD90 and CD133 have successfully been used to isolate CSCs, while transcription factors such as OCT4, SOX2 and Nanog are known tumorigenesis-driving players. Notably, such CSCs-specific markers are also highly functional in normal stem cells. Our study thus proposes a user-friendly method in identifying candidates through public data.

Here, we demonstrate the selection of stemness-related factors by identifying up-regulated genes in ESC and HCC samples respectively in microarray data via GEO2R. We reasoned that overlapping up-regulated genes from these two groups would shortlist factors conferring stemness properties by permitting continuous self-renewal and proliferative capabilities. Indeed, GO analysis of the shortlisted 139 factors displayed involvement in tumor-promoting signaling pathways and tumor cell growth. Moreover, previously characterized tumorigenesis-driving transcription factors such as MYCN and MYB were identified, indicating the robustness of our proposed method.

LncRNAs hold rich potential as therapeutic targets and prognostic markers as numerous studies indicate their contribution to self-renewal and proliferation of HCC CSCs. CUDR promotes malignant proliferation through multiple signaling pathways (22) while HOTAIR suppresses SETD2 in HCC CSCs (23). Thus, we selected the pseudogene-derived IncRNA ZNF204P for further investigation. To the best of our knowledge, our study is the first to publish experimental validation of ZNF204P, though it has been implicated as prognostic markers in various cancers (24) and a schizophrenia susceptibility allele (25). Given the complexities in perturbing IncRNAs (26), we opted to use CRISPRi/dCas9-KRAB system to block ZNF204P transcription, which entails proper annotation of transcription start sites (TSS). FANTOM CAT catalogues high-confidence TSSs through CAGE technique by capturing and single-molecule sequencing 5’ends of messenger RNAs (mRNAs) (27). Though with different transcription initiation evidence scores (TISEcore, Supplementary Table 2), we found that most of the ZNF204P transcripts corresponded to the same CAGE cluster, and thus designed sgRNAs +/− 200 relative to the defined TSS.

ZNF204P absence impaired cell survival while its knockdown reduced migratory/invasive properties. With ZNF204P up-regulated in TCGA-LIHC tumor samples, poor survival and enrichment of gene sets that drive cell proliferation and stem cell maintenance in ZNF204P-high group, we deduced that ZNF204P is a stemness-associated IncRNA with oncogenic capabilities. Furthermore, cytosolic ZNF204P shares a binding site with OCT4 and SOX2 for miRNA-145-5p, a well-studied tumor suppressor in cancer stem-like cells (20, 21). ZNF204P knockdown displays a corresponding decrease in OCT4 and SOX2, suggesting a miRNA-145-5p/stemness-regulating transcription factor axis. We acknowledge that more in-depth studies are required to substantiate our stand, such as direct perturbation of miRNA-145-5p to examine its effect on HepG2 cells, but present our preliminary findings on ZNF204P as evidence of a simple yet comprehensive method to derive stemness-related factors.

In conclusion, we have identified a stemness-related factor ZNF204P, a pseudogene-derived lncRNA that modulates cell proliferation and migration as an oncogene, and possibly functions as a decoy of tumor-suppressive miRNA-145-5p by interfering with its degradation of OCT4 and SOX2 in HepG2 cells. With our findings, we propose that such oncogenic lncRNAs provide another avenue to interrogate gene circuits that control stemness within tumor mass, enabling selection of appropriate candidates as therapeutic targets.

MATERIALS AND METHODS

Data acquisition and analysis
GSE25097 microarray expression data was downloaded from HCCDB (28) to select HCC samples that displayed combinatorial elevated expression (> HCC average) of cancer stemness markers EpCAM, Sall-4, Sox12 and BMI-1 (Supplementary Table 1). GSE50206 (3 ESC/3 HDF samples) (29) and GSE25097 (7 HCC/6 Healthy samples) (30) microarray datasets were analyzed on GEO2R available at http://www.ncbi.nlm.nih.gov/geo/geo2r/ to derive differentially expressed genes. The following filters were used to list significantly dysregulated genes (Supplementary Table 1): P-value < 0.05, log2 FC > 3 (GSE50206), log2 FC > 0.5 (GSE25097). Resulting genes were overlapped to derive stemness-related candidates, and GO analysis was conducted on Enrichr (31).

Mean differential expression of selected genes (ZNF204P, MYCN, MYB) between normal and tumor sets were visualized using TCGA-LIHC data, and Kaplan-Meier survival plots were taken from UALCAN portal (32).

Processed RNA-seq V2 for LIHC (HCC) was downloaded from public TCGA data portal, and the dataset contains 422 samples (50 normal samples, 372 tumor samples). Fragments Per Kilobase of exon per Million (FPKM) reads were used.

TIE scores for genomic region spanning from chromosome 6:27324485 to chromosome 6:27440849 (ZNF204P) on FANTOM CAT were downloaded for quantification (Supplementary Table 2). Association between miRNAs and target genes (OCT4, SOX2, Nanog, ZNF79, ZNF204P) were queried on mirSystem (33). Differential expression of resulting candidate miRNA-145-5p were visualized through UALCAN for TCGA-LIHC normal/tumor samples (49 normal, 369 tumor) (32).

Gene set enrichment analysis (GSEA)

GSEA software (version 4.1.0) was downloaded available at http://www.broadinstitute.org/gsea/downloads.jsp and expression data sets (.TXT format), phenotype labels (.CLS format) and gene sets (.GMX format) were created in accordance to GSEA specifications. Signature gene sets "Chiang Liver Cancer Subclass Proliferation Up" and "GOBP Regulation of Stem Cell Populations Maintenance" were downloaded from Molecular Signatures Database (MsigDB v7.4) (34). Input for basic fields were: Enrichment statistic = weighted, Metric for ranking genes = Difference of class, Number of permutations = 1,000.
Cloning strategies for CRISPRi-mediated knockdown

sgRNAs targeting the regions encompassing transcription start site (+/−200 relative to TSS) of ZNF204P was designed using CHOPCHOP v3 (35). The top and bottom strands of each targeting oligonucleotides were phosphorylated and annealed in thermocycler under settings: 37°C for 30 minutes, 95°C for 5 minutes, with ramping down to 25°C at a rate of 5°C/min. sgRNA sequences are included in Supplementary Table 1. Lenti_sgRNA_EFS_GFP (Addgene, #65656) vector (LRG) was digested with BsmBI (NEB) at 55°C for 1 hour, and subsequently annealed with each respective sgRNA duplex.

Cell culture and transfection

HCC cell-line HepG2 was maintained in DMEM (Invitrogen, USA) supplemented with fetal bovine serum 10% (Crytva, SH30084.03), 1% penicillin-streptomycin (Gibco, #15140122) in 37°C in 5% CO2 humidified incubator. Cells were passaged every 2-3 days and harvested using 0.25% Trypsin-EDTA (Invitrogen, #25200-056). For knockdown studies, 1 × 10^5 cells were seeded per well in a 24-well plate one day prior to transfection. sgRNA-encoding LRG (600 ng) and modified dCas9-KRAB-MeCP2 (300 ng) were mixed with 2 μl of Lipofectamine 3000 (Thermo Fisher Scientific, #1300015) and incubated at room temperature for 5 minutes. Transfection mixture was added to each well, and cells were harvested for downstream analysis at 24th and 48th hour time-points.

Cell viability assay

100 μl of cell media from each well (duplicates) in 24-well plate was collected and transferred to a 96-well plate at the 24th and 48th hour post-perturbation. 10 μl of CCK-8 solution (Dojindo, CK04-13) was added to each well, and incubated for 2 hours at 37°C, 5% CO2. The absorbance was read at 450 nm using TECAN microplate reader (Infinite 200 PRO), and the average value of duplicates were taken for reading of each group.

Migration and invasion assays

100 μl of cell suspension (DMEM-1% penicillin/streptomycin with no serum) containing 5 × 10^5 cells subjected to CRISPRi for 24 hours were dispensed into the Transwell insert (Corning, 3422) and incubated for 10 minutes at 37°C and 5% CO2. 600 μl of media solution (DMEM-1% penicillin/streptomycin with fetal bovine serum) was added to the bottom of lower chamber. 48 hours post-incubation, inserts were removed from plate for further analysis. For imaging, remaining cells from the membrane top were wiped off using a cotton-tipped applicator. Inserts were subsequently fixed in ice-cold 100% methanol for 15 minutes, washed in PBS and stained with 1% crystal violet solution (Sigma Aldrich) for 30 minutes at room temperature. After rinsing in water, inserts were left to dry, and images were taken at 10× magnification using Olympus IX51 light microscope. For crystal violet quantification, cut-out insert membranes were solubilized in 250 μl of 1% sodium dodecyl sulfate solution for 10 minutes at room temperature. 50 μl (in triplicates) were transferred to 96-well plate for OD reading at 595 nm.

The same steps were repeated for invasion assays, with the exception of using inserts pre-coated with Matrigel (Corning, 354234) at diluted concentration of 200 μg/ml.

Wound-healing assay

2 × 10^5 cells were seeded per well in 24-well plates pre-coated with 0.01% Poly-L-lysine (Sigma, P2636-25MG) one day prior to transfection. ‘0 hour’ time-point was established at 6 hours post-transfection, and each wound was made using a 200 μl pipette tip. Images were taken at 10x magnification using Olympus IX51 light microscope at stated time-points. Wound area was quantified through ImageJ available at https://imagej.nih.gov/ij/download.html using the polygon selection mode.

Colony formation assay

Transfected 2 × 10^3 cells per well were seeded in 6-well plates, and incubated for 10-14 days. All groups underwent PBS wash, 30 minute fixation with ice-cold 100% methanol, staining (0.5% crystal violet solution) for 2 hours at room temperature, and air-drying for at least 2 hours following washes with tap water. Images were taken at 10x magnification using Olympus IX51 light microscope at stated time points.

RNA isolation and quantitative real-time PCR

0.5-1 μg of RNA was used in synthesis of 1st strand complementary DNA (cDNA) using Superscript II reverse transcriptase (Thermo Fisher Scientific, #18064014) and oligo (d)T primers (Macrogen). The resulting cDNA was diluted × 10 in nuclease-free water and stored at −20°C. Quantitative PCR (qPCR) was carried out in triplicates using KAPA SYBR® FAST Bio-Rad iCycler 2X Qpcr Master Mix (Roche, KK4608). For RNA fractionation, 2 × 10^6 HepG2 cells were harvested for fractionated RNA isolation using PARIS kit (Ambion, AM1921), as per manufacturer’s protocol. All qPCR primer sequences are included in Supplementary Table 3.

ACKNOWLEDGEMENTS

This work was supported by the Collaborative Genome Program for Fostering New Post-Genome Industry through the National Research Foundation of Korea (NRF), funded by the Ministry of Science and ICT (NRF-2016M3C9A4921712) and by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education (NRF-2016R1D1A1B01015292).

AUTHOR CONTRIBUTIONS

J.-H.H and J.L performed the experiments, J.W.L and K.H.B.C analyzed the data. W.Y.C and M.J.K contributed to data interpretation and manuscript writing. J.-H.H, L.K.K and Y.-J.K conceived the study and wrote the manuscript. All authors read and approved the final manuscript.
CONFLICTS OF INTEREST

The authors have no conflicting interests.

REFERENCES

1. European Association for the Study of the Liver (2018) EASL Clinical Practice Guidelines: management of hepatocellular carcinoma. J Hepatol 69, 182-236
2. Nio K, Yamashita T, and Kameko S (2017) The evolving concept of liver cancer stem cells. Mol Cancer 16, 4
3. Ginestier C, Hur MH, Charafe-Jauffret E et al (2007) ALDH1 is a marker of normal and malignant human mammary stem cells and a predictor of poor clinical outcome. Cell Stem Cell 1, 555-567
4. Yan X, Shi P, Zhao G et al (2020) Targeting cancer stem cell pathways for cancer therapy. Signal Transduct Target Ther 5, 8
5. Kim WT and Ryu CJ (2017) Cancer stem cell surface markers on normal stem cells. BMB Rep 50, 285-298
6. Shi G and Jin Y (2010) Role of Oct4 in maintaining and regaining stem cell pluripotency. Stem Cell Res Ther 1, 39
7. Zhang H and Wang ZZ (2008) Mechanisms that mediate stem cell self-renewal and differentiation. J Cell Biochem 103, 709-718
8. Chen S, Zhu J, Wang F et al (2017) LncRNAs and their role in cancer stem cells. Oncotarget 8, 110685-110692
9. Fatica A and Bozzoni I (2014) Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 15, 7-21
10. Statello L, Guo CJ, Chen LL, and Huarte M (2020) Comprehensive genomic characterization of long non-coding RNAs across human cancers. Cancer Cell 28, 529-540
11. Statello L, Guo CJ, Chen LL, and Huarte M (2020) Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet 15, 7-21
12. Xie Z, Bailey A, Kuleshov MV et al (2021) Gene set knowledge discovery with enrichr. Curr Protoc 1, e90
13. Chandrashekar DS, Bashel B, Balasubramanya SAH et al (2017) UALCAN: a portal for facilitating tumor subgroup characterization (CRISPR) and CRISPR activation (CRISPRa) to explore the oncogenic IncRNA network. Methods Mol Biol 2348, 189-204
14. Hon GC, Ramilowski JA, Hashibarger J et al (2017) An atlas of human long non-coding RNAs with accurate 5′ ends. Nature 543, 199-204
15. Liu Q, Wang S, Zhang G et al (2018) HCCDB: a database of hepatocellular carcinoma expression atlas. Genomics Proteomics Bioinformatics 16, 269-275 (Available from: http://lifeome.net/database/hccdb/download.html)
16. Takahashi K, Tanabe K, Ohnuki M et al (2014) Induction of pluripotent stem cells derived from human somatic cells via a transient state resembling primitive streak-like mesoderm. Nat Commun 5, 3678
17. Lamb JR, Zhang C, Xie T et al (2011) Predictive genes in adjacent normal tissue are preferentially altered by sCNV during tumorigenesis in liver cancer and may rate limiting. PLoS One 6, e20090
18. Xie Z, Bailey A, Kuleshov MV et al (2021) Gene set knowledge discovery with enrichr. Curr Protoc 1, e90
19. Liberzon A, Birger C, Thorvaldottir H, Ghandi M, Mesirov JP, and Tamayo P (2015) The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst 1, 417-425
20. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, and Valen E (2019) CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res 47, W171-W174

ZNF204P as an oncogenic IncRNA in HCC
Ji-Hyun Hwang, et al.

http://bmbreports.org

http://bmbreports.org