LncSox4 promotes the self-renewal of liver tumour-initiating cells through Stat3-mediated Sox4 expression

Zhen-zhen Chen¹⁺, Lan Huang²⁺, Ya-hong Wu¹, Wen-jie Zhai¹, Ping-ping Zhu³ & Yan-feng Gao¹,⁴

Liver cancer has a tendency to develop asymptptomatically in patients, so most patients are diagnosed at a later stage. Accumulating evidence implicates that liver tumour-initiating cells (TICs) as being responsible for liver cancer initiation and recurrence. However, the molecular mechanism of liver TIC self-renewal is poorly understood. Here we discover that a long noncoding RNA (lncRNA) termed LncSox4 is highly expressed in hepatocellular carcinoma (HCC) tissues and in liver TICs. We find that LncSox4 is required for liver TIC self-renewal and tumour initiation. LncSox4 interacts with and recruits Stat3 to the Sox4 promoter to initiate the expression of Sox4, which is highly expressed in liver TICs and required for liver TIC self-renewal. The expression level of Sox4 correlates with HCC development, clinical severity and prognosis of patients. Altogether, we find that LncSox4 is highly expressed in liver TICs and is required for their self-renewal.
Epithelial carcinoma (HCC) is the most common type of liver cancer, the third leading cause of tumour-related death. Like other tumours, HCC tissues contain a heterogeneous mixture of cells. Tumour-initiating cells (TICs), a small subset within tumour bulk, are considered to be the source of tumours. Liver TICs are responsible for liver cancer formation, metastasis, drug-resistance and recurrence. Recently, several liver TIC markers have been identified, including EPCAM, CD133, CD13, CD90, CD24 and calcium channel subunit 

In addition, several signalling pathways are known to be involved in liver TIC self-renewal, including Notch, Wnt/β-catenin, Hedgehog and TGF/β signalling pathways. Given the critical role of TICs in tumorigenesis, intervention of TIC self-renewal can be a potential strategy for tumour treatment. However, the biological characters and molecular mechanisms of liver TICs formation and maintenance are still poorly understood.

Long noncoding RNAs (lncRNAs) are defined as a type of RNA molecules that are longer than 200 nucleotides (nt) without protein-coding potential. Due to the lack of corresponding proteins, lncRNAs were once considered as byproducts of RNA polymerase II, without biological functions. Recently the critical roles of lncRNAs in physiological and pathological processes have been frequently discovered. LncRNAs are involved in cell and organ development, stemness maintenance, cell differentiation and X chromatin inactivation. In addition, lncRNAs also participate in tumour proliferation, metastasis, invasion, energy regulation and TIC self-renewal. Mechanistically, lncRNAs exert their roles through cis-regulation or trans-regulation. LncRNAs can interact with chromatin DNA, mRNA or protein, thus regulating chromatin accessibility, mRNA stability and protein activity or stability. Though more and more lncRNAs have been discovered and critical roles of lncRNAs emerge, the biological function and molecular mechanism of lncRNAs are largely unknown.

Transcription factors, especially several stem factors have been reported in liver TIC regulation, including Oct4, Nanog, c-Myc and Zic2. Stat3, a well-known transcription factor, plays critical roles in development, inflammation, stemness regulation and tumorigenesis. Sox4, a Sox family transcription factor, is thought to be required for nervous system development, epigenetic reprogramming, stemness regulation and tumorigenesis. However, the function of Stat3 and Sox4 in liver tumorigenesis and liver TIC self-renewal is largely unknown.

Here we discovered a long noncoding RNA termed LncSox4 initiates liver TIC self-renewal through Stat3–Sox4 pathway. LncSox4 interacts with and recruits Stat3 to Sox4 promoter, initiating Sox4 expression and liver TIC self-renewal. We proved that the Stat3–Sox4 pathway participates in liver tumorigenesis and liver TIC self-renewal, thus offering a new potential target of TICs for eradicating liver cancer.

**Results**

**LncSox4 is highly expressed in liver cancer and liver TICs.** To understand how liver TICs are formed and maintained during tumorigenesis, we analysed several online-available data sets to understand how liver TICs are formed and maintained during liver tumorigenesis and liver TIC self-renewal, thus offering a new potential target of TICs for eradicating liver cancer.

Top 100 lncRNAs were selected and a heatmap was generated (Fig. 1a). We also analysed another patient cohort, Li’s cohort (GSE40144, microarray data of genes and lncRNAs, with disease-free survival and overall survival information of 59 hepatocellular carcinoma), and found hundreds of lncRNAs related to HCC severity and prognosis. We chose top 50 lncRNAs highly expressed in advanced samples and top 50 lncRNAs highly expressed in HCC patients with poor prognosis (Fig. 1b). On the basis of the overlap of these selections, we finally focused on 10 lncRNAs that are highly expressed in liver cancer and also related to both HCC severity and prognosis. To examine the effects of these lncRNAs on TIC self-renewal, we established lncRNA-silenced cells using pSiCoR lentivirus, followed by sphere formation, a standard approach to detect TIC self-renewal. We found the weakest sphere formation capacity of AL365371 (hereafter termed LncSox4) depleted cells, indicating the critical role of LncSox4 in liver TIC self-renewal (Fig. 1c). Furthermore, LncSox4 is highly expressed in HCC tissues (Fig. 1d), especially advanced HCC patients (Fig. 1e). Importantly, patients with high LncSox4 expression levels had poorer outcomes, and vice versa (Fig. 1f). We then examined the expression levels of LncSox4 using 24 peri-tumour, 12 early HCC and 12 advanced HCC samples by real-time PCR (Fig. 1g). Furthermore, in situ hybridization (ISH) assays were performed using 30 peri-tumour, 12 early HCC and 18 advanced HCC samples (Fig. 1h and Supplementary Fig. 1A), confirming high LncSox4 expression in liver tumours, especially in advanced liver tumours. We also observed similar results using northern blot (Fig. 1i). Above all, LncSox4 is highly expressed in liver tumours and related to clinical severity and prognosis.

Considering the critical role of LncSox4 in liver TIC self-renewal, we next examined LncSox4 expression levels in liver TICs. First, we analysed the expression profiles of LncSox4 and liver TIC surface markers. The samples were divided into two groups according to the expression levels of two widely accepted liver TIC surface markers, EPCAM and CD133 (mean expression levels of EPCAM and CD133 served as cut-off values, respectively), and then LncSox4 expression levels were analysed. Higher expression levels of LncSox4 were found in EPCAMhigh and CD133high samples, suggesting that LncSox4 is highly expressed in liver TICs (Supplementary Fig. 1B). The positive correlation between LncSox4 and TIC markers was confirmed in primary HCC samples, indicating that LncSox4 is related to liver TICs (Supplementary Fig. 1C). CD133 was used to enrich liver TICs and non-TICs. CD133+ cells showed vigorous sphere formation and tumour initiation capacities, conferring CD133+ cells as liver TICs (Supplementary Fig. 1D,E and Supplementary Table 1A). Then we enriched liver TICs from primary HCC samples by flow cytometry, detected LncSox4 expression levels in CD133 (non-TICs) and CD133+ (TICs) cells, and found higher expression of LncSox4 in liver TICs (Fig. 1j, upper panels). To determine LncSox4 expression in stem-like oncospheres and non-spheres derived from primary samples, we performed sphere formation assays using primary HCC cells and measured the expression levels of LncSox4 by real-time PCR, confirming the high expression of LncSox4 in stem-like oncospheres (Fig. 1j, lower panels). We verified the expression profiles of LncSox4 using fluorescence in situ hybridization (FISH) and confirmed high expression levels of LncSox4 in oncospheres (Fig. 1k). Finally we analysed LncSox4 expression in liver TICs using northern blot. Three samples were used for sphere formation and followed by northern blot, validating the high expression levels of LncSox4 in liver TICs (Fig. 1l). Another lncRNA (BC016831) showed comparable expression levels in liver TICs compared with non-TICs, and spheres compared with non-spheres, indicating specific LncSox4 expression in liver TICs (Supplementary
LncSox4 in liver cancer was confirmed by primary HCC samples. The indicated HCC samples were used for analysis of LncSox4 expression (Fig. 1F). To determine the location of LncSox4 in liver TICs, we performed nucleocytoplasmic separation using stem-like oncospheres, and examined the content of LncSox4 in nuclear and cytoplasmic fractions using real-time PCR (Supplementary Fig. 1G). The results showed that LncSox4 is located in the cell nucleus, consistent with the ISH results in Fig. 1h and FISH.

**Figure 1** | LncSox4 is highly expressed in liver cancer and liver TICs. (a) Expression profiles of lncRNAs in HCC samples were analysed using R language and Bioconductor according to Wang’s cohort. NL, normal liver; IL, inflammatory liver; CL, cirrhosis liver; eHCC, early HCC; aHCC, advanced HCC. (b) Li’s cohort was used to analyse lncRNAs that are related to clinical severity and prognosis. (c) LncSox4 plays an essential role in liver TIC self-renewal. 10 lncRNAs that are highly expressed in liver cancer and related to HCC severity and prognosis were selected. (d) High expression of LncSox4 in liver cancer. LncSox4 expression was analysed with R language and Bioconductor, and expression levels of LncSox4 were shown as mean ± s.d. (e,f) LncSox4 expression levels were related to clinical stages (e) and prognosis (f) according to Li’s cohort. Samples were divided into different groups according to clinical severity (e) or LncSox4 expression level (f), followed by expression analysis (e) or Kaplan–Meier survival analyses (f). (g,i) High expression of LncSox4 in liver cancer was confirmed by primary HCC samples. The indicated HCC samples were used for LncSox4 expression analysis using real-time PCR (g), *in situ* hybridization (h) and northern blot (i). (j) Histogram (mean ± s.d.) of LncSox4 in liver TICs and non-TICs. CD133+ liver TICs were sorted (upper panel) and oncospheres were derived (lower panel) from the indicated primary samples, followed by real-time PCR assays for LncSox4 transcript #1, primary sample #1. (k,l) High expression of LncSox4 was confirmed by FISH (k) and northern blot (l). FITC-conjugated LncSox4 probes and anti-CD133 antibody were used for non-sphere and sphere staining, followed by counterstain with DAPI. N, Non-sphere; S, Sphere. Scale bars, H, 100 μm; K, 10 μm. Data are shown as means ± s.d. Two-tailed Student’s *t*-test was used for statistical analysis. *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001; #P < 0.05; ##P < 0.01; NS, not significant. P < 0.05 was considered significant. Data are representative of at least three independent experiments. For i and l, the original images are shown in Supplementary Fig. 6.
results in Fig. 1k. Altogether, LncSox4 is highly expressed in HCC and liver TICs.

**LncSox4 is essential for liver TIC self-renewal.** Considering high LncSox4 expression in liver TICs, we further analysed its role in liver TIC self-renewal. First, we established LncSox4-silenced primary HCC cells using pSiCoR lentivirus, and the knockdown efficiency was confirmed using real-time PCR (Fig. 2a, left panels). Spheres with diameter >100 μm were counted 2 weeks later. LncSox4-depleted cells showed impaired sphere formation capacity, indicating the critical role of LncSox4 in liver TIC self-renewal (Fig. 2a, middle and right panels). Notably, there was no obvious difference of sphere size between LncSox4-silenced and control spheres (Supplementary Fig. 2A). We further detected the role of LncSox4 in liver TIC self-renewal using serial sphere formation assays. Four generations of sphere formation were performed, showing impaired long-term self-renewal with LncSox4 knockdown (Fig. 2b). To further examine the role of LncSox4 in liver TIC self-renewal, we silenced LncSox4 in 30 HCC samples available, and found impaired self-renewal in 23 samples using sphere formation assays (Supplementary Fig. 2B). What is more, we co-stained the spheres with antibodies against OCT4, a widely used stemness transcription factor, and Ki67, a well-known proliferation marker, showing impaired stemness and proliferation in LncSox4-silenced spheres (Supplementary Fig. 2C).

Then 1 × 10^6 LncSox4-silenced cells and control cells were subcutaneously injected into BALB/c nude mice, and tumour volumes were measured at the indicated time points. The tumour growth curves and tumour pictures indicate that LncSox4 is required for in vivo propagation of liver cancer (Fig. 2c). To detect tumour initiation, 10, 1 × 10^2, 1 × 10^3, 1 × 10^4 and 1 × 10^5 cells were subcutaneously injected into BALB/c nude mice and tumour initiation was evaluated 3 months later. Tumour-free mice ratios and established tumours were shown in Fig. 2d. TIC ratios were calculated using extreme limiting dilution analysis. The results showed that LncSox4-silenced cells contain remarkably low ratios of liver TICs, suggesting the essential role

![Figure 2](https://example.com/figure2.png)

**Figure 2 | LncSox4 is required for liver TIC self-renewal.** (a) LncSox4 plays an essential role in sphere formation. LncSox4-silenced primary HCC cells were established (left panel) using pSiCoR lentivirus, followed by sphere formation assays. Representative sphere pictures were shown in middle panels and calculated TIC ratios were shown in the right panels. Scale bars, 500 μm. (b) Serial sphere formation of LncSox4-silenced (shLnc) and control (shCtrl) primary HCC cells. For each sphere formation assay, 5,000 primary HCC cells (the 1st generation sphere formation) or signal sphere cells (the second, third, fourth generations) were used. (c) LncSox4 is required for HCC propagation. A total of 1 × 10^6 LncSox4-silenced or control cells were subcutaneously injected into BALB/c nude mice on the back, and tumour volumes were measured at the indicated time points. The tumour pictures were shown in right panels. (d) LncSox4 participates in HCC initiation. 10, 1 × 10^2, 1 × 10^3, 1 × 10^4 and 1 × 10^5 LncSox4-silenced or control cells were subcutaneously injected into BALB/c nude mice on the back and tumour formation was observed three months later. Tumour-free mice ratios and tumour pictures were shown. (e,f) Sphere formation (e) and tumour-initiation (f) assays of LncSox4-silenced liver TICs. CD133^+ liver TICs and CD133^− non-TICs were enriched from primary HCC cells using flow cytometry, followed by infection with LncSox4-silenced and control pSiCoR lentivirus. For sphere formation, 5,000 cells were incubated in FBS-free medium for 2 weeks. For tumour initiation, 10, 1 × 10^2, 1 × 10^3, 1 × 10^4 and 1 × 10^5 cells were subcutaneously injected into BALB/c nude mice on the back for 3 months. Data are shown as means ± s.d. Two tailed Student’s t-test was used for statistical analysis. *P < 0.05; **P < 0.01; ***P < 0.001; NS, not significant. P < 0.05 was considered significant. Data are representative of three independent experiments.
of LncSox4 in tumour initiation (Supplementary Table 1B). To further examine the role of LncSox4 in liver TICs, liver TICs and non-TICs were enriched and infected with LncSox4-silencing short hairpin RNA (shRNA) and control shRNA lentivirus. Sphere formation assays showed impaired self-renewal of LncSox4-silenced TICs, while, in non-TICs, LncSox4 knockout had impaired effect on sphere formation, indicating specific role of LncSox4 in liver TICs (Fig. 2e). Similarly, tumour initiation assays also confirmed the critical role of LncSox4 in liver TICs (Fig. 2f, Supplementary Table 1C).

We further verified the role of LncSox4 in liver TICs using LncSox4 overexpression cells (Supplementary Fig. 2D, upper panels). LncSox4 overexpression enhanced sphere formation (Supplementary Fig. 2D, middle and lower panels), but had no obvious impact on sphere size (Supplementary Fig. 2E), confirming the essential role of LncSox4 in liver TICs. We also detected long-term self-renewal of liver TICs using serial sphere formation, and found enhanced sphere formation capacities of LncSox4 overexpressing cells, confirming the critical role of LncSox4 in liver TIC self-renewal (Supplementary Fig. 2F). Then 1 × 10^6 LncSox4 overexpressed and control cells were subcutaneously injected into BALB/c nude mice, further confirming the critical role of LncSox4 in tumour propagation and initiation (Supplementary Fig. 2G). In summary, LncSox4 is essential for liver TIC self-renewal, tumour initiation and propagation.

LncSox4 promotes liver TIC self-renewal through Sox4. To explore the molecular mechanism of LncSox4 in liver TICs, we established LncSox4-silenced cells to identify its target genes. It is widely accepted that NFkB, Wnt/β-catenin, Notch and Hedgehog signalling pathways play critical roles in TIC self-renewal. Therefore, we checked the role of LncSox4 in the activation of these signalling pathways. We examined the expression levels of NFkB target genes (Vegf, Bcl2l1, Hif1a, Birc5, Mmp2, Twist1), Wnt/β-catenin target genes (Myc, Tiam, Kiaa, Ccdn1, Ccdn2, Sox4, Fnn4, Tcf1), Notch target genes (Hes6, Hey1, Hes1, Nrap) and Hedgehog target genes (Gli1, Patched, Gli3); and found impaired expression of transcription factor Sox4 on LncSox4 knockdown (Fig. 3a). To rule out off-target effects of shRNA treatment, we designed two independent LncSox4 shRNAs to confirm the target genes. The genes (such as Tcf7) downregulated by only one LncSox4 shRNA are probably off-target genes. The impaired Sox4 expression was further confirmed by western blot (Fig. 3b). We then analysed the expression correlation between LncSox4 and Sox4 using online-available data sets. The samples were divided into two groups according to LncSox4 or Sox4 expression levels (mean expression levels served as cut-off values), and we found high Sox4 expression in LncSox4high samples, and vice versa (Supplementary Fig. 3A,B). The positive correlation between LncSox4 and Sox4 expression was also confirmed using primary HCC samples (Supplementary Fig. 3C). These data indicated that LncSox4 could drive Sox4 expression.

Next, we examined whether LncSox4 promotes liver TIC self-renewal through Sox4. We silenced Sox4 expression and found reduced self-renewal by sphere formation assays, indicating the critical role of Sox4 in liver TIC self-renewal (Supplementary Fig. 3D). Then we rescued Sox4 expression in LncSox4-silenced cells and examined sphere formation. Sox4 rescue remarkably neutralized the impaired sphere formation induced by LncSox4 depletion, indicating LncSox4 drives liver TIC self-renewal in a Sox4-dependent manner (Fig. 3c). What is more, 10^5, 1 × 10^5, 1 × 10^4 and 1 × 10^3 LncSox4-silenced and Sox4 rescued cells were subcutaneously injected into BALB/c nude mice and tumour initiation was examined 3 months later. We found Sox4 largely rescued the impaired tumour initiation capacities induced by LncSox4 knockdown (Fig. 3d, Supplementary Table 1D). In addition, we established Sox4 knockout cells using CRISPR/Cas9 approach and overexpressed LncSox4 in Sox4 knockout cells, followed by tumour initiation assays using gradient transplant model. In wild-type (WT) cells, LncSox4 overexpression significantly promoted tumour initiation, while LncSox4 overexpression failed to promote tumour formation in Sox4 knockout cells, validating the critical role of Sox4 in LncSox4-mediated liver TIC activation (Fig. 3e, Supplementary Table 1E).

To examine whether LncSox4 regulates Sox4 expression through combination with Sox4 promoter, ChIRP assay was performed and enrichment of Sox4 promoter regions was measured using real-time PCR. The result suggested the interaction between LncSox4 and Sox4 promoter (Fig. 3f). We also examined Cox-2, Birc5, Ascl2 and Tcf7 promoters, but did not find an obvious enrichment, confirming the specific binding between LncSox4 and Sox4 promoter. Meanwhile, the failure of LncSox4-promoter binding conferred these genes as ‘off targets’ of LncSox4 shRNAs (Supplementary Fig. 3E). In addition, we predicted the structure of LncSox4 using thermodynamic ensemble and minimum free energy methods. Loops usually play critical roles in lncRNA function, so we focused on loops and found several common loops using two prediction methods. Surprisingly, the nucleotide sequence of one loop is highly complementary to the LncSox4 binding region of Sox4 promoter (Supplementary Fig. 3F), indicating the combination of LncSox4 and Sox4 promoter through sequence complement. We then deleted the LncSox4 binding region of Sox4 promoter using CRISPR/Cas9 approach (Supplementary Fig. 3G,H), and examined the effect of LncSox4 on Sox4 expression and liver TIC self-renewal. LncSox4 positively regulates Sox4 expression and liver TIC self-renewal in WT cells. Interestingly, LncSox4 overexpression had no longer impacted on Sox4 expression (Fig. 3g) and sphere formation (Fig. 3h) in LncSox4 binding region-deleted cells, suggesting that the interaction between LncSox4 and Sox4 promoter is essential for LncSox4-triggered Sox4 expression and liver TIC self-renewal.

Here, we discovered that LncSox4 binds to Sox4 promoter via sequence complement, which plays an essential role in LncSox4 function. LncRNA can bind to gene promoter to regulate gene expression in cis, but the molecular mechanism of this regulation mode is rarely reported. By CRISPR/Cas9-mediated editing of Sox4 promoter, we proved that LncSox4 binding to Sox4 promoter through sequence complement. Using Sox4 overexpression/knockout systems and sphere formation assays, we also proved that this complementary sequence is required for LncSox4 function.

LncSox4 interacts with Stat3. To further identify the molecular mechanism and binding partners of LncSox4 in liver TICs, we performed RNA pull-down assays followed by silver staining. LncSox4 specific bands were identified by mass spectrum, and Stat3 were found to interact with LncSox4 (Fig. 4a). The samples were further analysed using western blot, confirming the interaction of LncSox4 and Stat3 (Fig. 4b). We also verified this result using RNA immunoprecipitation (RIP) and found LncSox4 could be enriched in Stat3 precipitates, further proving that LncSox4 interacts with Stat3 (Fig. 4c). The interaction between LncSox4 and Stat3 was also confirmed by double FISH. LncSox4 is highly expressed in TICs and oncospheres and located within the cell nucleus. Notably, obvious co-localization of LncSox4 and Stat3 was found in liver TICs and spheres (Fig. 4d and Supplementary Fig. 4A). In addition, the interaction region of LncSox4 toward Stat3 was identified using LncSox4 truncates, revealing the
Figure 3 | LncSox4 drives liver TIC self-renewal through Sox4. (a,b) LncSox4 is required for Sox4 expression. mRNA expression of the indicated genes in LncSox4-silenced (shLncSox4) and control cells (shCtrl) were examined using real-time PCR (a). Impaired Sox4 expression in LncSox4-silenced cells was confirmed by western blot (b). (c,d) LncSox4 drives liver TIC self-renewal through Sox4 expression. Sox4 expression was rescued using PBPLV lentivirus in LncSox4-silenced cells, and then sphere formation (c) and tumour initiation (d) assays were performed. Five thousand indicated primary HCC cells were used for sphere formation, and 10, 1 x 10^2, 1 x 10^3, 1 x 10^4 and 1 x 10^5 cells were subcutaneously injected into BALB/c nude mice for tumour initiation. (e) Primary HCC cells were treated with CRISPR/Cas9 lentivirus for Sox4 deficiency, followed by LncSox4 overexpression (oeLnc). The indicated cells were used for in vivo tumour initiation, and the ratios of tumour-free mice were calculated 3 months later. (f) LncSox4 binds to Sox4 promoter. RNA ChIP assay was performed and fold enrichment was examined using real-time PCR. (g,h) The interaction of LncSox4 and Sox4 promoter was required for Sox4 function. The Sox4 promoter region for LncSox4 binding (−3,647 ~ −3,537) was deleted using CRISPR/Cas9 approach (Sox4KO), followed by LncSox4 overexpression. The established cells were examined for Sox4 expression and sphere formation capacities. Six samples were examined and similar results were found. Scale bars, C, H 500 μm. Data were shown as means ± s.d. Two-tailed Student’s t-test was used for statistical analysis. *p<0.05; **p<0.01; ***p<0.001. P<0.05 was considered significant. Data are representative of three independent experiments. For b, the original images are shown in Supplementary Fig. 6. NS, not significant.

Finally, we examined the biological relevance of the interaction between LncSox4 and Stat3. We detected the protein levels of Stat3 and found comparable protein levels in LncSox4-silenced and control cells, indicating that LncSox4 has no effect on Stat3 stability (Supplementary Fig. 4B). To examine the role of LncSox4–Stat3 association in LncSox4 function, we overexpressed LncSox4 and LncSox4 truncate without Stat3 binding region, and found impaired function of LncSox4 truncate in sphere formation (Fig. 4h). Altogether, LncSox4 associates with Stat3.

LncSox4 is required for Stat3-induced Sox4 expression. We then investigated the role of Stat3 in LncSox4-induced Sox4 expression, we used shRNA to silence Stat3 in LncSox4-overexpressing cells and examined Sox4 expression. We found that Stat3 silencing significantly reduced Sox4 expression, indicating that Stat3 is required for LncSox4-induced Sox4 expression.
expression. First, we detected whether Stat3 binds to Sox4 promoter using ChIP assay, and found that a specific region of the Sox4 promoter was enriched upon Stat3 immunoprecipitation, indicating that Stat3 binds to Sox4 promoter and probably regulates Sox4 expression (Fig. 5a). To further characterize the role of Stat3 in Sox4 expression, we deleted Stat3 using CRISPR/Cas9 approaches and found impaired activation of Sox4 promoter upon Stat3 deficiency, confirming that Stat3 regulates Sox4 expression (Fig. 5b). Interestingly, Stat3 bound the same region of Sox4 promoter as LncSox4 (Figs 3f and 5a), implying the possible function of LncSox4 in the interaction between Stat3 and Sox4 promoter. We then performed ChIP assay with Stat3 antibody using LncSox4-silenced cells and found remarkably decreased enrichment of Sox4 promoter upon LncSox4 depletion, proving the essential role of LncSox4 in Stat3 binding to Sox4 promoter (Fig. 5c).

To further verify the function of LncSox4 in Stat3-mediated Sox4 expression, we overexpressed Stat3 in LncSox4-silenced cells, and examined the activity of Sox4 promoter, chromatin accessibility and Sox4 expression. The Stat3/LncSox4 binding region of Sox4 promoter was cloned into PGL3 luciferase reporter plasmid. We found that Stat3 overexpression enhanced the luciferase activity in WT cells. However, Stat3 overexpression could not enhance the Sox4 promoter activity in LncSox4 knockdown cells (Fig. 5d). We also performed ChIP assays with antibodies against H3K4me3 and H3K27ac, two histone modification markers widely used for promoter activation examination. We found enhanced activation of Sox4 promoter on Stat3 overexpression in WT cells but not in LncSox4-silenced cells (Fig. 5e). Similarly, Stat3 overexpression significantly increased Sox4 expression in WT cells but not in LncSox4-deficient cells (Fig. 5f). In summary, Stat3 has little effect on Sox4 promoter activity, transcription activation of Sox4 promoter and Sox4 mRNA expression in LncSox4-depleted cells, strongly proving that LncSox4 is essential for Stat3-induced Sox4 promoter activation and Sox4 mRNA expression.
Stat3 and Sox4 are required for liver TIC self-renewal. Since LncSox4 primes liver TICs by promoting Stat3-induced Sox4 expression, we then explored the role of Stat3 and Sox4 in liver TIC self-renewal. The relationship between Sox4 and TIC surface markers were analysed using Wang’s cohort (GSE14520 (ref. 31)), with gene microarray data and survival information of two normal liver samples, 228 peri-tumour and 239 liver cancer samples) and Li’s cohort (GSE40144). The samples were divided into two groups according to EPCAM expression (mean expression level used as cut-off value), and comparable activation levels were observed in LncSox4-silenced cells. Since LncSox4 is required for Stat3-induced activation of Sox4 promoter, we then compared sphere formation capacity of Stat3-deleted cells, indicating the essential role of Stat3 in liver TICs (Fig. 6a). In addition, Stat3 knockout spheres were similar in size to control spheres (Supplementary Fig. 5D). We also silenced Stat3 in 30 primary samples using piSiCoR approaches and found impaired sphere formation capacities in almost all primary HCC samples upon Stat3 depletion (Supplementary Fig. 5E). Finally, we subcutaneously injected $1 \times 10^5$ Stat3 knockout or WT cells into BALB/c nude mice, and observed attenuated tumour expansion upon Stat3 deficiency (Supplementary Fig. 5F). BALB/c nude mice were subcutaneously injected with the indicated number of Stat3-knockout or WT cells to evaluate tumour initiation 3 months later. The impaired tumour initiation capacity was found in mice injected with Stat3 knockout cells (Supplementary Fig. 5G). More importantly, LncSox4 overexpression failed to promote liver tumour initiation, indicating LncSox4 drives liver TIC self-renewal through a Stat3-dependent manner (Supplementary Fig. 5G).

We then compared sphere formation capacity of Sox4-deficient and WT cells, and found fewer spheres formed in Sox4-deleted cells (Fig. 6b), without obvious changes in the sphere size (Supplementary Fig. 5H), indicating the critical role of Sox4 in sphere formation. In addition, four passages of sphere formation were performed to detect the function of Sox4 in liver TIC self-renewal. WT cells showed enhanced sphere formation capacities.
Figure 6 | Stat3/Sox4 is required for liver TIC self-renewal. (a) Stat3 is required for liver TIC self-renewal. Impaired Sox4 expression (left panels) and sphere formation capacities (middle and right panels) were found in Stat3-deficient cells. (b) Sox4-deficient cells were established and sphere formation assays were performed, and impaired sphere formation was observed. (c) Sox4 is essential in liver TIC self-renewal. Serial sphere formation assays were performed using Sox4-deficient or control cells. For each sphere formation assay, 5,000 cells were used. Representative images were shown in the left panel and calculated TIC ratios were shown in the right panel. (d) Sox4 drives liver cancer propagation. 1 × 10⁶ Sox4-deleted primary HCC cells were subcutaneously injected into BALB/c nude mice, and tumour volume was measured at the indicated time points. (e) A total 10, 1 × 10², 1 × 10³, 1 × 10⁴ and 1 × 10⁵ Sox4-deficient primary HCC cells were subcutaneously injected into BALB/c nude mice, and tumour formation was examined 3 months later. (f,g) Sphere formation (f) and tumour initiation (g) assays of Stat3- or Sox4-silenced liver TICs. CD133⁺ liver TICs were enriched from primary HCC cells using flow cytometer, followed by infection with Stat3- or Sox4-silenced pSiCoR lentivirus. For f, 5,000 cells were incubated in sphere formation medium for 2 weeks. For g, 10, 1 × 10², 1 × 10³, 1 × 10⁴ and 1 × 10⁵ cells were subcutaneously injected into BALB/c nude mice for tumour-initiation assay. Scale bars, 500 μm (a, b and c). Data were shown as means ± s.d. Two-tailed Student’s t-test was used for statistical analysis. *P < 0.05; **P < 0.01; ***P < 0.001. P < 0.05 was considered significant. Data are representative of three independent experiments. For a and b, the original images are shown in Supplementary Fig. 6. NS, not significant.

in advanced passages, while Sox4 deficiency impaired sphere formation with fewer spheres obtained in advanced passages. These results demonstrate the critical role of Sox4 in liver TIC self-renewal (Fig. 6c). We also silenced Sox4 in primary HCC cells of 30 samples using pSiCoR lentivirus, and examined sphere-formation capacities. Sox4 knockdown reduced liver TIC self-renewal in 25 HCC samples (Supplementary Fig. 5i). Finally, we detected the role of Sox4 in HCC propagation and initiation. A total 1 × 10⁶ Sox4-deficient primary HCC cells were injected into BALB/c nude mice, and tumour volume was measured every 4 days. Sox4-deficient cells formed smaller tumours compared with WT cells, indicating the essential role of Sox4 in liver cancer propagation (Fig. 6d). To examine tumour initiation, 10, 1 × 10², 1 × 10³, 1 × 10⁴ and 1 × 10⁵ primary HCC cells were subcutaneously injected into BALB/c nude mice and tumour formation was observed 3 months later. We found Sox4-deficient cells initiate fewer tumours compared with the same number of WT cells. Tumour-free mice ratios (Fig. 6e) and TIC ratios (Supplementary Table 1f) proved that Sox4 plays an essential role in tumour initiation.

We then enriched liver TICs by flow cytometer using anti-CD133 antibody, and silenced Stat3 or Sox4 in liver TICs and
non-TICs using pSiCoR lentivirus, followed by sphere formation and tumour initiation assays. Stat3 or Sox4 knockdown in TICs largely impaired sphere formation, while Stat3 or Sox4 depletion had little impact on self-renewal in non-TICs (Fig. 6f). In addition, Stat3 or Sox4 depletion in TICs reduced the tumour initiation and TIC ratios, while Stat3 or Sox4 knockdown in non-TICs had little effect on tumour initiation (Fig. 6g, Supplementary Table 1G). It has been reported that Nanog as a Stat3 target gene plays a critical role in TICs, so we compared the role of Nanog and Sox4 in Stat3-mediated liver TIC self-renewal. We rescued Nanog and Sox4 in Stat3-deficient cells, and performed sphere-forming assays. Both Nanog and Sox4 served as target genes of Stat3 and played a critical role in Stat3-mediated liver TIC self-renewal (Supplementary Fig. 5). Above all, the Stat3–Sox4 signalling plays an indispensable role in liver TIC self-renewal, liver cancer initiation and propagation.

**Sox4 is related to HCC severity and prognosis.** Finally we investigated the clinical significance of Sox4 expression in HCC tumorigenesis and progression. We analysed Wang’s cohort (GSE14520) using R language and Bioconductor, and found high Sox4 expression in liver tumours compared with peri-tumour tissues (Fig. 7a). Furthermore, high expression levels of Sox4 were found in metastasis samples and recurrent samples, indicating the potential role of Sox4 in HCC metastasis and recurrence (Fig. 7b,c). HCC progression was studied using TNM, BCLC and potential role of Sox4 in HCC metastasis and recurrence found in metastasis samples and recurrent samples, indicating the HCC samples were divided into two groups according to Sox4 expression. Higher expression of Sox4 was found in advanced HCC patients according to TNM stages (Fig. 7d), BCLC stages (Fig. 7e) and CLIP stages (Fig. 7f), indicating that Sox4 is associated with the severity of HCC. High Sox4 expression in HCC samples with poor prognosis was also confirmed by another cohort (Fig. 7g). To examine the relationship between Sox4 expression levels and HCC prognosis, the HCC samples were divided into two groups according to Sox4 expression levels (mean expression level served as cut-off value), and Kaplan–Meier survival analysis was performed. Sox4-high patients had a poor prognosis, and Sox4-low patients could survive longer, proving the critical role of Sox4 in HCC prognosis (Fig. 7h).

We then confirmed these results using primary HCC samples. We detected Sox4 expression levels using real-time PCR (Fig. 7i), western blot (Fig. 7j) and immunohistochemistry (IHC) (Fig. 7k), and confirmed high Sox4 expression in liver cancer, especially in advanced HCC tissues. The cohort data and our own results derived from primary samples endow Sox4 with a critical role in liver tumour. Besides its role in stemness maintenance and development regulation, we uncovered the critical role of Sox4 in liver cancer initiation, progress and prognosis prediction. Taken together, LncSox4 recruits Stat3 to Sox4 promoter and drives the expression of Sox4, which is required for liver TIC self-renewal and related to HCC progression, serving as a new target for eradicating liver TICs and a new marker for HCC prognosis (Fig. 7l).

**Discussion**

Liver tumour-initiating cells (TICs) account for HCC heterogeneity and recurrence. Many surface markers are used to identify and enrich liver TICs. In this study, we used CD133, a putative surface marker, to isolate liver TICs. CD133 is widely used as a stem marker for neural stem cells, embryonic stem cells, haemopoietic stem cell, glioblastomas and so on. The shared surface markers between tumour-initiating cells and other stem cells indicate the presence of similar self-renewal mechanism in these stem cells. To determine the self-renewal potential of liver TICs, in vitro sphere formation and in vivo tumour initiation, two widely used TIC functional assays, were performed in this study. Using these methods, we discovered the essential role of LncSox4, Sox4 and Stat3 in liver TICs.

LncRNAs are important for many physiological and pathological processes, including tumorigenesis and stemness, so it is reasonable that lncRNAs participate in TIC self-renewal. The functions of lncRNAs are frequently reported, while, their roles in liver TIC self-renewal are largely unknown. To explore specific lncRNAs in liver TIC self-renewal, we analysed online-available data sets with R language and performed ‘unbiased’ screening, and discovered many lncRNAs highly expressed in liver tumours and related to clinical prognosis. To determine their roles in liver TIC self-renewal, we performed a shRNA-based function screening and found LncSox4 related to liver TIC self-renewal. Our results indicate that LncSox4 promotes liver TIC self-renewal.

Through structure prediction and sequence alignment, we found obvious sequence complement between LncSox4 loop and Sox4 promoter. The interaction between Sox4 promoter and LncSox4 was confirmed by ChIRP assay. LncRNAs can work through interaction with chromatin DNA, RNA and protein. Many lncRNAs have been reported to interact with proteins and affect their stabilities or activities. Several lncRNAs can bind to gene-promoter regions, while the molecular mechanisms of this binding are rarely reported. Using CRISPR-Cas9 technique, we proved that LncSox4 binds to Sox4 promoter through sequence complement. We also proved that LncSox4 promotes Sox4 expression and liver TIC self-renewal through the sequence complement between LncSox4 and Sox4 promoter. Actually, the molecular mechanism of the interaction between LncSox4 and Sox4 promoter is still unclear. It is likely that the binding of LncSox4 and Sox4 promoter occurs during cell replication when the DNA double strands separate into signal strand. DNA repair and other mechanisms that can trigger the accessibility of double strands are also possible to drive the sequence complement.

Transcription factors are critical for gene transcription and cell fate determination. Using RNA pull-down assays, we found LncSox4 interacts with Stat3. Here we found Stat3 binds to Sox4 promoter and Sox4 is a new target gene of Stat3. Notably, LncSox4 is required for Stat3 binding to Sox4 promoter. Stat3 promotes Sox4 expression in WT cells, but not in LncSox4-silenced cells. Through these mechanisms, LncSox4 drives Sox4 expression and liver TIC self-renewal. We also found Stat3–Sox4 signalling pathway is required for liver TIC self-renewal. It has been well documented that Sox4 is crucial in tumorigenesis and stemness, but its function in liver TICs is unclear. We enriched CD133+ liver TICs using flow cytometer, and identified that Sox4 is highly expressed in liver TICs. Sox4 deficiency impaired liver TIC self-renewal in sphere formation assays in vitro and tumour-initiation experiments in vivo. Using online-available cohorts and primary samples, we found Sox4 expression levels related to HCC progression and prognosis, indicating its clinical significance. We also found the transcription factor Stat3 is involved in liver TIC self-renewal. In summary, using sphere formation assays in vitro and tumour-initiation assays in vivo, we found that the LncSox4–Stat3–Sox4 pathway plays a critical role in liver TIC self-renewal and serves as a target for liver TICs eradication.

Altogether, we found LncSox4 is highly expressed in liver TICs and required for liver TIC self-renewal. LncSox4 binds to Sox4 promoter and recruits Stat3, driving Sox4 promoter activation and Sox4 expression through LncSox4-dependent manner. Sox4 is required for liver TIC self-renewal and can serve as a target for liver TIC eradication. Sox4 expression levels are also related to...
Figure 7 | Sox4 expression levels are related to HCC severity and prognosis. (a) High expression of Sox4 in liver cancer tissues according to Wang’s cohort. R language and Bioconductor were used for gene expression analysis derived from Wang’s cohort. (b) Sox4 is highly expressed in metastasis samples. Tumour samples were divided into two groups according to metastasis properties, and Sox4 expression levels were analysed. (c) Sox4 expression levels are related to recurrence. Recurrent and non-recurrent samples were selected and Sox4 expression levels were compared. (d-f) High expression of Sox4 in serious HCC samples. The samples were divided according to TNM stage (d), BCLC stage (e) and CLIP stage (f), and Sox4 expression levels were analysed. BCLC, Barcelona Clinic Liver Cancer; CLIP, Cancer of the Liver Italian Program; TNM, tumour node metastasis. (g) Increased Sox4 expression levels in liver tumorigenesis. aHCC, advanced HCC; CL, cirrhosis liver; eHCC, early HCC; IL, inflammatory liver; NL, normal liver. (h) Sox4 expression is related to HCC prognosis. The samples were grouped according to Sox4 expression levels and Kaplan–Meier survival analyses were performed. (i) Increased Sox4 expression in HCC was confirmed by real-time PCR (i), western blot (j) and IHC (k). Primary HCC samples were obtained and marked according to obtaining time. mRNA was extracted for real-time PCR (i), cell lysate was used for western blot (j), and formalin-fixed tissue sections were used for IHC (k). Scale bars, 100 μm. (l) Working model. In liver TICs, LncSox4 is highly expressed and binds to Sox4 promoter. LncSox4 recruits Stat3 to Sox4 promoter and drives Sox4 expression, which is required for liver TIC self-renewal and related to HCC severity and prognosis. For a–f, data were shown as box and whisker plot. Box, interquartile range (IQR); whiskers, 5–95 percentiles; horizontal line within box, median. For g, i and k, data were shown as means ± s.d. Two-tailed Student’s t-test was used for statistical analysis. *P < 0.05; **P < 0.01; ***P < 0.001. P < 0.05 was considered significant. For j, the original images are shown in Supplementary Fig. 6. NS, not significant.

Liver cancer progression and prognosis, providing a new potential marker for HCC diagnosis and prognosis.

Methods

Cells and reagents. Human 293T cells were purchased from ATCC and maintained in DMEM medium supplemented with 10% fetal bovine serum (FBS; Invitrogen), 100 μg ml⁻¹ penicillin and 100 μl⁻¹ streptomycin. Primary HCC cells were obtained from surgical specimens from the first affiliated hospital of Zhengzhou University with informed consent and institutional approval. Primary samples were numbered according to the time we obtained. Tissues were cut into small pieces and then digested with collagenase IV for 60 min at 37 °C. Primary HCC cells were obtained after centrifugation. All the experiments were approved by the Institutional Committee of Zhengzhou University. HCC clinical information was shown in Supplementary Table 2. The cell lines used in this study were not contaminated by mycoplasma. Anti-β-Actin (catalogue A1978, 1:1,000) antibody was purchased from Sigma-Aldrich. Anti-Sox4 (catalogue 17919-1-AP, 1:200) antibody was purchased from Proteintech Group. Anti-Oct4 (catalogue 2750, 1:200), anti-Stat3 (catalogue 9139, 1:200), H3K4Me3 (catalogue 9751, 1:500) and H3K27Ac (catalogue 4353, 1:500) antibodies were purchased from Cell Signaling Technology. Phycocyanin-conjugated CD133 antibody (catalogue 130-098-826, 1:100) was purchased from Miltenyi Biotec. HRP-conjugated secondary antibodies (1:1,000) were purchased from Santa Cruz Biotechnology. Fluorescein-conjugated secondary antibodies were obtained from Molecular Probes Life Technologies.

Other reagents used in this study were: DAPI (catalogue 28718-90-3, Sigma-Aldrich), PEG5000 (catalogue 175233-46-2, Sigma-Aldrich), EGF (catalogue 4535, 1:300) antibodies were purchased from Cell Signaling Technology. Phycocyanin-conjugated CD133 antibody (catalogue 130-098-826, 1:100) was purchased from Miltenyi Biotec. HRP-conjugated secondary antibodies (1:1,000) were purchased from Santa Cruz Biotechnology. Fluorescein-conjugated secondary antibodies were obtained from Molecular Probes Life Technologies.

Other reagents used in this study were: DAPI (catalogue 28718-90-3, Sigma-Aldrich), PEG5000 (catalogue 175233-46-2, Sigma-Aldrich), EGF (catalogue 4535, 1:300) antibodies were purchased from Cell Signaling Technology. Phycocyanin-conjugated CD133 antibody (catalogue 130-098-826, 1:100) was purchased from Miltenyi Biotec. HRP-conjugated secondary antibodies (1:1,000) were purchased from Santa Cruz Biotechnology. Fluorescein-conjugated secondary antibodies were obtained from Molecular Probes Life Technologies.
Cohort analysis. Online-available data sets were downloaded from EBI (http://www.ebi.ac.uk) or NCBI (http://www.ncbi.nlm.nih.gov/gds/?term =). R language and Bioconductor were used for background correction, normalization, expression calculation and annotation29. The gene expression profiles were evaluated for further analyses using Excel, SPSS or Graphpad prism 5 software.

Three online available cohorts were used in this paper. Wang's cohort (GSE4234) shows microarray data of LncRNAs and genes, containing 10 normal liver samples, 10 inflammation liver samples, 10 cirrhosis liver samples, 13 early HCC samples and 13 advanced HCC samples. Li's cohort (GSE40144) has microarray data of genes and LncRNAs, with disease-free survival and overall survival information of 59 hepatocellular carcinoma. Wang's cohort (GSE14520) has gene microarray data and survival information of two normal liver samples, 228 peri-tumour and 239 cancer liver samples.

Nucleocytoplasmic separation. A total of 5 × 10⁶ oncosphere cells were resuspended in 0.5 ml resuspension buffer (10 mM HEPES, 1.5 mM MgCl₂, 10 mM KCl, 0.2% N-oxycetylglucoside, Protease inhibitor cocktail, RNA inhibitor, pH 7.9) for 10 min incubation, followed by homogenization. The cytoplasmic fraction was the supernatant after centrifugation (400g × 15 min). The pellet was resuspended in 0.2 ml phosphate-buffered saline (PBS), 0.2 ml nuclear isolation buffer (40 mM Tris-HCl, 20 mM MgCl₂, 4% Triton X-100, 1.28 M sucrose, pH 7.5) and 0.2 ml RNase-free H₂O, followed by 20 min incubation on ice to clean out the residual cytoplasmic faction. The pellet was nuclear fraction after centrifugation. RNA was extracted from nuclear and cytoplasmic fractions using RNA extraction kit (Tiangen Company, Beijing). LncSox4 content was examined by real-time PCR (ABI7300).

For LncSox4 content, standard reverse transcription was performed using reverse transcription kit (Promega). Notably, same amount of RNA and same volume of cDNA were required. In our experiment, 1 μg nuclear RNA and 1 μg cytoplasmic RNA were used, with the same final volume of nuclear and cytoplasmic cDNA (50 μl). Real-time PCR was performed using 1 μl nuclear cDNA or 1 μl cytoplasmic cDNA, with the same primers and ABI7300 profile. The relative LncSox4 contents were calculated using these formulae: nuclear ratio = 2^-ΔCt(nuclear)/(2^-ΔCt(nuclear) + 2^-ΔCt(cytoplasmic)); cytoplasmic ratio = 2^-ΔCt(cytoplasmic)/(2^-ΔCt(nuclear) + 2^-ΔCt(cytoplasmic)).

Flow cytometry. Primary HCC cells were incubated with phycoerythrin-conjugated anti-CD133 or control antibodies for 30 min at 4°C. Followed by washing twice in PBS supplemented with 1% FBS. The CD133⁺ and CD133⁻ cells were sorted using FACSariaII cytometer (BD).

Northern blot. Total RNA was extracted from primary HCC cells or oncospheres according to TRIZOL methods (Invitrogen) for northern blot. LncRNA and 18S fragments were cloned into PCDNA4 plasmid and northern probes were produced using Biotin RNA Labeling Mix (catalogue 20158, Thermo Scientific) for transcription using T7 RNA polymerase and Biotin RNA Labeling Mix (catalogue 11685597910, Roche). The following processes were similar to the knockdown protocol. For Sox4 promoter deficiency, a pair of guide RNAs was used simultaneously. The infected HCC cells were treated with puromycin and then monoclonalization was performed. The deletion of Sox4 promoter was confirmed by DNA sequencing.

FISH analysis. For fluorescence in situ hybridization (FISH), fluorescence-conjugated probes were designed by Biosearch Technologies and purchased from Life Technologies. The samples were treated with non-denaturing conditions, followed by adding fluorescence-conjugated probes. For double FISH, probes and secondary antibodies were added. The samples were counterstained with DAPI and observed using confocal microscopy.

Mouse models. Six-week-old male BALB/c nude mice were purchased from the HFK Biosciences maintained under specific pathogen-free conditions with approval by the Institutional Committee of Zhengzhou University. For tumour propagation, 10⁵ indicated tumour cells were subcutaneously injected into BALB/c nude mice on the back, and then the tumour volume was measured at the indicated time points. For tumour formation assays, 10, 1 × 10², 1 × 10³, 1 × 10⁴ and 1 × 10⁵ indicated cells were subcutaneously injected into BALB/c nude mice and the tumour formation was observed 3 months later. TIC ratios were obtained using extreme limiting dilution analysis30. Five BALB/c nude mice were used for every sample in tumour propagation assays, and six were used for tumour initiation assays.

ChIP and ChIRP. ChIP assays were performed following standard protocol (Upstate Biotechnology, Inc.). Primary HCC cells or spheres were treated with 1% formaldehyde for 10 min, cracked with SDS lysis buffer followed by ultrasonication, then incubated with proper antibodies (anti-Stat3, anti-H3K4Me3 or anti-H3K27Ac). After washing by high salt, low salt and LiCl buffer, the elution buffer was used to harvest the chromatin de-compact. Finally, the de-compact was performed and enrichment was examined using real-time PCR.

For ChIRP assays, biotin-labelled LncSox4 was obtained using in vitro transcription and incubated with crashed samples, followed by adding streptavidin-conjugated beads. The following processes were similar to ChIP assays. All buffers used for ChIRP were RNase-free.

Luciferase assay. The indicated regions of Sox4 promoter were cloned into pGL3 luciferase reporter plasmid and transfected into the indicated cells. One nanogram pRL-TK was co-transfected as loading control. Thirty-six hours later, the cells were cracked using lysis buffer and detected according to the manual of Dual-Luciferase Reporter Assay system (Promega).

RNA pull-down and mass spectrometry. Biotin-labelled LncSox4 was obtained using Biotin RNA Labeling Mix by T7 RNA polymerase, and then incubated with 1% formaldehyde-treated cell lysate for 4 h or overnight. The samples were separated using electrophoresis and LncSox4 specific bands were identified using mass spectrometry and retrieved in human proteomic library.

RIP assay. For RIP, the cells were treated with 1% formaldehyde and the crashed with RIPA buffer (150 mM NaCl, 0.5% sodium deoxycholate, 0.1% SDS, 1% NP40, 0.5% Triton X-100).
1 mM EDTA and 50 mM Tris (pH 8.0) supplemented with RNase inhibitors and proteinase inhibitors for 30 min, followed by centrifugation. The supernatants were incubated with the indicated antibodies for 4 h, and then protein A/G beads were added. The precipitates were washed with RIPA buffer followed by de-crosslinking. Finally RNA was extracted and LncSox4 enrichment was examined using real-time PCR.

**Statistics.** For box and whisker plots, figures were produced with GraphPad Prism 6 software. Box indicates interquartile range; whiskers indicate 5–95 percentiles; horizontal line within box denotes median value. Two-tailed Student's t-test was used for statistical analysis with Microsoft Excel, and P<0.05 was considered significant. For histograms, data were shown as means ± s.d. Two-tailed Student's t-test was used for statistical analysis with Microsoft Excel. *P<0.05; **P<0.01; ***P<0.001. P<0.05 was considered significant.

**Data availability.** The microarray data referenced during the study (GSE54238, GSE61444 and GSE14520) are available in a public repository from EBI (http://www.ebi.ac.uk/) or NCBI (http://www.ncbi.nlm.nih.gov/gds/?term = ). All the other data supporting the findings of this study are available within the article and its Supplementary information files. All other relevant source data are available from the corresponding author upon request.

**References**

1. Han, Z. G. Functional genomic studies: insights into the pathogenesis of liver cancer. *Ann. Rev. Genom. Hum. Gen.* 13, 171–205 (2012).
2. Bruix, J., Gores, G. J. & Mazzaferro, V. Hepatocellular carcinoma: clinical update. *JAMA* 312, 1020–1028 (2014).
3. Yuan, J. H. Long non-coding RNAs: new players in cell development and disease. *Cell Stem Cell* 17, 717–728 (2015).
4. Oishi, N., Yamashita, T. & Kaneko, S. Molecular biology of liver cancer stem cells. *Liver Cancer* 3, 71–84 (2014).
5. Zhao, W. et al. A mAb raised against recurrent tumor cells, targets liver tumor-initiating cells by binding to the calcium channel alpha 2 delta 1 subunit. *Cancer Cell* 23, 541–556 (2013).
6. Ma, S. et al. Identification and characterization of tumorigenic liver cancer stem/progenitor cells. *Gastroenterology* 132, 2542–2556 (2007).
7. Wang, Y. Y. et al. The IncTGFbeta promotes self-renewal of human liver cancer stem cells through activation of Wnt signaling. *Cell Stem Cell* 16, 413–425 (2015).
8. Yuan, J. H. et al. A long noncoding RNA activated by TGF-beta promotes the invasion-metastasis cascade in hepatocellular carcinoma. *Cancer Cell* 25, 666–681 (2014).
9. Zhu, P. P. et al. Csf1r negatively regulates self-renewal of liver cancer stem cells via suppression of NOTCH2 signalling. *Nat. Commun.* 6, 7122 (2015).
10. Zhu, P. et al. Inc-beta-Catm elicits EZH2-dependent beta-catenin stabilization and sustains liver CSC self-renewal. *Nat. Struct. Mol. Biol.* 23, 631–639 (2016).
11. Takebe, N. et al. Targeting Notch, Hedgehog, and Wnt pathways in cancer stem cells: clinical update. *Nat. Rev. Clin. Oncol.* 12, 445–464 (2015).
12. Fatica, A. & Bozzoni, I. Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21 (2014).
13. McHugh, C. A. et al. The Xist IncRNA interacts directly with SHARP to silence transcription through HDAC3. *Nature* 521, 232–236 (2015).
14. Wang, P. et al. The STAT3-binding long noncoding RNA Inc-DC controls human dendritic cell differentiation. *Science* 344, 310–313 (2014).
15. Kim, T. et al. Role of MYC-regulated long noncoding RNAs in cell cycle regulation and tumorigenesis. *J. Natl Cancer Inst.* 107, dju505 (2015).
16. Liu, B. et al. A cytoplasmic NF-kappaB interacting long noncoding RNA blocks IkappaB phosphorylation and suppresses breast cancer metastasis. *Cancer Cell* 27, 370–380 (2015).
17. Yang, F., Zhang, H., Mei, Y. & Wu, M. Reciprocal regulation of HIF-1alpha and methyltransferase DOT1L. *Immunity* 40, 772–784 (2014).

**Acknowledgements**

This work was supported by the National Natural Science Foundation of China (81373228, 91442121).

**Author contributions**

Z.-z.C. designed the experiments, performed the experiments, analysed the data and wrote the paper; L.H provided HCC samples and analysed the data; Y.-h.W. and W.-j.Z. initiated the study, designed and wrote the paper.

**Additional information**

Supplementary information accompanies this paper at http://www.nature.com/naturecommunications

**Competing financial interests:** The authors declare no competing financial interest.

**Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions**

**How to cite this article:** Chen, Z.-z. et al. LncSox4 promotes the self-renewal of liver tumour-initiating cells through Stat3-mediated Sox4 expression. *Nat. Commun.* 7:12598 doi: 10.1038/ncomms12598 (2016).

© The Author(s) 2016