ARTICLE





KLF5 and MYC modulated LINC00346 contributes to gastric cancer progression through acting as a competing endogeous RNA and indicates poor outcome

Tong-peng Xu¹ · Pei Ma¹ · Wen-yu Wang² · You Shuai¹ · Yan-fen Wang³ · Tao Yu¹ · Rui Xia⁴ · Yong-qian Shu¹

Received: 22 February 2018 / Revised: 10 October 2018 / Accepted: 31 October 2018 / Published online: 15 February 2019 © ADMC Associazione Differenziamento e Morte Cellulare 2018

Abstract

It was found in this study that long intergenic non-protein coding RNA 346 (LINC00346) was an lncRNA aberrantly expressed in gastric cancer (GC) based on multiple Gene Expression Omnibus (GEO) databases of GC cohorts. The LINC00346 gene was recurrently amplified and upregulated in GC, and its expression was positively correlated with poor pathologic stage, large tumor size, and poor prognosis. In addition, the oncogenic transcription factors KLF5 and MYC could bind to the LINC00346 promoter and enhance its expression. Gene Set Enrichment Analysis (GSEA) in the GEO datasets revealed that cell cycle and focal adhesion genes were enriched in patients with high LINC00346 expression. In vitro and in vivo assays of LINC00346 alterations revealed a complex integrated phenotype affecting cell growth, migration and invasion. Strikingly, high-throughput sequencing analysis after LINC00346 alterations highlighted alterations in cell cycle and focal adhesion pathways in GC cells. Mechanistically, argonaute 2 (Ago2) was recruited by LINC00346, which functioned as a molecular sponge for miR-34a-5p by antagonizing its ability to repress CD44, NOTCH1, and AXL protein translation. Taken together, our findings support a model in which the KLF5, MYC/LINC00346/miR-34a-5p cross-talk served as critical effectors in GC tumorigenesis and progression, suggesting a new therapeutic direction in the treatment of GC.

Edited by M. Piacentini.

Electronic supplementary material The online version of this article (https://doi.org/10.1038/s41418-018-0236-y) contains supplementary material, which is available to authorized users.

☑ Yong-qian Shu yongqian_shu@163.com shuyongqian@csco.org.cn

- ¹ Department of Oncology, the First Affiliated Hospital of Nanjing Medical University, No. 300 Guangzhou Road, Nanjing 210029 Jiangsu Province, People's Republic of China
- ² Cancer Research Institute, Seoul National University College of Medicine, 103. Daehak-ro, Jongno-gu, Seoul 03080, Republic of Korea
- ³ Department of Pathology, The Affiliated Hospital of Yangzhou University, Yangzhou University, Yangzhou, Jiangsu Province, People's Republic of China
- ⁴ Department of Medical Laboratory, Nanjing Chest Hospital, No. 215 Guangzhou Road, Nanjing 210029 Jiangsu Province, People's Republic of China

Introduction

Gastric cancer (GC) ranks the second leading cause of cancer mortality worldwide. Most GC cases were diagnosed at advanced stages and usually had poor clinical outcomes due to recurrence and metastasis [1]. Therefore, there is an urgent need to explore and identify specific genetic alterations associated with GC for the sake of seizing novel diagnostic and therapeutic opportunities [2]. Recently, large-scale cancer genomic studies have revealed many potential protein-coding gene biomarkers for GC, including TP53, KRAS, PIK3CA, ERBB2, and PTEN [3], but unfortunately they seem to offer little help in the diagnosis and treatment of GC.

The Human Genome Project (HGP) reveals that only 2% DNA sequences encode proteins and over 90% of them are actively transcribed, and therefore most non-coding transcripts are defined as noncoding RNAs (ncRNAs) [4]. Based on size, ncRNAs are classified as two groups: small noncoding RNAs (sncRNAs) and long noncoding RNAs (lncRNA). They have recently confirmed as critical regulators of gene expression. microRNAs (miRNAs) are a

significant subgroup of small sncRNAs and their functions in tumorigenesis have been demonstrated in multiple cancers [5, 6]. LncRNAs belong to a group of ncRNAs with more than 200 base sequences and were also uncovered to exert functions in modulation of distinct molecular processes [7]. The lncRNA expression profile is frequently dysregulated in GC, and specific lncRNAs, such as FENDRR. TINCR. and GAPLINC are correlated with cancer tumorigenesis [8–10]. Mechanism analysis revealed that lncRNA could bind to functional proteins, and regulate gene expression on transcriptional or post-transcriptional level [10]. Accumulating data show that lncRNAs can also act as decoys, enabling miRNAs to regulate target gene expression [9, 11]. Despite the above findings, our current knowledge on the expression profiles and biological functions of lncRNAs in GC remains limited.

In this study, we intended to re-annotate the probes of Agilent microarrays that match lncRNA sequences in GC based on two GEO datasets (GSE51575 and GSE65801) and identify deregulated lncRNAs in GC. It was found that most deregulated lncRNAs were associated with somatic copy-number alterations (SCNAs). Notably, a long intergenic non-coding RNA LINC00346 displayed considerable predictive effects in the diagnosis and prognosis of GC. We also demonstrated that LINC00346 gene was recurrently amplified, and upregulation of LINC00346 was associated with poor outcomes in GC patients. LINC00346 expression was also induced by the oncogenic transcription factor KLF5 or MYC, and strongly correlated with NOTCH1, AXL, and CD44 in GC tissues. Our findings may provide both in vitro and in vivo evidence to confirm that LINC00346 forms a molecular decoy for miR-34a-5p, which targets NOTCH1, CD44, and AXL for degradation. Also, we illuminated a model for LINC00346-mediated cell proliferation and migration in GC.

Results

IncRNA expression profiling in GC and deregulation of IncRNAs is associated with recurrent SCNAs

Two human lncRNA microarray datasets GSE51575 and GSE65801 (totally 58 paired cancer and noncancer tissues) using the Agilent-028004 SurePrint G3 Human GE 8 _ 60 K Microarray platform were downloaded to analyze expression profiles of different lncRNAs between GC and matched normal tissues. We found that 63 lncRNAs were consistently upregulated and 143 lncRNAs were consistently down-regulated in the two microarray datasets (fold change > 1.5, P < 0.05, FDR < 0.05; Supplementary Fig. 1a, b). Knowing that change in gene expression is frequently related to SCNAs in malignancies, we analyzed whether SCNAs also

existed in deregulated lncRNAs in GC. We re-assigned lncRNA gene loci in the genomic regions with SCNAs in GC downloaded from the Cancer Genome Atlas (TCGA) dataset. In all 206 deregulated lncRNAs, 11 (5.4%) were mapped to genomic loci with gained SCNAs in GC, and 60 (29.1%) to genomic loci with losed SCNAs (Supplementary Fig. 1c, Supplementary Table 3), suggesting that either SCNA gain or loss may contribute to deregulation of lncRNAs in GC.

Relationship between LINC00346 expression and clinical characteristics in GC

We focused on overexpressed lncRNAs due to their advantage as early diagnosis markers or therapeutic targets. As LINC00346 expression is increased most consistently in GC tissues with a remarkable P value and false discovery rate (FDR) in these microarray databases, it was chosen for further investigation in this study (Fig. 1a).

LINC00346 was identified with the full length of 6550 bp consisting of a poly (A)-positive sequence through rapid amplification of cDNA ends (RACE) and Sanger sequencing analysis after PCR (Supplementary Fig. 2a-c). TxCdsPredict in UCSC was used to evaluate protein coding potential of transcripts and a RNA with a score less than 800 was considered a ncRNA [12]. The txCdsPredict score for LINC00346 was 519, indicating that LINC00346 has no protein-coding potential. We confirmed that LINC00346 was indeed a ncRNA by ORF Finder analysis from the National Center for Biotechnology Information [13]. ORF Finder failed to create a protein of more than 165 amino acids (Supplementary Fig. 2d). In addition, LINC00346 does not contain a valid Kozak sequence, which further supports its no protein-coding feature.

To further verify the conclusion of the microarray datasets analysis, the LINC00346 expression was compared between matched GC and normal tissues (n = 106) derived from our cohort. Analysis of the tumor/non-tumor matched tissue (T/N) ratio for LINC00346 expression of the 106 patients revealed that LINC00346 expression was prominently increased in GC tissues (P < 0.001, Fig. 1b). LINC00346 enrichments in GC cell lines and normal gastric epithelial cell line GES1 were further quantified by reverse transcription (RT)-qPCR. It was found that the expression of LINC00346 in GC cell lines was significantly higher than that in GES-1 (Supplementary Fig. 3a). These results indicate that LINC00346 may exert important functions in gastric carcinogenesis.

GC patients were assigned to low or high subgroups according to whether LINC00346 expression was downregulated or up-regulated compared with the matched normal tissue samples, and clinicopathologic characteristics were further analyzed. The high-LINC00346 group



Fig. 1 lncRNA LINC00346 is over-expressed in gastric cancer (GC) and this overexpression is associated with a poor outcome. **a** LINC00346 pooled expression in GC and paired surrounding tissues in the GSE51575 and GSE65801 datasets (paired *t* test, P < 0.01). **b** LINC00346 expression in GC and matched surrounding tissues in the Nanjing dataset; n = 106, (Wilcoxon signed-rank test, P < 0.01). Data are represented as \log_2 fold change (tumor/normal). **c** Kaplan–Meier

analysis of DFS according to LINC00346 expression in total 106 patients (P < 0.0001). **d** Multivariable analysis was employed in the Nanjing cohort. **e** The ROC curve for predicting GC depended on LINC00346 expression, using the matched surrounding normal tissues as controls based on the GSE51575 and GSE65801 datasets (left panel) or Nanjing cohort (right panel)

(n = 79) exhibited greater depth of invasion (P < 0.001), more regional lymph-node metastasis (P = 0.004) and higher tumor stages (P < 0.001) as compared with the low-LINC00346 group (n = 27) (Supplementary Table 4).

Subsequently, the effects of LINC00346 enrichment and clinicopathological features on DFS were analyzed by Kaplan–Meier curve and log-rank test, showing that the recurrence rate was significantly higher in patients with high LINC00346 expression than that in those with low LINC00346 expression (median DFS 18 months vs. 36 months, P < 0.0001) (Fig. 1c). Univariate analyses of pathological characteristics regarded as potential predictors of DFS are listed in Supplementary Table 5. Further multivariate Cox proportional hazards analysis demonstrate that LINC00346 enrichment could serve as an independent indicator of DFS (hazard ratio [HR] = 2.002; 95% confidence interval [CI], 1.020–3.929; P = 0.044) in GC patients (Fig. 1d, Supplementary Table 5).

A receiver operating characteristic (ROC) curve with the normal tissues surrounding the tumor tissues as a control based on pooled of GSE51575 and GSE65801 databases



Fig. 2 KLF5, MYC, and LINC00346 are amplified in GC tissues, and KLF5 or MYC up-regulates LINC00346 expression in GC cells. a Oncoprint analysis data of MYC, Kruppel-like factor 5 (KLF5), and LINC00346 in GC patient samples are displayed. Data from the TCGA dataset containing 258 patients were extracted and used to generate an oncoprint plot using the cBioPortal. Gene amplification (red bars) and copy number gain (pink bars) are shown. b KLF5, MYC, and LINC00346 gene expression in GC with or without KLF5/ MYC/LINC00346 SCNAs. GC with KLF5/MYC/LINC00346 amplifications or gain expressed significantly higher transcript enrichment. c KLF5 binding on LINC00346 genomic promoters in KATO-III GC cells through analyzing GSE51705 database. The red bar highlights the LINC00346 genomic promoter region. d UCSC Genome Bioinformatics Site (http://genome.ucsc.edu/) showed high enrichment of MYC at the promoter of LINC00346. e Relationship between LINC00346 expression and KLF5 (left panel) or MYC (right panel)

was created to evaluate the diagnostic effect of LINC00346 expression. The cutoff value for distinguishing GC tissues from normal tissues was 6.33 (Normalized Intensity value). The area under the ROC curve (AUC) was 0.857 (95% CI = 0.779–0.915, P < 0.0001), with a specificity and sensitivity of 0.81 and 0.79, respectively (Fig. 1e, left). We next validated the result in our cohort, and showed that the cutoff value for differentiate the tumor tissues was 11.395 (Δ Ct) and AUC was 0.699 (95% CI = 0.632–0.760, P < 0.0001), with a specificity and sensitivity of 0.764 and 0.594 respectively (Fig. 1e, right).



mRNA levels (Normalized Intensity) in 52 GC tissues based on GSE51575. f Relationship between LINC00346 expression (ΔCt value) and KLF5 (left panel) or MYC (right panel) mRNA levels (ΔCt value) in 106 GC tissues in our cohort. g ChIP assay demonstrated endogenous MYC and KLF5 binding to the LINC00346 gene promoter (left). ChIP assay showed MYC/KLF5 enrichment on LINC00346 promoter transfected with MYC or KLF5 overexpression vector (middle) or siRNAs (right). *P < 0.05; **P < 0.01. h KLF5 and MYC protein expression, LINC00346 mRNA expression in cells transfected with corresponding siRNAs (left) or overexpression vector (right). i Dual-luciferase reporter assay was performed by cotransfection of the LINC00346 promoter fragment (pGL3-LINC00346) or LINC00346 promoters constructs with deletions in different binding elements for MYC and KLF5 with overexpression of MYC or KLF5. WT wild type, D deletion type. Luciferase activity was expressed as relative to that of the pGL3 vector (a promoter-less vector)

SCNAs and oncogenic transcription factors contribute to LINC00346 overexpression

To further identify the potential mechanism involved in LINC00346 overexpression in GC, we first analyzed the SCNA data available from TCGA datasets using cBioPortal [14]. As shown in Fig. 2a, LINC00346 gene copy number amplification or gain occurred in only 91 of the 258 samples (35.3%), and LINC00346 expression was high in tumors with high SCNAs as compared with tumors with no or low amplification (Fig. 2b and Supplementary Table 6),



indicating that SCNAs contributed partly to LINC00346 upregulation in GC.

We hypothesized that transcriptional factors (TFs) may play pivotal roles in modulating LINC00346 gene transcription. To confirm this hypothesis, we first obtained chromatin immunoprecipitation sequencing (ChIP-seq) of TFs from the ENCODE project by using UCSC Genome Bioinformatics Site (http://genome.ucsc.edu/) and GEO database (GSE51705), and found high enrichment of oncogenic TF MYC and KLF5 at the promoter of LINC00346 (Fig. 2c, d). Knowing that MYC and KLF5 are involved in GC cell survival or proliferation [15, 16], and were each recurrently amplified in patients with GC (Fig. 2a, Supplementary Table 7 and 8). We examined gene

✓ Fig. 3 GSEA and GO in LINC00346 high/low expression gastric cancer patients and cells. a GSEA comparison of the LINC00346 low (blue) and high (red) expression subgroups of GC patients in the GSE65801 dataset. Cytoscape and Enrichment map were used for the visualization of GSEA results (1% FDR, P = 0.005). Nodes represent enriched gene sets, grouped and annotated by similarity according to the related gene sets. Enrichment results were mapped as a network of gene sets (nodes). Node size is proportional to the total number of genes within the gene set. This network map was manually curated removing general and uninformative sub-networks, resulting in the simplified network map shown in (a). Enrichment plots are illustrated for a set of activated genes related to cell cycle, focal adhesion and Notch signaling pathway in the GSE65801 dataset. A positive value indicates a high correlation with the "LINC00346 high expression" group. NES normalized enrichment score. b Heatmap of altered genes in LINC00346 overexpressed BGC823 cells compared with control cells. c GO analysis for all genes with altered expressions. d, e Corroboration of changes in mRNA or protein levels of select genes by aRT-PCR and Western blot in cells with LINC00346 knockdown (siRNA1 and siRNA2) and overexpression. Bars: SD; n = 3, *P < 0.05, ***P* < 0.01

expression profiles of MYC and KLF5 in the TCGA GC cohort, where SCNA and gene expression information were available. It was found that tumors with either high MYC or KLF5 exhibited significantly increased expression of SCNAs ('copy number driven expression') (Fig. 2b). Meanwhile, correlation analysis revealed that LINC00346 was significantly and positively correlated with KLF5 and MYC in the GSE51575 dataset and our sample cohort (Fig. 2e, f). Clinicopathologic analysis revealed that upregulation of KLF5/MYC in tumor samples was significantly associated with lymph-node metastasis in our sample cohort (Supplementary Fig. 4a, b). In addition, our ChIP assays showed that MYC and KLF5 could bind to the LINC00346 promoter DNA regions (Fig. 2g, left). To further evaluate the biologic roles of KLF5 and MYC in LINC00346 expression, we performed loss- and gain-offunction methods, and revealed that the overexpression (Fig. 2g, middle) or depletion (Fig. 2g, right), up-regulated or decreased KLF5 and MYC enrichment on the LINC00346 promoter respectively, leading to LINC00346 upregulated or downregulated expression in GC cells respectively (Fig. 2h). In addition, luciferase report assays indicated that KLF5 bound to the E1 and E2 (-1612 bp) binding site, and MYC bound to the E3 binding site (Fig. 2i).

LINC00346 functions as a tumor inducer in vitro and in vivo

To explore the biologic pathways involved in GC pathogenesis, LINC00346 expression level was stratified by the median, and GSEA analysis was applied in GSE65801 datasets. Enrichment plots of GSEA showed that the gene signatures of cell cycle, focal adhesion, and Notch signaling

pathway were more involved in patients with LINC00346 higher expression relative to patients with lower expression in this datasets, indicating that LINC00346 function may be associated with cell proliferation and migration (Fig. 3a). To gain further insights into LINC00346-related pathway in accelerating GC progression, we evaluated the gene profiles in GC cells with altered LINC00346 expression. We knocked down LINC00346 using siRNAs or shRNAs, and overexpressed this gene by transfecting a LINC00346 overexpression plasmid in GC cells (Supplementary Fig. 3b-d). RNA transcriptome sequencing analysis was used to detect the altered gene expression profiles from control plasmid and pcDNA3.1-LINC00346 overexpression plasmid transfectants. A total of 1107 altered expression genes (>2-fold) were detected after LINC00346 overexpression in GC cells (Fig. 3b, Supplementary Table 9). Gene Ontology analysis revealed that changes in gene sets were related to cell proliferation, adhesion and regulation of the Notch signaling pathway in LINC00346-overexpressing cells (Fig. 3c), which is consistent with GC patients. Further qPCR data and Western blot validated that change in LINC00346 expression markedly affected the key gene signatures involved in tumorigenesis and disease progression (Fig. 3d, e), indicating that LINC00346 maybe a key regulator for maintenance of malignant phenotypes in GC.

To validate the GSEA analysis of LINC00346, gain and loss of function methods were performed to estimate the biological function of LINC00346 in GC cell lines. MTT ((3-(4, 5-dime thylthiazol-2-yl)-2, 5-diphenyltetrazolium bromide) tetrazolium) and colonyformation assays demonstrated that the growth and proliferation of GC cells transfected with siRNAs or shRNAs in BGC823 and SGC7901 cells with high LINC00346 expression but not scrambled oligonucleotides were suppressed significantly. Meanwhile, ectopic overexpression of LINC00346 by transfecting the MGC803 cells with relatively low LINC00346 expression, with the pcDNA3.1- LINC00346, selected by the addition of G418, induced GC cell proliferation in vitro (Fig. 4a, b and Supplementary Fig. 5a). We also tested the effect of LINC00346 on GC cell-cycle progression. As shown in Fig. 4c, LINC00346 depletion significantly blocked the cell cycle at the G1-S phase, while upregulation of LINC00346 slightly promoted cell cycle progression. These observations were further verified by EDU (red)/ DAPI (blue) immunostaining assay, showing that LINC00346 knockdown prominently repressed GC cell growth, and LINC00346 overexpression marginally promoted cell growth (Fig. 4d). Furthermore, wound healing and transwell assays were applied to detect the role of LINC00346 on GC cell migration and invasion (Fig. 4e, Supplementary Fig. 5b). The results showed that Fig. 4 LINC00346 promotes cell proliferation, migration, and invasion in vitro. **a**, **b** Colony-formation assays were performed to examine the proliferation ability of shRNA or overexpression plasmid treated GC cells. c FACS analysis of the effect of LINC00346 on cell cycle. d EdU immunofluorescence staining of differently treated GC cells showed the repressive effects of LINC00346. The graphs on the right show the percentage of EdU-positive nuclei. The data shown are the mean of three independent experiments. e Transwell assays showed that LINC00346 knockdown inhibited GC cell migration and invasion, while ectopic up-regulated LINC00346 promoted migration and invasion. Scale bars = 100 µm. Bars: SD; n = 3, *P <0.05, ***P* < 0.01



LINC00346 knockdown dramatically decreased cell migration and invasion, while LINC00346 overexpression promoted cell migration and invasion.

We then extended the research of the LINC00346 proliferation-induction role to an athymic(nu/nu) mouse model by injecting LINC00346 stable knockdown BGC823



Fig. 5 LINC00346 promotes GC cell growth and migration in vivo. **a** Representative images of mice photographed by the IVIS Imaging System and dissected tumors bearing from BGC823 cells transfection of scambled or shRNA groups, and **b** the tumor volume growth curves after injections in different groups. **c** LINC00346 transcription was quantitated using qRT-PCR upon transfection of BGC823 cells with scramble RNA or shRNAs, in subcutaneous tumors harvested at day 16. Data are mean \pm SD from three independent experiments. **P* < 0.05, ***P* < 0.01. **d** The metastatic animal model was established by injecting BGC823 cells stably knockdown LINC00346 into nude mice. Representative images of mice photographed by the IVIS Imaging System and lungs with tumor nodules from mice in each group

cells, LINC00346 overexpressing MGC803 cells, or control cells into nude mice to determine whether LINC00346 would affect GC cell tumorigenesis in vivo. The results showed that tumors grown from LINC00346 stable knockdown cells were smaller than those grown from control cells (Fig. 5a, b); however, tumors grown from LINC00346–overexpressing cells were slightly larger than those grown from control cells (Fig. 5e, f). We further explored the role of LINC00346 in lung colonization by inoculating cells directly into the tail vein of nude mice, and found that the lung metastatic burden was increased in LINC00346-overexpressing MGC803 cells, and

are shown. **e** Representative images of mice photographed by the IVIS Imaging System and dissected tumors bearing from MGC803 cells transfection of empty vector or pcDNA3.1-LINC00346 vector, and **f** the tumor volume growth curves after injections in two groups. **g** LINC00346 levels detected in tumors from two groups mice by qRT-PCR. **h** Metastastic animal model was performed by injecting MGC803 cells stably overexpressed LINC00346 into nude mice. Representative images of mice photographed by the IVIS Imaging System and lungs with tumor nodules from mice in each group are shown. Data are mean ± SD from three independent experiments. **P* < 0.05, ***P* < 0.01

decreased in LINC00346 stable knockdown BGC823 cells (Fig. 5d, h). To analyze the physiological function of LINC00346, we generated additional clones expressing LINC00346 at a similar level to tumor tissues, with the expression of LINC00346 20 times greater than the empty control expressing cells. We also validated these data by in vitro and in vivo assays (Supplementary Fig. 6).

In conclusion, LINC00346 appeared to have substantial impact on cell migration, invasion, and lung colonization, indicating that the impact of LINC00346 on patient prognosis might be mainly associated with the role in the metastatic spread.



✓ Fig. 6 LINC00346 promotes GC cell proliferation and invasion partly by modulating Notch1, CD44, and AXL. a Representative images of Notch1, CD44, and AXL protein expression detected by IHC analysis in primary tumor tissues (left panel) and lung colonies (right panel) formed from LINC00346 ectopic expression or control cells. The number above each image represents the IHC score. Scale bars = 100 µm. b qRT-PCR analysis of NOTCH1, AXL, and CD44 expression in primary gastric carcinoma with or without lymph node metastasis. The term -PCt was used to describe the expression level of NOTCH1, AXL, and CD44 $(-\Delta Ct = Ct(GAPDH)-Ct(NOTCH1/$ AXL/CD44)). A significant difference was detected in the mean value of NOTCH1, AXL, and CD44 expression between these two groups (Student's t-test). c Representative images of Notch1, CD44, and AXL protein expression detected by IHC analysis in primary gastric carcinoma with or without lymph node metastasis. The number above each image represents the IHC score. Scale bars = $100 \,\mu\text{m}$. d Notch1, CD44, and AXL protein levels were detected by Western blot analysis in BGC823 cells. e, f Transwell assays showed that knockdown of Notch1, CD44 or AXL partly reversed LINC00346-mediated invasion promotion. The statistical result was based on three independent experiments. Scale bars = 100 μ m. Bars: SD; n = 3, *P < 0.05, **P <0.01. g, h The pro-growth effect of LINC00346 was partly neutralized by suppressing Notch1 or AXL expression. GC cells were transfected with LINC00346 cDNA in the presence or absence of siRNAs or shRNA for Notch1 or AXL, and the viability of cells was determined by MTT assay (g) and colony-formation assays (h)

Notch1, AXL, and CD44 are key downstream targets of LINC00346

To further investigate the potential genes involved in LINC00346 function, we analyzed the RNA-sequencing and qPCR data, and confirmed that alterations in LINC00346 expression dramatically affected Notch1, CD44, and AXL gene enrichment (Fig. 3b, d, e), but LINC00346 knockdown slightly decreased CD44 mRNA and protein expression, possibly owing to the high basal expression of CD44 in GC cells. In agreement, immunohistochemistry (IHC) showed that LINC00346-transfected MGC803 cells exhibited higher Notch1, CD44, and AXL levels than control cells in the athymic (nu/nu) mouse model, especially in lung colonies (Fig. 6a). Previous studies demonstrated that Notch1, CD44, and AXL were involved in GC progression [17-19], and therefore we examined their expressions in human GC samples. IHC and qPCR showed that Notch1, CD44, and AXL expressions were increased in GC tumors with lymph node metastasis compared with primary GC tumors without lymph node metastasis (Fig. 6b, c).

To determine whether CD44, AXL, and Notch1 participated in increased GC cell proliferation and invasion induced by LINC00346, we performed a rescue experiment by cotransfected MGC803 cells with pcDNA3.1-LINC00346, siCD44, or siAXL, or siNotch1 (Fig. 6d, Supplementary Fig. 6a and Supplementary Fig. 7). Transwell assays showed that CD44, AXL, and Notch1 partially rescued the LINC00346–induced cell invasion (Fig. 6e, f and Supplementary Fig. 6b, c), and MTT and colonyformation assays showed that AXL and Notch1 partially rescued the LINC00346–induced cell proliferation in vitro, but not CD44 (Fig. 6g, h).

LINC00346 regulates CD44, Notch1, and AXL expression by competing for miR-34a-5p

To probe the mechanism for LINC00346-regulated CD44, Notch1, and AXL expression, we first examined its distribution in GC cells using RNA-fluorescence in situ hybridization (FISH) and subcellular fractionation. It was found that LINC00346 RNAs were more prevalent in the cytoplasm (Supplementary Fig. 8a, b), indicating that LINC00346 was likely to play a post-transcriptional regulation function.

Knowing that lncRNAs function as natural "microRNA sponges" that protect mRNAs by competing for their targeting miRNAs, it was worthy to investigate whether LINC00346 played a similar role. We predicted 645 miR-NAs targeting sites on LINC00346 (Supplementary Table 10) using the TargetScan and miRanda prediction algorithm. Of these miRNAs, miR-34a, miR-205, miR-181c, miR-143, and miR-23b have been reported to repress tumor proliferation and invasion, and with significant binding energy predicted by the extensively-used RNAup algorithm (Fig. 7a). Thus, we focused on these miRNAs and explored the interaction between LINC00346 in GC cell lines. Knockdown of LINC00346 significantly increased miR-34a-5p, and LINC00346 overexpression decreased its enrichment but not the expression of other miRNAs in GC cells (Supplementary Fig. 9a). Interestingly, miR-34a-5p, which was reported to target CD44, Notch1, and AXL [20-22], was found to act as a tumor suppressor in GC [23], and clinicopathologic analysis in our cohort revealed that downregulation of miR-34a in tumor samples was significantly associated with lymph-node metastasis (Supplementary Fig. 4c). Therefore we subsequently investigated whether LINC00346 acted as a sponge of miR-34-5p.

To verify the direct association between miR-34a and LINC00346 at endogenous levels, we applied RNA immunoprecipitation (RIP) to pull down endogenous miR-NAs bound to LINC00346. The result of qPCR showed that the LINC00346 RIP in BGC823 cells was significantly enriched for miR-34a-5p and LINC00346 as compared with the empty vector (MS2), GADPH, nontargeting microRNA (miR-21), LINC00346 with mutations in miR-34a-5p targeting sites (henceforth named LINC00346-mut1(miR-34a-5p)) (Fig. 7b). The specific binding between miR-34a-5p and LINC00346 was further verified by affinity pull-down of endogenous miR-34a-5p through pull down assay (Fig. 7c). For further validation, we constructed luciferase reporter vectors containing the 3' 600 nt or 700 nt of

LINC00346 containing the wild-type (WT) known as pLUC-LINC00346-1 and pLUC-LINC00346-2 or mutated miR-34a-5p binding sites known as pLUC-LINC00346-1mut and pLUC-LINC00346-2-mut. It was found that upregulation of miR-34a-5p decreased the luciferase activities of the reporter vector pLUC-LINC00346-1 but not pLUC-LINC00346-2, mutant reporter vector or empty vector in 293T cells (Fig. 7d). However, we found no significant difference in LINC00346 levels after overexpression of miR-34a-5p (Supplementary Fig. 9b). The miRNAs are confirmed to bind their targets and lead to RNA degradation and/or translational inhibition in an AGO2-dependent manner. To explore whether LINC00346 was modulated by miR-34a-5p in such a way, we conducted anti-AGO2 RIP in BGC823 cells transiently upregulating miR-34a-5p. Endogenous LINC00346 pull-down by AGO2 was specifically enriched in miR-34a-5p-treated cells (Fig. 7e), validating that miR-34a-5p was a bona fide LINC00346-targeted microRNA. These findings illustrate that miR-34a-5p was associated with LINC00346 but did not induce LINC00346 degradation.

To determine whether LINC00346 regulated CD44, Notch1, and AXL expression by competing for miR-34a-5p, we performed rescue assays. Ectopic expression of LINC00346 increased CD44, Notch1, and AXL protein levels and overexpression of miR-34a-5p reduced this increase, while inhibited of LINC00346 decreased CD44, Notch1, and AXL protein levels, and the use of inhibitor miR-34a reduced this decrease (Fig. 7f). Ectopic expression of miR-34a-5p remarkably inhibited cell migration and invasion, and reexpression of miR-34a-5p was able to reverse the effect of LINC00346 on cell migration/invasion in vitro (Fig. 7g).

To detect the effect of LINC00346 expression on miR-34a activity, we constructed three luciferase reporters to examine miR-34a activity by inserting the 3' untranslated region (UTR) of human CD44, Notch1 or AXL into 3'UTR of Renilla luciferase. As expected, miR-34a significantly decreased the luciferase signals of reporters of 3'UTR of CD44, Notch1 and AXL, and LINC00346 abrogated the knockdown efficiency exerted by miR-34a effectively (Fig. 7h). In addition, we performed FISH in GC cells to detect the subcellular distribution of miR-34a and LINC00346, and found a large degree of overlap between miR-34a and LINC00346 within the cytoplasm, indicating co-localization (Fig. 7i).

Discussion

In this study, we compared the lncRNA profiles of GC and surrounding normal gastric tissues by microarray data sets from GEO (GSE51575 and GSE65801), and found a novel lncRNA LINC00346 that displayed remarkable over-



expression in GC. We further confirmed this expression pattern in GC clinical samples and found that high LINC00346 expression was associated with poor prognosis

in our cohort. Specifically, LINC00346 enrichment could serve as a biomarker to distinguish cancer tissues from normal tissues. These results suggest that LINC00346

✓ Fig. 7 LINC00346 regulates Notch1, CD44 and AXL via interacting with miR-34a-5p. a Schematic outlining the predicted binding sites of miRNAs on LINC00346. b MS2-RIP followed by microRNA gRT-PCR to detect microRNAs endogenously associated with LINC00346. c BGC823 cell lysates were incubated with biotin-labeled LINC00346: after pull-down, microRNAs was extracted and assessed by qRT-PCR. d Luciferase activity in 293T cells were cotransfected with miR-34a and luciferase reporters containing nothing, LINC00346 or mutant transcript. Data are presented as the relative ratio of renilla luciferase activity to firefly luciferase activity. e Anti-AGO2 RIP was performed in BGC823 cells transiently overexpressing miR-34a or miR-21, followed by qRT-PCR to detect LINC00346 associated with AGO2. MiR-21 was used as a negative control. f Western blot assay was used to detect Notch1, CD44 or AXL protein expression in MGC803 cells upon miR-34a mimic or pcDNA3.1-LINC00346 treated (left), and in BGC823 cells in the absence of LINC00346 expression with or without anti-miR-34a (right). g Transwell assays showed that ectopic overexpression of miR-34a inhibited GC cell migration and invasion, while re-expression of miR-34a was able to reverse the promoting effect of LINC00346 on cell migration/invasion. Scale bars = $100 \,\mu m$. Bars: SD; n = 3, *P < 0.05, **P < 0.01. h The 3'UTR of miR-34a-5p target gene, Notch1, AXL or CD44, was inserted into pLUC vector (schematic representation of the target site depicted above). 293T cells were co-transfected with empty vector (EV) or pcDNA3-LINC00346 together with miR-34a-5p mimic or NC (negative control) and pLUC reporter plasmid containing the entire Notch1, AXL or CD44 3' UTR. After 48-h transfection, relative luciferase activity was measured. i In situ hybridization showed colocalization of LINC00346 (green) and miR-34a-5p (red). The rightmost graph shows the colocalized signals between the green and the red signals

expression could provide a prominently predictive and prognostic value for GC patients.

To exploit the biological function of LINC00346, we first identified the full poly (A)-positive sequence and found that it was 278 bp longer than that of NCBI or UCSC reported transcript by using RACE assay and Sanger sequencing analysis. GO analysis and GSEA demonstrated that the cell cycle, focal adhesion and Notch signaling pathway were dramatically enriched in response to LINC00346 alteration in GC patients and cells. In cultured GC cells and xenograft mouse models, silencing LINC00346 expression repressed cell growth and invasion, while LINC00346 up-regulation accelerated cell proliferation and metastasis. These findings consistently support that aggressive GC cells are addicted to higher LINC00346 expression, which may explain the pivotal role that LINC00346 plays in the development of GC.

The key cancer-related genes in GC are often located in the genomic regions with SCNAs, such as amplified genes (*ERBB2*, *EGFR*, *GATA4*, *MYC*, and *KLF5*) and deleted genes (*FHIT*, *RB1*, *CDKN2A/B*, and *WWOX*) [24]. Specifically, oncogenic transcription factors such as *MYC*, *MYCN*, *AR*, *MITF*, and *KLF5* are often activated, resulting in aberrant gene regulation by copy number amplifications and upregulation of their expressions during tumorigenesis [25, 26]. One third of the altered lncRNA genes were found to be located in recurrently deleted or amplified regions in the study, indicating a considerable contribution of genomic-level alteration

to the aberrant expression of lncRNAs in GC. Importantly, we found that oncogenic transcription factors KLF5 and MYC, along with LINC00346, were up-regulated through gene copy number amplification or gain in GC by analyzing 258 patients available from TCGA datasets. Many transcription factors have been uncovered to participate in aberrant genes modulation in GC [27]. In this study, our data demonstrated that LINC00346 overexpression in GC cells could also be activated by KLF5 and MYC. Thus, a KLF5/MYC/LINC00346 co-amplification and co-activated model in GC was first explored in this study.

By combining high-throughput data analysis and functional assays in LINC00346 higher/lower expression GC cells and patients, we identified three key target genes (Notch1, AXL, and CD44), which are known as wellcharacterized genes involved in GC proliferation and migration [9, 17, 19]. LncRNAs have distinct functions according to their subcellular distribution. We found that LINC00346 was a predominantly cytoplasm lncRNA in GC cells, suggesting that it plays a role in post-transcriptional modulation. LncRNA can interfere with miRNA pathways by competing for shared miRNA response elements and affect posttranscriptional modulation [11, 28]. It was found in our study that LINC00346 shared miR-34a-5p response elements with Notch1, AXL, and CD44. Then we combined RIP, RNA pulldown, RNA FISH and luciferase reporter's assays to validate that LINC00346 sponged miR-34a-5p and protected Notch1, AXL, and CD44 against degradation by miR-34a-5p, and caused up-regulation of Notch1, AXL, and CD44. We also identified that miR-34a-5p response element was specifically located at the regions spanning 625 to 645 nucleotides in LINC00346.

Here, we disclose a novel mechanism underlying GC development through a biological cross talk between LINC00346, KLF5, MYC, miR-34a-5p, Notch1, AXL, and CD44. The pleiotropic effects of LINC00346 on the proliferation and metastasis of GC cells suggest that LINC00346 could be an effective therapeutic target for GC.

Materials and methods

Computational analysis

Two datasets (GSE51575 and GSE65801) consisting of 58 paired gastric adenocarcinoma tissues and matched adjacent normal tissues were downloaded from GEO. GC gene expression data analysis and details of GC copy number alteration analysis are supplied in Supplementary Materials and Methods. A Gene Set Enrichment Analysis (GSEA) was performed to explore molecular pathways involved in GC pathogenesis by LINC00346 participation, as previously described [29].

5'and 3'rapid amplification of cDNA ends (RACE) analysis

5' and 3' RACE assays were performed using the 5'-Full RACE Kit with TAP (Takara, Dalian, China) and the 3'-Full RACE Core Set with PrimeScriptTM RTase Kit (Takara) according to the manufacturer's protocols.

Cell lines

The human gastric adenocarcinoma cell lines were maintained according to previously reported [8].

Tissue samples and clinical data collection

Included in this study were 106 patients who underwent primary GC resection at the 1st Affiliated Hospital of Nanjing Medical University (Nanjing, China). The study was approved by the Ethics Committee on Human Research of the hospital and written informed consent was obtained from all patients involved. The clinicopathological characteristics of the GC patients are supplied in Supplementary Table 1. All patients were followed up at 1–2 months intervals until April 2017, with a median follow-up period of 36 (range 20–48) months. Disease-free survival (DFS) was defined as the interval between the dates of surgery and recurrence; if recurrence was not diagnosed, patients were censored on the date of death or the last follow-up.

RNA preparation and quantitative real-time PCR

Total RNA extraction and subsequently quantitative realtime PCR were performed as previous described [8]. The primers used in the research are shown in Supplementary Table 2, and the miRNA primers were purchased from RiboBio Company (Guangzhou, China).

Vector construction, RNA interference, and transfection

Detailed descriptions of plasmid construction, RNA interference and cell transfection are available in Supplementary Materials and Methods and the sequences of these synthesized oligonucleotides for RNA interference (RNAi) are listed in Supplementary Table 2.

Cell proliferation assays

Details of MTT, colony-formation and 5-Ethynyl-2deoxyuridine (EDU) assays are available in previous studies [30].

Luciferase reporter assays

The luciferase assays were performed using a luciferase assay kit (Promega, Madison, WI, USA) based on the manufacturer's protocol. Details are available in Supplementary Materials and Methods.

Isolation of cytoplasmic and nuclear RNA

Cytoplasmic and nuclear RNAs were isolated and purified as previous described [10].

RNA FISH

In situ hybridization was performed using the QuantiGene ViewRNA miRNA ISH Cell Assay Kit (Affymetrix, Inc., Santa Clara, CA, USA) with probes targeting LINC00346 (VA4-20816) and probes targeting miR-34a-5p (VM1-10243) following the manufacturer's protocol. The FISH/IF assay, a combination of modified RNA FISH and IF, was performed as described previously [31].

RNA transcriptome sequencing

RNA transcriptome sequencing was performed as previously reported [10].

Chromatin immunoprecipitation assays

Chromatin immunoprecipitation (ChIP) experiments were performed using the MagnaChIP Kit (Millipore) according to the manufacturer's instructions as described previously [10]. The primer sequences are shown in Supplementary Table 2.

RIP and RNA pull-down assay

RIP and RNA pull-down experiments were performed as described previously [10].

Flow cytometry

Cell-cycle distribution was examined by flow cytometry as previously reported [10].

Western blot

Protein extraction and quantities in cells were performed as previously reported [10].

Wound healing, cell migration, and invasion assays

Wound healing, cell migration and invasion assays were performed as previously reported [8].

Tumor xenograft animal model

Athymic (nu/nu) mouse models were prepared as described in Supplementary Materials and Methods.

Immunohistochemistry

The immunohistochemical detection of CD44, Notch1, and AXL was performed according to a previously described method [8].

Statistical analysis

All statistical analyses were performed using SPSS 20.0 software (IBM, SPSS, Chicago, IL, USA), and details are supplied in Supplementary Materials and Methods.

Data availability

The Gene Expression Omnibus accession numbers for gene expression microarray and ChIP-seq data are GSE51575, GSE65801 and GSE51705, respectively.

Acknowledgements This study was supported by the National Natural Science Foundation of China (81602071 and 81672896); Jiangsu Provincial Clinical Science and Technology Projects (Clinical Research Center, BL2012008); the Natural Science Foundation of Jiangsu Province for Youth (BK20161066); and the Priority Academic Program Development of Jiangsu Higher Education Institutions (Public Health and Preventive Medicine, JX10231801).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

References

- 1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA. 2015;65:87–108.
- Ma Q, Reeves JH, Liberles DA, Yu L, Chang Z, Zhao J, et al. A phylogenetic model for understanding the effect of gene duplication on cancer progression. Nucleic Acids Res. 2014;42: 2870–8.
- Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 2014, 513: 202–9.
- 4. Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, et al. Landscape of transcription in human cells. Nature. 2012;489:101–8.
- Corney DC, Hwang CI, Matoso A, Vogt M, Flesken-Nikitin A, Godwin AK, et al. Frequent downregulation of miR-34 family in human ovarian cancers. Clin Cancer Res. 2010;16:1119–28.
- Mori M, Triboulet R, Mohseni M, Schlegelmilch K, Shrestha K, Camargo FD, et al. Hippo signaling regulates microprocessor and links cell-density-dependent miRNA biogenesis to cancer. Cell. 2014;156:893–906.
- Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. Nat Rev Genet. 2014;15:7–21.

- Xu TP, Huang MD, Xia R, Liu XX, Sun M, Yin L, et al. Decreased expression of the long non-coding RNA FENDRR is associated with poor prognosis in gastric cancer and FENDRR regulates gastric cancer cell metastasis by affecting fibronectin1 expression. J Hematol Oncol. 2014;7:63.
- Hu Y, Wang J, Qian J, Kong X, Tang J, Wang Y, et al. Long noncoding RNA GAPLINC regulates CD44-dependent cell invasiveness and associates with poor prognosis of gastric cancer. Cancer Res. 2014;74:6890–902.
- Xu TP, Liu XX, Xia R, Yin L, Kong R, Chen WM, et al. SP1induced upregulation of the long noncoding RNA TINCR regulates cell proliferation and apoptosis by affecting KLF2 mRNA stability in gastric cancer. Oncogene. 2015;34:5648–61.
- Yuan JH, Yang F, Wang F, Ma JZ, Guo YJ, Tao QF, et al. A long noncoding RNA activated by TGF-beta promotes the invasionmetastasis cascade in hepatocellular carcinoma. Cancer Cell. 2014;25:666–81.
- Li T, Xie J, Shen C, Cheng D, Shi Y, Wu Z, et al. Upregulation of long noncoding RNA ZEB1-AS1 promotes tumor metastasis and predicts poor prognosis in hepatocellular carcinoma. Oncogene. 2016;35:1575–84.
- Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. Nucleic Acids Res. 2007;35 (Web Server issue):W345–49.
- Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401–4.
- Chia NY, Deng N, Das K, Huang D, Hu L, Zhu Y, et al. Regulatory crosstalk between lineage-survival oncogenes KLF5, GATA4 and GATA6 cooperatively promotes gastric cancer development. Gut. 2015;64:707–19.
- Yang F, Xue X, Zheng L, Bi J, Zhou Y, Zhi K, et al. Long noncoding RNA GHET1 promotes gastric carcinoma cell proliferation by increasing c-Myc mRNA stability. FEBS J. 2014;281:802–13.
- Kim SJ, Lee HW, Baek JH, Cho YH, Kang HG, Jeong JS, et al. Activation of nuclear PTEN by inhibition of Notch signaling induces G2/M cell cycle arrest in gastric cancer. Oncogene. 2016;35:251–60.
- Zhu S, Chen Z, Katsha A, Hong J, Belkhiri A, El-Rifai W. Regulation of CD44E by DARPP-32-dependent activation of SRp20 splicing factor in gastric tumorigenesis. Oncogene. 2016;35:1847–56.
- Sawabu T, Seno H, Kawashima T, Fukuda A, Uenoyama Y, Kawada M, et al. Growth arrest-specific gene 6 and Axl signaling enhances gastric cancer cell survival via Akt pathway. Mol Carcinog. 2007;46:155–64.
- Park EY, Chang E, Lee EJ, Lee HW, Kang HG, Chun KH, et al. Targeting of miR34a-NOTCH1 axis reduced breast cancer stemness and chemoresistance. Cancer Res. 2014;74:7573–82.
- Mudduluru G, Ceppi P, Kumarswamy R, Scagliotti GV, Papotti M, Allgayer H. Regulation of Axl receptor tyrosine kinase expression by miR-34a and miR-199a/b in solid cancer. Oncogene. 2011;30:2888–99.
- Liu C, Kelnar K, Liu B, Chen X, Calhoun-Davis T, Li H, et al. The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. Nat Med. 2011;17:211–5.
- Zhou Y, Huang T, Siu HL, Wong CC, Dong Y, Wu F, et al. IGF2BP3 functions as a potential oncogene and is a crucial target of miR-34a in gastric carcinogenesis. Mol Cancer. 2017;16:77.
- 24. Deng N, Goh LK, Wang H, Das K, Tao J, Tan IB, et al. A comprehensive survey of genomic alterations in gastric cancer reveals systematic patterns of molecular exclusivity and co-occurrence among distinct therapeutic targets. Gut. 2012;61:673–84.

2192

- 25. Beroukhim R, Mermel CH, Porter D, Wei G, Raychaudhuri S, Donovan J, et al. The landscape of somatic copy-number alteration across human cancers. Nature. 2010;463:899–905.
- 26. Zhang X, Choi PS, Francis JM, Gao GF, Campbell JD, Ramachandran A, et al. Somatic super-enhancer duplications and hotspot mutations lead to oncogenic activation of the KLF5 transcription factor. Cancer Discov. 2018;8:108–125.
- 27. Kong X, Qian J, Chen LS, Wang YC, Wang JL, Chen H, et al. Synbindin in extracellular signal-regulated protein kinase spatial regulation and gastric cancer aggressiveness. J Natl Cancer Inst. 2013;105:1738–49.
- Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell. 2011;146:353–8.
- 29. Kogo R, Shimamura T, Mimori K, Kawahara K, Imoto S, Sudo T, et al. Long noncoding RNA HOTAIR regulates polycomb-dependent chromatin modification and is associated with poor prognosis in colorectal cancers. Cancer Res. 2011;71:6320–6.
- 30. Xu TP, Wang YF, Xiong WL, Ma P, Wang WY, Chen WM, et al. E2F1 induces TINCR transcriptional activity and accelerates gastric cancer progression via activation of TINCR/STAU1/ CDKN2B signaling axis. Cell Death Dis. 2017;8:e2837.
- Liu F, Yuan JH, Huang JF, Yang F, Wang TT, Ma JZ, et al. Long noncoding RNA FTX inhibits hepatocellular carcinoma proliferation and metastasis by binding MCM2 and miR-374a. Oncogene. 2016;35:5422–34.