

## Research paper

# Identification of ARGLU1 as a potential therapeutic target for gastric cancer based on genome-wide functional screening data



Fangyuan Li<sup>a,#</sup>, Jianfang Li<sup>a,#</sup>, Junxian Yu<sup>a</sup>, Tao Pan<sup>a</sup>, Beiqin Yu<sup>a</sup>, Qingqing Sang<sup>a</sup>, Wentao Dai<sup>a,b</sup>, Junyi Hou<sup>a</sup>, Chao Yan<sup>a</sup>, Mingde Zang<sup>c</sup>, Zhenggang Zhu<sup>a</sup>, Liping Su<sup>a</sup>, Yuan-Yuan Li<sup>a,b,\*</sup>, Bingya Liu<sup>a,\*\*</sup>

<sup>a</sup> Department of Surgery, Shanghai Key Laboratory of Gastric Neoplasms, Shanghai Institute of Digestive Surgery, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, PR China

<sup>b</sup> Shanghai Center for Bioinformation Technology, Shanghai Engineering Research Center of Pharmaceutical Translation, Shanghai 201203, PR China

<sup>c</sup> Department of Gastric Cancer Surgery, Fudan University Shanghai Cancer Center; Department of Oncology, Shanghai Medical College, Fudan University, 200032 Shanghai, PR China

## ARTICLE INFO

## Article History:

Received 28 January 2021

Revised 21 May 2021

Accepted 27 May 2021

Available online xxx

## Keywords:

Gastric cancer  
Prognostic model  
Therapeutic target  
ARGLU1

## ABSTRACT

**Background:** Due to the molecular mechanism complexity and heterogeneity of gastric cancer (GC), mechanistically interpretable biomarkers were required for predicting prognosis and discovering therapeutic targets for GC patients.

**Methods:** Based on a total of 824 GC-specific fitness genes from the Project Score database, LASSO—Cox regression was performed in TCGA-STAD cohort to construct a GC Prognostic (GCP) model which was then evaluated on 7 independent GC datasets. Targets prioritization was performed in GC organoids. ARGLU1 was selected to further explore the biological function and molecular mechanism. We evaluated the potential of ARGLU1 serving as a promising therapeutic target for GC using patients derived xenograft (PDX) model.

**Findings:** The 9-gene GCP model showed a statistically significant prognostic performance for GC patients in 7 validation cohorts. Perturbation of SXX4, DDX24, ARGLU1 and TTF2 inhibited GC organoids tumor growth. The results of tissue microarray indicated lower expression of ARGLU1 was correlated with advanced TNM stage and worse overall survival. Over-expression ARGLU1 significantly inhibited GC cells viability in vitro and in vivo. ARGLU1 could enhance the transcriptional level of mismatch repair genes including MLH3, MSH2, MSH3 and MSH6 by potentiating the recruitment of SP1 and YY1 on their promoters. Moreover, inducing ARGLU1 by LNP-formulated saRNA significantly inhibited tumor growth in PDX model.

**Interpretation:** Based on genome-wide functional screening data, we constructed a 9-gene GCP model with satisfactory predictive accuracy and mechanistic interpretability. Out of nine prognostic genes, ARGLU1 was verified to be a potential therapeutic target for GC.

**Funding:** National Natural Science Foundation of China.

© 2021 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>)

## 1. Introduction

Gastric cancer (GC) is the fifth most frequently diagnosed cancer and the third leading cause of cancer death worldwide, which is responsible for over 1 000 000 new cases and an estimated 783 000

deaths in 2018 [1–3]. Surgical resection and chemotherapy are the main treatment methods for the advanced GC. The 5-year survival rate of GC is less than 30% worldwide [4,5]. Most GC cases are diagnosed at advanced stages, with consequent poor outcome; treatment is mostly restricted to cytotoxic chemotherapy. Because of the great heterogeneity and complicated pathogenesis of GC, the current uniform treatment strategy used in virtually all patients seems suboptimal [6]. Therefore, it is urgent to improve our understanding of the pathogenesis of GC and to identify more effective, less toxic therapeutic strategies. The molecular complexity and heterogeneity of GC, both inter- and intra-tumor, are major obstacles for early detection, prognostic prediction and treatment of the disease. In this sense, biomarkers for predicting prognosis and therapeutic benefits are

\* Corresponding authors at: Shanghai Center for Bioinformation Technology, Shanghai Engineering Research Center of Pharmaceutical Translation, Shanghai 201203, PR China.

\*\* Corresponding authors at: Shanghai Institute of Digestive Surgery, and Shanghai Key Laboratory of Gastric Neoplasms, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai 200025, PR China.

E-mail addresses: [yyli@sbit.org](mailto:yyli@sbit.org) (Y.-Y. Li), [liubingya@sjtu.edu.cn](mailto:liubingya@sjtu.edu.cn) (B. Liu).

# These authors contributed equally to this work.

## Research in context

### Evidence before this study

Mechanistically interpretable biomarkers were required for predicting prognosis and discovering therapeutic targets for GC patients. However, current efforts are always focusing on predictive accuracy over explanatory power, and discovering more effective therapeutic targets for GC patients still demands a lot of efforts.

### Added value of this study

In the present study, we constructed a GCP model be of mechanistic interpretability and satisfactory predictive accuracy based on genome-wide functional screening data. We also identified four potential therapeutic targets including ARGLU1, DDX24, TTF2 and SSX4 for GC patients based on GCP model. The specific molecular mechanism of ARGLU1 inhibiting GC tumor growth was also further explored.

### Implications of all the available evidence

To the best of our knowledge, the present study is the first to integrate biomarkers with therapeutic targets selection. The GCP model proposed in our study, besides its prognostic efficacy, may help to stratify patients for treatment targeting to ARGLU1, DDX24, TTF2 or SSX4.

## 2.3. Construction of the gastric cancer prognostic (GCP) model

Univariate Cox regression was performed in TCGA-STAD training cohort to identify prognostic genes. And forty prognostic genes meeting the statistical significance ( $P < 0.05$ ) were then obtained to conduct the least absolute shrinkage and selection operator (LASSO) regression. Eventually, the prognostic model composed of  $\beta_i$  and gene expression level was constructed as follows: GCP score =  $\sum(\beta_i \cdot Exp_i)$ . Patients were classified into a high GCP score and low GCP score group by median. R package survivalROC was used to perform the time-dependent receiver operating characteristic (ROC) curve to evaluate the predictive value of the prognostic model for overall survival. The Kaplan-Meier survival analysis was conducted using R package survival.

## 2.4. Gene set enrichment analysis (GSEA)

For the mechanistic exploration of ARGLU1, gene set enrichment analysis (GSEA) was performed in TCGA-STAD cohort between high- and low ARGLU1 phenotypes to analyze the KEGG pathways between two groups. GSEA software with the Java platform and the MSigDB database as the reference set (c2.cp.kegg.v7.2.symbols.gmt gene sets) were downloaded. The top 5 positively and top 15 negatively correlated KEGG pathways were plotted by R package ggplot2. In addition, the enriched gene sets were considered to be significant when FDR was less than 0.25.

## 2.5. Cell culture

Eight GC cell lines including MKN28 (RRID: CVCL\_1416), MKN45 (RRID: CVCL\_Y476), NUGC3 (RRID: CVCL\_1612), MGC803 (RRID: CVCL\_5334), AGS (RRID: CVCL\_0139), KATOIII (RRID: CVCL\_0371), HS746T (RRID: CVCL\_0333) and HGC27 (RRID: CVCL\_1279) were cultured in RPMI-1640 medium (Gibco BRL, San Francisco, CA, USA) containing 10% fetal bovine serum (Invitrogen, Carlsbad, CA, USA) at 37 °C in a 5% CO<sub>2</sub>-humidified incubator. Cells used in this study were purchased from ATCC (USA) and authenticated by STR profiling according to the cell bank.

## 2.6. Patient derived organoid (PDO) culture

The tumor tissues and its paired non-tumor tissues from GC patients were taken out and washed in PBS, removing adipose and connective tissue. Tissues were cut with a scalpel into 2–4 mm pieces and transferred into 20 mL PBS/FBS in a 50 mL falcon tube.

The tissue samples were allowed to sediment and the supernatant was removed. Additionally, 10 mL of PBS/FBS was added to wash the tissue and the supernatant was removed. These steps were repeated until the supernatant was clear (approx. 6 times). The remaining tissue was digested in PBS + EDTA (2 mM) for 20 min at 4 °C with shaking. After digestion the supernatant was removed. The tissue was resuspended in 10 mL PBS/FBS and filtered through a 70 μm cell strainer (BD Biosciences) into a new 50 mL falcon tube (repeated several times). After centrifuging for 8 min, 800 rpm at 4 °C the supernatant was removed as completely as possible, and 50–100 μL Matrigel was added to the pellet avoiding air bubbles. Fifty μL Matrigel was plated in every well of a prewarmed 24-well plate. In every well 0.4 mL of crypt culture medium (CM) was added, which was composed of advanced DMEM/F12 (Life Technologies), supplemented with serum-free B27 (1 : 50, Life Technologies), N2 (1 : 100, Life Technologies), N-acetylcysteine (50 mM, Sigma), recombinant murine epithelial growth factor (EGF 50 ng/mL, Peprotech), Noggin (100 ng/mL, Peprotech), R-Spondin (1 μg/mL, Peprotech), Glutamax-I Supplement (1 : 100, Life Technologies), Penicillin/Streptomycin (400 μg/mL, Life Technologies), and HEPES (10 μM, Life Technologies). The cultures were incubated in a humidified incubator at 37 °C and 5% CO<sub>2</sub>.

indispensable for implementing cancer precision medicine [7]. The most widely adopted strategies for biomarker identification heavily depend on the identification of differentially expressed genes (DEGs) [8,9], and accordingly have limited mechanistic interpretability [10]. Identifying biomarkers with mechanistic interpretability as well as predictive accuracy for precision medicine has become a general problem for oncologists and biologists [10–12].

A very recent far-reaching research by Behan FM and colleagues [13] developed a CRISPR-Cas9 based functional genomics approach to discover cancer treatment targets. By integrating cell fitness effects with genomic features and target tractability for drug development, candidate therapeutic targets were systematically prioritized [13]. In this study, we attempted to identify biomarkers with mechanistic interpretability based on this genome-wide functional screening database and construct a GC prognostic model to predict patient's outcome and discover potential therapeutic targets for GC.

## 2. Methods

### 2.1. Data processing

Transcriptome profiling raw data with corresponding clinical information were downloaded from TCGA (<https://cancergenome.nih.gov/>) and the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) databases. From TCGA, we obtained 380 STAD (Stomach Adenocarcinoma) patients as training cohort (Table S1). Gene expression datasets GSE57303, GSE62254, GSE84437, GSE15459, GSE26253, GSE29272 and GSE34942 including 1614 GC patients from GEO database were acquired as validation cohorts (Table S3).

### 2.2. Gastric cancer specific fitness genes extraction

A total of 823 gastric cancer specific fitness genes (with the priority score  $\geq 20$ ) were retrieved from the Project Score database (<https://score.depmap.sanger.ac.uk/>).

## 2.7. Western blot

Cell sample were washed 3 times with 1 x PBS and protein extracts were prepared in RIPA cell lysis buffer (Kangwei, Beijing, China) supplemented with phosphatase inhibitor Cocktail III (Roche). The concentration of protein sample was quantified by using bicinchoninic acid protein assay kit (Pierce, Rockford, IL, USA) against a bovine serum albumin standard curve. A total of 20  $\mu\text{g}$  of protein was loaded onto a 10% sodium dodecyl sulfate polyacrylamide gel and transferred onto 0.22  $\mu\text{m}$  PVDF membranes (Millipore, MA, USA). The membranes were then blocked with 1 x TBST buffer containing 5% nonfat milk and incubated with corresponding antibodies at 4 °C overnight. Anti-GAPDH (CAT# HRP-60,004, use a concentration of 0.02  $\mu\text{g}/\text{mL}$ ) was purchased from Proteintech (Rosemont, IL, USA). Anti-ARGLU1 (CAT# ab272660, use a concentration of 0.1  $\mu\text{g}/\text{mL}$ ), anti-YY1 (CAT# ab109237, use a concentration of 0.1  $\mu\text{g}/\text{mL}$ ) and anti-SP1 (CAT# ab231778, use a concentration of 0.1  $\mu\text{g}/\text{mL}$ ) were purchased from Abcam and HRP-conjugated secondary antibodies (use a concentration of 0.02  $\mu\text{g}/\text{mL}$ ) were purchased from Proteintech (Rosemont, IL, USA). Membranes were then exposed to HRP-conjugated secondary antibody (32,460, Thermo Fisher Scientific) and developed with Thermo Pierce chemiluminescent (ECL) Western Blotting Substrate (Thermo, Waltham, MA, USA). Membranes were imaged with Tanon 5200 system (Tanon, Shanghai, China).

## 2.8. ChIP-Seq libraries and massive parallel sequencing

ChIP was performed as described above. DNA (15 ng) was quantified by fluorimetry, electrophoresis was resolved and fractions of 50–300 bp were extracted. DNA was blunted and phosphorylated with T4 DNA polymerase, Klenow DNA polymerase and T4 polynucleotide kinase. An A nucleotide was then added to the 3' end of the DNA fragments. Adaptor-ligated libraries were amplified by 20 cycles of PCR amplification. Purified DNA library was applied to an Illumina flow cell for cluster generation (TruSeq cluster generation kit v5) and subjected to massively parallel sequencing (Illumina Genome Analyzer IIx) following the manufacturer's protocols.

## 2.9. Chromatin immunoprecipitation (ChIP) and qPCR

GC cells were transfected with shARGLU1, saARGLU1 or with the nontargeting negative sh/sa-control and were grown in 10 cm plates. After formaldehyde fixation, ChIP was performed according to the manufacture's instruction (SimpleChIP, CAT# 9005, CST). Chromatin was then immunoprecipitated with anti-ARGLU1 (CAT# ab272660, Abcam, use a concentration of 10  $\mu\text{g}/\text{mL}$ ), anti-YY1 (CAT# 63227S, CST, use a concentration of 10  $\mu\text{g}/\text{mL}$ ), anti-SP1 (CAT# 9389S, CST, use a concentration of 10  $\mu\text{g}/\text{mL}$ ) or IgG antibody (CAT# 3900S, CST, use a concentration of 10  $\mu\text{g}/\text{mL}$ ) preincubated with protein G coated magnetic beads. The beads were washed, and immunoprecipitated DNA was eluted, de-crosslinked and purified. Primers were designed according to ChIP-seq peaks on target genes' promoter regions as templates (Supplementary Fig. 2b). Purified DNA was quantified by genomic qPCR using serial dilutions of the input as a standard curve, performed in triplicates. Results are expressed as the percent enrichment of bound DNA compared to each input.

## 2.10. Immunoprecipitation

Immunoprecipitation was performed using Pierce Crosslink Immunoprecipitation Kit according to manufacturer's protocol. Briefly, 5–10  $\mu\text{g}$  antibody (antibody against ARGLU1, YY1 or SP1, use a concentration of 20  $\mu\text{g}/\text{mL}$ ) was covalently crosslinked with magnetic protein A/G beads which were initially blocked by 3% Block ACE solution (Bio-Rad) for 3 h at room temperature with gently constant shaking. Subsequently, the beads were washed with TBS and

incubated with 1 mg of total cell lysate for 10 h at 4 °C with constant shaking. Antigen was eluted and subjected to protein blot analysis with antibody against ARGLU1, YY1 or SP1.

## 2.11. Immunoprecipitation followed by mass spectrometry (IP-MS)

Immunoprecipitation was performed by using antibody against ARGLU1 according to protocol described above. Antigen was eluted with a mixture of 74% acetonitrile, 25.9% water and 0.2% formic acid, mixed in equal volume with saturated sinapinic acid in 50% acetonitrile, 49% water and 1% trifluoroacetic acid, and spotted (2 ml) onto a ProteinChip Gold Array for analysis with a Bio-Rad Protein Chip System Series 4000 mass spectrometer.

## 2.12. Lentiviral constructs and infection

All shRNA and saRNA constructs in the pLV-EGFR(2A)-puro vector targeting ARGLU1, TTF2, DDX24 or SSX4 were purchased from the BioeGene (BioeGene Co. Ltd., Shanghai, China). Detailed information of these shRNA and saRNA sequences see Supplementary Table 4.

We produced shRNA/saRNA-expressing lentivirus with the third-generation packaging system in human embryonic kidney (HEK) 293T cells (ATCC, CAT# CRL-11,268, RRID:CVCL\_1926). Briefly, 60–80% confluent 293T cells in 6-well plate were transiently co-transfected with 5  $\mu\text{g}$  of lentiviral transfer vector, 1.67  $\mu\text{g}$  of pVSVG (envelope plasmid), 1.67  $\mu\text{g}$  of pRSV-Rev (packaging plasmid) and 1.67  $\mu\text{g}$  of pMDLg/pRRE (packaging plasmid) with Lipofectamine™ 2000 according to the manufacturer's instructions. Medium was replaced 24 h after transfection with RPMI-1640 medium containing 10% FBS, and virus supernatant was collected every 12 h for up to 3 days. Supernatant containing viral particles was used directly for cell infection or stored at –70 °C in aliquots.

For shRNA/saRNA lentivirus infection, target cells were seeded in a 6-well plate 24 h before infection and were grown to 60–80% confluency upon transduction. Culture medium was removed, and cells were incubated with virus supernatant along with 10  $\mu\text{g}/\text{ml}$  polybrene (Sigma) overnight. Virus-containing medium was replaced with fresh medium. Puromycin (Sigma) (2  $\mu\text{g}/\text{ml}$  for MGC803 and HGC27 cells, 12  $\mu\text{g}/\text{ml}$  for MKN28 and MKN24 cells) was applied to kill non-infected cells 48 h after infection to produce stably transfected cells (HGC27/Ctrl, HGC27/ARGLU1-shRNA#1, HGC27/ARGLU1-shRNA#3, MKN45/Ctrl and MKN45/ARGLU1-shRNA#1, MKN45/ARGLU1-shRNA#3, MGC803/Ctrl, MGC803/ARGLU1-saRNA#2, MGC803/ARGLU1-saRNA#3, MKN28/Ctrl, MKN28/ARGLU1-saRNA#2, MKN28/ARGLU1-saRNA#3) for further experiments.

For saRNA lentivirus infection in organoids, digested cells were seeded in a prewarmed 24-well plate and cultured with virus-containing medium. 6 h later, virus-containing medium was replaced with fresh medium to increase corresponding gene's mRNA transcription level.

## 2.13. Immunohistochemistry (IHC)

Gastric tumor and non-tumor tissues were collected from patients under surgery at Department of Surgery, Ruijin Hospital. None of the patients received chemotherapy or radiotherapy prior to surgery. A total of 96 pairs of tumor and non-tumor tissues were collected to construct tissues array for further immunobiological assay. Tissue array and tissue sections were deparaffinized in xylene and dehydrated before antigen retrieval for 10 min with an autoclave. Hydrogen peroxide was utilized to block endogenous peroxidase activity; nonspecific immunoglobulin binding sites were blocked by normal goat serum for 25 min at 37 °C. Tissue sections were incubated with antibodies against ARGLU1 (dilution 1:200, Abcam, CAT# ab272660), MLH3 (dilution 1:200, Abcam, CAT# ab4834), MSH2 (dilution 1:200, CST, CAT# 2017), MSH3 (dilution 1:200, Abcam, CAT# ab275928) and

MSH6 (dilution 1:200, CST, CAT# 5424) overnight at 4 °C. Biotinylated goat anti-mouse immunoglobulin G was used as a secondary antibody. After washing, the sections were incubated with streptavidin–biotin conjugated with HRP for 25 min, and the peroxidase reaction was performed with diaminobenzidine tetrahydrochloride. For the tissue array, staining intensity and the proportion of cell staining were used to score the overall tissue sections. The staining intensity was graded in 3 segments on a 3-point scale (staining scores): no staining (1 points), light brown (2 point), dark brown (3 points). The number of positive cells was divided into four grades (percentage scores): 0–25% (1), 26–50% (2), 51–75% (3) and 76–100% (4).

#### 2.14. Quantitative real-time PCR (qRT-PCR)

Total RNA was extracted utilizing TRIzol reagent (Invitrogen) and cDNA synthesis was performed by using a reverse transcription kit (Promega, Madison, WI, USA) according to the manufacturer's instructions. The mRNA expression levels of the 9 genes were measured using the SYBR Green PCR Master Mix (Applied Biosystems, Waltham, MA, USA) and the Applied Biosystems 7900HT sequence detection system (Applied Biosystems). And mRNA relative expression level was measured using the  $2^{-\Delta\Delta Ct}$  method and normalized to glyceraldehyde 3-phosphate dehydrogenase (GAPDH).

#### 2.15. Cell proliferation, colony formation assays

GC cells were seeded in 96-well plates at a density of 1000/well (200  $\mu$ l/well). A cell proliferation assay was conducted utilizing the cell counting kit8 (Dojindo, Kumamoto, Japan) according to the manufacturer's protocol. After being incubated with 20  $\mu$ l of CCK-8 reagent for 2 h, OD450 was then measured by spectrophotometry (BioTek, Vermont, USA).

For colony formation assay, cells were seeded at a density of 1000 cells/well into 6-well plates and incubated at 37 °C for 10 days. Cells were then washed twice with PBS and fixed in 100% methanol for 15 min, prior to staining with Giemsa solution for 20 min. The number of colonies containing  $\geq 50$  cells were counted under a microscope (IX71). The clone formation assay was applied to ARGLU1-over-expression/ARGLU1-knockdown (ARGLU1-OE/ARGLU1-KD) treated HGC27, MKN28, MKN45 and MGC803 cells and their corresponding negative control. All experiments were performed in triplicate.

#### 2.16. In vivo tumorigenesis

Male BALB/c nude mice (RRID: IMSR\_JCL:JCL:mID-0001, 4-weeks-old, purchased from Institute of Zoology, Chinese Academy of Sciences, Shanghai, China) were housed in a specific pathogen-free room in the Animal Experimental Center, Ruijin Hospital, Shanghai Jiao Tong University School of Medicine, China. Animal experiments were performed in accordance with the animal research principles and the Institution's guidelines. Twenty mice were randomly divided into four groups (five mice per group). The single blind method was adopted in our experiments. Mice were subcutaneously injected with  $2 \times 10^6$  tumor cells (HGC27/ARGLU1-shRNA, HGC27/ARGLU1-NC, MGC803/ARGLU1-saRNA, and MGC803/ARGLU1-NC) suspended in 150  $\mu$ l PBS (five mice per group). Tumor length (L) and width (W) were measured every 4 days using digital Vernier caliper. Tumor volume was determined using the following formula: volume = Length  $\times$  Width<sup>2</sup>/2. All mice were sacrificed under general anesthesia 4 weeks after injection. Tumor grafts were weighed and observed systematically.

#### 2.17. Reporter plasmid construction and activity assays

The fragments relative to the TSS (transcription start site) of human MSH2, MSH3, MSH6 and MLH3 genomic sequence (–2445 ~

+55) were synthesized by BioeGene (Shanghai, China). The fragments were cloned into the pGL3-basic vector (Promega, E1751) to generate reporters which were then verified by DNA sequencing. The relative luciferase activity was assessed by using luciferase assay detection kit (Beyotime Biotechnology, no. RG027). Briefly, cells were transfected with different reporter plasmids or pGL3-basic vector as a negative control. Cells were lysed after 48 h post-transfection, mixed with the dual luciferase assay reagent. Relative luciferase activity was calculated by normalizing the firefly luminescence to the Renilla luminescence.

#### 2.18. Patient derived xenograft (PDX) model construction and saRNA treatment

Fresh GC tissues were collected and washed with PBS containing  $1 \times$  Penicillin/Streptomycin for 3 times, followed by removing the muscle layer, adipose and connective tissue and gastric mucus. Tumor tissues were then cut into 1–2 mm pieces using scissors and inoculated subcutaneously in NOD-scid mice (RRID: BCBC\_1262) to establish PDX models. Routine pathological analyses were performed on successfully established PDX tumors and their parental tumor tissues. For further experiments, xenograft tumors were excised and cut into 1–2 mm pieces, followed by transplanting into nude mice (BALB/c nude mice, RRID: IMSR\_JCL:JCL:mID-0001). Ten Male BALB/c mice were subcutaneously transplanted xenograft tumor tissue to establish PDX models. Small activating RNA (saRNA) of ARGLU1 was designed and biosynthesized by BioeGene (Shanghai, China). For intratumoral delivery of lipidoid-formulated saRNA. Eight days after inoculation (tumor length > 2 mm), mice were randomly divided into two groups (saRNA-NC and saRNA-ARGLU1) each containing 5 animals. Lipidoid-formulated saRNA was administered via intratumoral injection at 5 mg/kg every 4 days for total 3 doses. Tumor length (L) and width (W) were measured every 4 days using digital Vernier caliper. Tumor volume was determined using the following formula: volume = Length  $\times$  Width<sup>2</sup>/2. All mice were sacrificed under general anesthesia 4 weeks after injection. Tumor grafts were weighed and observed systematically.

#### 2.19. Statistical analysis

Statistical analyses were performed with R (version 3.3.1) and GraphPad Prism 8.0. Two-tailed *t*-test was performed in clone formation assay and for CCK8 assay two-way ANOVA was conducted. Pearson correlation analysis was used to determine the correlation between two variables. Survival analysis was performed by using a log-rank test.  $P < 0.05$  was considered statistically significant.

#### 2.20. Ethics statement

Our study was approved by an institutional review board from the Human Research Ethics Committee of Ruijin Hospital (Approval No. 2020–115) and conducted in accordance with ethical guidelines (Declaration of Helsinki). Animal experiments were approved by the local Laboratory Animal Ethics Committee of Ruijin Hospital and conducted in accordance with animal use guidelines.

#### 2.21. Role of the funding source

This work was supported by grants from the National Natural Science Foundation of China (No. 81772509 and 82072605 to B Liu, No. 81902403 to MD Zang, No. 81,871,904 to ZG Zhu, No. 81871902 to L Su, No. 81902393 BQ Yu and No. 32000472 to W Dai), and Natural Science Foundation of Shanghai (No. 19ZR1431700 to BQ Yu). The funders were not involved in the study design, implementation, the analysis or the interpretation of data.

### 3. Results

#### 3.1. Constructing gastric cancer prognostic (GCP) model

The schematic overview of this study was shown in Fig. 1. We extracted 823 GC specific fitness genes with priority score not less than 20 in Project Score (<https://score.depmap.sanger.ac.uk/>) as a starting gene set for identifying potential therapeutic targets [13]. Functional enrichment analysis revealed that these 823 genes were mainly enriched in “Cell cycle”, “DNA transcription or replication”, and “RNA metabolism” (Fig. 2a). The molecular interaction network among the genes mapped to the enriched pathways was constructed by using Metascape (Fig. 2a).

Univariate cox regression analysis was first carried out in TCGA-STAD cohort (Table S1), leading to a total of 40 prognosis-related genes (Table S2). LASSO-Cox regression method was then performed on the 40 genes and generated a GC prognostic (GCP) model involving nine prognostic genes, ALCY, MRPL4, DCTN2, DDX24, ARGLU1, TTF2, POTEJ, SSX4, and VPS35 (Fig. 2b, c). The hazard ratio (HR) values with 95% confidence interval of the nine genes were listed in Fig. 2d. According to HR values, MRPL4, ARGLU1 and TTF2 were tumor suppressors (HR<1) in GC; while ALCY, DCTN2, DDX24, POTEJ, SSX4 and VPS35 functioned as proto-oncogenes (HR>1) in GC.

#### 3.2. Validating GCP model in independent cohorts

Seven independent datasets, which totally consisted of 1 614 GC patients (mean age  $\pm$  SD, 64.3  $\pm$  11.2; 42.97% I+II stage, 57.03% III+IV stage [III, IV]), were adopted to validate our GCP model (Table S3). GCP score for each patient was calculated with the formula: GCP score = 0.9145 \* Exp<sub>ALCY</sub> - 0.1134 \* Exp<sub>MRPL4</sub> + 0.2964 \* Exp<sub>DCTN2</sub> + 0.4580 \* Exp<sub>DDX24</sub> - 0.3401 \* Exp<sub>ARGLU1</sub> - 0.2921 \* Exp<sub>TTF2</sub> + 0.9307 \* Exp<sub>POTEJ</sub> + 1.2053 \* Exp<sub>SSX4</sub> + 0.5038 \* Exp<sub>VPS35</sub> (Fig. 2e). All eligible patients were divided into high-GCP and low-GCP score groups by the median GCP score. In line with the outcomes of the training cohorts, GC patients assigned to the high-GCP score group had significantly worse OS than those assigned to the low-GCP score group in validation cohorts ( $p < 0.05$ , log-rank test, Fig. 3a). The prediction efficacy of the GCP model was validated with the AUCs of ROC curves varying from 0.71 to 0.75 for 3-year OS events, and 0.64 to 0.73 for 5-year OS events on the seven independent cohorts (Fig. 3b).

#### 3.3. Prioritizing the nine potential therapeutic targets for GC

In order to prioritize the nine potential therapeutic targets involved in the GCP model, we collected flesh tissue samples from three GC patients and designed a workflow which was briefly shown in Fig. 4a. Initially, we detected relative mRNA levels of the nine genes in tumor tissue normalized to the paired non-tumor tissue (Fig. 4b), and selected four genes, ARGLU1, DDX24, TTF2 and SSX4, which were differentially expressed with  $p$  value less than  $10^{-3}$  at least in two patients (Fig. 4c). We then separately validated the functions of the four genes with patient derived organoid (PDO) models derived from the three GC patients by up-regulating the expression of tumor suppressor genes (ARGLU1 or TTF2) with small activating RNAs (saRNAs) or knocking-down the expression of proto-oncogenes (SSX4 or DDX24) with short-hairpin RNAs (shRNAs) (Fig. 4d). It was found that ARGLU1-induction, TTF2-induction or SSX4 knock-down could significantly inhibit tumor growth in two PDO models, while have no effect in the third PDO model (Fig. 4e). Next, we calculated GCP scores of the three patients by measuring relative mRNA expression levels of the nine genes normalized to GAPDH in tumor tissues and found that the two models with inhibition effects had relatively higher GCP score (Organoid 1, 2.93, Organoid 2, 3.61) while the third model with no significant effect have relatively lower GCP score

(Organoid 3, 0.67) (Fig. 4a). This observation is in line with the trend in Fig. 4e, where organoid 3 displayed the modest differential expression across the nine genes. We therefore propose that the ARGLU1, TTF2 and SSX4 could be potential therapeutic targets for GC.

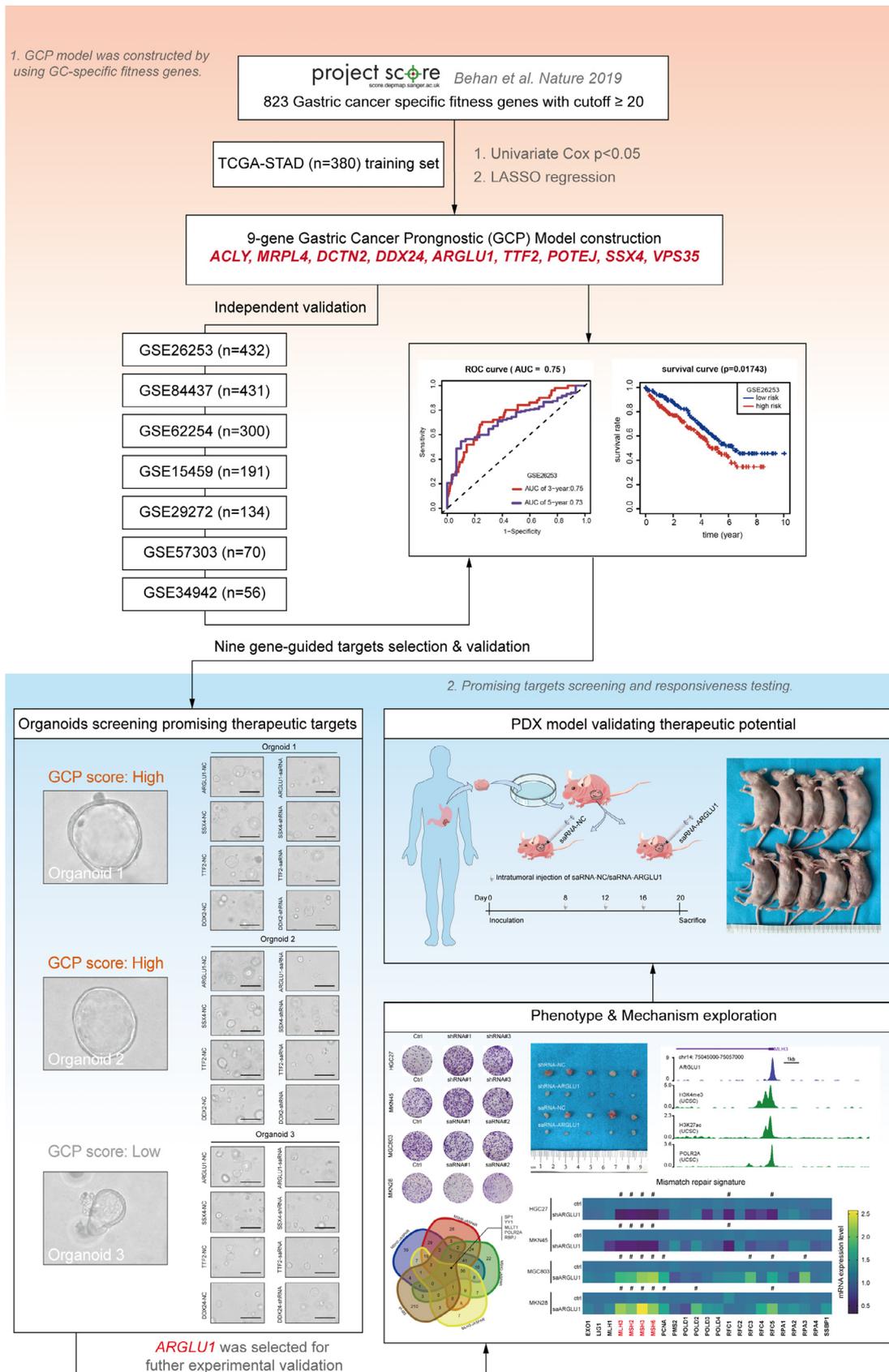
#### 3.4. Overexpression of ARGLU1 inhibited GC cells proliferation in vitro and in vivo

Among the above three potential targets, TTF2 [14, 15] and SSX4 [16] have been reported in several cancers, but there was few report on biological function of ARGLU1, which was as well the most differentially expressed gene across the three patients. We then focused on ARGLU1 for its role in GC pathogenesis. ARGLU1 mRNA expression was significantly down-regulated in tumor tissues compared with its paired non-tumor tissues in 28 out of 31 cancer types according to TCGA cohort + GTEX data, and among these cancer types, ARGLU1 expression in GC ( $N = 211$ ,  $T = 408$ ) was down-regulated (Fig. S1a). In GSE54129 dataset (Ruijin-cohort, mRNA level, 111 tumor samples of GC, and 21 normal stomach mucosa control), ARGLU1 was also found significantly down-regulated in tumor tissues (Fig. S1b). To reveal the clinicopathologic significance of ARGLU1 in GC, tissue microarray immunohistochemistry (IHC) assay for tumor tissues and its corresponding non-tumor tissues from 96 patients was performed (Ruijin-cohort protein level, Fig. 5a). ARGLU1 protein expression was very low in tumor tissues, but relatively high in the adjacent non-tumor tissues (Fig. 5b). According to Pearson chi-square test, we found that low expression of ARGLU1 was associated with advanced T and N stage ( $p < 0.05$ ,  $p < 0.001$ ) (Fig. 5c). Kaplan-Meier analysis showed a positive correlation between high expression of ARGLU1 and prolonged survival time both in TCGA-STAD cohort (transcriptional level) and Ruijin-cohort (protein level). This indicated that downregulated ARGLU1 in tumor tissues predicted a poor prognosis for GC patients (Fig. 5d).

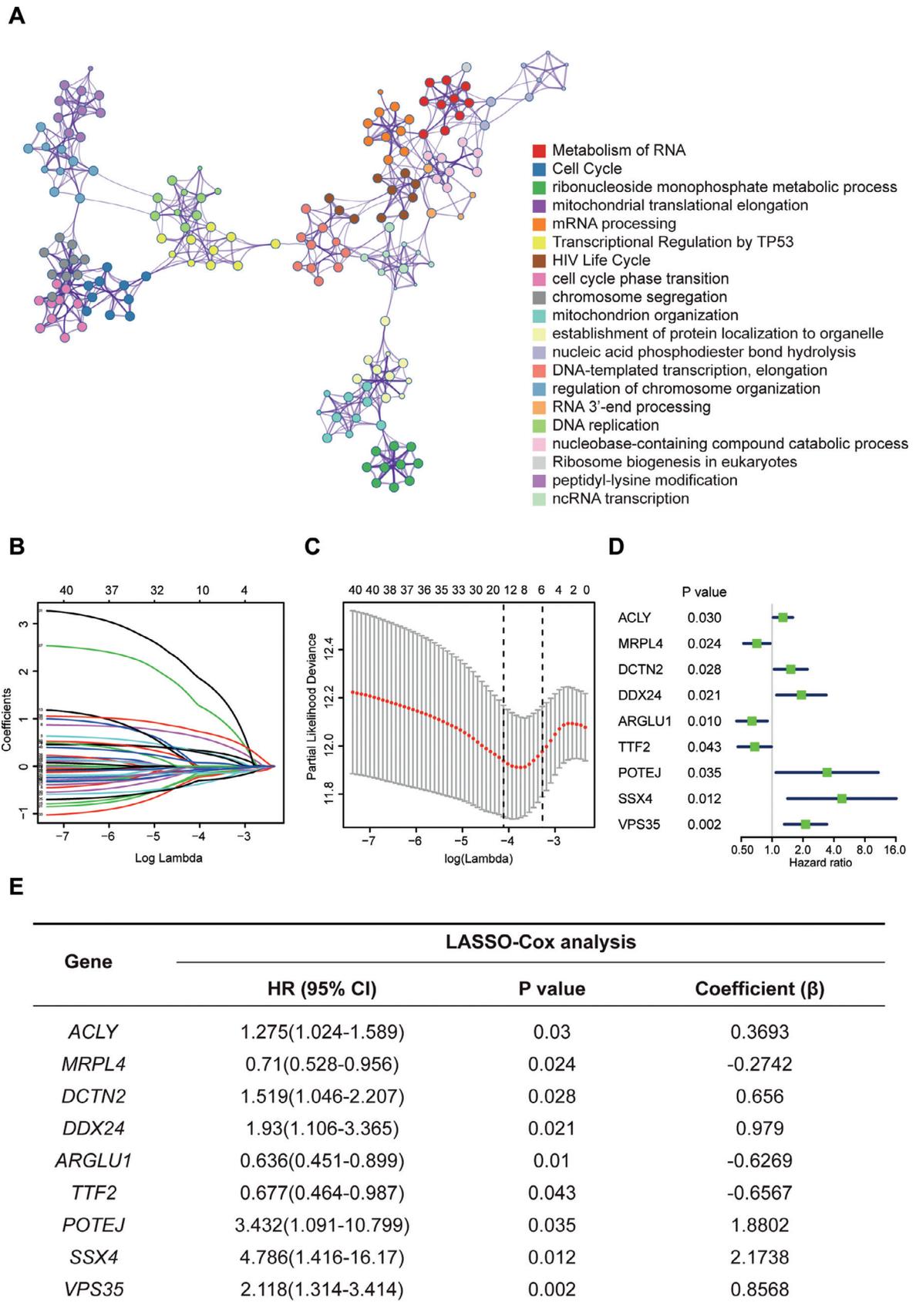
To investigate biological function of ARGLU1 in GC, we established ARGLU1 knock-down cell lines (HGC27/ARGLU1-shRNA#1, HGC27/ARGLU1-shRNA#3, MKN45/ARGLU1-shRNA#1 and MKN45/ARGLU1-shRNA#3) by lentivirus infection, and ARGLU1 over-expression cell lines by lentivirus infecting saRNA in MGC803 and MKN28 (MGC803/ARGLU1-saRNA#1, MGC803/ARGLU1-saRNA#2, MKN28/ARGLU1-saRNA#1 and MKN28/ARGLU1-saRNA#2) (Fig. S1c-e, Fig. 5e, f). We found that knock-down ARGLU1 could significantly increase GC cells viability, while upregulated ARGLU1 attenuated the proliferation of GC cells both in cell proliferation and clone formation assays (Fig. 5g, h). Moreover, we also subcutaneously transplanted the HGC27/ARGLU1-shRNA, HGC27/ARGLU1-NC, MGC803/ARGLU1-saRNA, and MGC803/ARGLU1-NC into the nude mice (Fig. 5i, Fig. S1f). The average weight of tumors showed that down-regulated ARGLU1 significantly accelerated tumor growth (HGC27/ARGLU1-shARGLU1 vs. HGC27/ARGLU1-NC) and ARGLU1 overexpression dramatically inhibited tumor growth (MGC803/ARGLU1-NC vs. MGC803/ARGLU1-saRNA) (Fig. 5j, k). Kaplan-Meier survival analysis revealed that down-regulated ARGLU1 was correlated with the worse overall survival of nude mice (Fig. S1g).

#### 3.5. ARGLU1 enhanced mismatch repair genes transcription in GC cells

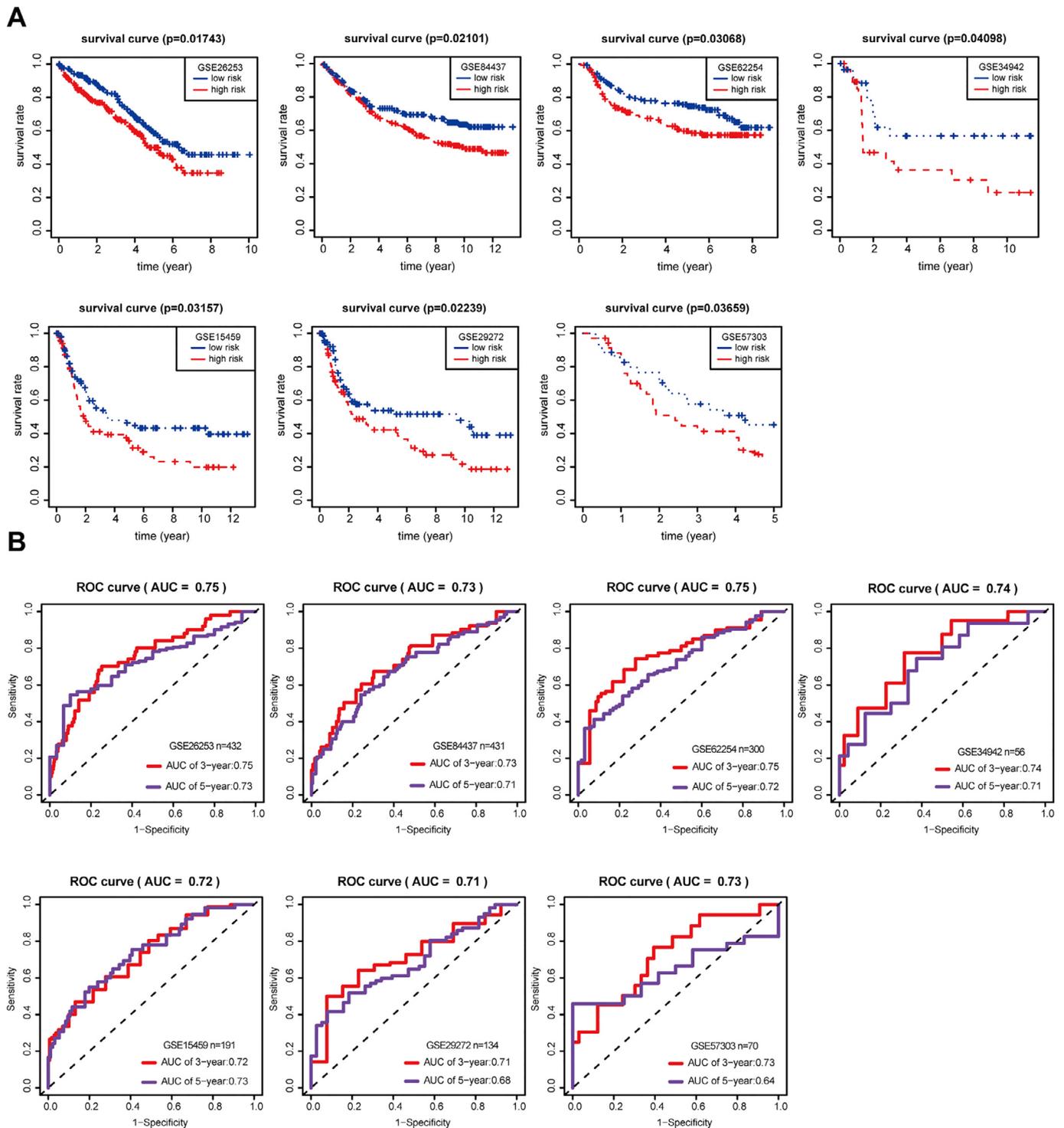
To understand the molecular mechanism of ARGLU1 in GC, GSEA was performed by using TCGA-STAD cohort (Fig. 6a). The results showed that ARGLU1 overexpression was positively correlated with spliceosome which was consistent with previous research [17,18] that ARGLU1 was a splicing regulator. Interestingly, “Mismatch repair (MMR)”, “Nucleotide excision repair” and “Base excision repair” were frequently enriched indicating ARGLU1 played a crucial role in repairing DNA mismatch repair. DNA mismatch repair (MMR) is an evolutionarily conserved process that is responsible for recognizing and repairing single-base mismatches and small insertion/deletion loops



**Fig. 1.** Flowchart of the study. Eight public gastric cancer datasets containing 1994 cases were included in this study. GCP model was developed in TCGA-STAD training set and validated in the other seven independent testing sets. Further, nine-gene based potential therapeutic targets screening was performed, and ARGLU1 was selected for biological function and mechanism exploration. PDX model was then used to validate the effect of ARGLU1-targeted therapeutic strategy.



**Fig. 2.** Construction of the GCP model. (a) GC specific fitness genes (823) obtained from Project Score database with the cutoff priority score  $\geq 20$ . (b, c) LASSO regression identified the 40 prognostic genes using TCGA-STAD as training cohort and the minimum account of genes was 9. (d, e) LASSO-Cox regression results of the GC-specific fitness genes (CI: confidence interval, HR: hazard ratio).



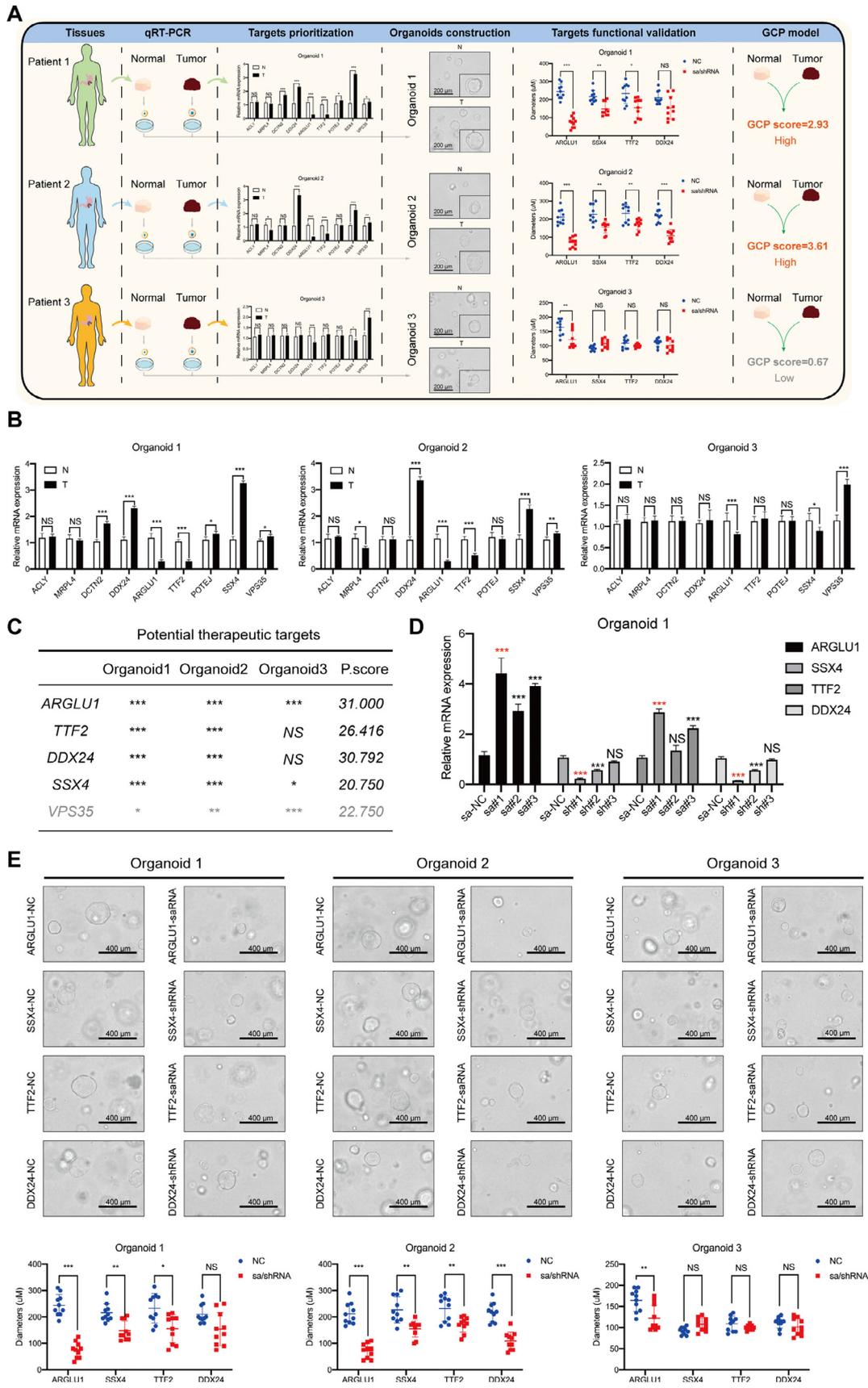
**Fig. 3.** Performance validation of GCP model in seven independent cohorts. (a) Kaplan-Meier analysis OS in testing cohorts (log-rank test). (b) The area under the curve (AUC) of receiver operating characteristic (ROC) curves predicting 3- and 5-year OS events in testing cohorts.

that are formed during DNA replication [19]. We detected MMR signature (23 genes in total) mRNA expression levels in 4 GC cells with ARGLU1 knock-down or overexpression by qRT-PCR and top 4 most ARGLU1-dependent genes including MLH3, MSH2, MSH3 and MSH6 were selected for further validation (#, |FoldChange| > 1.5,  $p < 0.01$ , two-tailed  $t$ -test, Fig. 6b). Chromatin immunoprecipitation followed by massive parallel sequencing (ChIP-Seq) in HGC27 cells using anti-ARGLU1 antibody, showed that ARGLU1 could bind the promoter regions (marked by H3K4me3, H3K27ac and POLR2A ChIP-peaks derived from UCSC) of MLH3, MSH2, MSH3 and MSH6 (Fig. 6c).

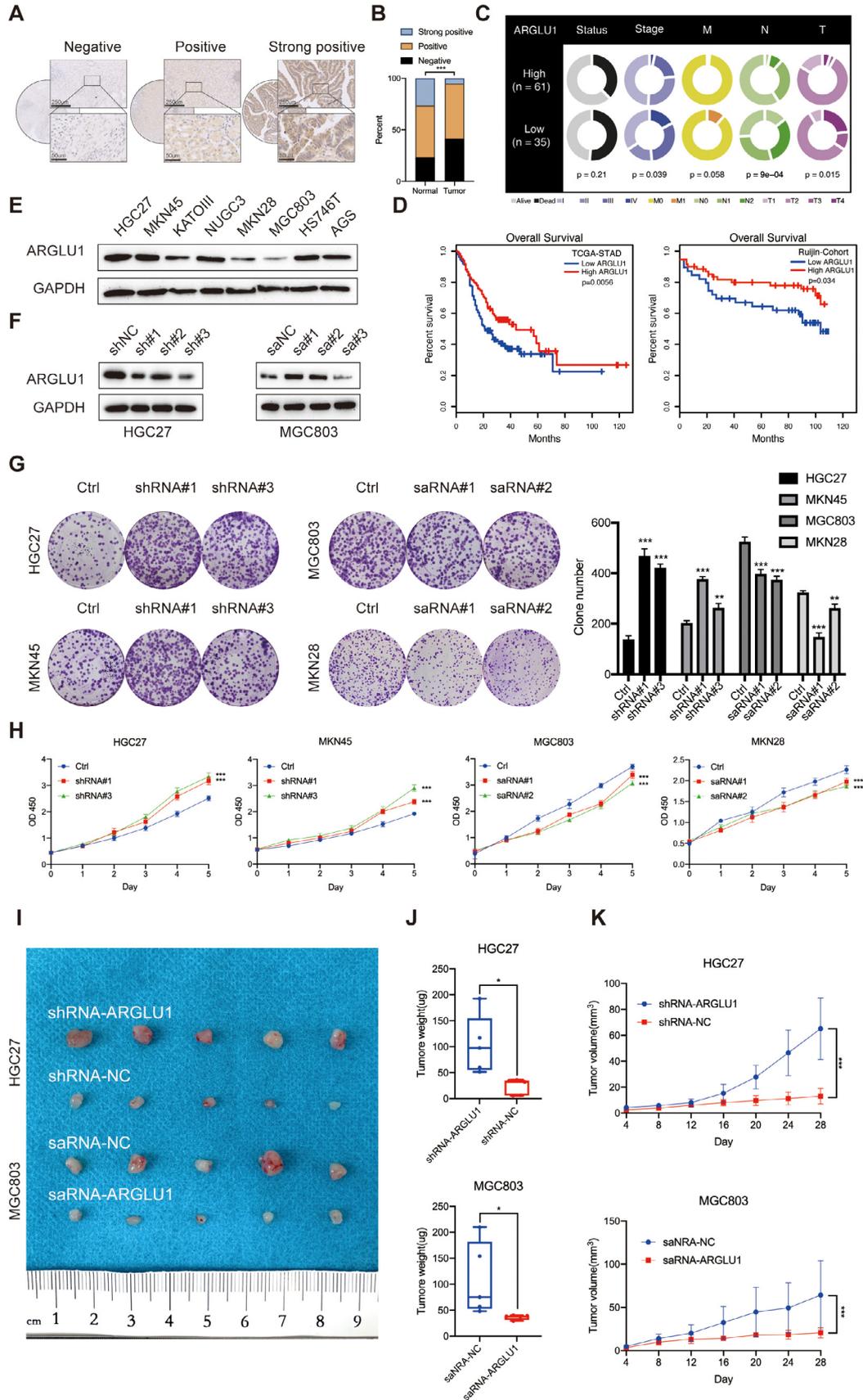
Furthermore, luciferase reporters assay in HGC27 and MGC803 cells with ARGLU1 knock-down or overexpression showed that the transcription of MLH3, MSH2, MSH3 and MSH6 was highly ARGLU1-dependent (Fig. S2a, Fig 6d).

### 3.6. ARGLU1 enhanced the transcriptional activity of AP1 and YY1 to MLH3, MSH2, MSH3 and MSH6

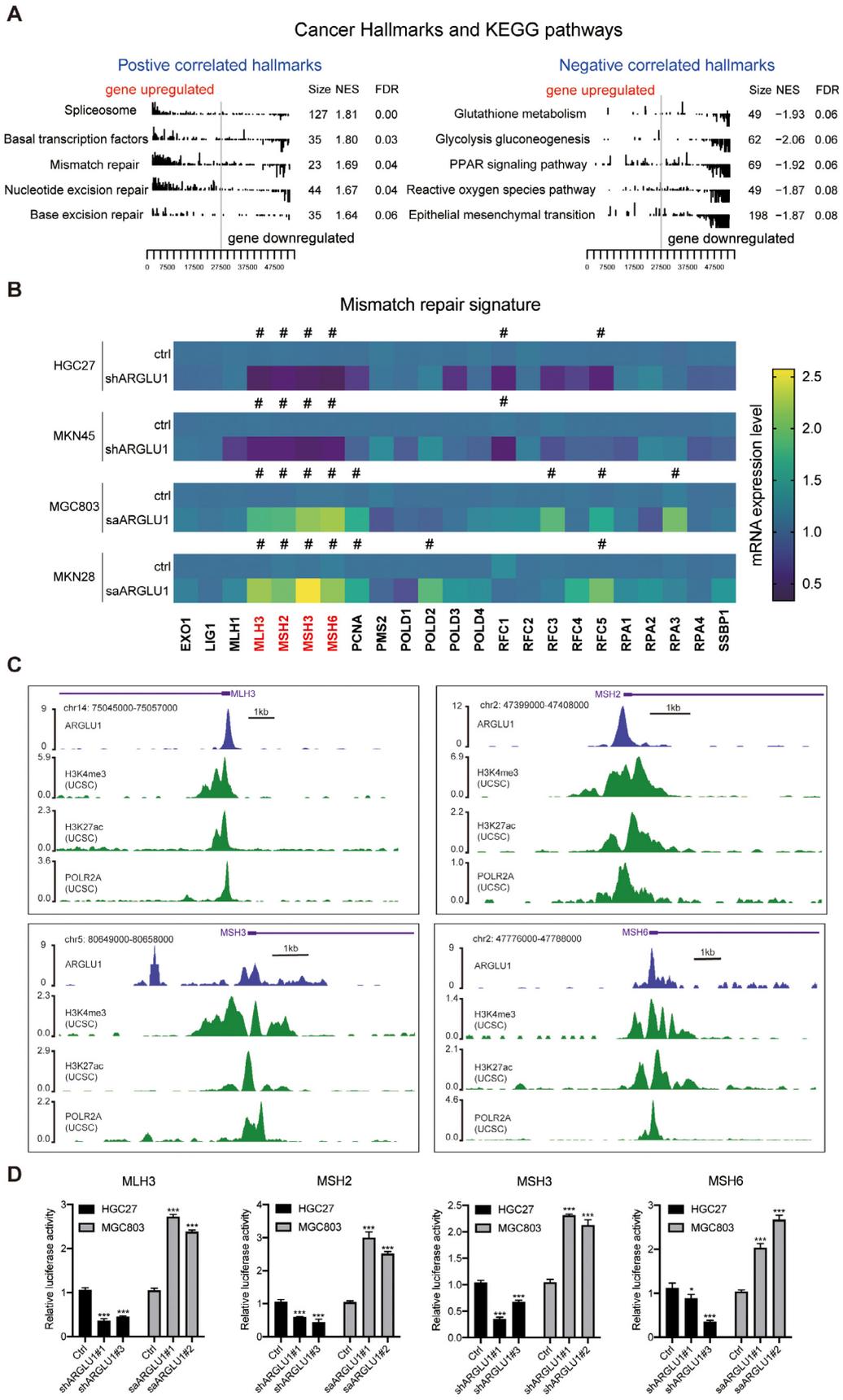
Previous research [17,18,20] demonstrated that ARGLU1 served not only as a crucial splicing regulator but also a transcriptional



**Fig. 4.** Prioritization and evaluation of the nine potential therapeutic targets in PDO models. (a) Flowchart of prioritization and evaluation targets in PDO models. (b) The nine genes' mRNA level of 3 GC patients' tumor and adjacent non-tumor tissues. (c) Prioritization of potential therapeutic targets. (d) qRT-PCR validation the 4 candidates' perturbation effect by saRNA/shRNA transfection on organoid 1. (e) Representative images of three GC organoids transfected with saRNA, shRNA or control (scale bars = 400  $\mu$ m, upper panel), and quantification of organoid diameters (bottom panel). The data are the means  $\pm$  SDs of three independent experiments (\*,  $p < 0.05$ , \*\*,  $p < 0.01$ , \*\*\*,  $p < 0.001$ , NS, not significant).

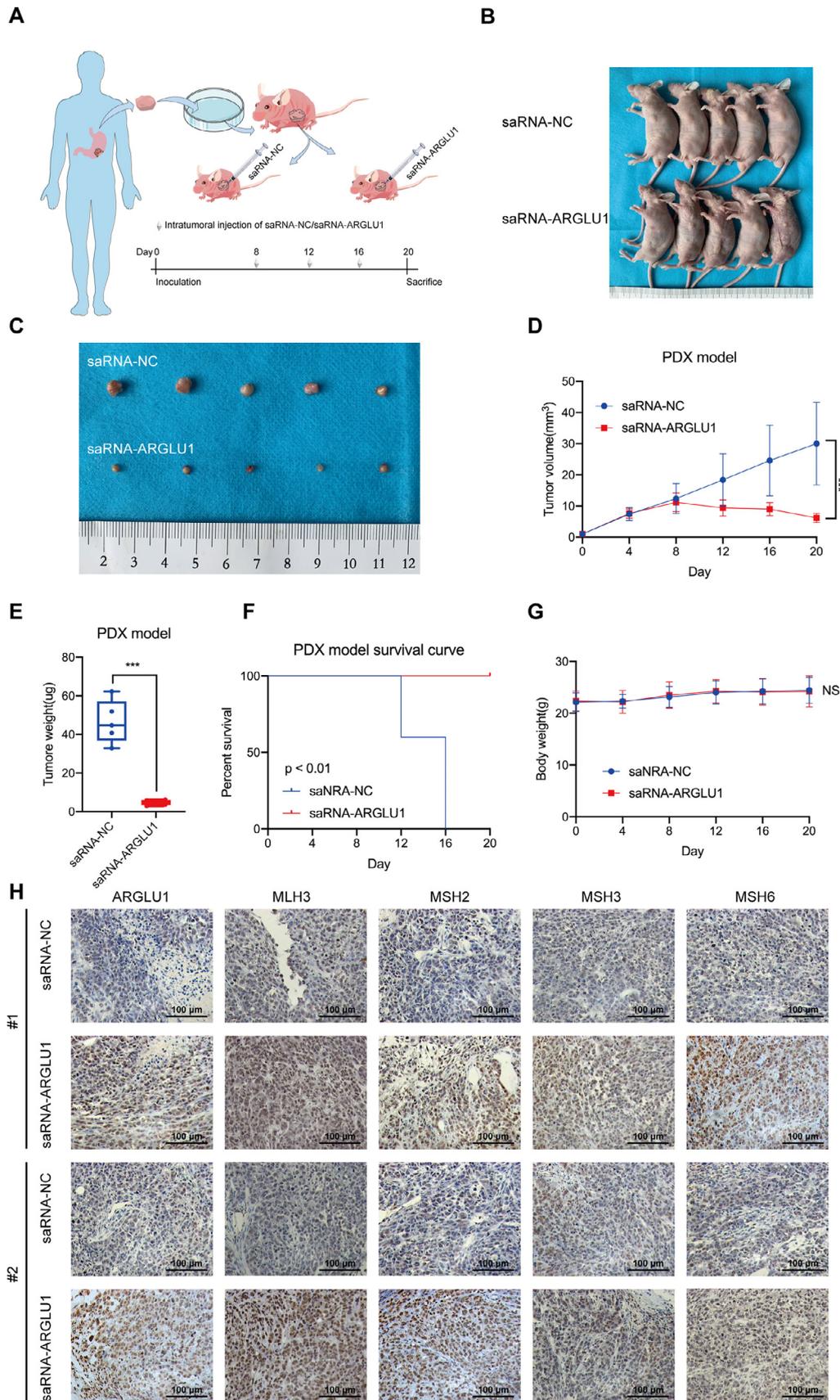


**Fig. 5.** ARGLU1 inhibited gastric cancer cells proliferation in vitro and tumor growth in vivo. (a) Representative IHC staining images for ARGLU1 expression in tissue microarray of 96 GC patients (Ruijin cohort). (b) Statistical graph of ARGLU1 protein level (IHC). (c) The correlation between ARGLU1 protein level and T stage and N stage in Ruijin cohort. (d) Survival curve of ARGLU1 in TCGA-STAD cohort and Ruijin cohort. (e) ARGLU1 protein level in 8 GC cell lines. (f) Western blot analysis of ARGLU1 over-expressed or knock-down stably transfected cell lines. (g, h) Clone formation and CCK8 assay of GC cell lines with ARGLU1 perturbation. 3 independent experiments were conducted, and data were shown with Mean  $\pm$  SD (two-tailed *t*-test, \*\*,  $p < 0.01$ , \*\*\*,  $p < 0.001$ ). (i) Representative images of resected subcutaneous tumors. (j) Tumor weight was recorded at time of harvest and plotted according to treatment group (Mean  $\pm$  SD, two-tailed *t*-test, \*,  $p < 0.05$ ). (k) Subcutaneous tumor dimensions were recorded using calipers at every 4 days. And tumor volume was calculated by formula: Length  $\times$  Width<sup>2</sup>/2 (mean  $\pm$  SD,  $n = 5$  for each group, one-way ANOVA, \*\*\*,  $p < 0.001$ ).



**Fig. 6.** ARGLU1 increased the transcriptional level of mismatch repair genes in GC cells. (a) GSEA analysis of ARGLU1 in TCGA-STAD cohort. (b) Quantitative RT-PCR analysis of MMR genes mRNA expression levels in four GC cells with ARGLU1 perturbation (#, |FoldChange|>1.5,  $p < 0.01$ ). (c) ChIP-seq data exhibited peaks of ARGLU1, H3K4me3, H3K27ac and POLR2A in the promoter region of MLH3, MSH2, MSH3 and MSH6. (d) Relative luciferase activity of MLH3, MSH2, MSH3 and MSH4 in different treatment groups of MGC803 and HGC27 cells. The data are the means  $\pm$  SDs of three independent experiments (two-tailed  $t$ -test, \*,  $p < 0.05$ , \*\*,  $p < 0.01$ , \*\*\*,  $p < 0.001$ , NS, not significant).





**Fig. 8.** ARGLU1 served as a promising therapeutic target for GC (PDX model). (a) Schematic diagram of GC PDX model treated with LNP-formulated saRNA-NC and saRNA-ARGLU1 via intratumoral injection. The experimental timeline corresponds to days following tumor inoculation. (b) Representative photographs of mice with subcutaneous tumors (Each tick mark on the ruler corresponds to 1 mm). (c) Representative photographs of tumors from PDX model (Each tick mark on the ruler corresponds to 1 mm). (d) Subcutaneous tumor growth curve (Tumor volume, mean  $\pm$  SD,  $n = 5$  for each group, one-way ANOVA, \*\*\*,  $p < 0.001$ ). (e) Tumor weight of two treatment groups (Mean  $\pm$  SD, ug, \*\*\*,  $p < 0.001$ ). (f) Kaplan-Meier survival curves of PDX model (death was defined as tumor length  $\geq 3$  mm). (g) Body weight of nude mice (NS, not significant). (h) Representative IHC staining images of ARGLU1 and MMR genes prepared from saRNA-NC and saRNA-ARGLU1 treatment xenografts under x200 magnitude.

coactivator. In order to illustrate the molecular mechanism of ARGLU1 increasing MMR genes' transcription, we conducted immunoprecipitation followed by mass spectrometry (IP-MS) in HGC27 cells using anti-ARGLU1 antibody and obtained 231 proteins potentially interacted with ARGLU1. By integrating IP-MS results with 4 MMR genes' potential transcription factors (TFs) predicted by JASPAR [21], we found that five proteins including SP1, YY1, MLLT1, POLR2A and RBPJ were potential ARGLU1 binding proteins (Fig. 7a). Subsequently, Co-IP performed in HGC27 cells confirmed that ARGLU1 could bind with SP1 and YY1 to form transcriptional complex (Fig. 7b). ChIP-qPCR in shARGLU1 or shNC treated HGC27 cells using antibodies against YY1 and SP1 found that ARGLU1 knock-down significantly attenuated ChIP signals on promoter regions of MLH3, MSH2, MSH3 and MSH6 (Fig. 7c, Fig. S2b). Furthermore, rescue assays were performed by transfecting shARGLU1 into siYY1, siSP1 or siNC treated HGC27 cells. We found that siSP1 or siYY1 treatment could slightly down-regulated the MMR genes' mRNA expression level, whereas ARGLU1 knock-down significantly reduced the transcriptional level' of MMR genes (Fig. 7d, Fig. S2c). Same conclusion could be made in rescue assays conducted in MGC803 cells suggesting that the MMR genes' transcription was more dependent on the expression level of ARGLU1 (Fig. 7e, Fig. S2d). These data suggested that ARGLU1 enhanced the transcriptional activity of SP1 and YY1 on MMR genes by potentiating the recruitment of SP1 and YY1 on the promoters of MMR genes (Fig. 7f).

### 3.7. Inducing ARGLU1 expression inhibited GC growth in PDX model

Given the favorable effect of inducing ARGLU1 expression on inhibition GC growth in vitro, in vivo and in PDO model, we further validated the clinical value of ARGLU1-induction by saRNA in GC using PDX model, which was established by fresh tumor tissues transplanting into nude mice subcutaneously. Eight days after tumor transplanted, mice were randomly divided into 2 groups (each group contains 5 mice) and treated with saRNA-NC or saRNA-ARGLU1 formulated by lipidoid-encapsulated nanoparticles (LNP) via intratumoral injection at 5 mg/kg every 4 days for three total treatments (Fig. 8a). All mice were subsequently euthanized, and xenograft tumors were harvested at 20 days after inoculation (Fig. 8b, c). Tumor volume was recorded every 4 days after inoculation, and analysis of tumor weight revealed that ARGLU1 induction could significantly inhibit tumor growth ( $4.74 \pm 1.40$  vs.  $46.52 \pm 11.17$ , Mean  $\pm$  SD, two-tailed *t*-test, Fig. 8d, e). Further analysis of survival curve revealed that the nude mice in ARGLU1-inducing treatment group showed a better outcome than that of control group (Death was defined as tumor length > 3 mm, Fig. 8f). Body weight of nude mice was measured every 4 days to evaluate possible toxic effect of ARGLU1 induction therapy and no significant difference was found between two groups (Fig. 8g). IHC staining revealed that the protein expression of ARGLU1 was increased in ARGLU1-saRNA-treated xenografts 2 weeks following initial doses. The expression of MLH3, MSH2, MSH3 and MSH6 in ARGLU1-induction group were significantly up-regulated in contrast to negative control group (Fig. 8h). These results suggested that ARGLU1 could serve as a potential therapeutic target for GC.

## 4. Discussion

The overall outcome of advanced GC is very poor and few molecular target has been proven effective for GC [22]. Biomarkers or signatures for predicting prognosis and therapeutic benefits are an indispensable part for implementing cancer precision medicine. The enhancements of prognostic and therapeutic benefits with the aid of biomarkers have been reported in colorectal cancer, breast cancer, lung cancer, etc. [23–26]. However, current efforts are always focusing on predictive accuracy over explanatory power [27–31]. On the

other hand, molecular genetic profiling of GC has yielded promising new candidate therapeutic targets such as RTKs or RAS and PI3-kinase signaling proteins [30,31]. Although various new agents are still being investigated for target therapy, several ongoing clinical trials are already targeting c-MET, STAT3, CLDN18.2, and mTOR [32–34]. Some preliminary results of these trials have been encouraging, but due to high heterogeneity of GC, only a little part of patients could get benefit from these target therapy. So far, trastuzumab is the only biomarker-driven therapy in clinical practice and especially in view of numerous negative clinical trials such as the TyTan and LOGiC trials of lapatinib (a tyrosine kinase inhibitor targeting HER2 and EGFR) [35,36] the RILOMET-1 trial of rilotumumab, and the METGastric trial of onartuzumab (targeting c-MET) [37,38]. Therefore, prognostic predicting biomarkers with mechanistic interpretability as well as effective molecular therapeutic targets for advanced GC remain the two urgent problems for oncologists. In this study, we aimed to construct a prognostic predicting model based on genome-wide functional screening data. And based on this model, we attempted to identify potential therapeutic targets for GC.

The far-reaching research presented by Fiona M. Behan performed CRISPR-Cas9 fitness screens in 339 cancer cell lines targeting over 18,000 genes to comprehensively catalog genes that are required for cancer cell fitness (defined as genes required for cell growth or viability). Genes required for cell fitness in specific molecular or histological contexts are likely to encode favorable drug targets, because of a reduced likelihood of inducing toxic effects in healthy tissues. In the present study, based on GC-specific fitness genes, we performed LASO—Cox regression in TCGA-STAD cohort and constructed a GC prognostic (GCP) model. The efficacy and stability of our GCP model was validated in seven independent cohorts.

Our GCP model consists of 9 prognostic genes which could be regarded as promising therapeutic targets for precision treatment in GC. There are plenty of experimental evidences have suggested that some of these 9 genes including DDX24 [39,40], TTF2 [14,15], SSX4 [16], ACLY [41–44], VPS35 [45–48], DCTN2 [49–52] and MRPL4 [53,54] played important roles in development and progression in many tumor types. However, there was few report about ARGLU1 functions in cancer. Our research found that ARGLU1 expression in tumor tissues of GC was down-regulated. Over expression of ARGLU1 inhibited GC cells proliferation in vitro and in vivo. Further experiments demonstrated that ARGLU1 could bind with SP1 and YY1 to enhance their transcriptional activity on MLH3, MSH2, MSH3 and MSH6. Higher expression of MMR proteins is more likely to have maintaining effect on genome stability and to play a causal role in delaying GC progression [55–57]. For translating these results to clinical practice, we induced ARGLU1 by LNP-formulated saRNA in PDX model. We found that intratumoral injection of LNP-formulated saRNA inducing ARGLU1 expression significantly inhibited tumor growth. These observations indicate that ARGLU1 could be a promising therapeutic target for GC.

The last but not least, our results showed that targeting ARGLU1, TTF2, SSX4 and DDX24 treatment could significantly inhibit GC growth in the PDO models with high GCP score. Although studies with more patients and more detailed evaluation are needed, the current data suggest that our GCP model, besides its prognostic efficacy, may help to stratify patients for treatment targeting to ARGLU1, DDX24, TTF2 or SSX4.

## Contributors

All authors read and approved the final version of the manuscript. Yuan-Yuan Li and Bingya Liu Conceptualization. Fanyuan Li Investigation and Writing - original draft; Jianfang Li Data curation; Junxian Yu and Tao Pan Formal analysis; Beiqin Yu and Junyi Hou Methodology; Chao Yan and Mingde Zang Resources; Qingqing Sang and

Wentao Dai Validation; Zhenggang Zhu Supervision; Liping Su Writing - review & editing.

### Data sharing

Transcriptome profiling raw data with corresponding clinical information were downloaded from TCGA (<https://cancergenome.nih.gov/>) and the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) databases. Gastric cancer specific fitness genes (with the priority score  $\geq 20$ ) were retrieved from the Project Score database (<https://score.depmap.sanger.ac.uk/>). ChIP-seq data in this study can be accessed in GEO (GSE174158).

### Declaration of Competing Interest

The authors have declared that no competing interest exists.

### Acknowledgements

We thank all study participants of the Shanghai Key Laboratory of Gastric Neoplasms. We thank Shanghai Center for Bioinformatics Technology for assistance in data storage and computation.

### Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:[10.1016/j.ebiom.2021.103436](https://doi.org/10.1016/j.ebiom.2021.103436).

### References

- [1] Siegel RL, Miller KD, Jemal A. Cancer statistics. *CA Cancer J Clin* 2019;69(1):7–34.
- [2] Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68(6):394–424.
- [3] Rawla P, Barsouk A. Epidemiology of gastric cancer: global trends, risk factors and prevention. *Prz Gastroenterol* 2019;14(1):26–38.
- [4] Thrift AP, El-Serag HB. Burden of Gastric Cancer. *Clin Gastroenterol Hepatol* 2020;18(3):534–42.
- [5] Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. *The Lancet* 2016;388(10060):2654–64.
- [6] Choi YY, Noh SH, Cheong J-H. Evolution of Gastric Cancer Treatment: from the Golden Age of Surgery to an Era of Precision Medicine. *Yonsei Med. J.* 2015;56(5):1177.
- [7] Vargas AJ, Harris CC. Biomarker development in the precision medicine era: lung cancer as a case study. *Nat Rev Cancer* 2016;16(8):525–37.
- [8] Li Q, Dai W, Liu J, Li YX, Li YY. Gene dysregulation analysis builds a mechanistic signature for prognosis and therapeutic benefit in colorectal cancer. *J Mol Cell Biol* 2020.
- [9] Cai WY, Dong ZN, Fu XT, Lin LY, Wang L, Ye GD, et al. Identification of a Tumor Microenvironment-relevant Gene set-based Prognostic Signature and Related Therapy Targets in Gastric Cancer. *Theranostics* 2020;10(19):8633–47.
- [10] Robinson WH, Lindstrom TM, Cheung RK, Sokolove J. Mechanistic biomarkers for clinical decision making in rheumatic diseases. *Nat Rev Rheumatol* 2013;9(5):267–76.
- [11] Topalian SL, Taube JM, Anders RA, Pardoll DM. Mechanism-driven biomarkers to guide immune checkpoint blockade in cancer therapy. *Nat Rev Cancer* 2016;16(5):275–87.
- [12] Lu Y, Fang Z, Li M, Chen Q, Zeng T, Lu L, et al. Dynamic edge-based biomarker non-invasively predicts hepatocellular carcinoma with hepatitis B virus infection for individual patients based on blood testing. *J Mol Cell Biol* 2019;11(8):665–77.
- [13] Behan FM, Iorio F, Picco G, Goncalves E, Beaver CM, Migliardi G, et al. Prioritization of cancer therapeutic targets using CRISPR-Cas9 screens. *Nature* 2019;568(7753):511–6.
- [14] Gao Y, Chen F, Niu S, Lin S, Li S. Replication and meta-analysis of common gene mutations in TTF1 and TTF2 with papillary thyroid cancer. *Medicine (Baltimore)* 2015;94(36):e1246.
- [15] Park E, Gong E-Y, Romanelli MG, Lee K. Suppression of estrogen receptor-alpha transactivation by thyroid transcription factor-2 in breast cancer cells. *Biochem Biophys Res Commun* 2012;421(3):532–7.
- [16] Schmitz-Winnenthal F.H., Galindo-Escobedo L.V., Rimoldi D., Geng W., Romero P., Koch M., et al. Potential target antigens for immunotherapy in human pancreatic cancer. 2007;252(2):290–8.
- [17] Koedoot E, Smid M, Foekens JA, Martens JW, Le Dévédec SE, van de Water B. Co-regulated gene expression of splicing factors as drivers of cancer progression. *Sci Rep* 2019;9(1):1–13.
- [18] Magomedova L, Tiefenbach J, Zilberman E, Florian Voisin V, Saikali M, et al. ARGLU1 is a transcriptional coactivator and splicing regulator important for stress hormone signaling and development. *Nucleic Acids Res.* 2019;47(6):2856–70.
- [19] Jiricny J. The multifaceted mismatch-repair system. *Nat Rev Mol Cell Biol* 2006;7(5):335–46.
- [20] Zhang D, Jiang P, Xu Q, Zhang X. Arginine and Glutamate-rich 1 (ARGLU1) Interacts with Mediator Subunit 1 (MED1) and Is Required for Estrogen Receptor-mediated Gene Transcription and Breast Cancer Cell Growth. *J Biol Chem* 2011;286(20):17746–54.
- [21] Khan A, Fornes O, Stigliani A, Gheorghie M, Castro-Mondragon JA, van der Lee R, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* 2018;46(D1):D260–D6.
- [22] Liu X, Meltzer SJ. Gastric cancer in the era of precision medicine. *Cell Mol Gastroenterol Hepatol* 2017;3(3):348–58.
- [23] Chen H, Sun X, Ge W, Qian Y, Bai R, Zheng S. A seven-gene signature predicts overall survival of patients with colorectal cancer. *Oncotarget* 2016;8(56):95054–65.
- [24] Khambata-Ford S, Garrett CR, Meropol NJ, Basik M, Harbison CT, Wu S, et al. Expression of Epireregulin and amphiregulin and K-ras mutation status predict disease control in metastatic colorectal cancer patients treated with cetuximab. *J Clin Oncol* 2007;25(22):3230–7.
- [25] Salazar R, Roepman P, Capella G, Moreno V, Simon I, Dreezen C, et al. Gene expression signature to improve prognosis prediction of stage II and III colorectal cancer. *J Clin Oncol* 2011;29(1):17–24.
- [26] Sparano JA, Gray RJ, Ravdin PM, Makower DF, Pritchard KI, Albain KS, et al. Clinical and genomic risk to guide the use of adjuvant therapy for breast cancer. *New Eng J Med* 2019;380(25):2395–405.
- [27] Cho JY, Lim JY, Cheong JH, Park YY, Yoon SL, Kim SM, et al. Gene expression signature-based prognostic risk score in gastric cancer. *Clin Cancer Res* 2011;17(7):1850–7.
- [28] Tan IB, Ivanova T, Lim KH, Ong CW, Deng N, Lee J, et al. Intrinsic subtypes of gastric cancer, based on gene expression pattern, predict survival and respond differently to chemotherapy. *Gastroenterology* 2011;141(2):476–85 e11.
- [29] Wang K, Kan J, Yuen ST, Shi ST, Chu KM, Law S, et al. Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer. *Nat Genet* 2011;43(12):1219–23.
- [30] Bass AJ, Thorsson V, Shmulevich I, Reynolds SM, Miller M, Bernard B, et al. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* 2014;513(7517):202–9.
- [31] Cristescu R, Lee J, Nebozhyn M, Kim K-M, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med* 2015;21(5):449–56.
- [32] Muro K, Chung HC, Shankaran V, Geva R, Catenacci D, Gupta S, et al. Pembrolizumab for patients with PD-L1-positive advanced gastric cancer (KEYNOTE-012): a multicentre, open-label, phase 1b trial. *Lancet Oncol* 2016;17(6):717–26.
- [33] Shitara K, Ohtsu A. Advances in systemic therapy for metastatic or advanced gastric cancer. *J Natl Compr Canc Netw* 2016;14(10):1313–20.
- [34] Becerra C, Stephenson J, Jonker DJ, Cohn AL, Asmis TR, Bekaii-Saab TS, et al. Phase Ib/II study of cancer stem cell (CSC) inhibitor BBI608 combined with paclitaxel in advanced gastric and gastroesophageal junction (GEJ) adenocarcinoma. *J Clin Oncol* 2015;33(15\_suppl):4069.
- [35] Hecht JR, Bang YJ, Qin SK, Chung HC, Xu JM, Park JO, et al. Lapatinib in combination with capecitabine plus oxaliplatin in human epidermal growth factor receptor 2-positive advanced or metastatic gastric, esophageal, or gastroesophageal adenocarcinoma: a RTO-013/LOGIC-A randomized phase III Trial. *J Clin Oncol* 2016;34(5):443–51.
- [36] Satoh T, Xu RH, Chung HC, Sun GP, Doi T, Xu JM, et al. Lapatinib plus paclitaxel versus paclitaxel alone in the second-line treatment of HER2-amplified advanced gastric cancer in Asian populations: tyTAN—a randomized, phase III study. *J Clin Oncol* 2014;32(19):2039–49.
- [37] Cunningham D, Tebbutt NC, Davidenko I, Murad AM, Al-Batran S-E, Ilson DH, et al. Phase III, randomized, double-blind, multicenter, placebo (P)-controlled trial of rilotumumab (R) plus epirubicin, cisplatin and capecitabine (ECX) as first-line therapy in patients (pts) with advanced MET-positive (pos) gastric or gastroesophageal junction (G/GEJ) cancer: R10MET-1 study. *J Clin Oncol* 2015;33(15\_suppl):4000.
- [38] Shah MA, Bang Y-J, Lordick F, Taberner J, Chen M, Hack SP, et al. METGastric: a phase III study of onartuzumab plus mFOLFOX6 in patients with metastatic HER2-negative (HER2-) and MET-positive (MET+) adenocarcinoma of the stomach or gastroesophageal junction (GEC). *J Clin Oncol* 2015;33(15\_suppl):4012.
- [39] Shi D, Dai C, Qin J, Gu W. Negative regulation of the p300-p53 interplay by DDX24. *Oncogene* 2016;35(4):528–36.
- [40] Oliver D, Ji H, Liu P, Gasparian A, Gardiner E, Lee S, et al. Identification of novel cancer therapeutic targets using a designed and pooled shRNA library screen. *Sci Rep* 2017;7(1):43023.
- [41] Chen Y, Li K, Gong D, Zhang J, Li Q, Zhao G, et al. ACLY: a biomarker of recurrence in breast cancer. *Pathol - Res Pract* 2020;216(9):153076.
- [42] Wen J, Min X, Shen M, Hua Q, Han Y, Zhao L, et al. ACLY facilitates colon cancer cell metastasis by CTNNB1. *J Exp Clin Cancer Res* 2019;38(1).
- [43] Granchi C. ATP citrate lyase (ACLY) inhibitors: an anti-cancer strategy at the crossroads of glucose and lipid metabolism. *Eur J Med Chem* 2018;157:1276–91.
- [44] Zaidi N, Swinnen JV, Smans K. ATP-citrate lyase: a key player in cancer metabolism. *Cancer Res.* 2012;72(15):3709–14.
- [45] An C.H., Kim Y.R., Kim H.S., Kim S.S., Yoo N.J., Lee S.H. Frameshift mutations of vacuolar protein sorting genes in gastric and colorectal cancers with microsatellite instability. 2012;43(1):40–7.

- [46] Liu Y, Deng H, Liang L, Zhang G, Xia J, Ding K, et al. Depletion of VPS35 attenuates metastasis of hepatocellular carcinoma by restraining the Wnt/PCP signaling pathway. *Genes Dis* 2020.
- [47] Kumar KR, Weissbach A, Heldmann M, Kasten M, Tunc S, Sue CM, et al. Frequency of the D620N mutation in VPS35 in Parkinson disease. *Arch Neurol* 2012;69(10):1360.
- [48] Zhang G, Tang X, Liang L, Zhang W, Li D, Li X, et al. DNA and RNA sequencing identified a novel oncogene VPS35 in liver hepatocellular carcinoma. *Oncogene* 2020.
- [49] Duijf PHG, Nanayakkara D, Nones K, Srihari S, Kalimutho M, Khanna KK. Mechanisms of genomic instability in breast cancer. *Trends Mol Med* 2019;25(7):595–611.
- [50] Yu Y, Gaillard S, Jude HT-C, Sneha N, et al. Inhibition of spleen tyrosine kinase potentiates paclitaxel-induced cytotoxicity in ovarian cancer cells by stabilizing microtubules. *Cancer Cell* 2015;28(1):82–96.
- [51] Rezaei-Tavirani M, Rezaei Tavirani M, Zamanian Azodi M, Moravvej Farshi H, Razzaghi M. Evaluation of skin response after erbium:yttrium–aluminum–garnet laser irradiation: a network analysis approach. *J Lasers Med Sci* 2019;10(3):194–9.
- [52] Wang S, Wang Q, Zhang X, Liao X, Wang G, Yu L, et al. Distinct prognostic value of dynactin subunit 4 (DCTN4) and diagnostic value of DCTN1, DCTN2, and DCTN4 in colon adenocarcinoma. *Cancer Manag Res* 2018;Volume 10:5807–24.
- [53] Bertucci F. Gene expression profiling for molecular characterization of inflammatory breast cancer and prediction of response to chemotherapy. *Cancer Res* 2004;64(23):8558–65.
- [54] Hermawan A, Putri H. Bioinformatics studies provide insight into possible target and mechanisms of action of nobiletin against cancer stem cells. *Asian Pacif J Cancer Prev* 2020;21(3):611–20.
- [55] Peltomaki P. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol* 2003;21(6):1174–9.
- [56] Leite M, Corso G, Sousa S, Milanezi F, Afonso LP, Henrique R, et al. MSI phenotype and MMR alterations in familial and sporadic gastric cancer. *Int J Cancer* 2011;128(7):1606–13.
- [57] Cho J, Kang SY, Kim K-M. MMR protein immunohistochemistry and microsatellite instability in gastric cancers. *Pathology* 2019;51(1):110–3.