

ORIGINAL RESEARCH

Meta-analysis of genome-wide association studies and functional assays decipher susceptibility genes for gastric cancer in Chinese populations

Caiwang Yan,^{1,2,3} Meng Zhu,^{1,2} Yanbing Ding,⁴ Ming Yang,⁵ Mengyun Wang,^{6,7} Gang Li,⁸ Chuanli Ren,⁹ Tongtong Huang,¹ Wenjun Yang,¹⁰ Bangshun He,¹¹ Meilin Wang,^{2,12} Fei Yu,¹ Jinchun Wang,¹ Ruoxin Zhang,^{6,7} Tianpei Wang,¹ Jing Ni,¹ Jiaping Chen,^{1,2} Yue Jiang,^{1,2} Juncheng Dai,^{1,2} Erbao Zhang,^{1,2} Hongxia Ma,^{1,2} Yanong Wang,¹³ Dazhi Xu,¹³ Shukui Wang,^{2,11} Yun Chen,^{2,14} Zekuan Xu ,^{2,15} Jianwei Zhou,^{2,16} Guozhong Ji,^{2,17} Zhaoming Wang,¹⁸ Zhengdong Zhang,^{2,12} Zhibin Hu ,^{1,2,3} Qingyi Wei,^{6,19,20} Hongbing Shen,^{1,2} Guangfu Jin 

► Additional material is published online only. To view please visit the journal online (<http://dx.doi.org/10.1136/gutjnl-2019-318760>).

For numbered affiliations see end of article.

Correspondence to

Dr Guangfu Jin, Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing 211166, China; guangfujin@njmu.edu.cn

CY, MZ, YD, MY, MW, GL and CR are joint first authors.

Received 22 March 2019

Revised 24 June 2019

Accepted 12 July 2019

Published Online First

5 August 2019



© Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.

To cite: Yan C, Zhu M, Ding Y, et al. *Gut* 2020;**69**:641–651.

ABSTRACT

Objective Although a subset of genetic loci have been associated with gastric cancer (GC) risk, the underlying mechanisms are largely unknown. We aimed to identify new susceptibility genes and elucidate their mechanisms in GC development.

Design We conducted a meta-analysis of four genome-wide association studies (GWASs) encompassing 3771 cases and 5426 controls. After targeted sequencing and functional annotation, we performed in vitro and in vivo experiments to confirm the functions of genetic variants and candidate genes. Moreover, we selected 33 promising variants for two-stage replication in 7035 cases and 8323 controls from other five studies.

Results The meta-analysis of GWASs identified three loci at 1q22, 5p13.1 and 10q23.33 associated with GC risk at $p < 5 \times 10^{-8}$ and replicated seven known loci at $p < 0.05$. At 5p13.1, the risk rs59133000[C] allele enhanced the binding affinity of NF- κ B1 (nuclear factor kappa B subunit 1) to the promoter of *PRKAA1*, resulting in a reduced promoter activity and lower expression. The knockout of *PRKAA1* promoted both GC cell proliferation and xenograft tumour growth in nude mice. At 10q23.33, the rs3781266[C] and rs3740365[T] risk alleles in complete linkage disequilibrium disrupted and created, respectively, the binding motifs of POU2F1 and PAX3, resulting in an increased enhancer activity and expression of *NOC3L*, while the *NOC3L* knockdown suppressed GC cell growth. Moreover, two new loci at 3q11.2 (OR=1.21, $p=4.56 \times 10^{-9}$) and 4q28.1 (OR=1.14, $p=3.33 \times 10^{-11}$) were associated with GC risk.

Conclusion We identified 12 loci to be associated with GC risk in Chinese populations and deciphered the mechanisms of *PRKAA1* at 5p13.1 and *NOC3L* at 10q23.33 in gastric tumourigenesis.

INTRODUCTION

Gastric cancer (GC) is the fifth most common cancer and the third leading cause of cancer death worldwide, and more than 40% of new cases and deaths each year occur in China.¹ The heritability

Significance of this study

What is already known on this subject?

- A subset of gastric cancer (GC) susceptibility loci have been identified by genome-wide association studies.
- Up to now, few loci have been experimentally confirmed for their biological mechanisms underlying the observed genetic associations.

What are the new findings?

- For the first time, we characterise the mechanisms underlying the association of loci of 5p13.1 and 10q23.33 with GC risk at the regulatory and functional levels.
- At the 5p13.1 locus, the C allele of rs59133000 reduces the promoter activity by enhancing the binding affinity of transcriptional repressor NF- κ B1 (nuclear factor kappa B subunit 1), leading to decreased expression of *PRKAA1*, weakening the suppressive effect of *PRKAA1* on tumourigenesis and eventually resulting in an increased risk of GC.
- At the 10q23.33 locus, the risk haplotype rs3781266[C]–rs3740365[T] can switch the binding motifs of POU2F1 and PAX3 in an enhancer element, resulting in increased enhancer activity and expression of *NOC3L* and ultimately promoting GC tumourigenesis.
- Here, we report 12 GC risk-related loci, including two novel ones of rs7624041 at 3q11.2 and rs10029005 at 4q28.1.
- Our pathway analysis highlights an important role of host immunity, especially the cellular response to interleukin 1, in GC susceptibility.

of GC has been estimated to be 22% in Nordic countries.² Familial aggregation of GC occurs in approximately 10% of the cases, due to heritable pathogenic mutations in the high-penetrance genes, including *CDH1*, *CTNNA1*, *BRCA2*, *STK11*,

Significance of this study

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

- ▶ Our results expand current understanding of genetic basis for GC risk and provide new deeper insights into the biology and mechanisms of susceptibility to GC.
- ▶ The identified risk variants may serve as the basis for risk assessment in the clinical setting.

PALB2, *BRCA1* and *RAD51C*.^{3–5} However, the vast majorities of GC cases are sporadic, without definitive germline mutations. A subset of common variants or susceptibility loci have been identified by genome-wide association studies (GWASs),^{6–13} including 1q22, 3q13.31, 5p13.1, 5q14.3, 6p21.1, 6p22.1, 8q24.3, 9q34.2, 10q23.33, 12q24.11–12 and 20q11.21. In addition, we and other groups have discovered low-frequency variants in *SPOCD1*¹³ and *ATM*¹⁴ with a higher penetrance for GC. Following the discovery of susceptibility loci in association studies, *MUC1* at 1q22⁹ and *PSCA* at 8q24.3^{15 16} have been experimentally confirmed as susceptibility genes for GC. However, the remaining loci await functional studies to further understand biological mechanisms underlying their observed genetic associations.

To comprehensively characterise susceptibility genes for GC, we performed a meta-analysis of four GWASs, with replication in other five studies, with a total of 10 806 cases and 13 749 controls from Chinese populations; the present study is thus the largest genetic study of GC in China to identify new genetic variants associated with GC risk. Through targeted sequencing, functional annotations and both in vitro and in vivo experiments, we deciphered the mechanisms by which the genetic variants at 5p13.1 and 10q23.33 influence gastric tumorigenesis.

MATERIALS AND METHODS**Genome-wide association studies**

Four GWASs were included in the initial meta-analysis: one new GWAS (Onco-GWAS) and three published GWASs (NJ-GWAS, BJ-GWAS and SX-GWAS).^{7 8 11} The characteristics of the participants are summarised in online supplementary table 1. The cases and cancer-free controls of each study were all ethnic Chinese from the same geographical area. All four studies used Illumina or Affymetrix chips for genotyping. We performed quality control procedures on genotyping data using the same protocol for all four GWAS datasets (online supplementary figure 1). The study details and quality control procedures are presented in online supplementary text. After quality control on GWAS datasets, cases and controls were roughly matched at the genetic level for all four studies (online supplementary figure 2). All participants provided a written informed consent. All studies were approved by the relevant institutional review boards.

Meta-analysis of GWASs

To facilitate the meta-analysis of four GWAS datasets, imputation was conducted separately for each GWAS dataset (details provided in online supplementary text and imputation quality shown in online supplementary figure 3). Single-variant association analysis was performed separately for each GWAS dataset by the use of SNPTEST software (V.2.5).¹⁷ Significant principal components of ancestry and demographic characteristics were included as covariates in the multivariate logistic regression model in estimating ORs and 95% CIs for the genetic variants.

An inverse variance-weighted fixed effects model was used in the meta-analysis as implemented in the GWAMA software (V.2.1).¹⁸ Quantile-quantile (Q-Q) plots and Manhattan plots were created for individual studies or the meta-analysis using R software (V.3.4.3) (online supplementary figures 4 and 5). Regional plots were generated by using LocusZoom,¹⁹ and the Bonferroni correction was used to control multiple testing in each region (0.05/number of variants tested). The details for targeted sequencing and functional annotations are provided in the online supplementary text.

Functional experiments

The activities of regulatory elements controlling potentially functional variants at 5p13.1 and 10q23.33 were experimentally validated by using luciferase reporter assays. Transcription factor binding site analysis was performed with the JASPAR 2018²⁰ and enhancer element locator algorithm.²¹ The binding of transcription factors was confirmed by electrophoretic mobility shift assay (EMSA)- and chromatin immunoprecipitation (ChIP)-based assays. RNA interference, plasmid overexpression, lentivirus-mediated knockdown and CRISPR/Cas9-mediated knockout, as well as cell viability, proliferation and 5-ethynyl-2'-deoxyuridine (EdU) incorporation assays, were used to evaluate the biological functions of *PRKAA1* and *NOC3L* in GC cells. The methods used for CRISPR/Cas9-mediated knockout of *PRKAA1* and lentivirus-mediated knockdown of *NOC3L* in the BGC823 cell line are shown in online supplementary figures 6 and 7. Cells (3.0×10^6) were injected subcutaneously into the bilateral armpit of 5- to 6-week-old male BALB/c nude mice to assess the growth of xenograft tumours after *PRKAA1* knockout or *NOC3L* knockdown. Protocols for these experimental procedures are available in the online supplementary text.

Replication studies

To identify new loci associated with GC risk, we performed a two-stage replication with detailed information summarised in online supplementary table 2. In the first stage (replication I), we genotyped 33 promising loci in 1710 GC cases and 1802 controls by using the iPLEX Sequenom MassARRAY platform (Sequenom, San Diego, California, USA) (online supplementary table 3). In the second stage (replication II), we further genotyped five variants that showed consistent associations in replication I ($p < 0.05$) in 5325 cases and 6521 controls from Jiangsu (replication II-JS), Shanghai (replication II-SH), Shandong (replication II-SD) and Ningxia (replication II-NX) by using a TaqMan allelic discrimination assay with an ABI 7900 system (Applied Biosystems, Carlsbad, California, USA). Logistic regression analysis was applied to estimate the ORs and 95% CIs for each replication study independently, and all results from replication studies and GWAS datasets were combined with a fixed-effects meta-analysis.

Gene-based and pathway analyses

We performed gene-based and pathway analyses on the basis of single-variant p values from the meta-analysis as implemented in MAGMA (Multi-marker Analysis of GenoMic Annotation).²² A total of 4622 curated gene sets representing known biological and metabolic pathways were derived from the KEGG (Kyoto Encyclopedia of Genes and Genomes)²³ and GO (Gene Ontology)²⁴ databases, which were catalogued by and obtained from MSigDB V.6.20.²⁵ The details are provided in the online supplementary text.

Table 1 Gastric cancer susceptibility loci at 1q22, 5p13.1 and 10q23.33 with a p value less than 5×10^{-8} in the meta-analysis of GWASs (3771 cases and 5426 controls) and two new loci identified in replication studies (7035 cases and 8323 controls)

Locus	Associated gene	Variant*	Position (build 37)	Alleles (risk/other)	RAF† (case, control)	Study	OR (95% CI)‡	P value‡
Known loci								
1q22	<i>MUC1</i>	Lead: rs760077	1:155 178 782	T/A	0.89, 0.85	GWAS	1.38 (1.26 to 1.51)	2.09×10^{-12}
		Functional: rs4072037	1:155 162 067	T/C	0.87, 0.84	GWAS	1.33 (1.22 to 1.45)	1.15×10^{-10}
5p13.1	<i>PRKAA1</i>	Lead: rs6897169	5:40 726 138	C/T	0.57, 0.51	GWAS	1.25 (1.17 to 1.33)	5.48×10^{-12}
		Functional: rs59133000	5:40 798 974	C/T	0.55, 0.50	GWAS	1.24 (1.16 to 1.32)	1.17×10^{-11}
10q23.33	<i>NOC3L</i>	Lead: rs10509671	10:96 069 054	C/A	0.19, 0.15	GWAS	1.34 (1.23 to 1.45)	2.51×10^{-12}
		Functional: rs3781266	10:96 052 747	C/T	0.19, 0.15	GWAS	1.31 (1.21 to 1.42)	1.03×10^{-10}
		Functional: rs3740365	10:96 053 239	T/A	0.19, 0.15	GWAS	1.31 (1.21 to 1.42)	1.00×10^{-10}
New loci								
3q11.2	<i>NSUN3</i>	Lead: rs7624041	3:94 108 663	G/A	0.11, 0.09	GWAS	1.21 (1.09 to 1.34)	4.57×10^{-4}
					0.10, 0.08	Replication	1.21 (1.11 to 1.32)	5.00×10^{-6}
						Combined	1.21 (1.13 to 1.28)	4.56×10^{-9}
4q28.1	<i>LOC285419, ANKRD50</i>	Lead: rs10029005	4:125 451 364	A/G	0.36, 0.33	GWAS	1.16 (1.09 to 1.24)	4.52×10^{-6}
					0.36, 0.33	Replication	1.13 (1.08 to 1.19)	1.03×10^{-6}
						Combined	1.14 (1.10 to 1.19)	3.33×10^{-11}

*Lead variant is the most significant variant at the locus in the results of meta-analysis; functional variant is previously reported (for 1q22) or newly identified in the present study (for 5p13.1 and 10q23.33).

†RAF was calculated by using combined data of samples in GWASs or replication studies.

‡OR, 95% CI and p value were derived from the fixed-effects inverse variance-weighted meta-analysis. GWAS, genome-wide association study; RAF, risk allele frequency.

RESULTS

Meta-analysis of GC GWAS datasets and targeted sequencing of chromosomes 1q22, 5p13.1 and 10q23.33

An overview of the analysis strategy is presented in online supplementary figure 1. After quality control, 3771 GC cases and 5426 controls were retained in four GWAS datasets of Chinese ancestry (online supplementary figure 1), and more than 6 million variants were included in the genetic association analysis. Q-Q plots did not show evidence of a substantial inflation rate ($\lambda=1.00-1.07$, online supplementary figure 4). Manhattan plots from individual GWAS and meta-analysis are shown in online supplementary figure 5 and Figure 1A, respectively. Genetic variants of three regions were found to be associated with GC risk at the genome-wide association level ($p < 5 \times 10^{-8}$), including 1q22 (lead variant rs760077: $p=2.09 \times 10^{-12}$), 5p13.1 (rs6897169: $p=5.48 \times 10^{-12}$) and 10q23.33 (rs10509671: $p=2.51 \times 10^{-12}$) (table 1). Moreover, an additional seven known loci for GC susceptibility, including 3q13.31, 5q14.3, 6p21.1, 6p22.1, 8q24.3, 9q34.2 and 20q11.21 were replicated in the meta-analysis with p values less than 0.05 (online supplementary table 4).

The regional association results showed that the most strongly risk-associated variant rs760077 at 1q22 was previously reported, and rs6897169 at 5p13.1 and rs10509671 at 10q23.33 were highly correlated with previously reported variants ($r^2=0.96$ and 1.00 for rs10074991 and rs3781264, respectively) (online supplementary figure 8). Conditional logistical regression analysis and targeted sequencing of 1q22, 5p13.1 and 10q23.33 showed that there were no secondary signals with independent effects from the top hits (online supplementary figures 8 and 9).

A promoter variant of *PRKAA1* at 5p13.1 modulates its tumour-suppressing role and modifies GC risk

To delineate the functional variants mapped to 5p13.1, we evaluated potential functions of 59 variants in strong linkage

disequilibrium (LD) ($r^2 > 0.6$) with the lead variant rs6897169 of this region (online supplementary table 5), all of which are in non-coding regions. Expression quantitative trait loci (eQTL) analysis showed the strongest correlation between rs6897169 genotypes and *PRKAA1* expression levels in normal stomach tissues ($p=7.2 \times 10^{-4}$) (online supplementary table 6; figure 1B). We further focused on variants that are located within the annotated stomach-related *cis*-regulatory elements. The three most strongly supported variants (ie, rs3805495, rs5913300 and rs10065570) are located at the centre of DNase I hypersensitivity site (DHS) peaks as well as within regions harbouring promoter or enhancer histone marks (eg, H3K4me1, H3K4me3 and H3K27ac) in stomach tissues and transcription factor binding sites (figure 1C). Given the considerable evidence for allelic differences in *PRKAA1* expression levels, we measured the effect of rs5913300 alleles on promoter activity and rs3805495 as well as rs10065570 alleles on the enhancer activity by luciferase reporter assays. The results revealed that the C allele of rs5913300, which was associated with an increased GC risk (OR=1.24, 95% CI: 1.16 to 1.32, $p=1.17 \times 10^{-11}$, table 1) and correlated with decreased expression of *PRKAA1*, consistently reduced the promoter activity of the reporter gene, as compared with the T allele (figure 2A). However, no allelic differences in the enhancer activity were observed for rs3805495 and rs10065570 (online supplementary figure 10). ChIP assays further confirmed H3K4me3 at the rs5913300 site (figure 2B). These results suggested that the C allele of rs5913300 might reduce the promoter activity and result in decreased expression of *PRKAA1*.

Subsequently, we sought to clarify whether the base change of rs5913300 might influence the preference for transcription factors. The EMSA assay showed that the C allele had a stronger binding band than the T allele (figure 2C). The competitive EMSA assay indicated that a 100-fold excess of unlabeled probe containing the C allele could completely abolish binding with the T allele, whereas the unlabeled T allele could not

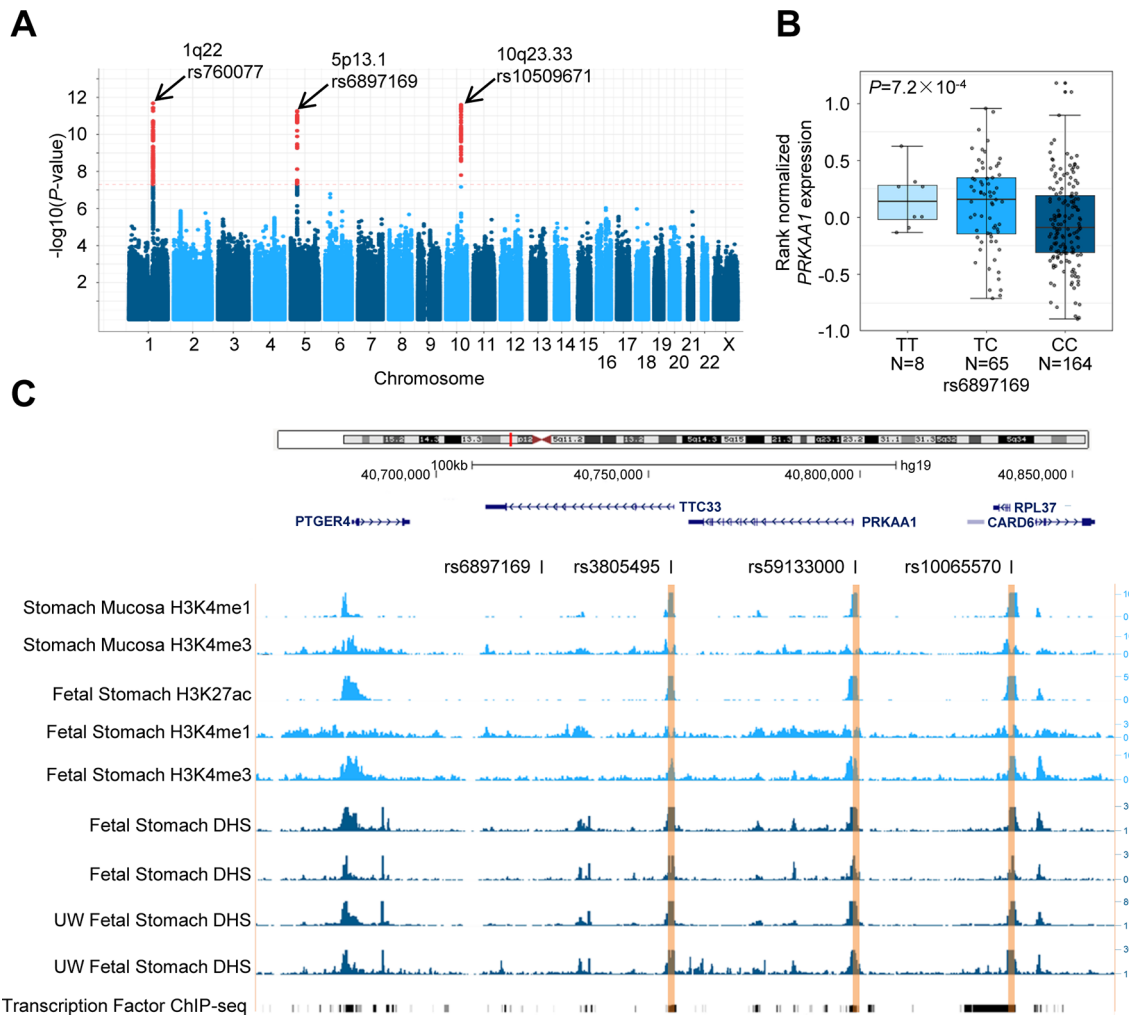


Figure 1 Manhattan plot for meta-analysis of gastric cancer genome-wide association studies and functional annotations on the locus of 5p13.1. (A) The associations ($-\log_{10}(p)$ values, y-axis) are plotted against genomic position (y-axis by chromosome and the chromosomal position); the variants at 1q22, 5p13.1 and 10q23.33 with $p < 5 \times 10^{-8}$ are plotted in red. (B) The risk allele C in the lead variant rs6897169 at 5p13.1 was correlated with decreased expressions of *PRKAA1* in normal stomach tissues. (C) ChIP-Seq data show enrichments of histone marks, DHSs and transcription factors in the sites of three variants (rs3805495, rs5913300 and rs10065570) in strong linkage disequilibrium with the lead variant rs6897169 ($r^2=1.00, 1.00$ and 0.76 , respectively) (see also online supplementary table 5). ChIP, chromatin immunoprecipitation; DHS, DNase I hypersensitivity site; UW, University of Washington.

abolish binding with the C allele (figure 2C). Further in silico analysis revealed that rs59133000 was mapped within the binding motif of the transcription factor nuclear factor kappa B subunit 1 (NF- κ B1) (figure 2D, left; online supplementary table 7). Using NCI-N87 cells heterozygous for rs59133000, we found the preferential binding of NF- κ B1 to the risk allele C through ChIP followed by quantitative PCR (ChIP-qPCR) and Sanger sequencing of the PCR amplifications (figure 2D, online supplementary figure 11). NF- κ B1 antibody supershift assays further revealed that NFKB1 is an allele-preferential binder of rs59133000 (figure 2E). Inhibiting the expression of NF- κ B1 with specific small interfering RNA (siRNA) duplexes resulted in an increased expression of *PRKAA1* (online supplementary figure 12A). After stimulation by tumour necrosis factor α (TNF- α), activation of NF- κ B pathway first resulted in upregulation of *PRKAA1* expression, while treatment for 3 hours, the expression of *PRKAA1* gradually decreased with the increased concentration of TNF- α (online supplementary figure 12B). Taken together, these results suggest that the C allele of rs59133000 may enhance the binding affinity

of NF- κ B1 to the promoter of *PRKAA1* as a transcriptional repressor.

By using the expression data from TCGA, we found that *PRKAA1* was overexpressed in GC tumours (online supplementary figure 13A). We further measured *PRKAA1* expression in gastric normal tissues or precancerous lesions and found that *PRKAA1* expression was significantly reduced in intestinal metaplasia and dysplasia tissues as compared with superficial gastritis or atrophic gastritis (online supplementary figure 13B). To further explore biological significance of *PRKAA1* in gastric carcinogenesis, we knocked down *PRKAA1* by using siRNAs in BGC823 and SGC7901 cells that expressed relatively high levels of *PRKAA1* (online supplementary figure 14), and we found that *PRKAA1* knockdown promoted cellular malignant phenotypes in both the BGC823 and SGC7901 cell lines (online supplementary figure 15). Subsequently, by establishing a *PRKAA1* knockout BGC823 cell line by using the CRISPR/Cas9 system (supplementary figure 6), we found that the deletion of *PRKAA1* significantly increased proliferation (figure 3A,B) and clonogenicity (figure 3C) in these cell lines. Then, by rescuing the expression of

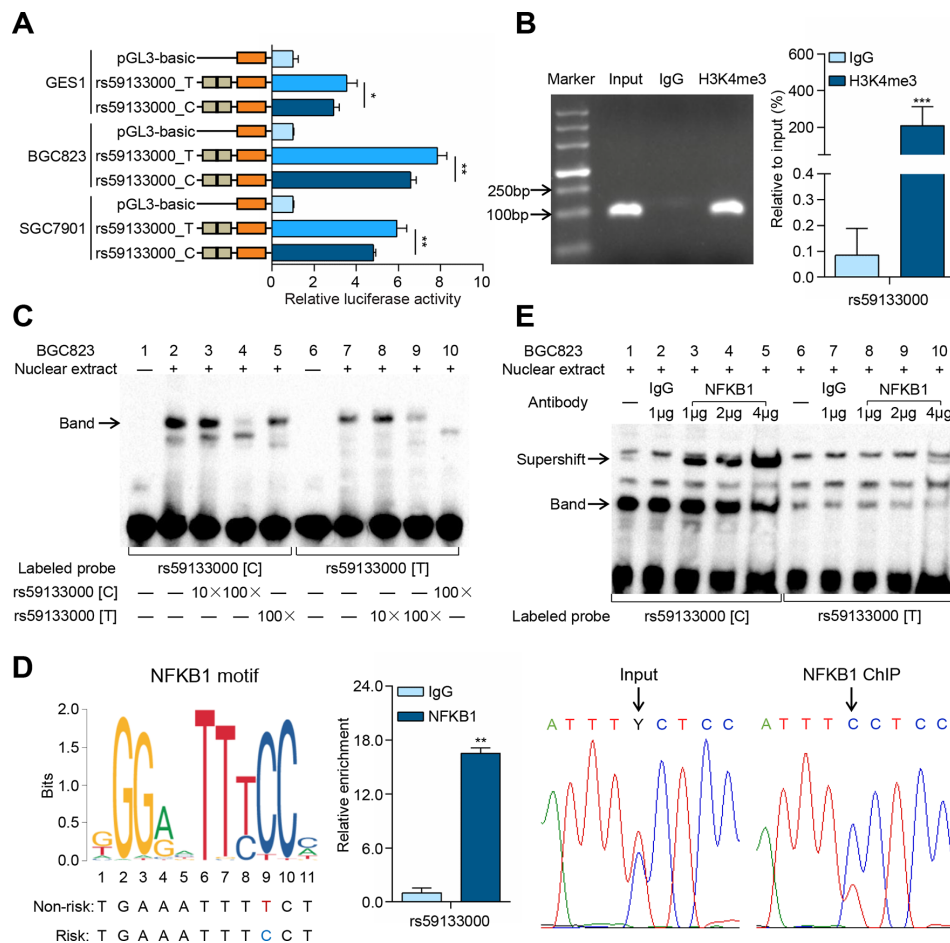


Figure 2 The variant rs59133000 at 5p13.1 regulates transcriptional activity of *PRKAA1* by modulating the binding affinity of nuclear factor kappa B subunit 1 (NF- κ B1) to the promoter. (A) Allele-specific constructs containing the putative promoter sequence flanking rs59133000 were cloned into the pGL3-basic luciferase reporter vector and transfected into the GES1, BGC823 and SGC7901 cell lines. The results of luciferase activity were normalised to those of pGL3 basic (n=6). (B) DNA electropherogram of the PCR fragment generated using primers flanking the variant rs59133000 following ChIP (left); based on ChIP followed by quantitative PCR (ChIP-qPCR) against H3K4me3, enrichment was quantified relative to the amount of input DNA (right, n=3), and antibody directed against IgG was used as a negative control. (C) EMSA with biotin-labelled oligonucleotides containing the rs59133000[C] or rs59133000[T] allele and nuclear extracts from BGC823 cells. Lanes 1 and 6 show the mobility of the labelled oligonucleotides without nuclear extracts; lanes 2 and 7 show the mobility of the labelled oligonucleotides with nuclear extracts in the absence of the competitor oligonucleotide and lanes 3–5 and 8–10 show the mobility of the labelled oligonucleotides with nuclear extracts in the presence of unlabelled rs59133000[C] or rs59133000[T] competitors. The arrow indicates a DNA–protein complex. (D) In silico predicted preferential binding of NF- κ B1 to the risk allele C of rs59133000 (left); after stimulated by tumour necrosis factor α (100 ng/mL) for 3 hours, the ChIP enrichment of NF- κ B1 as determined by ChIP-qPCR (middle, n=2) and Sanger sequencing of the PCR amplifications (right). (E) Supershift assays using the NF- κ B1 antibody in BGC823 cells. IgG was used as a negative control. In (A), (B) and (D), the data are presented as the means \pm SDs, and the p values were calculated using Student's t-test. * p <0.05, ** p <0.01 and *** p <0.001. ChIP, chromatin immunoprecipitation.

PRKAA1 in the established *PRKAA1* knockout BGC823 cell line, we found that the cell proliferation and clonogenicity decreased (figure 3D–F). Furthermore, by performing xenograft tumour assays with *PRKAA1* knockout BGC823 cell lines, we found that *PRKAA1* knockout reduced the growth of xenograft tumours in nude mice (figure 4). Taken together, these results suggest that *PRKAA1* functions as a tumour suppressor in GC development.

A risk haplotype at 10q23.33 activates an enhancer of *NOC3L* that promotes the GC tumourigenesis

To determine potentially functional variants at 10q23.33, we performed functional annotations for 78 variants that were in strong LD (r^2 >0.6) with the lead variant rs10509671 (online supplementary table 8). Three missense variants located in the exons of *PLCE1* and *NOC3L* were predicted to be tolerant or benign (online supplementary table 9). Subsequent eQTL

analysis indicated a strong correlation between the genotypes of the lead variant rs10509671 and *NOC3L* expression levels ($p=8.4\times 10^{-8}$) (Figure 5A; online supplementary table 10). Then, we focused on two non-coding regions: region 1 for three variants (ie, rs11187842, rs3781266 and rs3740365 at a distance of 728 bp in perfect LD, pairwise $r^2=1.00$) and region 2 for two variants (ie, rs7903902 and rs12220125 at a distance of 782 bp, $r^2=0.91$), which were marked by the H3K4me1 and H3K27ac histone modifications and by the DHS in stomach tissues (figure 5B). Further luciferase assays indicated that compared with the wild-type haplotype in region 1, the variant haplotype of rs11187842[A]–rs3781266[C]–rs3740365[T], which was highly correlated with the risk allele rs10509671[C] and significantly associated with an increased GC risk (table 1), produced a higher enhancer activity of the reporter gene (figure 5C). However, there were no differences in enhancer activities between the wild-type

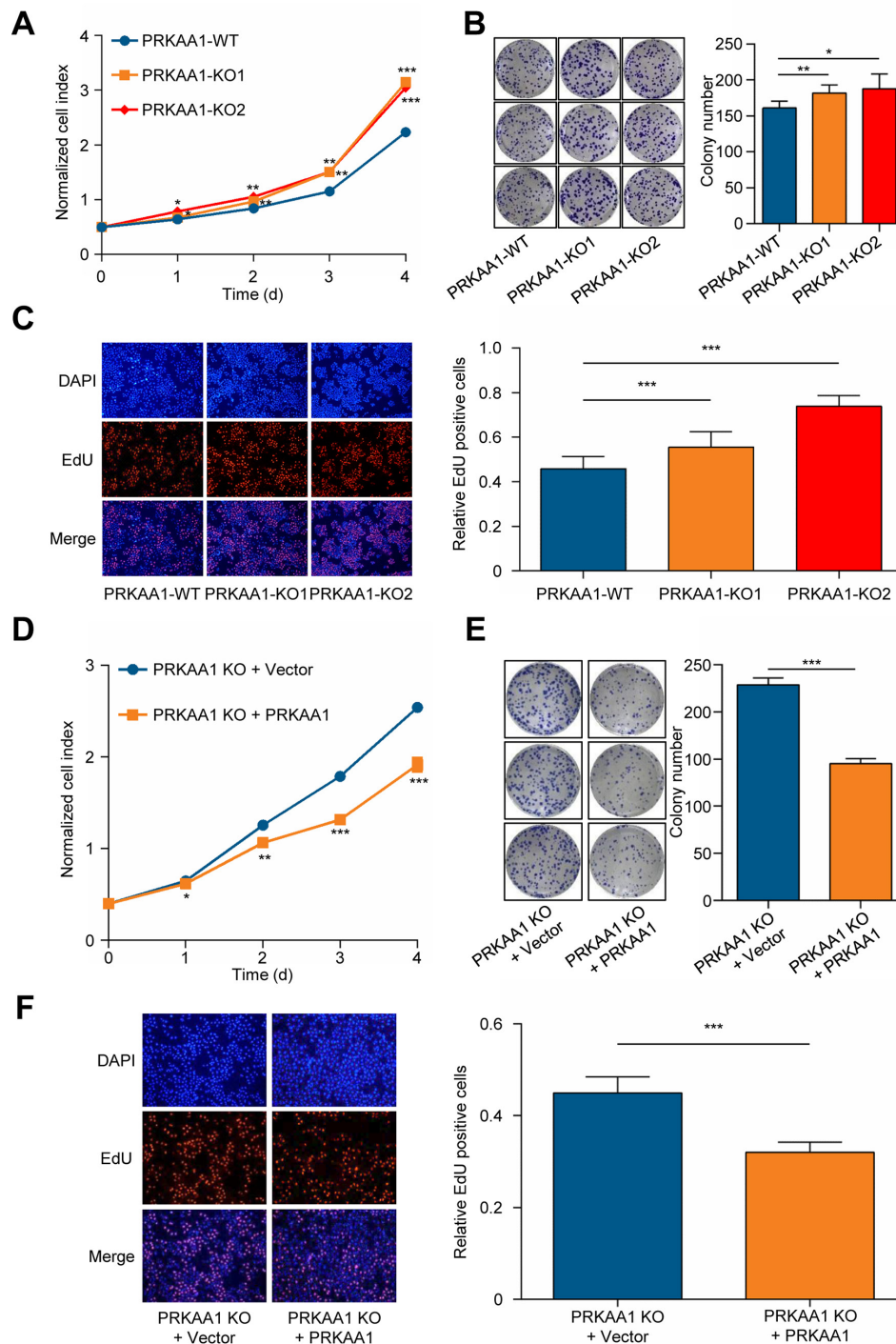


Figure 3 *PRKAA1* knockout promotes gastric malignant cellular phenotypes. (A) Proliferation curves before (WT) or after CRISPR/Cas9-mediated knockout (KO1 and KO2) of *PRKAA1* in BGC823 cells. (B) Colony formation assay with BGC823 cells with *PRKAA1* knockout (KO1 and KO2) and negative control cells (WT). The numbers of colonies were counted and are presented in the histogram. (C) EdU proliferation analysis before (WT) or after knockout of *PRKAA1* (KO1 and KO2) in BGC823 cells. (D–F) Recovery of *PRKAA1* expression in *PRKAA1* knockout BGC823 cells reduced cell proliferation. The data are presented as the means \pm SDs, and data analysis was conducted using Student's *t* test. **p* < 0.05, ***p* < 0.01, ****p* < 0.001. DAPI, 4',6-diamidino-2-phenylindole; WT, wild type.

and variant haplotypes in region 2 (figure 5D). In region 1, the ChIP assays also confirmed the H3K4me1 and H3K27ac modifications to the enhancer segment (online supplementary figure 16). Taken together, these results indicate that the risk haplotype of rs11187842[A]–rs3781266[C]–rs3740365[T] may elevate the enhancer activity at 10q23.33.

Next, we performed EMSA assays that showed stronger binding bands for the labelled probes containing either rs3781266[T] or rs3740365[T] allele, but probes for different alleles of rs11187842 did not exhibit differential bands (figure 5E). Competitive EMSA assays confirmed stronger binding affinities for the rs3781266[T] and rs3740365[T] alleles

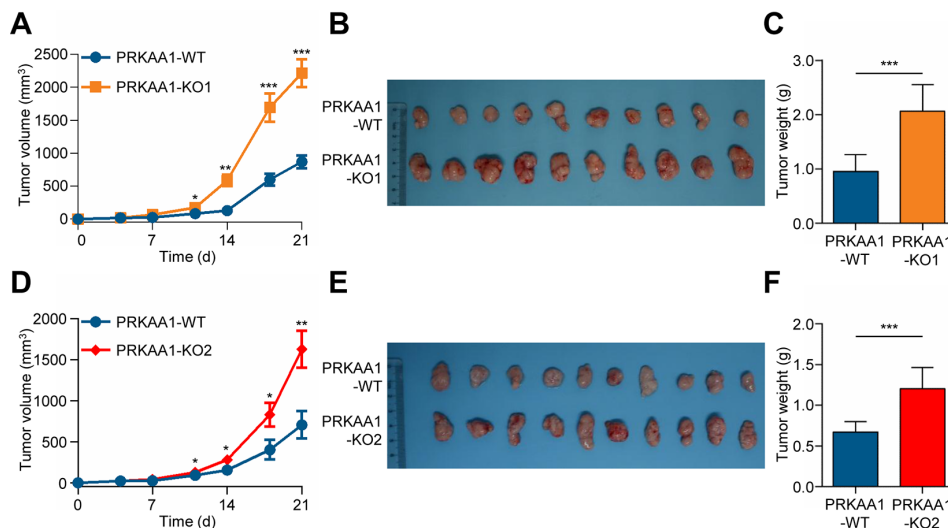


Figure 4 *PRKAA1* knockout promotes the tumorigenicity of gastric cancer cells in nude mice. In vivo xenograft tumour formation assays were performed using *PRKAA1* knockout (KO1 and KO2), and negative control (WT) cells (3.0×10^6) subcutaneously injected into the bilateral armpit (left: WT; right: *PRKAA1* KO1 or KO2) of 5- to 6-week-old male BALB/c nude mice. Tumour growth was measured twice weekly. On day 21, mice were sacrificed and tumours were photographed. (A, D) Mean volume of xenograft tumours in the WT or *PRKAA1* knockout (KO1 and KO2) groups. The data are presented as the means \pm SEMs. (B, E) Photographs of excised tumour tissues from mice (top, tumours from the WT group; bottom, tumours from the *PRKAA1* knockout group). (C, F) Average weight of harvested tumours in the WT and *PRKAA1* knockout groups. The data are presented as the means \pm SDs. The p values were calculated using Student's t-test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. WT, wild type.

(figure 5E, online supplementary figure 17). These two variants were in silico predicted to reside within binding motifs of multiple transcription factors in region 1 (online supplementary table 11). Further EMSAs using competitor DNA targeted to predicted transcription factors suggested that POU2F1 bound to the rs3781266 site, while PAX3 bound to the rs3740365 site (online supplementary figure 18A). The binding band was effectively competed away with an unlabeled intact sequence but not by oligonucleotides with a mutated POU2F1 or PAX3-binding motif (online supplementary figure 18B). Additional ChIP assays verified that POU2F1 and PAX3 in vivo bound to the DNA sequences containing rs3781266 and rs3740365, respectively (figure 5F,G; online supplementary figure 19). Taken together, these results suggest that the risk haplotype of rs3781266[C]-rs3740365[T] may disrupt the motif for POU2F1 binding but create the motif for PAX3 binding to the enhancer element at 10q23.33.

The expression data from TCGA showed that *NOC3L* was significantly overexpressed in GC tumours (online supplementary figure 20A). Subsequently, we knocked down *NOC3L* in BGC823 and SGC7901 cells, which expressed relatively high levels of *NOC3L* (online supplementary figure 20B) by RNA interference and observed a reduction in malignant cellular phenotypes, including behaviours such as cell proliferation and colony formation (online supplementary figure 21). We failed to establish a *NOC3L* knockout cell line by using the CRISPR/Cas9 system possibly due to the previously reported lethality²⁶; thus, we performed the short hairpin RNA (shRNA)-mediated knockdown of *NOC3L* in BGC823 cells (online supplementary figure 7). *NOC3L* knockdown cells exhibited significantly lower cell proliferation than the wild-type counterparts (figure 6A,C), and the phenotype was restored after rescuing the expression of *NOC3L* (online supplementary figure 22). Furthermore, by performing xenograft tumour assays, we found that *NOC3L* knockdown suppressed the growth of xenograft tumours in nude mice (figure 6D,E). Collectively, these results suggest that *NOC3L* may promote GC development.

Population replications identify two new loci at 3q11.2 and 4q28.1 associated with GC risk

To discover additional loci related to GC risk, we performed a two-stage replication study to assess 33 promising variants not previously reported. In the stage of replication I, we identified five variants consistently associated with GC risk at $p < 0.05$ in 1710 GC cases and 1802 controls (online supplementary table 3). In the stage of replication II, we further genotyped these five variants in an additional 5325 GC cases and 6521 controls and confirmed the associations between two variants (rs7624041 and rs10029005) and GC risk (table 1). Finally, we found that rs7624041 at 3q11.2 (OR=1.21, 95% CI: 1.13 to 1.28, $p = 4.56 \times 10^{-9}$) and rs10029005 at 4q28.1 (OR=1.14, 95% CI: 1.10 to 1.19, $p = 3.33 \times 10^{-11}$) were significantly associated with GC risk, without heterogeneity among studies (table 1 and online supplementary figure 23).

Gene-based and pathway analysis identify pathways related to immune response contributing to GC susceptibility

We identified 1379 genes to be associated with GC risk at a p value of < 0.05 (online supplementary figures 24 and 34) genes passing correction for multiple testing ($p_{FDR} < 0.05$) (online supplementary table 12). Among these genes, 28 were located in the regions 1q22, 5p13.1 and 10q23.33, and the other six new genes were *DAGLB*, *GFR1*, *OR10G2*, *IRGC*, *YDJC* and *CCDC116*.

In the pathway analysis, 222 pathways satisfied $p < 0.05$, of which 5 reached the significance level of $p < 1 \times 10^{-3}$, that is, cellular response to interleukin 1 ($p = 4.34 \times 10^{-6}$), response to platelet-derived growth factor ($p = 3.88 \times 10^{-4}$), immune response ($p = 4.57 \times 10^{-4}$), positive regulation of T-cell-mediated immunity ($p = 7.64 \times 10^{-4}$) and Toll-like receptor signaling pathway ($p = 8.28 \times 10^{-4}$) (online supplementary table 13). After permutation-based multiple testing correction, the cellular response to interleukin 1 pathway remained significant ($p = 0.022$). These results highlight the importance of pathways related to the immune response in GC susceptibility.

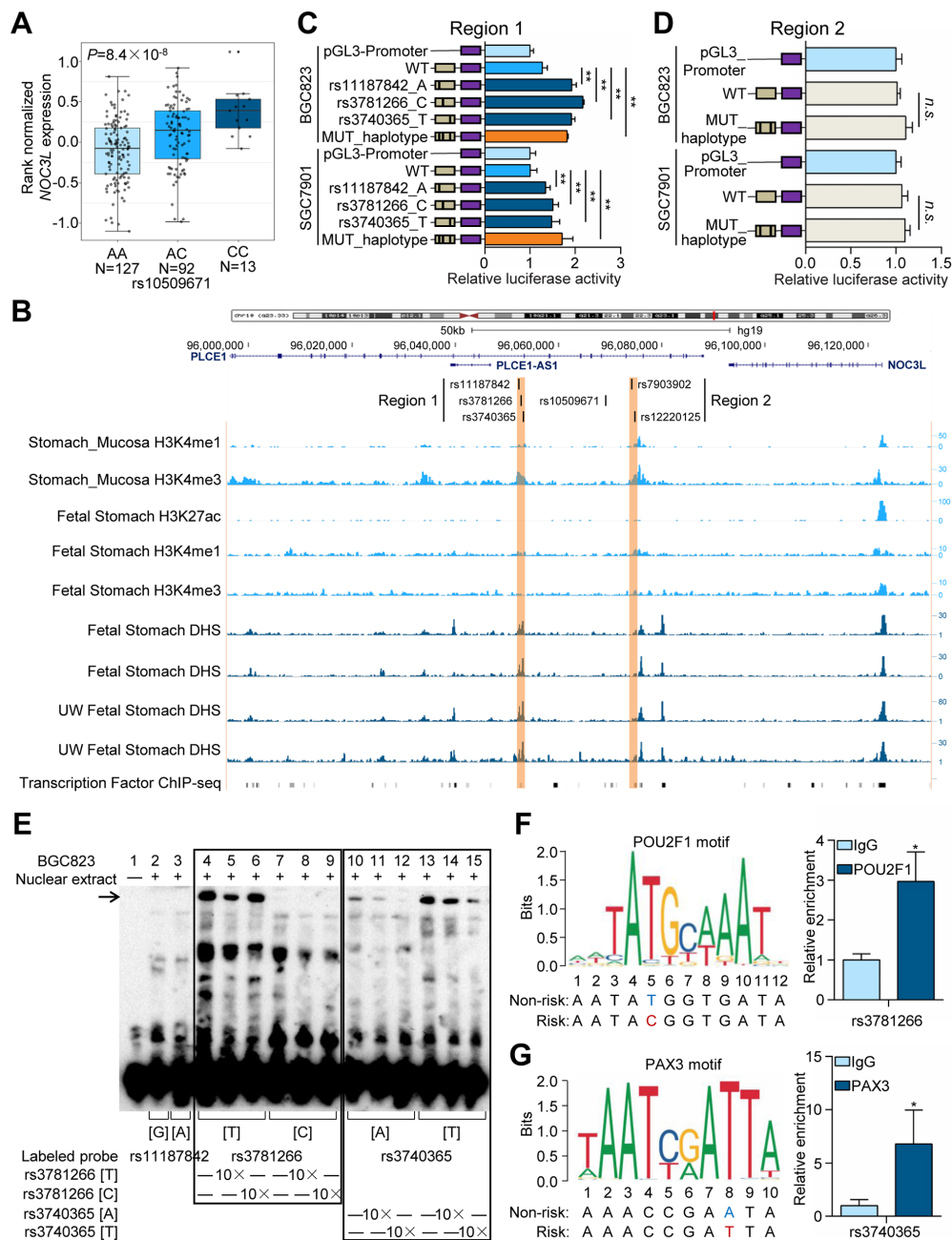


Figure 5 The haplotype at 10q23.33 regulates enhancer activity and expression of *NOC3L* by modulating the binding affinity of POU2F1 and PAX3. (A) The risk allele C of the lead variant rs10509671 was correlated with increased expressions of *NOC3L* in normal stomach tissues. (B) ChIP-Seq data show enrichment of histone marks, DHSs and transcription factors in two regions that include variants (region 1: rs11187842, rs3781266 and rs3740365; region 2: rs7903902 and rs12220125) in strong LD with the lead variant rs10509671 ($R^2 > 0.6$) (see also online supplementary table 8). (C–D) Allele-specific constructs containing the two putative enhancer regions (region 1 and region 2) were cloned into the pGL3-promoter luciferase reporter vector and transfected into BGC823 and SGC7901 cells. (C) Luciferase activity of the 900 bp DNA segment in the region 1 haplotype. Three variants within region 1, which were perfectly correlated with each other (pairwise $r^2=1.00$), were mutated to modify their common alleles (G, T and A) to the risk alleles (A, C and T, respectively). (D) Luciferase activity of the 800 bp DNA segment in the region 2 haplotype. The haplotype contained two variants in strong LD: rs7903902 and rs12220125 ($r^2=0.91$). The luciferase activity values were normalised to that of pGL3-promoter. The data are presented as the means \pm SDs, $n=6$. (E) EMSA with biotin-labelled oligonucleotides containing the rs11187842, rs3781266 and rs3740365 alleles in region 1 and nuclear extracts from BGC823 cells. Lane 1 shows the mobility of the labelled oligonucleotides without nuclear extracts; lanes 2–4, 7, 10 and 13 show the mobility of the labelled oligonucleotides with nuclear extracts in the absence of competitor oligonucleotides; and lanes 5–6, 8–9, 11–12 and 14–15 show the mobility of the labelled oligonucleotides with nuclear extracts in the presence of the associated unlabelled competitors. The arrow indicates a DNA–protein complex for rs3781266 and rs3740365. See also online supplementary figure 17. (F) The risk allele C of rs3781266 disrupts the POU2F1-binding motif. Left: Predicted preferential binding of POU2F1 to the non-risk allele T of rs3781266. Right: ChIP enrichment of POU2F1 as determined by ChIP-qPCR ($n=3$). (G) The risk allele T of rs3740365 creates a Pax3-binding motif. Left: Predicted preferential binding of Pax3 to the risk allele T of rs3740365. Right: ChIP enrichment of Pax3 as determined by ChIP-qPCR ($n=3$). In (C), (D), (F) and (G), the data are presented as the means \pm SDs, and the p values were calculated using Student's t -test. * $p < 0.05$, ** $p < 0.01$. ChIP, chromatin immunoprecipitation; DHS, DNase I hypersensitivity site; n.s., not significant; UW, University of Washington; WT, wild type.

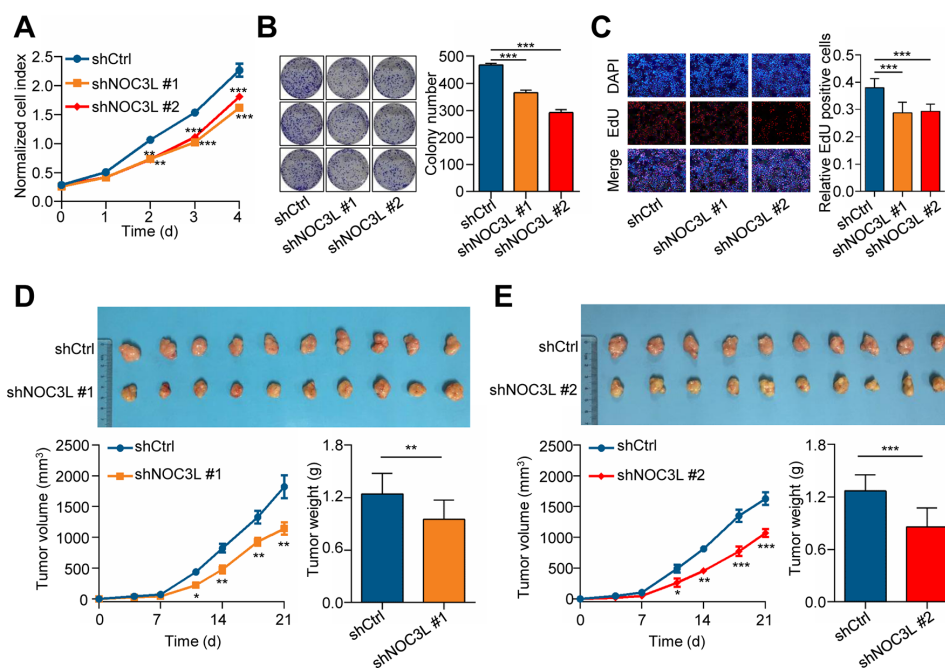


Figure 6 *NOC3L* knockdown suppresses gastric cancer cell growth. (A) Proliferation curve of *NOC3L* knockdown (shNOC3L #1 and #2) or negative control (shCtrl) BGC823 cells. (B) Colony formation assay of *NOC3L* knockdown or negative control BGC823 cells. The numbers of colonies were counted and were presented in the histogram. (C) EdU proliferation analysis indicating the effect of *NOC3L* on the growth of *NOC3L* knockdown or negative control BGC823 cells. (D, E) In vivo xenograft tumour formation assays were performed using *NOC3L* knockdown (shNOC3L #1 and #2) and negative control (shCtrl) cells (3.0×10^6) subcutaneously injected into the bilateral armpit (left: shCtrl; right: shNOC3L #1 or #2) of 5- to 6-week-old male BALB/c nude mice. Tumour growth was measured twice weekly. On day 21, mice were sacrificed, and excised tumours were photographed in the shCtrl or *NOC3L* knockdown groups. Mean volume of xenograft tumours is presented as the means \pm SEMs. Average weight of harvested tumours is presented as the means \pm SDs. The p values were calculated using Student's t-test. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

DISCUSSION

In 2011, our group first reported chromosome 5p13.1 as a susceptibility locus for GC in Chinese populations.⁸ Since then, several studies have replicated this association in populations of different ethnicities, including Chinese,¹⁰ Korean²⁷ and European¹⁴ populations. However, the functional mechanism of this locus remains unknown to date. In the present study, we found that the risk C allele of rs59133000 may reduce the promoter activity by enhancing the binding affinity of NF- κ B1 and also lower the expression of *PRKAA1*. NF- κ B1 (ie, p50), together with the other NF- κ B family members (RelA/p65, RelB, c-Rel and NF- κ B2/p52), forms homodimer and heterodimer that regulate many genes involved in inflammatory responses.²⁸ Among them, the commonly studied heterodimer p50:p65 can induce gene expression and provide a rapid molecular switch for responses to pathogens or inflammatory stimuli.²⁹ However, p50 homodimers can repress inflammation by competing with activating NF- κ B dimers and preventing them from binding to the sites on the promoters of target genes.³⁰ Moreover, the NF- κ B signalling is one of the important triggers in inflammation-induced carcinogenesis, including that in GC.³¹ Therefore, we presented an allele-specific regulation of *PRKAA1* mediated by NF- κ B1 binding to the promoter, which may have accounted for the observed differences in GC risk among individuals.

PRKAA1 encodes AMP-activated protein kinase (AMPK) subunit 1, AMPK α 1, which is required for the formation of the AMPK heterotrimeric serine/threonine kinase complex with the β - and γ -subunits.³² AMPK is a central regulator of the cellular metabolism and energy homeostasis, but the role of AMPK α 1 in tumour development has not been extensively investigated. In the present study, the knockout and recovery of *PRKAA1* in GC

cells and xenograft tumour assays in nude mice demonstrated the role of *PRKAA1* in suppressing GC cell growth. Recently, it was also shown that phosphorylation and activation of AMPK α 1 suppresses gastric tumourigenesis through the LKB1-AMPK-PGC1 α axis.³³ Notably, the use of metformin, the activator of AMPK, could reduce the risk of GC,³⁴ but these results remain controversial in observational studies.³⁵ In the present study, we found that the C allele of rs59133000, by lowering *PRKAA1* expression (naturally mimicking the opposite condition to metformin use), was associated with an increased GC risk, providing a genetic evidence supporting the protective effect of metformin use in the absence of confounding bias. Therefore, our results not only elucidated the mechanism by which genetic variants at 5p13.1 modify GC risk but also pinpointed the chemopreventive role of metformin in GC risk at the genetic level, suggesting that metformin has clinical implications and translational value to benefit populations at high risk of GC.

The 10q23.33 locus was initially found to be associated with the risk of both GC and oesophageal squamous cell carcinoma (ESCC) by two independent GWASs,^{7, 36} which was replicated in the subsequent studies.^{8, 37, 38} *PLCE1* was previously considered an ESCC susceptibility gene accounting for the signal of 10q23.33, although its reported functional roles were conflicting in recent studies.^{39–41} However, the biological mechanisms of 10q23.33 in GC development have not been investigated. Of particular interest, the lead variant rs10509671 at 10q23.33 was not correlated with *PLCE1* expression in normal stomach tissues (online supplementary table 10) but instead was strongly correlated to the expression of *NOC3L*. *NOC3L* encodes an NOC3-like DNA replication regulator, also known as a factor for adipocyte differentiation 24 (*fad24*). In previous studies,

fad24 was reported to be a positive regulator of adipogenesis by controlling DNA replication and promoting adipocyte differentiation.^{42,43} However, biological function of *NOC3L* in cancer remains unknown. The present study, for the first time, demonstrated that *NOC3L* could promote GC cell growth. It has been reported that genetic variants at 10q23.33 may have a stronger effect on gastric cardia cancer than on non-cardia cancer.⁴⁴ Obesity and high BMI have been considered risk factors for gastric cardia cancer via adipokine pathophysiology and systemic inflammation mechanism.⁴⁵ Notably, adiponectin, which is the most abundant adipokine negatively correlated with BMI, can activate the AMPK signalling but inhibit the NF- κ B signalling, resulting in a negative effect on carcinogenesis.⁴⁶ Therefore, we speculate that *NOC3L* may promote gastric tumourigenesis through promoting adipogenesis and increasing systemic inflammation. Moreover, the 10q23.33 locus was found to switch the binding of POU2F1 and PAX3 to a regulatory element by base changes at two variant sites (rs3781266 and rs3740365, respectively), which, however, are in complete LD. On the other hand, the mutant haplotype of rs3781266[C]–rs3740365[T] can upregulate *NOC3L* expression to promote GC cell growth, leading to an increased GC risk. Taken together, our findings suggest a model in which two variants at 10q23.33 may synergistically regulate the oncogenic role of *NOC3L* in GC development via binding affinity switching of the transcription factors POU2F1 and PAX3.

Our gene-based analysis showed that the majority of the significant genes (28/34) map to 1q22, 5p13.1 and 10q23.33. This finding again strengthens the dominant effects of these three loci on GC susceptibility. Of particular interest, the pathway analysis revealed that GC susceptibility genes were enriched in a series of pathways related to the immune response, especially the cellular response to the interleukin 1 pathway. This enrichment is consistent with the established link between chronic inflammation and GC development, because *Helicobacter pylori* infection causes chronic gastritis by promoting proinflammatory cytokine release and achlorhydria and follows a stepwise cascade of events from metaplasia to dysplasia before malignancy.⁴⁷ Therefore, the associations between genetic variants in inflammation-related genes and GC risk were extensively examined via a candidate gene approach previously.^{38,48} The most well-known example is polymorphisms in the interleukin-1 gene cluster,⁴⁹ including *IL1B-31* or *IL1B-511* and *IL1RN*, which exhibit a crucial role in the response to *H. pylori* infection and gastric neoplastic progression. Our pathway analysis results again highlighted the critical role of the host immunity in determining GC risk.

In summary, the present study reported 12 GC risk-related loci, including two novel ones. Overall, these loci can explain about 1.25% of the phenotypic variance in GC risk, whereas three loci of 1q22, 5p13.1 and 10q23.33 accounted for 0.76%. Importantly, for the first time, we characterised the mechanisms underlying the role the risk loci of 5p13.1 and 10q23.33 play in GC at the regulatory and functional levels. At the 5p13.1 locus, the C allele of rs59133000 reduces the promoter activity by enhancing the binding affinity of the transcriptional repressor NF- κ B1, leading to downregulated expression of *PRKAA1*, weakening the suppressive effect of *PRKAA1* on tumourigenesis and eventually resulting in an increased risk of GC. At the 10q23.33 locus, the risk haplotype rs3781266[C]–rs3740365[T] can switch the binding motifs of POU2F1 and PAX3 in an enhancer element, resulting in an increased enhancer activity and upregulated expression of *NOC3L* and ultimately promoting the tumourigenesis of GC. Moreover, the results of our pathway analysis also highlight an important role of host immunity, especially the

cellular response to interleukin 1, in GC susceptibility. Therefore, the present study expands our current understanding of the genetic basis of GC risk and provides new evidence for genes and biological pathways involved in the tumourigenesis of GC.

Author affiliations

¹Department of Epidemiology, Center for Global Health, School of Public Health, Nanjing Medical University, Nanjing, China

²Jiangsu Key Lab of Cancer Biomarkers, Prevention and Treatment, Collaborative Innovation Center for Cancer Medicine, Nanjing Medical University, Nanjing, China

³State Key Laboratory of Reproductive Medicine, China International Cooperation Center for Environment and Human Health, Nanjing Medical University, Nanjing, China

⁴Department of Gastroenterology, The Affiliated Hospital of Yangzhou University, Yangzhou, China

⁵Shandong Provincial Key Laboratory of Radiation Oncology, Cancer Research Center, Shandong Cancer Hospital Affiliated to Shandong University, Shandong Academy of Medical Sciences, Jinan, China

⁶Cancer Institute, Fudan University Shanghai Cancer Center, Shanghai, China

⁷Department of Oncology, Shanghai Medical College, Fudan University, Shanghai, China

⁸Department of General Surgery, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, The Affiliated Cancer Hospital of Nanjing Medical University, Nanjing, China

⁹Department of Laboratory Medicine, Clinical Medical College of Yangzhou University, Yangzhou, China

¹⁰Key Laboratory of Fertility Preservation and Maintenance, The General Hospital, Ningxia Medical University, Yinchuan, China

¹¹General Clinical Research Center, Nanjing First Hospital, Nanjing Medical University, Nanjing, China

¹²Key Lab of Modern Toxicology of Ministry of Education, School of Public Health, Nanjing Medical University, Nanjing, China

¹³Department of Gastric Cancer and Soft Tissue Sarcomas, Fudan University Shanghai Cancer Center, Shanghai, China

¹⁴Department of Immunology, Nanjing Medical University, Nanjing, China

¹⁵Department of General Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

¹⁶Department of Molecular Cell Biology and Toxicology, School of Public Health, Nanjing Medical University, Nanjing, China

¹⁷Institute of Digestive Endoscopy and Medical Center for Digestive Diseases, Second Affiliated Hospital of Nanjing Medical University, Nanjing, China

¹⁸Department of Epidemiology and Cancer Control, St. Jude Children's Research Hospital, Memphis, Tennessee, USA

¹⁹Duke Cancer Institute, Duke University Medical Center, Durham, North Carolina, USA

²⁰Department of Population Health Sciences, Duke University School of Medicine, Durham, North Carolina, USA

Correction notice This article has been corrected since it published Online First. An author note has been added.

Acknowledgements The authors thank all study participants, researchers and clinicians who contributed to this study.

Contributors Gual, HS, QW and ZH designed the study and edited the manuscript; CY and MZ performed statistical analysis, conducted experiments and wrote the manuscript; TH, FY, JW, TW, JN, JC and YJ performed the genotyping and experiments; YD, MY, MenW, GL, CR, WY, BH and MeiW, RZ, YW and DX contributed to sample collection; EZ, JD, HM, SW, YC, ZX, JZ, GuoJ, ZW and ZZ reviewed data and provided critical comments or suggestions; Gual had primary responsibility for final content.

Funding This work was supported by grants from the National Major Research and Development Programme (2016YFC1302703); National Natural Science Foundation of China (81872702, 81521004 and 81573228); 333 High-Level Talents Cultivation Project of Jiangsu Province (BRA2018057); Jiangsu Outstanding Youth Fund (BK20160095); Project funded by China Postdoctoral Science Foundation (2019TQ0157).

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval This study was approved by the Institutional Review Board of Nanjing Medical University. Animal care and handling procedures were performed in accordance with the National Institutes of Health's Guide for the Care and Use of Laboratory Animals and were approved by the Committee on the Ethics of Animal Experiments of Nanjing Medical University (Nanjing, China).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Author note Dr Qingyi Wei is a visiting professor at Fudan University.

ORCID iDs

Zekuan Xu <http://orcid.org/0000-0001-5179-4128>

Zhibin Hu <http://orcid.org/0000-0002-8277-5234>

Guangfu Jin <http://orcid.org/0000-0003-0249-5337>

REFERENCES

- Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. *CA Cancer J Clin* 2016;66:115–32.
- Mucci LA, Hjeltnborg JB, Harris JR, et al. Familial risk and heritability of cancer among twins in Nordic countries. *JAMA* 2016;315:68–76.
- Hansford S, Kaurah P, Li-Chang H, et al. Hereditary diffuse gastric cancer syndrome: CDH1 mutations and beyond. *JAMA Oncol* 2015;1:23–32.
- Sahasrabudhe R, Lott P, Bohorquez M, et al. Germline mutations in PALB2, BRCA1, and RAD51C, which regulate DNA recombination repair, in patients with gastric cancer. *Gastroenterology* 2017;152:983–6.
- Fewings E, Larionov A, Redman J, et al. Germline pathogenic variants in PALB2 and other cancer-predisposing genes in families with hereditary diffuse gastric cancer without CDH1 mutation: a whole-exome sequencing study. *Lancet Gastroenterol Hepatol* 2018;3:489–98.
- Sakamoto H, Yoshimura K, Saeki N, et al. Genetic variation in PSCA is associated with susceptibility to diffuse-type gastric cancer. *Nat Genet* 2008;40:730–40.
- Abnet CC, Freedman ND, Hu N, et al. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. *Nat Genet* 2010;42:764–7.
- Shi Y, Hu Z, Wu C, et al. A genome-wide association study identifies new susceptibility loci for non-cardia gastric cancer at 3q13.31 and 5p13.1. *Nat Genet* 2011;43:1215–8.
- Saeki N, Saito A, Choi IJ, et al. A functional single nucleotide polymorphism in mucin 1, at chromosome 1q22, determines susceptibility to diffuse-type gastric cancer. *Gastroenterology* 2011;140:892–902.
- Hu N, Wang Z, Song X, et al. Genome-Wide association study of gastric adenocarcinoma in Asia: a comparison of associations between cardia and non-cardia tumours. *Gut* 2016;65:1611–8.
- Wang Z, Dai J, Hu N, et al. Identification of new susceptibility loci for gastric non-cardia adenocarcinoma: pooled results from two Chinese genome-wide association studies. *Gut* 2017;66:581–7.
- Tanikawa C, Kamatani Y, Toyoshima O, et al. Genome-Wide association study identifies gastric cancer susceptibility loci at 12q24.11–12 and 20q11.21. *Cancer Sci* 2018;109:4015–24.
- Zhu M, Yan C, Ren C, et al. Exome array analysis identifies variants in SPOCD1 and BTN3A2 that affect risk for gastric cancer. *Gastroenterology* 2017;152:2011–21.
- Helgason H, Rafnar T, Olafsdottir HS, et al. Loss-Of-Function variants in ATM confer risk of gastric cancer. *Nat Genet* 2015;47:906–10.
- Saeki N, Ono H, Yanagihara K, et al. rs2294008T, a risk allele for gastric and gallbladder cancers, suppresses the PSCA promoter by recruiting the transcription factor YY1. *Genes Cells* 2015;20:382–91.
- Sung H, Yang HH, Hu N, et al. Functional annotation of high-quality SNP biomarkers of gastric cancer susceptibility: the Yin Yang of PSCA rs2294008. *Gut* 2016;65:361–4.
- Marchini J, Howie B, Myers S, et al. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007;39:906–13.
- Mägi R, Morris AP. GWAMA: software for genome-wide association meta-analysis. *BMC Bioinformatics* 2010;11:288.
- Pruim RJ, Welch RP, Sanna S, et al. LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* 2010;26:2336–7.
- Khan A, Fornes O, Stigliani A, et al. JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework. *Nucleic Acids Res* 2018;46:D260–D266.
- Hallikas O, Palin K, Sinjushina N, et al. Genome-Wide prediction of mammalian enhancers based on analysis of transcription-factor binding affinity. *Cell* 2006;124:47–59.
- de Leeuw CA, Mooij JM, Heskes T, et al. MAGMA: generalized gene-set analysis of GWAS data. *PLoS Comput Biol* 2015;11:e1004219.
- Kanehisa M, Goto S. Kegg: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res* 2000;28:27–30.
- Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. the gene ontology Consortium. *Nat Genet* 2000;25:25–9.
- Liberzon A, Subramanian A, Pinchback R, et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 2011;27:1739–40.
- Ochiai N, Nishizuka M, Miyamoto T, et al. Targeted disruption of fad24, a regulator of adipogenesis, causes pre-implantation embryonic lethality due to the growth defect at the blastocyst stage. *Biochem Biophys Res Commun* 2013;438:301–5.
- Hwang J-Y, Kim D-H, Ji Y-I, et al. Recapitulation of previous genome-wide association studies with two distinct pathophysiological entities of gastric cancer in the Korean population. *J Hum Genet* 2013;58:233–5.
- Sokolova O, Naumann M. NF- κ B Signaling in Gastric Cancer. *Toxins* 2017;9. doi:10.3390/toxins9040119.
- Zhang Q, Lenardo MJ, Baltimore D. 30 years of NF- κ B: a Blossoming of relevance to human pathobiology. *Cell* 2017;168:37–57.
- Collins PE, Kiely PA, Carmody RJ. Inhibition of transcription by B cell leukemia 3 (Bcl-3) protein requires interaction with nuclear factor κ B (NF- κ B) p50. *J Biol Chem* 2014;289:7059–67.
- Tan P, Yeoh K-G. Genetics and molecular pathogenesis of gastric adenocarcinoma. *Gastroenterology* 2015;149:1153–62.
- Hardie DG. AMP-activated/SNF1 protein kinases: conserved guardians of cellular energy. *Nat Rev Mol Cell Biol* 2007;8:774–85.
- Li W, Wong CC, Zhang X, et al. CAB39L elicited an anti-Warburg effect via a LKB1-AMPK-PGC1 α axis to inhibit gastric tumorigenesis. *Oncogene* 2018;37:6383–98.
- Cheung KS, Chan EW, Wong AYS, et al. Metformin use and gastric cancer risk in diabetic patients after *Helicobacter pylori* eradication. *J Natl Cancer Inst* 2019;111:484–9.
- Zhou X-L, Xue W-H, Ding X-F, et al. Association between metformin and the risk of gastric cancer in patients with type 2 diabetes mellitus: a meta-analysis of cohort studies. *Oncotarget* 2017;8:55622–31.
- Wang L-D, Zhou F-Y, Li X-M, et al. Genome-Wide association study of esophageal squamous cell carcinoma in Chinese subjects identifies susceptibility loci at PLCE1 and C20orf54. *Nat Genet* 2010;42:759–63.
- Wu C, Hu Z, He Z, et al. Genome-Wide association study identifies three new susceptibility loci for esophageal squamous-cell carcinoma in Chinese populations. *Nat Genet* 2011;43:679–84.
- Mocellin S, Verdi D, Pooley KA, et al. Genetic variation and gastric cancer risk: a field synopsis and meta-analysis. *Gut* 2015;64:1209–19.
- Wei L, Shao M, Zhao Y, et al. Functional role of PLCE1 intronic insertion variant associated with susceptibility to esophageal squamous cell carcinoma. *Carcinogenesis* 2018;39:191–201.
- Zhai S, Liu C, Zhang L, et al. Plce1 promotes esophageal cancer cell progression by maintaining the transcriptional activity of snail. *Neoplasia* 2017;19:154–64.
- Chen Y, Wang D, Peng H, et al. Epigenetically upregulated oncoprotein PLCE1 drives esophageal carcinoma angiogenesis and proliferation via activating the PI-PLC ϵ -NF- κ B signaling pathway and VEGF-C/ Bcl-2 expression. *Mol Cancer* 2019;18:1.
- Johmura Y, Osada S, Nishizuka M, et al. Fad24 acts in concert with histone acetyltransferase HBO1 to promote adipogenesis by controlling DNA replication. *J Biol Chem* 2008;283:2265–74.
- Johmura Y, Watanabe K, Kishimoto K, et al. Fad24 causes hyperplasia in adipose tissue and improves glucose metabolism. *Biol Pharm Bull* 2009;32:1656–64.
- He Y, Wang C, Wang Z, et al. Genetic variant PLCE1 rs2274223 and gastric cancer: more to be explored? *Gut* 2016;65:359–60.
- Hayakawa Y, Sethi N, Sepulveda AR, et al. Esophageal adenocarcinoma and gastric cancer: should we mind the gap? *Nat Rev Cancer* 2016;16:305–18.
- Renehan AG, Zwahlen M, Egger M. Adiposity and cancer risk: new mechanistic insights from epidemiology. *Nat Rev Cancer* 2015;15:484–98.
- Ajani JA, Lee J, Sano T, et al. Gastric adenocarcinoma. *Nat Rev Dis Primers* 2017;3.
- McLean MH, El-Omar EM. Genetics of gastric cancer. *Nat Rev Gastroenterol Hepatol* 2014;11:664–74.
- El-Omar EM, Carrington M, Chow WH, et al. Interleukin-1 polymorphisms associated with increased risk of gastric cancer. *Nature* 2000;404:398–402.

© 2020 Author(s) (or their employer(s)) 2020. No commercial re-use. See rights and permissions. Published by BMJ.