Long noncoding RNA CCAT1 rs67085638 SNP contribution to the progression of gastric cancer in a Polish population

Tomasz Olesiński1, Anna Lutkowska2, Adam Balcerek2, Anna Sowińska3, P Piotrowski4, Tomasz Trzeciak5, Tomasz Maj1, Piotr Hevelke1 & Pawel P. Jagodziński2*

The role of the long noncoding RNA CCAT1 NC_000008.10:g.128220661C > T (rs67085638) in the development of colon cancer has been reported. Therefore, we assessed the prevalence of rs67085638 in patients with gastric cancer (GC). We also evaluated the effect of rs67085638 on B-cell-specific Moloney leukaemia virus insertion site 1 (BMI1) transcripts in primary GC and counterpart histopathologically confirmed disease-free margin tissue. Using high-resolution melting analysis, we evaluated rs67085638 frequency in patients with the GC genotype (n = 214) and controls (n = 502) in a Polish Caucasian population. qRT-PCR was used to determine BMI1 transcripts. We observed the trend of rs67085638 association in all patients with GC ($p_{\text{trend}} = 0.028$), a strong risk of the GC genotype in male ($p_{\text{trend}} = 0.035$) but not female ($p_{\text{trend}} = 0.747$) patients, and the association with non-cardia GC ($p_{\text{trend}} = 0.041$), tumour stages T3 ($p_{\text{trend}} = 0.014$) and T4 ($p_{\text{trend}} = 0.032$), differentiation grading G3 ($p_{\text{trend}} = 0.009$), lymph node metastasis stage N3 ($p_{\text{trend}} = 0.005$) and metastasis stage M0 ($p_{\text{trend}} = 0.027$). We found that significantly increased BMI1 transcripts were associated with the primary GC genotype classified as grade G3 ($p = 0.011$) and as lymph node metastasis N3 ($p = 0.010$) and counterpart marginal tissues ($p = 0.026$, $p = 0.040$, respectively) from carriers of the T/T versus C/C genotypes. rs67085638 may contribute to increased BMI1 transcripts and the progression and rapid growth of GC.

Gastric cancer (GC) is a deadly disease resulting in more than 841,000 deaths each year1. The development of GC is due to the interaction of environmental and genetic factors. Environmental carcinogenic factors encompass low socioeconomic status, tobacco smoking, radiation, Helicobacter pylori infection, low consumption of fruits and vegetables and high intake of salty and smoked food, male sex, obesity and older age2–6. However, not everyone is exposed to the environmental carcinogenic factors that develop into GC, which suggests a strong genetic background contributing to its occurrence6–8. A genetic study demonstrated the different variants of genes that contribute to the development of GC8. Genome-wide association studies (GWAS) have revealed many loci as major genetic components in the development of GC9–11. Many previously studied single nucleotide polymorphisms (SNPs) did not reach statistical significance in GWAS, but they may force the development and progression of GC6–8.

Long noncoding RNAs (lncRNAs) are a class of transcripts that are longer than 200 nucleotides in length and are not translated to proteins12. In the nuclei and cytosol, lncRNAs play critical roles in various mechanisms regulating gene expression, which can be related to development, progression and drug resistance in cancer13–15. Abnormal lncRNA expression has been demonstrated in many malignancies, including GC15. Recently, studies have suggested that colon cancer-associated transcript-1 (CCAT1), also designated LOC100507056 or the cancer-associated region lncRNA-5, plays a role in the growth and invasion of GC16–19. CCAT1 includes a region 2,628 base pairs in length, and its expression results in two isoforms of IncRNAs:

1Department of Oncological Gastroenterology, Maria Sklodowska-Curie National Research Institute of Oncology, Warsaw, Poland. 2Department of Biochemistry and Molecular Biology, Poznań University of Medical Sciences, 6 Święcickiego St., 60-781 Poznan, Poland. 3Department of Computer Science and Statistics, Poznań University of Medical Sciences, Poznan, Poland. 4Molecular Biology Department, National Institute of Geriatrics, Rheumatology and Rehabilitation, Warsaw, Poland. 5Department of Orthopedics and Traumatology, Poznan University of Medical Sciences, Poznan, Poland. *email: pjagodzi@ump.edu.pl
two generations were excluded from the control group. We also obtained body mass index (BMI) data for the Poznań, Poland, Subjects who were pregnant, had a history of cancer or had blood relatives with GC going back ing routine medical examinations at the Department of Radiotherapy of the Greater Poland Cancer Center in

---

**Table 1.** Available clinicopathological characteristics of patients with gastric cancer. T tumour, N node, M metastasis, G grade. *Age at first diagnosis.

| Sex          | Male | Female |
|--------------|------|--------|
| No. of patients | 139  | 75     |

*Mean age (years) ± SD*  
56.1 ± 11.8  
57.5 ± 12.1

| GC localization | Cardia | Non-cardia |
|-----------------|--------|------------|
| No. of patients | 54     | 160        |

| Histological type | Diffuse | Intestinal | Mixed | Undetermined |
|-------------------|---------|------------|-------|--------------|
| No. of patients   | 111     | 79         | 18    | 6            |

| Tumour stage | T1 | T2 | T3 | T4 | T5 |
|--------------|----|----|----|----|----|
| No. of patients | 24 | 32 | 86 | 72 |    |

| Lymph node metastasis stage | N0 | N1 | N2 | N3 |
|-----------------------------|----|----|----|----|
| No. of patients              | 78 | 35 | 42 | 59 |

| Metastasis stage | M0 | M1 |
|------------------|----|----|
| No. of patients | 192 | 22 |

| Histological grading | G1 | G2 | G3 |
|----------------------|----|----|----|
| No. of patients | 7  | 66 | 141|

**Materials and methods**

**Study population.** The studied subjects included 214 patients with gastric adenocarcinoma (GC, 54 cardia and 160 non-cardia) who were diagnosed by oncologists. All cases were gastric adenocarcinoma and were confirmed by pathological examination with routine evaluation of differentiation stage and grade according to the staging system of the tumour node metastasis (TNM) classification system and the World Health Organization (Table 1). Patient data and primary GC tissue samples were obtained from subjects enrolled between July 2016 and January 2019 at the Department of Oncological Gastroenterology, Maria Skłodowska-Curie Oncology Center, Warsaw, Poland. The inclusion criteria were age over 18 years, new diagnosis of gastric adenocarcinoma, and ability to endure surgical treatment or chemotherapy. The exclusion criteria encompassed patients with a non-adenocarcinoma and the presence of an extragastric tumour. Subjects with severe systemic disorders were also excluded from the group of patients with GC.

The control group was matched to patients by age and sex and included 175 healthy females with a mean age of 56.9 ± 12.3 years and 327 healthy males with a mean age of 55.8 ± 11.8 years randomly selected during routine medical examinations at the Department of Radiotherapy of the Greater Poland Cancer Center in Poznań, Poland, Subjects who were pregnant, had a history of cancer or had blood relatives with GC going back two generations were excluded from the control group. We also obtained body mass index (BMI) data for the controls and patients with GC as well as the presence of diabetes in the GC patients (Supplementary Table 1).

**Tissue samples.** We found the greatest association of rs67085638 with G3 differentiation grading and with N3 lymph node metastasis in patients with GC. Therefore primary GC tissue and counterpart histopathologically confirmed disease-free margin tissue from these patients were collected to determine the influence of rs67085638 on the level of BMI1 transcript. Primary GC tissue samples and distal surgical resection margin histopathologically confirmed disease-free tissue samples were obtained from 42 patients with a mean age of 51.3 ± 7.6 years and classified as grade G3 at the time of surgery and from 42 patients with a mean age of 52.4 ± 6.6 years and classified as N3 at the time of surgery at the Department of Oncological Gastroenterology, Maria Skłodowska-Curie Oncology Center, Warsaw. A portion of the tissue sample was immediately snap-frozen in liquid nitrogen and stored at − 80 °C until RNA isolation was performed.

**Genetic analysis.** DNA was isolated from peripheral blood cells via a salting-out procedure. The primers were designated using Oligo 7.6 software (DBA Oligo, Inc., Colorado Springs, CO). The NC_000008.10:g.128220661C > T (rs67085638) polymorphism DNA fragment (140 bp) was amplified using the following primers: forward 5′-GCTGTAATAACGCTGAT3′ and reverse 5′AACTGAATGATGAAAGG3′. The rs67085638 SNP was then genotyped via high-resolution melting (HRM) curve analysis previously described using HOT FIREPol EvaGreen (Solis BioDyne, Tartu, Estonia) with a LightCycler 480 system (Roche)
Diagnosics, Mannheim, Germany). The presence of this SNP was reanalysed by Sanger sequencing analysis of arbitrarily chosen samples, comprising 10% of the samples from both cases and controls. The concordance rate between HRM and sequencing was 100%.

Reverse transcription and quantitative real-time PCR (qRT-PCR) analysis of BMI1 transcript levels in GC tissue and histopathologically confirmed disease-free margin tissue. Frozen primary GC and distal surgical resection counterpart margin histopathologically confirmed disease-free tissues were homogenized, and total RNA was isolated according to the method of Chomczynski and Sacchi27. RNA quality was determined spectrophotometrically and by agarose gel electrophoresis. RNA samples were treated with DNase I, quantified, and reverse-transcribed into complimentary DNA (cDNA) (Supplementary Table 2). Quantitative analysis of BMI1 cDNA (Supplementary data 1) was performed by the Light Cycler480 II Real-Time PCR System (Roche Diagnostics GmbH, Mannheim, Germany) with SYBR Green I as the detection dye. BMI1 cDNA was quantified using the relative quantification method with a calibrator (Supplementary Table 2). The quantity of the BMI1 transcript in each sample was standardized by the geometric mean of reference transcript levels: hydroxymethylbilane synthase (HMBS) and beta-2-microglobulin (B2M) (Supplementary Table 2). The BMI1 mRNA levels were expressed as multiples of these cDNA concentrations in the calibrator.

Statistical analysis. The distinction in genotypic prevalence between the patients and controls and their genotype deviation from Hardy–Weinberg (HW) equilibrium were evaluated using a χ² test. The rs67085638 genotype deviation from HW equilibrium were evaluated using a χ² test. The rs67085638 SNP with GC, and the p-value of c<0.05 was considered statistically significant. Statistical analysis comparing BMI1 transcript levels between the T/T versus C/C and C/T versus C/C genotype carriers was evaluated using the Kruskal–Wallis test with Dunn’s post hoc test. Statistical analyses were conducted using Statistica version 10, 2011 (Stat Soft, Inc., Tulsa, USA), as previously described28.

Ethical approval. The study procedures were approved by the Local Ethical Committee of the Poznań University of Medical Sciences (reference number of ethical approval: 673/15 and 190/19). The study was carried out in accordance with the approved guidelines. Informed consent was obtained from all individual participants included in the study.

Results

Distribution of the CCAT1 rs67085638 SNP between the group of patients with GC and the control group. The χ² test of HW equilibrium showed values of 0.973 and 0.949 for the patients with GC and the control group, respectively. The genotype distribution of the rs67085638 SNP in the group of patients is presented in Table 2. We found a significant association of the rs67085638 SNP with all the patients with GC, and the p-trend value calculated for the rs67085638 polymorphism was p_trend = 0.028. The logistic regression analysis, which was adjusted for the effects of age, BMI and presence of diabetes, demonstrated an association of the rs6983267 SNP with T/T as well as the T/T + C/T genotype and T allele for all the patients with GC. For the T/T versus C/C genotypes, the adjusted OR was 1.360 (95% CI 1.011–1.828, p = 0.041); for the C/T + T/T versus C/C genotypes, the adjusted OR was 1.423 (95% CI 1.025–1.977, p = 0.041); and for T versus C, the adjusted OR was 1.373 (95% CI 1.056–1.784, p = 0.018). However, there was no significant association of the rs6983267 SNP with T/T versus C/C; the adjusted OR was 1.320 (95% CI 0.982–1.773, p = 0.065), and for C/T versus C/C, the adjusted OR was 1.252 (95% CI 0.922–1.985, p = 0.122).

Prevalence of the CCAT1 rs67085638 SNP among patients with cardia and non-cardia localization of GC. In the patients with cardia localization of GC, p_trend = 0.253, and for T/T versus C/C, the adjusted OR was 1.334 (95% CI 0.839–2.123, p = 0.222); for C/T versus C/C, the adjusted OR was 1.257 (95% CI 0.664–2.377, p = 0.481); for C/T + T/T versus C/C, the adjusted OR was 1.411 (95% CI 0.792–2.512, p = 0.241); and for T versus C, the adjusted OR was 1.279 (95% CI 0.820–1.993, p = 0.277).

In the patients with non-cardia localization of GC, we found a significant association of the rs67085638 SNP with GC, and the p-trend value calculated for the rs67085638 polymorphism was p_trend = 0.041. The logistic regression analysis demonstrated an association of the rs6983267 SNP with the C/T + T/T genotype as well as the T allele. For C/T + T/T versus C/C, the adjusted OR was 1.576 (95% CI 1.051–2.364, p = 0.028), and for T versus C, the adjusted OR was 1.404 (95% CI 1.047–1.881, p = 0.023). However, there was no significant association of the rs6983267 SNP with T/T versus C/C; the adjusted OR was 1.320 (95% CI 0.982–1.773, p = 0.065), and for C/T versus C/C, the adjusted OR was 1.252 (95% CI 0.922–1.985, p = 0.122).
In the patients with mixed histological types of GC, \( p = 0.091 \); and for T versus C, the adjusted OR was 1.402 (95% CI 0.999–1.966, \( p = 0.009 \); and for T versus C, the adjusted OR was 1.522 (95% CI 1.039–2.177, \( p = 0.022 \); for T/T versus C/C, the adjusted OR was 1.716 (95% CI 1.133–2.578, \( p = 0.009 \); and for T/T versus C/C, the adjusted OR was 1.505 (95% CI 1.048–2.159, \( p = 0.027 \). However, we did not observe an association of rs67085638 with T4 tumour stage; for T/T versus C/C, the adjusted OR was 1.403 (95% CI 0.924–2.133, \( p = 0.111 \).

We did not find a significant association of rs67085638 with T1 and T2 tumour stages of GC (Table 4). In the patients with T3 tumour stage GC, \( p_{\text{trend}} = 0.014 \), and for T/T versus C/C, the adjusted OR was 1.522 (95% CI 1.039–2.230, \( p = 0.031 \); for C/T versus C/C, the adjusted OR was 1.807 (95% CI 1.087–3.006, \( p = 0.022 \); for C/T + T/T versus C/C, the adjusted OR was 1.716 (95% CI 1.022–2.879, \( p = 0.041 \); and for T versus C, the adjusted OR was 1.565 (95% CI 1.116–2.194, \( p = 0.009 \). In patients with T4 tumour stage GC, \( p_{\text{trend}} = 0.032 \), and for C/T versus C/C, the adjusted OR was 2.190 (95% CI 1.205–3.979, \( p = 0.010 \); for C/T + T/T versus C/C, the adjusted OR was 2.164 (95% CI 1.215–3.855, \( p = 0.009 \); and for T versus C, the adjusted OR was 1.505 (95% CI 1.048–2.159, \( p = 0.027 \). However, we did not observe an association of rs67085638 with T4 tumour stage; for T/T versus C/C, the adjusted OR was 1.403 (95% CI 0.924–2.133, \( p = 0.111 \).

Prevalence of the CCAT1 rs67085638 SNP in patients with different tumour grading of GC. We observed an association of rs67085638 with G3 differentiation grading in the patients with GC (Table 4). In the patients with G3 differentiation grading of GC, \( p_{\text{trend}} = 0.009 \), and for T/T versus C/C, the adjusted OR was 1.468

| Genotype | Patients (frequency) | Controls (frequency) | Odds ratio (95%CI) | \( p^t \) | Adjusted odds ratio (95%CI) | \( p \) | \( p_{\text{trend}} \) |
|----------|---------------------|---------------------|-------------------|--------|-----------------------------|--------|-------------------|
| All      |                     |                     |                   |        |                             |        |                   |
| C/C      | 85 (0.40)           | 259 (0.48)          | Referent          | \( p = 0.028 \) |                             |        |                   |
| C/T      | 101 (0.47)          | 217 (0.43)          | 1.309 (0.9295–1.843) | 0.123  | 1.430 (0.981–2.085)         | 0.063  |                   |
| T/T      | 28 (0.13)           | 46 (0.09)           | 1.712 (1.006–2.911) | 0.0458 | 1.360 (1.011–1.828)         | 0.041  |                   |
| C/T + T/T| 129 (0.60)          | 263 (0.52)          | 1.379 (0.9964–1.909) | 0.0522 | 1.423 (1.025–1.977)         | 0.035  |                   |
| MAF       | 0.37                | 0.31                | 1.303 (1.027–1.653) | 0.029  | 1.373 (1.056–1.784)         | 0.018  |                   |

Table 2. Prevalence of the CCAT1 rs67085638 polymorphism among patients with GC and controls. Significant results are highlighted in bold font. \( ^{a} \chi^2 \). \( ^{b} \) ORs were adjusted by age, BMI and presence of diabetes. \( ^{c} \) Minor allele frequency.
Prevalence of the \textit{CCAT1} rs67085638 SNP in patients with various lymph node and GC metastasis stages. We found a significant association of rs67085638 with N3 lymph node metastasis in GC (Table 5). In the patients with N3 lymph node metastasis of GC, \(p\)\textsubscript{trend} = 0.0005, and for T/T versus C/C, the adjusted OR was 9.134 (95% CI 1.250–2.929, \(p = 0.003\)); for C/T versus C/C, the adjusted OR was 3.081 (95% CI 1.492–6.362, \(p = 0.002\)); for C/T + T/T versus C/C, the adjusted OR was 2.703 (95% CI 1.462–4.997, \(p = 0.001\)); and for T versus C, the adjusted OR was 1.980 (95% CI 1.345–2.915, \(p = 0.001\)). However, we did not observe an association between rs67085638 and the patients with N0, N1 and N2 lymph node metastasis (Table 5). In
the patients with N0 lymph node metastasis of GC, \( p_{\text{trend}} = 0.626 \), and for T/T versus C/C, the adjusted OR was 1.120 (95% CI 0.639–1.808, \( p = 0.642 \)); for C/T versus C/C, the adjusted OR was 1.229 (95% CI 0.739–2.072, \( p = 0.311 \); for C/T + T/T versus C/C, the adjusted OR was 1.319 (95% CI 0.769–2.262, \( p = 0.314 \)); and for T versus C, the adjusted OR was 1.192 (95% CI 0.798–1.781, \( p = 0.390 \). In the patients with N1 lymph node metastasis, \( p_{\text{trend}} = 0.909 \), and for T/T versus C/C, the adjusted OR was 1.152 (95% CI 0.648–2.048, \( p = 0.629 \)); for C/T versus C/C, the adjusted OR was 0.895 (95% CI 0.410–1.953, \( p = 0.780 \)); for C/T + T/T versus C/C, the adjusted OR was 1.029 (95% CI 0.514–2.059, \( p = 0.935 \)); and for T versus C, the adjusted OR was 1.082 (95% CI 0.639–1.830, \( p = 0.769 \). In the patients with N2 lymph node metastasis, \( p_{\text{trend}} = 0.475 \), and for T/T versus C/C, the adjusted OR was 1.120 (95% CI 0.300–1.669, \( p = 0.428 \); for C/T versus C/C, the adjusted OR was 0.649 (0.283–1.489, \( p = 0.304 \); for C/T + T/T versus C/C, the adjusted OR was 0.669 (0337–1.328, \( p = 0.335 \); and for T versus C, the adjusted OR was 1.022 (0.593–1.764, \( p = 0.937 \). In the patients with N3 lymph node metastasis, \( p_{\text{trend}} = 0.721 \), and for T/T versus C/C, the adjusted OR was 1.150 (0.648–2.048, \( p = 0.629 \)); for C/T versus C/C, the adjusted OR was 0.895 (0.410–1.953, \( p = 0.780 \)); for C/T + T/T versus C/C, the adjusted OR was 1.029 (0.514–2.059, \( p = 0.935 \)); and for T versus C, the adjusted OR was 1.082 (95% CI 0.639–1.830, \( p = 0.769 \).
was 1.324 (95% CI 0.810–2.167, \( p = 0.261 \)); for C/T versus C/C, the adjusted OR was 0.998 (95% CI 0.448–2.222, \( p = 0.997 \)); for C/T + T/T versus C/C, the adjusted OR was 1.118 (95% CI 0.562–2.225, \( p = 0.750 \)); and for T versus C, the adjusted OR was 1.204 (95% CI 0.719–2.017, \( p = 0.480 \)).

We also found a contribution of the rs67085638 SNP to the M0 metastasis stage of GC (Table 5). In the patients with M0 metastasis stage GC, we observed \( p_{\text{trend}} = 0.027 \), and for T/T versus C/C, the adjusted OR was 1.349 (95% CI 1.023–1.778, \( p = 0.033 \)); for C/T + T/T versus C/C, the adjusted OR was 1.438 (95% CI 1.021–2.025, \( p = 0.037 \)); and for T versus C, the adjusted OR was 1.357 (95% CI 1.033–1.782, \( p = 0.028 \)). However, we did not observe a significant association of rs67085638 with the M0 metastasis stage.

We did not find a contribution of the rs67085638 SNP to the M1 metastasis stage of GC (Table 5). In the patients with M1 metastasis stage, we observed \( p_{\text{trend}} = 0.638 \), and for T/T versus C/C, the adjusted OR was 1.175 (95% CI 0.458–2.267, \( p = 0.297 \)); for C/T versus C/C, the adjusted OR was 1.291 (95% CI 0.624–4.794, \( p = 0.291 \)); and for T versus C, the adjusted OR was 1.431 (95% CI 0.703–2.913, \( p = 0.322 \)).

### Table 5. Prevalence of the CCAT1 rs67085638 polymorphism among various lymph nodes and stages of metastasis of GC. Significant results are highlighted in bold font. a \( \chi^2 \) or b Fisher’s exact test. c ORs were adjusted by age, BMI and presence of diabetes. d Minor allele frequency.

| Lymph node metastasis stage | Genotype | Patients (frequency) | Controls (frequency) | Odds ratio (95% CI) | \( p^a \) | Adjusted odds ratio (95% CI)c | \( p^p \) | \( p_{\text{trend}} \) |
|----------------------------|----------|---------------------|---------------------|---------------------|----------|------------------------------|----------|----------------|
| N0 C/C                     | 34 (0.43)| 239 (0.48)          | Referent            | –                   | Referent | –                            | –        | 0.626         |
| C/T                       | 37 (0.47)| 217 (0.43)          | 1.199 (0.727–1.977) | 0.478               | 1.224    | (0.739–2.027)                | 0.431    |
| T/T                       | 7 (0.09) | 46 (0.09)           | 1.070 (0.447–2.560) | 0.880               | 1.120    | (0.639–1.808)                | 0.642    |
| C/T + T/T                 | 44 (0.56)| 263 (0.52)          | 1.176 (0.727–1.902) | 0.508               | 1.319    | (0.769–2.262)                | 0.314    |
| N1 C/C                     | 17 (0.48)| 239 (0.48)          | Referent            | –                   | Referent | –                            | –        | 0.909         |
| C/T                       | 14 (0.40)| 217 (0.43)          | 0.907 (0.437–1.888) | 0.794               | 0.895    | (0.410–1.953)                | 0.780    |
| T/T                       | 4 (0.11) | 46 (0.09)           | 1.223 (0.393–3.801) | 0.759               | 1.152    | (0.648–2.048)                | 0.629    |
| C/T + T/T                 | 18 (0.51)| 263 (0.52)          | 0.962 (0.485–1.910) | 0.912               | 1.029    | (0.514–2.059)                | 0.935    |
| N2 C/C                     | 19 (0.45)| 239 (0.48)          | Referent            | –                   | Referent | –                            | –        | 0.475         |
| C/T                       | 17 (0.40)| 217 (0.43)          | 0.985 (0.499–1.945) | 0.966               | 0.998    | (0.448–2.222)                | 0.997    |
| T/T                       | 6 (0.14) | 46 (0.09)           | 1.641 (0.622–4.33)  | 0.313               | 1.324    | (0.810–2.168)                | 0.261    |
| C/T + T/T                 | 23 (0.55)| 263 (0.52)          | 1.100 (0.584–2.071) | 0.768               | 1.118    | (0.562–2.225)                | 0.750    |
| N3 C/C                     | 15 (0.25)| 239 (0.48)          | Referent            | –                   | Referent | –                            | –        | 0.0005        |
| C/T                       | 33 (0.56)| 217 (0.43)          | 2.423 (1.281–4.584) | 0.0053              | 3.081    | (1.492–6.362)                | 0.002    |
| T/T                       | 11 (0.19)| 46 (0.09)           | 3.810 (1.645–8.824) | 0.001               | 9.134    | (1.250–2.929)                | 0.003    |
| C/T + T/T                 | 44 (0.74)| 263 (0.52)          | 2.666 (1.446–4.915) | 0.0012              | 2.703    | (1.462–4.997)                | 0.001    |
| M0 C/C                     | 76 (0.39)| 239 (0.48)          | Referent            | –                   | Referent | –                            | –        | 0.027         |
| C/T                       | 90 (0.47)| 217 (0.43)          | 1.304 (0.9132–1.863)| 0.1435              | 1.348    | (0.940–1.934)                | 0.104    |
| T/T                       | 26 (0.14)| 46 (0.09)           | 1.777 (1.030–3.068) | 0.0373              | 1.349    | (1.023–1.778)                | 0.033    |
| C/T + T/T                 | 116 (0.60)| 263 (0.52)          | 1.387 (0.989–1.945) | 0.0575              | 1.438    | (1.021–2.025)                | 0.037    |
| M1 C/C                     | 9 (0.41) | 239 (0.48)          | Referent            | –                   | Referent | –                            | –        | 0.638         |
| C/T                       | 11 (0.50)| 217 (0.43)          | 1.346 (0.547–3.311) | 0.516               | 1.175    | (0.608–5.060)                | 0.297    |
| T/T                       | 2 (0.09) | 46 (0.09)           | 1.155 (0.242–5.520) | 0.695               | 1.019    | (0.458–2.267)                | 0.962    |
| C/T + T/T                 | 13 (0.59)| 263 (0.52)          | 1.313 (0.5510–3.127)| 0.538               | 1.729    | (0.624–4.794)                | 0.291    |
| MAF^d                      | 0.34     | 0.31                | 1.163 (0.615–2.201) | 0.621^*              | 1.431    | (0.703–2.913)                | 0.322    |

\( ^a \chi^2 \) or \(^b\) Fisher’s exact test. \( ^c\) ORs were adjusted by age, BMI and presence of diabetes. \( ^d\) Minor allele frequency.
Influence of the rs67085638 polymorphism on BMI1 transcript levels in GC tissue and counterpart histopathologically confirmed disease-free margin tissue. We found significantly increased BMI1 transcript levels in the primary GC tissue classified as grade G3 ($p = 0.011$) and counterpart histopathologically confirmed disease-free grade G3 ($p = 0.026$, respectively) in the carriers of the T/T versus C/C genotypes but not for the carriers of the C/T versus C/C genotypes ($p = 0.056$, $p = 0.11$, respectively) (Fig. 1A, B). There was also a statistically significant increase in the BMI1 transcript levels in the primary GC tissue classified as lymph node metastasis N3 ($p = 0.010$) and counterpart histopathologically confirmed disease-free margin tissue ($p = 0.040$) in the carriers of the T/T versus C/C genotypes but not in the carriers of the C/T versus C/C genotypes ($p = 0.087$, $p = 0.23$, respectively) (Fig. 1C, D).

Discussion

CCAT1 and CCAT2 are located in the 8q24.21 region, which is frequently amplified in colorectal cancer. In the 8q24 region, several complex molecular interactions are tissue-specific between CCAT1 and MYC, which regulate CCAT1 and MYC gene expression. CCAT1 is situated inside a super-enhancer region of the MYC promoter. CCAT1-L protects against interactions between MYC and its enhancers, recruiting CCCTC-binding factor or CTCF. Moreover, CCAT1 lncRNA functions as a sponge for some tumour suppressor miRNAs to titrate the MYC protein. In this context, MYC binds to the promoter of CCAT1 and increases its transcription.

Overexpression of CCAT1 has been observed in various malignancies compared to normal counterpart tissue, including GC. CCAT1 displays oncogenic properties in various cancers, and in the cytoplasm or nucleus, it supports biological processes in cancer, such as proliferation, migration, cell cycle progression, apoptosis, and chemoresistance. Moreover, CCAT1 positively correlated with the clinicopathological characteristics,
progression and treatment outcomes of various cancers, including advanced TNM stage, vascular invasion, overall survival and recurrence-free survival, among others\(^9\).

In our studies, we observed a trend of rs67085638 SNP association in all patients with GC. rs67085638 was a strong risk factor for GC in male but not female patients. To date, the rs67085638 and rs7013433 polymorphisms have been demonstrated as risk factors for colorectal cancer\(^7\). However, we did not find an association between the rs7013433 polymorphism and either the occurrence of GC or the clinicopathological characteristics of GC (data not shown). The increased risk of rs67085638 in males can be due to the additive effect of environmental factors, including tobacco smoking, regular alcohol consumption, and limited consumption of fruits and vegetables, which are common in males in our country. Our observation is consistent with findings in other populations in which male sex is an important factor contributing to the development of GC\(^6,10\).

To date, the role of CCAT1 lncRNA in the development of GC has been intensively studied. CCAT1 was significantly elevated in primary GC tissue compared with normal gastric tissue and significantly correlated with the progression of GC and supports the proliferation and migration of GC cells\(^16\). A detailed study revealed that CCAT1 is involved in the negative regulation of the miR-219-1 and miR-490/hnRNP A1 axes, contributing to malignant transformation, progression and migration of GC cells\(^19,25,27\). CCAT1 expression was associated with larger gastric tumour size, lymphatic metastasis and advanced TNM stage\(^18\). CCAT1-L is also involved in the epithelial-mesenchymal transition and metastasis of gastric adenocarcinoma\(^18\).

We found that rs67085638 was associated with tumour stages T3 and T4, histological differentiation G3, lymph node metastasis stage N3 and metastasis stage M0. In addition, we determined that homozygous carriers of the CCAT1 rs67085638 T allele exhibited increased BMI1 transcript levels compared to homozygous carriers of the CCAT1 rs67085638 C allele from primary GC tissues as well as in histopathologically confirmed cancer-free margin tissue. Our findings suggest that the CCAT1 rs67085638 T allele elevating BMI1 expression might support increased rapid tumour growth compared to lower-grade tumour cells and the expansion of cancer cells into the outer lining of the stomach or other organs and lymph node metastasis.

BMI1 is an oncogene and catalytic component of polycomb group proteins involved in epigenetic gene silencing. BMI1 displays a critical function in tissue-specific regulation of gene expression and consequently several elementary cellular processes\(^21\). SNP rs6983267 has been shown to be located in the enhancer region of MYC, which plays a role in the regulation of CCAT1 expression\(^21\). Recently, it was demonstrated that the presence of the T allele of rs67085638 increased CCAT1 expression\(^22\). Moreover, CCAT1 expression correlated with BMI1 mRNA and protein levels in both GC cells in vitro and in vivo in murine tumour models\(^23\). CCAT1 knockdown repressed the proliferation and invasion of gallbladder cancer cells via miR-218-5p controlling BMI1 transcript translation\(^27\). In cigarette smoke extract-exposed human bronchial epithelial cells, CCAT1 negatively regulates miR-218, which modulates BMI1 expression and cell cycle progression\(^28\). BMI1 has been found to be a characteristic marker of poor prognosis in patients with breast cancer, nasopharyngeal carcinoma, oesophageal squamous cell carcinoma and gastric carcinoma\(^6,30,40\). Several studies have shown that BMI1 is associated with the progression, epithelial-mesenchymal transition and metastasis of GC\(^41\). The reduced expression of BMI1 inhibits epithelial-mesenchymal transition and spreading of melanoma cells\(^41\).

**Conclusion**

Our study is the first to demonstrate that the CCAT1 rs67085638 SNP is a risk factor for gastric carcinogenesis in Caucasian Polish individuals. We also found that homozygous carriers of CCAT1 rs67085638 T were associated with increased BMI1 transcript levels and increased rapid growth compared to lower-grade tumour cells, spread of cancer cells to neighbouring tissues and lymph node metastasis. However, our studies have some limitations. In patients with tumour stage T4, we observed the greatest contribution of C/T but not T/T to GC, which may be due to our relatively small group of patients with GC. This study should be repeated in other independent cohorts. Moreover, we need precisely explain the mechanism by which rs67085638 may regulate the transcription of BMI1.

Received: 18 June 2020; Accepted: 7 July 2021
Published online: 28 July 2021

**References**

1. Global Burden of Disease Cancer Collaboration et al. The global burden of cancer 2013. *JAMA Oncol.* 1(4), 505–527 (2015).
2. Karimi, P., Islami, F., Anandasabapathy, S., Freedman, N. D. & Kamangar, F. Gastric cancer: Descriptive epidemiology, risk factors, screening, and prevention. *Cancer Epidemiol. Biomark. Prev.* 23, 700–713 (2014).
3. Ma, S. H. et al. Impact of alcohol drinking on gastric cancer development according to *Helicobacter pylori* infection status. *Br. J. Cancer* 113, 1381–1388 (2015).
4. Nomura, A. M., Wilkens, L. R., Henderson, B. E., Epplein, M. & Kolonel, L. N. The association of cigarette smoking with gastric cancer: The multiethnic cohort study. *Cancer Causes Control* 23, 51–58 (2012).
5. Camargo, M. C., et al. Improved survival of gastric cancer with tumour Epstein–Barr virus positivity: An international pooled analysis. *Gut* 63, 236–243 (2014).
6. Fang, X. et al. Landscape of dietary factors associated with risk of gastric cancer: A systematic review and dose-response meta-analysis of prospective cohort studies. *Eur. J. Cancer* 51, 2820–2832 (2015).
7. Choi, Y. J. & Kim, N. Gastric cancer and family history. *Korean J. Intern. Med.* 31, 1042–1053 (2016).
8. Gieck, O. et al. Genetic variants in gastric cancer: Risks and clinical implications. *Exp. Mol. Pathol.* 103, 101–111 (2017).
9. Helgason, H. et al. Loss-of-function variants in ATM confer risk of gastric cancer. *Nat. Genet.* 47, 906–910 (2015).
10. Cai, M. et al. Environmental factors, seven GWAS-identified susceptibility loci, and risk of gastric cancer and its precursors in a Chinese population. *Cancer Med.* 6, 708–720 (2017).
11. Hishida, A. et al. GWAS analysis reveals a significant contribution of PSMA to the risk of *Helobacter pylori*-induced gastric atrophy. *Carcinogenesis* 40, 661–668 (2019).
This study was funded by Grant No. 502-01-01124182-07474 from Poznań University of Medical Sciences and by Maria Skłodowska-Curie National Research Institute of Oncology subsidy from Ministry of Science and

Author contributions
T.O. Research idea, wrote the main manuscript text, Development of research method, prepared Fig. 1, supervisor molecular biology technician, performing experiments. A.L. Collection data, supervisor molecular biology technician. A.B. Performing experiments. A.S. Statistics. P.P. Collection data, description of the material section. T.T. Prepared supplementary data, supervisor molecular biology technician. T.M. Collection data, description of the material section. P.H. Collection data, description of the material section. P.J. Wrote the main manuscript text, approval of manuscript. All authors reviewed the manuscript.

Funding
This study was funded by Grant No. 502-01-01124182-07474 from Poznań University of Medical Sciences and by Maria Skłodowska-Curie National Research Institute of Oncology subsidy from Ministry of Science and
Higher Education. The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-021-94576-9.

Correspondence and requests for materials should be addressed to P.P.J.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

© The Author(s) 2021