의학석사 학위논문

Assessment of copy number in proto-oncogenes are predictive of poor survival in advanced gastric cancer

진행성 위암에서 원발암유전자의 복제수 평가에 의한 낮은 생존율의 예측

2021년 8월

서울대학교 대학원

의과대학 병리학 전공

리 메 이 휘

Assessment of copy number in proto-oncogenes are predictive of poor survival in advanced gastric cancer

지도 교수 강 경 훈

이 논문을 의학석사 학위논문으로 제출함
2021년 4월

서울대학교 대학원
의과대학 병리학 전공
리 메 이 휘

리메이휘의 의학석사 학위논문을 인준함
2021년 7월

Chair _____
Vice Chair _____
Examiner _____

Abstract

MEIHUI LI

Department of Pathology, College of medicine

The Graduate School

Seoul National University

The copy number (CN) gain of proto-oncogenes is a frequent finding in gastric carcinoma (GC), but its prognostic implication remains elusive. The study aimed to characterize the clinicopathological features, including prognosis, of GCs with copy number gains in multiple protooncogenes. Three hundred thirty-three patients with advanced GC were analyzed for their gene ratios in *EGFR*, *GATA6*, *IGF2*, and *SETDB1* using droplet digital PCR (ddPCR) for an accurate assessment of CN changes in target genes. The number of GC patients with 3 or more genes with CN gain was 16 (4.8%). Compared with the GCs with 2 or less genes with CN gain, the GCs with 3 or more CN gains displayed more frequent venous invasion, a lower density of tumor-infiltrating lymphocytes (TILs), and lower methylation levels of L1 or SAT-alpha. Microsatellite instability-high tumors or Epstein-Barr virus-positive tumors were not found in the GCs with 3 or more genes with CN gain. Patients of this groups also showed the worst clinical outcomes for both overall survival and recurrence-free survival, which was persistent in the multivariate survival analyses. Our findings suggest that the ddPCR-based detection of multiple CN gain of protooncogenes might help to

identify a subset of patients with poor prognosis.

Keyword : Copy number gain, droplet digital PCR, gastric cancer, proto-oncogene, prognosis

Student Number : 2019-27243

Table of Contents

List of Tables

List of Figures

Introduction

Gastric carcinoma (GC) is one of the most common malignancies in Eastern Asia and one of the leading causes of cancer-related deaths. TNM cancer staging provides prognostic information, but clinical outcomes vary among patients with GC of the same cancer stage. For patients with GC of the same cancer stage, further prognostic information could be gained from biomarkers including pathological parameters, such as lymph vascular invasion, perineural invasion, and tumor-infiltrating lymphocytes (TILs). Molecular markers might provide information about the prognostic features of the tumor. The Cancer Genome Atlas (TCGA) project has defined four subtypes of GC, including GCs with microsatellite instability (MSI), Epstein-Barr virus (EBV), genomic stability (GS), and chromosomal instability (CIN), which have been associated with different prognoses [1,2]. The EBV subtype was associated with the best prognosis, while the GS subtype was associated with the worst prognosis. Although the CIN subtype fell in between the above two subtypes, it demonstrated the greatest survival benefit from adjuvant chemotherapy [1], which indicates that the molecular subtyping of GCs might provide prognostic and predictive value.

CIN consists of numerical and/or structural aberrations in chromosomes. Numerical abnormality refers to the gain or loss of whole chromosomes, whereas structural abnormalities include the amplification, loss, translocation, and inversion of chromosomal regions of various sizes ranging from a single gene to an arm. Through the

TCGA project, many genes have been found to be amplified or undergo copy number gain, including *EGFR*, *FGFR1*, *GATA6*, *HER2* (*ERBB2*), *IGF2*, *MYC*, and *SETDB1* in GCs [2]. Although copy number gains of these genes are expected to occur mainly in the CIN subtype of GC, the prognostic significance of the copy number gains of these genes has not yet been clarified.

Digital polymerase chain reaction (dPCR) is a method that provides quantitative information about copy number changes in probed genes without the need for standard curves. DNA samples obtained from formalin-fixed archival tissues contain inhibitors for PCR and formalin-induced interstrand crosslinking, which can result in errors in the analysis of copy number variation by quantitative PCR (qPCR). However, dPCR can provide more accurate results because it does not use the comparison of PCR rates relied on by qPCR but instead uses the determination of whether amplification above a threshold has occurred. In the present study, we aimed to elucidate whether copy number changes in seven genes (*EGFR*, *FGFR1*, *GATA6*, *HER2*, *IGF2*, *MYC*, and *SETDB1*) are related to the survival of patients with advanced GC and might serve to detect a subset of GC cases with poor prognosis. The genes included in this study belong to those which are most frequently amplified in TCGA. We used droplet dPCR (ddPCR) to evaluate the copy number changes of the seven genes in formalin-fixed, paraffin embedded tissue samples of advanced GC.

Materials and Methods

Patient samples

A total of 333 formalin-fixed, paraffin-embedded GC tissues were retrieved from the surgical files of the Department of Pathology, Seoul National University Hospital, Seoul, Korea. The patients underwent surgery and extended lymph node dissection (D2) for advanced GC (T2-T4) from 2007 to 2008. Patients were included in the study according to the following criteria: age at diagnosis >18 years, advanced GC, adenocarcinoma histology, and availability of formalin-fixed paraffin-embedded (FFPE) cancer tissues. The exclusion criteria included patients who had a history of other primary malignancies (except for papillary thyroid cancer) within 5 years or received chemotherapy before surgical resection. Clinical and histological information was obtained from electronic medical records, including Lauren histology, tumor subsite within the stomach, lymphatic embolus, venous invasion status, perineural invasion status, and tumor-node-metastasis (TNM) stage (American Joint Committee on Cancer, 7th edition). The patients were previously analyzed for their EBV infection, MSI, and tumoral L1 and SAT-alpha methylation statuses [3,4]. A tissue microarray was constructed from the tumor center and immune stained against CD3 and CD8. TILs were counted in the CD3 and CD8-immunostained cores (0.2 cm in diameter), and CD3 and CD8 TIL densities were determined in a previous study [5]. For the ddPCR assay, on glass slides, we marked the tumor areas with the highest tumor purity and the most representative histology

of the case under the microscope and then manually dissected the corresponding tumor areas on three to five unstained serial sections (10 μm-thick). The dissected tissues were subjected to DNA extraction and purification using a QIAamp DNA FFPE Tissue kit (Qiagen Gmbh, Hilden, Germany). Purified DNA was quantified with a Qubit® 2.0 fluorometer (Thermo Scientific, Wilmington, DE). As control DNA, normal genomic DNA was extracted from the nonneoplastic gastric mucosa of GC patients and from the white blood cells of healthy volunteers using the QIAamp DNA Mini Kit (Qiagen). This study was approved by the Institutional Review Board of Seoul National University Hospital, which waived the requirements to obtain informed patient consent (approval no. H-1312-051-542). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Design of primers and probes

The primers and probe sequences of one protooncogene (*EGFR*) and reference gene (Ribonuclease P RNA component H1 (*RPPH1*)) are listed in Table 1 [6] and were synthesized by Integrated DNA Technologies (Coralville, IA, USA). The primers and probes of the other six genes, including *MYC*, *HER2* (*ERBB2*), *FGFR1*, *GATA6*, *IGF2*, and *SETDB1*, were purchased from Life Technologies (Carlsbad, CA, USA). *RPPH1* was used as a reference locus because

it is a highly conserved region that is present at 1 copy per haploid genome.

Droplet digital PCR

Both the target gene and *RPPH1* loci were amplified simultaneously in duplex PCR. PCR mixtures were made with ddPCR Supermix for Probes (Bio-Rad, Hercules, CA) according to the manufacturer's protocol. Each 20 μl reaction mixture contained 100 ng of DNA, 900 nM of forward and reverse primers, and 250 nM of probes. We omitted the restriction enzyme digestion of DNA because formalin-fixed DNA tends to be highly fragmented. dPCR was performed on the QX200 ddPCR system (Bio-Rad). In brief, 20 μl of the PCR mixture was partitioned into an emulsion of approximately 20,000 uniformly sized droplets via a QX200 droplet generator. The droplets were transferred to a 96-well PCR plate, heat-sealed, and placed in a conventional thermal cycler (T100, Bio-Rad). The thermal cycling conditions were 95°C for 10 min; 40 cycles of 94°C for 30 s, 57°C for 50 s and 72°C for 30 s; 98°C for 10 min; and a 12°C hold. After PCR, the plate was loaded on a QX200 droplet reader for the automatic detection of the fluorescence in each well. Analysis of the ddPCR data was performed with QuantaSoft software (Bio-Rad).

Pyrosequencing methylation assay of L1 and SAT-alpha

After bisulfite modification of the extracted DNA, the modified DNA was subjected to pyrosequencing methylation assays of L1 and

SAT-alpha. The detailed procedures and determination of methylation levels were described in a previous study [4].

Quantitative real-time RT-PCR (qRT-PCR)

Total RNA was extracted from cells using RNeasy Plus Mini Kit (QIAGEN). First-strand cDNA was synthesized with LeGene Premium Express 1st Strand cDNA Synthesis System (LeGene Biosciences, San Diego, CA, USA) and either stored at -20°C or used immediately. Quantitative RT-PCR (RT-qPCR) reaction and analysis were performed using Bio-Rad iQ5 System (BioRad, Hercules, CA, USA). SYBRTM Green PCR Master Mix (Thermo Fisher Scientific, Waltham, MA, USA) was used for SYBR Green-based RT-qPCR according to the manufacturer's protocol. Primers used in RT-qPCR reactions were purchased from the Integrated DNA Technologies (Coralville, IA, USA) and listed in Table 2. Each qPCR analysis was done in triplicate, and the average of the triplicate values represents every single result of the qPCR analysis. Fold changes in gene expression of test and control samples were determined by using the $2^{-\Delta\Delta Ct}$ method. Relative quantity (RQ) is $2^{-\Delta\Delta Ct}$ and copy number variation (CNV) is 2 x RQ.

Cell culture

Cell lines SNU-1, SNU-5, SNU-16, SNU-216, SNU-484, SNU-601, SNU-620, SNU-638, SNU-668, SNU-719, MKN-28, MKN-45, and MKN74 were cultured in RPMI-1640 (Welgene Co., Daegu, Korea)

supplemented with 10% heat-inactivated FBS (Fetal bovine serum) (Gibco, Grand Island, NY, USA), 100 U/ml penicillin, and 100 μg/ml streptomycin (Gibco). FU-97 cell lines were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% heat-inactivated FBS (Fetal bovine serum) (Gibco, Grand Island, NY, USA), 100 U/ml penicillin, 100 μg/ml streptomycin (Gibco), and 10mg/L Insulin (Sigma-Aldrich, St. Louis, MO, USA). All cells were grown in an incubator with 5% CO_2 at 37°C.

Statistical analysis

Statistical analyses were performed using SPSS software for Windows, version 25.0 (IBM, Armonk, NY, USA). Two-sided *P*-values of less than 0.05 were considered statistically significant. To identify whether the gene ratios were normally distributed in GC tissue samples, normalization tests were performed for the gene ratios, which revealed that the gene ratios were not normally distributed. The mean values of the gene ratios across two groups or across three or more groups were compared using both Student's t-test and the Mann-Whitney test and both ANOVA and the Kruskal-Wallis test, respectively. The clinical outcome data were last updated in December 2019. Of the included 333 patients, 14 patients were lost to follow-up. Recurrence-free survival (RFS) was measured from the date of surgery for advanced GC to the date of the first documented recurrence or the date of death from any cause, whichever occurred first. Overall survival (OS) was calculated from

the date of resection to the date of death from any cause or the last clinical follow-up time. Survival curves were assessed using the Kaplan-Meier method and the log rank test. Multivariate comparisons of survival rates were performed with the Cox proportional hazards regression model, and baseline characteristics were adjusted using a backward stepwise regression model including covariates of prognostic value.

Table 1. Oligonucleotide sequences for the primers and probes

		5' mode Gene name	Sequence	3′ mode	Amplicon
HER2 (ERBB2)	forward		AAGCTAAG AATAAGGCCAGATGG		75 bp
	reverse		CGCACAGCACCAAGGAAAAG		
	probe	FAM	CAGCAGAACAACGCAGCCCTC CCT	BHQ1	
RPPH1 (reference)	forward		GCGGATGCCTCCTTTGC		73 bp
	reverse		ACCTCACCTCAGCCATTGAA CT		
	probe	HEX	CTTGGAACAGACTCACGGCC AGCG	BHQ1	

Paired primers and probes for *EGFR*, *FGFR1*, *GATA6*, *IGF2*, *MYC*, and *SETDB1* were purchased from ThermoFisher Scientific. Assay IDs for the six genes are Hs07528418_cn, Hs02882334_cn, Hs06475245_cn, Hs04392053_cn, Hs02602824_cn, and Hs01643975_cn, respectively.

Table 2. The primer sequences for qRT-PCR

Gene	Primer Sequence (5′~3′)	NCBI Gene Number	AT(°C)
GATA6	F : CTCTACAGCAAGATGAACGG	NM_005257	55
	R : CCATAAGGTGGTAGTTGTGG		
EGFR	F : TAACAAGCTCACGCAGTTGG	NM_001346897.2	58
	R : GTTGAGGGCAATGAGGACAT		
IGF2	F: CTGGAGACGTACTGTGCTA	NM_000612.6	55
	R : GCTTCCAGGTGTCATATTGG		
SETDB1	F : AGGAACTTCGGCATTTCATCG	NM_001145415	55
	R : TGTCCCGGTATTGTAGTCCCA		
GAPDH 1	F : TGGTAAAGTGGATATTGTTG C	NM_001256799.3	58
	R: GCCATGGGTGGAATCATA		

Abbreviation: AT, annealing temperature; F, Forward; R, Reverse.

Results

A total of 333 advanced GC patients were analyzed for their gene ratios using ddPCR. The demographic findings are summarized in Table 3. The mean age of the patients was 60.8 years, with a median age of 61 years (ranging from 29 to 86 years). The male to female ratio was 223:110. The cancer stage was IB in 30 patients, II in 108 patients, and III in 195 patients. Regarding the tumor subsite within the stomach, 92 cases involved the upper one-third (high body and cardia), while the others did not involve the upper one-third. Lauren's histology was intestinal type in 125 cases, diffuse type in 161 cases, mixed type in 43 cases, and unclassified in 4 cases. The molecular subtype was the MSI subtype in 42 cases, the Epstein-Barr virus subtype in 26 cases, and the non-MSI/non-EBV subtype in 265 cases.

Table 3. Demographical findings

Parameters	
Age (mean, median, range) (yrs)	60.8 (61) (29-86)
Sex (M:F)	223 : 110
Site (not involving cardia; involving cardia)	241 : 92
Lauren (intestinal: diffuse: mixed; unclassified)	125 : 161 : 43 : 4
Lymphatic emboli (absent: present)	105 : 228
Venous invasion (absent: present)	239 : 94
Perineural invasion (absent: present)	138 : 195
Cancer staging (I: II: III: IV)	30 : 103 : 152 : 48
N category (N0; N1; N2: N3a: N3b)	91 : 56 : 68 : 72 : 46
T category (T2: T3: T4a: T4b)	61 : 121 : 133 : 18
M category (M0: M1)	284 : 49
Molecular subtype (MSS/EBV-: MSI-H: EBV+)	265 : 42 : 26

Gene ratios in gastric cancer tissues and their normality tests

To identify whether the gene ratio ranged from approximately 1 in normal cells, we analyzed the statuses of the seven genes in the peripheral blood leukocytes of normal volunteer subjects (n=20). The average values of the probed gene ratio ranged from 0.89 to 1.30 with standard deviation values less than 0.196. However, for GC tissue samples, the average values of the seven gene ratios ranged from 0.84 to 43.5 (Table 4). To identify whether the gene ratios were normally distributed in the GC tissue samples, a normalization test was performed using the Shapiro-Wilk test, which showed that the gene ratios were non-normally distributed.

Table 4. Gene ratios of seven genes relative to *RPPH1* in peripheral blood leukocytes and advanced gastric cancers

	Peripheral blood leukocyte (n=20)			Advanced gastric cancer (n=333)		
	Average	Median	SD	Average	Median	SD
EGFR	1.30	1.28	0.196	2.69	1.87	5.574
FGFR1	1.03	1.03	0.027	1.07	0.93	1.206
GATA6	1.15	1.14	0.077	1.23	1.07	0.736
HER2 (*ERBB2*)	1.01	1.01	0.025	43.5	1.69	506.366
IGF2	1.18	1.17	0.121	0.84	0.79	0.489
MYC	1.06	1.07	0.043	2.09	1.69	2.005
SETDB1	0.89	0.89	0.052	1.68	1.6	0.698

Gene ratios and survival

For the survival analysis, the GC patients were grouped into 10 equal-sized subsets (i.e., each group has approximately the same number of patients), from subset 1 to subset 10, according to the increasing order of the gene ratios of the individual genes. With the Kaplan-Meier log rank test, *EGFR*, *FGFR1*, *GATA6*, *IGF2*, and *SETDB1* showed lower survival in subset 10 than in the other subsets (Fig. 1 & 2). When the patients were divided into subset 10 and the other subsets, *EGFR*, *FGFR1*, *GATA6*, *IGF2*, and *SETDB1* exhibited significant differences in survival time between subset 10 and the other subsets in the Kaplan-Meier log rank test (Fig. 3 & 4). The clinicopathological and molecular features that were found to be statistically significant in univariate survival analysis included tumor subsite, Lauren classification, T stage, N stage, M stage, venous invasion, lymphatic embolus, and perineural invasion. When the individual genes were included in multivariate survival analysis with clinicopathological factors that were found to be significantly associated with survival, the *EGFR* and *IGF2* gene ratios were independent prognostic parameters associated with poor prognosis in terms of both OS and RFS (Table 5). The *GATA6* gene ratio was found to be a significant risk factor in the multivariate analysis of OS only, and the *SETDB1* gene ratio was found to be a significant risk factor in the multivariate analysis of RFS only.

To evaluate the additive effect of copy number gains in four genes (*EGFR*, *GATA6*, *IGF2*, and *SETDB1*) on prognostication power, a

tumor was scored “1” or “0” when the specific gene ratio belonged to subset 10 or the other subsets, respectively. The sum of scores for the four genes ranged from 0 to 4 in each tumor. Although the survival curves of the sum scores were significantly different for OS and RFS (Kaplan-Meier log rank test), the survival curves of sum scores 1 and 2 were similar, and those of sum scores 3 and 4 were similar (Fig. 5 A& B). Thus, the GC patients were classified into 3 subsets, including a subset with sum score 0, a subset with sum score 1 or 2, and a subset with sum score 3 or 4 (Fig. 1 C& D). The sum scores of 3 and 4 were also independent prognostic factors of OS (HR = 3.805, 95% CI = 2.014 - 7.188, $P < 0.001$) and RFS (HR = 3.709, 95% CI = 1.953 - 7.042, $P < 0.001$) in GC patients regardless of tumor subsite, Lauren histology, venous invasion, lymphatic invasion, perineural invasion, and T, N, and M categories (Table 6).

To identify whether CNV determined by ddPCR was correlated with expression levels of mRNA in four genes, we measured mRNA expression levels of four genes in 14 gastric cancer cell lines, using RT-qPCR, which were analyzed for their CNV in four genes using ddPCR. Four genes showed significant correlations between RT-qPCR and ddPCR values (Fig. 6).

Figure 1

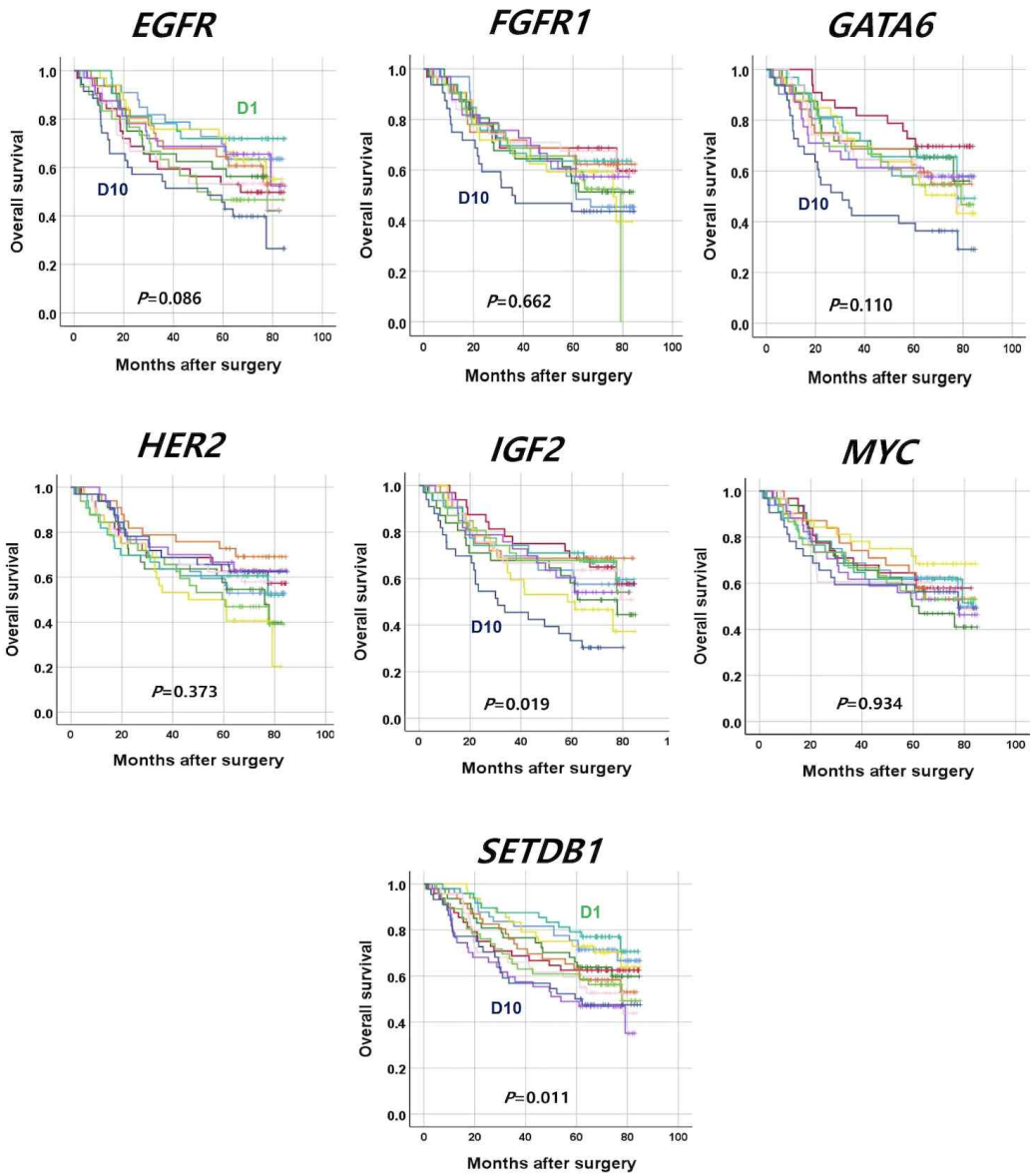

EGFR
D1
D10
P=0.086
FGFR1
D10
P=0.662
GATA6
D10
P=0.110
HER2
P=0.373
IGF2
D10
P=0.019
MYC
P=0.934
SETDB1
D1
D10
P=0.011
Overall survival
Months after surgery

Figure 2

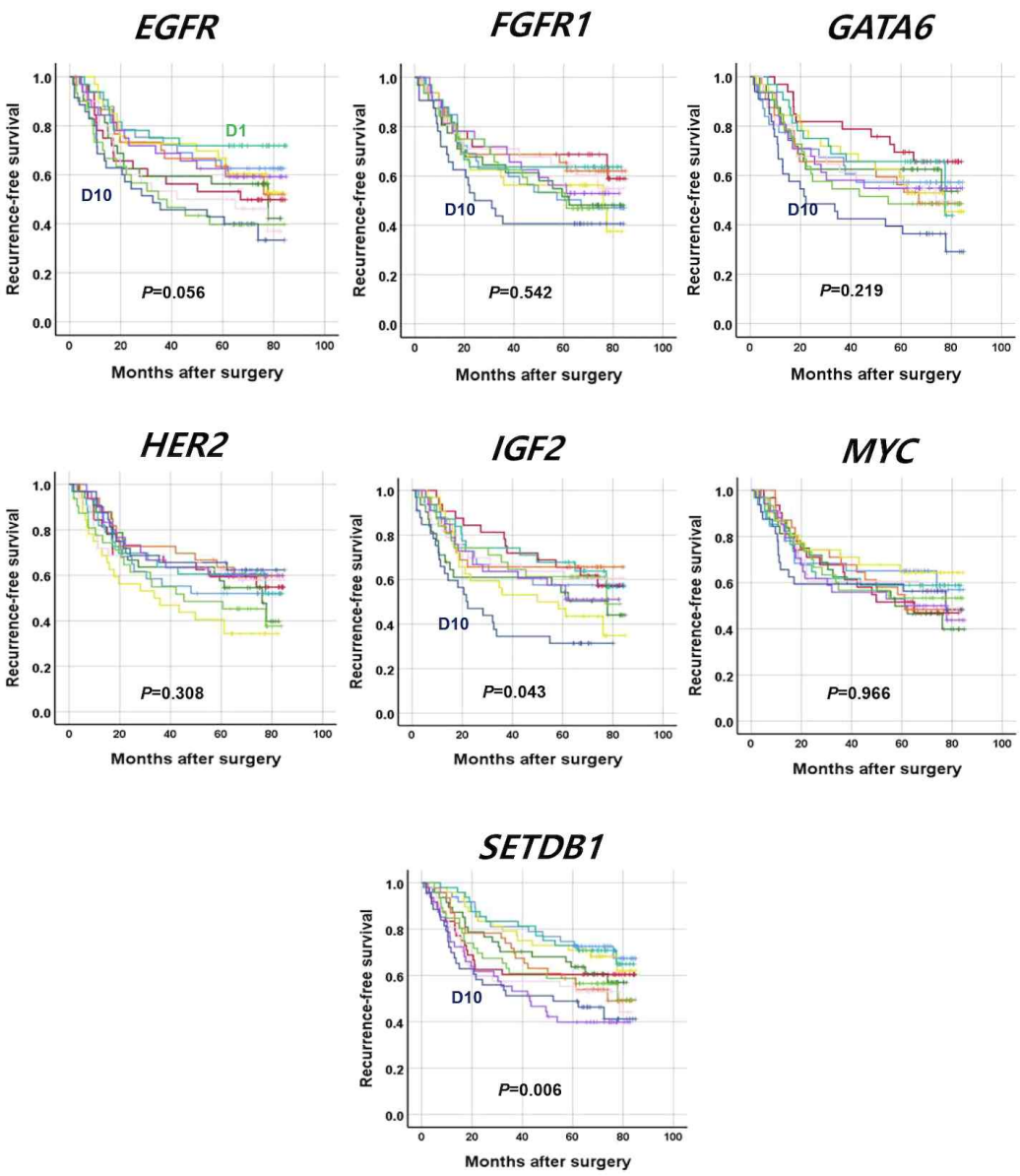

EGFR
D1
D10
P=0.056
FGFR1
D10
P=0.542
GATA6
D10
P=0.219
HER2
P=0.308
IGF2
D10
P=0.043
MYC
P=0.966
SETDB1
D10
P=0.006
Recurrence-free survival
Months after surgery

Figure 3

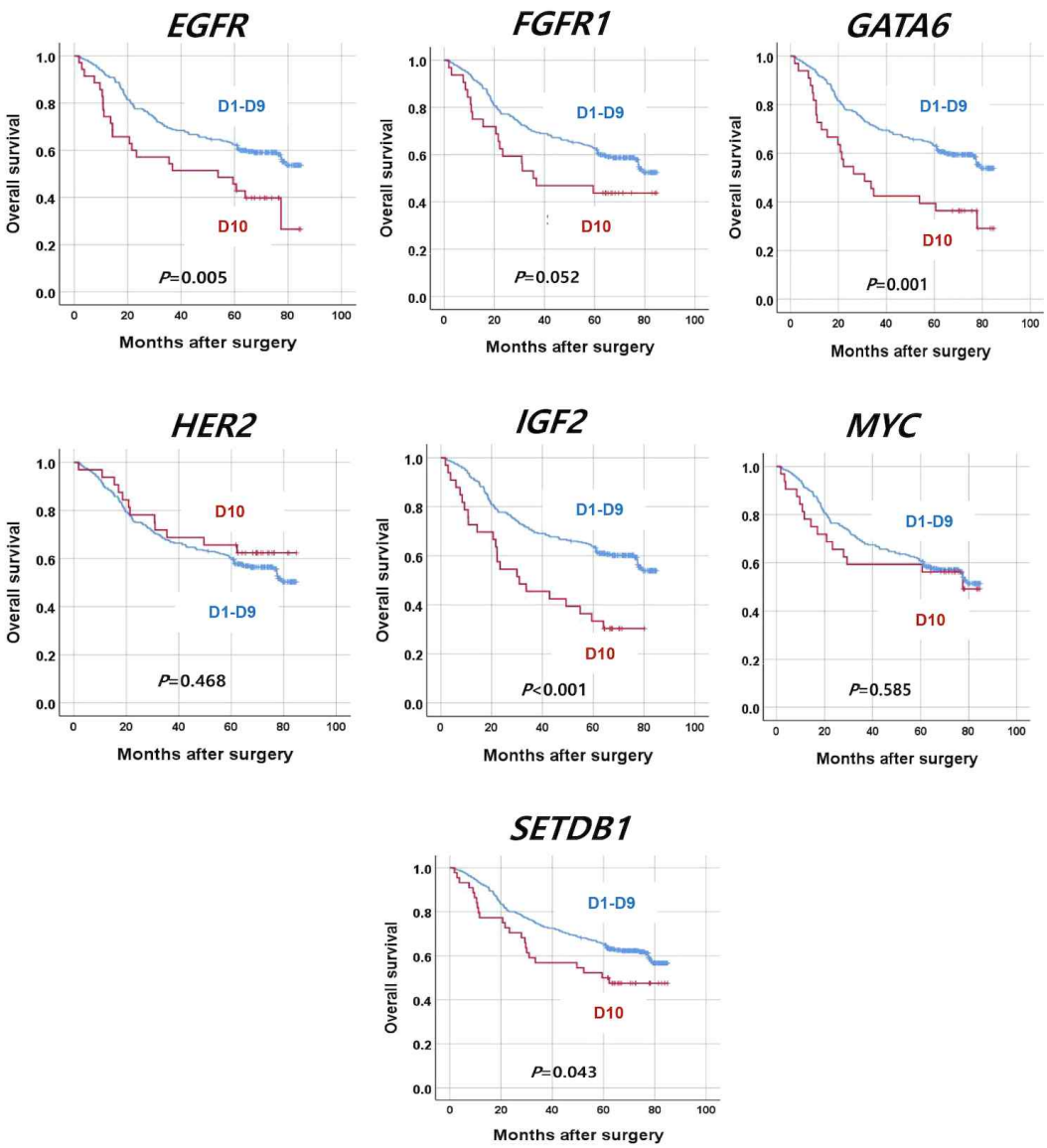

EGFR
D1-D9
D10
P=0.005
FGFR1
D1-D9
D10
P=0.052
GATA6
D1-D9
D10
P=0.001
HER2
D10
D1-D9
P=0.468
IGF2
D1-D9
D10
P<0.001
MYC
D1-D9
D10
P=0.585
SETDB1
D1-D9
D10
P=0.043
Overall survival
Months after surgery

Figure 4

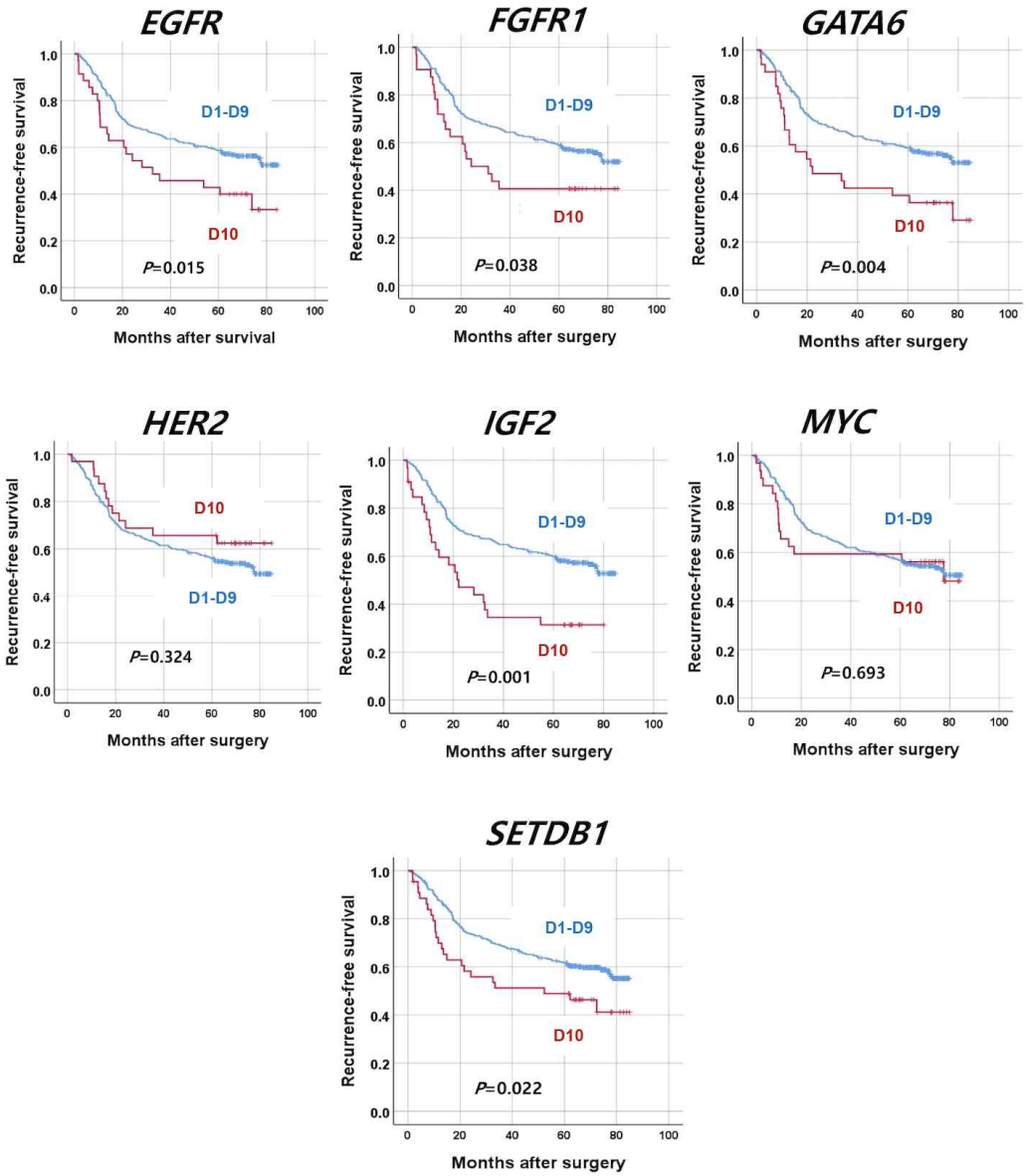

EGFR
Recurrence-free survival
D1-D9
D10
P=0.015
Months after survival
FGFR1
Recurrence-free survival
D1-D9
D10
P=0.038
Months after surgery
GATA6
Recurrence-free survival
D1-D9
D10
P=0.004
Months after surgery
HER2
Recurrence-free survival
D10
D1-D9
P=0.324
Months after surgery
IGF2
Recurrence-free survival
D1-D9
D10
P=0.001
Months after surgery
MYC
Recurrence-free survival
D1-D9
D10
P=0.693
Months after surgery
SETDB1
Recurrence-free survival
D1-D9
D10
P=0.022
Months after surgery

Table 5. Univariate and multivariate Cox regression analyses for overall survival and recurrence-free survival in individual genes

	Univariate analysis			Multivariate analysis[a]		
	HR	95% C.I.	*P*-value	HR	95% C.I.	*P*-value
Overall survival						
EGFR (subset 10 vs. subsets 1-9)	1.895	1.202-2.986	0.006	1.960	1.224-3.140	0.005
FGFR1 (subset 10 vs. subsets 1-9)	1.626	0.991-2.666	0.054	1.153	0.665-2.000	0.613
GATA6 (subset 10 vs. subsets 1-9)	2.104	1.335-3.316	0.001	1.927	1.182-3.140	0.008
HER2 (subset 10 vs. subsets 1-9)	0.804	0.445-1.452	0.469	1.025	0.546-1.922	0.939
IGF2 (subset 10 vs. subsets 1-9)	2.328	1.487-3.644	0.000	2.770	1.730-4.435	<0.001
MYC (subset 10 vs. subsets 1-9)	1.161	0.680-1.982	0.586	0.999	0.569-1.754	0.998
SETDB1 (subset 10 vs. subsets 1-9)	1.564	1.011-2.418	0.044	1.482	0.954-2.303	0.080
Recurrence-free survival						
EGFR (subset 10 vs. subsets 1-9)	1.745	1.109-2.747	0.016	1.677	1.053-2.669	0.029

FGFR1 (subset 10 vs. subsets 1-9)	1.658	1.024-2.686	0.040	1.184	0.690-2.029	0.540
GATA6 (subset 10 vs. subsets 1-9)	1.908	1.213-3.002	0.005	1.605	0.993-2.594	0.054
HER2 (subset 10 vs. subsets 1-9)	0.744	0.412-1.343	0.326	0.923	0.495-1.718	0.800
IGF2 (subset 10 vs. subsets 1-9)	2.182	1.385-3.437	0.001	2.562	1.591-4.125	<0.001
MYC (subset 10 vs. subsets 1-9)	1.113	0.653-1.899	0.693	0.955	0.549-1.660	0.870
SETDB1 (subset 10 vs. subsets 1-9)	1.638	1.069-2.509	0.023	1.574	1.021-2.429	0.040

a, Cox proportional hazards regression model, adjusted for tumor subsite, Lauren histology, lymphatic emboli, venous invasion, perineural invasion, T category, N category, and M category.

Figure 5

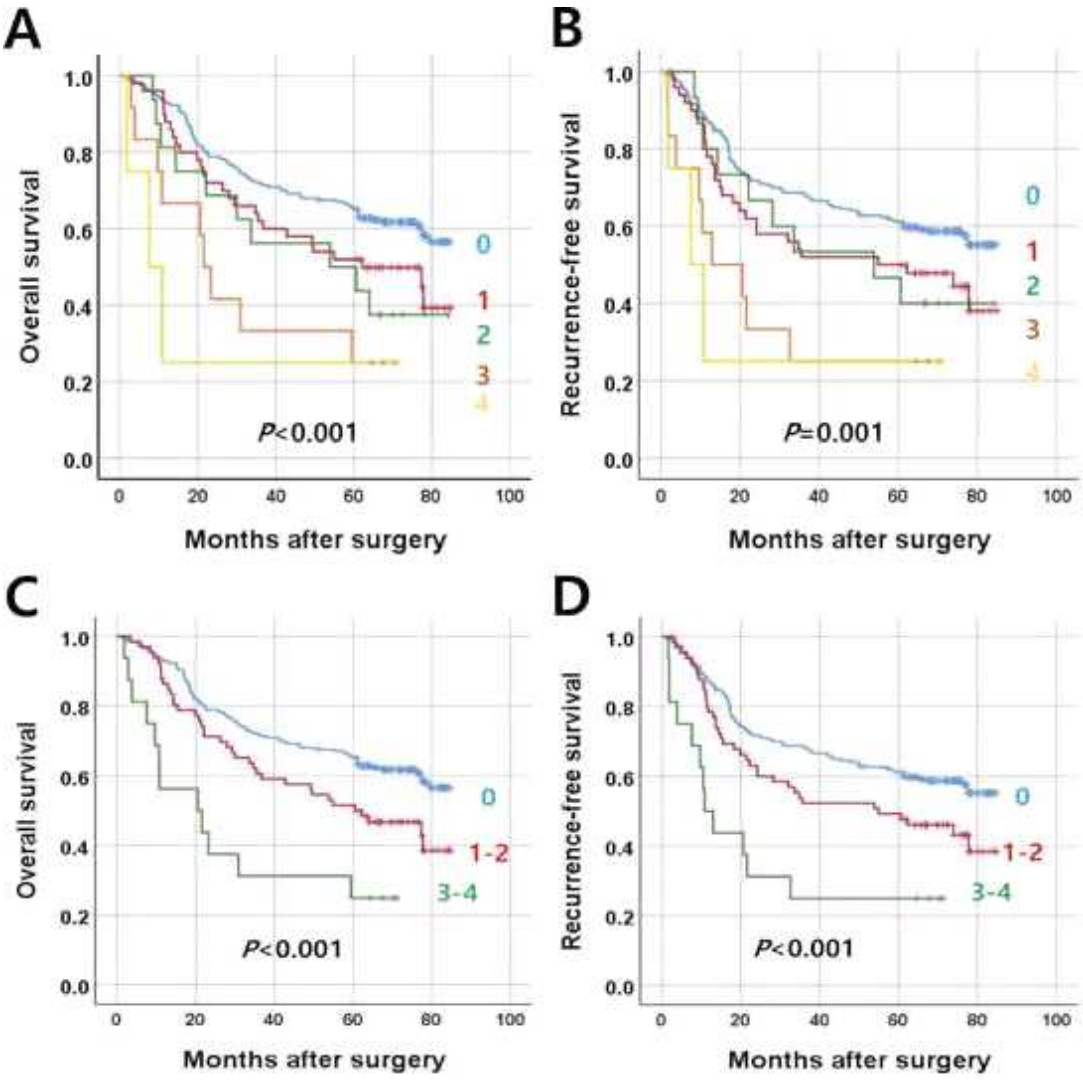
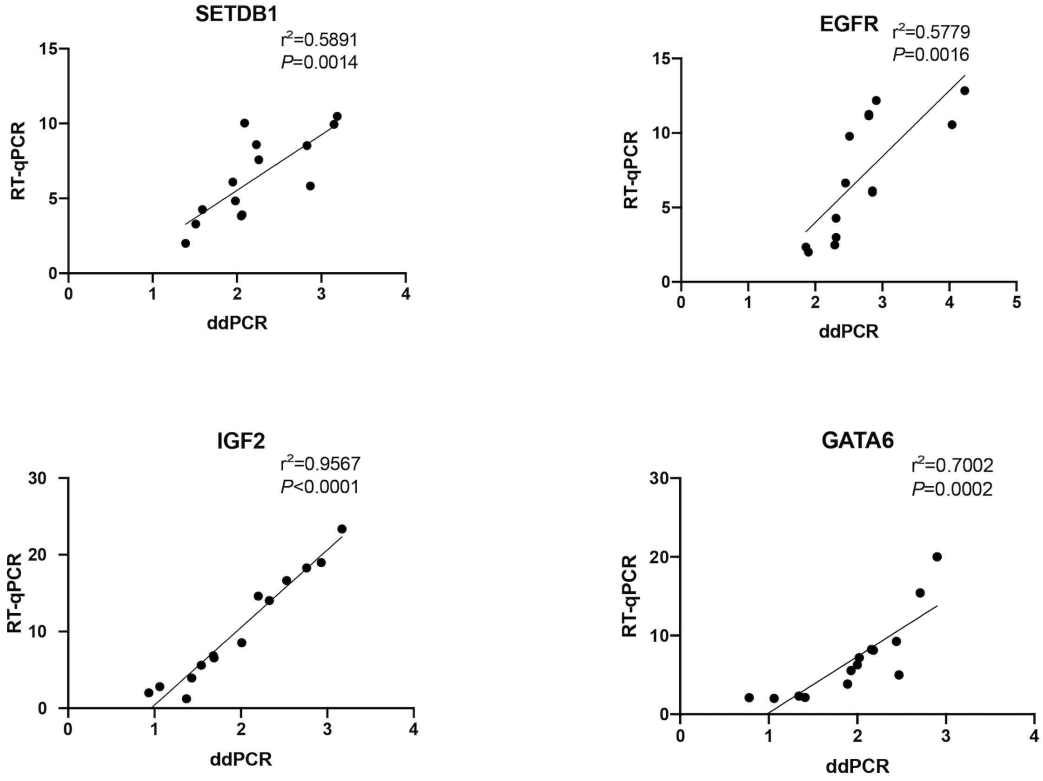

A
Overall survival
1.0
0.8
0.6
0.4
0.2
0.0
0
20
40
60
80
100
Months after surgery
P<0.001
0
1
2
3
4
B
Recurrence-free survival
P=0.001
C
Overall survival
P<0.001
0
1-2
3-4
D
Recurrence-free survival
P<0.001

Table 6. Univariate and multivariate Cox regression analyses for overall survival and recurrence-free survival in subsets

	Univariate analysis		Multivariate analysis [a]	
	HR (95% C.I.)	*P*-value	HR (95% C.I.)	*P*-value
Overall survival				
Sum score		<0.001*		<0.001*
0 (n=245)	Ref		Ref	
1, 2 (n=72)	1.584 (1.082-2.317)	0.018*	1.522 (1.022-2.266)	0.039*
3, 4 (n=16)	3.402 (1.861-6.222)	<0.001*	3.805 (2.014-7.188)	<0.001*
Recurrence-free survival				
Sum score		<0.001*		<0.001*
0 (n=245)	Ref		Ref	
1, 2 (n=72)	1.512 (1.037-2.205)	0.032*	1.329 (0.893-1.978)	0.160
3, 4 (n=16)	3.235 (1.773-5.902)	<0.001*	3.709 (1.953-7.042)	<0.001*

a, Cox proportional hazards regression model, adjusted for tumor subsite, Lauren histology, lymphatic emboli, venous invasion, perineural invasion, T category, N category, and M category (Tables 4 & 9).

Figure 6. Linear regression analysis indicated a significant correlation between qPCR and ddPCR values



Gene ratios and clinicopathological features

Table 7 summarizes the relationships between the sum scores and clinicopathological features. The sum score was higher in tumors with venous invasion than in tumors without venous invasion. The sum score tended to be higher in tumors with N3b than in tumors without nodal metastasis. However, no differences in the sum score were found in association with age, sex, tumor subsite, Lauren histology, lymphatic emboli, perineural invasion, tumor depth, distant metastasis, or molecular subtype. When TIL density was compared among GCs with different sum scores, CD3 TIL and CD8 TIL densities were highest in tumors with sum scores of 0 and lowest in tumors with sum scores of 3-4 (Table 8). When the methylation levels of repetitive DNA elements, including L1 and SAT-alpha, were compared among the three subsets, the L1 or SAT-alpha methylation level was higher in the subset with a sum score of 0 than in the subsets with a sum score of 1-2 or a sum score of 3-4 (Table 8). However, because EBV and MSI subtypes were not classified into the copy number gain type, copy number gain status needs to be analyzed for correlation with TIL densities and L1 or SAT-alpha methylation level in non-MSI/non-EBV subtype. Not only CD3 TIL and CD8 TIL densities but also L1 or SAT-alpha methylation levels were highest in GCs with a sum score of 0 and lowest in GCs with a sum score of 3-4 (Tables 9).

Table 7. Comparison of clinicopathological features according to the sum scores

		Sum scores			*P*-value
		0	1-2	3-4	
Age					0.115
<62 years	167	125 (51.0%)	38 (22.8%)	4 (25.0%)	
≥62 years	166	120 (49.0%)	34 (47.2%)	12 (75.0%)	
Sex					0.430
M	223	161 (65.7%)	49 (68.1%)	13 (81.3%)	
F	110	84 (34.3%)	23 (31.9%)	3 (18.8%)	
Site					0.191
Not involving	241	183 (74.7%)	46 (63.9%)	12 (75.0%)	
Involving cardia	92	62 (25.3%)	26 (36.1%)	4 (25.0%)	
Lauren					0.111
Intestinal	125	82 (33.5%)	35 (48.6%)	8 (50.0%)	
Diffuse	151	129 (52.7%)	27 (37.5%)	5 (31.3%)	
Mixed	43	32 (13.1%)	8 (11.1%)	3 (18.8%)	
Unclassified	4	2 (0.8%)	2 (2.8%)	0	
Lymphatic emboli					0.114
Absent	105	85 (34.7%)	16 (22.2%)	4 (25.0%)	
Present	228	160 (65.3%)	56 (77.8%)	12 (75.0%)	
Venous invasion					0.005*
Absent	239	186 (75.9%)	46 (63.9%)	7 (43.8%)	

Present	94	59 (24.1%)	26 (36.1%)	9 (56.3%)	
Perineural invasion					
Absent	138	101 (41.2%)	31 (43.1%)	6 (37.5%)	0.912
Present	195	144 (58.8%)	41 (56.9%)	10 (62.5%)	
N category					0.058
N0	91	76 (31.0%)	12 (16.7%)	3 (18.8%)	
N1	56	38 (15.5%)	15 (20.8%)	3 (18.8%)	
N2	68	50 (20.4%)	17 (23.6%)	1 (6.3%)	
N3a	72	52 (21.2%)	17 (23.6%)	3 (18.8%)	
N3b	46	29 (11.8%)	11 (15.3%)	6 (37.5%)	
T category					0.336
T2	61	51 (20.8%)	10 (13.9%)	0	
T3	121	90 (36.7%)	24 (33.3%)	7 (43.8%)	
T4a	133	91 (37.1%)	34 (47.2%)	8 (50.0%)	
T4b	18	13 (5.3%)	4 (5.6%)	1 (6.3%)	
M category					
M0	284	211 (86.1%)	57 (79.2%)	16 (100.0%)	0.080
M1	49	34 (13.9%)	15 (20.8%)	0	
Molecular subtype					0.161
MSS/EBV-	265	189 (77.1%)	60 (83.3%)	16 (100.0%)	

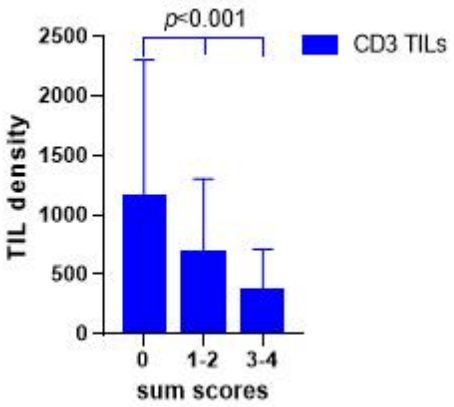
MSI-H	42	33 (13.5%)	9 (12.5%)	0	
EBV+	26	23 (9.4%)	3 (4.2%)	0	

Abbreviations: MSS, microsatellite-stable; EBV-, EBV-negative; MSI-H, high level of microsatellite instability; EBV+, EBV-positive

Table 8. Density of tumor-infiltrating lymphocytes and methylation levels of L1 and SAT-alpha in gastric carcinomas according to the sum scores

		Sum scores			*P*-value	*P*-value
		0	1-2	3-4	ANOVA	Kruskal-Wallis
		(n=245)	(n=72)	(n=16)		
TIL density						
CD3 TILs at tumor center	Mean (SD)	1147.0 (1155.54)	675.2 (628.28)	358.1 (353.31)	<0.001	<0.001
CD8 TILs at tumor center	Mean (SD)	887.5 (1050.00)	437.0 (399.30)	278.3 (281.98)	<0.001	<0.001

Abbreviations: TIL, tumor infiltrating lymphocyte



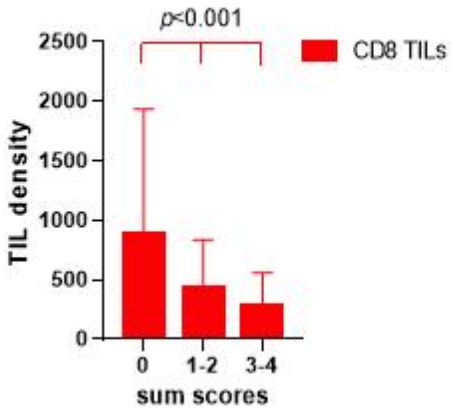


Table 9. Methylation levels of L1 and SAT-alpha in non-MSI/non-EBV gastric carcinomas according to the sum scores

		Sum scores			*P*-value	*P*-value
		0	1-2	3-4	ANOVA	Kruskal-Wallis
		(n=189)	(n=60)	(n=16)		
Methylation level of repetitive DNA elements						
L1	Mean (SD)	71.5% (7.35)	67.4% (8.89)	67.0% (11.57)	0.001	0.005
SAT-alpha	Mean (SD)	62.8% (8.72)	58.3% (9.39)	58.4% (10.59)	0.002	0.002

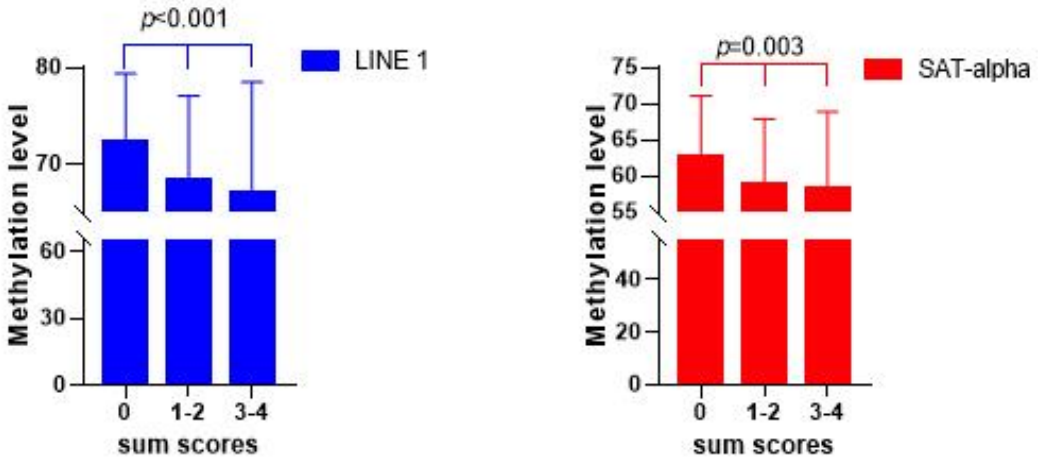


Table 10. Univariate and multivariate Cox regression analysis for overall survival

	Univariate analysis		Multivariate analysis	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Tumor subsite (involving cardia vs. not involving cardia)	2.242 (1.603-3.135)	<0.001	1.677 (1.053-2.669)	0.029
Lauren histology		0.005		0.114
Intestinal type	Ref			
Diffuse type	1.439 (1.004-2.063)	0.047	1.142 (0.775-1.683)	0.503
Mixed type	0.841 (0.464-1.526)	0.570	0.916 (0.487-1.723)	0.787
Unclassified	5.466 (1.700-17.575)	0.004	5.241 (1.316-20.867)	0.019
Lymphatic emboli (present vs. absent)	3.254 (2.081-5.088)	<0.001	1.538 (0.912-2.594)	0.107
Venous invasion (present vs. absent)	1.859 (1.321-2.615)	<0.001	0.985 (0.673-1.440)	0.936
Perineural invasion (present vs. absent)	1.601 (1.131-2.266)	0.006	1.011 (0660-1.551)	0.959
CD3 TIL density (high vs. low)	0.500 (0.358-0.699)	<0.001	1.105 (0.677-1.804)	0.690
CD8 TIL density (high vs. low)	0.463 (0.329-0.650)	<0.001	0.576 (0.402-0.826)	0.003
T category		<0.001		0.001
T2	Ref			
T3	1.184 (0.657-2.134)	0.574	0.639 (0.341-1.195)	0.161
T4	3.580 (2.086-6.144)	<0.001	1.460 (0.809-2.635)	0.209
T5	6.170 (2.956-12.880)	<0.001	1.805 (0.803-4.055)	0.153
N category		<0.001		<0.001
N0	Ref			
N1	1.841 (0.958-3.538)	0.067	1.310 (0.665-2.582)	0.435
N2	2.710 (1.521-4.829)	0.001	1.284 (0.670-2.460)	0.451
N3a	4.080 (2.331-7.141)	<0.001	2.349 (1.232-4.480)	0.009
N3b	9.973 (5.623-17.689)	<0.001	3.989 (2.022-7.870)	<0.001
M category (M1 vs. M0)	5.535 (3.824-8.012)	<0.001	3.128 (2.057-4.756)	<0.001

Table 11. Univariate and multivariate Cox regression analysis for recurrence-free survival

	Univariate analysis		Multivariate analysis		
	HR (95% CI)	*P*-value	HR	95% CI	*P*-value
Tumor subsite (involving cardia vs. not involving cardia)	2.171 (1.561-3.021)	<0.001	1.812 (1.259-2.608)		0.001
Lauren histology		0.011			0.197
Intestinal type	Ref				
Diffuse type	1.468 (1.028-2.097)	0.035	1.092 (0.742-1.606)		0.656
Mixed type	0.951 (0.540-1.674)	0.861	0.899 (0.491-1.646)		0.731
Unclassified	4.867 (1.514-15.646)	0.008	4.284 (1.082-16.970)		0.038
Lymphatic emboli (present vs. absent)	3.074 (1.999-4.728)	<0.001	1.363 (0.825-2.254)		0.227
Venous invasion (present vs. absent)	1.701 (1.215-2.379)	0.002	0.887 (0.610-1.291)		0.531
Perineural invasion (present vs. absent)	1.644 (1.168-2.314)	0.004	1.013 (0.669-1.533)		0.952
CD3 TIL density (high vs. low)	0.465 (0.334-0.646)	<0.001	1.025 (0.622-1.688)		0.924
CD8 TIL density (high vs. low)	0.435 (0.311-0.608)	<0.001	0.524 (0.368-0.744)		<0.001
T category		<0.001			<0.001
T1	Ref		Ref		
T2	1.380 (0.760-2.503)	0.290	0.733 (0.392-1.371)		0.331
T3	3.985 (2.293-6.923)	<0.001	1.792 (0.987-3.252)		0.055
T4	5.554 (2.633-11.715)	<0.001	1.530 (0.675-3.469)		0.309
N category		<0.001			<0.001
N0	Ref		Ref		
N1	1.700 (0.892-3.239)	0.107	1.381 (0.713-2.675)		0.338
N2	2.747 (1.557-4.848)	<0.001	1.739 (0.964-3.138)		0.066

N3a	4.396 (2.559-7.555)	<0.001	3.248 (1.834-5.753)	<0.001
N3b	9.071 (5.172-15.909)	<0.001	4.386 (2.401-8.013)	<0.001
M category (M1 vs. M0)	4.239 (2.940-6.111)	<0.001	2.179 (1.444-3.288)	<0.001

Table 12. Mean comparison (median value).
P-values were obtained with Student's t test or ANOVA, and if the p-values were less than 0.1, nonparametric tests were conducted with the Mann-Whitney test or Kruskal-Wallis test.

	n	EGFR ratio	*P*-value	GATA6 ratio	*P*-value	IGF2 ratio	*P*-value	SETDB1 ratio	*P*-value
Age			0.900		0.936		0.069		0.221
<62 years	167	2.73		1.23		0.79	0.194	1.65	
≥62 years	167	2.65		1.24		0.89		1.72	
Sex			0.054		0.813		0.060		0.344
M	224	2.28	0.167	1.24		0.88	0.093	1.71	
F	110	3.57		1.22		0.77		1.64	
Site			0.092		0.313		0.851		0.456
Not involving	241	2.37	0.767	1.21		0.84		1.70	
Involving cardia	93	3.52		1.30		0.85		1.65	
Lauren			<0.001		0.722		0.126		0.014
Intestinal	125	2.29	0.181	1.28		0.89		1.81	0.179
Diffuse	162	2.51		1.22		0.80		1.63	
Mixed	43	3.26		1.14		0.81		1.53	
Unclassified	4	16.43		1.31		1.29		1.77	
Lymphatic emboli			0.329		0.192		0.662		0.112
Absent	105	2.25		1.16		0.86		1.62	
Present	229	2.89		1.27		0.83		1.72	
Venous invasion			0.787		0.107		0.683		0.001
Absent	239	2.64		1.19		0.83		1.62	0.014
Present	95	2.82		1.34		0.86		1.87	
Perineural invasion			0.239		0.398		0.071		0.344
Absent	138	3.12		1.19		0.90	0.104	1.72	
Present	196	2.39		1.26		0.80		1.65	
N stage			0.086		0.116		0.025		0.008
N0	91	2.07	0.088	1.12		0.80	0.073	1.54	<0.001
N1 (1-2)	56	2.26		1.20		0.89		1.69	
N2 (3-6)	68	2.26		1.18		0.74		1.77	
N3a (7-15)	72	2.89		1.34		0.82		1.69	
N3b (>15)	47	4.71		1.42		1.03		1.90	
T stage			0.072		0.145		0.673		0.034
T2	61	1.83	0.154	1.09		0.79		1.54	<0.010
T3	121	2.18		1.21		0.83		1.68	
T4a	134	3.22		1.34		0.88		1.78	
T4b	18	5.08		1.14		0.82		1.78	
M stage			0.011		0.292		0.419		0.931

M0	285	2.37	0.698	1.22		0.85		1.69	
M1	49	4.55		1.34		0.79		1.68	
Molecular subtype			0.456		0.001		0.238		<0.001
MSS/EBV-	265	2.88		1.31	<0.001	0.86		1.75	<0.001
MSI-H	43	2.14		0.99		0.73		1.42	
EBV+	26	1.68		0.87		0.81		1.42	

Table 13. Multivariate Cox regression analysis for overall survival and recurrence-free survival

	Overall survival[a]		Recurrence-free survival[a]	
	HR (95% CI)	*P*-value	HR (95% CI)	*P*-value
Sum score		0.001		0.001
0 (n=237)	Ref		Ref	
1, 2(n=66)	1.274 (0.849-1.911)	0.243	1.065 (0.709-1.600)	0.762
3, 4(n=16)	3.261 (1.728-6.157)	<0.001	3.328 (1.762-6.287)	<0.001

a, Cox proportional hazards regression model, adjusted for tumor subsite, Lauren histology, lymphatic emboli, venous invasion, perineural invasion, CD3 TIL density, CD8 TIL density, T category, N category, M category, L1 methylation level, and SAT-alpha methylation level.

Discussion

In the present study, we analyzed the gene ratios of 7 genes, including *MYC*, *EGFR*, *ERBB2*, *FGFR1*, *GATA6*, *IGF2*, and *SETDB1*, in advanced GC patients using ddPCR. To determine the cut-off value of the gene ratios with prognostic utility, we partitioned the GC patients into 10 subsets according to the gene ratios and then performed survival analysis, which revealed that subset 10 with the highest gene ratios for *EGFR*, *FGFR1*, *GATA6*, *IGF2*, and *SETDB1* was associated with worse clinical outcomes in patients with GC. Of these five genes, *FGFR1* was not found to be an independent prognostic parameter in multivariate analysis. To assess the additive effect of copy number gains in the four genes (*EGFR*, *GATA6*, *IGF2*, and *SETDB1*), we calculated the sum score; in other words, we counted, in each case, the number of genes for which the gene ratio belonged to subset 10. According to survival curves, the GC cases could be grouped into GCs with a sum score of 0, a sum score of 1 or 2, and a sum score of 3 or 4. GCs with sum scores of 3 or 4 were found to be associated with worse survival in GC patients (OS, hazard ratio of 3.320, 95% CI = 1.756 - 6.278, $P < 0.001$; RFS, hazard ratio of 3.285, 95% CI = 1.736 - 6.217, $P < 0.001$) in the multivariate analysis, regardless of tumor subsite, Lauren histology, venous invasion, lymphatic invasion, perineural invasion, and T, N, and M categories.

Our study demonstrated that the sum score was inversely

associated with the CD3 or CD8 TIL density, which indicates that the copy number gain of multiple protooncogenes is associated with decreased infiltration of CD3 or CD8 TIL density. Our finding is in line with findings of recent studies in which the amplification of *MYC*, *NOTCH2*, and *FGFR1* was inversely associated with the expression of genes related to cytotoxic T cell function in pancreatic ductal adenocarcinoma [7,8]. Not only the amplification but also the SNV mutations of protooncogenes have been demonstrated to be associated with decreased cytotoxic T cell infiltration in tumor areas. For lung cancers, *EGFR* mutations have been linked with decreased cytotoxic T cell infiltration [9,10], whereas for colorectal cancers, *KRAS* mutations have been associated with increased marrow-derived suppressor cell infiltration and the subsequent decreased infiltration of cytotoxic T cells [11,12]. Based on the association between the copy number gain of multiple protooncogenes and the decreased infiltration of CD3 and CD8 TILs, it might be questioned whether the prognostic value of the sum score is bestowed by the decreased density of TILs. However, in the multivariate analysis, both the sum score and CD8 TILs were found to be independent prognostic parameters for both OS and RFS (Tables 10 &11).

When we correlated the sum scores with clinicopathological features, GCs with high sum scores showed an association with venous invasion but did not show associations with lymphatic emboli and nodal stage. At present, the reason why GCs with high sum scores are more likely to invade veins rather than lymphatic vessels

is unclear. Whether GC cells intravasate into either blood or lymphatic vessels might be related to several factors, including the physical differences between lymphatic and blood vessels, the more favorable conditions for tumor cell survival in lymphatic vessels because of the low-shear system of fluid transport [13], and the active molecular mechanisms attracting malignant cells more towards blood or lymphatic vessels [13,14]. In the present study, when we correlated the copy number gain of the four individual genes with venous invasion, we found that the *SETDB1* gene ratio was significantly higher in GCs with venous invasion than in GCs with no venous invasion (Table 12). The *SETDB1* (KMT1E) gene encodes a histone methyltransferase that methylates Lys-9 of histone H3 up to trimethylation. The *SETDB1* gene is located on chromosome 1q21, which shows copy number gains in several tissue types of human cancers, including breast cancer [15], melanoma [16], lung cancer [17,18], and liver cancer [19]. An oncogenic role of SETDB1 has been demonstrated in lung cancer and prostate cancer, in which SETDB1 is involved in the positive stimulation of WNT signaling [20,21]. The downregulation of the *SETDB1* gene has been found to decrease the migration and invasion of prostate cancer cells and inhibit the growth of prostate cancer cells by inducing G0/G1 cell cycle arrest [22]. Significant relationships between higher SETDB1 protein expression and shorter survival times have been demonstrated in patients with lung cancer [17,23], liver cancer [19], and colon cancer [24]. Although the copy number gain in GC can be referred to in the COSMIC and TCGA databases,

little information is available in the literature regarding relationships between the higher expression of SETDB1 protein or the copy-number gain of *SETDB1* and the clinicopathological features of GC.

Tumoral L1 hypomethylation and SAT-alpha hypomethylation have been shown to be associated with shortened survival in patients with advanced GC [4]. Tumoral L1 and SAT-alpha hypomethylation occurs in the background of diffuse genomic hypomethylation, which is closely associated with chromosomal instability. Thus, the copy number gain of multiple genes is expected to occur in GCs with L1 hypomethylation or SAT-alpha hypomethylation. In a previous study, we determined L1 and SAT-alpha methylation statuses using pyrosequencing methylation assays, so we used the previous data of L1 and SAT-alpha methylation levels and compared L1 and SAT-alpha methylation levels among different sum scores, which revealed a significant difference between GCs with sum scores of 0 and GCs with sum scores of 1-2 or sum scores of 3-4 (Table 8). To identify whether the prognostic significance of the sum score could be affected by L1 and SAT-alpha methylation statuses, we performed multivariate analysis with the inclusion of L1 and SAT-alpha methylation statuses and other prognostic variables that were found to be statistically significant in the univariate survival analysis (Table 13). The sum score was found to be an independent prognostic parameter for both OS and RFS.

In conclusion, copy number gains in three or four of the *EGFR*,

GATA6, *IGF2*, and *SETDB1* genes were found to be associated with venous invasion, decreased TIL densities, decreased levels of DNA methylation in L1 or SAT-alpha, and shortened rates of both OS and RFS. A high sum score was found to be an independent prognostic parameter associated with poor prognosis in patients with advanced GC. An independent study is needed to validate the prognostic value of high sum scores in the four genes.

References

1. Sohn, B. H. *et al.* Clinical Significance of Four Molecular Subtypes of Gastric Cancer Identified by The Cancer Genome Atlas Project. *Clin Cancer Res*, doi:10.1158/1078-0432.CCR-16-2211 (2017).

2. Cancer Genome Atlas Research, N. Comprehensive molecular characterization of gastric adenocarcinoma. *Nature* **513**, 202-209, doi:10.1038/nature13480 (2014).

3. Song, Y. S. *et al.* Methylation status of long interspersed element-1 in advanced gastric cancer and its prognostic implication. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* **19**, 98-106, doi:10.1007/s10120-015-0463-6 (2016).

4. Kim, Y. *et al.* Combinatory low methylation statuses of SAT-alpha and L1 are associated with shortened survival time in patients with advanced gastric cancer. *Gastric cancer : official journal of the International Gastric Cancer Association and the Japanese Gastric Cancer Association* **22**, 37-47, doi:10.1007/s10120-018-0852-8 (2019).

5. Kim, Y. *et al.* Combination of L1 methylation and tumor-infiltrating lymphocytes as prognostic marker in advanced gastric cancer. *Gastric Cancer*, doi:10.1007/s10120-019-01025-8 (2019).

6. Whale, A. S. *et al.* Comparison of microfluidic digital PCR and conventional quantitative PCR for measuring copy number variation. *Nucleic Acids Res* **40**, e82, doi:10.1093/nar/gks203 (2012).

7. Bailey, P. *et al.* Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* **531**, 47-52, doi:10.1038/nature16965 (2016).

8. Balli, D., Rech, A. J., Stanger, B. Z. & Vonderheide, R. H. Immune Cytolytic Activity Stratifies Molecular Subsets of Human Pancreatic Cancer. *Clin Cancer Res* **23**,3129-3138, doi:10.1158/1078-0432.CCR-16-2128 (2017).

9. Concha-Benavente, F. & Ferris, R. L. Reversing EGFR Mediated Immunoescape by Targeted Monoclonal Antibody Therapy. *Front Pharmacol* **8**, 332, doi:10.3389/fphar.2017.00332 (2017).

10 Busch, S. E. *et al.* Lung Cancer Subtypes Generate Unique Immune Responses. *J Immunol* **197**, 4493-4503, doi:10.4049/jimmunol.1600576 (2016).

11. Liao, W. *et al.* KRAS-IRF2 Axis Drives Immune Suppression and Immune Therapy Resistance in Colorectal Cancer. *Cancer Cell* **35**, 559-572 e557, doi:10.1016/j.ccell.2019.02.008 (2019).

12. Lal, N. *et al.* KRAS Mutation and Consensus Molecular Subtypes 2 and 3 Are Independently Associated with Reduced Immune Infiltration and Reactivity in Colorectal Cancer. *Clin Cancer Res* **24**, 224-233, doi:10.1158/1078-0432.CCR-17-1090 (2018).

13. Wong, S. Y. & Hynes, R. O. Lymphatic or hematogenous dissemination: how does a metastatic tumor cell decide? *Cell Cycle* **5**, 812-817, doi:10.4161/cc.5.8.2646 (2006).

14. Aleskandarany, M. A., Sonbul, S. N., Mukherjee, A. & Rakha, E. A. Molecular Mechanisms Underlying Lymphovascular Invasion in Invasive Breast Cancer. *Pathobiology* **82**, 113-123, doi:10.1159/000433583 (2015).

15. Liu, L., Kimball, S., Liu, H., Holowatyj, A. & Yang, Z. Q. Genetic

alterations of histone lysine methyltransferases and their significance in breast cancer. *Oncotarget* **6**, 2466-2482, doi:10.18632/oncotarget.2967 (2015).

16. Ceol, C. J. *et al.* The histone methyltransferase SETDB1 is recurrently amplified in melanoma and accelerates its onset. *Nature* 471, 513-517, doi:10.1038/nature09806 (2011).

17. Inoue, Y. *et al.* Clinicopathological and Survival Analysis of Japanese Patients with Resected Non-Small-Cell Lung Cancer Harboring NKX2-1, SETDB1, MET, HER2, SOX2, FGFR1, or PIK3CA Gene Amplification. *J Thorac Oncol* **10**, 1590-1600, doi:10.1097/JTO.0000000000000685 (2015).

18. Rodriguez-Paredes, M. *et al.* Gene amplification of the histone methyltransferase SETDB1 contributes to human lung tumorigenesis. *Oncogene* **33**, 2807-2813, doi:10.1038/onc.2013.239 (2014).

19. Wong, C. M. *et al.* Up-regulation of histone methyltransferase SETDB1 by multiple mechanisms in hepatocellular carcinoma promotes cancer metastasis. *Hepatology* **63**, 474-487, doi:10.1002/hep.28304 (2016).

20. Sun, Q. Y. *et al.* SETDB1 accelerates tumourigenesis by regulating the WNT signalling pathway. *J Pathol* **235**, 559-570, doi:10.1002/path.4482 (2015).

21. Li, Q. *et al.* Quantitative proteomic study of human prostate cancer cells with different metastatic potentials. *Int J Oncol* **48**, 1437-1446, doi:10.3892/ijo.2016.3378 (2016).

22. Sun, Y. *et al.* Histone methyltransferase SETDB1 is required for

prostate cancer cell proliferation, migration and invasion. *Asian J Androl* **16**, 319-324, doi:10.4103/1008-682X.122812 (2014).

23. Lafuente-Sanchis, A. *et al.* Prognostic value of ERCC1, RRM1, BRCA1 and SETDB1 in early stage of non-small cell lung cancer. *Clin Transl Oncol* **18**, 798-804, doi:10.1007/s12094-015-1440-6 (2016).

24. Ho, Y. J. *et al.* Significance of histone methyltransferase SETDB1 expression in colon adenocarcinoma. *APMIS* **125**, 985-995, doi:10.1111/apm.12745 (2017).

Abstract in Korean

Protooncogenes의 유전자 카피수 (copy number, CN) 증가는 위암 (gastric carcinoma, GC)에서 빈번하게 발견되지만 그 예후적 의미는 여전히 파악하기 어렵다. 본 연구는 위암종에서 변화를 보이는 여러 protooncogenes에서 다중 유전자 CN의 증가를 통해 위암의 예후를 포함한 임상 병리학적 특징을 특성화하고자 하였다. 위암 환자 333명의 EGFR, GATA6, IGF2 및 SETDB1 유전자의 비율로 CN 변화를 평가하기 위해 droplet digital PCR (ddPCR)을 사용하여 분석하였다. CN 획득 유전자가 3개 이상인 위암 환자는 16명(4.8%)이며 이는 CN 획득 유전자가 2개 이하인 위암 환자에 비해 빈번한 정맥 침윤 (venous invasion)과 낮은 tumor-infiltrating lymphocytes (TILs) 밀도, 그리고 낮은 L1 또는 SAT-alpha의 메틸화 수준을 보였다. CN 획득 유전자가 3개 이상인 위암 환자에서는 미소부수체 (microsatellite) 불안정성이 높은 종양이나 엡스타인-바-바이러스 (Epstein-Barr virus) 양성 종양이 발견되지 않았다. 다변량 생존 분석에서 CN 획득 유전자가 3개 이상은 생존율 및 재발 생존 가능성이 가장 나빴다. 본 연구를 통해, 암유전자에 대한 ddPCR을 기반으로 한 다중 CN 획득 검출이 불량한 예후를 가진 위암환자를 식별하는데 도움이 되고 있음을 제시하였고, 추가적인 연구를 통해 유용성을 검증할 필요가 있다