ORIGINAL ARTICLE

Early Gastric Cancer: identification of molecular markers able to distinguish submucosa-penetrating lesions with different prognosis

Chiara Molinari[1] · Gianluca Tedaldi[1] · Francesca Rebuzzi[1] · Paolo Morgagni[2] · Laura Capelli[1] ·
Sara Ravaioli[1] · Maria Maddalena Tumedei[1] · Emanuela Scarpi[3] · Anna Tomezzoli[4] · Riccardo Bernasconi[4] ·
Maria Raffaella Ambrosio[5,6] · Alessia D'Ignazio[7] · Leonardo Solaini[2,8] · Francesco Limarzi[1] · Giorgio Ercolani[2] ·
Giovanni Martinelli[9] · Paola Ulivi[1] · Luca Saragoni[10]

Abstract

Background Early Gastric Cancer (EGC) reaches 25% of the gastric cancers surgically treated in some areas of Northeastern Italy and is usually characterized by a good prognosis. However, among EGCs classified according to Kodama's criteria, Pen A subgroup is characterized by extensive submucosal invasion, lymph node metastases and worse prognosis, whereas Pen B subgroup by better prognosis. The aim of the study was to characterize the differences between Pen A, Pen B and locally advanced gastric cancer (T3N0) in order to identify biomarkers involved in aggressiveness and clinical outcome.

Methods We selected 33 Pen A, 34 Pen B and 20 T3N0 tumors and performed immunohistochemistry of mucins, copy number variation analysis of a gene panel, microsatellite instability (MSI), *TP53* mutation and loss of heterozygosity (LOH) analyses.

Results Pen A subgroup was characterized by MUC6 overexpression ($p = 0.021$). Otherwise, the Pen B subgroup was significantly associated with the amplification of *GATA6* gene ($p = 0.002$). The higher percentage of MSI tumors was observed in T3N0 group ($p = 0.002$), but no significant differences between EGC types were found. Finally, *TP53* gene analysis showed that 32.8% of Pen tumors have a mutation in exons 5–8 and 50.0% presented LOH. Co-occurrence of *TP53* mutation and LOH mainly characterized Pen A tumors ($p = 0.022$).

Conclusions Our analyses revealed that clinico-pathological parameters, microsatellite status and frequency of *TP53* mutations do not seem to distinguish Pen subgroups. Conversely, the amplification of *GATA6* was associated with Pen B, as well as the overexpression of MUC6 and the *TP53*$^{mut/LOH}$ significantly characterized Pen A.

Keywords Early Gastric Cancer · Submucosa-penetrating tumors · MUC6 · *GATA6* · *TP53*

✉ Gianluca Tedaldi
gianluca.tedaldi@irst.emr.it

1 Biosciences Laboratory, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) IRCCS, via P. Maroncelli 40, 47014 Meldola, FC, Italy

2 Department of Surgery, Morgagni-Pierantoni Hospital, Forlì, Italy

3 Biostatistics and Clinical Trials Unit, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) IRCCS, Meldola, Italy

4 Department of Pathology, University of Verona, Verona, Italy

5 Pathology Unit, University of Siena, Siena, Italy

6 Pathology Unit, Azienda USL Toscana Nord-Ovest, Pisa, Italy

7 Surgery Unit, University of Siena, Siena, Italy

8 Department of Medical and Surgical Sciences, University of Bologna, Bologna, Italy

9 Department of Medical Oncology, Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST) IRCCS, Meldola, Italy

10 Pathology Unit, Morgagni-Pierantoni Hospital, Forlì, Italy

Introduction

Gastric cancer (GC) is one of the most frequent tumors worldwide, being the fifth for incidence and the third for mortality, with about 1,033,701 new cases and 782,685 deaths in 2018 [1].

The term "early gastric cancer" (EGC), as explained by the Japanese Gastric Cancer Association, identifies carcinomas limited to gastric mucosa and/or submucosa, regardless of lymph node status [2, 3]. In Romagna, an area of Northeastern Italy, EGCs account for about 25% of all GCs surgically treated [4]. This percentage of "early lesions" is considered very high for Western countries' standards, but still not comparable with that of Eastern countries, where EGCs represent more than 50% of all tumors [5–7].

The definition of EGC has been debated over the years, given that, despite a generally good prognosis (5-year survival > 90%), tumors with different outcomes belong to this group. Numerous studies have focused on key parameters that could be associated with adverse prognosis, risk of lymph node metastases or treatment failure in EGC [4, 8–12], but no consensus has been reached. Of note, the number of positive lymph nodes and the infiltrative growth pattern according to Kodama's classification [13] have been shown to be independent prognostic factors in 530 EGC patients [14]. In this classification, tumors lesser than 4 cm and invading deeply the submucosa are defined submucosa-penetrating EGCs (Pen) and divided in Pen A, characterized by elevated lesions, extensive submucosal invasion and nodular masses, high incidence of vascular invasion and lymph node metastases, and Pen B, associated with depressed lesions, fenestrated invasion of submucosa and low incidence of vascular invasion and lymph node metastasis [13]. More importantly, Pen A and Pen B show different prognosis, being Pen A type associated with worse prognosis and thus a candidate for a surgical treatment similar to advanced tumors [4, 10, 11, 14–17].

Independently from clinico-pathological characteristics, other parameters could be implicated in the different prognosis of EGC subtypes, but to date the molecular characteristics of each EGC subgroup have not been deeply explored yet. At the beginning of eighties, Inokuchi found that a different cell nuclear DNA distribution pattern correlates with malignancy, being Pen A characterized by aneuploid and high-ploidy DNA range, similarly to advanced carcinoma [18]. On the other hand, recent studies on tumor microenvironment and the genomic features suggest that cellular and molecular markers may be useful in determining the aggressiveness of EGC [19, 20].

In this context, the purpose of this study was to investigate the molecular characteristics of ECG. We focused the analyses on two different types of Pen tumors, Pen A and Pen B, known to have different prognosis, with the final aim of finding biomarkers useful in the treatment decision making for these patients.

Methods

Study design and case series

This was a retrospective multicenter study conducted at the Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori on cases recruited by three member centers of the Italian Gastric Cancer Research Group (GIRCG): Morgagni-Pierantoni Hospital of Forlì, University of Verona and University of Siena (Italy). The case series included GC patients who underwent radical surgery with histological diagnosis of Pen A ($n = 33$) and Pen B ($n = 34$), characterized by an invasion depth > 500 μm, according to the definition of submucosa-penetrating EGCs (Pen) [2]. Patients were selected from a case series consecutively enrolled from 1990 to 2014, based on the availability of tissue samples. Patients were matched for age, gender and lymph nodes status. Moreover, a group of T3N0 tumors ($n = 20$), matched with EGC for gender and age, was analyzed. All the cases analyzed have been evaluated blinded by two different pathologists. Histotypes were defined according to Lauren's classification [21] and the lymphovascular invasion has been evaluated with hematoxylin and eosin staining, except for doubtful cases evaluated by immunohistochemistry. The study was approved by the Local Ethics Committee (C.E.ROM. protocol number: IRSTB044) and written informed consent was obtained from all patients.

Immunohistochemistry

Immunohistochemistry (IHC) was performed on 4-μM-thick sections of surgical specimens fixed in formalin and embedded in paraffin (FFPE) using VENTANA Bench-Mark Ultra platform and Optiview DAB Detection kit (Ventana Medical Systems Inc, Tucson, AZ, USA). Sections were stained with prediluted mouse monoclonal antibodies against MUC6 (Cell Marque Cat# 293 M-95, RRID:AB_1160664), MUC2 (Cell Marque Cat# 291 M-14, RRID:AB_1160632) and MUC5AC (Cell Marque Cat# 292 M-98, RRID:AB_1160659) according to manufacturer's instructions. Epidermal Growth Factor Receptor (EGFR) detections were performed by staining the sections with mouse monoclonal antibody against EGFR (Genemed Cat# 61-0027, RRID: AB_11204237) diluted 1:100 in Antibody diluent (Ventana Medical Systems). For all biomarkers, slides were finally automatically counterstained with Hematoxylin II (Ventana Medical Systems). Stained sections were

evaluated in blind by an expert pathologist. All images were captured using Axiocam ERc5s digital camera on Axioscope microscope (Carl Zeiss, Oberkochen, Germany). Expression values of biomarkers were considered as categorical variables, using the following cuts off: < 1%: negative; 1–49%: weak positive; ≥ 50% positive. Only the diffusely positive cases (≥ 50%) have been considered as positive for the purpose of our study.

DNA extraction

The DNA extraction from FFPE tissues was performed using QIAamp DNA FFPE Tissue kit (Qiagen, Hilden, Germany). The DNA was quantified with Qubit® dsDNA BR Assay on Qubit 3.0 fluorometer (Thermo Fisher Scientific) and diluted for the subsequent molecular analyses.

Microsatellite instability analysis

The microsatellite instability (MSI) analysis was performed using 20 ng of DNA and the Genequality CC-MSI kit (AB Analitica, Padova, Italy), which amplifies 10 microsatellite markers (Supplementary Table S1). The fragment analysis was carried out using the automated sequencer 3130 Genetic Analyzer and (Applied Biosystems, Foster City, CA, USA). The results were visualized with GeneMapper software (Applied Biosystems). According to the guidelines [22], samples with at least four unstable microsatellites were classified as tumors with high microsatellite instability (MSI-H), while cases with 1–3 unstable microsatellites as tumors with low microsatellite instability (MSI-L). The others were classified as tumors with microsatellite stability (MSS).

Multiplex ligation-dependent probe amplification

The Multiplex Ligation-dependent Probe Amplification (MLPA) analysis was performed with the SALSA MLPA probemix P458-B1 Gastric Cancer (MRC-Holland, Amsterdam, The Netherlands), starting from 100 ng of DNA of each sample. This kit analyzes the copy number variations (CNVs) of 16 genes implicated in GC (*CCNE1*, *CCND1*, *CDK6*, *ERBB2*, *EGFR*, *MET*, *FGFR2*, *KRAS*, *MYC*, *PTP4A3*, *PIK3CA*, *KLF5*, *TOP2A*, *FGFR1*, *GATA4*, *GATA6*), through the amplification of multiple probes. The list of probes and chromosomal locations is reported in Supplementary Table S2. Amplicons were separated by capillary electrophoresis using the 3130 Genetic Analyzer (Applied Biosystems), and data analysis was carried out through Coffalyser MLPA-DATA version 9.4 software (MRC-Holland) comparing each tumor tissue with a pool of healthy gastric tissues. To determine the CNVs of all genes, normalized probe ratios that were less than 0.7 were classified as "loss", between 0.7 and 1.3 as "normal" and higher than 1.3

as "gain", according to the manufacturer's instructions and as previously established [23]. Genes with at least one probe with "gain" were classified as "amplified".

Sequencing analysis of *TP53*

TP53 exons 5–8 were amplified, starting from 20 ng of DNA, with Multiplex PCR kit (Qiagen) and custom primers (sequences are available upon request) and PCR products were sequenced using BigDye Terminator v.3.1 cycle sequencing kit (Thermo Fisher Scientific) on an ABI-3130 Genetic analyzer (Applied Biosystems). The mutations identified have been reported using the Human Genome Variation Society (HGVS) nomenclature [24], and their pathogenicity has been assessed in accordance with the guidelines [25, 26]. In particular, we distinguished the variants in Pathogenic (class 5), Likely-Pathogenic (class 4), Variants of Uncertain Significance (class 3), Likely-Benign (class 2) and Benign (class 1), reporting in the results only variants of classes 3–5.

Loss of heterozygosity analysis of *TP53* gene

Loss of heterozygosity (LOH) analysis of *TP53* gene was performed starting from 20 ng of DNA, through amplification and fragment analysis of two microsatellites closed to *TP53* locus: D17S796 and D17S1353 [27]. The amplification of microsatellite regions was performed using Multiplex PCR kit (Qiagen) and custom primers (sequences are available upon request), with forward ones 5-end labeled with FAM or HEX fluorescent dies (Integrated DNA Technologies, Coralville, IA, USA). Subsequently, PCR products were separated by capillary electrophoresis on an ABI-3130 Genetic analyzer (Applied Biosystems). The LOH analysis was performed evaluating, for each microsatellite, the height of the peaks corresponding to the 2 alleles in the tumor sample compared to the normal counterpart. The microsatellites with homozygous alleles (same number of repetitions) were excluded from the analysis. The microsatellite was considered presenting LOH when the ratio of the 2 alleles in the tumor was reduced by more than 30% with respect to the healthy counterpart, as previously reported [28]. The sample was defined as "tumor with LOH" when both the microsatellites presented LOH and "tumor with no LOH" when both of them did not present LOH. The samples with no concordance between the 2 microsatellites were classified as uninformative, as well as the samples with both the microsatellites in homozygosity.

Statistical analysis

This was a retrospective multicenter study for the identification of phenotypical and molecular characteristics of Pen A

and Pen B, also in comparison with a series of T3N0 GCs. No sample size calculation was performed due to the explorative nature of the study. Descriptive statistics were reported as proportions and median values (range). Chi-square test or Fisher test, as appropriate, were used to evaluate the association of clinico-pathological characteristics and gene amplification between tumor subgroups. Non-parametric ranking test (Median test) was used to compare continuous data. Due to the explorative nature of the study, no multiple testing correction was made. All *p* values were based on two-sided testing, and statistical analysis was carried out using SAS software, version 9.4 (SAS Institute, Cary, NC, USA).

Results

Clinico-pathological features do not distinguish submucosa-penetrating EGC subgroups

A case series of 87 patients with GC was collected for the study. In particular, 33 Pen A, 34 Pen B and 20 locally advanced (T3N0) tumors, matched for age, gender and lymph node status, were analyzed.

As shown in Table 1, overall, the patients had a median age of 70 years, 20.9% presented lymph nodes involvement, the most frequent site of cancer onset was the antrum of the

Table 1 Clinico-pathological characteristics of patients

	Total (n=87) n (%)	Pen A (n=33) n (%)	Pen B (n=34) n (%)	T3N0 (n=20) n (%)	p
Median age at diagnosis (years): (range)	70 (41–87)	72 (41–87)	68 (44–86)	71 (46–87)	0.460
Sex					
Males	46 (53.5)	17 (51.5)	17 (51.5)	12 (60.0)	
Females	40 (46.5)	16 (48.5)	16 (48.5)	8 (40.0)	
Unknown/missing	1	0	1	0	0.801
Site					
Antrum	51 (58.7)	19 (57.6)	20 (60.6)	12 (60.0)	
Body	30 (34.5)	13 (39.4)	12 (36.4)	5 (25.0)	
Fundus	4 (4.6)	0	1 (3.0)	3 (15.0)	
Cardias	1 (1.2)	1 (3.0)	0	0	
Unknown/missing	1	0	1	0	0.189
Histotype					
Intestinal	72 (86.8)	30 (93.8)	26 (78.8)	16 (88.8)	
Diffuse	6 (7.2)	2 (6.2)	3 (9.1)	1 (5.6)	
Mixed	5 (6.0)	0	4 (12.1)	1 (5.6)	
Unknown/missing	4	1	1	2	0.323
Histological grade					
1	9 (10.7)	4 (12.1)	4 (12.5)	1 (5.3)	
2	32 (38.1)	12 (36.4)	15 (46.9)	5 (26.3)	
3	43 (51.2)	17 (51.5)	13 (40.6)	13 (68.4)	
Unknown/missing	3	0	2	1	0.427
Lymph nodes					
N − (0)	68 (79.1)	24 (72.7)	24 (72.7)	20 (100.0)	
N + (> 0)	18 (20.9)	9 (27.3)	9 (27.3)	0	
Unknown/missing	1	0	1	0	0.032
Lymphovascular invasion					
Absent	63 (74.1)	23 (69.7)	23 (71.9)	17 (85.0)	
Present	22 (25.9)	10 (30.3)	9 (28.3)	3 (15.0)	
Unknown/missing/NA	2	0	2	0	0.437
Median size-cm (range)	3.0 (0.6–8.0)	2.0 (0.6–6.0)	2.0 (1.0–4.0)	4.7 (2.5–8.0)	
≤ 2	34 (39.5)	17 (51.5)	17 (51.5)	0	
> 2	52 (60.5)	16 (48.5)	16 (48.5)	20 (100.0)	
Unknown/missing	1	0	1	0	0.0002

stomach and tumors mostly belonged to the intestinal-type according to Lauren classification [21].

Regarding the histological grade, 51.5%, 40.6% and 68.4% of Pen A, Pen B and T3N0 tumors were G3, respectively, whereas 36.4%, 46.9% and 26.3% were G2, respectively. The 72.7% of either Pen A and Pen B did not present metastatic lymph nodes, even if 30.3% and 27.3% of them, respectively, showed lymphovascular invasion of the gastric wall, as well as 15.0% of the T3N0 tumors. Finally, the median size of Pen A and Pen B tumors was 2.0 cm, while for T3N0 tumors the median size was 4.7 cm.

Overall, no significant association was found between clinico-pathological parameters and tumor subgroups, except for lymph node involvement, due to the specific T3N0 selection, and for the size of the tumor ($p = 0.0002$), but no differences were observed between Pen A and Pen B (Table 1).

Pyloric gland-type MUC6 is significantly associated with Pen A subgroup

Mucins are high molecular weight glycoproteins expressed in many epithelial tissues, including gastric tissue. In our study, we performed the evaluation of MUC2 (intestinal-type mucin) and MUC6/MUC5AC (gastric-type mucins) on the complete case series ($n = 87$; Supplementary Table S3) by IHC.

The 33.3% and 20.0% of Pen A and T3N0 tumors, respectively, showed a high expression of MUC6, while only 2.9% of Pen B tumors showed a strong expression of the protein. Overall, MUC6 protein levels showed a statistically significant association with the tumor type, being more frequently overexpressed in Pen A tumors ($p = 0.021$, Fig. 1). Moreover, smaller Pen B tumors resulted more frequently negative for MUC6 ($p = 0.020$).

Similarly, MUC5AC was more expressed in Pen A tumors (51.5%) than in Pen B (32.3%), being its expression into T3N0 tumors at intermediate values, although this difference was not statistically significant.

The intestinal-type MUC2 resulted poorly expressed in all tumor subgroups. In fact, 75.8% of Pen A, 70.6% of Pen B and 75.0% of T3N0 tumors were negative, while only 9.1% of Pen A and 5.9% of Pen B were diffusely positive. No T3N0 tumors positive for MUC2 were observed.

Finally, EGFR protein levels were also evaluated. EGFR was expressed only in 27.3% of Pen A, 20.6% of Pen B and 35.0% of T3N0 tumors, but no association with subgroups was found.
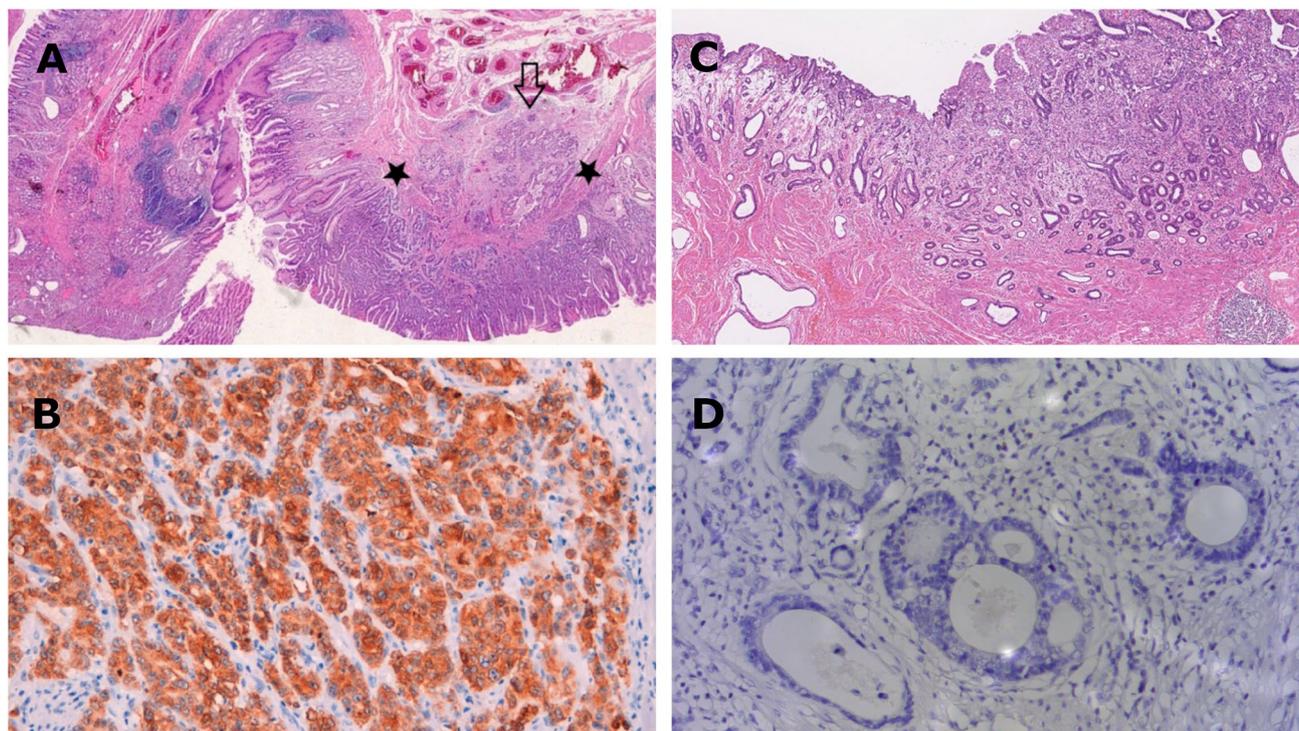
Fig. 1 Immunohistochemistry results of MUC6 protein expression in Pen A and Pen B tumors. **a** Early Gastric Cancer, Kodama Pen A (Hematoxylin and Eosin, × 100). Pen A invades the muscularis mucosae (stars) and penetrates the submucosa layer with a nodular pattern (arrow). **b** MUC6 diffusely positive in Pen A (MUC6, × 400). **c** Early Gastric Cancer, Kodama Pen B (Hematoxylin and Eosin, × 200). **d** MUC6 negative in Pen B (MUC6, × 400). The images shown above are taken from different cases analyzed during the study

Table 2 Results of the MSI analysis on the whole case series (87 tumors)

	Total (n = 87) n (%)	Pen A (n = 33) n (%)	Pen B (n = 34) n (%)	T3N0 (n = 20) n (%)	p
Microsatellite status					
MSS	68 (78.2)	28 (84.8)	30 (88.2)	10 (50.0)	
MSI-H	19 (21.8)	5 (15.2)	4 (11.8)	10 (50.0)	0.002

Table 3 Amplification frequency for each gene in the tumors analyzed by MLPA (47 tumors)

Gene	Pen A (n = 15) n. amplified (%)	Pen B (n = 18) n. amplified (%)	T3N0 (n = 14) n. amplified (%)	p
PTP4A3	10 (66.7)	13 (72.2)	6 (42.9)	0.212
KLF5	10 (66.7)	7 (38.9)	7 (50.0)	0.281
GATA4	10 (66.7)	7 (38.9)	6 (42.9)	0.244
MYC	7 (46.7)	9 (50.0)	5 (35.7)	0.710
GATA6	5 (33.3)	12 (66.7)	1 (7.1)	0.002
MET	4 (26.7)	4 (22.2)	8 (57.1)	0.090
CCND1	7 (46.7)	7 (38.9)	2 (14.3)	0.158
ERBB2	5 (33.3)	6 (33.3)	5 (35.7)	0.988
FGFR1	6 (40.0)	6 (33.3)	3 (21.4)	0.555
TOP2A	5 (33.3)	6 (33.3)	4 (28.6)	0.950
EGFR1	5 (33.3)	6 (33.3)	3 (21.4)	0.717
FGFR2	4 (26.7)	4 (22.2)	4 (28.6)	0.913
CCNE1	4 (26.7)	5 (27.8)	3 (21.4)	0.913
KRAS	5 (33.3)	3 (16.7)	3 (21.4)	0.519
CDK6	4 (26.7)	5 (27.8)	1 (7.1)	0.304
PIK3CA	3 (20.0)	4 (22.2)	2 (14.3)	0.847

GATA6 amplification have been predominantly found in Pen B tumors

High-level MSI and chromosomal instability (CIN) characterized by a large amount of CNV are some of the most important characteristics that distinguish molecular subtypes of sporadic GC. The MSI analysis revealed that 21.8% of the 87 tumors showed MSI-H. The remaining tumors resulted to be MSS, since no tumors with MSI-L were found. With regard to the subgroups, 15.2%, 11.8% and 50.0% of Pen A, Pen B and T3N0 tumors, respectively, were classified as MSI-H (p = 0.002, Table 2). No significant differences were found between EGC subgroups.

The CNV analysis was performed by MLPA on a subset of 47 tumors (15 Pen A, 18 Pen B and 14 T3N0 tumors), due to the poor availability of tissue for some patients.

All the 16 genes present in the panel (Supplementary Table S2) were amplified in at least one sample; on the opposite, no deletions were found. The genes more frequently amplified were PTP4A3 (29/47 tumors, 61.7%), KLF5 (24/47 tumors, 51.1%), GATA4 (23/47 tumors, 48.9%) and MYC (21/47, 44.7%), whereas the less frequently amplified were PIK3CA (9/47 tumors, 19.1%), CDK6 (10/47 tumors, 21.3%) and KRAS (11/47 tumors, 23.4%) (Table 3). Overall, Pen A tumors resulted amplified at similar, or even higher, frequency with respect to Pen B tumors. Interestingly, the only gene with an opposite trend is GATA6 (GATA binding protein 6), which is also the only gene with a copy number significantly different in the 3 subgroups of tumors. In our 47 analyzed samples, GATA6 resulted amplified mainly in Pen B (66.7%) than in Pen A (33.3%) and T3N0 tumors (7.1%) (p = 0.002). A significant association was found between tumor dimensions and GATA6 amplification (p = 0.003), with smaller tumors more frequently amplified.

TP53 gene alterations in early gastric cancer

TP53 is the most recurrent mutated gene in gastric adenocarcinomas, and its alteration has been found early in the progression of this malignancy. In the present study, mutation analysis of TP53 gene hotspot regions (exons 5–8) was performed on 30 Pen A and 31 Pen B tumors of the case series, whereas the LOH analysis was performed on 26 Pen A and 26 Pen B samples (tumor tissues and healthy counterparts), due to the lack of material.

We detected 21 TP53 mutations of classes 3–5 in 20/61 (32.8%) of our EGCs. Only one patient showed 2 different variants. Mutations most frequently involved exon 5 with 11/21 mutations (52.4%), followed by exon 8 with 4/21 mutations (19.0%), while exons 6 and 7 presented both 3/21 mutations (14.3%). Regarding mutation type, 12/21 (57.1%) were missense variants, 5/21 (23.8%) were deletions, 2/21 (9.5%) were splicing variants, 1/21 (4.8%) was a synonymous variant and 1/21 (4.8%) was a duplication. Regarding the pathological significance, 8/21 (38.1%) of the mutations identified in the case series were classified as pathogenic variants (class 5), while 3/21 (14.3%) as likely pathogenic variants (class 4) and 10/21 (47.6%) as variants of uncertain significance (class 3) (Supplementary Table S4).

No significant differences emerged between the two EGC subpopulations, being 9/30 (30.0%) of Pen A and 11/31 (36.7%) of Pen B mutated in TP53 gene, with a Pen B tumor with 2 different mutations. No statistically significant differences in terms of exon, mutation type and pathological significance were identified in the TP53 mutations of the two subgroups (Supplementary Fig. S1a, b and c).

The correlation analysis between TP53 mutation status and the other parameters showed a significant association between TP53 mutation and MSS, indeed the 95.0% of TP53-mutated tumors showed MSS (p = 0.030). A trend was also observed regarding the association between TP53

mutation and lymph nodes involvement ($p = 0.0548$), as well as lymphovascular invasion ($p = 0.0685$), being mutations mainly observed in Pen A tumors.

The LOH analysis was performed by using two microsatellite loci, D17S1353 and D17S796, located upstream and downstream of *TP53* gene, respectively. In 10 out of 52 samples, both microsatellites were homozygous and in 2 samples they were discordant. Consequently, the *TP53* LOH analysis was informative for 40 samples (18 Pen A and 22 Pen B). Among the informative tumors, 20 samples showed LOH (6 Pen A and 14 Pen B). In particular, 66.7% of Pen A with *TP53* LOH had also a *TP53* mutation and 33.3% did not. On the opposite, 35.7% of Pen B with *TP53* LOH had also a *TP53* mutation and 64.3% did not. All the LOH analysis results are summarized in Table 4.

Discussion

EGCs are defined as "adenocarcinomas limited to the gastric mucosa and/or submucosa, regardless of lymph node status" [2]. Even by definition EGCs have good prognosis, a percentage of them is characterized by a survival of 70% at 5 years. The only independent prognostic factors so far identified are the presence of lymph node metastases and Pen A subtype, according to Kodama classification, whereas no useful molecular markers have been pinpointed in this regard [2, 15, 29].

Two classes of submucosa-penetrating EGC with different prognosis, Pen A and Pen B, were analyzed and compared with locally advanced tumors (T3N0), to identify new markers capable of discriminating the different tumor classes.

Our study revealed that analyzed tumors expressed mainly gastric-type mucins, with the pyloric gland-type mucin MUC6 significantly more expressed in Pen A ($p = 0.021$). Mucins are the major components of gastric mucus involved in the protection of gastric mucosa against infection and inflammation and secreted gastric-type mucins are normally expressed in non-neoplastic gastric tissue [30]. Nevertheless, despite their tumor suppressor role in the intestinal tract, their overexpression has been associated with various cancers, as a result of decreased activation of innate and adaptive immune responses [31]. During neoplastic transformation, mucin levels may be altered and reflect the histological origin of the tumor, being MUC5AC mainly secreted by foveolar cells of the stomach antrum and body, while MUC6 by pyloric glands [32–34]. Although published data are often discordant, the expression of gastric-type mucins seems to be associated with adverse clinico-pathological parameters and decreased survival [35–37]. In our study, the overexpression of MUC6 found in Pen A tumors could be an index of a pseudopyloric origin of this specific EGC subtype. Moreover, even if MUC6 levels have not been found to be associated with lymph nodes metastasis, its role in the aggressiveness of this EGC subtype could be further investigated in a larger case series.

The assessment of microsatellite instability in our case series revealed a percentage of MSI tumors similar to other studies [38], with locally advanced tumors having a more frequently unstable profile, probably due to the selection of tumors without nodal involvement [39]. Nevertheless, MSI phenotype could be established in cancer cells even at early stages of tumorigenesis, despite MSI seems to not have a prognostic role in EGC [40, 41]. In our study, we found 13% of MSI-H tumors among EGCs, but no differences were observed between Pen A and Pen B.

GCs belonging to the molecular CIN subtype harbor many CNVs [38], some of which found early in gastric carcinogenesis [42]. We analyzed a panel of 16 genes suggested to be of diagnostic and/or clinical relevance in GC and found that amplification was the most frequent alteration, mainly involving *KLF5* (13q21.1), *GATA4* (8p23.1), *MYC* (8q24.21) and *PTP4A3* (8q24.3) genes. Moreover, *CCND1*, and to a lesser extent also *CCNE1*, were confirmed to be amplified since the early stages, as shown by Ooi et al. [43].

Interestingly, even if the largest number of alterations was observed in Pen A tumors, the only gene more frequently amplified in Pen B, *GATA6*, was also the only one able to significantly distinguish Pen A and Pen B tumors ($p = 0.002$). This gene is located on chromosome 18q11.2 and encodes a transcription factor mainly involved in embryogenesis and development of tissues, including those of the gastrointestinal tract. Its oncogenic role has been demonstrated in many neoplasms [44, 45], including gastric, esophageal and pancreatic adenocarcinomas, in

Table 4 Frequency of *TP53* mutations in relation to *TP53* LOH status in informative EGC samples

	Pen A + Pen B			Pen A			Pen B		
	Total ($n = 40$)	*TP53 mut* ($n = 11$)	*TP53 wt* ($n = 29$)	Total ($n = 18$)	*TP53 mut* ($n = 5$)	*TP53 wt* ($n = 13$)	Total ($n = 22$)	*TP53 mut* ($n = 6$)	*TP53 wt* ($n = 16$)
TP53 LOH (%)	20 (100)	9 (45.0)	11 (55.0)	6 (30.0)	4 (66.7)	2 (33.3)	14 (70.0)	5 (35.7)	9 (64.3)
No TP53 LOH (%)	20 (100)	2 (10.0)	18 (90.0)	12 (60.0)	1 (8.3)	11 (91.7)	8 (40.0)	1 (12.5)	7 (87.5)
p	0.031			0.022			0.351		

which *GATA6* is frequently amplified or overexpressed. However, in recent years, conflicting results on the function of this gene have also been published and a dual role of GATA6 is emerging, albeit largely to be clarified. In ductal pancreatic cancer, the amplification and overexpression of *GATA*6 were related to an increased tumor proliferation in a group of patients, whereas in other patients the high expression of GATA6 seemed to be associated with better prognosis, supporting its context-dependent role into inflammatory and cancer-related pathways [45–47]. GATA6 overexpression is observed already from the earliest phases of gastroesophageal tumorigenesis and reflects the presence of intestinal differentiation, acting on targets involved in cell replication and self-renewal, differentiation and transcriptional control [48]. Moreover, KLF5/GATA4/GATA6 collaborate to maintain a pro-oncogenic transcriptional regulatory network in GC cells, thus promoting GC development [49]. On the opposite, Liu et al. have recently demonstrated that GATA6 is significantly downregulated in metastatic tumors and its overexpression, in cellular models, suppresses migration, invasion and metastasization of gastric cancer cells [50]. In the present study, we showed that the amplification of *GATA6* significantly correlated with reduced tumor size ($p = 0.003$) and was associated with Pen B tumors, an EGC subtype that is known to be characterized by better prognosis compared to Pen A tumors [17]. It is also interesting to note that in Pen B subgroup, *GATA6* was amplified more frequently in N0 and low-grade tumors, although no statistical significance was reached, supporting the possible role of this gene in preventing tumor dissemination and in maintaining epithelial cells in a differentiated state.

TP53 is the most commonly altered gene in GC [38, 51]. It has an important role in many fundamental cellular processes such as cell cycle arrest, DNA repair, apoptosis and senescence [52], and alterations in this gene at early tumor stages seem to engage environmental exposures to promote gastric premalignancy [53].

The genetic analysis of the *TP53* gene performed on Pen A and Pen B tumors of our case series revealed an overall mutation frequency of 32.8%, consolidating the idea that alterations in this gene are a driver event detectable in the early stages of gastric carcinogenesis. The *TP53* mutations were significantly associated with MSS ($p = 0.03$), suggesting the presence of different molecular subgroups of GC in our case series, in particular CIN, mainly associated with *TP53* mutations, and MSI, characterized by microsatellite instability.

No significant differences between Pen A and Pen B were observed in *TP53* mutations. According to the literature, most of the identified mutations affect exon 5 of the *TP53* gene and consist of missense variants, followed by deletions and duplications [54]. Furthermore, it has been confirmed that most of the *TP53* variants are transitions G:C > A:T (approximately 61.9%).

Interestingly, it has been shown that *TP53* hotspot mutations co-occurring with LOH (*TP53*$^{mut/LOH}$) result in protein stabilization and oncogenic gain of function that contribute to acceleration of metastasization, increased drug resistance and worse survival [55]. In the present study, we found 50.0% of Pen tumors presenting *TP53* LOH, mainly belonging to Pen B subtype. Moreover, 27% of EGCs showed only allelic loss at *TP53* gene-level without mutations in hotspot regions with 41% of Pen B and 11% of Pen A having this characteristic. On the other hand, despite the low number of cases, *TP53*$^{mut/LOH}$ co-occurrence was significantly associated with Pen A subtype ($p = 0.022$). This is of particular interest, since data recently published highlighted that *TP53*$^{mut/LOH}$ was more frequently observed in EGC with poor prognosis, but not in advanced tumors, demonstrating its prognostic significance especially in patients with early disease [20].

The present study has several limitations. First, this is a retrospective study. Second, the small number of cases analyzed, due to the low frequency of Pen B lesions. Furthermore, given the number of analyses performed in this study and the consequent necessity of a considerable quantity of biological material, the number of cases analyzed for some markers was reduced.

Conclusions

Our results highlighted for the first time some molecular and cellular markers capable of distinguishing early-stage stomach tumors, Pen A and Pen B. Despite the small number of cases, the copy number of *GATA6* and the expression of MUC6 significantly distinguished the two tumor subtypes. Moreover, even if further investigations are needed, the co-occurrence of *TP53* mutations and LOH could be a marker of the more aggressive Pen A subtype. It will be interesting to confirm the data obtained in a wider case series of EGCs and to correlate them with prognosis of patients affected by submucosa-penetrating gastric lesions, with the final aim of finding suitable biomarkers to be used in the pre-surgical phase, able to distinguish patients with different prognosis and usable to decide the best therapeutic strategy.

Author contributions CM, PM, GM, PU and LSa conceived the idea for and designed the study; CM, GT, FR, LC, SR and MMT were responsible for data acquisition and interpretation; ES performed the statistical analysis; PM, AT, RB, MRA, ADI, LSo, FL, GE and LSa

collected samples and interpreted the clinical patient data; AT, RB, MRA, FL and LSa performed the histo-pathological examination of the tissue specimens; CM, GT and PU were major contributor in writing the manuscript; PM, GM and LSa critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Compliance with ethical standards

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.
2. Japanese Gastric Cancer Association. Japanese classification of gastric carcinoma English edition. Gastric Cancer. 2011;14:101–12.
3. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. Lancet (London, England). 2016;388:2654–64.
4. Saragoni L, Gaudio M, Vio A, Folli S, Nanni O, Saragoni A. Early gastric cancer in the province of Forlì: follow-up of 337 patients in a high risk region for gastric cancer. Oncol Rep. 1998;5:945–8.
5. Sano T, Sasako M, Kinoshita T, Maruyama K. Recurrence of early gastric cancer. Follow-up of 1475 patients and review of the Japanese literature. Cancer. 1993;72:3174–8.
6. Bollschweiler E, Berlth F, Baltin C, Mönig S, Hölscher AH. Treatment of early gastric cancer in the Western World. World J Gastroenterol. 2014;20:5672–8.
7. Sano T, Katai H, Sasako M, Maruyama K. The management of early gastric cancer. Surg Oncol. 2000;9:17–22.
8. Abe S, Yoshimura H, Nagaoka S, Monden N, Kinugasa S, Nagasue N, et al. Long-term results of operation for carcinoma of the stomach in T1/T2 stages: critical evaluation of the concept of early carcinoma of the stomach. J Am Coll Surg. 1995;181:389–96.
9. Sano T, Kobori O, Muto T. Lymph node metastasis from early gastric cancer: endoscopic resection of tumour. Br J Surg. 1992;79:241–4.
10. Folli S, Dente M, Dell'Amore D, Gaudio M, Nanni O, Saragoni L, et al. Early gastric cancer: prognostic factors in 223 patients. Br J Surg. 1995;82:952–6.
11. Saragoni L, Gaudio M, Morgagni P, Folli S, Vio A, Scarpi E, et al. The role of growth patterns, according to Kodama's classification, and lymph node status, as important prognostic factors in early gastric cancer: analysis of 412 cases. Gastric Cancer. 2000;3:134–40.
12. Jung DH, Huh CW, Kim J-H, Hong JH, Park JC, Lee YC, et al. Risk-Stratification model based on lymph node metastasis after noncurative endoscopic resection for early gastric cancer. Ann Surg Oncol. 2017;24:1643–9.
13. Kodama Y, Inokuchi K, Soejima K, Matsusaka T, Okamura T. Growth patterns and prognosis in early gastric carcinoma. Superficially spreading and penetrating growth types. Cancer. 1983;51:320–6.
14. Saragoni L, Morgagni P, Gardini A, Marfisi C, Vittimberga G, Garcea D, et al. Early gastric cancer: diagnosis, staging, and clinical impact. Evaluation of 530 patients. New elements for an updated definition and classification. Gastric Cancer. 2013;16:549–54.
15. Saragoni L. Upgrading the definition of early gastric cancer: better staging means more appropriate treatment. Cancer Biol Med. 2015;12:355–61.
16. Saragoni L, Scarpi E, Ravaioli A, Morgagni P, Roviello F, Vindigni C, et al. Early gastric cancer: clinical behavior and treatment options. Results of an Italian multicenter study on behalf of the Italian Gastric Cancer Research Group (GIRCG). Oncologist. 2018;23:852–8.
17. Morgagni P, Solaini L, Saragoni L, Bencivenga M, Fumagalli Romario U, Graziosi L, et al. An aggressive early gastric cancer: Kodama's PenA type. Eur J Surg Oncol. 2018;44:1186–90.
18. Inokuchi K, Kodama Y, Sasaki O, Kamegawa T, Okamura T. Differentiation of growth patterns of early gastric carcinoma determined by cytophotometric DNA analysis. Cancer. 1983;51:1138–41.
19. Chen D, Chen G, Jiang W, Fu M, Liu W, Sui J, et al. Association of the collagen signature in the tumor microenvironment with lymph node metastasis in early gastric cancer. JAMA Surg. 2019;154:e185249.
20. Datta J, Da Silva EM, Kandoth C, Song T, Russo AE, Hernandez JM, et al. Poor survival after resection of early gastric cancer: extremes of survivorship analysis reveal distinct genomic profile. Br J Surg. 2020;107:14–9.
21. Lauren P. The two histological main types of gastric carcinoma: diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. Acta Pathol Microbiol Scand. 1965;64:31–49.
22. Umar A, Boland CR, Terdiman JP, Syngal S, de la Chapelle A, Rüschoff J, et al. Revised Bethesda guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability. J Natl Cancer Inst. 2004;96:261–8.
23. Bunyan DJ, Eccles DM, Sillibourne J, Wilkins E, Thomas NS, Shea-Simonds J, et al. Dosage analysis of cancer predisposition genes by multiplex ligation-dependent probe amplification. Br J Cancer. 2004;91:1155–9.
24. den Dunnen JT, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, et al. HGVS Recommendations for the description of sequence variants: 2016 update. Hum Mutat. 2016;37:564–9.
25. Plon SE, Eccles DM, Easton D, Foulkes WD, Genuardi M, Greenblatt MS, et al. Sequence variant classification and reporting: recommendations for improving the interpretation of cancer susceptibility genetic test results. Hum Mutat. 2008;29:1282–91.
26. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–24.
27. Saeki H, Kitao H, Yoshinaga K, Nakanoko T, Kubo N, Kakeji Y, et al. Copy-neutral loss of heterozygosity at the p53 locus in carcinogenesis of esophageal squamous cell carcinomas associated with p53 mutations. Clin Cancer Res. 2011;17:1731–40.
28. Oki E, Zhao Y, Yoshida R, Egashira A, Ohgaki K, Morita M, et al. The difference in p53 mutations between cancers of the upper and lower gastrointestinal tract. Digestion. 2009;79(Suppl 1):33–9.
29. Everett SM, Axon AT. Early gastric cancer in Europe. Gut. 1997;41:142–50.

30. Ho SB, Shekels LL, Toribara NW, Kim YS, Lyftogt C, Cherwitz DL, et al. Mucin gene expression in normal, preneoplastic, and neoplastic human gastric epithelium. Cancer Res. 1995;55:2681–90.

31. Kufe DW. Mucins in cancer: function, prognosis and therapy. Nat Rev Cancer. 2009;9:874–85.

32. Bartman AE, Buisine MP, Aubert JP, Niehans GA, Toribara NW, Kim YS, et al. The MUC6 secretory mucin gene is expressed in a wide variety of epithelial tissues. J Pathol. 1998;186:398–405.

33. Saito A, Shimoda T, Nakanishi Y, Ochiai A, Toda G. Histologic heterogeneity and mucin phenotypic expression in early gastric cancer. Pathol Int. 2001;51:165–71.

34. Cavalcanti E, De Michele F, Lantone G, Panarese A, Caruso ML. Mucin phenotype of differentiated early gastric cancer: an immunohistochemistry study supporting therapeutic decision making. Cancer Manag Res. 2019;11:5047–54.

35. Kocer B, Soran A, Kiyak G, Erdogan S, Eroglu A, Bozkurt B, et al. Prognostic significance of mucin expression in gastric carcinoma. Dig Dis Sci. 2004;49:954–64.

36. Koseki K, Takizawa T, Koike M, Ito M, Nihei Z, Sugihara K. Distinction of differentiated type early gastric carcinoma with gastric type mucin expression. Cancer. 2000;89:724–32.

37. Kerckhoffs KGP, Liu DHW, Saragoni L, van der Post RS, Langer R, Bencivenga M, et al. Mucin expression in gastric- and gastro-oesophageal signet-ring cell cancer: results from a comprehensive literature review and a large cohort study of Caucasian and Asian gastric cancer. Gastric Cancer. 2020;23:765–79.

38. Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513:202–9.

39. Polom K, Marrelli D, Pascale V, Ferrara F, Voglino C, Marini M, et al. The pattern of lymph node metastases in microsatellite unstable gastric cancer. Eur J Surg Oncol. 2017;43:2341–8.

40. Mizuguchi A, Takai A, Shimizu T, Matsumoto T, Kumagai K, Miyamoto S, et al. Genetic features of multicentric/multifocal intramucosal gastric carcinoma. Int J cancer. 2018;143:1923–34.

41. Kim DG, An JY, Kim H, Shin S-J, Choi S, Seo WJ, et al. Clinical implications of microsatellite instability in early gastric cancer. J Gastric Cancer. 2019;19:427–37.

42. Sugai T, Eizuka M, Arakawa N, Osakabe M, Habano W, Fujita Y, et al. Molecular profiling and comprehensive genome-wide analysis of somatic copy number alterations in gastric intramucosal neoplasias based on microsatellite status. Gastric Cancer. 2018;21:765–75.

43. Ooi A, Oyama T, Nakamura R, Tajiri R, Ikeda H, Fushida S, et al. Gene amplification of CCNE1, CCND1, and CDK6 in gastric cancers detected by multiplex ligation-dependent probe amplification and fluorescence in situ hybridization. Hum Pathol. 2017;61:58–67.

44. Tsuji S, Kawasaki Y, Furukawa S, Taniue K, Hayashi T, Okuno M, et al. The miR-363-GATA6-Lgr5 pathway is critical for colorectal tumourigenesis. Nat Commun. 2014;5:3150.

45. Zhong Y, Wang Z, Fu B, Pan F, Yachida S, Dhara M, et al. GATA6 activates Wnt signaling in pancreatic cancer by negatively regulating the Wnt antagonist Dickkopf-1. PLoS ONE. 2011;6:e22129.

46. Martinelli P, Carrillo-de Santa Pau E, Cox T, Sainz B, Dusetti N, Greenhalf W, et al. GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. Gut. 2017;66:1665–76.

47. Martinelli P, Madriles F, Cañamero M, Pau ECS, Del PN, Guerra C, et al. The acinar regulator Gata6 suppresses KrasG12V-driven pancreatic tumorigenesis in mice. Gut. 2016;65:476–86.

48. Sulahian R, Casey F, Shen J, Qian ZR, Shin H, Ogino S, et al. An integrative analysis reveals functional targets of GATA6 transcriptional regulation in gastric cancer. Oncogene. 2014;33:5637–48.

49. Chia N-Y, Deng N, Das K, Huang D, Hu L, Zhu Y, et al. Regulatory crosstalk between lineage-survival oncogenes KLF5, GATA4 and GATA6 cooperatively promotes gastric cancer development. Gut. 2015;64:707–19.

50. Liu H, Du F, Sun L, Wu Q, Wu J, Tong M, et al. GATA6 suppresses migration and metastasis by regulating the miR-520b/CREB1 axis in gastric cancer. Cell Death Dis. 2019;10:35.

51. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012;2:401–4.

52. Bellini MF, Cadamuro ACT, Succi M, Proença MA, Silva AE. Alterations of the TP53 gene in gastric and esophageal carcinogenesis. J Biomed Biotechnol. 2012;2012:891961.

53. Sethi NS, Kikuchi O, Duronio GN, Stachler MD, McFarland JM, Ferrer-Luna R, et al. Early TP53 alterations engage environmental exposures to promote gastric premalignancy in an integrative mouse model. Nat Genet. 2020;52:219–30.

54. Olivier M, Hollstein M, Hainaut P. TP53 mutations in human cancers: origins, consequences, and clinical use. Cold Spring Harb Perspect Biol. 2010;2:a001008.

55. Alexandrova EM, Mirza SA, Xu S, Schulz-Heddergott R, Marchenko ND, Moll UM. p53 loss-of-heterozygosity is a necessary prerequisite for mutant p53 stabilization and gain-of-function in vivo. Cell Death Dis. 2017;8:e2661.