1 **1) Article type: Original article**

2

3 **2) Title: Indolent feature of *Helicobacter pylori*-uninfected intramucosal signet**

4 **ring cell carcinomas with *CDH1* mutations**

5

6 **3) Authors:**

7 Mitsuhiro Nikaido[1,13], Nobuyuki Kakiuchi[1,2,3,13], Shin'ichi Miyamoto[1,4], Tomonori

8 Hirano[1,2,3], Yasuhide Takeuchi[2,5], Taro Funakoshi[6], Akira Yokoyama[6], Tatsuki

9 Ogasawara[2,3], Yoshihiro Yamamoto[6], Atsushi Yamada[6], Takeshi Setoyama[1,7],

10 Takahiro Shimizu[1], Yukari Kato[8], Suguru Uose[8], Takaki Sakurai[5,9], Sachiko

11 Minamiguchi[5], Kazutaka Obama[10], Yoshiharu Sakai[10,11], Manabu Muto[6], Tsutomu

12 Chiba[1,8], Seishi Ogawa[2,3,12], and Hiroshi Seno[1]

13

14 **4) Affiliations:**

15 1. Department of Gastroenterology and Hepatology, Kyoto University Graduate

16   School of Medicine, Kyoto, Japan

17 2. Department of Pathology and Tumor Biology, Kyoto University Graduate School of

18   Medicine, Kyoto, Japan

19 3. Institute for the Advanced Study of Human Biology (WPI-ASHBi), Kyoto, Japan

20   4. Department of Gastroenterology, National Hospital Organization Kyoto Medical

21   Center, Kyoto, Japan

22   5. Department of Diagnostic Pathology, Kyoto University Graduate School of

23   Medicine, Kyoto, Japan

24   6. Department of Therapeutic Oncology, Kyoto University Graduate School of

25   Medicine, Kyoto, Japan

26   7. Department of Gastroenterology, Osaka Red Cross Hospital, Osaka, Japan

27   8. Department of Gastroenterology and Hepatology, Kansai Electric Power Hospital,

28   Osaka, Japan

29   9. Department of Pathology, Kansai Electric Power Hospital, Osaka, Japan

30   10. Department of Surgery, Kyoto University Graduate School of Medicine, Kyoto,

31   Japan

32   11. Department of Surgery, Osaka Red Cross Hospital, Osaka, Japan

33   12. Department of Medicine, Center for Hematology and Regenerative Medicine,

34   Karolinska Institute, Stockholm, Sweden

35   13. These authors contributed equally: Mitsuhiro Nikaido and Nobuyuki Kakiuchi.

36

37   **5) Corresponding author information:**

38   Shin'ichi Miyamoto

39 Department of Gastroenterology, National Hospital Organization Kyoto Medical Center,

40 1-1 Fukakusa-Mukaihata-Cho, Fushimi, Kyoto, 612-8555, Japan

41 (Phone) 81-(75)-641-9161, (Fax) 81-(75)-643-4325

42 Email: shmiyamo@kuhp.kyoto-u.ac.jp

43

44 **6) A short running head:** *Hp*-uninfected signet ring cell carcinoma

45

46 **7) Word count: 3873**

47

Abstract

Background: In *Helicobacter pylori* (*Hp*)-uninfected individuals, diffuse-type gastric cancer (DGC) was reported as the most common type of cancer. However, the carcinogenic mechanism of *Hp*-uninfected sporadic DGC is largely unknown.

Methods: We performed whole-exome sequencing of *Hp*-uninfected DGCs and *Hp*-uninfected normal gastric mucosa. For advanced DGCs, external datasets were also analyzed.

Results: Eighteen patients (aged 29–78 years) with DGCs and nine normal subjects (28–77 years) were examined. The mutation burden in intramucosal DGCs (10–66 mutations per exome) from individuals aged 29–73 years was not very different from that in the normal gastric glands, which showed a constant mutation accumulation rate (0.33 mutations/exome/year). Unbiased dN/dS analysis showed that *CDH1* somatic mutation was a driver mutation for intramucosal DGC. *CDH1* mutation was more frequent in intramucosal DGCs (67%) than in advanced DGCs (27%). In contrast, *TP53* mutation was more frequent in advanced DGCs (52%) than in intramucosal DGCs (0%). This discrepancy in mutations suggests that *CDH1*-mutated intramucosal DGCs make a relatively small contribution to advanced DGC formation. Among the 16 intramucosal DGCs (median size, 6.5 mm), 15 DGCs were pure signet ring cell carcinoma (SRCC) with reduced E-cadherin expression and a low proliferative

4

74 capacity (median Ki-67 index, 2.4%). Five SRCCs reviewed endoscopically over 2–5

75 years showed no progression.

76 Conclusions: Impaired E-cadherin function due to *CDH1* mutation was considered as

77 an early carcinogenic event of *Hp*-uninfected intramucosal SRCC. Genetic and clinical

78 analyses suggest that *Hp*-uninfected intramucosal SRCCs may be less likely to

79 develop into advanced DGCs.

80

81 **Mini abstract:** The *CDH1* somatic mutation was considered as an early carcinogenic

82 event in *Hp*-uninfected intramucosal SRCC. *Hp*-uninfected SRCCs may progress

83 slowly and be less likely to develop into advanced DGCs.

84

85 **Keywords**: stomach neoplasm, signet ring cell carcinoma, *Helicobacter pylori*, CDH1,

86 whole-exome sequencing

87

Introduction

Epidemiological studies have reported that *Helicobacter pylori* (*Hp*) infection is the greatest risk factor for gastric cancer [1, 2]. Chronic gastritis caused by *Hp* infection was shown to promote oncogenic mutations [3, 4]. Although most gastric cancers arise in *Hp*-infected stomachs, gastric cancers in *Hp*-uninfected stomachs also occur, with their prevalence reported to account for 0.42–5.4% of all gastric cancers [5]. The prevalence of *Hp* infection is lower in developed countries than in developing countries [6]. In Japan, the prevalence of *Hp* infection and incidence of gastric cancer was previously high. However, in recent decades, its prevalence and mortality rate by gastric cancer have decreased [7-9]. The incidence of *Hp*-uninfected gastric cancers is expected to increase as the *Hp* infection rate declines. However, the molecular basis of gastric cancer in the *Hp*-uninfected stomach is largely unknown. Recent studies showed that somatic mutations accumulate in normal tissues with age [10-14]; however, mutation accumulation in the *Hp*-uninfected normal gastric mucosa has not been reported.

Histologically, gastric cancer is classified into intestinal-type gastric cancer and diffuse-type gastric cancer (DGC) [15]. DGCs include poorly differentiated adenocarcinomas (PDAs) and signet ring cell carcinomas (SRCCs). In *Hp*-uninfected patients, gastric cancers are most often found as intramucosal DGCs, which are

107   typically 'pure' SRCCs that contain only SRCC cells and no PDA components [5, 16].

108   *Hp*-uninfected intramucosal DGC is smaller in size than *Hp*-infected intramucosal

109   DGC [16, 17]. Additionally, T1 stage (UICC classification) DGCs were shown to be

110   significantly less invasive in patients without *Hp* infection than in those with *Hp*

111   infection [17].

112   Hereditary diffuse gastric cancer (HDGC) is mainly caused by *a CDH1* germline

113   mutation [18, 19]. E-cadherin, encoded by *CDH1*, is a cell adhesion molecule that

114   regulates signaling pathways associated with cell proliferation, survival, invasion, and

115   migration. Based on studies of HDGC, DGCs are thought to arise from SRCCs

116   because of reduced cell adhesion associated with the loss of E-cadherin function [19].

117   Although somatic *CDH1* mutations have been found in 29–56% of sporadic DGCs in

118   previous genetic analyses, the mechanisms of sporadic DGC development remain

119   unclear [20-22]. *RHOA* mutations and *CLDN18-ARHGAP* fusion have also been

120   reported to cause DGC [21-24]. These results were mainly obtained in studies of

121   advanced gastric cancer, whereas few studies have examined intramucosal cancer.

122   Furthermore, previous studies either mainly consisted of *Hp*-infected patients or did

123   not assess the *Hp* infection status. In the present study, we examined the

124   clinicopathological and genetic features of *Hp*-uninfected DGCs. We performed whole-

125   exome sequencing (WES) to investigate the genetic alterations in early and advanced

126 stage DGCs. These genetic alterations were then compared with those in the *Hp*-

127 uninfected normal gastric mucosa.

128

129 **Materials and methods**

130 **Patient and sample collection**

131 Tumor samples, paired blood samples, and clinical data were collected from patients

132 diagnosed with *Hp*-uninfected DGC and enrolled in the Kyoto University Hospital and

133 Kansai Electric Power Hospital between January 2016 and March 2020. As normal

134 controls, we collected biopsied gastric tissues and peripheral blood from nine subjects

135 without *Hp*-infection or gastric cancer who underwent diagnostic endoscopy for upper

136 gastrointestinal symptoms. All patients and normal subjects provided written informed

137 consent to participate in the study. The study protocol conformed to the ethical

138 guidelines of the Declaration of Helsinki and was approved by the ethics committee of

139 Kyoto University Hospital and institutional board of Kansai Electric Power Hospital.

140 Certified pathologists (S.M. and T.S.) histologically examined the surgically- or

141 endoscopically-resected or biopsied specimens according to the third English edition

142 of Japanese classification of gastric carcinoma and Lauren's classification [15, 25].

143

144 ***Hp* infection status evaluation**

145 Diagnostic criteria for *Hp*-uninfected gastric cancer have not been established. These

146 criteria must be strict to rule out naturally *Hp*-eradicated cases [5]. Patients were

147 defined as *Hp*-uninfected when they met all of the following criteria: (i) no history of *Hp*

148 eradication therapy; (ii) negative results in the *Hp* infection test (serum anti-*Hp*

149 antibody <3 U/mL and/or negative urea breath test result); (iii) absence of *Hp* bacterial

150 body in normal gastric mucosa identified by immunohistochemistry with an anti-*Hp*

151 antibody (Dako, Glostrup, Denmark), Giemsa staining, or hematoxylin and eosin

152 (H&E) staining of formalin-fixed paraffin-embedded (FFPE) sections; and (iv) no

153 endoscopic findings of mucosal atrophy.

154

155 **Immunohistochemistry**

156 Immunohistochemistry analysis was performed on FFPE specimens using primary

157 antibodies against E-cadherin (clone NCH-38, Dianova, Hamburg, Germany) and Ki-

158 67 (clone MIB-1, Dako). The proportion of tumor cells positive for Ki-67 was measured

159 at the site with the highest number of labeled nuclei. Investigators were blinded to the

160 tumor genotypes.

161

162 **Tumor-DNA extraction**

163 Ten-micrometer-thick sections were sliced from the FFPE tissue onto membrane

164 frame slides (Leica, Wetzlar, Germany). Laser capture microdissection of tumor cells

165 was performed using a Leica LMD 7000 instrument. To perform WES, genomic DNA

166 was extracted from the microdissected tumor cells and matched peripheral blood using

167 the GeneRead DNA FFPE Kit and the QIAamp DNA Mini Kit (Qiagen, Hilden,

168 Germany), respectively.

169

170 **Gastric single gland isolation**

171 We performed WES on *Hp*-uninfected normal gastric epithelia for comparison with *Hp*-

172 uninfected DGCs. Normal gastric epithelium is composed of single-layer columnar

173 cells that are compacted into numerous small replication units known as glands. Each

174 gland is replenished by stem cells located at the neck or base of the gland [26]. To

175 evaluate the precise somatic mutation in the *Hp*-uninfected normal gastric epithelia,

176 genomic DNA extracted from fresh single glands was subjected to WES as previously

177 described [11]. Briefly, two biopsies of normal gastric mucosa were obtained from the

178 fundic gland area, at the gastric angle, for each *Hp*-uninfected subject without a history

179 of gastric cancer. A total of 18 samples were obtained from nine subjects. Gastric

180 epithelia were manually dissociated from the lamina propria after treating the gastric

181 mucosa with 20 mM EDTA in PBS at 4°C for 20 min. Subsequently, single glands were

182 picked up under a stereomicroscope. Genomic DNA isolated from each gland was split

183 into two aliquots, each of which was independently subjected to whole-genome

184 amplification (WGA) using the REPLI-g Single Cell Kit (Qiagen). Both amplified DNA

185 samples were subjected to WES independently. Variants commonly detected in the

186 two split samples were considered as true somatic mutations.

187

188 **Whole-exome sequencing**

189 WES libraries were prepared using SureSelect Human All Exon V5 (Agilent

190 Technologies, Santa Clara, CA, USA) or xGen Exome Research Panel (Integrated

191 DNA Technologies, Coralville, IA, USA), followed by sequencing of enriched exon

192 fragments using a HiSeq 2500 or NovaSeq 6000 system (Illumina, San Diego, CA,

193 USA) in 100–150-bp paired-end mode as previously described [27]. The target depth

194 was 100×, and the actual depth was 108× on average. Sequencing reads were aligned

195 to the human reference genome (GRCh37), and mutation calling was performed using

196 the Genomon2 pipeline (v.2.6) as previously described [11]. Candidate mutations were

197 adopted using the Empirical Bayesian Mutation Calling (EBCall) algorithm with the

198 following filtering criteria: (i) a sufficient number of reads (total reads ≥8 and variant

199 reads ≥3); (ii) variant allele frequencies (VAFs) ≥0.05 (for FFPE samples), ≥0.25 (for

200 single-gland WGA samples), and <0.02 (for germline control); (iii) a strand ratio not

201 equal to 0 or 1; and (iv) p value by EBCall ≤$10^{-3.5}$. Putative germline variants were also

202 excluded by comparing VAFs with matched controls using Fisher's exact test ($p \leq 10^{-1}$).

203 Candidate mutations in tumor samples were validated by PCR-based amplicon deep

204 sequencing [27]. Because of the limited amount of DNA available, validation was

205 performed with three samples, and the true positive rate was 95% (39 of 41).

206 Driver genes were investigated based on dN/dS, which is the ratio of the number of

207 nonsynonymous substitutions per nonsynonymous site to the number of synonymous

208 substitutions per synonymous site, using dNdScv [28]. We also adopted 69 driver

209 genes previously reported in comprehensive genetic analyses of gastric cancer [21-

210 24] (Table S1). Because mutations in the driver gene had a high prior probability of

211 being true mutations, we included the driver gene mutations found by loosening the

212 filter criteria: (i) VAFs >0.04; (ii) p value by EB call $\leq 10^{-3}$; and (iii) a strand ratio = 0–1.

213 Copy number abnormalities were evaluated using WES data with our in-house pipeline

214 'CNACS' [12].

215

216 **External dataset**

217 Because we collected only two cases of advanced DGC, somatic mutation data

218 identified by MuTect and clinical information from patients with DGC were downloaded

219 from The Cancer Genome Atlas (TCGA) data portal [23]. Thirty-seven DGC cases

220 from TCGA were reported as *Hp*-negative, and all had T2–T4 tumors (UICC TNM

221   classification 6th edition). As hypermutated tumors were considered genetically

222   different from our samples, tumors with more than 12 mutations/Mb were excluded

223   from the comparative data [23]. We also used 30 *Hp*-negative advanced DGC data

224   from Hong Kong [22], all of which were DGCs without microsatellite instability. In total,

225   data for 60 *Hp*-uninfected advanced DGCs (two from our cohort, 28 from TCGA, and

226   30 from Hong Kong dataset) were collected and compared with data for nine

227   intramucosal DGC cases from our cohort. The Hong Kong dataset could not be used

228   for mutation number and mutational signature analyses because of a lack of

229   information about synonymous mutations. Next, 282 single-nucleotide variants (SNVs)

230   from 11 DGC samples, 260 SNVs from 18 normal glands, and 2596 SNVs from 28

231   DGCs from TCGA were allocated to 65 'the Catalogue Of Somatic Mutations In Cancer'

232   (COSMIC) single base substitution signatures (SBS) using SigProfilerExtractor [29].

233

234   **Statistical analysis**

235   All tests were two-tailed, and $p < 0.05$ was considered as significant. The linearity of

236   the number of mutations in the normal gastric glands, intramucosal DGCs, and

237   advanced DGCs with age was evaluated based on Pearson's correlation coefficient in

238   a linear regression model that assumed a zero intercept, because the number of

239   somatic mutations in an exome region at 0 years of age was assumed to be nearly 0,

240   as previously described [11]. The Mann–Whitney *U* test was used to compare normal

241   glands, intramucosal DGCs, and advanced DGCs. Driver mutation rates between

242   intramucosal and advanced DGCs were compared by Fisher's exact test with

243   Benjamini-Hochberg adjustment. $q < 0.05$ was considered to indicate significance. All

244   statistical analyses were performed using R software (version 3.6.3).

245

246   **Results**

247   **Clinical information**

248   The clinical information of 18 patients with *Hp*-uninfected DGCs (median age, 57

249   years; range 29–78) and nine subjects without *Hp* infection (median age, 44 years;

250   range 28–77) was collected (Table 1). Sixty-seven percent of patients with DGC had

251   a smoking history, and 33% were current smokers (Table 2). Fifty percent of patients

252   with DGC had alcohol consumption habit. Among the 18 patients, 16 were diagnosed

253   with intramucosal DGC (median age, 60 years; range 29–73) and two with advanced

254   DGC. Sixteen intramucosal DGCs were found as flat pale areas upon endoscopy, all

255   of which were resected endoscopically and were found to contain an SRCC

256   component. One DGC localized in the gastric cardia consisted of a mixture of SRCC

257   and PDA components (Fig. 1a and lower panels of Fig. 1b), and the remaining 15

258   DGCs localized near the gastric angle were 'pure' SRCCs scattered around the neck

259   of the gland (Figs. 1a and upper panels of 1b). These intramucosal 'pure' SRCCs had

260   a low proliferation capacity, with a median Ki-67 labeling index of 2.4% (range, 0–

261   15.4%). Five cases with intramucosal 'pure' SRCC showed no progression over 2–5

262   years according to endoscopic image review (Fig. 2).

263

264   **Genomic analysis of intramucosal DGCs**

265   Tumor cells were dissected by laser microdissection to increase the tumor cell content

266   because the intramucosal DGCs were small (median, 6.5 mm; range, 3–14 mm) and

267   their tumor cells were scattered around the necks of glands. Small tumors with low

268   densities of signet ring cells did not yield adequate amounts of DNA to perform WES.

269   Accordingly, WES was performed for only nine intramucosal DGCs (Table 2). The

270   median amount of extracted DNA was 67 ng (range, 10–309 ng). A total of 239

271   mutations (median, 20 mutations/exome; range, 10–66) were found (Table S2). The

272   most recurrently mutated gene was *CDH1*; *CDH1* mutations were found in six of nine

273   intramucosal DGCs (67%) and were considered as positively selected by dN/dS

274   analysis ($q < 0.001$). All six cases with *CDH1* mutations had 'pure' SRCCs, which were

275   near the gastric angle. For the other genes, dN/dS analysis showed no significant

276   results ($q > 0.05$). However, a search for 69 previously reported driver genes revealed

277   mutations in three genes, *RNF43*, *TGFBR2*, and *FAT4* (Fig. 3a, Tables S1–S3). All

278   nine cases of *Hp*-uninfected intramucosal DGC had two or fewer driver mutations.

279   *CDH1* mutation is generally known as a loss-of-function mutation [30]. A missense

280   mutation at L581R and loss of heterozygosity were found in Case 2 (Fig. 3a). A

281   missense mutation at L581 was previously reported in a sporadic DGC [20]. Two-hit

282   *CDH1* mutations, including truncating mutations, were observed in Cases 5 and 6 (Fig.

283   3a). Cases 3 and 5 harbored a mutation at I250 (L249_T253del, I250N, respectively),

284   which was reported as a mutation hotspot in sporadic DGCs and resulted in impaired

285   cellular aggregation *in vitro* [21] (Fig. 3b). Hotspot splice site mutations (c.531+2T>A,

286   c.687+1_687+4del) reported in sporadic DGCs were observed in Cases 4 and 7 [21,

287   24] (Fig. 3b). These splice site mutations are thought to induce exon truncations in

288   extracellular domain 1 to prevent the homodimerization of E-cadherin according to

289   computer models [21]. Case 13, the only case with PDA and SRCC located at the

290   cardia, did not harbor *CDH1* mutations. Among the three cases without *CDH1*

291   mutations, we detected a frameshift *RNF43* mutation (A146fs) in Case 10, but we

292   detected no driver mutations in Cases 12 and 13. We also searched for germline

293   mutations in *CDH1*, *CTNNA1*, *PALB2*, *BRCA1*, and *RAD51C*, all of which had been

294   reported as causative genes of HDGC using blood DNA [19, 31, 32]. None of the nine

295   subjects possessed pathogenic germline mutations in these genes. These results

296   indicate that somatic *CDH1* mutations can be considered as an early event in

297 intramucosal SRCC carcinogenesis. E-cadherin expression was

298 immunohistochemically reduced in all eight cases with intramucosal 'pure' SRCCs

299 near the gastric angle but was maintained in DGC without *CDH1* mutations at the

300 gastric cardia (Figs. 1a, b). These data suggest that reduced E-cadherin expression

301 contributes to SRCC development near the gastric angle, whereas the one case of

302 PDA/SRCC at the gastric cardia may have been caused by different mechanisms.

303

Analysis of advanced DGCs

305 Genomic DNA extracted from laser capture microdissected SRCC cells in the mucosal

306 part of two advanced DGCs was subjected to WES. For technical reasons, some PDA

307 cells were included. *A TP53* missense mutation at R248W and loss of heterozygosity

308 was found in Case 11 (Fig. 3a, Table S4). Four driver mutations in *CDH1*, *TGFBR2*,

309 *ARID1A*, and *RHOA* were found in Case 14 (Fig. 3a, Tables S3, S4). The *CDH1*

310 mutation at D402V has been reported in sporadic DGCs [20, 33] (Fig. 3b). Neither of

311 the two cases had pathogenic germline mutations in the HDGC-causative genes

312 described above.

313 To explore the differences in gene alternations between *Hp*-uninfected intramucosal

314 and advanced DGCs, we examined nine intramucosal DGCs and 60 advanced DGCs,

315 including two from our cohort and 58 from external datasets (Table S5). *CDH1*

316   mutations were more frequent in intramucosal DGCs (6 of 9, 67%) than in advanced

317   DGCs (16 of 60, 27%) ($p = 0.02$, $q = 0.11$, Fig 4a). In contrast, *TP53* mutations were

318   significantly more frequent in advanced DGCs (31 of 60, 52%) than in intramucosal

319   DGCs (0 of 9, 0%) ($p < 0.01$, $q = 0.03$). *CDH1* and *TP53* co-mutated DGC occurred in

320   only 10% (6 of 60) of advanced DGC (Fig. 4b). *CDH1*-mutated/*TP53* wild-type

321   advanced DGC accounted for 17% (10 of 60) and *TP53*-mutated/*CDH1* wild-type

322   advanced DGC accounted for 42% (25 of 60) of total advanced DGCs. The difference

323   of *CDH1* and *TP53* mutation frequencies between intramucosal and advanced DGCs

324   suggests that *Hp*-uninfected advanced DGCs may be more likely develop from

325   precursor lesions other than *CDH1*-mutated intramucosal DGCs.

326

327   **Comparison of *Hp*-uninfected normal gastric mucosa, intramucosal DGC, and**

328   **advanced DGC**

329   To investigate the mechanism of *Hp*-uninfected intramucosal DGC development, we

330   compared the mutations and mutational signatures between intramucosal DGCs and

331   normal gastric glands. WES also indicated that mutations accumulated with age at a

332   frequency of 0.33 mutations/exome/year in the normal gastric gland (Fig. 5a, Table

333   S6). The number of mutations in intramucosal DGCs was not significantly higher than

334   that in normal glands when the age of the patient was considered ($p = 0.40$, Fig. 5a).

18

335   In contrast, advanced DGCs had a significantly higher number of mutations than

336   intramucosal DGCs and normal glands (p = 0.001, p < 0.001, respectively, Fig. 5a).

337   Next, we performed mutational signature analysis, as different mutational processes

338   may contribute to the accumulation of mutations in a cell, with each imprinting a

339   mutational signature on the cell genome. COSMIC SBSs 1, 5, and 40 accounted for

340   the majorities of normal glands and intramucosal DGCs (Fig. 5b). SBS 1 is known as

341   an age-related signature [34]. The underlying mechanism of SBS 1 is likely

342   deamination of 5-methylcytosine at CpG sites, leading to a C>T transition. The relative

343   contributions of SBS 1 did not significantly differ between the normal gland and

344   intramucosal DGC (p = 0.40). In addition, the number of mutations allocated to SBS 1

345   did not differ between the normal gland and intramucosal DGC when considering the

346   age of the patient (p = 0.92). SBS 5 is characterized by C>T and T>C transitions [34].

347   SBS 40 is similar to SBS 5, making it difficult to estimate their separate contributions

348   [13, 29]. Therefore, we counted SBS 5 and SBS 40 together (designated as SBS 5/40).

349   SBS 5/40 also accumulates with age, although the underlying mutational processes

350   are not well understood [34]. The contribution of SBS 5/40 tended to be higher in

351   intramucosal DGCs than in normal glands (p = 0.07, Fig. 5c). This trend did not change

352   when the number of mutations allocated to SBS 5/40 was compared while considering

353   the age of the patient (p = 0.08, Fig. 5d). This difference in mutational signature spectra

19

354 suggests that intramucosal DGC and normal glands undergo different mutational

355 processes despite similar numbers of mutations.

356   To further investigate the mechanism underlying the progression from intramucosal

357 DGC to advanced DGC, we compared the mutational signatures. The contribution of

358 SBS 5/40 was lower and that of SBS 17b was higher in advanced DGCs than in

359 intramucosal DGCs (Fig. 5b). Although the etiology of SBS 17b is not well understood,

360 its possible link to damage inflicted by reactive oxygen species has been reported [35].

361 This result suggests that other or additional mutational processes are involved in the

362 development of advanced DGCs compared to intramucosal DGCs.

363

364 **Discussion**

365 In the present study, we found that 67% (6 of 9) of *Hp*-uninfected intramucosal DGCs

366 harbored *CDH1* somatic mutations, and the six DGCs with *CDH1* somatic mutations

367 were histologically 'pure' SRCCs without PDA components. *Hp*-uninfected

368 intramucosal DGCs harbored similar numbers of mutations as normal gastric glands

369 and few driver mutations other than those in *CDH1*. These results suggest that *CDH1*

370 somatic mutations are driver mutations for the development of *Hp*-uninfected 'pure'

371 SRCCs.

372   The median Ki-67 index of 'pure' SRCCs was 2.4%, which is similar to that in a

373   previous report, indicating their low proliferative capacity [16]. Interestingly, five cases

374   with 'pure' SRCCs whose endoscopic images could be reviewed showed no

375   progression over 2–5 years. These results suggest that *Hp*-uninfected 'pure' SRCCs

376   have an indolent feature that makes them less likely to become invasive. Yorita et al.

377   also reported in a retrospective study that *Hp*-uninfected early SRCCs were less likely

378   to be invasive cancers [17]. Furthermore, all 15 'pure' SRCCs, including those that

379   were not genetically analyzed, were located near the gastric angle and characterized

380   by reduced E-cadherin expression. In animal studies, the loss of E-cadherin function

381   due to *Cdh1* deficiency resulted in the development of SRCC-like atypical cells, but it

382   was not sufficient for invasive cancer formation [26, 30]. Factors such as chronic

383   stimulation by *Helicobacter* infection or nitroso compounds and *Trp53* mutations are

384   required for the invasion of SRCC in *Cdh1* knockout mice [26, 30, 36, 37]. The present

385   study showed that 10% of advanced DGCs harbored both *CDH1* and *TP53* mutations

386   in contrast to none of the *CDH1*-mutated intramucosal SRCCs harboring *TP53*

387   mutation. These results suggest that some *CDH1*-mutated intramucosal SRCCs

388   become advanced DGCs with the addition of *TP53* mutations, as shown in animal

389   experiments [26, 36]. This hypothesis is supported by a study showing that p53 is not

390   aberrantly expressed in intramucosal HDGCs, whereas its expression is altered in

391   invasive HDGCs [38]. However, the discrepancy in the *CDH1* and *TP53* mutation

392   frequencies between intramucosal and advanced DGC suggests that *CDH1*-mutated

393   intramucosal DGCs make a relatively small contribution to advanced DGC formation

394   in the *Hp*-uninfected stomach and that unknown *TP53*-mutated precancerous lesions

395   might exist.

396   In addition to *Hp* infection, smoking has been reported to increase the risk of DGC

397   compared to in non-smokers [39, 40]. Previous studies reported that the prevalence

398   of smoking was high in patients with *Hp*-uninfected DGCs [41, 42]. However, SBS 4,

399   the signature associated with smoking, was not observed in the present study (Fig.

400   5b), which is consistent with the findings of a recent genetic study of DGC [43]. Further

401   analysis is required to clarify the carcinogenic effect of smoking on DGC development.

402   Recent studies reported alcohol-related mutagenesis in gastric cancers in patients

403   with alcohol consumption habit and *ALDH2*-defective alleles [12, 43]. However, SBS

404   16, an alcohol-related mutational signature, was not detected in our cohort (Fig. 5b),

405   possibly because of the small number of mutations detected by WES analysis (Table

406   2, Fig. 3a).

407   We confirmed that somatic mutations accumulate with age at a frequency of 0.33

408   mutation/exome/year in the *Hp*-uninfected normal gastric epithelia. This mutation

409   frequency is lower than that in the normal colon, as previously determined by WES

410   [11]. The predominant mutational signature of normal gastric epithelia without *Hp*

411   infection was SBS 5/40, followed by SBS 1. This result differs from that observed for

412   the normal colon and small intestine, in which SBS 1 is most predominant [11, 14].

413   This study had several limitations. First, the cohort size was small. Second, the size

414   of intramucosal DGC lesion in this study was small. The small lesion size may have

415   contributed to the very slow disease progression, such as that observed for SRCC foci

416   found in HDGC [44]. Further studies are needed to evaluate larger-sized early DGCs.

417   Third, even cases without *CDH1* mutation showed reduced E-cadherin expression.

418   We could not determine the mechanism by which E-cadherin expression was

419   attenuated despite the absence of *CDH1* mutations and copy number alternations.

420   Although we performed *CDH1* promoter methylation analysis, we failed to obtain

421   informative results. The methylation rates in intramucosal DGCs were generally low,

422   although they tended to be higher than those in the adjacent normal mucosa (data not

423   shown). Recent genetic analysis of *Hp*-uninfected DGC reported that *CDH1* mutations

424   were found in only one of seven cases, in contrast to the high prevalence of *CDH1*

425   mutations observed in our study [42]. The low amount of input DNA (10 ng in every

426   case) and the long period of paraffin embedding (median 6 years) may have affected

427   the quality of the analysis, resulting in a low frequency of detection of *CDH1* mutations

428   [45].

429   In conclusion, we demonstrated that *CDH1* somatic mutations are positively

430 selected as driver mutations in *Hp*-uninfected sporadic intramucosal SRCCs and that

431 *Hp*-uninfected intramucosal SRCCs may progress slowly and be less likely to develop

432 into advanced DGCs.

433

444

445 **Disclosure of conflicts of interest**: The authors declare no potential conflicts of

446 interest.

447   **References**

448   1.   Parsonnet J, Friedman GD, Vandersteen DP, Chang Y, Vogelman JH,

449   Orentreich N, et al. Helicobacter pylori infection and the risk of gastric

450   carcinoma. N Engl J Med 1991;325:1127-1131.

451   2.   Nomura A, Stemmermann GN, Chyou PH, Kato I, Perez-Perez GI, Blaser MJ.

452   Helicobacter pylori infection and gastric carcinoma among Japanese

453   Americans in Hawaii. N Engl J Med 1991;325:1132-1136.

454   3.   Shimizu T, Marusawa H, Matsumoto Y, Inuzuka T, Ikeda A, Fujii Y, et al,

455   Accumulation of somatic mutations in TP53 in gastric epithelium with

456   Helicobacter pylori infection. Gastroenterology 2014;147:407-417.

457   4.   Matsumoto Y, Marusawa H, Kinoshita K, Endo Y, Kou T, Morisawa T, et al.

458   Helicobacter pylori infection triggers aberrant expression of activation-induced

459   cytidine deaminase in gastric epithelium. Nat Med 2007;13:470-476.

460   5.   Yamamoto Y, Fujisaki J, Omae M, Hirasawa T, Igarashi M. Helicobacter pylori-

461   negative gastric cancer: characteristics and endoscopic findings. Dig Endosc

462   2015;27:551-561.

463   6.   Zamani M, Ebrahimtabar F, Zamani V, Miller WH, Alizadeh-Navaei R, Shokri-

464   Shirvani J, et al. Systematic review with meta-analysis: the worldwide

465   prevalence of Helicobacter pylori infection. Aliment Pharmacol Ther

466   2018;47:868-876.

467   7.   Kamada T, Haruma K, Ito M, Inoue K, Manabe N, Matsumoto H, et al.   Time

468   Trends in Helicobacter pylori Infection and Atrophic Gastritis Over 40 Years in

469   Japan. Helicobacter 2015;20:192-198.

470   8.   Inoue M. Changing epidemiology of Helicobacter pylori in Japan. Gastric

471   Cancer 2017;20:3-7.

472   9.   Bertuccio P, Chatenoud L, Levi F, Praud D, Ferlay J, Negri E, et al. Recent

473   patterns in gastric cancer: a global overview. Int J Cancer 2009;125:666-673.

474   10.   Jaiswal S, Fontanillas P, Flannick J, Manning A, Grauman PV, Mar BG, et al.

475   Age-related clonal hematopoiesis associated with adverse outcomes. N Engl J

476   Med 2014;371:2488-2498.

477   11.   Kakiuchi N, Yoshida K, Uchino M, Kihara T, Akaki K, Inoue Y, et al. Frequent

478   mutations that converge on the NFKBIZ pathway in ulcerative colitis. Nature

479   2020;577:260-265.

480   12.   Yokoyama A, Kakiuchi N, Yoshizato T, Nannya Y, Suzuki H, Takeuchi Y, et al.

481   Age-related remodelling of oesophageal epithelia by mutated cancer drivers.

482   Nature 2019;565:312-317.

483   13.   Moore L, Leongamornlert D, Coorens THH, Sanders MA, Ellis P, Dentro SC, et

484   al. The mutational landscape of normal human endometrial epithelium. Nature

485   2020;580:640-646.

486   14.   Blokzijl F, de Ligt J, Jager M, Sasselli V, Roerink S, Sasaki N, et al. Tissue-

487   specific mutation accumulation in human adult stem cells during life. Nature

488   2016;538:260-264.

489   15.   Lauren P. The two histological main types of gastric carcinoma: diffuse and so-

490   called intestinal-type carcinoma. An attempt at a histo-clinical classification.

491   Acta Pathol Microbiol Scand 1965;64:31-49.

492   16.   Horiuchi Y, Fujisaki J, Yamamoto N, Shimizu T, Miyamoto Y, Tomida H, et al.

493   Biological   behavior   of   the   intramucosal   Helicobacter   pylori-negative

494   undifferentiated-type early gastric cancer: comparison with Helicobacter pylori-

495   positive early gastric cancer. Gastric Cancer 2016;19:160-165.

496   17.   Yorita N, Ito M, Boda T, Kotachi T, Nagasaki N, Abuduwaili M, et al. Potential of

497   Helicobacter pylori-uninfected signet ring cell carcinoma to invade the

498   submucosal layer. J Gastroenterol Hepatol 2019;34:1955-1962.

499   18.   Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, et al. E-

500   cadherin germline mutations in familial gastric cancer. Nature 1998;392:402-

501   405.

502   19.   van der Post RS, Vogelaar IP, Carneiro F, Guilford P, Huntsman D,

503   Hoogerbrugge N, et al. Hereditary diffuse gastric cancer: updated clinical

504   guidelines with an emphasis on germline CDH1 mutation carriers. J Med Genet

505   2015;52:361-374.

506   20.   Machado JC, Oliveira C, Carvalho R, Soares P, Berx G, Caldas C, et al. E-

507   cadherin gene (CDH1) promoter methylation as the second hit in sporadic

508   diffuse gastric carcinoma. Oncogene 2001;20:1525-1528.

509   21.   Cho SY, Park JW, Liu Y, Park YS, Kim JH, Yang H, et al. Sporadic Early-Onset

510   Diffuse Gastric Cancers Have High Frequency of Somatic CDH1 Alterations,

511   but Low Frequency of Somatic RHOA Mutations Compared With Late-Onset

512   Cancers. Gastroenterology 2017;153:536-549.

513   22.   Wang K, Yuen ST, Xu J, Lee SP, Yan HH, Shi ST, et al. Whole-genome

514   sequencing and comprehensive molecular profiling identify new driver

515   mutations in gastric cancer. Nat Genet 2014;46:573-582.

516   23.   Network CGAR. Comprehensive molecular characterization of gastric

517   adenocarcinoma. Nature 2014;513:202-209.

518   24.   Kakiuchi M, Nishizawa T, Ueda H, Gotoh K, Tanaka A, Hayashi A, et al.

519   Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma.

520   Nat Genet 2014;46:583-587.

521   25.   Association JGC. Japanese classification of gastric carcinoma: 3rd English

522   edition. Gastric Cancer 2011;14:101-112.

523   26.   Hayakawa Y, Ariyama H, Stancikova J, Sakitani K, Asfaha S, Renz BW, et al.

524   Mist1 Expressing Gastric Stem Cells Maintain the Normal and Neoplastic

525   Gastric Epithelium and Are Supported by a Perivascular Stem Cell Niche.

526   Cancer Cell 2015;28:800-814.

527   27.   Yoshida K, Sanada M, Shiraishi Y, Nowak D, Nagata Y, Yamamoto R, et al.

528   Frequent pathway mutations of splicing machinery in myelodysplasia. Nature

529   2011;478:64-69.

530   28.   Martincorena I, Raine KM, Gerstung M, Dawson KJ, Haase K, Van Loo P, et al.

531   Universal Patterns of Selection in Cancer and Somatic Tissues. Cell

532   2017;171:1029-1041.

533   29.   Alexandrov LB, Kim J, Haradhvala NJ, Huang MN, Tian Ng AW, Wu Y, et al.

534   The repertoire of mutational signatures in human cancer. Nature 2020;578:94-

535   101.

536   30.   Mimata A, Fukamachi H, Eishi Y, Yuasa Y. Loss of E-cadherin in mouse gastric

537   epithelial cells induces signet ring-like cells, a possible precursor lesion of

538   diffuse gastric cancer. Cancer Sci 2011;102:942-950.

539   31.   Majewski IJ, Kluijt I, Cats A, Scerri TS, de Jong D, Kluin RJ, et al. An α-E-catenin

540   (CTNNA1) mutation in hereditary diffuse gastric cancer. J Pathol 2013;229:621-

541   629.

542   32.   Sahasrabudhe R, Lott P, Bohorquez M, Toal T, Estrada AP, Suarez JJ, et al.

543   Germline Mutations in PALB2, BRCA1, and RAD51C, Which Regulate DNA

544   Recombination Repair, in Patients With Gastric Cancer. Gastroenterology

545   2017;152:983-986.

546   33.   Kuboki Y, Yamashita S, Niwa T, Ushijima T, Nagatsuma A, Kuwata T, et al.

547   Comprehensive   analyses   using   next-generation   sequencing   and

548   immunohistochemistry enable precise treatment in advanced gastric cancer.

549   Ann Oncol 2016;27:127-133.

550   34.   Alexandrov LB, Jones PH, Wedge DC, Sale JE, Campbell PJ, Nik-Zainal S, et

551   al. Clock-like mutational processes in human somatic cells. Nat Genet

552   2015;47:1402-1407.

553   35.   Tomkova M, Tomek J, Kriaucionis S, Schuster-Böckler B. Mutational signature

554   distribution varies with DNA replication timing and strand asymmetry. Genome

555   Biol 2018;19:129. doi: 10.1186/s13059-018-1509-y.

556   36.   Shimada S, Mimata A, Sekine M, Mogushi K, Akiyama Y, Fukamachi H, et al.

557   Synergistic tumour suppressor activity of E-cadherin and p53 in a conditional

558   mouse model for metastatic diffuse-type gastric cancer. Gut 2012;61:344-353.

559   37.   Humar B, Blair V, Charlton A, More H, Martin I, Guilford P. E-cadherin deficiency

560   initiates gastric signet-ring cell carcinoma in mice and man. Cancer Res

561   2009;69:2050-2056.

562   38.   van der Post RS, Gullo I, Oliveira C, Tang LH, Grabsch HI, O'Donovan M, et al.

563   Histopathological, Molecular, and Genetic Profile of Hereditary Diffuse Gastric

564   Cancer: Current Knowledge and Challenges for the Future. Adv Exp Med Biol

565   2016;908:371-391.

566   39.   Ye W, Ekström AM, Hansson LE, Bergström R, Nyrén O. Tobacco, alcohol and

567   the risk of gastric cancer by sub-site and histologic type. Int J Cancer

568   1999;83:223-229.

569   40.   Koizumi Y, Tsubono Y, Nakaya N, Kuriyama S, Shibuya D, Matsuoka H, et al.

570   Cigarette smoking and the risk of gastric cancer: a pooled analysis of two

571   prospective studies in Japan. Int J Cancer 2004;112:1049-1055.

572   41.   Horiuchi Y, Fujisaki J, Ishizuka N, Omae M, Ishiyama A, Yoshio T, et al. Study

573   on Clinical Factors Involved in Helicobacter pylori-Uninfected, Undifferentiated-

574   Type Early Gastric Cancer. Digestion 2017;96:213-219.

575   42.   Kiso M, Urabe Y, Ito M, Masuda K, Boda T, Kotachi T, et al. Clinical and genomic

576   characteristics of mucosal signet-ring cell carcinoma in Helicobacter pylori-

577   uninfected stomach. BMC Gastroenterol 2020; 20:243. doi: 10.1186/s12876-

578   020-01387-9.

579   43.   Suzuki A, Katoh H, Komura D, Kakiuchi M, Tagashira A, Yamamoto S, et al.

580   Defined lifestyle and germline factors predispose Asian populations to gastric

581   cancer. Sci Adv 2020;6:eaav9778. doi: 10.1126/sciadv.aav9778.

582 44. Guilford P, Humar B, Blair V. Hereditary diffuse gastric cancer: translation of

583   CDH1 germline mutations into clinical practice. *Gastric Cancer* 2010,13:1-10.

584 45.   Do H, Dobrovic A. Sequence artifacts in DNA from formalin-fixed tissues:

585   causes and strategies for minimization. Clin Chem 2015;61:64-71.

Figure Legends

Figure 1 Clinicopathological features of 18 *Hp*-uninfected DGCs

(a) Tumor locations and pathological features. Sixteen cases of intramucosal DGC and two cases of advanced DGC were sampled. The red dots represent intramucosal DGCs. Fifteen of sixteen intramucosal DGCs were concentrated near the gastric angle and were 'pure' SRCCs with reduced E-cadherin immunoreactivity. One cardiac intramucosal DGC consisted of SRCC and PDA components with maintained E-cadherin immunoreactivity. Green ovals represent advanced DGCs. Advanced DGCs consisted of SRCC and PDA components with reduced E-cadherin immunoreactivity.

(b) Representative endoscopic and histological images of intramucosal DGC cases. Upper panels: Images of Case 5. A pale flat lesion was observed at the anterior wall of the gastric antrum, close to the gastric angle, upon endoscopy (yellow arrowhead). H&E staining of the endoscopic submucosal dissection specimen showed that SRCCs were confined to the proliferative zone of the mucosa. Immunohistochemistry for E-cadherin showed weak immunoreactivity in SRCCs. Immunohistochemistry for Ki-67 showed only a few SRCCs positive for nuclear staining.

Lower panels: Images of Case 13. A pale flat lesion was observed at the greater curvature of the gastric cardia upon endoscopy (yellow arrowhead). H&E staining of the endoscopic submucosal dissection specimen showed SRCCs and PDAs in the

mucosa. Immunohistochemistry for E-cadherin showed almost the same immunoreactivity in carcinoma cells as in surrounding epithelial cells. Immunohistochemistry for Ki-67 showed that all carcinoma cells were negative for nuclear staining.

Scale bar indicates 50 µm.

Figure 2 Changes in endoscopic findings in *Hp*-uninfected intramucosal signet ring cell carcinoma over time

Top panels: Endoscopic images of Case 8. A biopsy from a tiny pale area at the greater curvature of the gastric angle revealed an SRCC in the year 2018. After biopsy, the lesion became indistinct. In 2020, the lesion became visible, and the size of the lesion was almost the same as that observed in 2018.

The second, third, fourth and fifth panels: Endoscopic images of Cases 10, 16, 17, and 18 respectively. Upon retrospective review of the endoscopic images, SRCCs were found that remained unchanged for 5, 3, 2, and 3 years, respectively. Yellow arrowheads show the lesions. The number in the lower right-hand corner represents the year in which the image was taken.

Figure 3 Mutational landscape and *CDH1* mutation sites

624 (a) Mutational landscape of *Hp*-uninfected nine intramucosal DGCs and two advanced

625 DGCs. Tumor invasion depth, location, histological type, and E-cadherin

626 immunoreactivity are indicated. LOH, loss of heterozygosity.

627 (b) *CDH1* mutation sites. Alterations in our sequenced tumor samples are plotted

628 above the *CDH1* protein. Somatic mutations in sporadic DGCs that have been

629 repeatedly reported in previous studies are plotted below the *CDH1* protein [Refs. 20,

630 21, 24, 33].

631 Sig, signal peptide; Precursor, precursor sequence; EC, extracellular domain; TM,

632 transmembrane domain; Cytoplasmic, cytoplasmic domain; # reported as a germline

633 mutation of HDGC.

634

635 **Figure 4 Driver gene mutation frequencies in *Hp*-uninfected intramucosal and**

636 **advanced DGCs**

637 (a) Comparison of mutation frequencies in driver genes between *Hp*-uninfected

638 intramucosal DGCs (*n* = 9) and advanced DGCs (*n* = 60). *p* and *q* values were

639 provided for *CDH1* and *TP53* calculated by two-tailed Fisher's exact test with

640 Benjamini-Hochberg adjustment. There were no significant differences in other genes.

641 (b) Detailed *CDH1/TP53* mutation profiles for *Hp*-uninfected intramucosal and

642 advanced DGCs.

643

Figure 5 Mutation number and mutational signature analyses of *Hp*-uninfected normal gastric glands, intramucosal DGCs, and advanced DGCs

(a) Correlation between age and number of mutations in *Hp*-uninfected normal glands, intramucosal DGCs and advanced DGCs. The number of mutations in single glands from normal gastric epithelium without *Hp* infection was plotted against patient age (n = 18, blue dots). A regression line (blue dotted line) assuming an intercept of zero is shown, with R^2 and coefficient values. Orange and gray dots indicate the number of mutations in intramucosal and advanced DGCs respectively. The two-tailed Mann–Whitney U test for comparison between normal glands and intramucosal DGCs showed no significant difference (p = 0.40). In contrast, advanced DGCs had a significantly higher number of mutations than intramucosal DGCs and normal glands (p = 0.001, p < 0.001, respectively).

(b) Relative contribution of COSMIC SBSs in *Hp*-uninfected normal glands, intramucosal DGCs, and advanced DGCs.

(c) Proportion of mutations allocated to SBS 5/40 in intramucosal DGCs and normal glands. SBS 5/40 in intramucosal DGCs was relatively higher than that in normal glands (p = 0.07, two-tailed Mann–Whitney U test).

(d) Correlation between age and number of mutations allocated to SBS 5/40 in normal

662   glands (blue dots) and intramucosal DGCs (orange dots). Blue dotted line shows

663   regression line assuming an intercept of zero for normal glands. Intramucosal DGCs

664   showed a relatively larger number of mutations than normal glands (p = 0.08, two-

665   tailed Mann–Whitney U test).

666

Table 1. Patient characteristics

	Intramucosal DGC (n=16)	Advanced DGC (n=2)	Normal gland (n=9)
Sex			
Male	9	1	6
Female	7	1	3
Age, years, median (range)	57 (29-78)	44.5 (37-52)	44 (28-77)
Tumor location			
Cardia	1	0	N/A
Body	3	1	N/A
Antrum	12	1	N/A
Size of tumor, mm, median (range)	6.5 (3-14)	97.5 (50-145)	N/A
Treatment			N/A
Endoscopic resection	16	0	N/A
Gastrectomy	0	1	N/A
Chemotherapy	0	1	N/A
Histology			
Pure SRCC	15	0	N/A
PDA/SRCC	1	2	N/A
Depth of invasion*			
M	16	0	N/A
SE	0	2	N/A
Cases met HDGC testing criteria	6	1	N/A
Cases subjected to WES	9	2	9

DGC, diffuse-type gastric cancer; SRCC, signet ring cell carcinoma; PDA, poorly differentiated carcinoma; HDGC, hereditary diffuse gastric cancer; WES, whole exome sequencing; N/A, not applicable

* According to the third English edition of Japanese classification of gastric carcinoma: M, Tumor confined to the mucosa; SE, Tumor invasion is contiguous to or exposed beyond the serosa

Table 2. Details of DGC patients

Case number	Age	Sex	Family history of GC[a]	HDGC testing criteria	Location	Histology	Size	Depth of invasion[b]	E-cadherin expression	Ki-67 index	WES	Smoking history	Alcohol intake habit[c]	ALDH2 allele deficiency
1	48	M	no		Antrum	SRCC	3	M	reduced	0.0%	N/A	former	light	N/A
2	61	M	no		Antrum	SRCC	13	M	reduced	1.8%	○	current	light	no
3	63	M	yes	○	Antrum	SRCC	14	M	reduced	3.4%	○	current	heavy	hetero
4	56	F	no		Antrum	SRCC	10	M	reduced	0.7%	○	never	never/rare	no
5	73	M	no		Antrum	SRCC	7	M	reduced	0.0%	○	former	moderate	no
6	50	M	no		Antrum	SRCC	12	M	reduced	3.7%	○	former	light	hetero
7	40	M	yes	○	Antrum	SRCC	6	M	reduced	2.9%	○	current	never/rare	homo
8	65	F	no		Antrum	SRCC	4	M	reduced	15.4%	N/A	former	never/rare	N/A
9	55	F	yes	○	Body	SRCC[d]	4	M[d]	N/A	N/A	N/A	never	light	N/A
10	51	M	no		Body	SRCC	5	M	reduced	1.6%	○	current	light	hetero
11	52	F	no		Body	PDA/SRCC	145	SE	reduced	1.0%	○	current	never/rare	no
12	29	F	no	○	Antrum	SRCC	8	M	reduced	0.9%	○	never	never/rare	no
13	61	F	unknown		Caridia	PDA/SRCC	6	M	maintained	0.0%	○	unknown	unknown	no
14	37	M	no	○	Antrum	PDA/SRCC	50	SE	reduced	0.0%	○	never	never/rare	no
15	63	F	no		Body	SRCC	9	M	reduced	9.8%	N/A	never	never/rare	N/A
16	58	M	no		Antrum	SRCC	6	M	reduced	4.4%	N/A	former	never/rare	N/A
17	71	M	yes	○	Antrum	SRCC	6	M	reduced	1.9%	N/A	current	heavy	N/A
18	78	F	yes	○	Antrum	SRCC	6	M	reduced	5.7%	N/A	never	light	N/A

GC, Gastric cancer; HDGC, hereditary diffuse gastric cancer; SRCC, signet ring cell carcinoma; PDA, poorly differentiated carcinoma; N/A, not applicable
[a] Family history of GC at any age in first- or second-degree relatives; [b] According to the third English edition of Japanese classification of gastric carcinoma: M, Tumor confined to the mucosa; SE, Tumor invasion is contiguous to or exposed beyond the serosa; [c] Never/rare, <1U ethanol/week; light, 1-8.9 U/week; moderate, 9-18 U/week; heavy, >18 U/week (1 U =22g ethanol); [d] The biopsy specimen showed SRCC, but the endoscopic resection specimen showed no evidence of tumor, suggesting that the lesion was intramucosal and was completely removed by biopsy.

Figure1

a
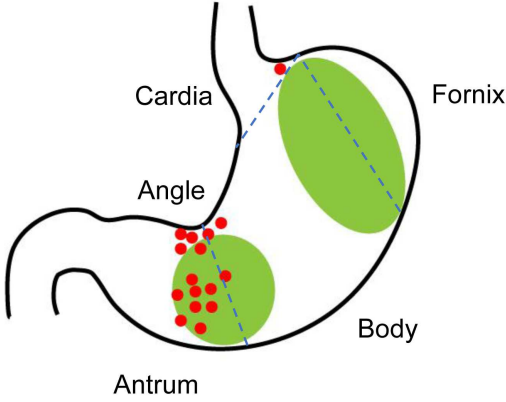
Intramucosal DGC: 16 cases

Cardia : 1 case, PDA/SRCC, E-cadherin maintained
Body : 3 cases, SRCC, E-cadherin reduced
Antrum : 12 cases, , SRCC, E-cadherin reduced

Advanced DGC: 2 cases

Body : 1 case, PDA/SRCC, E-cadherin reduced
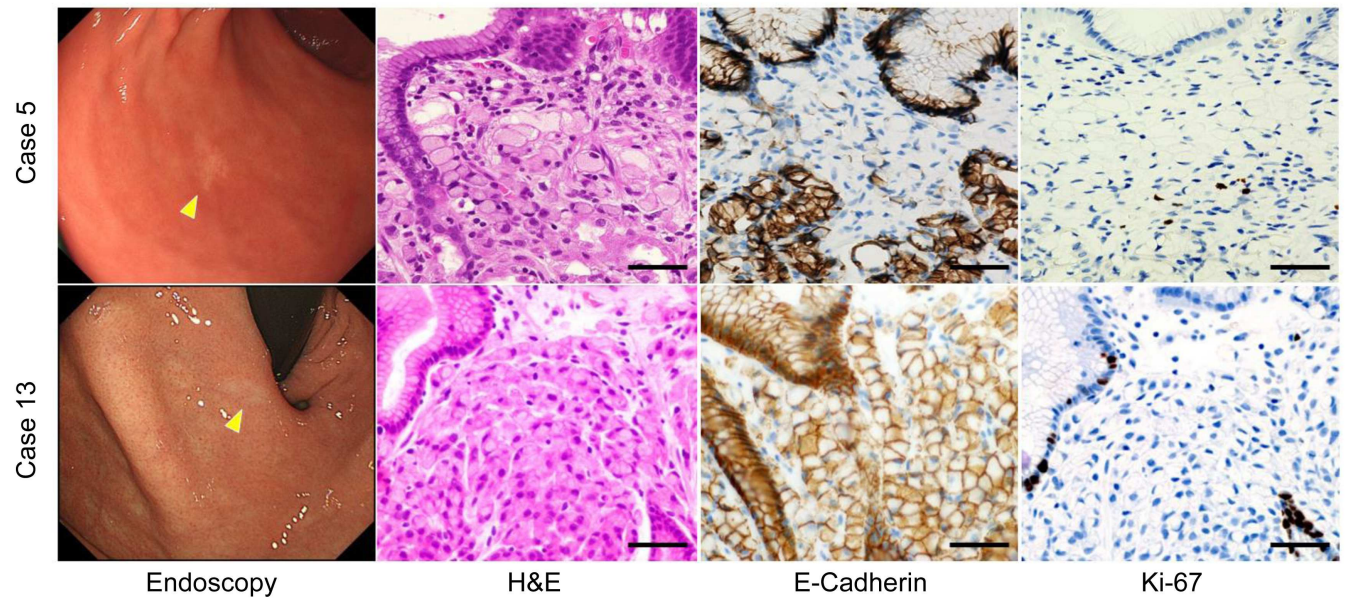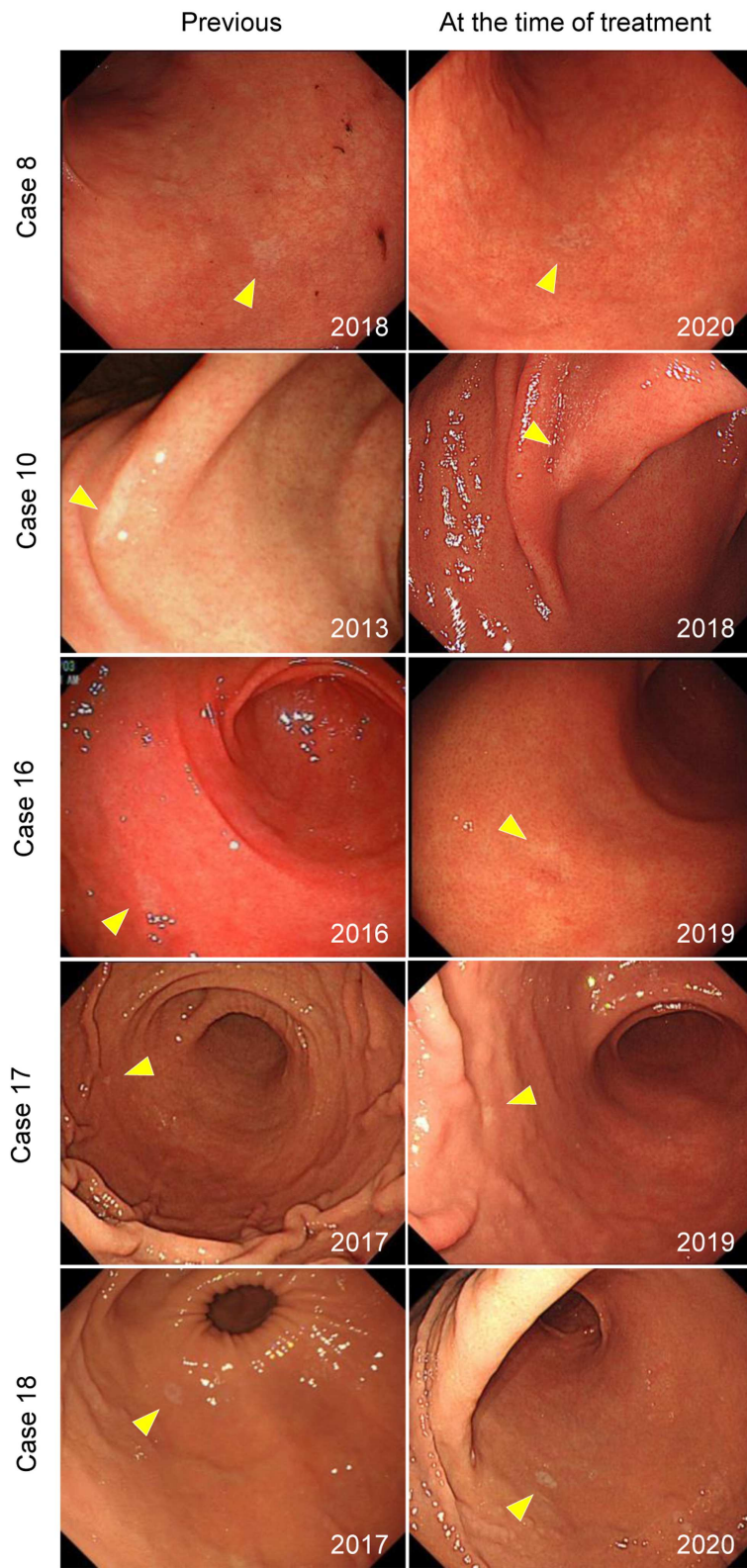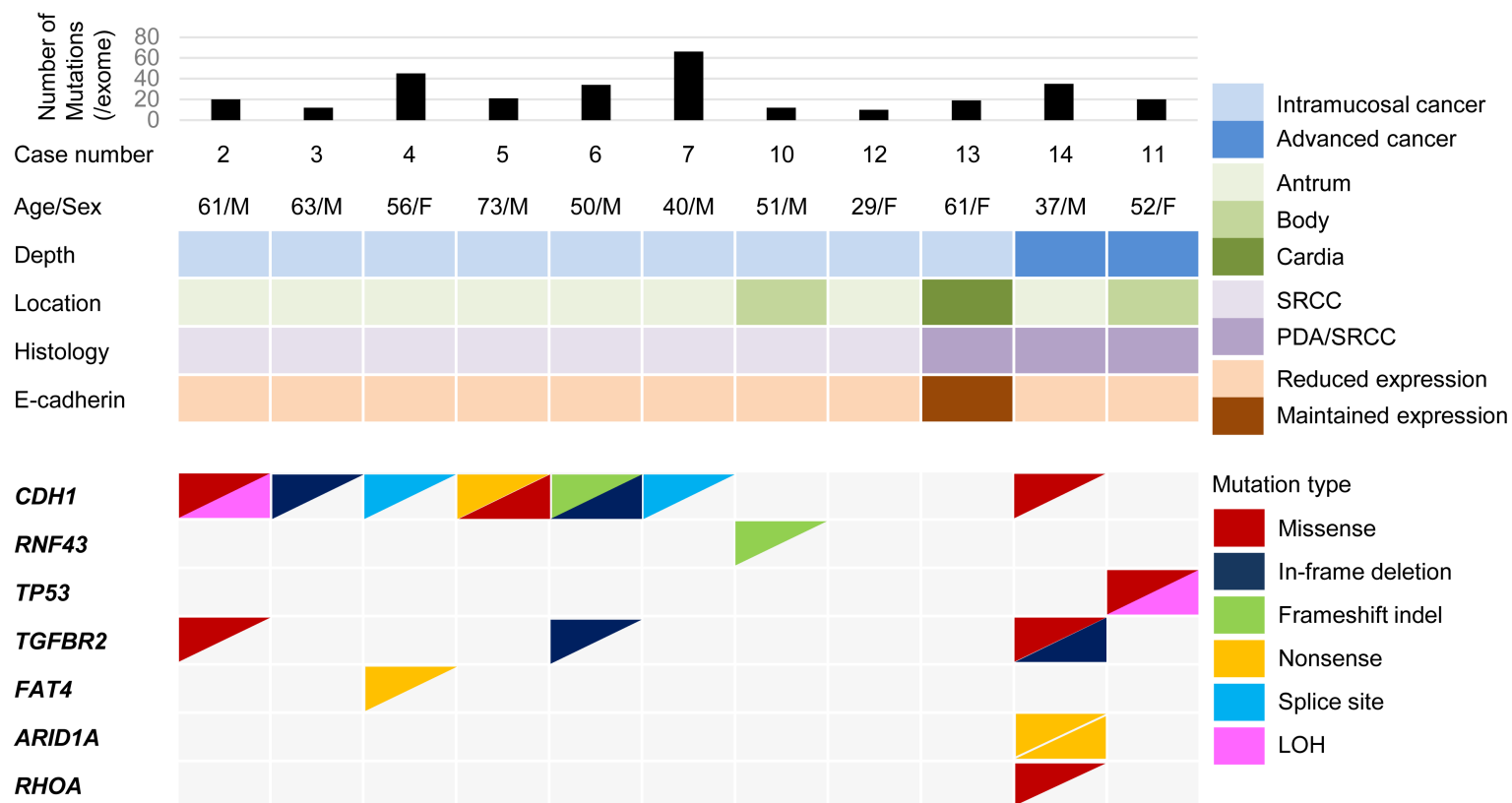Antrum : 1 case, PDA/SRCC, E-cadherin reduced

b

| | Endoscopy | H&E | E-Cadherin | Ki-67 |

Figure2

Figure3

Figure4

a

p = 0.02
q = 0.11

p < 0.01
q = 0.03

67%

27%

22%

7%

11% 12%

11%

2%

52%

0%

0%

13%

0%

12%

0%

12%

0%

12%

CDH1   *TGFBR2*   *FAT4*   *RNF43*   *TP53*   *SPTA*   *ARID1A*   *RHOA*   *GLI3*

■ Intramucosal DGC   ■ Advanced DGC

b

■ CDH1/TP53 wild-type
■ CDH1 wild-type/TP53-mutated
■ CDH1-mutated/TP53 wild-type
■ CDH1/TP53 co-mutated

Intramucosal DGC   Advanced DGC

Figure5

a

b

c

d