CDH1 mutations in gastric cancers are not influenced by family history

CURRENT STATUS: UNDER REVIEW

BMC Medical Genetics  BMC Series

Sangjoon Choi
Samsung Medical Center
ORCiD: 0000-0003-2108-0575

Jiryeon Jang
Sungkyunkwan University

You Jeong Heo
Sungkyunkwan University

So Young Kang
Samsung Medical Center

Seung Tae Kim
Samsung Medical Center

Jeeyun Lee
Samsung Medical Center

Won Ki Kang
Samsung Medical Center

Jong-Won Kim
Samsung Medical Center

Kyoung-Mee Kim  kkmkys@skku.edu
Corresponding Author
ORCiD: 0000-0002-1162-9205

DOI:
10.21203/rs.2.23340/v1

SUBJECT AREAS
Medical Genetics

KEYWORDS
CDH1, Gastric cancer, Mutation, Sequencing, Family history
Abstract

Background: CDH1 mutation is the most frequent genetic alteration in hereditary diffuse gastric cancer (GC) and early onset diffuse GC patients. However, the incidence of CDH1 mutations in sporadic GC with or without family history has not been studied.

Methods: This retrospective study includes a total of 993 Korean patients with primary advanced GC who underwent surgery and received palliative chemotherapy. Targeted deep sequencing was performed in all cases and family history of GC was searched with survival analysis.

Results: We found CDH1 alterations in 146 of 993 patients (14.7%) and 8 were germline (0.8%). Out of 146 patients with CDH1 mutations, 25 (17.1%) had a family history of GC in one of their first relatives, and 12 patients (8.2%) were diagnosed with familial GC (FGC). All cases with FGC were diffuse type by Lauren classification, and only one harbored a previously reported germline mutation of CDH1 (c.2638G>A) and the remaining 11 harbored known somatic CDH1 mutations. Among all patients with CDH1 mutation, there was no significant survival difference between patients with family history or FGC. In the 847 patients without CDH1 mutation, 189 (22.3 %) had a family history of GC and 92 patients (10.9%) were FGC. CDH1 mutations were more frequent in patients with early onset (<45 years) GC (45.5%) compared with patients with late onset GC (10.9%) (p = 0.001), but were not significantly associated with the family history of GC (p > 0.05).

Conclusions: CDH1 mutations are mostly somatic and typically are not associated with family history.

Background

CDH1 gene encodes the protein E-cadherin, which is a cell-cell adhesion protein that plays an important role in tumor development, epithelial invasion, and progression of tumor
cells in gastric cancer (GC). Germline mutation of CDH1 is the most frequently found genetic alteration in hereditary diffuse gastric cancer (HDGC) patients and is detected in 25–30% of cases fulfilling the clinical criteria of International Gastric Cancer Linkage Consortium [1]. A total of 122 CDH1 germline mutations has been reported worldwide, and the types of mutations include deletions, insertions and splice site, nonsense, and missense mutations [2]. Most previous studies identified germline mutations through polymerase chain reaction and sequencing [3–7] with multiplex ligation-dependent probe amplification [8–11] or high-resolution melting analyses [12] using DNA from patients with familial gastric cancer (FGC), HDGC, or early-onset gastric cancer (EOGC). From these studies, CDH1 germline mutations were identified in 10% of FGC cases [10, 11], 0 ~ 19% of HDGC cases [3–5, 8], and 0 ~ 8.9% [5–7, 9] of EOGC cases with diffuse or mixed histologic types [6]. Germline mutations of CDH1 were also reported in 1.7% of patients with sporadic GC [12].

The recent advent of and use of multigene cancer panel testing using next-generation sequencing (NGS) has dramatically improved testing by enabling simultaneous evaluation of many patients for multiple high-risk cancer syndromes as well as moderate-risk cancer-associated genes such as CDH1 [13]. While this strategy has resulted in significant increases in identification of individuals at risk, it has also led to clinical conundrums regarding management of nonsyndromic carriers of an unexpected or unanticipated genetic mutation [14], such as patients with CDH1 mutation [15]. With NGS, somatic mutations of CDH1 are found in 9%~36% of GC cases [16–22] and are highly prevalent in sporadic early-onset diffuse-type GC (EODGC) (53.2% of cases) [17]. Somatic mutations of CDH1 have been suggested as a poor prognostic marker in diffuse-type GC [22, 23]. Although a few studies have investigated the frequency and impact of CDH1 mutations in HDGC and EODGC [9, 10, 17, 24] the incidence of CDH1 mutations in sporadic GC with or
without family history has not been studied. Although familial aggregation is found in about 10% of GC cases, the prevalence of CDH1 alterations in patients with familial aggregation is unclear [25].

In the present study, we studied CDH1 alterations with deep targeted sequencing and investigated the association with family history in 993 GC patients in a GC-prevalent country.

**Methods**

**Patients**

This study involved retrospective analysis of patients with GC at Samsung Medical Center in Seoul, South Korea. All patients were diagnosed with GC by pathologic examination and provided written informed consent. The 993 patients included 647 male and 346 female patients, and the patient age ranged from 18 to 80 years (mean 60.8). By Lauren classification, diffuse type was found in 463 cases. Patients < 45 years old were defined as EOGC, and 110 cases in the total patient group were classified as EOCG. Among these 110 cases, 81 were diffuse type and were diagnosed as EODGC.

Patients who had family history of GC in first- or second-degree relatives were considered as GC with family history. Patients with family history were defined as FGC if they met the following criteria: (1) GC in 2 or more first/second-degree relatives, with at least one diagnosis before age 50 or (2) GC in 3 or more first/second-degree relatives, independent of age. The remaining patients were classified as sporadic GC, as previously described [26].

**DNA extraction and targeted deep sequencing**

Genomic DNA was extracted from fresh tissues using QIAamp DNA mini kits (Qiagen, Valencia, CA, USA) and from formalin-fixed paraffin-embedded (FFPE) tissues using the
QIAamp DNA FFPE Tissue kit according to the manufacturer’s instruction. DNA concentration was measured by spectrophotometry (ND1000, Nanodrop Technologies, Thermo Scientific, MA, USA). Sample purity was assessed using a Qubit Fluorometer (Life Technologies, Grand Island, NY, USA).

Genomic DNA (~ 250 ng) from each tissue was sheared in a Covaris S220 ultrasonicator (Covaris, Woburn, MA, USA) and used with CancerSCAN™ probes and a SureSelect XT reagent kit HSQ (Agilent Technologies) for construction of a library according to the manufacturer’s protocol. This panel is designed to enrich 381 exons, covering 366.2 kb of the human genome [23]. After enriched exome libraries were multiplexed, the libraries were sequenced on a HiSeq 2500 sequencing platform (Illumina). Briefly, a paired-end DNA sequencing library was prepared through gDNA shearing, end-repair, A-tailing, paired-end adaptor ligation, and amplification. After hybridization of the library with bait sequences for 27 h, the captured library was purified and amplified with an index barcode tag, and the library quality and quantity were assessed. Sequencing of the exome library was performed using the 100-bp paired-end mode of the TruSeq Rapid PE Cluster Kit and TruSeq Rapid SBS Kit (Illumina).

Variant detection

Sequence reads were mapped to the human genome (hg19) using Burrows-Wheeler Aligner (BWA) [27]. Duplicate read removal was performed using Picard and SAMtools [28]. Local alignment was optimized using the Genome Analysis Toolkit (GATK) [29]. To detect single nucleotide variants (SNVs), we integrated the results of three variant callers to increase sensitivity [30–32]. For insertion and deletions (indels), Pindel was used [33]. Copy number variations were calculated for targeted regions by dividing read depth per exon by estimated normal reads per exon using an in-house reference.

Single nucleotide polymorphisms (SNPs) of CDH1 were removed based on dbSNP
Analyses of CDH1 variants

Lists of CDH1 variants were assembled from the literature and a search using the ClinVar [34], Human Gene Mutation Database (HGMD) [35], the Genome Aggregation Database (05gnomAD) (https://gnomad.broadinstitute.org/), the single nucleotide polymorphism database (dbSNP) (https://www.ncbi.nlm.nih.gov/snp/), 1000 genome project, Exac03, ESP5400, and KRDGB_1100. For somatic alterations, Catalogue of Somatic Mutations in Cancer (COSMIC v85) and TCGA databases were used. The predisposition discovery process of CDH1 variants in the 993 GC cases is described in Fig. 1.

Statistical analysis

SPSS ver. 23.0 statistical software program (SPSS Inc., Chicago, IL, USA) was used for statistical analyses. Chi-square and linear by linear test were performed to analyze correlations of clinicopathological features between patients with or without family history, and FGC patients. Overall survival (OS) rates were calculated by the Kaplan-Meier method and differences in survival were compared using log-rank test. A p-value < .05 indicated a statistically significant result.

Results

Mutations types of CDH1

We identified 146 (14.7%) CDH1 mutations from 993 GC cases by NGS analysis. The 146 patients with CDH1 mutations included 69 male and 77 female patients, and patient age ranged from 18 to 80 years (mean 49.6 years). Diffuse-type by Lauren classification was identified in 127 cases (Table 1). Patients < 45 years were defined as EOGC; a total of 50 cases was classified as EOGC, and 47 of these cases were EODGC. CDH1 mutations were
more frequent in patients with EOGC (45.5%) compared with patients with late onset GC (10.9%) (p = < 0.001) (Table 2).

| Table 1 | Characteristics of patients with gastric cancer harboring CDH1 mutation (n = 146) |
|---------|-----------------------------------------------------------------------------|
|         | Familial gastric cancer (n = 12) | GC with family history (n = 13) | GC without family history (n = 121) | p value |
| Age, year | Median (range) | 47 (30–68) | 57 (50–71) | 50 (18–80) | 0.485 |
| Sex | | Male | 5 | 10 | 54 | 0.050 |
|     | | Female | 7 | 3 | 67 | |
| Differentiation | | Well or moderate | 0 | 5 | 5 | |
|     | | Poor | 9 | 4 | 61 | |
|     | | Signet ring cell | 3 | 4 | 46 | |
|     | | Others | 0 | 0 | 9 | |
|     | | Lauren histology | | | | 0.081 |
|     | | Intestinal | 0 | 5 | 5 | |
|     | | Diffuse | 12 | 8 | 107 | |
|     | | Mixed | 0 | 0 | 6 | |
|     | | Indeterminate | 0 | 0 | 3 | |
|     | | CDH1 mutations | | | | 0.229 |
|     | | Germline | 1 | 2 | 5 | |
|     | | Somatic | 11 | 11 | 116 | |

| Table 2 | Characteristics of gastric cancer patients with or without CDH1 mutation |
|---------|---------------------------------------------------------------------|
|         | GC with CDH1 mutation (n = 146) | GC without CDH1 mutation (n = 847) | p value |
| Age | | | | < 0.001 |
| < 45 | 50 | 60 | 96 | 787 | 0.370 |
| ≥ 45 | 96 | 787 | Absent | 12 | 92 | |
| FGC | 12 | 92 | GC with FHx | 13 | 97 | |
| Absent | 121 | 658 | |

To rule out the possibility of deamination effects of formalin, the nucleotide changes were analyzed, and the prevalence of C > G alterations (2.74%) was not significantly frequent (p > 0.05) (Fig. 2A). Mutations of CDH1 included nonsynonymous SNVs (n = 99, 67.8%), stopgain SNV (n = 14, 9.6%), frameshift deletion (n = 12, 8.2%), and frameshift insertion (n = 4, 2.7%), frameshift substitution (n = 1, 0.7%), nonframeshift deletion (n = 15, 10.3%), and nonframeshift substitution (n = 1, 0.7%). The distribution of CDH1 mutation is shown in Fig. 2B. The distributions of somatic mutations within the CDH1 gene are depicted in Fig. 3.
With a cut-off value for variant allele frequency (VAF) of > 50% [36], germline mutations were identified in 8 patients (0.8%) and 2 of them were novel: c.1424T > A in exon 10 and c.1839_1841delCAT in exon 12. The two novel CDH1 mutations were validated as germline mutations by Sanger sequencing of paired normal tissue. Remaining 6 mutations have been reported previously as uncertain significance. The precise mutation profiles of germline CDH1 alteration with clinicopathologic characteristics are described in Table 3.

**Table 3**

| Case | Age/Sex | Exon | cDNA sequence changes | Function | Clinical significance in ClinVar | Laurent’s classification | Histologic type | Family history |
|------|---------|------|-----------------------|----------|---------------------------------|--------------------------|----------------|---------------|
| 1    | 65/M    | exon10 | c.1424T > A*          | nonsynonymous SNV | Uncertain significance | Diffuse          | SRCC           | No            |
| 2    | 53/M    | exon10 | c.1418T > A           | nonsynonymous SNV | Uncertain significance | Diffuse          | Tubular PD     | No            |
| 3    | 45/F    | exon10 | c.1478T > C           | nonsynonymous SNV | Uncertain significance | Diffuse          | Tubular PD     | No            |
| 4    | 59/M    | exon12 | c.1839_1841delCAT*    | nonframeshift deletion | Uncertain significance | Diffuse          | SRCC           | No            |
| 5    | 47/M    | exon16 | c.2638G > A           | nonsynonymous SNV | Uncertain significance | Diffuse          | Tubular PD     | FGC           |
| 6    | 47/F    | exon05 | c.546A > C            | nonsynonymous SNV | Uncertain significance | Diffuse          | Mucinous ADC   | No            |
| 7    | 50/M    | exon14 | c.2246G > A           | nonsynonymous SNV | Uncertain significance | Intestinal        | Papillary WD   | FHx           |
| 8    | 70/F    | exon02 | c.76G > C             | nonsynonymous SNV | Uncertain significance | Diffuse          | Tubular PD     | No            |

*These two mutations are considered as novel ones

AA Amino acid, F/U Follow up, SNV Single nucleotide variant, SRCC Signet ring cell carcinoma, DOD Dead of disease, PD poorly differentiated, Del Deletion, FGC Familial gastric cancer, ADC Adenocarcinoma, FHx Family history

Mutations of CDH1 in patients with and without family history

In the 146 patients with CDH1 mutations, 25 (17.2%) had a familial history of GC in one of their first relatives, and 12 patients (8.2%) were diagnosed with FGC with CDH1 mutation (Table 1). All cases with FGC were diffuse-type by Lauren classification, and only one harbored germline mutations of CDH1 (c.2638G > A); the remaining 12 had known somatic CDH1 mutations. No FGC patients met the clinical criteria or histologic features of HDGC.
Unexpectedly, we found germline mutations of CDH1 in 6 patients without a family history of GC. Among all patients with CDH1 mutation, there was no significant survival difference between patients with family history or FGC (Fig. 4). In the 847 patients without CDH1 mutation, 189 (22.3%) had a family history of GC in one of their first relatives, and 92 patients (10.9%) were diagnosed as FGC (Table 2). We could not find any statistical significance in CDH1 mutations with family history of GC (p > 0.05).

Discussion

Although familial aggregation is found in about 10% of GC cases [25], the incidence of CDH1 alterations in patients with or without family history of GC is unclear. We explored CDH1 alterations with NGS and investigated the association with family history in 993 GC patients. We found that occurrence of CDH1 mutation was not significantly associated with family history.

After the first report of germline CDH1 mutation in GC [37], various types of CDH1 mutations have been reported in HDGC or diffuse type GC by polymerase chain reaction and sequencing [3–7] aided with multiplex ligation-dependent probe amplification [8–11] or high-resolution melting analyses [30] from patients with FGC, HDGC, or EOGC. CDH1 germline mutations were identified in 10% of FGC cases[10, 11], 0 ~ 19% of HDGC cases [3–5, 8] and 0 ~ 8.9% [5–7, 9] of EOGC cases [6], as well as in 1.7% of patients with sporadic GC [12]. With NGS, somatic mutations of CDH1 are found in 9%~36% of GC cases[16–22] and are highly prevalent in sporadic EODGC (53.2% of cases). Somatic CDH1 mutation has also been suggested as a poor prognostic marker in diffuse-type GC [17, 22, 23]. The previously reported germline and somatic CDH1 mutations detected in patients with gastric cancer by various technologies are summarized in Table 4. Although a few studies have investigated the frequency and impact of CDH1 mutations in HDGC and EODGC [9, 10, 17, 24], the incidence of CDH1 mutations in sporadic GC patients with or
without family history is unclear [25]. To the best of our knowledge, this is the first study exploring CDH1 mutations in association with family history. Our results showed that CDH1 mutations are not influenced by family history.

In the present study, we identified 146 CDH1 mutations in 993 patients. Eight CDH1 mutations were germline (0.8%, 8/993), and this incidence is similar to a previous study in a Chinese population (1.7%, 4/236), the only study investigating CDH1 germline mutation in sporadic GC patients [12]. In previous studies in EODGC patients, CDH1 germline mutation rates were higher in Korean (8%, 2/25) [9] and Italian populations (7.2%, 19/264).
In a meta-analyses study, GC occurring in a high-prevalence area harbored less frequent CDH1 germline mutations than that in a low-incidence area [2]. Based on these observations, our low incidence of CDH1 germline mutations may be due to patient age (more frequent CDH1 germline mutations in younger patients) and because the present study was performed in a GC-prevalent area.

In GC-prevalent areas, the ratio of missense mutations was higher than that of non-missense (deletion, insertion, truncating, and nonsense) mutations [2]. In addition, non-missense mutations were reported to be likely pathogenic compared with missense mutations. We also found that the frequency of missense mutation was higher (67.8%) compared with that of non-missense mutation, and this result is similar to the previous analysis. A recent comprehensive genomic study demonstrated that RhoGAP domain-containing fusions or PPAPDC1A fusions, which are mostly somatic mutations, are common in GC of diffuse-type [38]. In addition, germline mutations in genes such as CTNNA1 were observed at a similar frequency as CDH1 germline mutation [39]. These findings suggest that GC might be frequently caused by somatic mutations rather than germline alterations of CDH1 in GC-prevalent areas such as Korea, as previously suggested [2].

Unlike previous studies that focused exclusively on HDGC patients, we compared the distribution of CDH1 alterations in patients with and without family history of GC. As a result, most CDH1 alterations in both groups were somatic mutations. Thus, CDH1 somatic mutation appears to play a more important role than germline mutation regardless of family history;

The limitation of this study is that the patient's family history was investigated only through electronic medical records. Therefore, the age of onset and histological information of GC in the patient's family could not be accurately identified. In addition, there is a possibility of missed records. Further studies might be required to analyze the
family information about GC more in detail with pedigree analysis.

Conclusions

In conclusion, we found 146 CDH1 mutations in 14.7% of GC patients. CDH1 mutations were not associated with family history. Rare germline mutations in sporadic and familial GC patients suggest that somatic CDH1 mutation may play an important role in pathogenesis of GC in GC-prevalent areas.

Abbreviations

GC: Gastric cancer; HDGC: hereditary diffuse gastric cancer; FGC: Familial gastric cancer; EOGC: Early-onset gastric cancer; NGS: Next-generation sequencing; EODGC: Early-onset diffuse-type gastric cancer; FFPE: Formalin-fixed paraffin-embedded; SNV: Single nucleotide variants; Indels: Insertion and deletions; SNP: Single nucleotide polymorphisms; NCBI: National Center for Biotechnology Information; HGMD: Human Gene Mutation Database; gnomAD: Genome Aggregation Database; dbSNP: single nucleotide polymorphism database; COSMIC: Catalogue of Somatic Mutations in Cancer; TCGA: The Cancer Genome Atlas; OS: Overall survival; VAF: Variant allele frequency;

Declarations

Acknowledgements

We would like to thank the patients for participating in this study

Author’s contribusions

JWK and KMK conceived and designed the study. JJ, YJH, and SYK performed data analysis. SC generated tables and figures. SC and JJ drafted the manuscript. STK, JL, WKK, JWK, and KMK revised the manuscript and contributed to knowledge.

Funding

This work was supported by the Basic Science Research Program through the National
Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017R1E1A1A01075005 and NRF-2017R1A2B4012436).

Availability of data and materials
All data generated or analyzed during this study are included in this published article.

Ethics approval and consent to participate
The Institutional Review Board of Samsung Medical Center approved this study and waived informed consent.

Consent for publication
Not applicable

Competing interests
The authors declare no competing interests.

References
1. Guilford P, Humar B, Blair V: Hereditary diffuse gastric cancer: translation of CDH1 germline mutations into clinical practice. Gastric Cancer 2010, 13(1):1-10.

2. Corso G, Marrelli D, Pascale V, Vindigni C, Roviello F: Frequency of CDH1 germline mutations in gastric carcinoma coming from high- and low-risk areas: metanalysis and systematic review of the literature. BMC Cancer 2012, 12:8.

3. van der Post RS, Vogelaar IP, Manders P, van der Kolk LE, Cats A, van Hest LP, Sijmons R, Aalfs CM, Ausems MG, Gomez Garcia EB et al: Accuracy of Hereditary Diffuse Gastric Cancer Testing Criteria and Outcomes in Patients With a Germline Mutation in CDH1. Gastroenterology 2015, 149(4):897-906 e819.

4. Jakubowska A, Lawniczak M, Wojnarska B, Cybulski C, Huzarski T, Byrski T, Toloczko-Grabarek A, Jaworska K, Durda K, Starzynska T et al: CDH1 gene mutations do not contribute in hereditary diffuse gastric cancer in Poland. Fam Cancer 2010,
5. Oliveira C, Ferreira P, Nabais S, Campos L, Ferreira A, Cirnes L, Alves CC, Veiga I, Fragoso M, Regateiro F et al: E-Cadherin (CDH1) and p53 rather than SMAD4 and Caspase-10 germline mutations contribute to genetic predisposition in Portuguese gastric cancer patients. Eur J Cancer 2004, 40(12):1897-1903.

6. Corso G, Pedrazzani C, Pinheiro H, Fernandes E, Marrelli D, Rinnovati A, Pascale V, Seruca R, Oliveira C, Roviello F: E-cadherin genetic screening and clinico-pathologic characteristics of early onset gastric cancer. Eur J Cancer 2011, 47(4):631-639.

7. Suriano G, Oliveira C, Ferreira P, Machado JC, Bordin MC, De Wever O, Bruyneel EA, Moguilevsky N, Grehan N, Porter TR et al: Identification of CDH1 germline missense mutations associated with functional inactivation of the E-cadherin protein in young gastric cancer probands. Hum Mol Genet 2003, 12(5):575-582.

8. Hansford S, Kaurah P, Li-Chang H, Woo M, Senz J, Pinheiro H, Schrader KA, Schaeffer DF, Shumansky K, Zogopoulos G et al: Hereditary Diffuse Gastric Cancer Syndrome: CDH1 Mutations and Beyond. JAMA Oncol 2015, 1(1):23-32.

9. Kim S, Chung JW, Jeong TD, Park YS, Lee JH, Ahn JY, Kim DH, Choi KD, Lee W, Song HJ et al: Searching for E-cadherin gene mutations in early onset diffuse gastric cancer and hereditary diffuse gastric cancer in Korean patients. Fam Cancer 2013, 12(3):503-507.

10. Yoon KA, Ku JL, Yang HK, Kim WH, Park SY, Park JG: Germline mutations of E-cadherin gene in Korean familial gastric cancer patients. J Hum Genet 1999, 44(3):177-180.

11. Richards FM, McKee SA, Rajpar MH, Cole TR, Evans DG, Jankowski JA, McKeown C, Sanders DS, Maher ER: Germline E-cadherin gene (CDH1) mutations predispose
to familial gastric cancer and colorectal cancer. *Hum Mol Genet* 1999, 8(4):607-610.

12. Chen QH, Deng W, Li XW, Liu XF, Wang JM, Wang LF, Xiao N, He Q, Wang YP, Fan YM: Novel CDH1 germline mutations identified in Chinese gastric cancer patients. *World J Gastroenterol* 2013, 19(6):909-916.

13. Kurian AW, Ford JM: Multigene Panel Testing in Oncology Practice: How Should We Respond? *JAMA Oncol* 2015, 1(3):277-278.

14. Robson ME, Storm CD, Weitzel J, Wollins DS, Offit K, American Society of Clinical O: American Society of Clinical Oncology policy statement update: genetic and genomic testing for cancer susceptibility. *J Clin Oncol* 2010, 28(5):893-901.

15. Ford JM: Totally unexpected: nonsyndromic CDH1 mutations and hereditary diffuse gastric cancer syndrome. *JCO Precis Oncol* 2017, 1:1-2.

16. Shu Y, Zhang W, Hou Q, Zhao L, Zhang S, Zhou J, Song X, Zhang Y, Jiang D, Chen X et al: Prognostic significance of frequent CLDN18-ARHGAP26/6 fusion in gastric signet-ring cell cancer. *Nat Commun* 2018, 9(1):2447.

17. Cho SY, Park JW, Liu Y, Park YS, Kim JH, Yang H, Um H, Ko WR, Lee BI, Kwon SY et al: Sporadic Early-Onset Diffuse Gastric Cancers Have High Frequency of Somatic CDH1 Alterations, but Low Frequency of Somatic RHOA Mutations Compared With Late-Onset Cancers. *Gastroenterology* 2017, 153(2):536-549 e526.

18. Wang K, Yuen ST, Xu J, Lee SP, Yan HH, Shi ST, Siu HC, Deng S, Chu KM, Law S et al: Whole-genome sequencing and comprehensive molecular profiling identify new driver mutations in gastric cancer. *Nat Genet* 2014, 46(6):573-582.
**RHOA in diffuse-type gastric carcinoma.** *Nat Genet* 2014, **46**(6):583-587.

20. Wang K, Kan J, Yuen ST, Shi ST, Chu KM, Law S, Chan TL, Kan Z, Chan AS, Tsui WY et al: **Exome sequencing identifies frequent mutation of ARID1A in molecular subtypes of gastric cancer.** *Nat Genet* 2011, **43**(12):1219-1223.

21. Cancer Genome Atlas Research N: **Comprehensive molecular characterization of gastric adenocarcinoma.** *Nature* 2014, **513**(7517):202-209.

22. Li X, Wu WK, Xing R, Wong SH, Liu Y, Fang X, Zhang Y, Wang M, Wang J, Li L et al: **Distinct Subtypes of Gastric Cancer Defined by Molecular Characterization Include Novel Mutational Signatures with Prognostic Capability.** *Cancer Res* 2016, **76**(7):1724-1732.

23. Cho J, Ahn S, Son DS, Kim NK, Lee KW, Kim S, Lee J, Park SH, Park JO, Kang WK et al: **Bridging genomics and phenomics of gastric carcinoma.** *Int J Cancer* 2019, **145**(9):2407-2417.

24. Choi HJ, Ki CS, Suh SP, Kim JW: **Presymptomatic identification of CDH1 germline mutation in a healthy korean individual with family history of gastric cancer.** *Ann Lab Med* 2014, **34**(5):386-389.

25. Choi YJ, Kim N: **Gastric cancer and family history.** *Korean J Intern Med* 2016, **31**(6):1042-1053.

26. Kluijt I, Sijmons RH, Hoogerbrugge N, Plukker JT, de Jong D, van Krieken JH, van Hillegersberg R, Ligtenberg M, Bleiker E, Cats A et al: **Familial gastric cancer: guidelines for diagnosis, treatment and periodic surveillance.** *Fam Cancer* 2012, **11**(3):363-369.

27. Li H, Durbin R: **Fast and accurate long-read alignment with Burrows-Wheeler transform.** *Bioinformatics* 2010, **26**(5):589-595.

28. Yang J, Ding X, Sun X, Tsang SY, Xue H: **SAMSVM: A tool for misalignment**
filtration of SAM-format sequences with support vector machine. *J Bioinform Comput Biol* 2015, 13(6):1550025.

29. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M *et al*: *The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data*. *Genome Res* 2010, 20(9):1297-1303.

30. Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G: *Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples*. *Nat Biotechnol* 2013, 31(3):213-219.

31. Wilm A, Aw PP, Bertrand D, Yeo GH, Ong SH, Wong CH, Khor CC, Petric R, Hibberd ML, Nagarajan N: *LoFreq: a sequence-quality aware, ultra-sensitive variant caller for uncovering cell-population heterogeneity from high-throughput sequencing datasets*. *Nucleic Acids Res* 2012, 40(22):11189-11201.

32. Wei Z, Wang W, Hu P, Lyon GJ, Hakonarson H: *SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data*. *Nucleic Acids Res* 2011, 39(19):e132.

33. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z: *Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads*. *Bioinformatics* 2009, 25(21):2865-2871.

34. Landrum MJ, Lee JM, Riley GR, Jang W, Rubinstein WS, Church DM, Maglott DR: *ClinVar: public archive of relationships among sequence variation and human phenotype*. *Nucleic Acids Res* 2014, 42(Database issue):D980-985.

35. Stenson PD, Mort M, Ball EV, Shaw K, Phillips A, Cooper DN: *The Human Gene Mutation Database: building a comprehensive mutation repository for*
clinical and molecular genetics, diagnostic testing and personalized genomic medicine. *Hum Genet* 2014, **133**(1):1-9.

36. Li MM, Datto M, Duncavage EJ, Kulkarni S, Lindeman NI, Roy S, Tsimberidou AM, Vnenck-Jones CL, Wolff DJ, Younes A et al: *Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of American Pathologists*. *J Mol Diagn* 2017, **19**(1):4-23.

37. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, Taite H, Scoular R, Miller A, Reeve AE: *E-cadherin germline mutations in familial gastric cancer*. *Nature* 1998, **392**(6674):402-405.

38. Yang H, Hong D, Cho SY, Park YS, Ko WR, Kim JH, Hur H, Lee J, Kim SJ, Kwon SY et al: *RhoGAP domain-containing fusions and PPAPDC1A fusions are recurrent and prognostic in diffuse gastric cancer*. *Nat Commun* 2018, **9**(1):4439.

39. Wong SS, Kim KM, Ting JC, Yu K, Fu J, Liu S, Cristescu R, Nebozhyn M, Gong L, Yue YG et al: *Genomic landscape and genetic heterogeneity in gastric adenocarcinoma revealed by whole-genome sequencing*. *Nat Commun* 2014, **5**:5477.

Figures
Flow charts of CDH1 variants filtered by public databases

Figure 1
Figure 2

(A) Distribution of nucleic acid changes in CDH1 mutations. (B). Distribution of amino acid change
Figure 3

Distributions of somatic mutations within the CDH1 gene. The figure was created using the cbioportal mutation mapper program.

(https://www.cbioportal.org/mutation_mapper)
Overall survival of gastric cancer patients with CDH1 mutation with and without family history

Supplementary Files

This is a list of supplementary files associated with the primary manuscript. Click to download.
