

Investigation on Gastric Cancer Susceptibility Genes in Chinese Early-onset Diffuse Gastric Cancer

Yi Feng

Peking University Cancer Hospital & Institute <https://orcid.org/0000-0002-2445-1290>

Meng Zhang

Peking University Cancer Hospital & Institute

Wenmei Li

Peking University Cancer Hospital & Institute

Sisi Li

Peking University Cancer Hospital & Institute

Changda Qu

Peking University Cancer Hospital & Institute

Meizhu Meng

Peking University Cancer Hospital & Institute

Meiying Ye

Peking University Cancer Hospital & Institute

Weijia Chen

Peking University Cancer Hospital & Institute

Xiaofang Xing

Peking University Cancer Hospital & Institute

Shuqin Jia (✉ shuqin_jia@hsc.pku.edu.cn)

<https://orcid.org/0000-0001-8092-1739>

Research article

Keywords: hereditary diffuse gastric cancer, early-onset gastric cancer, gastric cancer susceptibility gene, pathogenic germline variant candidate

Posted Date: July 10th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-39794/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at Annals of Oncology on October 1st, 2019.
See the published version at <https://doi.org/10.1093/annonc/mdz247.134>.

Investigation on gastric cancer susceptibility genes in Chinese early-onset diffuse gastric cancer

Yi Feng, Meng Zhang, Wenmei Li, Sisi Li, Changda Qu, Meizhu Meng, Meiyong Ye, Weijia Chen, Xiaofang Xing, Shuqin Jia*

Peking University Cancer Hospital & Institute, Key laboratory of Carcinogenesis and Translational Research (Ministry of Education / Beijing), Department of Molecular Diagnostics, No. 52 Fucheng road, Haidian District, 100142, Beijing, China.

* correspondence author

Email address of the correspondence author: shuqin_jia@hsc.pku.edu.cn

Abstract

Background

Diffuse gastric cancer (DGC) is known as gastric cancer with histological form of diffuse type morphology. Early onset DGC is an indicator of suspicious hereditary diffuse gastric cancer (HDGC). HDGC patients are often less sensitive to chemotherapy and suffer from highly invasive late stage malignancy with poor prognosis. Such hereditary cancer syndrome is closely related to deleterious *CDH1* germline variant. In addition, potential pathogenic germline variants in other candidate genes were also observed in HDGC families. However, the profile of gastric cancer (GC) susceptibility gene in Chinese HDGC patients is yet to be elucidated.

Methods

To investigate gastric cancer susceptibility genes in Chinese gastric cancer patients, we collected peripheral blood samples from 29 patients fulfilled both HDGC clinical diagnosis and genetic testing criteria updated on 2015 by IGCLC. Genomic DNA was extracted from peripheral blood followed by Whole Exome Sequencing (WES) with average sequencing depth of at least 100X. Gastric cancer predisposing SNV and Indel candidates were filtered based on population allele frequencies in public nucleotide polymorphism databases and pathogenic variant filtering process was performed to identify potential gastric cancer predisposing variant candidates. Immunohistochemistry was conducted on biopsies from patient who carries potential pathogenic *CDH1* germline variant.

Results

In general, 336296 germline non-synonymous variants were detected from 29 GC patients. According to the pathogenic variant filtering process, 25 germline variant candidates in 22 genes were identified as gastric cancer predisposing variant candidates. In addition, a novel splice-site variant, c.2165-1G>A (NM_004360.4), in *CDH1* was detected in a Chinese early-onset HDGC patient who developed ovarian metastasis. Furthermore, IHC analysis on the ovarian cancer tissue from this patient demonstrated weakly to moderate staining of E-cadherin compared with positive control. Moreover, another two variants, *CTNNA1*

c.1975_1976del (NM_001903.2), *APC* c.-19+1G>A (NM_001127511.2), were suspicious of gastric cancer predisposing variants.

Conclusions

This study suggested that *CDH1* c.2165-1G>A may act as a gastric cancer predisposing variant. In addition, to further investigate molecular mechanisms of early-onset gastric cancer, one may consider 22 genes observed in this study. Furthermore, the inconclusive results in this study warrant future investigation on gastric cancer susceptibility gene discovery that cohort selection may require more stringent conditions.

Key words: hereditary diffuse gastric cancer, early-onset gastric cancer, gastric cancer susceptibility gene, pathogenic germline variant candidate

Background

Gastric cancer (GC) is the fifth most common malignancy and third leading cause of cancer-related mortality worldwide (1, 2). Moreover, GC is the most prevalent cancer in eastern Asia including China (3). In general, most of the GCs are sporadic, whereas 1-3% is caused by inherited cancer predisposition syndromes (4). The best characterised inherited GC is hereditary diffuse gastric cancer (HDGC) which is diffuse-type with multiple foci of signet-ring cell carcinoma (SRCC) underneath an intact surface epithelium (5). Patients with this manifestation are likely to be highly invasive late stage cancers that are not sensitive to chemotherapy, and are often suffered from poor prognosis.

Previous studies demonstrated that germline mutations in *CDH1* is one of the risk factors related with HDGC, due to reduced or even loss activity of normal protein functions of E-cadherin (5). *CDH1* is a tumour suppressor gene which maps to chromosome 16q22.1 and consists of 16 exons. The genomic length of *CDH1* is approximately 100kb, and it is transcribed into 4.5-kb mRNA followed by being translated to a 120kDa E-cadherin, a highly conserved transmembrane glycoprotein whose function is to maintain calcium-dependent cell-cell adhesion through association of cytosolic protein complex mainly formed by catenins (6, 7). Therefore, susceptibility for gastric cancer in *CDH1* mutation carriers is suggested to be attributed to aberrant E-cadherin function or protein loss, since a normal functioning E-cadherin is highly required by various signalling pathways and cross talks to maintain the balance of cell proliferation, motility and polarity (7). In addition, the estimated cumulative risk for *CDH1* mutation carriers to develop GC by age 80 years is 67% for men and 83% for women (8). Therefore, screening for *CDH1* germline pathogenic variants was included in the clinical criteria for HDGC diagnosis by the International Gastric Cancer Linkage Consortium (IGCLC) (9, 10).

The global incidence of gastric cancer is highly diverse and the reason for this geographic difference is remained to be elucidated. Additionally, such discrepancy remains in terms of incidence of identified *CDH1* germline variants. There are more than 100 different germline variants characterised in *CDH1* coding region and splice site region from families with different ethnic background worldwide (4, 11, 12). However, less *CDH1* germline variants were detected from families in areas with middle/high gastric cancer incidence, namely Italy (Tuscany), Portugal (North), Japan, China, Korea, South America, Lithuania and Poland compared to low incidence countries (12). In addition, reports indicated that in countries of low incidence, the detection rate of pathogenic *CDH1* variants in patients fulfilling the IGCLC screening criteria ranges from 10-18%, whereas such number dropped to less than 10% in high-incidence areas (4) This may due to a potential genetic difference between diverse ethnic groups that attributes to different genetic background and environmental effects, and may suggest that the direct implementation of clinical management guidelines for HDGC, primarily based on Caucasian population, may not be appropriate for Asian population. Considering that China contributes about 42% gastric cancer incidence worldwide, investigations in Chinese population is in urgent need. Furthermore, not all families fulfilling these criteria have mutations in *CDH1*, indicating the existence of other DGC predisposition genes (4). For instance, a germline truncating variant in *CTNNA1* was present in two invasive diffuse gastric cancer patients (13), and Hansford *et al.* suggested that apart from

pathogenic germline variant in *CTNNA1*, a gene involved in E-cadherin signalling, other genes such as *BRCA2*, *STK11*, *SDHB*, *PRSS1*, *ATM*, *MSR1* and *PALB2* may also potentially contribute to gastric cancer susceptibility (11). In order to improve the clinical management of such malignancy, studies on exploiting novel pathogenic variations in *CDH1* and new susceptibility genes for HDGC in Chinese population are required to further elucidate the molecular mechanism of HDGC. Given that early-onset gastric cancers showed advanced stage of disease with high proportion of poorly differentiated carcinomas (14), and that early age of diagnosis potentially indicate presence of familial and hereditary cancer which may be attributable to genetic factors, therefore, investigating on such group of cancer patients may assist to identify susceptibility genes to gastric cancer.

Our study was performed to aid the need of clinical data from Chinese population based on early onset Chinese GC patients, and to clarify potential susceptibility gene candidates, as well as underpinning potential association between genotypic information and clinical features in this group of patients.

Methods

Patient information

A total of 29 patients included in this study were recruited from Peking University Cancer Hospital. To investigate pathogenic variations in *CDH1* and potential susceptibility genes in Chinese HDGC patients, peripheral blood samples were collected from patients. Family history was self-reported by patients, and clinicopathological information was collected from electronic medical record system in Peking University Cancer Hospital (Table 1). In this cohort, 28 patients were diagnosed as DGC with onset age younger than 40-year, whereas one patient, a first degree relative of one of the 28 patients, was confirmed diagnosis of DGC at age of 49 years. All of the 29 patients fulfilled HDGC clinical diagnosis criteria updated on 2015 by IGCLC. Written consent was signed by patients and the study was approved by the Medical Ethics Committee, Beijing Cancer Hospital.

Table 1 Patient clinicopathological characteristics

Characteristic	Number (%)
Gender	
Female	20 (68.97)
Male	9 (31.03)
Age at diagnosis of HDGC, y	
Mean \pm SD	30 \pm 5
Range, y	20-49
\leq 30	15 (51.72)
30-40	13 (44.83)
>40	1 (3.45)
Lauren classification	
Diffuse	26 (89.66)
Mixed	3 (10.34)
Cancer family history	
gastric cancer	7 (24.14)
colorectal cancer	2 (6.90)
gastrointestinal cancer	14 (48.28)
Pan-cancer	18 (62.07)
HP infection	
Yes	8 (27.59)
No	19 (65.52)
NA	2 (6.90)
EBV infection	
Yes	10 (34.48)
No	12 (41.38)
NA	7 (24.14)

Whole exome sequencing and potential pathogenic germline variant filtering process

Genomic DNA was extracted from peripheral blood samples and whole exome regions were captured followed by library enrichment. Qualified library was sequenced for 150bp paired-

end Whole Exome Sequencing with average sequencing depth of at least 100X using the Illumina HiSeq 2500 platform. Reads were aligned to the reference human genome GRCh37 and Genome Analysis tool kit (GATK) was used for realignment and base recalibration. Variant Call Format file was further generated by a standard pipeline.

Non-synonymous SNV and InDel variants with allele frequency (AF) below 0.01 or not included in public nucleotide polymorphism databases (1000_AF, 1000_EAS, NHLBI-ESP_AA/EA, gnomAD_AF, gnomAD_EAS) were preserved. Pathogenic variant candidates were then filtered by process indicated as follows (Figure 1):

1. Variants classified as pathogenic / likely pathogenic (P/LP) in ClinVar or as disease-causing mutations (DM) in HGMD were directly characterised as candidate variants;
2. SNVs predicted to both have the highest-level impact on protein function by two *in silico* tools (SIFT, Polyphen2) were retained;
3. The maintained SNVs, InDels and splice site variants were filtered by candidate gene set consist of gastric cancer susceptibility genes indicated in literatures and in NCCN guidelines (Table 2);
4. The filtered splice site variants were subsequently tested by splicing signal prediction tool (HSF), and the one predicted to be “disturbing splicing” was preserved as deleterious variant candidates.

Table 2. Candidate genes with potential implication of gastric cancer susceptibility

Candidate gene	Evidence suggestive of potential implication in GC susceptibility	Reference
<i>CDH1</i>	known to be associated with hereditary diffuse gastric cancer occurrence	(11)
<i>CTNNA1</i>	Potential deleterious germline mutation in <i>CTNNA1</i> was observed in a HDGC family	(13)
<i>BRCA1</i>	frameshift germline variants detected in diffuse gastric cancer patients	(15)
<i>BRCA2</i>	pathogenic truncating variants found in <i>BRCA2</i> in an east Asian HDGC patient	(11)
<i>RAD51C</i>	deleterious truncating germline mutation observed in diffuse gastric cancer patient	(15)
<i>MHL1</i>	cancer susceptibility genes, recommended by NCCN guideline, attribute to hereditary cancer syndromes that associated with increased gastric cancer risk	(16)
<i>MHS2</i>		
<i>MHS6</i>		
<i>PMS2</i>		
<i>EPCAM</i>		
<i>APC</i>		
<i>SMAD4</i>		
<i>BMPR1A</i>		
<i>PTEN</i>	potential deleterious <i>PTEN</i> germline missense variant identified in patient with gastric adenocarcinoma	(17)
<i>STK11</i>	likely pathogenic missense variant in <i>STK11</i> , a Peutz-Jeghers syndrome susceptibility gene, observed in 3 HDGC patients	(11)
<i>SDHB</i>	likely pathogenic missense variant in <i>SDHB</i> , a Cowden-like syndrome associated gene, detected in HDGC patient with early-onset lobular breast cancer history	(11, 18)
<i>PRSS1</i>	possible pathogenic truncating mutation discovered in <i>PRSS1</i> , a gene associated with hereditary pancreatitis, in HDGC patient	(11, 19, 20)
<i>ATM</i>	potential deleterious mutations found in <i>ATM</i> in HDGC and FGC patients	(11, 21, 22)
<i>MSR1</i>	possible pathogenic truncating mutation discovered in <i>MSR1</i> , a gene associated with increased risk of esophageal cancer and prostate cancer, in HDGC patients	(11, 23, 24)
<i>PALB2</i>	potential pathogenic variants found in HDGC patients	(11, 15, 25)
<i>TP53</i>	pathogenic variant in <i>TP53</i> observed in an early-onset diffuse type gastric cancer patient with breast cancer family history	(22)
<i>ROHA</i>	<i>RHOA</i> germline variants found in DGCs with presence of poorly differentiated adenocarcinomas	(26)

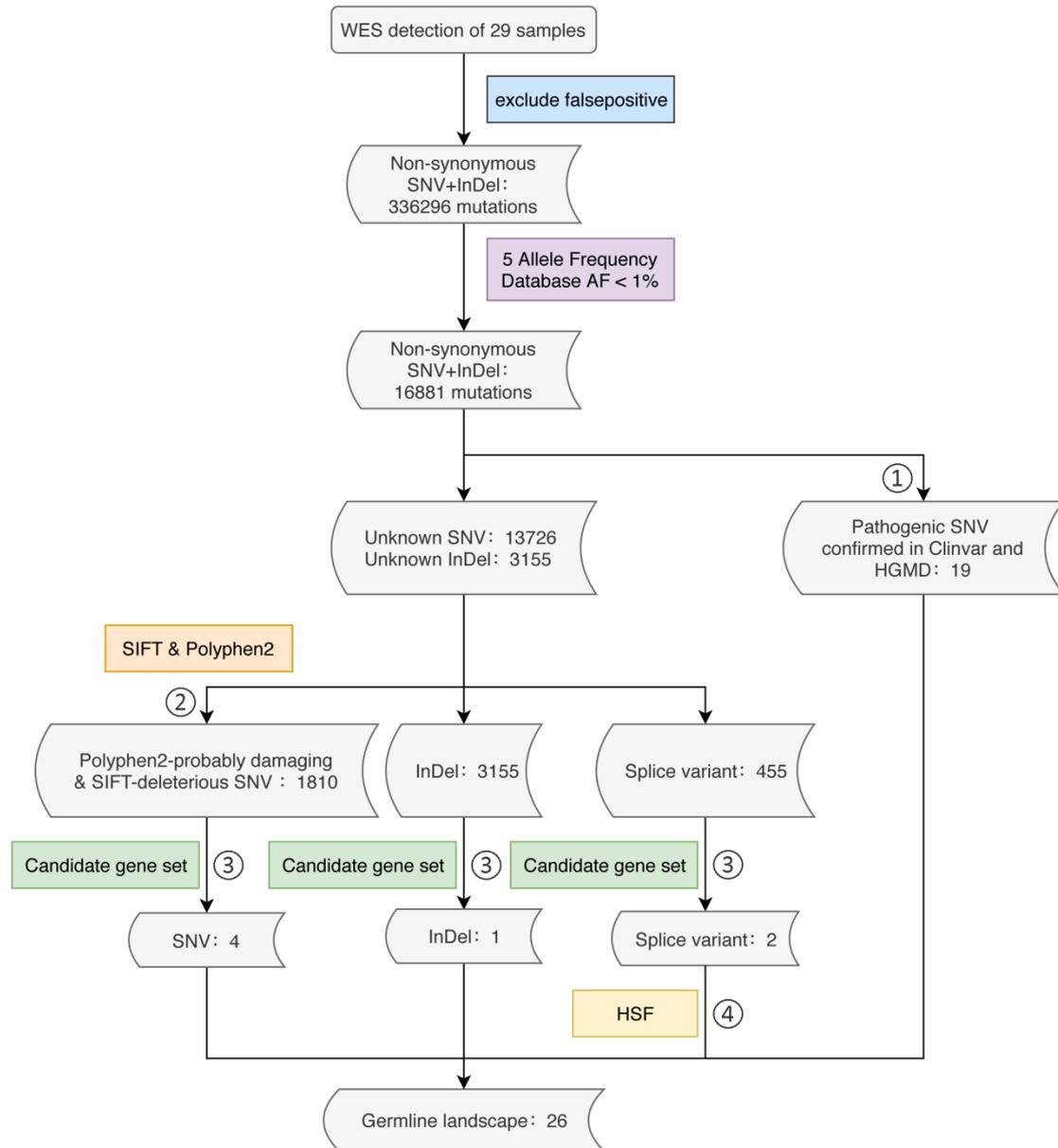


Figure 1. Process of identifying potential pathogenic germline variant candidates

False positive variants were excluded and 5 reference allele frequency databases (1000_AF, 1000_EAS, NHLBI-ESP_AA/EA, gnomAD_AF, gnomAD_EAS) were used to initialise non-synonymous variants filtering. Variants with allele frequency less than 0.01 or not included in reference database were preserved. The preserved variants were firstly filtered by ClinVar and HGMD, and the one categorised P/LP in ClinVar or DM in HGMD was directly classified as pathogenic germline variant candidate.

The remaining unidentified variants were then sorted as SNV, InDel and splice-site variant. Furthermore, SNVs with highest-level impact on protein function, predicted as deleterious in SIFT and probably damaging in Polyphen2, were retained. InDels, splice variants and the retained SNVs, were further filtered by candidate gene set constitute of gastric cancer predisposing genes reported in literatures and in NCCN guidelines (Table 2). Splicing signal prediction tool (HSF) was employed to analyse the potential effect of filtered splice site

variants to further assist variant candidate identification. (P: pathogenic, LP: likely pathogenic, DM: disease-causing mutations)

Immunohistochemistry

The slides were baked at 60°C for 2 h prior to the high throughput IHC procedure subsequently followed by being deparaffinized via sequential washing with xylene, graded ethanol, and water. Antigens were retrieved for 15 min at 95°C. Endogenous peroxidase was blocked with 3% H₂O₂ for 30 min. Nonspecific staining was blocked using 10% normal goat serum (in 1× PBS) for 1 h at 37°C. The slides were incubated overnight at 4°C with CDH1 antibody (diluted in 1× PBS). The enhancing step, followed by incubation with the secondary antibody (1 h at room temperature) and the diaminobenzidine (DAB) substrate (5 min at room temperature), were performed according to the protocol of the ABC kit (DAKO). Haematoxylin was used as a counterstain in the last step. The slides were then rinsed, cleared, and mounted. The staining was optimized based on negative and positive controls.

The expression of CDH1 protein were estimated into four levels as 0 (negative), 1 (mild), 2 (moderate), and 3 (intense) by three pathologists independently.

Results

Clinicopathological characteristics of HDGC gastric cancer patients

The family history and clinical characteristics of the 29 patients in this cohort were summarised in Table 3. All patients met the IGCLC criteria (version 2015) for HDGC clinical diagnosis and for *CDH1* pathogenic germline variant screening. Twenty eight patients were diagnosed as GC before age of 40 and one, DGC patient (#RB797) diagnosis at 49-year, is a first degree relative of a 26-year old DGC female (#RB651) in this cohort. The mean age at diagnosis of DGC for all patients included in this study cohort was 30 years (SD 5, range 20-49). According to pathological reports, 26 patients had DGC, whereas 3 were mixed-type GCs. Signet-ring cell carcinoma was observed in 24 patients and 25 had poorly differentiated adenocarcinoma.

Identification of potential pathogenic variant candidates and genes

In general, approximately 336296 non-synonymous variants (SNV and Indels) were detected from 29 GC patients. Germline variants with allele frequency less than 0.01 and not included in five public population databases (1000_AF, 1000_EAS, NHLBI-ESP_AA/EA, gnomAD_AF, gnomAD_EAS) were preserved for further potential pathogenic variant candidates investigation. The filtered variants that are identified as disease causing in HGMD or pathogenic/likely pathogenic in ClinVar were prioritised as potential deleterious variant candidates, and such group of variants was consisted of 17 missense and 2 splice donor variants located in 18 genes. Furthermore, association between gastric cancer susceptibility and the prioritised 18 candidate genes were manually reviewed based on GeneCards, NCBI database and literature review. The remaining unidentified variants were further analysed according to the pathogenic germline variant candidates filtering process indicated in materials and methods. Therefore, 3 SNV, 1 Indel and 2 splice site variants in 4 genes were selected as potential germline pathogenic variant candidates for GC. In total, 24 germline variant candidates in 22 genes observed in 20 GC patients (Table 4).

Potential deleterious *CDH1* splice site variant

No known pathogenic variant in *CDH1* was discovered in 29 GC patients. However, a splice site variant, c.2165-1G>A (NM_004360.4), was found in an EBV and HP negative mixed type early on-set (age of diagnosis: 31 years) GC patient who developed peritoneal and ovarian metastasis. This variant was reported as a pathogenic somatic mutation associated with breast carcinomas in somatic mutation database-COSMIC, suggesting its likelihood of pathogenicity. However, it is neither recorded in ClinVar nor any other population database. Furthermore, two splice site variants in the same position, c.2165-1 G>T and c.2165-1G>C, has been classified as likely pathogenic in ClinVar both by single submitters (27, 28).

Since this variant occurs in 5' untranslated region of *CDH1*, the normal splice function might be influenced by such alteration, affecting normal protein expression of E-cadherin.

Therefore, we performed immunohistochemistry to detect E-cadherin expression on the metastatic ovarian cancer tissues of this patient (#RB809). Compared with the positive control in GC tissues with normal E-cadherin expression, there are sporadic weak to moderate staining on cell membrane in patient carrying heterozygous germline c.2165-1G>A variant (Figure 2).

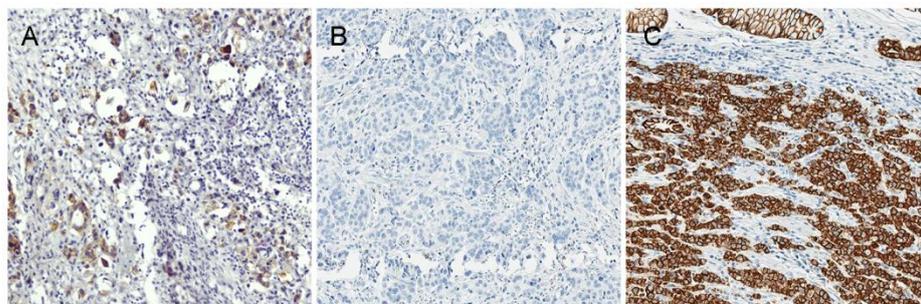


Figure 2. Expression of E-cadherin in HDGC patient (#RB809) carrying potential deleterious variant

Immunohistochemistry staining on (A) metastatic ovarian cancer tissue from gastric cancer patient carrying germline variant in *CDH1* splice (c.2165-1G>A); (B) Negative control; (C) GC tissues with normal E-cadherin expression (Positive control).

Other potential deleterious variants identified in candidate genes associated with GC susceptibility

Another candidate variant identified in *CDH1* was a heterozygous missense variants (c.2053G>A p.Val685Met) located in extracellular domain. This variant used to be identified in 1 of 188 healthy Hispanic individuals in a whole genome sequencing project (29), however, given that participants in this study were younger than 50 years, and that pathogenic variants in *CDH1* only confers increased life-time cancer risk, the unaffected status of such individual may not be representative of control population. Moreover, this variant is absent in Chinese population in 1000 Gnomes Project. Nevertheless, despite the damaging effect predicted by both *in silico* tools (SIFT and Polyphen2), functional studies are required to further identify the potential impact on normal E-cadherin function.

A frameshift variant in *CTNNA1* (c.1975_1976del, p.Glu659ArgfsTer42) was observed in a male diffuse type gastric cancer patient (#RB684) diagnosed on age of 28 years. This variant is neither recorded in ClinVar nor HGMD, and is not present in population databases. In addition, a study reported two *CTNNA1* variants (nonsense and frameshift) in 2 unrelated HDGC families both showed loss of α -catenin protein expression and preserved E-cadherin expression. It is known that α -catenin is a critical component of an adaptor protein complex to bridge E-cadherin and filaments of cytoskeleton, and maintains cell stability as well as inhibiting cell motility (7). Therefore, the variant in *CTNNA1* detected in this study may hold the potential to impair normal function of E-cadherin-catenin complex and contribute to diffuse type pathology.

APC is a colon cancer susceptibility gene whose pathogenic germline mutation may also confer carrier increased GC risk. A 26 year-old female diffuse gastric cancer patient (#RB651) with poorly differentiated phenotype carries a heterozygous splice donor variant in *APC* (c.-19+1G>A), which is not included in either ClinVar or HGMD. No population database recorded this variant either. However, her mother, a DGC patient (#RB797) diagnosed on age of 49 years does not carry such variant. Although this variant was predicted to disrupt normal splice by a functional predictive tool, its potential pathogenicity still needs to be confirmed by *in vitro* study.

Two missense variants in *BRCA1* (c.5518G>T, c.3448C>T) discovered in another two early-onset diffuse GC patients (#RB805, #RB665 respectively) in this cohort were filtered as pathogenic variant candidates. One of the highlighted variants, *BRCA1* c.3448C>T, has been reviewed by ENIGMA expert panel and been classified as “benign” in ClinVar database, whereas the other variant was also recorded in ClinVar yet classified by single submitted as variant of uncertain significance (VUS) due to lack of functional study to verify its pathogenicity.

Potential pathogenic SNVs identified in genes with little evidence of GC predisposition

Among the prioritised 18 genes filtered by step 1 (**Figure 1**), only *POLH* has been reported to associate with a cancer-prone syndrome, which could increase gastric cancer risk (30), whereas five genes (*GBA*, *SBDS*, *PRDM16*, *SPINK1*, *VWF*) exhibited potential implication in gastric cancer. In addition, 6 genes (*LIPH*, *GSS*, *SBF1*, *SLC22A5*, *TYR*, *HNF1A*) were reported to be associated with several solid tumours. However, their implication in gastric cancer has yet to be elucidated. Moreover, there is little evidence showed association between cancer and the remaining 5 genes (*ATP7B*, *MYH7*, *F8*, *KCNQ4*, *PTS*). Therefore, interpretation of pathogenicity on the highlighted candidate variants in these 18 genes should be cautious. *In vitro* functional study or segregation analyses are required to further investigate their potential deleterious effect on gastric cancer predisposition.

POLH

POLH encodes polymerase η that belongs to the Y-family of DNA polymerases, which functions to enable replication on damaged template DNA in an relatively error-free manner (31). Deleterious germline mutations on *POLH* may lead to a cancer-prone syndrome, xeroderma pigmentosum, variant type (XPV; MIM #278750) that follows an autosomal recessive inheritance manner (32). A germline missense variant in *POLH* gene was observed in a 35-year-old breast cancer female whose uncle developed breast cancer at age of 43 years (33). However, according to a case-control study on investigating association between candidate genes and breast cancer susceptibility, individuals carrying rare germline missense variants in *POLH* showed decreased risk of breast cancer (34). In addition, another case-control study reported a common germline variant in *POLH* (c.1783A>G) that was associated with risk of melanoma in Caucasian origin population (35).

GBA

GBA is a protein coding gene located in chromosome 1 that encodes a lysosomal member protein (36). Mutations in *GBA* that leads to aberrant function may cause Gauche's disease and Parkinson disease (37). Study hypothesised that *GBAP1*, acting as competing endogenous RNA, could promote *GBA* expression in GC through the association of miRNA-212-3p, and suggested that high expression of *GBA* is associated with poor GC prognosis (38). In addition, according to a GWAS study, a SNP (rs4460629) in *GBA* was speculated as causal variant for GC (39). However, to the best of our knowledge, *GBA* have not been reported as GC susceptibility gene.

SBDS

SBDS encodes a highly conserved protein that plays an important role in ribosome biogenesis. Biallelic mutations in *SBDS* gene have been mostly associated with Shwachman-Diamond syndrome (SDS, OMIM #260400), an autosomal recessive disorder with high risk of myelodysplastic syndrome (MDS) and acute myeloid leukaemia (AML) (40). However, Jason *et al.* speculated that germline mutations in *SBDS* gene may increase cancer predisposition (41). A case report introduced a female diagnosed as SDS three month after birth, and she developed poorly differentiated ductal breast carcinoma on age of 30 years. This patient carries two germline variants in *SBDS* gene, c.183_184TA>CT and c.258+2T>C, both inherited from her parents (42). Additionally, a nonsense germline variant in *SBDS* was detected in an early-onset gastric cancer patient with diffuse type adenocarcinoma (22). These evidences suggested the deleterious *SBDS* germline mutation may hold the potential to increase cancer susceptibility.

PRDM16

PRDM16 is a protein coding gene that resides in chromosome 1p36. *PRDM16* plays critical roles to the maintenance of brown adipocytes, adult hematopoietic and neural stem cells. Moreover, *PRDM16* has been reported abnormally expressed and rearranged in various cancers including GC (43-46). Additionally, inhibition of *PRDM16* by hypoxia-induced-miR-24 expression could promote GC cell proliferation and migration (47).

SPINK1

SPINK1 resides in chromosomal region 5q32, and encodes a 79 amino acid peptide that includes a 23 amino acid signal peptide. It is widely expressed in a variety of cancer malignancies, and associated with poor prognosis in majority of solid tumours (48). The serum levels of *SPINK1* may help to distinguish high-risk groups of gastric cancer and discriminate advanced gastric cancer (49). Additionally, protein expression of *SPINK1* in tumour tissue may have prognostic value on gastric cancer patients (50).

VWF

VWF encodes a protein called Von Willebrand factor, a large multimeric glycoprotein found in blood plasma, platelet α -granules, and subendothelial connective tissue (51). *VWF* plays important roles in hemostasis through mediating adhesion of platelets to subendothelial connective tissue, and by association with blood clotting factor VIII. Therefore, dysregulation of *VWF* could lead to hemostatic disorders or complications including including stroke,

myocardial infarction and diabetes (52). Study showed that cancer cell-derived VWF promoted gastric cancer metastasis (53).

LIPH

LIPH is a member of the mammalian triglyceride lipase family. The expression of LIPH was found to be involved in papillary thyroid carcinoma development (54) and mRNA level of LIPH may act as a prognostic marker for esophageal adenocarcinoma, lung cancer and breast cancer (55-57). Study demonstrated that CIRH1A plays regulatory roles in proliferation and apoptosis of colorectal cancer cell (58).

GSS

GSS, a protein coding gene located in chromosome 20, encodes glutathione synthetase, whose deficiency may lead to inborn error of glutathione metabolism with autosomal recessive inheritance pattern (59). GSS is important in glutathione synthesis in several tissues and under stressful conditions (60). Seven intronic polymorphic variants in *GSS* gene were reported to correlate with lung cancer survival (61). Additionally, GSS protein expression was increased in colonic tumour tissue compare with normal mucosa (62), whereas the protein level of GSS may assist the early detection of colorectal cancer (63). Moreover, SNPs in *GSS* gene expressed prognostic and predictive value in treatment of bladder cancer (64).

SBF1

SBF1 encodes a member of the protein-tyrosine phosphatase family. Mutations in this gene have been reported to be associated with Charcot-Marie-Tooth disease 4B3 (65). Few studies reported the *SBF1* implication in cancer. An *In vitro* study indicated that *SBF1* expression correlated with gemcitabine sensitivity in head and neck squamous cell carcinoma (66).

SLC22A5

SLC22A5 (also known as *OCTN2*), located within a cluster region on chromosome 5q, is a sodium dependent carnitine transporter and expressed in various human tissues (67). Genetic mutations in *SLC22A5* that impairs normal expression of *OCTN2* may cause systemic carnitine deficiency (68). Case-control study demonstrated that G risks allele at SNP rs27437 in *SLC22A5* gene was associated with colorectal cancer risk (69). Study on GIST patients suggested that investigation on minor allele of *SLC22A5* may assist to optimise imatinib therapy (70, 71).

TYR

TYR is a protein coding gene located in chromosome 11q14.3. It encodes tyrosinase, a membrane glycoprotein widely expressed in mammalian tissues and essential in melanin production (72). Mutations in *TYR* that result in abnormal expression or dysfunction of tyrosinase often lead to oculocutaneous albinism with different levels of severity in different ethnic groups (73). Studies on identification of potential cancer susceptibility variants

suggested that a minor allele variant p.R402Q on *TYR* gene associates with increased risk of cutaneous cancer (74-76), whereas heterozygous polymorphism of *TYR* gene in codon 192 correlated with decreased risk of prostate cancer metastases (77) as well as associated with elevated squamous cell skin carcinoma risk in Caucasian population (78).

HNF1A

HNF1A, a protein coding gene, was initially discovered in the liver and was subsequently identified to be widely expressed in several tissues including the pancreas, kidney, intestine and stomach (79). Mutations in *HNF1A* are a common cause of maturity-onset diabetes of the young (MODY) (80). Study on investigating colorectal cancer predisposing gene through panel sequencing found a germline variant of *HNF1A* c.92G>A in colorectal patient who met Amsterdam II criteria (81). Moreover, another germline variant in *HNF1A* c.1018C>G was detected in early-onset female lung cancer patient (82). In addition, a germline deletion spanning exon2 to 3 was observed in a family with MODY type 3 and primary liver cell neoplasia (83). A study on families with familial liver adenomatosis also found heterozygous truncating germline variants in *HNF1A* (84).

Table 3. Characteristics of 29 affected individuals with HDGC in 28 families

Subject ID	Age of diagnosis, y	Cancers family history	Differentiation	Cellular morphology	TNM stage	Lauren classification	IGCLC criteria
RB809	31	SDR Liver cancer 70~y, SDR Leukaemia, TDR Liver cancer 70~y	moderately/poorly	signet-ring cell	T4NxM1	Mixed	2
RB802	31	SDR colon cancer 60y	poorly	signet-ring cell	TxNxM1	Diffuse	2
RB665	32	NA	poorly	signet-ring cell	T4N2M0	Diffuse	2
RB666	29	NA	poorly	signet-ring cell	T4aN2M0	Mixed	2
RB798	28	SDR Lung cancer 50~y, SDR esophageal cancer 70~y	poorly	NA	T4N3M0	Diffuse	2
RB694	35	NA	poorly	signet-ring cell	T3N2M0	Diffuse	2
RB659	29	NA	poorly	signet-ring cell	T4aN3bM0	Mixed	2
RB651	26	FDR Gastric cancer 49y, FDR Lung cancer 50y, TDR Gastric cancer 40y	poorly	NA	T4N3M0	Diffuse	2
RB808	26	NA	poorly	signet-ring cell	T4N2M0	Diffuse	2
RB803	34	NA	poorly	NA	TxNxM1	Diffuse	2
RB799	31	SDR Gastric cancer 30~y, SDR Renal carcinoma 50~y	poorly	signet-ring cell	TxNxM1	Diffuse	2
RB800	34	NA	NA	NA	T4aN3M0	Diffuse	2
RB662	34	FDR Colon cancer 60y, TDR Gastric cancer 80y, TDR Gastric cancer 80y	poorly	signet-ring cell	T4N3M0	Diffuse	2
RB810	28	SDR cancer 60~y, SDR Gastric cancer 50~y	poorly	signet-ring cell	T3N0M0	Diffuse	2
RB805	29	SDR esophageal cancer 60~y	poorly	signet-ring cell	T4N0M0	Diffuse	2
RB650	32	NA	poorly	signet-ring cell	T2N0M0	Diffuse	2
RB692	20	SDR Breast cancer 48y	poorly	signet-ring cell	T4N2M0	Diffuse	2
RB672	24	NA	poorly	signet-ring cell	T2N0M0	Diffuse	2
RB661	34	NA	poorly	signet-ring cell	T4N2M0	Diffuse	2

RB804	29	SDR Lung cancer 60y	poorly	signet-ring cell	T4N1M0	Diffuse	2
RB801	34	FDR Leiomyosarcoma 55y	poorly	NA	T3N0M0	Diffuse	2
RB684	28	SDR Gastric cancer 60~y	NA	signet-ring cell	T1N0M0	Diffuse	2
RB811	24	NA	poorly	signet-ring cell	T4N1M0	Diffuse	2
RB680	31	FDR Liver cancer 52y	poorly	signet-ring cell	T1N0M0	Diffuse	2
RB797	49	FDR Gastric cancer 26y, SDR cancer 60y, TDR cancer	NA	signet-ring cell	T1N0M0	Diffuse	1
RB681	35	SDR Pancreatic cancer 50~y	poorly	signet-ring cell	T3N3M0	Diffuse	2
RB685	25	TDR Gastric cancer 40~y	poorly	signet-ring cell	T1N0M0	Diffuse	2
RB807	22	SDR Liver cancer	poorly	signet-ring cell	T3N3M0	Diffuse	2
RB806	29	SDR nasopharyngeal cancer	NA	signet-ring cell	T1N0M0	Diffuse	2

IGCLC criteria:

1. 2 GC cases regardless of age, at least one confirmed DGC
2. One case of DGC < 40
3. Personal or family history of DGC and lobular breast cancer, one diagnosis < 50

SDR: second-degree relative; TDR: third-degree relative, FDR: first-degree relative; “~” indicates approximate age of diagnosis; “NA” in Differentiation column and “x” in TNM stage column represent information not available.

Table 4. Potential deleterious variant candidates observed in 29 gastric cancer patients

Patient ID	Gene	Mutation type	cDNA change	Protein change	Het/ Hom	Refseq ID	Classification criteria
RB809	<i>ATP7B</i>	missense	c.2755C>G	p.Arg919Gly	het	NM_000053.3	HGMD-DM
RB666	<i>GBA</i>	missense	c.476G>A	p.Arg159Gln	het	NM_001005741.2	HGMD-DM
RB650	<i>POLH</i>	missense	c.1603A>G	p.Lys535Glu	het	NM_006502.2	HGMD-DM
RB662	<i>MYH7</i>	missense	c.5342G>A	p.Arg1781His	het	NM_000257.2	ClinVar-Pathogenic
RB694	<i>CIRH1A</i>	missense	c.1693C>T	p.Arg565Trp	het	NM_032830.2	ClinVar-Pathogenic
RB799/ RB681	<i>F8</i>	missense	c.3169G>A	p.Glu1057Lys	het	NM_000132.3	ClinVar-Pathogenic
RB806	<i>GSS</i>	missense	c.941C>T	p.Pro314Leu	het	XM_005260406.1	ClinVar-Pathogenic
RB798	<i>KCNQ4</i>	missense	c.546C>G	p.Phe182Leu	het	NM_004700.3	ClinVar-Pathogenic
RB665	<i>LIPH</i>	missense	c.742C>A	p.His248Asn	het	NM_139248.2	ClinVar-Pathogenic
RB811	<i>PRDM16</i>	missense	c.3301G>A	p.Val1101Met	het	NM_022114.3	ClinVar-Pathogenic
RB650	<i>PTS</i>	missense	c.259C>T	p.Pro87Ser	het	NM_000317.2	ClinVar-Pathogenic
RB801	<i>SBF1</i>	missense	c.4768A>G	p.Thr1590Ala	het	NM_002972.2	ClinVar-Pathogenic
RB662	<i>SLC22A5</i>	missense	c.51C>G	p.Phe17Leu	het	NM_003060.3	ClinVar-Pathogenic
RB810	<i>SPINK1</i>	Splice donor	c.194+2T>C	.	het	XM_005268501.1	ClinVar-Pathogenic
RB665	<i>TYR</i>	missense	c.230G>A	p.Arg77Gln	het	NM_000372.4	ClinVar-Pathogenic
RB808	<i>VWF</i>	missense	c.2561G>A	p.Arg854Gln	het	NM_000552.3	ClinVar-Pathogenic
RB804	<i>HNF1A</i>	missense	c.1854C>G	p.Ile618Met	het	NM_000545.5	ClinVar-Likely pathogenic
RB802	<i>SBDS</i>	Splice donor	c.258+2T>C	.	het	NM_016038.2	ClinVar-Likely pathogenic
RB651	<i>APC</i>	Splice donor	c.-19+1G>A	.	het	NM_001127511.2	Candidate gene-affecting splicing
RB805	<i>BRCA1</i>	missense	c.5518G>T	p.Asp1840Tyr	het	NM_007294.3	Candidate gene-affecting protein function
RB665	<i>BRCA1</i>	missense	c.3448C>T	p.Pro1150Ser	het	NM_007294.3	Candidate gene-affecting protein function
RB803	<i>CDH1</i>	missense	c.2053G>A	p.Val685Met	het	NM_004360.4	Candidate gene-affecting protein function

RB809	<i>CDH1</i>	Splice acceptor	c.2165-1G>A	.	het	NM_004360.4	Candidate gene-affecting protein function
RB684	<i>CTNNA1</i>	Frameshift	c.1975_1976del	p.Glu659ArgfsTer42	het	NM_001903.2	Candidate gene-truncation

Het: heterozygosity, Hom: homozygosity, HGMD: Human Gene Mutation Database, DM: disease-causing mutations

Discussion

HDGC was firstly described in 1964, and it is one of the familial forms of gastric carcinoma, characterised by multiple microscopic foci of intramucosal signet-ring cell and was considered as a highly invasive tumour with poor prognosis (85, 86). It was confirmed by linkage and mutational analysis that germline mutations in *CDH1* play an important role in multi-generational incidence of such deleterious malignancy (87). Subsequently, approximately a hundred other HDGC families were reported one after another in the next decades (85). Moreover, studies suggested that the cumulative lifetime risks of developing gastric cancer by 80 years for male and female individuals carrying pathogenic *CDH1* mutations are 40% to 67% and 63% to 83% respectively. In addition, female carriers may have 39% to 52% risk of developing breast cancer, predominantly characterised as lobular type (8, 11, 88). Considering the high penetrance of pathogenic *CDH1* mutation carriers, with the purpose of identification of this dominantly inherited familial cancer syndrome, a consensus statement of clinical diagnosis and management of HDGC was formed to improve clinical practice. Therefore, three major characteristics, early-onset, diffuse type classification and family clustering, have been widely accepted to be clinical diagnostic criteria for HDGC and screening conditions for pathogenic germline variants in *CDH1* (9). Such consensus was later revised and updated to recommend patients with DGC diagnosed before 40 years or two GC in first and second degree relatives with one confirmed DGC before age of 50-year or personal / family history of DGC and lobular breast cancer with one diagnosed before 50 years to be considered for clinical diagnosis of HDGC and further evaluated by genetic testing to identify *CDH1* genetic status (4, 10). However, a study on exploiting pathogenic germline *CDH1* mutation in Korean HDGC patients demonstrated that the detection rate of *CDH1* carrier is only 4.3%, which is no more than half of the detection rate (10%-18%) in low prevalence gastric cancer countries (4, 89). This evidence suggested that the previously established *CDH1* germline mutation screening criteria established based on low prevalence of gastric cancer population may not be appropriate to identify *CDH1* pathogenic variants in HDGC patients in East Asian countries with high incidence of gastric cancer.

In current study, with the aim of investigating novel susceptibility genes and potential pathogenic germline variants that may contribute to occurrence and development of early-onset diffuse-type gastric cancer, 28 GC patients in this cohort had confirmed diagnosis of GC with diffuse-type before 40-year, whereas one patient, diagnosed as DGC at 49 years, is a first degree relative to one of the 28 GC patients. All 29 GC patients fulfilled the 2015 HDGC clinical diagnosis criteria. In this study, 25 pathogenic variant candidates in 22 genes were highlighted by our filtering process. Since diffuse-type GC may attribute to genetic abnormalities in cell adhesion function in relation to E-cadherin and WNT signalling (36), four variants discovered in this study, *CDH1* c.2165-1G>A (NM_004360.4), *CDH1* c.2053G>A (NM_004360.4), *CTNNA1* c.1975_1976del (NM_001903.2) APC c.-19+1G>A (NM_001127511.2), were suspicious to be gastric cancer predisposing genetic factors. Although previously reported *CDH1* pathogenic mutations were not observed in this study, a novel splice site variant c.2165-1G>A in *CDH1* with likelihood of pathogenicity was discovered in one patient, suggesting a low detection rate of potential *CDH1* pathogenic variants, 3.4% (1/29), in Chinese HDGC patients. This result is consistent with the above

study mentioned Korean population study (89), further indicating that genetic testing for germline mutation in *CDH1* gene alone may not be enough to identify genetic and molecular basis of HDGC in high incidence gastric cancer area, especially in China.

In addition, convincing pathogenic mutations in predisposing genes associated with increased gastric cancer were neither discovered. Therefore, one potential explanation could be susceptibility genes that increase gastric cancer risks in Chinese population may differ from that of in Caucasian population. In addition, many cancer-predisposing gene candidates other than *CDH1*, such as *CTNNA1*, *INSR*, *FBXO24*, *DOT1L*, *BRCA1*, *PALB2*, *RAD51C*, *MSH2*, *ATR*, *NBN* and *RECQL5* were previously discovered and reported in families with HDGC with various ethnic background (11, 13, 15, 25, 90). However, the pathogenicity of such variant candidates require for further verification through functional studies and co-segregation analysis on family members. In addition, families without *CDH1* germline pathogenic mutation who met HDGC clinical diagnosis criteria may be benefit from cancer-predisposing-gene panel screening. Therefore, considering the low detection rate of *CDH1* germline mutation in early-onset DGC patients, genetic screening for Chinese gastric cancer patients who meet the HDGC clinical diagnosis criteria using a multi-gene panel including other gastric cancer associated predisposing-gene are likely to be more cost effective.

Nevertheless, the information of structural aberration and large deletion and insertion were not able to be precisely acquired by whole exome sequencing analysis in this study, therefore, using a combination of multiplex ligation-dependent probe amplification (MLPA) and RT-PCR analysis may potentially help identify more genetic alteration in Chinese HDGC patients, since such methodology has ever discovered a large genomic deletion (c.164-?_387+?del) in a Japanese familial gastric cancer family, which results in the loss of exon 3 in *CDH1* and a E-cadherin truncating product (91). Such hypothesis requires for large number cohort study verification containing consecutive Chinese gastric cancer patients.

Furthermore, the inconclusive results in this study warrant future investigation on gastric cancer susceptibility gene discovery that cohort selection may require more stringent conditions rather than early-onset diffuse-type GC patients.

Conclusions

This study suggested that *CDH1* c.2165-1G>A may act as a gastric cancer predisposing variant. In addition, to further investigate molecular mechanisms of early-onset gastric cancer, one may consider 22 genes observed in this study. Furthermore, the inconclusive results in this study warrant future investigation on gastric cancer susceptibility gene discovery that cohort selection may require more stringent conditions rather than early-onset diffuse-type GC patients.

List of abbreviations

Diffuse gastric cancer (DGC), hereditary diffuse gastric cancer (HDGC), gastric cancer (GC), whole exome sequencing (WES), signet-ring cell carcinoma (SRCC), the International Gastric Cancer Linkage Consortium (IGCLC), Genome Analysis tool kit (GATK), allele frequency (AF), pathogenic / likely pathogenic (P/LP), disease-causing mutations (DM), variant of uncertain significance (VUS)

Declarations

Ethics approval and consent to participate

A written informed consent to participate in the study was obtained from each subject in accordance with the declaration of Helsinki principles. Each patient or the patient's family was fully informed of the investigational nature of this study and provided their written, informed consent. The study protocol was approved by the Medical Ethics Committee, Beijing Cancer Hospital (approve number #2018KT64).

Consent for publication

Not applicable.

Availability of data and materials

The datasets used and analysed during the current study are available from the corresponding author on reasonable request

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by the National Key R&D Program of China (No. 2018YFC0910700), Shen Shuxiang Medical Research Fund (No. 910247), Three-year-rotating Budget Program, Beijing Municipal Commission of Health and Family Planning, Clinical Medicine Development Special Funding of Beijing Municipal Administration (ZYLX201701), Science Foundation of Peking University Cancer Hospital (2020-15).

Acknowledgments

The authors thank all patients who participated in this study.

Authors' contributions

YF participated in the data analysis of the study and drafted and finalised the manuscript, MZ helped with study data collection and analysis and manuscript preparation, WL, SL, CQ, MM, MY, WC, helped with study sample and data collection, XX participated in the design of the study and manuscript editing, SJ conceived the study, and participated in its design and coordination and helped to draft and edit the manuscript. All authors have read and approved the final manuscript.

Reference

1. de Martel C, Forman D, Plummer M. Gastric cancer: epidemiology and risk factors. *Gastroenterology clinics of North America*. 2013;42:219-40.
2. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology*. 2014;23:700-13.
3. Bray F, Ren JS, Masuyer E, Ferlay J. Global estimates of cancer prevalence for 27 sites in the adult population in 2008. *International journal of cancer Journal international du cancer*. 2013;132:1133-45.
4. van der Post RS, Vogelaar IP, Carneiro F, Guilford P, Huntsman D, Hoogerbrugge N, et al. Hereditary diffuse gastric cancer: updated clinical guidelines with an emphasis on germline CDH1 mutation carriers. *Journal of medical genetics*. 2015;52:361-74.
5. Guilford P, Blair V, More H, Humar B. A short guide to hereditary diffuse gastric cancer. *Hereditary cancer in clinical practice*. 2007;5:183-94.
6. Pedrazzani C, Corso G, Marrelli D, Roviello F. E-cadherin and hereditary diffuse gastric cancer. *Surgery*. 2007;142:645-57.
7. Shenoy S. CDH1 (E-Cadherin) Mutation and Gastric Cancer: Genetics, Molecular Mechanisms and Guidelines for Management. *Cancer management and research*. 2019;11:10477-86.
8. Pharoah PD, Guilford P, Caldas C, International Gastric Cancer Linkage C. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. *Gastroenterology*. 2001;121:1348-53.
9. Caldas C, Carneiro F, Lynch HT, Yokota J, Wiesner GL, Powell SM, et al. Familial gastric cancer: overview and guidelines for management. *Journal of medical genetics*. 1999;36:873-80.
10. Fitzgerald RC, Hardwick R, Huntsman D, Carneiro F, Guilford P, Blair V, et al. Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. *Journal of medical genetics*. 2010;47:436-44.
11. Hansford S, Kaurah P, Li-Chang H, Woo M, Senz J, Pinheiro H, et al. Hereditary Diffuse Gastric Cancer Syndrome: CDH1 Mutations and Beyond. *JAMA oncology*. 2015;1:23-32.
12. 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.
13. Majewski IJ, Kluijt I, Cats A, Scerri TS, de Jong D, Kluin RJ, et al. An alpha-E-catenin (CTNNA1) mutation in hereditary diffuse gastric cancer. *The Journal of pathology*. 2013;229:621-9.
14. Rona KA, Schwameis K, Zehetner J, Samakar K, Green K, Samaan J, et al. Gastric cancer in the young: An advanced disease with poor prognostic features. *Journal of surgical oncology*. 2017;115:371-5.
15. Sahasrabudhe R, Lott P, Bohorquez M, Toal T, Estrada AP, Suarez JJ, et al. Germline Mutations in PALB2, BRCA1, and RAD51C, Which Regulate DNA Recombination Repair, in Patients With Gastric Cancer. *Gastroenterology*. 2017;152:983-6 e6.
16. (NCCN) NCCN. Clinical Practice Guidelines in Oncology. Gastric cancer, Version 4. 2019. [cited 2020 March 17]; Available from: https://www.nccn.org/professionals/physician_gls/pdf/gastric.pdf
17. Marques M, Ramalho R, Baldaque-Silva F, Macedo G. Novel mutation identified in Cowden syndrome presenting as a gastric adenocarcinoma. *Clinics and research in hepatology and gastroenterology*. 2013;37:e131-2.

18. Ni Y, Zbuk KM, Sadler T, Patocs A, Lobo G, Edelman E, et al. Germline mutations and variants in the succinate dehydrogenase genes in Cowden and Cowden-like syndromes. *American journal of human genetics*. 2008;83:261-8.
19. Le Marechal C, Masson E, Chen JM, Morel F, Ruzniewski P, Levy P, et al. Hereditary pancreatitis caused by triplication of the trypsinogen locus. *Nature genetics*. 2006;38:1372-4.
20. Felderbauer P, Schnekenburger J, Lebert R, Bulut K, Parry M, Meister T, et al. A novel A121T mutation in human cationic trypsinogen associated with hereditary pancreatitis: functional data indicating a loss-of-function mutation influencing the R122 trypsin cleavage site. *Journal of medical genetics*. 2008;45:507-12.
21. Huang DS, Tao HQ, He XJ, Long M, Yu S, Xia YJ, et al. Prevalence of deleterious ATM germline mutations in gastric cancer patients. *Oncotarget*. 2015;6:40953-8.
22. Slavin T, Neuhausen SL, Rybak C, Solomon I, Nehoray B, Blazer K, et al. Genetic Gastric Cancer Susceptibility in the International Clinical Cancer Genomics Community Research Network. *Cancer genetics*. 2017;216-217:111-9.
23. Orloff M, Peterson C, He X, Ganapathi S, Heald B, Yang YR, et al. Germline mutations in MSR1, ASCC1, and CTHRC1 in patients with Barrett esophagus and esophageal adenocarcinoma. *Jama*. 2011;306:410-9.
24. Xu J, Zheng SL, Komiya A, Mychaleckyj JC, Isaacs SD, Hu JJ, et al. Germline mutations and sequence variants of the macrophage scavenger receptor 1 gene are associated with prostate cancer risk. *Nature genetics*. 2002;32:321-5.
25. Fewings E, Larionov A, Redman J, Goldgraben MA, Scarth J, Richardson S, et al. Germline pathogenic variants in PALB2 and other cancer-predisposing genes in families with hereditary diffuse gastric cancer without CDH1 mutation: a whole-exome sequencing study. *The lancet Gastroenterology & hepatology*. 2018;3:489-98.
26. Kakiuchi M, Nishizawa T, Ueda H, Gotoh K, Tanaka A, Hayashi A, et al. Recurrent gain-of-function mutations of RHOA in diffuse-type gastric carcinoma. *Nature genetics*. 2014;46:583-7.
27. National Center for Biotechnology Information. ClinVar; [VCV000532459.2]. [cited 2020 March 16]; Available from: <https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000617777.1>
28. National Center for Biotechnology Information. ClinVar; [VCV000617777.1]. [cited 2020 March 23]; Available from: <https://www.ncbi.nlm.nih.gov/clinvar/variation/VCV000617777.1>
29. Bodian DL, McCutcheon JN, Kothiyal P, Huddleston KC, Iyer RK, Vockley JG, et al. Germline variation in cancer-susceptibility genes in a healthy, ancestrally diverse cohort: implications for individual genome sequencing. *PloS one*. 2014;9:e94554.
30. Lichon V, Khachemoune A. Xeroderma pigmentosum: beyond skin cancer. *Journal of drugs in dermatology : JDD*. 2007;6:281-8.
31. Johnson RE, Prakash S, Prakash L. Efficient bypass of a thymine-thymine dimer by yeast DNA polymerase, Poleta. *Science*. 1999;283:1001-4.
32. Opletalova K, Bourillon A, Yang W, Pouvelle C, Armier J, Despras E, et al. Correlation of phenotype/genotype in a cohort of 23 xeroderma pigmentosum-variant patients reveals 12 new disease-causing POLH mutations. *Human mutation*. 2014;35:117-28.
33. Rostami P, Zendejdel K, Shirkoohi R, Ebrahimi E, Ataei M, Imanian H, et al. Gene Panel Testing in Hereditary Breast Cancer. *Archives of Iranian medicine*. 2020;23:155-62.
34. Girard E, Eon-Marchais S, Olasso R, Renault AL, Damiola F, Dondon MG, et al. Familial breast cancer and DNA repair genes: Insights into known and novel susceptibility genes from the GENESIS study, and implications for multigene panel testing. *International journal of cancer Journal international du cancer*. 2019;144:1962-74.

35. Di Lucca J, Guedj M, Lacapere JJ, Fagnoli MC, Bourillon A, Dieude P, et al. Variants of the xeroderma pigmentosum variant gene (POLH) are associated with melanoma risk. *European journal of cancer*. 2009;45:3228-36.
36. Barneveld RA, Keijzer W, Tegelaers FP, Ginns EI, Geurts van Kessel A, Brady RO, et al. Assignment of the gene coding for human beta-glucocerebrosidase to the region q21-q31 of chromosome 1 using monoclonal antibodies. *Human genetics*. 1983;64:227-31.
37. Velayati A, Yu WH, Sidransky E. The role of glucocerebrosidase mutations in Parkinson disease and Lewy body disorders. *Current neurology and neuroscience reports*. 2010;10:190-8.
38. Ma G, Liu H, Du M, Zhang G, Lin Y, Ge Y, et al. A genetic variation in the CpG island of pseudogene GBAP1 promoter is associated with gastric cancer susceptibility. *Cancer*. 2019;125:2465-73.
39. Lee JH, Kim Y, Choi JW, Kim YS. Genetic variants and risk of gastric cancer: a pathway analysis of a genome-wide association study. *SpringerPlus*. 2015;4:215.
40. Warren AJ. Molecular basis of the human ribosomopathy Shwachman-Diamond syndrome. *Advances in biological regulation*. 2018;67:109-27.
41. Stumpff J, Ghule PN, Shimamura A, Stein JL, Greenblatt M. Spindle microtubule dysfunction and cancer predisposition. *Journal of cellular physiology*. 2014;229:1881-3.
42. Singh SA, Vlachos A, Morgenstern NJ, Ouansafi I, Ip W, Rommens JM, et al. Breast cancer in a case of Shwachman Diamond syndrome. *Pediatric blood & cancer*. 2012;59:945-6.
43. Meng X, Zhao Y, Liu J, Wang L, Dong Z, Zhang T, et al. Comprehensive analysis of histone modification-associated genes on differential gene expression and prognosis in gastric cancer. *Experimental and therapeutic medicine*. 2019;18:2219-30.
44. White PS, Thompson PM, Gotoh T, Okawa ER, Igarashi J, Kok M, et al. Definition and characterization of a region of 1p36.3 consistently deleted in neuroblastoma. *Oncogene*. 2005;24:2684-94.
45. Martinez-Climent JA, Alizadeh AA, Segraves R, Blesa D, Rubio-Moscardo F, Albertson DG, et al. Transformation of follicular lymphoma to diffuse large cell lymphoma is associated with a heterogeneous set of DNA copy number and gene expression alterations. *Blood*. 2003;101:3109-17.
46. Fei LR, Huang WJ, Wang Y, Lei L, Li ZH, Zheng YW, et al. PRDM16 functions as a suppressor of lung adenocarcinoma metastasis. *Journal of experimental & clinical cancer research : CR*. 2019;38:35.
47. Yang L, Zhang W, Wang Y, Zou T, Zhang B, Xu Y, et al. Hypoxia-induced miR-214 expression promotes tumour cell proliferation and migration by enhancing the Warburg effect in gastric carcinoma cells. *Cancer letters*. 2018;414:44-56.
48. Rasanen K, Ikonen O, Koistinen H, Stenman UH. Emerging Roles of SPINK1 in Cancer. *Clinical chemistry*. 2016;62:449-57.
49. Kemik O, Kemik A, Sumer A, Almali N, Gurluler E, Gures N, et al. The relationship between serum tumor-associated trypsin inhibitor levels and clinicopathological parameters in patients with gastric cancer. *European review for medical and pharmacological sciences*. 2013;17:2923-8.
50. Wiksten JP, Lundin J, Nordling S, Kokkola A, Stenman UH, Haglund C. High tissue expression of tumour-associated trypsin inhibitor (TATI) associates with a more favourable prognosis in gastric cancer. *Histopathology*. 2005;46:380-8.
51. Sadler JE. Biochemistry and genetics of von Willebrand factor. *Annual review of biochemistry*. 1998;67:395-424.
52. Xiang Y, Hwa J. Regulation of VWF expression, and secretion in health and disease. *Current opinion in hematology*. 2016;23:288-93.
53. Yang AJ, Wang M, Wang Y, Cai W, Li Q, Zhao TT, et al. Cancer cell-derived von Willebrand factor enhanced metastasis of gastric adenocarcinoma. *Oncogenesis*. 2018;7:12.

54. Li Y, Zhou X, Zhang Q, Chen E, Sun Y, Ye D, et al. Lipase member H is a downstream molecular target of hypoxia inducible factor-1alpha and promotes papillary thyroid carcinoma cell migration in BCPAP and KTC-1 cell lines. *Cancer management and research*. 2019;11:931-41.
55. Ishimine H, Zhou R, Sumitomo K, Ito Y, Seki Y, Yoshida Y, et al. Lipase member H frequently overexpressed in human esophageal adenocarcinomas. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2016;37:2075-81.
56. Cui M, Jin H, Shi X, Qu G, Liu L, Ding X, et al. Lipase member H is a novel secreted protein associated with a poor prognosis for breast cancer patients. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2014;35:11461-5.
57. Seki Y, Yoshida Y, Ishimine H, Shinozaki-Ushiku A, Ito Y, Sumitomo K, et al. Lipase member H is a novel secreted protein selectively upregulated in human lung adenocarcinomas and bronchioloalveolar carcinomas. *Biochemical and biophysical research communications*. 2014;443:1141-7.
58. Guo F, Chen JJ, Tang WJ. CIRH1A augments the proliferation of RKO colorectal cancer cells. *Oncology reports*. 2017;37:2375-81.
59. Xia H, Ye J, Wang L, Zhu J, He Z. A case of severe glutathione synthetase deficiency with novel GSS mutations. *Brazilian journal of medical and biological research = Revista brasileira de pesquisas medicas e biologicas / Sociedade Brasileira de Biofisica* [et al]. 2018;51:e6853.
60. Lu SC. Regulation of glutathione synthesis. *Molecular aspects of medicine*. 2009;30:42-59.
61. Sun Z, Chen J, Aakre J, Marks RS, Garces YY, Jiang R, et al. Genetic variation in glutathione metabolism and DNA repair genes predicts survival of small-cell lung cancer patients. *Annals of oncology : official journal of the European Society for Medical Oncology / ESMO*. 2010;21:2011-6.
62. Kim AD, Zhang R, Han X, Kang KA, Piao MJ, Maeng YH, et al. Involvement of glutathione and glutathione metabolizing enzymes in human colorectal cancer cell lines and tissues. *Molecular medicine reports*. 2015;12:4314-9.
63. Strohkamp S, Gemoll T, Humborg S, Hartwig S, Lehr S, Freitag-Wolf S, et al. Protein levels of clusterin and glutathione synthetase in platelets allow for early detection of colorectal cancer. *Cellular and molecular life sciences : CMLS*. 2018;75:323-34.
64. Ke HL, Lin J, Ye Y, Wu WJ, Lin HH, Wei H, et al. Genetic Variations in Glutathione Pathway Genes Predict Cancer Recurrence in Patients Treated with Transurethral Resection and Bacillus Calmette-Guerin Instillation for Non-muscle Invasive Bladder Cancer. *Annals of surgical oncology*. 2015;22:4104-10.
65. Nakhro K, Park JM, Hong YB, Park JH, Nam SH, Yoon BR, et al. SET binding factor 1 (SBF1) mutation causes Charcot-Marie-Tooth disease type 4B3. *Neurology*. 2013;81:165-73.
66. Zhu MH, Ji SL, Zhang CY, Cui L, Xiong L, Zheng HL. DNA microarray reveals ZNF195 and SBF1 are potential biomarkers for gemcitabine sensitivity in head and neck squamous cell carcinoma cell lines. *International journal of clinical and experimental pathology*. 2014;7:1514-23.
67. Koepsell H. The SLC22 family with transporters of organic cations, anions and zwitterions. *Molecular aspects of medicine*. 2013;34:413-35.
68. Koepsell H, Lips K, Volk C. Polyspecific organic cation transporters: structure, function, physiological roles, and biopharmaceutical implications. *Pharmaceutical research*. 2007;24:1227-51.

69. Zou D, Lou J, Ke J, Mei S, Li J, Gong Y, et al. Integrative expression quantitative trait locus-based analysis of colorectal cancer identified a functional polymorphism regulating SLC22A5 expression. *European journal of cancer*. 2018;93:1-9.
70. Angelini S, Pantaleo MA, Ravegnini G, Zenesini C, Cavrini G, Nannini M, et al. Polymorphisms in OCTN1 and OCTN2 transporters genes are associated with prolonged time to progression in unresectable gastrointestinal stromal tumours treated with imatinib therapy. *Pharmacological research : the official journal of the Italian Pharmacological Society*. 2013;68:1-6.
71. Qiu HB, Zhuang W, Wu T, Xin S, Lin CZ, Ruan HL, et al. Imatinib-induced ophthalmological side-effects in GIST patients are associated with the variations of EGFR, SLC22A1, SLC22A5 and ABCB1. *The pharmacogenomics journal*. 2018;18:460-6.
72. Wang N, Hebert DN. Tyrosinase maturation through the mammalian secretory pathway: bringing color to life. *Pigment cell research*. 2006;19:3-18.
73. Dolinska MB, Kus NJ, Farney SK, Wingfield PT, Brooks BP, Sergeev YV. Oculocutaneous albinism type 1: link between mutations, tyrosinase conformational stability, and enzymatic activity. *Pigment cell & melanoma research*. 2017;30:41-52.
74. Hu HH, Guedj M, Descamps V, Jouary T, Bourillon A, Ezzedine K, et al. Assessment of tyrosinase variants and skin cancer risk in a large cohort of French subjects. *Journal of dermatological science*. 2011;64:127-33.
75. Gudbjartsson DF, Sulem P, Stacey SN, Goldstein AM, Rafnar T, Sigurgeirsson B, et al. ASIP and TYR pigmentation variants associate with cutaneous melanoma and basal cell carcinoma. *Nature genetics*. 2008;40:886-91.
76. Ibarrola-Villava M, Hu HH, Guedj M, Fernandez LP, Descamps V, Basset-Seguín N, et al. MC1R, SLC45A2 and TYR genetic variants involved in melanoma susceptibility in southern European populations: results from a meta-analysis. *European journal of cancer*. 2012;48:2183-91.
77. Luscombe CJ, French ME, Liu S, Saxby MF, Jones PW, Fryer AA, et al. Outcome in prostate cancer associations with skin type and polymorphism in pigmentation-related genes. *Carcinogenesis*. 2001;22:1343-7.
78. Nan H, Kraft P, Hunter DJ, Han J. Genetic variants in pigmentation genes, pigmentary phenotypes, and risk of skin cancer in Caucasians. *International journal of cancer Journal international du cancer*. 2009;125:909-17.
79. Pontoglio M, Barra J, Hadchouel M, Doyen A, Kress C, Bach JP, et al. Hepatocyte nuclear factor 1 inactivation results in hepatic dysfunction, phenylketonuria, and renal Fanconi syndrome. *Cell*. 1996;84:575-85.
80. Colclough K, Bellanne-Chantelot C, Saint-Martin C, Flanagan SE, Ellard S. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha and 4 alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia. *Human mutation*. 2013;34:669-85.
81. Martín-Morales L, Rofes P, Diaz-Rubio E, Llovet P, Lorca V, Bando I, et al. Novel genetic mutations detected by multigene panel are associated with hereditary colorectal cancer predisposition. *PloS one*. 2018;13:e0203885.
82. Donner I, Katainen R, Sipila LJ, Aavikko M, Pukkala E, Aaltonen LA. Germline mutations in young non-smoking women with lung adenocarcinoma. *Lung cancer*. 2018;122:76-82.
83. Willson JS, Godwin TD, Wiggins GA, Guilford PJ, McCall JL. Primary hepatocellular neoplasms in a MODY3 family with a novel HNF1A germline mutation. *Journal of hepatology*. 2013;59:904-7.
84. Bacq Y, Jacquemin E, Balabaud C, Jeannot E, Scotto B, Branchereau S, et al. Familial liver adenomatosis associated with hepatocyte nuclear factor 1alpha inactivation. *Gastroenterology*. 2003;125:1470-5.

85. Gurzu S, Jung I, Orlowska J, Sugimura H, Kadar Z, Turdean S, et al. Hereditary diffuse gastric cancer--An overview. *Pathology, research and practice*. 2015;211:629-32.
86. Blair V, Martin I, Shaw D, Winship I, Kerr D, Arnold J, et al. Hereditary diffuse gastric cancer: diagnosis and management. *Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association*. 2006;4:262-75.
87. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, et al. E-cadherin germline mutations in familial gastric cancer. *Nature*. 1998;392:402-5.
88. Kaurah P, MacMillan A, Boyd N, Senz J, De Luca A, Chun N, et al. Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. *Jama*. 2007;297:2360-72.
89. Kim S, Chung JW, Jeong TD, Park YS, Lee JH, Ahn JY, et al. Searching for E-cadherin gene mutations in early onset diffuse gastric cancer and hereditary diffuse gastric cancer in Korean patients. *Familial cancer*. 2013;12:503-7.
90. Donner I, Kiviluoto T, Ristimaki A, Aaltonen LA, Vahteristo P. Exome sequencing reveals three novel candidate predisposition genes for diffuse gastric cancer. *Familial cancer*. 2015;14:241-6.
91. Yamada H, Shinmura K, Ito H, Kasami M, Sasaki N, Shima H, et al. Germline alterations in the CDH1 gene in familial gastric cancer in the Japanese population. *Cancer science*. 2011;102:1782-8.

Figures

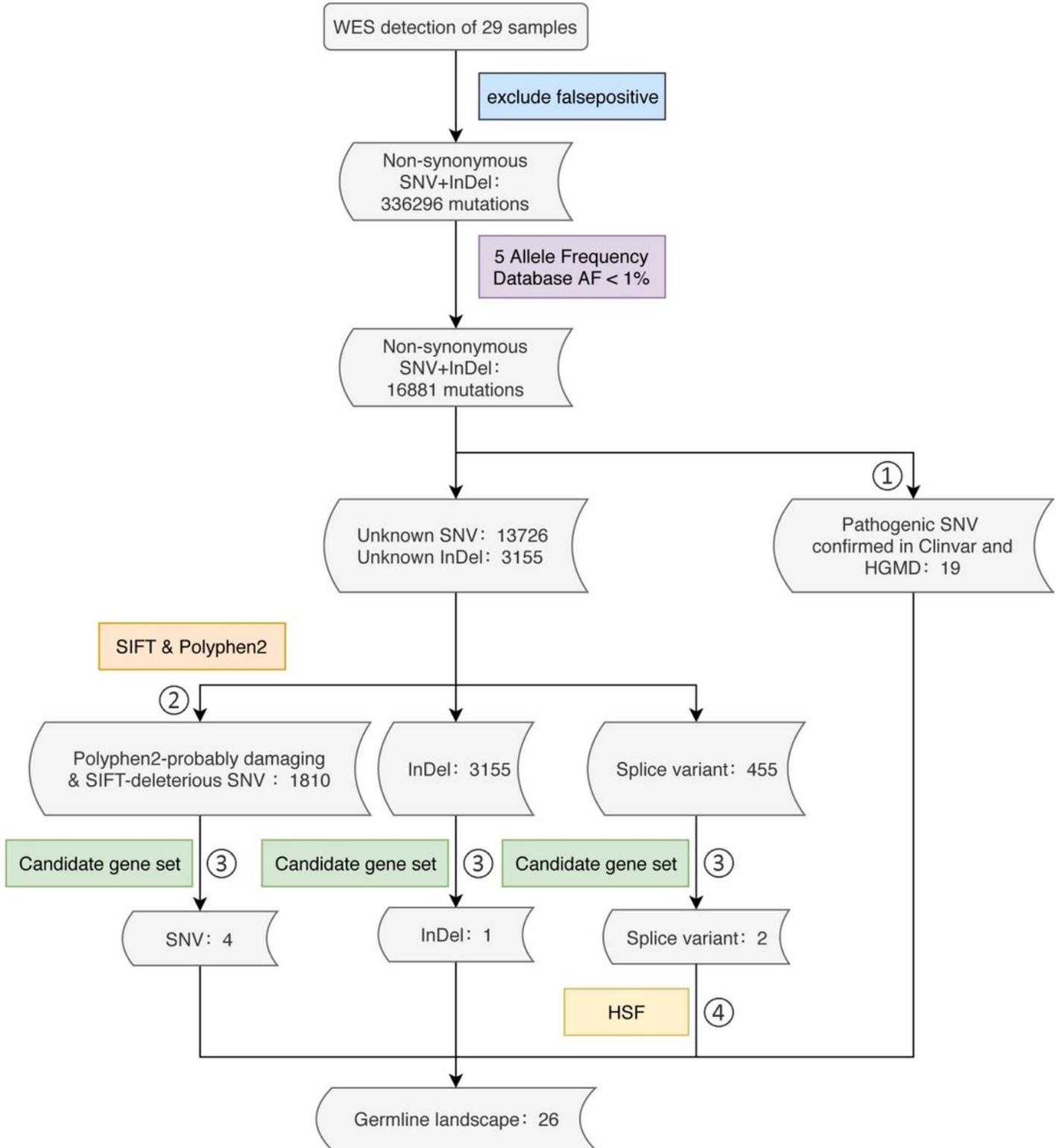


Figure 1

Process of identifying potential pathogenic germline variant candidates

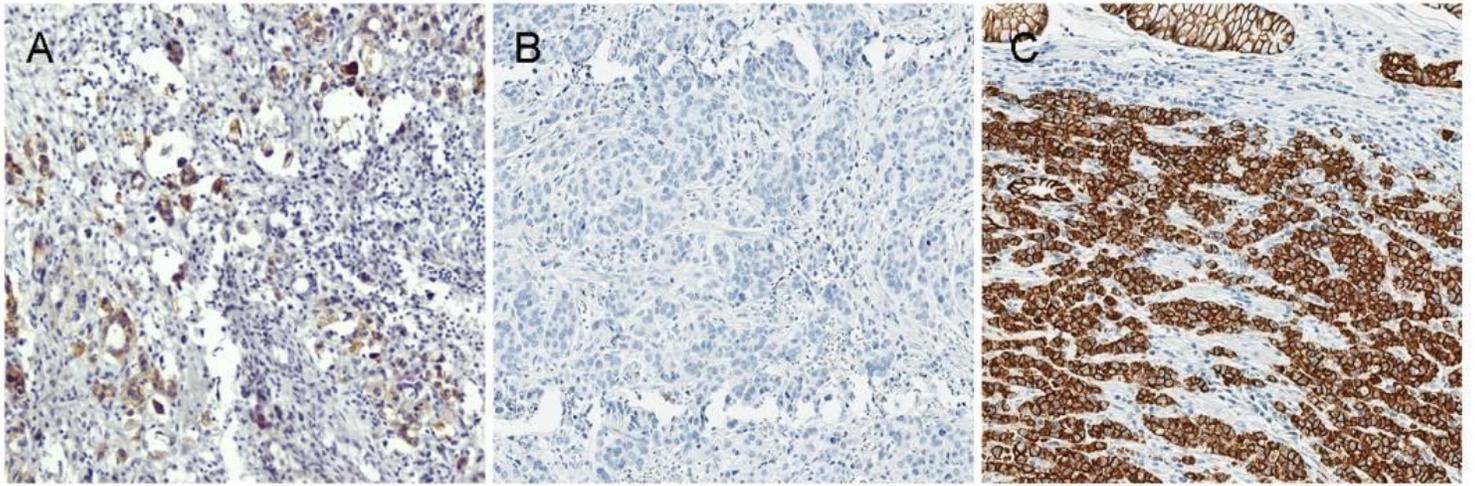


Figure 2

Expression of E-cadherin in HDGC patient (#RB809) carrying potential deleterious variant

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [STROBEchecklistcohort20200702.doc](#)