Contents available at **PubMed**
www.clinicalandtranslationalinvestigation.com

ORIGINAL ARTICLE

Low Expression of E-Cadherin and *CDH1* Variants Associated with Diffuse Gastric Cancer

Azaria García-Ruvalcaba[1,2], Katia C. Vázquez-Ibarra[1,2], María T. Magaña-Torres[1], Lourdes del C. Rizo de-la-Torre[1], Lennon Meléndez-Aranda[1], Gabriela López-Armas[3], José A. Cruz-Ramos[4], Jorge Peregrina-Sandoval[5], Esther Espinoza-Jiménez[6], María E. Rosales-Gradilla[7], and Josefina Y. Sánchez-López[1]*

[1]Western Biomedical Research Center, Instituto Mexicano del Seguro Social (IMSS), Guadalajara; [2]Doctorate Programme in Human Genetics, University Center for Health Sciences, University of Guadalajara, Guadalajara; [3]Industrial Technical Education Center, Guadalajara; [4]Immunobiology Laboratory, Department of Cellular and Molecular Biology, Jalisco Institute of Cancerology, Guadalajara; [5]University Center for Biological and Agricultural Sciences, University of Guadalajara, Guadalajara; [6]Department of Pathology, General Regional Hospital 110, IMSS, Guadalajara; [7]Department of Pathology, National Western Medical Center, IMSS, Guadalajara, Jal., Mexico

ABSTRACT

Background: Reduced or null expression of E-cadherin protein is a frequent cause of diffuse gastric cancer (DGC). More than 50% of patients with DGC present somatic variants in *CDH1* gene. **Objectives:** The objectives of this study were to study E-cadherin expression and identify variants in the *CDH1* gene in gastric tumors of patients with DGC. **Methods:** We studied 18 Mexican DGC patients who attended a hospital of the Mexican Social Security Institute; E-cadherin expression was determined by immunohistochemistry, and variants were identified by Sanger sequencing in promoter and coding regions. Predictive analysis was performed using PolyPhen-2 and HOPE software. **Results:** We found that 56% of DGC patients showed reduced expression of E-cadherin. All patients carried *CDH1* variants; overall, 12 different *CDH1* variants were identified. Predictive analysis revealed that the rs114265540 variant was probably damaging, with a value of 0.985, indicating a functional impact on the E-cadherin protein. Variants rs34939176 and rs33964119 were identified as risk factors for DGC (odds' ratios [OR] = 31.3, 95% CI 6.3-154.0, p < 0.001; OR = 6.1, 95% CI 2.0-19.0, p < 0.001, respectively) given their elevated frequency and by comparing it with those reported for MXL population in the 1000 Genomes Project database. **Conclusions:** In this Mexican population, the percentage of diffuse gastric tumors with reduced expression of E-cadherin was similar to that reported in other populations. All gastric tumors of DGC patients studied had somatic *CDH1* gene variants; however, the rs114265540, rs34939176, and rs33964119 variants were importantly related to DGC. (REV INVEST CLIN. 2023;75(1):37-44)

Keywords: Diffuse gastric cancer. Gastric tumors. Immunohistochemistry. CDH1 variants. E-cadherin.

INTRODUCTION

Gastric cancer (GC) is classified under three types: intestinal, diffuse, and mixed type, according to Lauren's criteria. Intestinal and diffuse GC exhibit numerous differences in pathology, epidemiology, and etiology. Diffuse GC (DGC) lacks cell adhesion and is characterized by the presence of signet-ring cells[1,2]. Some factors that promote DGC development are the loss of E-cadherin expression and pathogenic variants in the *CDH1* gene[3].

The E-cadherin protein is a transmembrane glycoprotein responsible for maintaining stable contacts between epithelial cell types. It is encoded by the *CDH1* gene, located on chromosome 16q22.1; the gene has 16 exons and a length of 98,250 bp, and the most common isoform is translated by mRNA variant 1 with 4.5 kb and 882 aa (NCBI Reference Sequence: NM_004360.3)[4,5]. E-cadherin expression has been studied by immunohistochemistry (IHC) in DGC patients of different ethnicities, and the percentage of cases with reduced expression of this protein observed worldwide is in the range of 41.7-100% (Table 1). In Mexico, only one study reported that 90% of DGC cases had reduced expression of E-cadherin[6].

On the other hand, over 700 variants in the *CDH1* gene have been described in DGC patients in different populations at the somatic and germline levels[7-10], providing valuable information on the spectrum of variants in this gene related to GC pathology. To date, 17 variants in the *CDH1* gene have been reported in Mexico, most of which are not pathogenic[6,11]. Although GC ranks fifth in terms of cancer mortality rates in Mexico[12], it has not been sufficiently studied. This study aimed to analyze the expression of E-cadherin and identify *CDH1* gene variants in diffuse gastric tumors of Mexican patients.

METHODS

We analyzed samples of paraffin-embedded gastric tumors that were obtained by biopsy through endoscopy belonging to 18 patients with a histopathological diagnosis of DGC with the presence of signet ring cell and/or poorly differentiated. All patients with early, locally advanced, or metastatic DGC were included

Table 1. Cases with low expression of E-cadherin in patients with diffuse gastric cancer (DGC) from different populations

Population	DGC cases (n)	Cases with low expression (%)	Reference
Mexican	18	55.6	This study
Mexican	69	90.0	6
Chinese	75	65.3	25
Egyptian	11	41.7	26
Egyptian	22	64.0	27
English	24	54.0	28
German	115	54.8	29
Greek	21	55.0	30
Indian	30	50.0	31
Iranian	4	100	32
Italian	29	62.0	33
Japanese	47	63.8	34
Japanese	75	55.0	35
Japanese	76	67.1	36
Korean	25	80.0	37
Peruvian	90	60.0	38
Polish	30	53.3	39
Romanian	61	82.4	40
Ukrainian	33	65.8	41
American	116	58.0	42

in the study. Samples with intestinal type or mixed type of GC were not included in the study. Samples that did not have sufficient quality and quantity of cells (< 60% tumoral cells) for analysis were excluded from the study. The inclusion process of the cases was consecutive and samples were collected between 2012 and 2017 in the pathology service of Hospital 110 of the Mexican Social Security Institute (IMSS) in the city of Guadalajara, Jalisco, Mexico. This research was approved by the Local Ethics and Research Committees (Comité de Ética en Investigación 13058 and Comité Local de Investigación en Salud 1305) of Western Biomedical Research Center, IMSS.

The paraffin-embedded tissue samples were cut on a microtome into 5 μm-thick sections. Immediately, they were dewaxed in the stove at 60°C for 1 h, followed by a xylol and ethanol series and finally water.

The slides described above underwent heat-mediated antigen retrieval using the pressure cooker method with citrate buffer at pH 6.0 for 25 min. Tissues were washed in 0.1 M phosphate-buffered saline (PBS) 3 times for 5 min each time. Slides were incubated in PolyDetector AP Blocker (cat. no. BSB 0055; Bio SB, Inc., Santa Barbara, CA, United States of America [USA]) for 5 min and washed in 0.1 M PBS buffer 3 times for 5 min each time. Antibody incubation was carried out using Novocastra™ Liquid Mouse Monoclonal Antibody anti-E-cadherin (Leica Biosystems Nussloch GmbH, Germany) overnight at 4°C. Tissues were then washed in 0.1 M PBS buffer 3 times for 5 min each time. The immunohistochemical process was continued using a Dako LSAB System-HRP system (K0675; Agilent Technologies, USA); the biotinylated secondary antibody and streptavidin-conjugated horseradish peroxidase were used according to the manufacturer's protocol. Immunodetection was visualized using diaminobenzidine tetrahydrochloride (DAB; Sigma-Aldrich; Merck Millipore, Darmstadt, Germany). Slides were counterstained with hematoxylin for 2 min (cat no. BSB 0024; Bio-SB, Inc.) and mounted with Entellan® (cat. no. 107960; Merck-Millipore). The staining results were scored on a semiquantitative scale established by a pathologist based on the percentage of cells with loss of membranous staining and were classified into two categories: normal (≥ 91% of stained cells) and reduced (≤ 90% stained cells).

Tumoral DNA was isolated using an Invisorb® Spin Tissue Mini Kit (Invitek Molecular GmbH, Germany). The promoter region and 16 exons of the *CDH1* gene were analyzed by PCR using previously described primers[11,13]. The resulting fragments were cleaned with ExoSAP-IT™ PCR Product Cleanup Reagent (Thermo Fisher, USA). For Sanger sequencing, a Big Dye Terminator kit v.3.1 and ABI 310 Genetic Analyzer (Applied Biosystems, Foster City, CA, USA) were used for capillary electrophoresis; the Chromas Lite v.2.6.6 program was used for data analysis. The proportion of cases with reduced E-cadherin expression and frequencies of *CDH1* variants was obtained by direct counting and Chi-square test; a haplotype analysis was performed by Arlequin v3.5 software. Predictive analysis was performed in the variants with unknown pathogenicity or uncertain reports. Four programs were used, NNSplice and Net-Gene2 for intronic variants[14,15], and HOPE and Poly-Phen-2 for missense variants[16,17].

RESULTS

Overall, 18 DGC patients were analyzed, ten women (56%) and eight men (44%). The age range was 28-76 years, with an average of 59 ± 14.3 years. Reduced E-cadherin expression was observed in 56% of the DGC patients, whose values of stained cells ranged between 2% and 90% ($\bar{X}$ = 54.5%). The remaining 44% showed normal expression, with a percentage of stained cells between 95% and 100% ($\bar{X}$ = 98.3%).

Through *CDH1* gene sequencing, 12 different SNPs were identified: c.−285C > A rs16260, c. −273G > A rs1330727101, c.48 + 6C > T rs3743674, c.388−83G > A rs140449923, c.388−48A > T,G rs766910353, c.531 + 4G > A rs1172442868, c.1321−10T > C,G rs1597897718, c.2076T > C rs1801552, c.2164 + 17dupA rs34939176, c.2253C > T rs33964119, c.2331C > G rs114265540, and c.*54C > T rs1801026 (note that new alleles were observed in our study) (Table 2). Each patient presented a minimum of three variants and a maximum of seven, with c.2076T > C being the most common variant observed in a homozygous state in all patients.

A comparative analysis was performed between the genotype frequency distribution of each of the variants and E-cadherin protein expression. Results showed that the genotypes were similarly distributed among patients with normal and reduced expression (p > 0.05); no variant was related to the reduced expression of E-cadherin.

In addition, haplotypes were established with the 12 identified variants (ordered from 5' to 3') to relate them to the expression of E-cadherin. The results showed 21 different combinations, and the most frequent was CGTGAGTCACCC (36.1%, n = 13 chromosomes). This haplotype comprises wild-type alleles at all sites, except the variants c.48 + 6C > T and c.2076T > C (third and eighth positions). The other haplotypes were observed with very low frequencies (<6%). The distribution of haplotype frequencies between patients with normal versus reduced E-cadherin expression did not show significant differences (p = 1.0).

Genotypic and allelic frequencies were compared with the data reported in a population with Mexican

Table 2. Allelic and genotypic frequencies of the identified variants in *CDH1* gene in tumors of patients with diffuse gastric cancer

Location	Variant	This study		1KGP MXL		Clinvar	New alleles
		Genotypic frequency	Allelic frequency	Genotypic frequency	Allelic frequency		
Promoter	c.−285C>A rs16260	CC:0.72 CA:0.17 AA:0.11	C:0.81 A:0.19	CC:0.60 CA:0.31 AA: 0.09 p = 0.48	C:0.75 A:0.25 p = 0.49	B HDGC	NR
	c.−273G>A rs1330727101	GG:0.89 GA: 0.11 AA: 0.00	G:0.94 A: 0.06	NR	NR	NR[a]	NR
Intron 1	c.48+6C>T rs3743674	CC: 0.06 CT: 0.50 TT: 0.44	C: 0.31 T: 0.69	CC: 0.03 CT: 0.53 TT: 0.44 p = 0.96	C: 0.30 T: 0.70 p = 0.92	B HDGC	NR
Intron 3	c.388−83G>A rs140449923	GG: 0.56 GA: 0.44 AA:0.00	G: 0.78 A: 0.22	GG: 1.00 GA 0.00 AA:0.00 p < 0.001	G: 1.00 A 0.00 p < 0.001	NR	NR
	c.388−48A>G rs766910353	AA:0.94 AG:0.06 GG:0.00	A:0.97 G:0.03	NR	NR	NR	G
Intron 4	c.531+4G>A rs1172442868	GG:0.94 GA:0.06 AA:0.00	G:0.97 A:0.03	NR	NR	U	NR
Intron 9	c.1321−10T>G rs1597897718	TT:0.83 TG:0.17 GG:0.00	T:0.92 G:0.08	NR	NR	PB	G
Exon 13	c.2076T>C, p.Ala692 = rs1801552	TT:0.00 TC:0.00 CC:1.00	T: 0.00 C: 1.00	TT:0.11 TC:0.42 CC:0.47 p < 0.001	T:0.32 C:0.68 p < 0.001	B	NR
Intron 13	c.2164+17dupA rs34939176	AA:0.11 AAA:0.89 AAAA:0.00	A:0.56 AA:0.44	AA:0.80 AAA:0.2 AAAA:0.00 p < 0.001	A:0.90 AA:0.10 p < 0.001	B	NR
Exon 14	c.2253C>T, p.Asn751 = rs33964119	CC:0.39 CT:0.50 TT:0.11	C:0.64 T:0.36	CC:0.80 CT:0.20 TT:0.00 p = 0.008	C:0.90 T:0.10 p < 0.001	B	NR
Exon 15	c.2331C>G, T, p.Asp777Glu rs114265540	CC:0.83 CG:0.17 GG:0.00	C:0.92 G:0.08	CC:1.00 CG:0.00 GG:0.00 p = 0.03	C:1.00 G:0.00 p = 0.009	U	NR
UTR 3'	c.*54C>T rs1801026	CC:0.56 CT:0.44 TT:0.00	C:0.78 T:0.22	CC:0.66 CT:0.31 TT:0.03 p = 0.78	C:0.81 T:0.19 p = 0.64	PB/B	NR

1KGP: 1000 Genomes Project; B: benign; NR: not reported; PB: probably benign; U: uncertain; HDGC: hereditary diffuse gastric cancer. [a]Variant previously analyzed (58); the allele "A" could modify the binding sites of the transcription factors IRF3 and SP2.

ancestry residing in Los Angeles, USA (MXL, n = 71 individuals) in the 1000 Genomes Project (1KGP)[18] to determine whether any of the identified variants could be associated with the risk of developing DGC. Five variants, c.388−83G > A, c.2076T > C, c.2164 + 17dupA, c.2253C > T, and c.2331C > G, showed statistically significant differences (p < 0.05) (Table 2). The odds' ratios (OR) were calculated using the Cochran-Armitage trend test[19] under classical Mendelian inheritance patterns (dominant, recessive and codominant). The association or risk was evaluated with the possible genotypes (wildtype, heterozygous, or homozygous) for each variant. The results were significant for c.2164 + 17dupA (OR = 31.4, 95% CI 6.39-154.08, p = 0.00001) under the dominant model (dupA/dupA or dupA/A vs. A/A), and for c.2253C>T (OR = 6.16, 95% CI 2.00-19.01, p = 0.0008 and OR = 3.92; 95% CI 1.30-11.86; p = 0.0120) under dominant (T/T or T/C vs. C/C) and codominant (T/C vs. T/T or C/C) models, respectively.

Predictive analysis results showed that the intronic variants c.388−83G > A, C and c.388−48A > G did not affect splicing (delta value= 0), whereas for the c.531 + 4G > A variant, there was an increase in the donor site and a delta value = −0.58, which might interfere with the recognition of the splice donor site of exon 4. On the other hand, PolyPhen-2 revealed that the missense variant Asp777Glu (c.2331C > G) was probably damaging, with a value of 0.985, indicating a functional impact on the protein E-cadherin; this variant was observed in two patients with reduced E-cadherin expression and one with normal expression. The HOPE program, which analyzes the structural and functional effect of a variant on the protein, showed that the secondary structure could be destabilized and be damaging.

DISCUSSION

The most well-known alteration associated with DGC development is loss of a functional *CDH1* gene, whose product is the protein E-cadherin. The *CDH1* gene is commonly inactivated by somatic mutations, methylation of promoters or enhancers, or chromosomal rearrangements; consequently, E-cadherin protein shows an abnormal distribution or is reduced or absent[20-24]. To date, we found in the literature 19 reports worldwide that have identified reduced E-cadherin expression in DGC patients using IHC (Table 1)[6,25-42]. The percentage of cases with reduced expression of this protein is in the range of 41.7-100%. In this study, we found that 56% of DGC patients had reduced E-cadherin protein; we compared our results only with those reported for other populations, where they used a scale similar to ours. Our results are similar to studies performed in Indian (50.0%, p = 0.729)[31], Japanese (63.8%, p = 0.539)[34], Korean (80.0%, p = 0.085)[37], and Romanian populations (82.4%, p = 0.087)[40]; however, they are statistically different from results in Iranian (100%, p = 0.018)[32] and Mexican (90%) populations[6]. These differences are probably due to the distinct types of scales used to evaluate E-cadherin expression; for instance, the categories normal, abnormal (atypical or heterogeneous), and negative were used in the previous Mexican study[6], whereas in this work, the categories normal and reduced were used. The methodology used in the studies may be the primary source of heterogeneity in the results since the type of antibody and dilution is different in each study; therefore, the sensitivity of IHC may depend on this factor. In addition, the cutoff point that determines whether the expression of E-cadherin is preserved depends on the pathologist's criterion and usually varies between 5% and 90%[43]. Visual assessment of immunohistochemistry is subjective and has limited dynamic ranges in measurements since the nature of the visually obtained variables assessment is categorical, and multiple studies have demonstrated low inter- and intraobserver agreement[44].

Digital pathology can help to eliminate bias when counting to assess immunohistochemistry markers and has been shown to increase significantly interobserver (between pathologists) and intraobserver (within the same pathologist) agreement[45]. Some algorithms for evaluating membrane-type stains such as E-cadherin and HER2 are based on color deconvolution and segmentation to distinguish specifically membrane decoration by the antibody. Algorithms with different approaches to immunohistochemistry quantification such as cell-based H-score, pixel-based H-score, and average threshold method have also been developed[46]. However, coordination and governance through international guidelines is also required for quantification with digital pathology methods. This involves the implementation of common worldwide algorithms for the quantification of immunohistochemistry markers,

taking into account the reagents and methods used for staining and the use of robust algorithms for quantification.

Regarding *CDH1* variants, it has been reported that up to 38% of sporadic DGC tumors have mutations in the *CDH1* gene[47]. In our study, all patients had *CDH1* variants, and a minimum of 3 and a maximum of 7 variants were identified. About 12 SNPs were identified, all of them previously reported in the SNP database, and none was related to reduced E-cadherin expression. This suggests the presence of other mechanisms to silence the gene in these patients, probably methylation, which has also been reported in the *CDH1* gene as a cause of reduced E-cadherin expression[11,48,49]. Chi-square test[50] was used to compare the frequencies of *CDH1* variants c.388−83G > A, c.2076T > C, c.2164 + 17dupA, c.2253C > T, and c.2331C > G with those found in the MXL population (Mexican ancestry in Los Angeles, California) reported in the 1000 Genomes database[18]. Results showed an association with DGC (p < 0.05). Moreover, the variants c.2164 + 17dupA and c.2253C > T were linked with increased risks for DGC; therefore, it is important to continue studying these variants in our population, as well as in other populations. In the literature, some of these variants are reported in association with DGC. For instance, the variant c.2253C > T in Chinese population may contribute to a predisposition to GC in groups with a family history (OR = 3.19, 95% CI = 1.29-7.91)[51], but its association is controversial, because, in Taiwanese and Iranian populations, no such association was observed[52,53]. For other variants, an association has not been shown, such as c.2076T > C p. Ala692 = rs1801552[52-54] and c.2164 + 17dupA[55,56].

Seven of the 12 identified variants are reported in ClinVar[8] as benign or probably benign (Table 2). The rs16260 variant was previously associated to risk of DGC in our population, with an OR of 1.98 (95% CI 1.01-3.98) for heterozygotes and 6.5 (95% CI 2.1-19.6) for homozygotes[57], and the others are reported as having uncertain significance (c.531 + 4G > A and c.2331C > G) or have not been reported (c. −273G > A, c.388−83G > A,C, and c.388−48A > G), all of which were included in the predictive analysis. We reported in a previous study that allele A of variant c. −273G > A could modify the binding sites of the transcription factors IRF3 and SP2[58]. Prediction for

the intronic variants c.388−83G > A,C and c.388−48A > G with the NNSplice and NetGene2 programs showed that splicing would not be affected. Analysis of variant c.531 + 4G > A located in intron 4 indicated a possible loss of splice site, although *in vitro* and *in vivo* studies are needed. However, analysis in cell lines of a variant located at the adjacent site (c.531 + 3A > G) showed that splicing is not altered in the presence of the allele G[59]. Considering these results, it is probable that variant c.531 + 4G > A does not modify the splice site. On the other hand, the missense variant c.2331C > G (p.Asp777Glu) was predicted by our analysis to be damaging, as the secondary structure resulting from the change of aspartic acid to glutamic acid at residue 777 could destabilize the structure of the E-cadherin protein; moreover, the change of the medium (Asp) to large (Glu) side chain could probably interfere with the interaction with p120 catenin (amino acids 736−781), whose function is to stabilize and prevent the entry of E-cadherin into degrading endocytic pathways. Then, in the absence of p120 catenin, E-cadherin interacts with Hakai protein (binding sites at amino acids 734−798), promoting internalization, and targeting the plasma membrane or degrading E-cadherin[60]. Finally, all gastric tumors of DGC patients studied here had somatic *CDH1* gene variants; however, the c.2164 + 17dupA, c.2253C > T, and c.2331C > G variants were importantly related to DGC.

ACKNOWLEDGMENTS

REFERENCES

1. Ansari S, Gantuya B, Tuan VP, Yamaoka Y. Diffuse gastric cancer: a summary of analogous contributing factors for its molecular pathogenicity. Int J Mol Sci. 2018;19:2424.
2. Sitarz R, Skierucha M, Mielko J, Offerhaus GJ, Maciejewski R, Polkowski WP. Gastric cancer: epidemiology, prevention, classification, and treatment. Cancer Manag Res. 2018;10: 239-48.
3. Korivi BR, Faria S, Aly A, Sun J, Patnana M, Jensen CT, et al. Intestinal and diffuse gastric cancer: a retrospective study comparing primary sites. Clin Imaging. 2019;56:33-40.
4. Reinhold WC, Reimers MA, Lorenzi P, Ho J, Shankavaram UT, Ziegler MS, et al. Multifactorial regulation of E-cadherin expression: an integrative study. Mol Cancer Ther. 2010;9:1-16.
5. Graziano F. The E-cadherin gene, structure and function. In: Corso G, Roviello F, editors. Spotlight on Familial and Hereditary Gastric Cancer. 1st ed. Dordrecht: Springer; 2013. p. 27-33.

6. Gamboa-Domínguez A, Domínguez-Fonseca C, Chavarri-Guerra Y, Vargas R, Reyes-Gutiérrez E, Green D, et al. E-cadherin expression in sporadic gastric cancer from Mexico: exon 8 and 9 deletions are infrequent events associated with poor survival. Hum Pathol. 2005;36:29-35.

7. Fokkema IF, Taschner PE, Schaafsma GC, Celli J, Laros JF, den Dunnen JT. LOVD v.2.0: the next generation in gene variant databases. Hum Mutat. 2011;32:557-63.

8. Landrum MJ, Lee JM, Benson M, Brown GR, Chao C, Chitipiralla S, et al. ClinVar: improving access to variant interpretations and supporting evidence. Nucleic Acids Res. 2018;46:D1062-7.

9. Tate JG, Bamford S, Jubb HC, Sondka Z, Beare D, Bindal N, et al. COSMIC: the catalogue of somatic mutations in cancer. Nucleic Acids Res. 2019;47:D941-7.

10. Kopanos C, Tsiolkas V, Kouris A, Chapple CE, Aguilera MA, Meyer R, et al. VarSome: the human genomic variant search engine. Bioinformatics. 2019;35:1978-80.

11. Bustos-Carpinteyro AR, Oliveira C, Sousa A, Oliveira P, Pinheiro H, Carvalho J, et al. *CDH1* somatic alterations in Mexican patients with diffuse and mixed sporadic gastric cancer. BMC Cancer. 2019;19:69.

12. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71:209-49.

13. Corso G, Carvalho J, Marrelli D, Vindigni C, Carvalho B, Seruca R, et al. Somatic mutations and deletions of the E-cadherin gene predict poor survival of patients with gastric cancer. J Clin Oncol. 2013;31:868-75.

14. Brunak S, Engelbrecht J, Knudsen S. Prediction of human mRNA donor and acceptor sites from the DNA sequence. J Mol Biol. 1991;220:49-65.

15. Reese MG, Eeckman FH, Kulp D, Haussler D. Improved splice site detection in Genie. J Comput Biol. 1997;4:311-23.

16. Adzhubei I, Jordan DM, Sunyaev SR. Predicting functional effect of human missense mutations using PolyPhen-2. In: Current Protocols in Human Genetics. Ch. 7. Hoboken: Wiley; 2103.

17. Venselaar H, Te Beek TA, Kuipers RK, Hekkelman ML, Vriend G. Protein structure analysis of mutations causing inheritable diseases. an e-science approach with life scientist friendly interfaces. BMC Bioinformatics. 2010;11:548.

18. 1000 Genomes Project Consortium, Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, et al. A global reference for human genetic variation. Nature. 2015;526:68-74.

19. Ziegler A, König IR. A Statistical Approach to Genetic Epidemiology. Concepts and Applications. 1st ed. Weinheim, Germany: Wiley-VCH; 2006. p. 199-214.

20. Humar B, Blair V, Charlton A, More H, Martin I, Guilford P, et al. E-cadherin deficiency initiates gastric signet-ring cell carcinoma in mice and man. Cancer Res. 2009;69:2050-6.

21. Grady WM, Willis J, Guilford PJ, Dunbier AK, Toro TT, Lynch H, et al. Methylation of the *CDH1* promoter as the second genetic hit in hereditary diffuse gastric cancer. Nat Genet. 2000;26: 16-7.

22. Oliveira C, De Bruin J, Nabais S, Ligtenberg M, Moutinho C, Nagengast FM, et al. Intragenic deletion of *CDH1* as the inactivating mechanism of the wild-type allele in an HDGC tumour. Oncogene. 2004;23:2236-40.

23. Oliveira C, Sousa S, Pinheiro H, Karam R, Bordeira-Carriço R, Senz J, et al. Quantification of epigenetic and genetic 2nd hits in *CDH1* during hereditary diffuse gastric cancer syndrome progression. Gastroenterology. 2009;136:2137-48.

24. Tedaldi G, Molinari C, José CS, Barbosa-Matos R, André A, Danesi R, et al. Genetic and epigenetic alterations of *CDH1* regulatory regions in hereditary and sporadic gastric cancer. Pharmaceuticals (Basel). 2021;14:457.

25. Zhao C, Feng Z, He H, Zang D, Du H, Huang H, et al. Protein expression-based classification of gastric cancer by immunohistochemistry of tissue microarray. PLoS One. 2020;15:e0238836.

26. Saad AA, Awed NM, Abd Elkerim NN, El-Shennawy D, Alfons MA, Elserafy ME, et al. Prognostic significance of E-cadherin expression and peripheral blood micrometastasis in gastric carcinoma patients. Ann Surg Oncol. 2010;17:3059-67.

27. Hamed MK, Hanna HH, Fawzy G. Role of E-cadherin protein expression and E-cadherin gene promoter hypermethylation in gastric carcinoma. Egypt J Surg. 2018;38:361-8.

28. Jawhari A, Jordan S, Poole S, Browne P, Pignatelli M, Farthing MJ. Abnormal immunoreactivity of the E-cadherin-catenin complex in gastric carcinoma: relationship with patient survival. Gastroenterology. 1997;112:46-54.

29. Gabbert HE, Mueller W, Schneiders A, Meier S, Moll R, Birchmeier W, et al. Prognostic value of E-cadherin expression in 413 gastric carcinomas. Int J Cancer. 1996;69:184-9.

30. Skarlos DV, Bai M, Goussia A, Samantas E, Galani E, Tsavdaridis D, et al. Expression of a molecular marker panel as a prognostic tool in gastric cancer patients treated postoperatively with docetaxel and irinotecan. a study of the Hellenic cooperative oncology group. Anticancer Res. 2007;27:2973-83.

31. Sisodiya N, Jagani R. Clinicopathological significance of E-cadherin immunoexpression in gastric carcinoma. Int J Contemp Med Res. 2016;3:2431-3.

32. Torabizadeh Z, Nosrati A, Saravi SN, Charati JY, Janbabai G. Evaluation of E-cadherin expression in gastric cancer and its correlation with clinicopathologic parameters. Int J Hematol Oncol Stem Cell Res. 2017;11:158-64.

33. Graziano F, Mandolesi A, Ruzzo A, Bearzi I, Testa E, Arduini F, et al. Predictive and prognostic role of E-cadherin protein expression in patients with advanced gastric carcinomas treated with palliative chemotherapy. Tumour Biol. 2004;25:106-10.

34. Kawanishi K, Doki Y, Shiozaki H, Yano M, Inoue M, Fukuchi N, et al. Correlation between loss of E-cadherin expression and overexpression of autocrine motility factor receptor in association with progression of human gastric cancers. Am J Clin Pathol. 2000;113:266-74.

35. Yashiro M, Nishioka N, Hirakawa K. Decreased expression of the adhesion molecule desmoglein-2 is associated with diffuse-type gastric carcinoma. Eur J Cancer. 2006;42:2397-403.

36. Uchikado Y, Okumura H, Ishigami S, Setoyama T, Matsumoto M, Owaki T, et al. Increased Slug and decreased E-cadherin expression is related to poor prognosis in patients with gastric cancer. Gastric Cancer. 2011;14:41-9.

37. Joo YE, Park CS, Kim HS, Choi SK, Rew JS, Kim SJ. Prognostic significance of E-cadherin/catenin complex expression in gastric cancer. J Korean Med Sci. 2000;15:655-66.

38. Arévalo F, Monge E, Morales D. E-cadherin molecular expression in the diffuse and intestinal types of gastric adenocarcinoma. a report from Lima, Peru. Rev Esp Enferm Dig. 2007;99:325-9.

39. Czyzewska J, Guzińska-Ustymowicz K, Ustymowicz M, Pryczynicz A, Kemona A. The expression of E-cadherin-catenin complex in patients with advanced gastric cancer: role in formation of metastasis. Folia Histochem Cytobiol. 2010;48:37-45.

40. Lazăr D, Tăban S, Ardeleanu C, Dema A, Sporea I, Cornianu M, et al. The immunohistochemical expression of E-cadherin in gastric cancer; correlations with clinicopathological factors and patients' survival. Rom J Morphol Embryol. 2008;49:459-67.

41. Bazas VM, Lukyanova NY, Demash DV, Galakhin KO, Myasoedov DV. Relation between cell-to-cell adhesion and angiogenesis and clinico-morphological prognostic factors in patients with gastric cancer. Exp Oncol. 2008;30:235-9.

42. Xia M, Xie Y, Zan L, Reddy S, Tan C, Li J, et al. Membranous staining of β-catenin and E-cadherin expression in patients with gastric cancer. Int J Clin Exp Pathol. 2017;10:8980-90.

43. Xing X, Tang YB, Yuan G, Wang Y, Wang J, Yang Y, et al. The prognostic value of E-cadherin in gastric cancer: a meta-analysis. Int J Cancer. 2013;132:2589-96.

44. Gomes DS, Porto SS, Balabram D, Gobbi H. Inter-observer variability between general pathologists and a specialist in breast pathology in the diagnosis of lobular neoplasia, columnar cell lesions, atypical ductal hyperplasia and ductal carcinoma in situ of the breast. Diagn Pathol. 2014;9:121.

45. Brügmann A, Eld M, Lelkaitis G, Nielsen S, Grunkin M, Hansen JD, et al. Digital image analysis of membrane connectivity is a robust measure of HER2 immunostains. Breast Cancer Res Treat. 2012;132:41-9.

46. Ram S, Vizcarra P, Whalen P, Deng S, Painter CL, Jackson-Fisher A, et al. Pixelwise H-score: a novel digital image analysis-based metric to quantify membrane biomarker expression from immunohistochemistry images. PLoS One. 2021;16:e0245638.

47. García-Peláez J, Barbosa-Matos R, Gullo I, Carneiro F, Oliveira C. Histological and mutational profile of diffuse gastric cancer: current knowledge and future challenges. Mol Oncol. 2021;15:2841-67.

48. Yamamoto E, Suzuki H, Takamaru H, Yamamoto H, Toyota M, Shinomura Y. Role of DNA methylation in the development of diffuse-type gastric cancer. Digestion. 2011;83:241-9.

49. Usui G, Matsusaka K, Mano Y, Urabe M, Funata S, Fukayama M, et al. DNA methylation and genetic aberrations in gastric cancer. Digestion. 2021;102:25-32.

50. Preacher KJ. Calculation for the Chi-Square Test: An Interactive Calculation Tool for Chi-Square Tests of Goodness of Fit and Independence [Software]. 2001. Available from: https://www.quantpsy.org [Last accessed on 2022 Nov 29].

51. Zhang Y, Liu X, Fan Y, Ding J, Xu A, Zhou X, et al. Germline mutations and polymorphic variants in MMR, E-cadherin and MYH genes associated with familial gastric cancer in Jiangsu of China. Int J Cancer. 2006;119:2592-6.

52. Chu CM, Chen CJ, Chan DC, Wu HS, Liu YC, Shen CY, et al. *CDH1* polymorphisms and haplotypes in sporadic diffuse and intestinal gastric cancer: a case-control study based on direct sequencing analysis. World J Surg Oncol. 2014;12:80.

53. Moridnia A, Tabatabaiefar MA, Zeinalian M, Minakari M, Kheirol-lahi M, Moghaddam NA. Novel variants and copy number variation in *CDH1* gene in Iranian patients with sporadic diffuse gastric cancer. J Gastrointest Cancer. 2019;50:420-7.

54. Humar B, Graziano F, Cascinu S, Catalano V, Ruzzo AM, Magnani M, et al. Association of *CDH1* haplotypes with susceptibility to sporadic diffuse gastric cancer. Oncogene. 2002;21:8192-5.

55. Caggiari L, Miolo G, Buonadonna A, Basile D, Santeufemia DA, Cossu A, et al. Characterizing metastatic HER2-positive gastric cancer at the *CDH1* haplotype. Int J Mol Sci. 2017;19:47.

56. Guindalini RS, Cormedi MC, Maistro S, Pasini FS, Branas PC, Dos Santos L, et al. Frequency of *CDH1* germline variants and con-tribution of dietary habits in early age onset gastric cancer patients in Brazil. Gastric Cancer. 2019;22:920-31.

57. Medina-Franco H, Ramos-De la Medina A, Vizcaino G, Medina-Franco JL. Single nucleotide polymorphisms in the promoter region of the E-cadherin gene in gastric cancer: case-control study in a young Mexican population. Ann Surg Oncol. 2007;14:2246-9.

58. García-Ruvalcaba A, Del C Rizo de la Torre L, Magaña-Torres MT, Prado-Montes-de-Oca E, Ruiz-Ramírez AV, Rangel-Villalobos H, et al. Predicting pathogenicity of *CDH1* gene variants in patients with early-onset diffuse gastric cancer from Western Mexico. Rev Invest Clin. 2021;73:172-81.

59. Vogelaar IP, Figueiredo J, van Rooij IA, Simões-Correia J, van der Post RS, Melo S, et al. Identification of germline mutations in the cancer predisposing gene *CDH1* in patients with orofacial clefts. Hum Mol Genet. 2013;22:919-26.

60. Figueiredo J, Söderberg O, Simões-Correia J, Grannas K, Suriano G, Seruca R. The importance of E-cadherin binding partners to evaluate the pathogenicity of E-cadherin missense mutations associated to HDGC. Eur J Hum Genet. 2013;21:301-9.