Title
Resolving pathogenicity classification for the CDH1 c.[715G>A] (p.Gly239Arg) Variant

Journal
European Journal of Human Genetics, 29(7)

ISSN
1018-4813

Authors
Yelskaya, Zarina
Arnold, Angela G
Marcell, Vanessa J
et al.

Publication Date
2021-07-01

DOI
10.1038/s41431-021-00825-w

Peer reviewed

European Journal of Human Genetics (2021) 29:1103–1109
https://doi.org/10.1038/s41431-021-00825-w

ARTICLE

Resolving pathogenicity classification for the *CDH1* c.[715G>A] (p.Gly239Arg) Variant

Zarina Yelskaya[1] · Angela G. Arnold[2] · Vanessa J. Marcell[2] · Laura H. Tang[1] · Erin E. Salo-Mullen[2] · Vivian E. Strong[3] · Zsofia K. Stadler[2] · Liying Zhang[1,4]

Received: 28 July 2020 / Revised: 25 January 2021 / Accepted: 29 January 2021 / Published online: 22 February 2021

Abstract

Hereditary Diffuse Gastric Cancer (HDGC) syndrome is associated with *CDH1* germline likely pathogenic/pathogenic variants. Carriers of *CDH1* germline likely pathogenic/pathogenic variants are predisposed to diffuse gastric cancer and lobular breast cancer. This study aims to classify the *CDH1* c.[715G>A] missense variant identified in a diffuse gastric cancer prone family by performing splicing studies. RT-PCR and subsequent cloning experiments were performed to investigate whether this variant completely disrupts normal splicing. This variant preferentially abolishes normal splicing through activation of a cryptic 3′ acceptor splice site within exon 6 of *CDH1*, presumably leading to a premature protein truncation within first extracellular domain repeat of E-cadherin protein. Our results contributed to evidence necessary to resolve pathogenicity classification of this variant, indicating that this variant is to be classified as pathogenic.

Introduction

Gastric cancer (GC) is the fifth most common cancer worldwide and the third most common cause of cancer deaths with nearly 841,000 deaths per year. GC has been histopathologically described into two subtypes: intestinal type and diffuse type [1, 2]. Intestinal type gastric cancer accounts for 60% of GC cases, is characterized as a well to moderately differentiated adenocarcinoma with cohesive tumor cell groups having a well-defined glandular architecture. Diffuse type gastric cancer (DGC) is characterized by signet ring cell morphology without intercellular adhesions or gland formation and diffuse infiltrating growth pattern, which accounts for 30% of GC. Sporadic GC accounts for the majority of GC cases, while GC with familial presentation accounts for 10%, of which 1–3% are hereditary [2]. Hereditary diffuse gastric cancer (HDGC) (OMIM #137215) is one of the main syndromes in hereditary GC cases with autosomal dominant susceptibility, and is characterized by early-onset DGC and lobular breast cancer [3]. HDGC is associated with germline likely pathogenic/pathogenic variants within *CDH1* gene (NM_004360.5), which encodes E(epithelial)-cadherin, a transmembrane glycoprotein functioning in homophilic calcium-dependent cell-to-cell adhesion [4, 5].

The cumulative risk of GC in *CDH1* likely pathogenic/pathogenic variants carriers is 42% in males and 33% in females by 80 years of age, and the risk of breast cancer was 55% in females [6]. Currently, 155 *CDH1* germline variants have been reported across multiple ethnicities, with 126 reported to have pathogenic implications and the rest are considered variants of uncertain significance due to insufficient data at this time for further classification. Majority of germline variants described in literature are as a result of truncated proteins or inadequate mRNA expression leading to a straight-forward classification of their pathogenicity. Even though genotype-phenotype correlations for truncating variants are straight forward, missense variants require greater scrutiny. The ClinGen *CDH1* VCEP (variant classification expert panel) decides that

✉ Liying Zhang
LiyingZhang@mednet.ucla.edu

1 Department of Pathology, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

2 Department of Medicine, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

3 Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY 10065, USA

4 Department of Pathology and Laboratory Medicine, David Geffen School of Medicine, University of California, Los Angeles (UCLA), Los Angeles, CA 90095, USA

functional assays for missense alterations should not be used for *CDH1* variant classification the currently published in vitro and in vivo functional studies could not be confidently used to predict pathogenicity of E-cadherin missense variants. However, new assays that can be used to better predict the clinical effects of *CDH1* missense variants could be developed [7, 8]. Hence, clinical and genetic studies should include family histories, allelic frequency in healthy individuals, co-segregation within families and unrelated HDGC families, in silico predictions on its potential effects on normal splicing and in certain cases in vitro functional studies are required in assessing missense variant pathogenicity [9].

Previous studies reported on minigene assays of *CDH1* c.[715G>A] (p.Gly239Arg) (previous publication may refer to this variant as G239R) variant resulting in the deletion of 29 base pairs creating a preferential splice site in exon 6 [1, 10]. However, pathogenicity interpretations of this variant remain conflicting, with six likely pathogenic and two pathogenic classifications reported in ClinVar (Variation ID: 132709) at the time of manuscript submission. Given the life-altering recommendation for prophylactic gastrectomy in carriers of *CDH1* c.[715G>A], it is important to ascertain better pathogenicity of this variant, particularly if this variant is found in families with less striking histories of GCs as compared to the current family and for the quantification of mutant transcripts in variant carriers. We hypothesized that by performing splicing studies on a HDGC-prone family with *CDH1* c.[715G>A] variant will aid in resolving this variant's pathogenicity classification. We have submitted our data to ClinVar database (https://submit.ncbi.nlm.nih.gov/subs/clinvar_wizard/SUB8719376/overview), which can be found under submission ID: SUB8719376. Here we report that due to the disruption of splicing as well as strong segregation clinical data, c.[715G>A] variant should be classified as pathogenic, according to the 2015 American College of Medical Genetics and Association for Molecular Pathology (ACMG/AMP) Standards and Guidelines [8, 11].

Materials (Subjects) and methods

Patient samples

Studies were conducted on multiple individuals of a large DGC prone family, depicted in a four-generation pedigree, Fig. 1A. Genomic DNA extracted from peripheral blood samples of individuals III-2, 3, 5, 6, 7 and IV-1, 2 all tested heterozygous for *CDH1* c.[715G>A] variant. All participants elected to participate in a research study to further characterize the variant and provided written informed consent for the research protocol that was reviewed and approved by Memorial Sloan Kettering Cancer Center (MSKCC) institutional review board. A peripheral blood sample from the seven participants mentioned above was collected using the PAXgene Blood RNA tube and submitted for analysis. Control RNAs were extracted from eight unrelated individuals seen at MSKCC who do not carry the *CDH1* c.[715G>A] variant.

In silico analysis

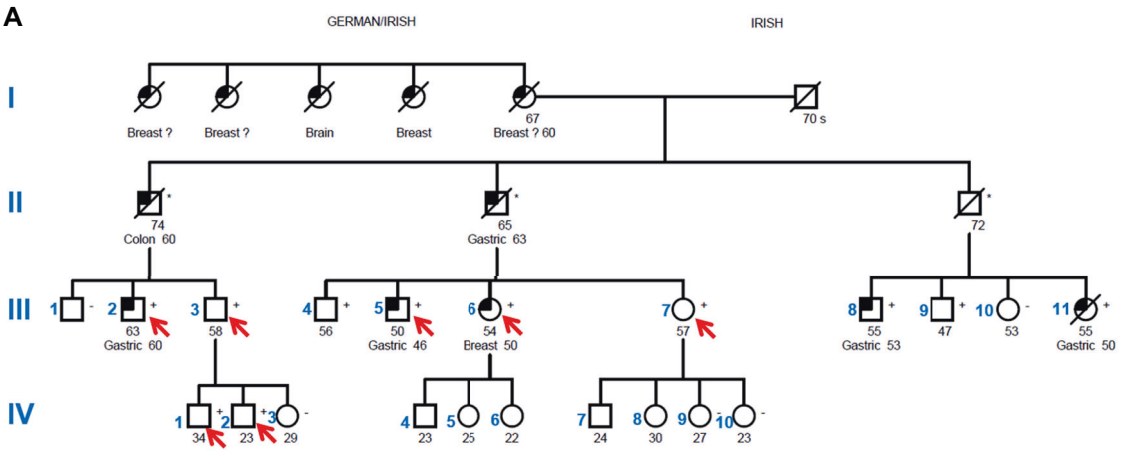
Sequence data spanning the *CDH1* locus for Homo sapiens [Chromosome 16: 68,737,292-68,835,541] was obtained from the Ensembl Genome Browser (https://www.ensembl.org/index.html), using exon numbering as in Ensembl transcript ID: ENST00000261769.10. Primers were designed using the Primer 3 software (https://bioinfo.ut.ee/primer3-0.4.0/). In silico evaluation of evolutionary conservation and slicing predictions were performed using Alamut (Interactive Biosoftwar), which includes SSF, MaxEnt, NNSPLICE, GeneSplicer, and HSF tools.

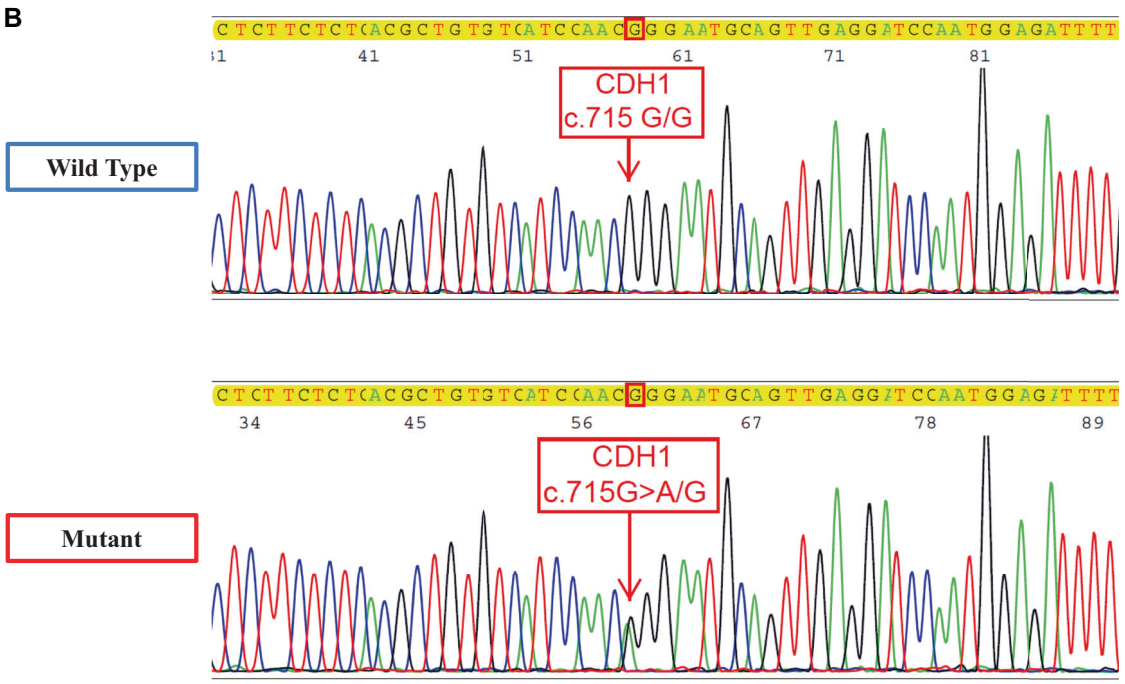
cDNA Synthesis

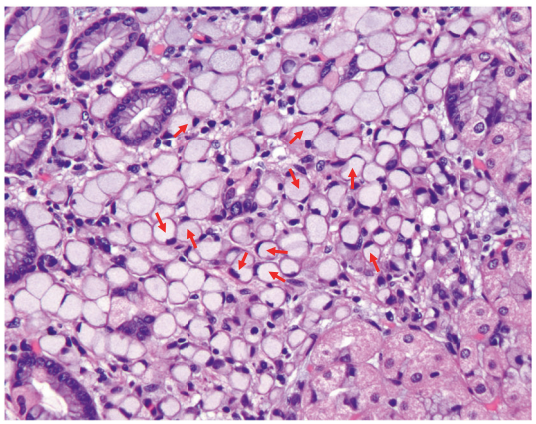
Total RNA was extracted from whole blood using the PAXgene Blood RNA Kit (PreAnalytiX, Qiagen, Valencia, CA), and was subsequently used for cDNA synthesis (SuperScript III First-Strand Synthesis System, Invitrogen Life Technologies, Carlsbad, CA). Forward primer in Exon 5, sequence: 5′-AAACAGGATGGCTGAAGGTG-3′ and the reverse primer in Exon 6, sequence: 5′-ATTCGGGCTT GTTGTCATTC-3′. Each PCR reaction consisted of 17.5 µl SIGMA REDTaqReady Mix 1x, 1 µl of 10 µM forward and reverse primers, 3 µl of cDNA and water to a final volume of 27 µl. Cycling conditions were 96 °C for 5 min, 94 °C for 30 s [45×], 55 °C for 45 s [45×], and 72 °C for 60 s [45×], with a final extension at 72 °C for 5 min.

Cloning

To investigate mutant allele impact on splicing, RT-PCR products were cloned into pCR4 TOPO vectors (Invitrogen, Carlsbad, CA), following manufacturer procedures (Invitrogen, Carlsbad, CA). DNA from colonies was amplified using the *CDH1* 10 µM primer sequences detailed above and subjected to direct DNA sequencing analysis using forward and reverse primers (BigDye Terminator v3.1 Cycle Sequencing kit and 3730 Genetic analyzer, Applied Biosystems, Foster City, CA).


A
GERMAN/IRISH
IRISH
I
II
III
IV
Breast ?
Breast ?
Brain
Breast
67
Breast ? 60
70 s
74
Colon 60
65
Gastric 63
72
63
Gastric 60
58
56
50
Gastric 46
54
Breast 50
57
55
Gastric 53
47
53
55
Gastric 50
34
23
29
23
25
22
24
30
27
23


B
Wild Type
CDH1
c.715 G/G
Mutant
CDH1
c.715G>A/G


C

◀ **Fig. 1 Diffuse gastric cancer in a HDGC family with *CDH1* c.[715G>A] variant. A** Four-generation HDGC-prone family pedigree. Asterisks in the second generation indicate obligate carriers. Individuals tested heterozygous for wild-type and mutant allele (*CDH1 c.715G/A*) mutant allele are indicated with (+), while individuals tested negative (*CDH1* c.715G/G) for mutant allele are indicated with (−). Red arrows point to participants in our study. Individuals affected by cancer are indicated by type on cancer and age of diagnosis for individuals whose information was available. **B** Germline *CDH1* sequence within Exon 6 demonstrating the wild-type sequence of G/G at c.715 position, top panel and the mutant sequence of A/G at c.715 position, bottom panel. **C** Hematoxylin and eosin (H&E) stained patient's III-2 biopsy specimen showing poorly differentiated infiltrating adenocarcinoma with mucinous and signet ring cell features. Arrows are pointing to signet ring cells, which are characteristic of HDGC.

Results

Patient and family history

A large DGC prone family, depicted in a four-generation pedigree, Fig. 1A was seen at MSKCC and multiple family members tested positive for *CDH1* c.[715G>A] variant. For the first-generation individuals, genetic testing is unavailable; however, from the maternal lineage it is evident that all the siblings have had cancer, although the types of cancer are unknown/unconfirmed for some individuals. In the second generation, the three deceased brothers would have to be obligate carriers of the *CDH1* c.[715G>A] variant since at least one progeny from each brother tested positive for this variant. In addition, second generation shows the first sign of HDGC in this family as one brother was diagnosed with GC at 63 years of age. All 11 family members in the third generation have had genetic testing with genetic test results for the *CDH1* c.[715G>A] variant depicted in the pedigree (Fig. 1A). For individual III-2, invasive adenocarcinoma with signet ring cell or poor cohesive type was evident at gastrectomy at 60 years of age, Fig. 1C, and individual III-5 was reported to have diffused T3 N0 adenocarcinoma at gastrectomy at 46 years of age. Individuals III-8 and III-11 were diagnosed with diffuse GC at age 53 and 50 years, respectively, and individual III-11 has died of GC at age 55. Pathology reports from individuals III-3, 6, 7, and 9, stated no malignancy at gastrectomy, however, individual III-6 has been diagnosed with invasive ductal carcinoma (and carcinoma in situ) breast cancer at age 50. Some members of the fourth generation have had genetic testing with IV-1 and IV-2 individuals testing positive for the *CDH1* familial variant.

Segregation analysis

The pattern of diffuse GC within the family described in Fig. 1A, is consistent with autosomal dominant inheritance of HDGC with the presence of a pathogenic *CDH1* variant. Our family studies revealed that members within the second generation were obligate carriers of the *CDH1 c.715G>A* variant and one was diagnosed with GC at 63 years of age. Four family members within the third generation have had diffuse GC and carry the *CDH1 c.715G>A* variant, individuals III-2, 5, 8, and 11 (Fig. 1A). Individual III-11 has died of GC at age 55. Therefore, *CDH1 c.715G>A* variant segregates with the GCs in this family.

Variant in silico analysis

Multi-species comparative genomic analysis was used to identify sequence homology at the *CDH1* p.Gly239 variant position in ten distantly related species: Human, Chimpanzee, Northern white-cheeked gibbon, Olive Baboon, Rat, Mouse, Dog, Platypus, Chicken, and Frog. Analysis indicated that the *CDH1* p.Gly239 is conserved across these species (data not shown). For in silico evaluation of *CDH1* c.[715G>A] potential effects on splicing, we utilized Alamut software, which incorporates five tools to predict the potential effects of *CDH1* c.[715G>A], on normal mRNA splicing. All five tools predicted that the missense variant would create a novel 3′ acceptor splice site within exon 6 of *CDH1* gene, which is consistent with the CAG/G consensus splice site (Fig. 2A). Additional in silico data shows high prediction scores for the creation of this novel splice site at c.717 nucleotide of *CDH1* gene (Fig. 2B).

Cloning analysis of RNA transcripts

To determine whether the *CDH1 c.715A* mutant can generate any *CDH1* exon 6 full-length transcripts, i.e., whether the mutant transcript completely disrupts normal splicing, we cloned the RT-PCR products into the TOPO sequencing vector. Analysis from a total of 99 colonies from 7 family members revealed that the wild-type sequence with *CDH1* r.906g (c.715G) was detected in 45 clones (Fig. 3A). Another transcript, *CDH1* r.879_907del, deletion of 29 nucleotides from the 5′ end of exon 6, was seen in 53 aberrantly spliced clones (Fig. 3B) indicating that the mutant transcripts preferentially disrupts normal splicing. In addition, one clone contained the mutant transcript, *CDH1* r.906g>a (c.[715G>A]), and was not aberrantly spliced (Fig. 3C). These results indicate that the mutant "A" transcript preferentially abolishes normal splicing through activation of a cryptic 3′ acceptor splice site within exon 6 of *CDH1*. Interestingly, *CDH1* r.906g>a (c.715A) transcript may result in full-length product with the mutant "A" transcript at *CDH1 c.715* position in about 1% of the transcripts.

Discussion

Our splicing studies and segregation analysis on a HDGC-prone family with *CDH1* c.[715G>A], variant produced the

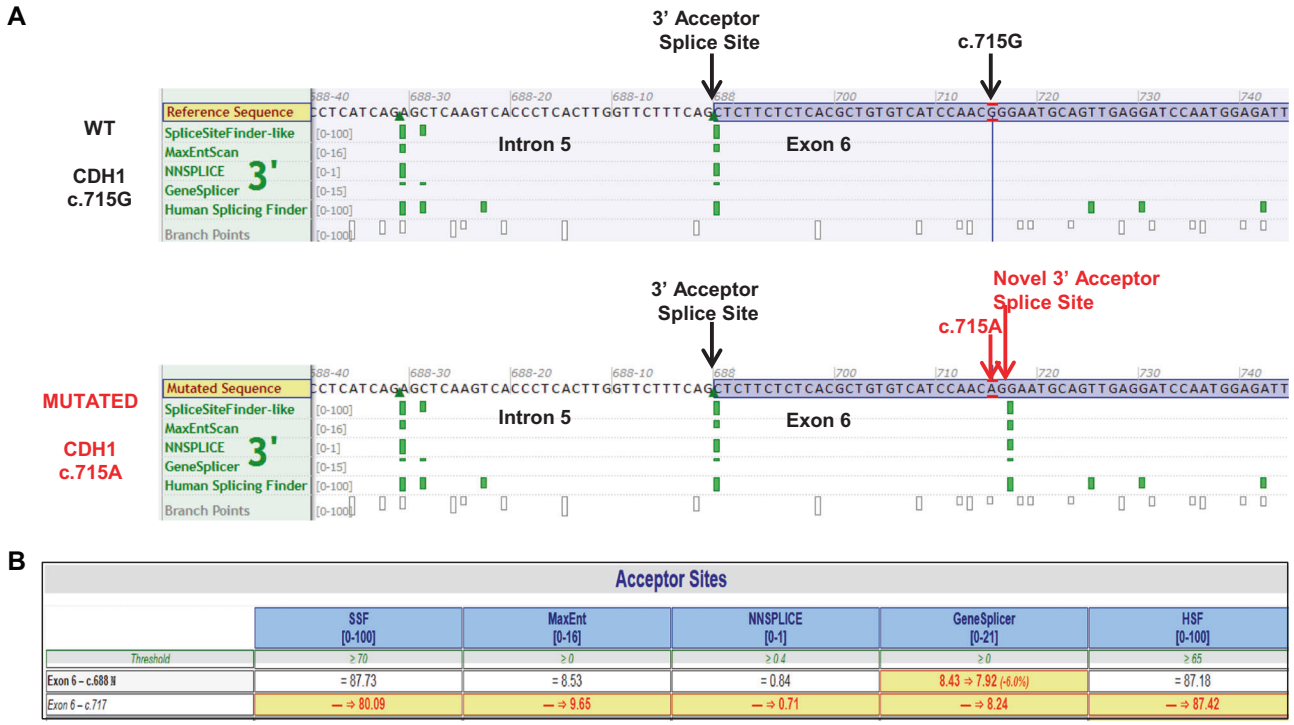


Fig. 2 In silico tools predict *CDH1* c.[715G>A] variant results in novel 3′ splice site. Prediction of the potential effects of the *CDH1* c.[715G>A], alteration on normal mRNA splicing. All five tools from the Alamut software predicted that the alteration would create a strong, novel 3′ acceptor splice site. **A** The green bars indicate variant would affect normal splicing and the height represents how closely the sequence is to the consensus splice site. Top panel showing reference sequence, bottom panel showing mutant sequence splicing predictions. **B** Splicing score predictions for the cryptic 3′ acceptor splice site.

necessary data for resolution of this variant's pathogenicity classification. We've determined that the deletion of 29 nucleotides from the 5′ end of exon 6, was seen in 54% of analyzed clones, indicating that the mutant allele preferentially disrupts normal splicing via the creation of cryptic 3′ acceptor splice site. The deletion of 29 nucleotides presumably results in a premature termination codon at position 233. The aberrantly spliced product results in a truncated E-cadherin protein within the first extracellular cadherin repeat, and consequently without the further translation of extracellular cadherin repeats 2–4, membrane-proximal extracellular domain, transmembrane domain and cytoplasmic domain, which are present in the wild-type full-length 882 amino acid E-cadherin. Thus, if the 29 base pair deletion transcripts do not undergo non-sense mediated decay, and are translated, would result in a non-functional E-cadherin protein.

We did observe ~1% of clones that contained the mutant transcript, *CDH1* c.[715A], and was not aberrantly spliced, which indicates that it is a leaky cryptic site with exhibited residual normal splicing. However, since more than half (54%) of transcripts had the 29 bp deletion, we believe the extremely low level of residual normal splicing transcript (~1%) and the normal E-cadherin generated from it are insufficient to maintain its normal function. We cannot discount that the level of leakage may vary among individuals which may explain the differences of age of onset and cancer types (gastric vs. breast) in the family members. Putative effects on protein function may be explained with a study by Petrova et al., providing evidence that suggests *CDH1* p.Gly239Arg change may not affect extracellular domain ability to undergo necessary adhesion activation changes, but instead results in E-cadherin resistance to the typical intracellular activation signaling events. The following mechanism was proposed: p.Gly239Arg change appears to uncouple adhesive binding functions of the ectodomain from p120-catenin-related signaling functions in the cytoplasm; hence, interfering with the transduction of signaling across the membrane, in the process known as inside-out signaling [12].

ACMG/AMP describe standards and guidelines for the assessment of sequence variants in their 2015 publication. These standards and guidelines provide a process for utilizing various types of variant evidence to classify variants identified in Mendelian gene disorders into the recommended five categories: pathogenic, likely pathogenic, uncertain significance, likely benign, and benign. Use of term "likely" was defined to be applied to variants where evidence support a greater than 90% certainty of a variant either being disease-causing or benign. For "likely

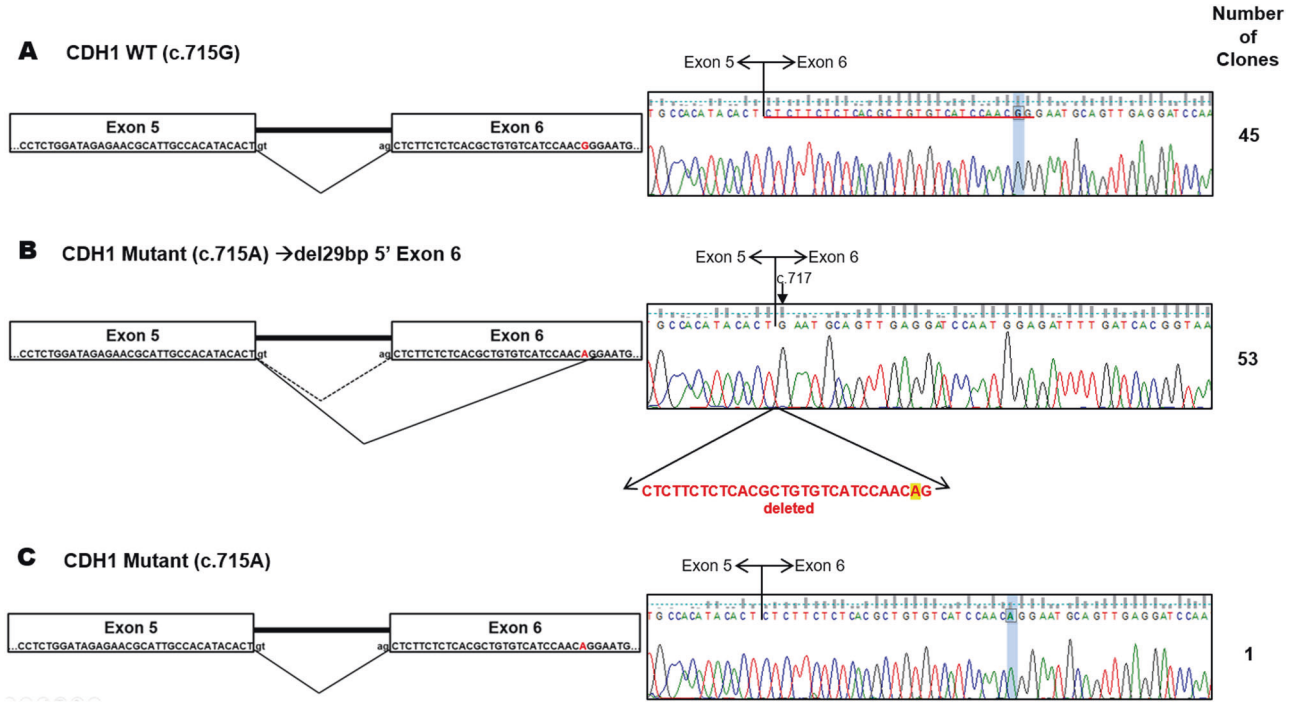


Fig. 3 *CDH1* c.[715G>A] mutant transcripts with novel acceptor splice site deletes 29 nucleotides from the 5′ end of exon 6. A Wild-type sequence with *CDH1* c.[715G] was detected in 45 clones. **B** Aberrantly spliced sequence with the deletion of 29 nucleotides from the 5′ end of exon 6, was seen in 53 clones. **C** One clone was not aberrantly spliced, but did contain the mutant transcript, *CDH1* c.[715A]. A total of 99 clones were used for this study. Blue highlighted region on the electropherograms, show *CDH1 c.715* position.

pathogenic" classification, this indicates a 90% confidence in pathogenicity and that the likelihood of new evidence resulting in downgrade reclassifications would be rare, which is sufficient in guiding patient care and permitting physicians to take action [11, 13]. Given that the guidance of patient care, disease management, and recommended medical interventions rely on variant classification, it is imperative to accurately and with as much thoroughness as possible, classify these variants. *CDH1* is considered a clinically actionable gene with pathogenic variants predisposing to fatal cancers. For patients with HDGC, average age of clinical presentation is 38 years and in the majority of cases patients are asymptomatic. However, upon presentation of specific symptoms, the disease has poor prognosis and is typically in advanced metastatic stages. For high-risk individuals, screening and surveillance upper endoscopy protocols are recommended, and individuals who carry a pathogenic *CDH1* germline variant are recommended a prophylactic total gastrectomy, which still remains the only effective clinical option for these individuals and family members. Furthermore, it is advised that individuals with positive pathogenic germline variants consider prophylactic total gastrectomies regardless of endoscopy findings due to the high, variable, unpredictable penetrance and poor detection of cancerous foci in *CDH1* carriers. Previous reports describe asymptomatic carriers with *CDH1* variants, although without evidence of disease on preoperative endoscopies, showed evidence of HDGC in resected gastrectomy specimens [14]. While "pathogenic" variant classification is applied to a variant that is considered to be a well-established cause of the patient's disease, "likely pathogenic" variant classification is considered the probable cause of the patient's disease, carrying some degree of uncertainty and is typically applied cautiously towards clinical decisions-making. Hence, "pathogenic" variant classification decreases the degree of uncertainty, providing physicians, variant carrying individuals, and family members better guidance towards disease management decisions.

In addition to our segregation analysis, in which we determined that *CDH1 c.715G>A* variant segregates within the HDGC-prone family presented here, there have been a number of previous reports of *CDH1 c.715G>A* observed in individuals with personal and family history of early-onset diffuse GC [1, 10, 15], and one with lobular breast cancer [16] (ACMG/AMP PS4) [8]. Previous RT-PCR and minigene assays demonstrated that this variant causes abnormal splicing with a deletion of 29 base pairs from the 5′ end of exon 6 [1]. In addition, *CDH1* c.[715G>A] variant was not observed in the major population databases of germline variation (1000G, ESP, and gnomAD) indicating it is not a common benign variant (ACMG/AMP PM2). Previous studies did not provide sufficient data for pathogenic classification. However, the addition of data presented here with transcript quantitation data, establishes that *CDH1* c.[715G>A] variant preferentially

abolishes normal splicing through activation of a cryptic splice site within exon 6 of *CDH1* gene (ACMG/AMP PS3) and qualifies this variant for pathogenic classification based on the ACMG/AMP standards and guidelines [8, 11]. Resolution of pathogenicity classification will help guide individuals and family members testing positive for *CDH1* c.[715G>A] variant in life-altering, yet potentially life-saving, surgical decisions.

Funding This study was funded by Department of Pathology, Memorial Sloan Kettering Cancer Center.

Compliance with ethical standards

Conflict of interest LZ: Honoraria (Future Technology Research LLC, BGI, Illumina); Honoraria and Travel and accommodation expenses (Roche Diagnostics Asia Pacific). Family members hold leadership position and ownership interest of Decipher Medicine. ZKS: Immediate family member serves as a consultant for Allergan, Adverum Biotechnologies, Alimera Sciences, Fortress Biotech, Genentech/Roche, Novartis, Optos, Regeneron, Regenxbio, and Spark Therapeutics.

References

1. 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.
2. 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.
3. 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. Oncol 2015;1:23–32.
4. Corso G, Marrelli D, Roviello F. E-cadherin germline missense mutations in diffuse gastric cancer. OA Cancer. 2013;1:4.
5. van Roy F, Berx G. The cell-cell adhesion molecule E-cadherin. Cell Mol Life Sci. 2008;65:3756–88.
6. Roberts ME, Ranola JMO, Marshall ML, Susswein LR, Graceffo S, Bohnert K, et al. Comparison of CDH1 penetrance estimates in clinically ascertained families vs families ascertained for multiple gastric cancers. JAMA Oncol. 2019;5:1325–31.
7. Figueiredo J, Melo S, Carneiro P, Moreira AM, Fernandes MS, Ribeiro AS, et al. Clinical spectrum and pleiotropic nature of CDH1 germline mutations. J Med Genet. 2019;56:199–208.
8. Lee K, Krempely K, Roberts ME, Anderson MJ, Carneiro F, Chao E, et al. Specifications of the ACMG/AMP variant curation guidelines for the analysis of germline CDH1 sequence variants. Hum Mutat. 2018;39:1553–68.
9. Oliveira C, Pinheiro H, Figueiredo J, Seruca R, Carneiro F. Familial gastric cancer: genetic susceptibility, pathology, and implications for management. Lancet Oncol. 2015;16:e60–70.
10. 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. Fam Cancer. 2013;12:503–7.
11. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17:405–24.
12. Petrova YI, Schecterson L, Gumbiner BM. Roles for E-cadherin cell surface regulation in cancer. Mol Biol Cell. 2016;27:3233–44.
13. Harrison SM, Rehm HL. Is 'likely pathogenic' really 90% likely? Reclassification data in ClinVar. Genome Med. 2019;11:72.
14. Shenoy S. CDH1 (E-Cadherin) mutation and gastric cancer: genetics, molecular mechanisms and guidelines for management. Cancer Manag Res. 2019;11:10477–86.
15. More H, Humar B, Weber W, Ward R, Christian A, Lintott C, et al. Identification of seven novel germline mutations in the human E-cadherin (CDH1) gene. Hum Mutat. 2007;28:203.
16. Susswein LR, Marshall ML, Nusbaum R, Vogel Postula KJ, Weissman SM, Yackowski L, et al. Pathogenic and likely pathogenic variant prevalence among the first 10,000 patients referred for next-generation cancer panel testing. Genet Med. 2016;18:823–32.