CDH1 and CTNNA1 Genetic Screening in Tunisian Patients with Hereditary Diffuse Gastric Carcinoma

Jihenne Ben Aissa-Haj, Maria Kabbage, Houcemeddine Othmen, Patrick Saulnier, Azer Ferah, Amira Jaballah-Gabteni, Mouna Medhioub, Amel Khsiba, Afifa Maaloul, Houda Belfkih, et al.

CDH1 and *CTNNA1* Genetic Screening in Tunisian Patients with Hereditary Diffuse Gastric Carcinoma

Jihenne BEN AISSA-HAJ (jihenne.benaissa@pasteur.utm.tn)
Institut Pasteur de Tunis https://orcid.org/0000-0002-1368-7169

Maria Kabbage
Institut Pasteur de Tunis

Houcemeddine Othmen
University of the Witwatersrand

Patrick Saulnier
Institut Gustave Roussy: Gustave Roussy

Azer Ferah
Institut Pasteur de Tunis

Amira Jaballah-Gabteni
Institut Pasteur de Tunis

Mouna Medhioub
Hospital Mohamed Taher Maamouri

Amel Khsiba
Hospital Mohamed Taher Maamouri

Afifa Maaloul
Institut Pasteur de Tunis

Houda Belfkih
Hospital Mohamed Taher Maamouri

Hassen Touinsi
Hospital Mohamed Tahar Maamouri

Sonia Ben Nasr
Hospital of military

Lamine Hamzaoui
Hospital Mohamed Tahar Maamouri

Emna Chalbi
Hospital of Mohamed Tahar Maamouri

Sonia Abdelhak
Institut Pasteur de Tunis

Mohamed Samir Boubaker
Institut Pasteur de Tunis

Mousaddak Azzouz

Hospital of Mohamed Tahar Maamouri

Etienne Rouleau

Gustave Roussy Institute: Gustave Roussy

Research Article

Keywords: CDH1, CTNNA1, germline mutations, Hereditary Diffuse Gastric Cancer, large rearrangements and Tunisian patients

Posted Date: June 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-617841/v1

Abstract

Background: Mutational screening of the *CDH1* gene is a standard treatment for patients who meet the criteria for Hereditary Diffuse Gastric Cancer (HDGC). In this framework, the classification of variants found in this gene is a crucial step for the clinical management of patients at high risk for HDGC. The aim of this study was to identify *CDH1* as well as *CTNNA1* mutation profiles predisposing to HDGC from Tunisia.

Methods: Thirty four cases were included for this purpose with a mean age at diagnosis of 48 years old. We performed Sanger Sequencing for the entire coding sequence of both genes and MLPA (Multiplex Ligation Probe Amplification) assay to investigate large rearrangements of the *CDH1* gene.

Results: As a result, three cases (8.82%) carried probably pathogenic variants in the *CDH1* gene. These variants involves a novel splice alteration, a missense located in exon 14 detected by Sanger Sequencing and a large rearrangement detected by MLPA assay.

Conclusion: Our results suggest that the *CDH1* p.G761R variant is probably pathogenic and involved in the conformational space shift of the protein. Molecular modeling analysis highlights a putative influence on the conformational properties of the Juxta-Membrane Domain core (JMD) that could result in destabilizing the protein-protein complexes and therefore impacting the downstream pathways. Also, a large deletion from the 5' locus including exons 1 and 2 of the *CDH1* gene implicating the signal peptide and a part of the precursor domain of the protein. These findings highlight the critical importance of screening for large *CDH1* rearrangements as well as mutations for the management of HDGC families and individuals at high risk for more personalised medicine. We therefore suggest a revision of the status of p.G761R mutation from Variant of Unknown Significance (VUS) to likely pathogenic.

Background

Gastric Carcinoma (GC) is the fifth most common cancer worldwide with approximately one million new cases registered in 2018 (5,7%) with a wide variation in geographical distribution. It represents the third leading cause of death from cancer worldwide, causing 783,000 deaths in 2018, accounting for 8.2 % of all cancer deaths [1, 2]. Gastric tumors are histologically and genetically heterogeneous, likely due to the exposure of populations to different environmental risk factors and different genetic predispositions. Despite a decline in incidence and mortality, the burden of GC remains relatively high [3]. Incidence predominates in populations from certain geographic regions and socioeconomic groups [4, 5]. High incidence areas include East Asia, Eastern Europe, Central and South America, Japan and Korea, while low incidence rates are observed in South Asia, North and East Africa, and North America [6, 7].

In Tunisia, GC represents the seventh most frequently diagnosed cancer with an incidence of 4% (637 new cases per year) and the fifth cause of death with a rate of 5.8% [1, 2]. A recent Tunisian study reported a significant increase in the diffuse type with a concomitant decrease in *Helicobacter Pylori* (HP)

in the last decade [8]. However, epidemiological data on hereditary forms are not available in Tunisia or in North Africa.

Histologically, GC is divided into three main subtypes intestinal diffuse and mixed which have different epidemiological and prognostic features [9–11]. Sporadic Gastric tumors represent 90% and Familial Clustering is rare and represents about 10%. Only 1 to 3% are hereditary [12, 13] including several syndromal forms such as Familial Intestinal Gastric Cancer (FIGC), Proximal Polyposis of the Stomach (GAPPS), Lynch Syndrome (LS) and Hereditary Diffuse Gastric Carcinoma (HDGC) (OMIM: 137215). HDGC is an autosomal dominant inherited disorder caused by germline mutations of the *CDH1* gene with a risk of developing a diffuse type starting at age 45 [14].

CDH1 is located in the 16q22.1 locus with 16 exons and mutation carriers have 70–80% a lifetime risk of developing GC [15, 16] with an estimated cumulative risk by age 80 years of 67–70% for males and 56–83% for females. Furthermore, the cumulative risk of breast cancer (BC) in women with *CDH1* mutations was 39% (95% CI, 12-84) and the combined risk of GC and BC was 90% by age 80 [17, 18]. To date, more than 155 mutations resulting in loss of function of the *CDH1* gene have been reported worldwide [17, 19, 20]. However, no hotspots have been characterized.

E-Cadherin (OMIM: 192090), *CDH1* gene product belongs to the Cadherin superfamily, is a calcium-dependent cell-cell adhesion molecule that plays a critical role in the establishment of epithelial architecture, maintenance of cell polarity and differentiation. It consists of a single transmembrane domain linked to a cytoplasmic domain, and an extracellular domain consisting of five tandemly repeated domains called EC1–EC5, which are exclusive to cadherins family [13, 21, 22]. According to the International Gastric Cancer Linkage Consortium (IGCLC), patients who meet the inclusion criteria for HDGC must be tested for *CDH1* germline mutations [17]. Although approximately 14–50% of cases meeting IGCLC inclusion criteria are carriers of pathogenic germline mutations of the *CDH1* gene [17, 19, 20], several families meeting HDGC criteria have no detectable pathogenic *CDH1* variant. Recently, whole exome sequencing (WES) analysis identified a truncating variant in the *CTNNA1* gene in an HDGC family that has no detectable pathogenic variant of *CDH1* gene [23]. *CTNNA1* encodes for α-catenin, which is a protein that interacts with E-cadherin. It plays an important role in the cell adhesion process by linking E-cadherin located on the plasma membrane to actin filaments. Germline *CTNNA1* truncating mutations have been reported in patients with HDGC highlighting its important role in this disease risk susceptibility [17, 23–26].

To the best of our knowledge, no study on hereditary GCs has been conducted so far with the aim of identifying the mutation spectrum, neither in Tunisia nor in North Africa. In the present study, our aim is to identify the genetic profiles of *CDH1* and *CTNNA1* in Tunisian patients with DGC to find a new tool for molecular screening of individuals at high risk in the Tunisian population. To do so, we selected a cohort of 34 cases of DGC with suspected HDGC meeting or not IGCLC inclusion criteria.

Materials And Methods

Study population

This study was conducted in accordance with the Declaration of Helsinki and with the approval of the Institutional Review Board (IRB) of Institut Pasteur de Tunis. It is a retrospective and consecutive study that included 34 unrelated Tunisian consent patients between 2009 and 2019. Of these included cases, 22 fulfil the HDGC definition Criteria [19] (1) Two or more GC cases regardless of age, at least one confirmed DGC, in first-degree and second-degree relatives (2) One case of DGC before 40 years old (3) Personal or familial history of DGC and lobular breast cancer (LBC), one diagnosed before 50 years old. Blood samples were collected from 33 index cases and their consenting relatives in the gastroenterology department of Hospital Mahmoud Matri Nabeul and one case in the oncology department of Military Hospital of Tunisia and sent to our laboratory for molecular analysis.

DNA extraction

Total genomic DNA (gDNA) was isolated from peripheral blood before any treatment using the salting-out method or the DNeasy blood kit from Qiagen according to the manufacturer's instructions. Somatic DNA (sDNA) was performed using the All Prep DNA/RNA/Protein Mini Kit from QIAGEN according to the manufacturer's instructions (Qiagen GmbH, Hilden, Germany). DNA quality and concentration were measured using a NanoDrop™ spectrophotometer and Qubit.

Primer design

Primers covering all coding exons and border regions of the *CDH1* and *CTNNA1* genes were designed using Primer Express and amplified by PCR. Forward and reverse primers contained the extensions 18F tail (ACCGTTAGTTAGCGATTT) and 18R tail (CGGATAGCAAGCTCGT) at their 5' ends.

Genetic analysis of *CDH1* and *CTNNA1*

Screening of the coding and flanking regions of both *CDH1* and *CTNNA1* genes was performed on gDNA extracted from peripheral blood. Sanger sequencing was used to screen *CDH1* and *CTNNA1* genes for index cases and their voluntary relatives. Polymerase chain reactions (PCR) were performed in a volume of 10µl containing 10ng gDNA.

Amplification of *CDH1* exons 1-10 and exons 14-16 was performed using GoTaq Hot Start Polymerase (Promega) and exons 11-13 using Hot Start Taq Polymerase (Qiagen). Amplification of *CTNNA1* exons 2-13 was performed with GoTaq Hot Start Polymerase (Promega) and exons 14-18 using Hot Start Taq Polymerase (Qiagen). The settings of all primers of *CDH1* and *CTNNA1* (sequences, lengths and concentrations) used in the present study are shown in **Additional files 1 and 2** respectively. PCR products were purified using ExoSAP-IT, sequenced using a BigDye terminator, purified using Sephadex G50 and then Sanger Sequencing analyzed using an automated sequencer (ABI 3730; Applied Biosystems, Foster City, CA). Generated Data was analyzed using SeqScape version 3.2 and BioEdit Sequence Alignment Editor Version 7.2.5. Variants found in our study were described using the recommendations of

the Human Genome Variation Society 'HGVS' [27], and interpretations were based on the American College of Medical Genetics and Genomics 'ACMG' guidelines [28, 29].

Search for large deletions/duplications of the *CDH1* gene using multiplex ligation dependent probe amplification (MLPA) assay

A total of 28 gDNAs and 10 sDNA were screened for copy number variations (CNV) using Multiplex Ligation Dependent Probe Amplification (MLPA). The assay was performed using the SALSA P083-D2 *CDH1* MLPA kit (MRC-Holland) according to the manufacturer's instructions. The kit contains 35 probes with amplification products ranging from 130 to 140 nucleotides. These include 20 probes for the *CDH1* region, one upstream flanking probe and one downstream flanking probe. Samples containing 50 ng DNA were analyzed by MLPA using probe mix P002. Female control DNA was obtained from Promega. Blood DNA from an individual known to have a deletion of exon 11 of the *CDH1* gene was used as a positive control. MLPA products were run on the ABI Prism 3730 xl Genetic Analyzer (Applied Biosystems) and analyzed using the Peak Scanner™ software v1.0 (Applied Biosystems).

In silico prediction tools

The predicted effects of all identified variants were evaluated using several *in silico* prediction tools to support functional effect and pathogenicity, such as UMD Predictor (http://umd-predictor.eu/), Sorting Intolerant From Tolerant (SIFT) (http://sift.jcvi.org/) used to examine the degree of conservation for amino acid residues across species and find changes in protein structure and function, PolyPhen-2 (http://genetics.bwh.harvard.edu/pph2/), Protein Variation Effect Analyzer (PROVEAN) (http://provean.jcvi.org/) to filter sequence variants to identify non-synonymous or indel variants thought to be functionally important. Mutation Taster (http://www. mutationtaster.org/) was used to assess the impact of mutations on protein function and to study the effects on splice sites and mRNA expression, FATHMM (http://fathmm.biocompute.org.uk/) to predict the functional consequences of coding variants (Non-Synonymous single nucleotide variants) and noncoding variants and Varsome (https://varsome.com/), which is a variant knowledge community, data aggregator and variant data discovery tool. All identified variants were classified based on their pathogenicity (benign, likely benign, pathogenic or likely pathogenic). All rare and novel variants were cross-referenced with general mutation databases (ClinVar) (https://www.ncbi.nlm.nih.gov/clinvar/), Leiden Open Variant Database (LOVD) (https://www.lovd.nl/) and UniProt as well as published reports to prioritize them for processing workup.

In silico analysis method of c.2281G>A mutation effect

Preparing the structures

The structure of Cadherin-1 (E-Cadherin) has been partially solved. To date, there are 14 available crystal and Cryo-EM structures from the Protein Data Bank (PDB), from which we selected the one containing the mutation. We used MODELLER [30] to generate the structure of the mutant. To investigate the functional effects of the mutation, different protocols were applied. The FlexPepDock method from the ROSETTA

package was used to refine peptide-protein complexes. The protocol retains 300 structures of the low and high-resolution stages before calculating the energy score. In addition we used MODPEP [31] to generate an ensemble of conformations that are likely to bind the target for the wild-type and mutant forms. Within the MODPREP workflow, psipred was applied to assign the secondary structure, whereupon the structure of the peptide wasassembled using experimentally collected data. Inthe final stage, molecular dynamics was applied to refine the structures. The ensemble consists of 1000 conformations, which are then processed for analysis. Finally, we run an *in silico* alanine scanning protocol from ROSETTA [32] to calculate the variation in the binding energy (DDG) between two partners after mutating each residue to Alanine. Data from the *in silico* study were analyzed using the MDTraj python library [33].

Predicting the change in consensus splice sites

To predict the change of consensus splice sites, we used SPiCE [34]. It combines *in silico* predictions from Splice Site Finder-like (SSF-like) and MaxEntScan (MES) (2, 3) and uses logistic regression to define two optimal decision thresholds: optimal sensitivity threshold (ThSe) and optimal specificity threshold (ThSp), 0.115 and 0.749 respectively.

Immunohistochemistry

To evaluate the expression of E-cadherin, we performed immunohistochemical staining on Formalin Fixed Paraffin Embedded (FFPE) samples corresponding to gastric tumor tissue. Immunostaining was performed with a primary mouse monoclonal against E-Cadherin (NCL-L-E-Cad, clone 36B5, Novocastra TM, Biopole), recognizing the external Nt domain, followed by incubation with a peroxidase coupled Post Primary Rabbit anti-mouse IgG according according to the manufacturer's instructions using Novolink MPolymer Detection Systems Kit (Biopole). Detailed protocols are available upon request. The final Immuno Reactive score (IRS) ranged from 0-3 Scores 0-1 for the group with negative to weak expression, score 2 for the group with moderate expression and score 3 for the group with high expression considered normal expression.

Results

Characteristics of the study population

All tumors were classified as Diffuse Carcinomas by two independent pathologists. As shown in **Additional file 3 and table 1**, *CDH1* and *CTNNA1* were sequenced for 34 unrelated Tunisian GC patients. The cohort included 13 (38,24%) males and 21 (61,76%) females with a mean age of 48 years at diagnosis (range 23-82 years). There were two patients with a family history of DGC in first or second degree relatives and 15 patients had DGC at an age less than 50 years. Majority of the patients (14: 41.18%) had an advanced stage of disease (T3 and T4) (**Table 1**). According to family history, some families had other cancers such as BC (37,5%), CCR (37,55%) and other tumors (25%).

CDH1 Genetic Testing

A total of 34 Tunisian patients with DGC were selected for *CDH1* germline mutations screening. In a first step, a total of 27 *CDH1* variants (**Additional file 4**) were identified and filtered using some exclusion criteria 1) Do not consider polymorphisms and synonymous variants 2) Exclude variants reported in Clinvar as Benign or Likely Benign. Of the 27 variants, two were novel (c.765G>A and c.1565+3_1565+4delinsGT) and 10 were classified as polymorphisms, because the Minor Allele Frequency (MAF) in the 1000Genomes database was greater than 1% (c.48+6C>T, c.531+10G>C, c.1320+45G>C , c.1566-80C>G, c.1712-52G>C ,c.1896C>T, c.1937-13T>C , c.2076T>C, c.2164+17dupA, c.2439+52 G>A and c.2634C>T). Two out of 27 (7.41%) variants, previously identified by Sanger Sequencing, were predicted to be deleterious by various *in silico* tools (**Table 2**).

The distribution of identified *CDH1* variants is shown in **Fig. 1.** Exonic coding variants represented 10 out of 27 variants (37.04%) and based on the ClinVar database, variants were classified as benign or likely benign (62.96%), one was described as VUS (3.7%) and two were novel variants (7.4%). Indeed the splice site variant (intron 10) seems to have a deleterious effect. The VUS “c.2281G>A” located at Exon 14 corresponds to the cytoplasmic domain of E-Cadherin.

Variant of uncertain significance

The index case « JI-007» had a *CDH1* germline mutation in the cytoplasmic domain of E-cadherin, at nucleotide position 2281 (c.2281G>A). This mutation causes a change in protein structure at position 761 from amino acid G to R. No information was found regarding linkage disequilibrium related to this variant. It is a rare variant “rs779648243” with a MAF of 0.0012 in the general population with an "Uncertain significance" in ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/). VUS, related to a variant not yet classified as benign or pathogenic, is a common challenge in genetics and a good candidate for computational predictors [35]. Several *in silico* prediction tools have been tested to evaluate the potential functional effect of this variant. It is located in a conserved protein domain throughout several species. Using Mutation taster this variant has been described as “Disease causing”. Both Polyphen and SIFT described it as “Damaging”. It was classified by UMD predictor as “Pathogenic” and by Provean as “Deleterious” as well as by other complementary online prediction tools (FATHMM, MutationAssessor, LRT). Although this variant has been previously reported [36], but this is the first time it is identified in a Tunisian patient.

Novel variants

In the current study, we found two novel variants (7.4%) c.765G>A in exon 6, resulting in a synonymous variant (p.Gln255=) and c.1565+3_1565+4delinsGT at intron 10. These variants have not been previously reported in the literature or in a variant-tracking database.

In total, three potentially pathogenic mutations predicted by *in silico* tools were found in our study. These mutations were carried by three different patients who meet the 2015 HDGC test criteria (3/22 representing 13.64 %) (**Fig. 2**).

The first case « JI-014 » had the c.1565+3_1565+4delinsGT in intron 10. She was referred for a molecular screening of *CDH1* as she was suspected of having HDGC from the Oncology Department of Military hospital in Tunisia .She was a 42 years old woman diagnosed with DGC (T4N1M1) at Antro pyloric and treated with Palliative Chemotherapy.. Her brother and paternal uncle were diagnosed with GC and died at the age of 25 and 80 years respectively «**Fig. 2A**». This index case showed a loss of E-cadherin expression in gastric tumor tissue. This indel is predicted to affect splice sites. Indeed, the donor Site is decreased 3 bps upstream with a percentage of -44.5% (MaxEnt:-64%; NN SPLICE: -25.1%, SSF: -16.8%) resulting in a cryptic site "**Fig. 3**".

The second index case « JI-007 » was a man diagnosed with DGC (T4N0M1) at the age of 25 and died at the same age. He has a silent pedigree «**Fig. 2B**» without GC or other cancers family history. This patient carried the predicted deleterious variant (c.2281G>A) in exon 14 and showed loss of E-cadherin expression in gastric tumor tissue "**Fig.7CD**".

The third index case « JI-020 » carrying the large heterozygous deletion detected by MLPA assay was a 79 years old woman diagnosed with DGC and treated with Total Gastrectomy. She has a sister diagnosed with BC at age 50 and died at the same age. She also has a daughter diagnosed with CCR at the age of 48 «**Fig. 2C**». Due to the unavailability of her tumor tissue, we were unable to study the immunohistochemistry of E-Cadherin.

Screening of large deletions/duplications in the *CDH1* gene unsig multiplex ligation-dependent probe amplification (MLPA) assay

Since heterozygous large deletions or duplications may remain undetected by conventional sequencing, we searched for possible rearrangements of the *CDH1* locus using the MLPA assay [37]. Results were analyzed using the Coffalyser software (MRC-Holland, Amsterdam, Holland): for normal sequences a probe dosage ratio of 1.0 is expected; probes with a dosage ratio <0.7 or >1.3 indicate deletions or duplications in the corresponding exons respectively. By comparing the control probes with the studied cases, we found that only the index case JI-020 had abnormal multiplex ligation-dependent probe amplification (MLPA) features with more than 45% reduction in signal in one or more exons of the *CDH1* gene, indicating the presence of large deletions in the *CDH1* locus. As shown in **Fig. 4 and Additional file 5** the "JI-020" had deletions of 4161 pb at the 5'-end of the gene, spanning at least exons 1 and 2 from position 67,325572 to 67,329733. No other abnormalities were observed in the remaining patients.

In silico analysis of the c.2281G>A mutation

The mutation c.2281G>A occurs in the cytoplasmic tail of E-Cadherin whose role is to regulate downstream cell-cell adhesion signalling (**Fig. 5A**). The corresponding amino acid was solved as part of the Juxta-Membrane Domain core region (JMD core) [38] which interacts with p120 catenin (p120) (**Fig. 5B**). In the co-crystal structure, it corresponds to an 18 amino acid peptide (residues 756-773) that interacts with the Armadillo (ARM) domain of p120. G761 interacts with the depth of the concavity formed by p120.

We first refined the s JMD core_WT/ARM and JMD core_R761/ARM complexes to evaluate whether the mutation would significantly affect the peptide-protein interface. The complexes with the best ROSETTA scores show only a low Root Mean Square Deviation (RMSD) of 0.16 Angstroms. The refined wild type model shows more favorable ROSETTA scores calculated from the 10 best conformations obtained with a median value of -640.55 and a standard deviation of 2.95. The mutant shows a less favorable median value of -574.784 and a standard deviation of 0.86. In addition, *in silico* alanine scan analysis did not reveal that position 761 is a hotspot residue for interaction with the ARM domain (**Additional file 6**). However, we found that the R761 mutation induces intrachain salt-bridge formation in the JMD core by pairing with E759, which partially interacts with K574. The latter pairs only with K574 of the p120 ARM domain to form a salt bridge in the wild-type form (**Fig. 5C**).

We then investigated the hypothesis that the conformational properties of the JMD core are affected by the mutation. We generated a trajectory of 1000 putative bound conformations for the WT and the mutant forms using MODPREP. We found that WT structure was able to capture more conformations similar to the bound crystal shape after structural adjustment (**Fig. 5D**). For example, seven conformations showed an RMSD value lesser than 2.5 Angstroms while the number increased to 30 at Angstroms at a cutoff of 3 Angstroms. On the other hand, we reported 0 and 2 conformations respectively for the same RMSD thresholds of the mutant form. From the ensemble, we calculated the Root Mean Square Fluctuation (RMSF) per amino acid of the JMD core (**Fig. 5E**). We found that the WT form is more stable while the mutant form shows an increase in the flexibility for the R761 and the G763-D868 segment. In addition, we found that R761 in the mutant form was able to form transient salt bridges with 8 acidic residues of the JMD core including D756, E757, E758, E762, E763, D764, D766, and D768 accounting for 5% of the total ensemble sampled. These residues represent the total acidic amino acids of the JMD core.

CTNNA1 Genetic Testing

A total of 34 Tunisian patients with DGC were selected to screen for *CTNNA1* germline mutations. All identified variants are summarized in **Additional file 7**.The distribution of identified *CTNNA1* variants is summarized in **Fig. 6**. All identified coding variants were synonymous, representing 8 out of 15 (53.33%). Based on the ClinVar database, the variants were classified as benign (53.33%) and a novel variant identified in two patients (JI-001 and JI-006) in intron 16 (c.2193-68C>T) was predicted to be a polymorphism.

Immunohistochemistry

The E-Cadherin expression pattern was investigated by IHC in only 23/34 GC cases, for which the FFPE tumor tissues were available. **Table 1 and Additional file 8** summarize the clinicopathologic features of the studied patients. Our results showed negative E-Cadherin immunostaining in 30.43% (7/23) cases versus 69.57 % (16/23) positive cases. The expression groups were classified into negative to weak expression (score 0-1) representing 39.13 % (9/23) with a normal membranous E-cadherin expression pattern in cryptes and adjacent glandular cells (**Fig. 7B**). The moderate expression group (score2)

included 21.74% (5/23) cases and the high expression group (score3) included 39.13 % (9/23) cases. “Abnormal” E-cadherin expression pattern includes both lost/reduced membranous expressions (**Fig7CD**).

Discussion

In the current study, we screened 34 DGC patients from unrelated families of North-East Tunisian region with suspected HDGC to shed light on the molecular basis of GC. This region is known to have a relatively high proportion of digestive cancer syndromes and diffuse gastric tumors. However, no epidemiological data are available. To do so, we performed a screening of the full coding and flanking regions of both *CDH1* and *CTNNA1* genes as well as *CDH1* large rearrangements. An IHC was used to investigate the E-cadherin protein expression profile in the available GC FFPE tissues as well. To our knowledge, this is the first study investigating the genetic mutation profile of patients with HDGC syndrome in Tunisia and in North Africa. As results, we identified three probably pathogenic mutations as predicted by *in silico* prediction tools (a splice alteration, a large rearrangement and a missense probably pathogenic). Approximately 10 to 20% of pathogenic variants are found in the *CDH1* gene from families that meet the IGCLC criteria [39, 40] which is partially consistent with our results as we found probably deleterious *CDH1* variants in 13.64 % of HDGC screened patients. Comparing with the literature, approximately 92% were already reported as described in the **Additional file 9**.

The c.2281G>A is a very rare variant (PM2). Structural bioinformatic analysis showed evidence in favor of a deleterious effect for the c.2281G>A mutation. According to our results, the c.2281G>A mutation can cause a shift in the conformational space of the protein. The mutation allows only a handful of conformations relevant to binding to be visited while the free energy landscape is scanned according to similar mechanisms described earlier[41, 42]. This is consistent with Glycine being endowed with more flexibility compared to Arginine. This could allow more efficient sampling of functionally relevant structures including the bound form. Since Glycine is able to form intrachain salt bridges with the acidic residues of the JMD core (**Fig5.AB**), the mutation could have a significant impact on the conformational space of the protein, thus also explaining the flexibility of the mutant form. Such a property would have a significant consequence by restricting the plasticity of the mutant form to conformations other than that of the WT form. Moreover, G761 has been shown to be highly conserved in the JMD core and the GGG motif (residues 759-763) is crucial for the formation of a rotational structure that interacts with residues F437, W477, and N478 of p120 (PP2-PP3) [43].

For the “JI-007” index case with the c.2281G>A *CDH1* mutation, we observed a loss of E-Cadherin protein expression by IHC in Gastric cancer FFPE tissue. Indeed, the impairment of the protein-protein complex induced by the mutation may explain reduced E-cadherin function as predicted by *in silico* modeling analysis which probably leads to HDGC. This is a major hallmark of tumor malignancy [44] which is induced by a variety of factors including transcriptional regulation, mutation, and aberrant cadherin internalization (**Fig.8**) [45]. The Ubiquitin -dependent endocytosis of E-cadherin caused by E3 ubiquitin ligase Hakai [46] and PS1/g-secretase-mediated cleavage of E-cadherin (**Fig.8**) [47] were associated with

the depletion of E-cadherin from the cell surface [38] highlted by the loss of membranous staining of E-cadherin in tumor cells in our results (**Fig.7CD**).

The 761G is the third amino acid in a peptide sequence composed of 12 amino acids (from 758 to 769) which is crucial for the link of E-cadherin cytoplasmic domain (PM1) to PS1 and p120 (**Fig.8**). In fact, PS1 stimulates cell adhesion via its interaction with E-Cadherin, Beta-catenin, p120. So p.G761R clives the E-Cadherin and causes the destabilization and the disassembly of this complex (**Fig.8 Step1**) resulting in an increase of the cytoplasmic pool of Beta-Catenin (**Fig.8 Step 3**) and thus negatively regulates the signaling pathways (**Fig.8 Step4**) [50].

Moreover, JI-007's sDNA, examined by Sanger sequencing, showed a loss of heterozygosity for this variant. This is an additional criterion for classifying the variant as probably pathogenic (PP4). Unfortunately, this variant was not tested in index case's relative to verify familial segregation, as they did not give their consent.

Currently, this variant has PM1, PM2, PP2, PP3 and PP4 criteria according to the ACMG classification [29]. In light of these findings, we suggest reconsidering the clinvar classification of this variant from VUS (Class3) to likely pathogenic (Class4). Further functional studies or cosegregation analysis should be performed to confirm its pathogenicity.

On the other hand, MLPA analysis showed that "JI-020" contained a large deletion from the 5' locus including exons 1 and 2 of the *CDH1* gene implicating the signal peptide and part of the precursor domain of the E-cadherin protein. Large *CDH1* deletions are rare and occur in only 4% of HDGC families by mechanisms mainly involving non-allelic homologous recombination in Alu repeat sequences [51]. A recent study reported that the 5' breakpoint was 279 bp away from a breakpoint associated with exons 1-2 deletion, which could be a recombination hotspot due to two Alu sequences being very close to each other. The patient containing this deletion had bilateral LBC with metastases at the age of 32 and died of DGC three years later. She came from a large family with eight siblings, but none of them had cancer. E-cadherin expression was not detected in the bilateral LBC by IHC [51]. Importantly, the immature molecule contains a short signal peptide and a precursor region preceding the extracellular domain prior protein processing [52]. Signal peptides serve as docking sites for the signal recognition particle, the main molecule responsible for detecting the translocation code of secretory and membrane proteins [53–55]. It was reported that the *CDH1* signal peptide core is essential for E-cadherin synthesis and delivery to extracytoplasmic regions. The failure in this checkpoint leads to loss of protein expression and function, and ultimately to disease [56]. Due to the unavailability of the tumour tissue, we were unable to perform an E cadherin IHC to confirm this result for this index case. These findings highlight the critical importance of screening for large rearrangements of *CDH1* as well as *CDH1* mutations for the management of HDGC families and individuals at high risk.

The clinical utility of identifying the *CDH1* mutation spectrum determines whether unaffected relatives are at risk for developing DGC or LBC. Regarding carriers of the *CDH1* pathogenic variant, the updated recommendations are total prophylactic, reduced emphasis on prophylactic total gastrectomy for weak

family history and total gastrectomy for positive biopsies. If there is a family history of LBC, annual breast surveillance is recommended, and bilateral risk-reducing mastectomy with or without reconstruction should be considered [57].

In addition to *CDH1* mutations ,pathogenic variants in *CTNNA1* are known to occur in a small proportion of families with HDGC [23, 57]. Nevertheless, little is known about the penetrance of the *CTNNA1* gene [24]. In the current study, no pathogenic *CTNNA1* variants were found. All coding variants were synonymous. Based on the ClinVar database, variants were classified as benign (53.33%). A novel variant was identified in intron 16 (c.2193-68C>T) in two patients (JI-001 and JI-006) which was predicted to be benign by all *in silico* prediction tools.

Our results indicate that the genetic mutation profile of studied patients with suspected HDGC is different from other families in other populations, as we did not find any reported pathogenic *CDH1* and *CTNNA1* mutations. These findings could be explained by the significant variability in GC frequency worldwide as well as risk factors [1]. Our results highlight the particular genetic background of the Tunisian population compared to others. Indeed, several papers published in Tunisia have reported the particular genetic background of our population [58–61].

Conclusions

The identification of hereditary cancer susceptibility genes is an indispensable step in understanding the basic molecular events of tumorigenesis and for the clinical management of affected families. The identification of *CDH1* mutations in HDGC and the emergence of gene-directed gastrectomy as a therapeutic strategy represent the culmination of a successful collaboration between molecular biologists, geneticists, oncologists, gastroenterologists and surgeons. In this first Tunisian *CDH1* study, the frequency of identified mutations is comparable to that reported in the literature with the presence of a large *CDH1* rearrangement.

List Of Abbreviations

HDGC	Hereditary Diffuse Gastric Carcinoma
MLPA	Multiplex Ligation Probe Amplification
IHC	Immunohistochemistry
GC/DGC	Gastric Carcinoma / Diffuse Gastric Carcinoma
FIGC	Familial Intestinal Gastric Cancer
GAPPS	Proximal Polyposis Of the Stomach
LS	Lynch Syndrome
BC/LBC	Breast Cancer/Lobular Breast Cancer
IGCLC	International Gastric Cancer Linkage Consortium
gDNA	Genomic DNA
sDNA	Somatic DNA
ACMG	American College of Medical Genetics and Genomics
HP	*Helicobacter Pylori*
PDB	Protein Data Bank
SSF	Splice Site Finder-Like
ThSe	Sensitivity Threshlod
ThSp	Specificity Threshold
FFPE	Formalin Fixed Paraffin Embedded
IRS	Immuno Reactive Score
CCR	Colorectal Cancer
5'UTR	5' Untranslated Region
B/LB	Benign/Likely Benign
VUS	Variant of Uncertain Significance
NR	Not Reported
NI	Non Indicated
ARM	Armadillo
JMD core	Juxta-Membrane Domain Core region
RMSF	Root Mean Square Fluctuation
RMSD	Root Mean Square Deviation

WES	Whole Exome Sequencing
PM1/PM2	Moderate
PP2/PP3/PP4	Supporting

Declarations

Ethics approval and consent to participate

Written informed consents were obtained from all participants. Ethical approval according to the Declaration of Helsinki Principles was obtained from the biomedical ethics committee of Institut Pasteur de Tunis (IRB) (2017/6F/I/Gastric Cancer-V1).

Consent for publication

Not applicable

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its Additional file.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgment

The authors are extremely grateful to the patients whose participation made this work possible, to Docteur Jean Yves Scozec and Docteur.Ludovic Lacroix from Institut Gustave Roussy for their collaboration.

Funding

This work was supported by the Tunisian Ministry of Public Health, the Tunisian Ministry of Higher Education and Scientific Research (LR16IPT05).

Conflict of interest statement

The authors declare that they have no competing interests. All authors have approved the submission of this article in its current version.

Authors' Information

Jihenne BEN AISSA-HAJ: jihennebenaissa@gmail.com/ jihenne.benaissa@pasteur.utm.tn/ jihenne.ben-aissa@gustaveroussy.fr

Maria KABBAGE: maria.kabbage79@gmail.com

Houcemeddine Othman: houcemoo@gmail.com

Patrick Saulnier: Patrick.saulnier@gustaveroussy.fr

Azer Ferah: azer.farah@gmail.com

Amira Jaballah-Gabteni: amirajaballah@gmail.com

Mouna Medhioub: Medhioub.mouna@yahoo.fr

Amel Khsiba: amal.khsiba@yahoo.fr

Afifa Maaloul: jheyna@live.com

Houda Belfkih: hbelfkih@yahoo.com

Hassen Touinsi: hassentouinsi2@yahoo.fr

Sonia Ben Nasr: sonia.bennasr_res@yahoo.fr

Lamine Hamzaoui: lamine015@yahoo.fr

Emna Chalbi: emnachelbi1@gmail.com

Sonia Abdelhak: sonia.abdelhak@pasteur.utm.tn

Samir Boubaker: boubaker.samir@yahoo.fr

Mousaddak Azzouz: mm.azzouz@rns.tn

Etienne Rouleau: etienne.rouleau@gustaveroussy.fr

Authors' Contribution

JBAH led the study and the writing of the paper.

JBAH, MK, ER, HO, and SB contributed to the experimental design and conception of the study.

Sample and clinical data acquisition were performed by MMe, AK, HB, HT, LH, EC, SBN and MA.

JBAH, MK, HO and ER contributed to data analysis and result interpretation.

JBAH, MK, HO and ER provided the Bioinformatics study analysis.

JBAH, MK, PS, AF, AJG and AMa provided the technical support for the experimental study.

Involvement in the drafting of the manuscript: MK and HO.

SA and SB provided the critical analysis of the paper.

Submission procedure: JBAH.

All authors read and approved the final manuscript.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68:394–424.

2. Rawla P, Barsouk A. Epidemiology of gastric cancer: global trends, risk factors and prevention. Przegla□d Gastroenterol. 2019;14:26–38. doi:10.5114/pg.2018.80001.

3. Van Cutsem E, Sagaert X, Topal B, Haustermans K, Prenen H. Gastric cancer. The Lancet. 2016;388:2654–64. doi:10.1016/S0140-6736(16)30354-3.

4. Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology, risk factors, screening, and prevention. Cancer Epidemiol Biomark Prev Publ Am Assoc Cancer Res Cosponsored Am Soc Prev Oncol. 2014;23:700–13.

5. Kelley JR, Duggan JM. Gastric cancer epidemiology and risk factors. J Clin Epidemiol. 2003;56:1–9.

6. Asombang AW. Gastric cancer in Africa: Current management and outcomes. World J Gastroenterol. 2014;20:3875. doi:10.3748/wjg.v20.i14.3875.

7. Crew KD, Neugut AI. Epidemiology of gastric cancer. World J Gastroenterol. 2006;12:354–62.

8. Elghali MA, Gouader A, Bouriga R, Mahjoub M, Jarrar MS, Ziadi S, et al. Gastric Adenocarcinomas in Central Tunisia: Evolution Specificities through Two Decades and Relation with Helicobacter pylori. Oncology. 2018;95:121–8.

9. Lauren P. THE TWO HISTOLOGICAL MAIN TYPES OF GASTRIC CARCINOMA: DIFFUSE AND SO-CALLED INTESTINAL-TYPE CARCINOMA. AN ATTEMPT AT A HISTO-CLINICAL CLASSIFICATION. Acta Pathol Microbiol Scand. 1965;64:31–49.

10. Laurén PA, Nevalainen TJ. Epidemiology of intestinal and diffuse types of gastric carcinoma a time-trend study in finland with comparison between studies from high-and low-risk areas. Cancer. 1993;71:2926–33.

11. MA J, SHEN H, KAPESA L, ZENG S. Lauren classification and individualized chemotherapy in gastric cancer. Oncol Lett. 2016;11:2959–64. doi:10.3892/ol.2016.4337.

12. 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. doi:10.1038/32918.

13. 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.

14. Park J-G, Yang H-K, Kim WH, Caldas C, Yokota J, Guilford PJ. Report on the first meeting of the International Collaborative Group on Hereditary Gastric Cancer. J Natl Cancer Inst. 2000;92:1781–2.

15. Fitzgerald RC, Hardwick R, Huntsman D, Carneiro F, Guilford P, Blair V, et al. International Gastric Cancer Linkage Consortium: Hereditary diffuse gastric cancer: updated consensus guidelines for clinical management and directions for future research. J Med Genet. 2010;47:436–44.

16. 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. J Med Genet. 2015;52:361–74.

17. 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. doi:10.1001/jamaoncol.2014.168.

18. Pharoah PD, Guilford P, Caldas C, International Gastric Cancer Linkage Consortium. Incidence of gastric cancer and breast cancer in CDH1 (E-cadherin) mutation carriers from hereditary diffuse gastric cancer families. Gastroenterology. 2001;121:1348–53.

19. 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. J Med Genet. 2015;52:361–74. doi:10.1136/jmedgenet-2015-103094.

20. Seevaratnam R, Coburn N, Cardoso R, Dixon M, Bocicariu A, Helyer L. A systematic review of the indications for genetic testing and prophylactic gastrectomy among patients with hereditary diffuse gastric cancer. Gastric Cancer Off J Int Gastric Cancer Assoc Jpn Gastric Cancer Assoc. 2012;15 Suppl 1:S153-163.

21. Stemmler MP. Cadherins in development and cancer. Mol Biosyst. 2008;4:835–50.

22. Van Roy F, Berx G. The cell-cell adhesion molecule E-cadherin. Cell Mol Life Sci. 2008;65:3756–88.

23. Majewski IJ, Kluijt I, Cats A, Scerri TS, Jong D de, Kluin RJ, et al. An α-E-catenin (CTNNA1) mutation in hereditary diffuse gastric cancer. J Pathol. 2013;229:621–9.

24. Clark DF, Michalski ST, Tondon R, Nehoray B, Ebrahimzadeh J, Hughes SK, et al. Loss-of-function variants in CTNNA1 detected on multigene panel testing in individuals with gastric or breast cancer. Genet Med. 2020;22:840–6.

25. Benusiglio PR, Colas C, Guillerm E, Canard A, Delhomelle H, Warcoin M, et al. Clinical implications of CTNNA1 germline mutations in asymptomatic carriers. Gastric Cancer. 2019;22:899–903. doi:10.1007/s10120-018-00907-7.

26. Marwitz T, Hüneburg R, Spier I, Lau J-F, Kristiansen G, Lingohr P, et al. Hereditary Diffuse Gastric Cancer: A Comparative Cohort Study According to Pathogenic Variant Status. Cancers. 2020;12:3726.

27. Dunnen JT den, Dalgleish R, Maglott DR, Hart RK, Greenblatt MS, McGowan-Jordan J, et al. HGVS Recommendations for the Description of Sequence Variants: 2016 Update. Hum Mutat. 2016;37:564–9. doi:10.1002/humu.22981.

28. 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. doi:10.1002/humu.23650.

29. 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 Off J Am Coll Med Genet. 2015;17:405–24. doi:10.1038/gim.2015.30.

30. Sali A, Blundell TL. Comparative protein modelling by satisfaction of spatial restraints. J Mol Biol. 1993;234:779–815.

31. Raveh B, London N, Schueler-Furman O. Sub-angstrom modeling of complexes between flexible peptides and globular proteins. Proteins Struct Funct Bioinforma. 2010;78:2029–40. doi:https://doi.org/10.1002/prot.22716.

32. Kortemme T, Kim DE, Baker D. Computational Alanine Scanning of Protein-Protein Interfaces. Sci STKE. 2004;2004:pl2–pl2. doi:10.1126/stke.2192004pl2.

33. McGibbon RT, Beauchamp KA, Harrigan MP, Klein C, Swails JM, Hernández CX, et al. MDTraj: A Modern Open Library for the Analysis of Molecular Dynamics Trajectories. Biophys J. 2015;109:1528–32. doi:10.1016/j.bpj.2015.08.015.

34. Leman R, Gaildrat P, Le Gac G, Ka C, Fichou Y, Audrezet M-P, et al. Novel diagnostic tool for prediction of variant spliceogenicity derived from a set of 395 combined in silico/in vitro studies: an international collaborative effort. Nucleic Acids Res. 2018;46:7913–23.

35. Wagih O, Galardini M, Busby BP, Memon D, Typas A, Beltrao P. A resource of variant effect predictions of single nucleotide variants in model organisms. Mol Syst Biol. 2018;14. doi:10.15252/msb.20188430.

36. Momozawa Y, Iwasaki Y, Parsons MT, Kamatani Y, Takahashi A, Tamura C, et al. Germline pathogenic variants of 11 breast cancer genes in 7,051 Japanese patients and 11,241 controls. Nat Commun. 2018;9:4083.

37. Schouten JP, McElgunn CJ, Waaijer R, Zwijnenburg D, Diepvens F, Pals G. Relative quantification of 40 nucleic acid sequences by multiplex ligation-dependent probe amplification. Nucleic Acids Res. 2002;30:e57–e57. doi:10.1093/nar/gnf056.

38. Ishiyama N, Lee S-H, Liu S, Li G-Y, Smith MJ, Reichardt LF, et al. Dynamic and Static Interactions between p120 Catenin and E-Cadherin Regulate the Stability of Cell-Cell Adhesion. Cell. 2010;141:117–28. doi:10.1016/j.cell.2010.01.017.

39. Benusiglio PR, Colas C, Rouleau E, Uhrhammer N, Romero P, Remenieras A, et al. Hereditary diffuse gastric cancer syndrome: improved performances of the 2015 testing criteria for the identification of probands with a CDH1 germline mutation. J Med Genet. 2015;52:563–5.

40. van der Post RS, Vogelaar IP, Manders P, van der Kolk LE, Cats A, van Hest LP, et al. Accuracy of Hereditary Diffuse Gastric Cancer Testing Criteria and Outcomes in Patients With a Germline Mutation in CDH1. Gastroenterology. 2015;149:897-906.e19.

41. Grünberg R, Leckner J, Nilges M. Complementarity of Structure Ensembles in Protein-Protein Binding. Structure. 2004;12:2125–36. doi:10.1016/j.str.2004.09.014.

42. Paul F, Weikl TR. How to Distinguish Conformational Selection and Induced Fit Based on Chemical Relaxation Rates. PLOS Comput Biol. 2016;12:e1005067. doi:10.1371/journal.pcbi.1005067.

43. Ishiyama N, Lee S-H, Liu S, Li G-Y, Smith MJ, Reichardt LF, et al. Dynamic and Static Interactions between p120 Catenin and E-Cadherin Regulate the Stability of Cell-Cell Adhesion. Cell. 2010;141:117–28. doi:10.1016/j.cell.2010.01.017.

44. Hanahan D, Weinberg RA. The Hallmarks of Cancer. Cell. 2000;100:57–70.

45. Mosesson Y, Mills GB, Yarden Y. Derailed endocytosis: an emerging feature of cancer. Nat Rev Cancer. 2008;8:835–50.

46. Fujita Y, Krause G, Scheffner M, Zechner D, Leddy HEM, Behrens J, et al. Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. Nat Cell Biol. 2002;4:222–31.

47. Marambaud P, Shioi J, Serban G, Georgakopoulos A, Sarner S, Nagy V, et al. A presenilin-1/gamma-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions. EMBO J. 2002;21:1948–56.

48. The Role of E-Cadherin-Catenin Complex in Prostate Cancer Progression | IntechOpen. https://www.intechopen.com/books/advances-in-prostate-cancer/the-role-of-e-cadherin-catenin-complex-

in-prostate-cancer-progression. Accessed 7 Jun 2021.

49. Tian X, Liu Z, Niu B, Zhang J, Tan TK, Lee SR, et al. E-Cadherin/β-Catenin Complex and the Epithelial Barrier. J Biomed Biotechnol. 2011;2011:e567305. doi:10.1155/2011/567305.

50. Marambaud P, Shioi J, Serban G, Georgakopoulos A, Sarner S, Nagy V, et al. A presenilin-1/gamma-secretase cleavage releases the E-cadherin intracellular domain and regulates disassembly of adherens junctions. EMBO J. 2002;21:1948–56.

51. Oliveira C, Senz J, Kaurah P, Pinheiro H, Sanges R, Haegert A, et al. Germline CDH1 deletions in hereditary diffuse gastric cancer families. Hum Mol Genet. 2009;18:1545–55. doi:10.1093/hmg/ddp046.

52. Paredes J, Figueiredo J, Albergaria A, Oliveira P, Carvalho J, Ribeiro AS, et al. Epithelial E- and P-cadherins: Role and clinical significance in cancer. Biochim Biophys Acta BBA - Rev Cancer. 2012;1826:297–311.

53. Ehsan A, Mahmood K, Khan YD, Khan SA, Chou K-C. A Novel Modeling in Mathematical Biology for Classification of Signal Peptides. Sci Rep. 2018;8:1039.

54. Halic M, Becker T, Pool MR, Spahn CMT, Grassucci RA, Frank J, et al. Structure of the signal recognition particle interacting with the elongation-arrested ribosome. Nature. 2004;427:808–14.

55. Janda CY, Li J, Oubridge C, Hernández H, Robinson CV, Nagai K. Recognition of a signal peptide by the signal recognition particle. Nature. 2010;465:507–10.

56. Figueiredo J, Melo S, Gamet K, Godwin T, Seixas S, Sanches JM, et al. E-cadherin signal sequence disruption: a novel mechanism underlying hereditary cancer. Mol Cancer. 2018;17:112.

57. Blair VR, McLeod M, Carneiro F, Coit DG, D'Addario JL, van Dieren JM, et al. Hereditary diffuse gastric cancer: updated clinical practice guidelines. Lancet Oncol. 2020;21:e386–97.

58. Boujemaa M, Hamdi Y, Mejri N, Romdhane L, Ghedira K, Bouaziz H, et al. Germline copy number variations in BRCA1/2 negative families: Role in the molecular etiology of hereditary breast cancer in Tunisia. PLOS ONE. 2021;16:e0245362. doi:10.1371/journal.pone.0245362.

59. Hamdi Y, Boujemaa M, Ben Rekaya M, Ben Hamda C, Mighri N, El Benna H, et al. Family specific genetic predisposition to breast cancer: results from Tunisian whole exome sequenced breast cancer cases. J Transl Med. 2018;16:158. doi:10.1186/s12967-018-1504-9.

60. Jaballah-Gabteni A, Tounsi H, Kabbage M, Hamdi Y, Elouej S, Ben Ayed I, et al. Identification of novel pathogenic MSH2 mutation and new DNA repair genes variants: investigation of a Tunisian Lynch syndrome family with discordant twins. J Transl Med. 2019;17:212. doi:10.1186/s12967-019-1961-9.

61. Mighri N, Hamdi Y, Boujemaa M, Othman H, Ben Nasr S, El Benna H, et al. Identification of Novel BRCA1 and RAD50 Mutations Associated With Breast Cancer Predisposition in Tunisian Patients. Front Genet. 2020;11. doi:10.3389/fgene.2020.552971.

Tables

Table 1: Clinical-pathological Characteristics of 34 selected Patients.

	Total	
	N	%
Total	34	
Gender		
Male	13	38,24
Female	21	61,76
Age at diagnosis		
<=40	14	41,18
>40	20	58,82
Tumour subtype		
Diffuse	34	100
HP Status		
Present	17	50
Absent	7	20,59
NI	7	20,59
Criteria IGCLC 2015*		
None	12	35,29
1	2	5,88
2	15	44,12
3	5	14,71
Stage		
NI	2	5,88
1	14	41,18
2	4	11,76
3	9	26,47
4	5	14,71

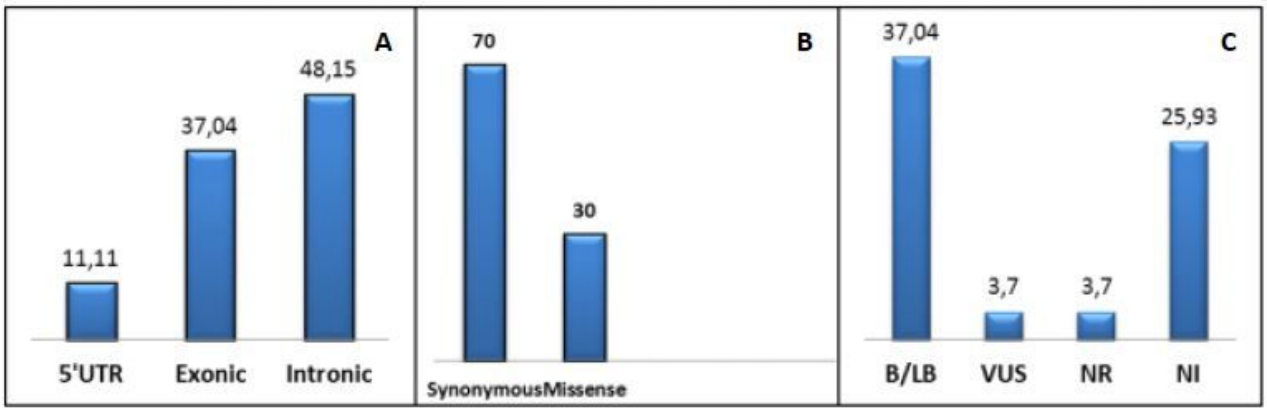
*(1) Two or more GC cases regardless of age, at least one confirmed DGC, in first-degree and second-degree relatives (2)One case of DGC before 40 years old (3) Personal or familial history of DGC and LBC, one diagnosed before 50 years old.

Table 2: Characteristics of the predicted deleterious variants

CDH1 gene		Exon 1-2	Intron 10	Exon 14
Zoom in gene region		chr16: 67,325,572-67,239,733	c.1565+3_1565+4delinsGT	c.2281G>A
Method of identification		MLPA assay	Sanger Sequencing	S S Sanger Sequencing
Type of mutation		Deletion of 4,161pb	Indel variant	Missense variant
Variant's reference		Novel	Novel	rs779648243
Clinvar classification		NR	NR	VUS
Index Case		JI-020	JI-014	JI-007
Clinicopathogical caracteristics of the patient	**Age at diagnosis/ Sex**	79/F	42/F	26/M
	TNM	T3N2M0	T4N1M1	T4N0M1
	Localisation	NI	AP	F
	Personal history	DGC	DGC	DGC
	Familial history	CCR-BC	GC	No history
	IGCLC 2015	3	2	2
E-Cadherin Expression		NA	Heterogeneous Loss	Homogenous Loss
Protein change		-	NA	p.G761R
Classification		PD	PD	PD

F:Female, M: Male, VUS : Variant of Uncertain Significance;;NR : Not reported ;NA : Non Applicable; PD:Probably Deleterious, BC: Breast Cancer; CCR: Colorectal Cancer; GC: Gastric Cancer, F:Fundic; AP: AntroPyloric

Figures

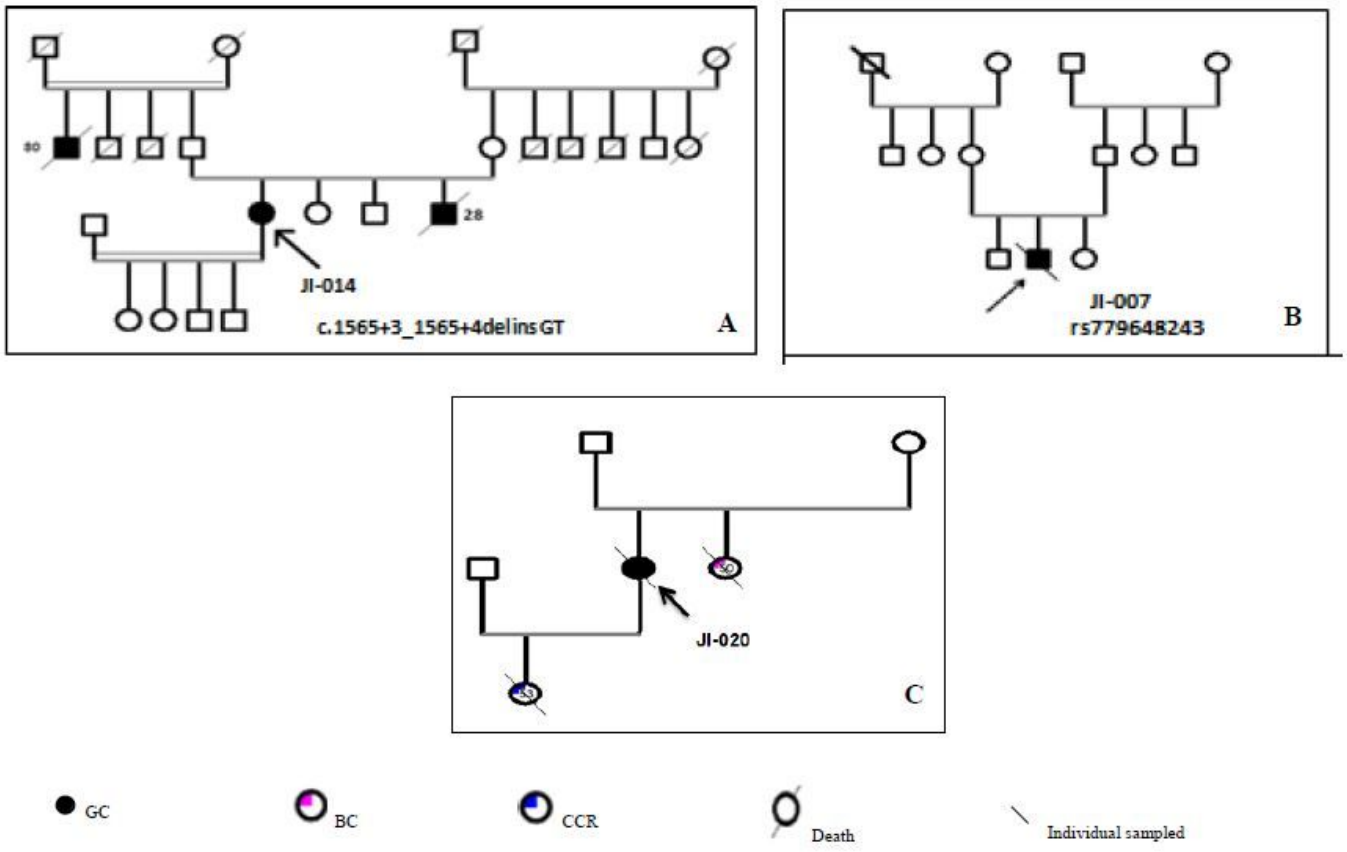


(A) SNV/Indel distribution **(B)** Distribution of coding variants **(C)** Clinvar classification,

5'UTR: 5' Untranslated Region; B/LB: Benign/Likely Benign; VUS: Variant of Uncertain Significance, NR: Not reported; NI: Non Indicated

Figure 1

Percentage of distribution of 27 CDH1 identified variants:



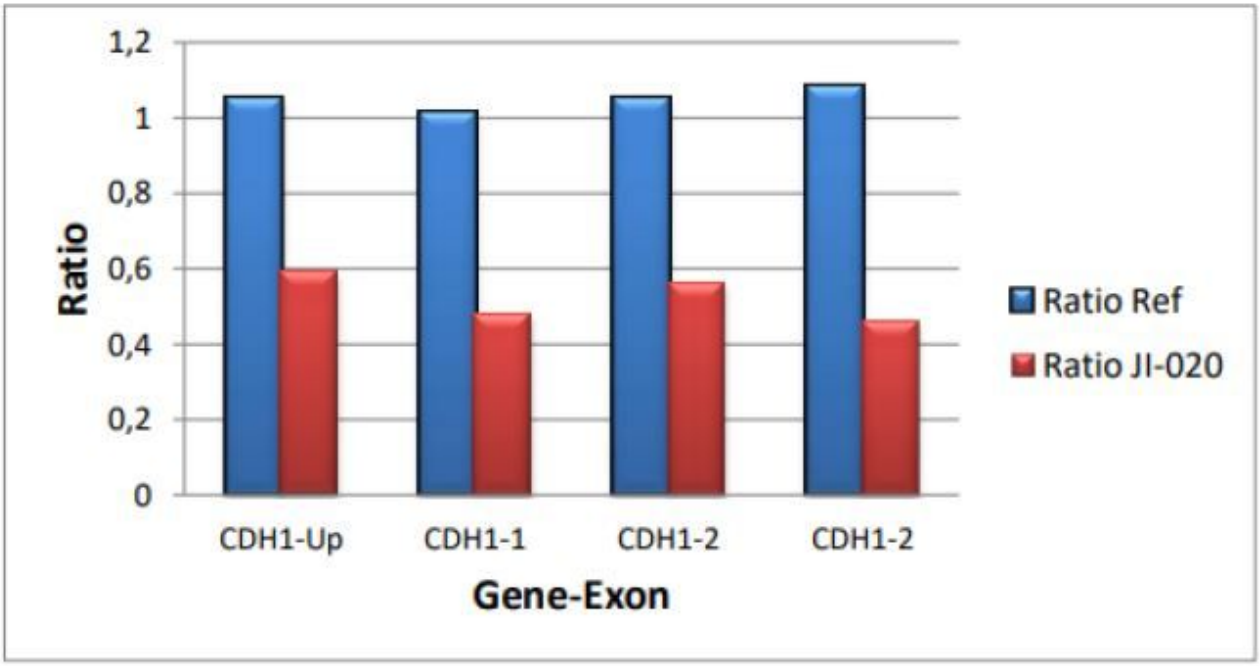
CG: Gastric Carcinoma, BC: Breast Carcinoma, CCR: Colorectal Carcinoma

Figure 2

Family History of index cases carrying selected variants.

Figure 3

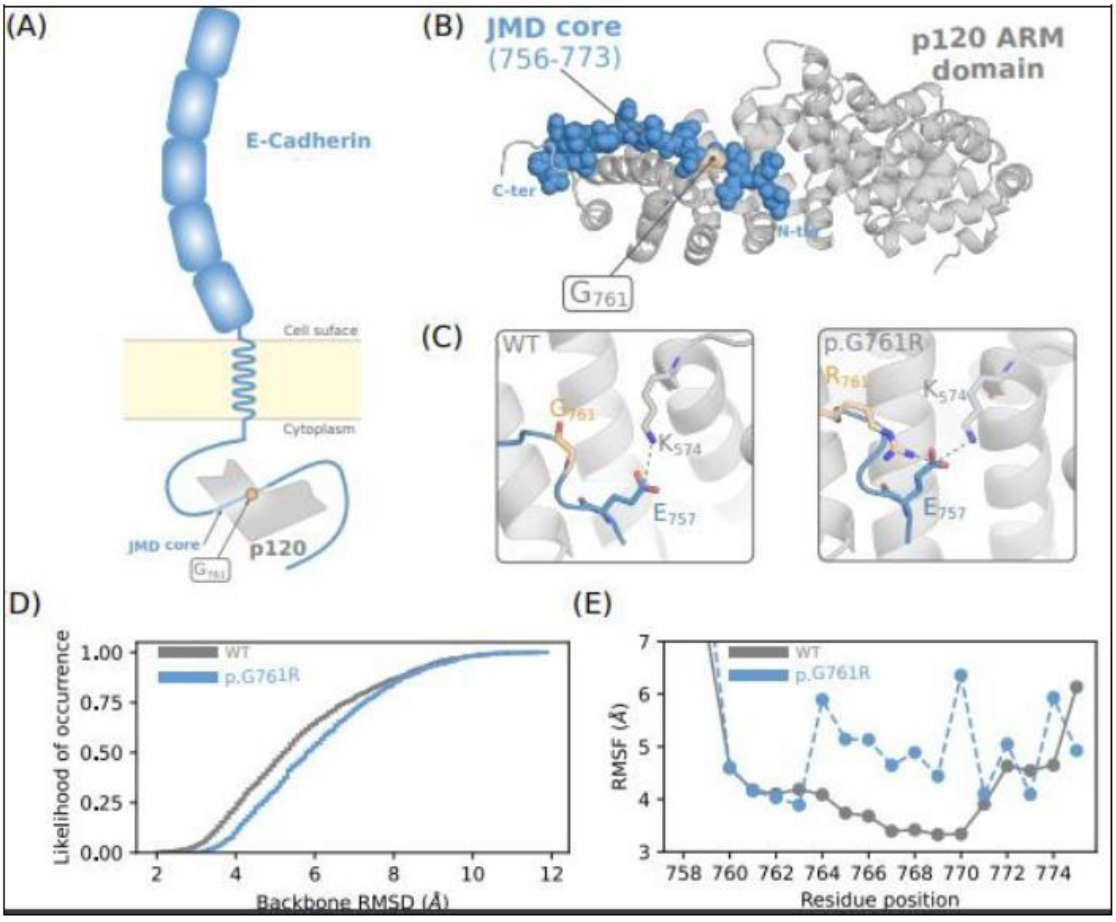
Indel effect for the index case “JI-014”as predicted by Alamut Visual Interactive Biosoftware



(Bleu) Control probes / (Red) Index cases harboring Exons deletions (*CDH1* Exons 1 and 2).

Figure 4

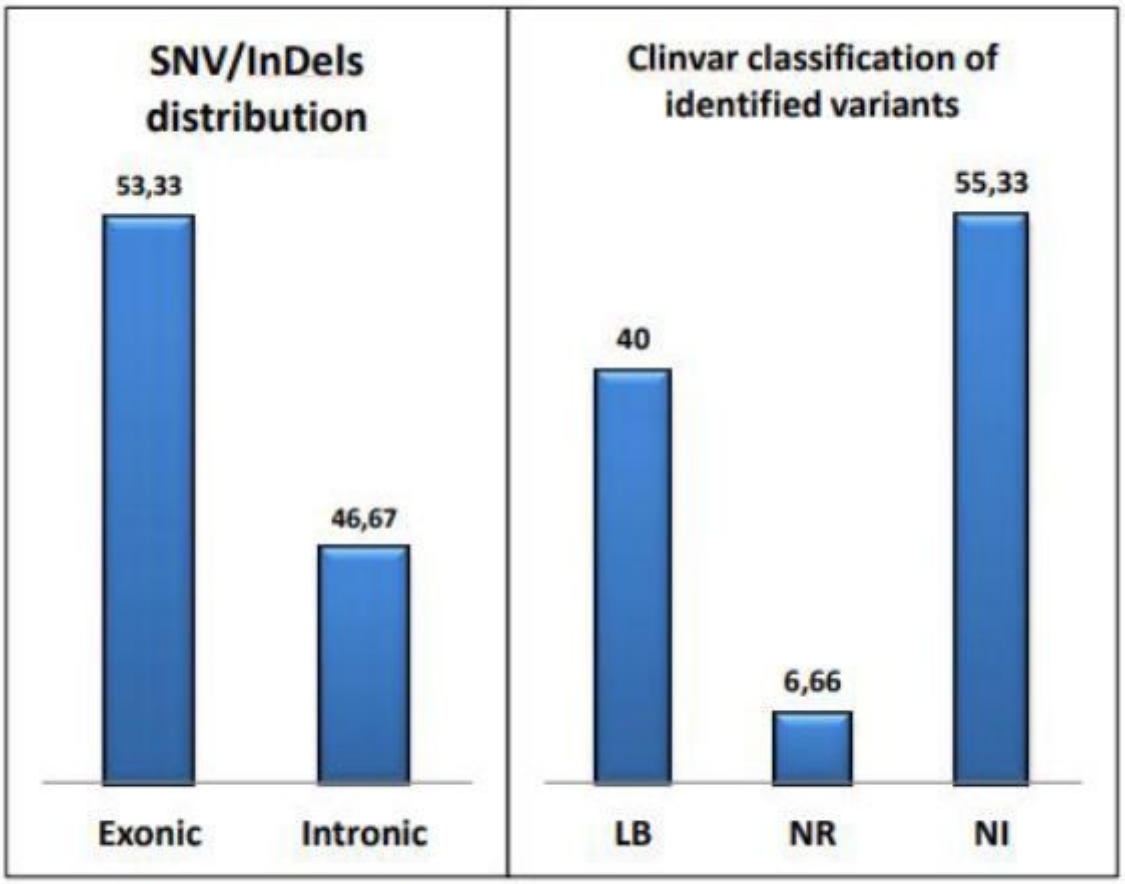
Detection of CDH1 exon deletions by MLPA assay.



(A) Schematic representation of the E-cadherin/p120 complex that includes the JMD core and the position of the mutation. **(B)** Co-crystal structure of the JMD core with p120 ARM domain showing the position of the mutated residue (light orange). **(C)** Interaction of G761 and R761 with the nearby amino acids in the WT form and the mutant form respectively. **(D)** Cumulative likelihood of occurrence as a function of the backbone RMSD of the JMD core. All the structures of the ensembles were first fitted to the bound conformation of JMD core prior to the calculation of the RMSD. **(E)** Root Mean Square Fluctuation (RMSF) profiles of the JMD core residues calculated for the WT and the mutant forms.

Figure 5

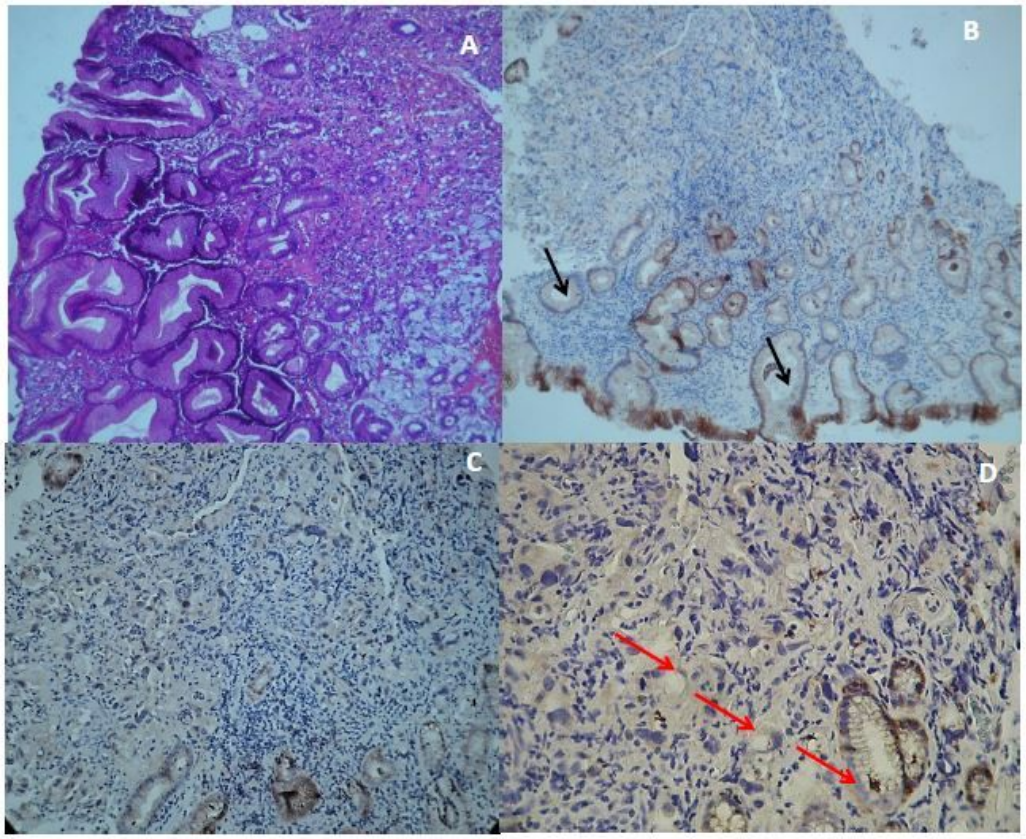
In silico analysis of the mutation c.2281G>A on E-Cadherin.



B Benign; NR: Not reported; NI: Non Indicated

Figure 6

Distribution of 15 identified CTNNA1 variants.



A) H&E staining of JI-014 tumor tissue (X100)
B) E-Cadherin immunostaining expression in gastric tumor tissue (X100)
Black Arrow shows normal membranous E-cadherin staining in crypte and glandular cells
C) Loss of membranous E-cadherin expression in tumor cells (X200)
D) Red Arrow shows a loss/reduced of E-cadherin expression in tumor cells and residual glands (X400)

Figure 7

E-Cadherin expression status in tumor gastric tissue.

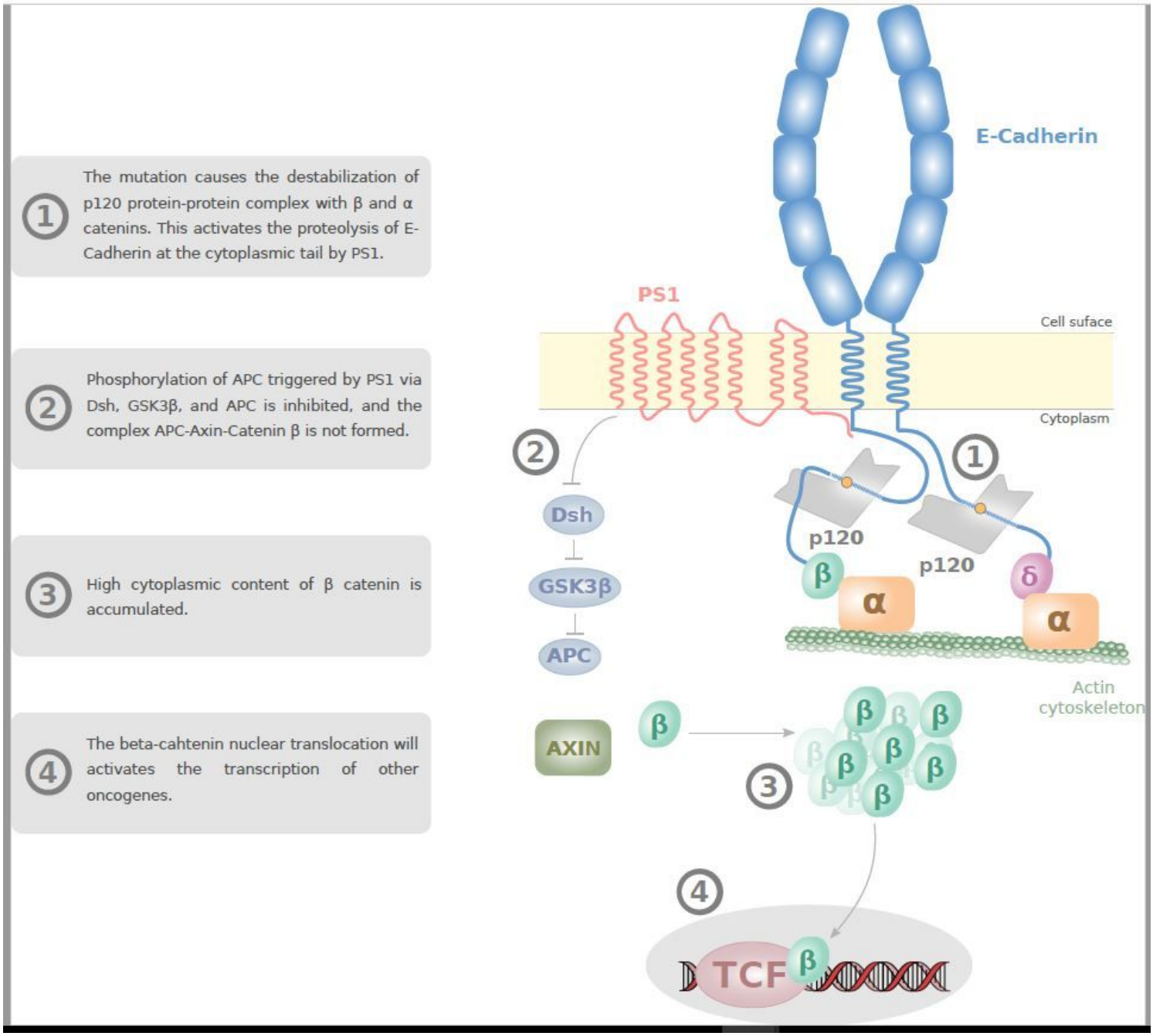


Figure 8

E-Cadherin/Beta-Catenin signaling pathway alteration in the presence of p.G761R (Inspired from [48, 49]).

Supplementary Files

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

- AdditionalFile1.xls
- AdditionalFile2.xls
- AdditionalFile3.xls

- AdditionalFile4.xls
- AdditionalFile5.pdf
- AdditionalFile6.xls
- AdditionalFile7.xls
- AdditionalFile8.xls
- AdditionalFile9.pdf