A rare MSH2 variant as a candidate marker for Lynch Syndrome II screening in Tunisia: A case of Diffuse Gastric Carcinoma

Maria Kabbage (✉ maria.kabbage@pasteur.utm.tn)

Laboratory of Biomedical Genomics and Oncogenetics. Institut Pasteur de Tunis

<https://orcid.org/0000-0001-5653-2936>

Jihenne Ben Aissa-Haj

Institut Pasteur de Tunis

Houcemedine Othman

Sydney Brenner Institute for Molecular Bioscience

Amira Jaballah-Gabteni

Institut Pasteur de Tunis

Sarra Laarayedh

Institut Pasteur de Tunis

Sahar Elouej

Aix Marseille University, Marseille Medical Genetics

Mouna Medhioub

Mohamed Taher Maamouri Hospital de Nabeul

Haifa Tounsi

Institut Pasteur de Tunis

Amal Khessiba

Mohamed Taher Maamouri Hospital de Nabeul

Moufida Mahmoudi

Hospital Mohamed Taher Maamouri, Nabeul

Houda BelFekih

Hospital Mohamed Taher Maamouri, Nabeul

Afifa Maaloul

Institut Pasteur de Tunis

Hassen Touinsi

Hospital Mohamed Taher Maamouri, Nabeul

Lamine Hamzaoui

Hospital Mohamed Taher Maamouri, Nabeul

Emna Chelbi

Hospital Mohamed Taher Maamouri, Nabeul

Sonia Abdelhak

Institut Pasteur de Tunis

Samir Boubaker

Institut Pasteur de Tunis

Moussadak Azzouz

Jamal Mohamed College

Research article

Keywords: HDGC, CDH1, Lynch Syndrome II, Genetic screening, Target gene Sequencing, DNA Repair genes, MSH2 and Tunisian Family

Posted Date: August 17th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-806482/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

BACKGROUND

Inherited syndromic forms of digestive cancers with age occurrence before 50 seem to be relatively in higher proportions in Tunisia suggesting genetic susceptibility. Several syndromic forms are known to predispose to early onset gastric tumors such as Hereditary Diffuse Gastric Cancer (HDGC) and Lynch Syndrome (LS). LS II is an extracolonic cancer syndrome characterized by a tumor spectrum including gastric cancer (GC). In the current work, our main aim was to identify the mutational spectrum underlying the genetic predisposition to diffuse gastric tumors occurring in a Tunisian family suspected with both HDGC and LS II syndromes.

PATIENTS AND METHODS

The index case was a woman diagnosed with Diffuse Gastric Carcinoma (DGC) fulfilling the international guidelines for both HDGC and LSII syndromes. Genomic DNA was used for the search of *CDH1* and *CTNNA1* germ mutations. We performed, also a DNA repair custom panel targeting candidate genes recovering the four DNA repair systems. Structural bioinformatics analysis was conducted to predict the effect of the revealed mutations.

RESULTS

No deleterious variants have been identified neither in *CDH1* nor in *CTNNA1* coding and their flanking regions. DNA repair genes panel screening identified two variants: a rare *MSH2* c.728G > A classified as a variant with uncertain significance and a novel *FANCD2* variant c.1879G > T. The structural prediction model of *MSH2* mutation and electrostatic potential calculation showed for the first time that *MSH2* c.728G > A is likely pathogenic and is involved in the MSH2-MLH1 complex stability. This mutation appears to affect the MSH2-MLH1 complex as well as DNA-complex stability. The c.1879G > T *FANCD2* variant was predicted to destabilize the protein structure.

CONCLUSIONS

Our results showed that MSH2 p.R243Q variant is likely pathogenic and is involved in the MSH2-MLH1 complex stability with a putative impact on the binding with MLH1 by disrupting the electrostatic potential suggesting the revision of its status from VUS to likely pathogenic. This variant seems to be a shared variant in Mediterranean region. These findings emphasize the importance of testing DNA repair genes for patients diagnosed with DGC with suspicion of LSII as well as colorectal cancer allowing better clinical surveillance for more personalized medicine.

Background

Gastric Carcinoma (GC) is the fifth Cancer worldwide with approximately one million new cases registered in 2018 (5.7%) with significant geographical distribution variations. It represents the third largest cause of cancer related death with 783,000 deaths worldwide in 2018 which represents 8.2 % of all cancer deaths. In Tunisia as in other countries, the sporadic form of GC is the most frequent estimated at 637 new cases per year and ranks the 7th among all diagnosed cancers [1]. Inherited syndromic forms of digestive cancers with age occurrence before the age of 50 (Gastric and Colorectal) seem to be relatively in higher proportions in Tunisian population suggesting hence a genetic susceptibility [2]. However no epidemiological data are available and incidence rates of hereditary forms are lacking. Several syndromic forms are known to predispose to gastric tumors such as Hereditary Diffuse Gastric Cancer (OMIM: 137215), Lynch Syndrome (OMIM: 120435), Peutz Jegher Syndrome (OMIM: 175200) and Li Fraumeni Syndrome (OMIM: 151623) [3].

Indeed, 3–5 % of GCs are caused by an autosomal dominant inherited mutations and familial aggregation occurs in approximately 10% of the cases [4]. Hereditary Diffuse Gastric Cancer (HDGC) syndrome is known as the most frequent autosomal dominant form of GC and is a diffuse and poorly differentiated morphologic type [5]. Approximately 40 % of HDGC cases have germ mutations in *CDH1* (E-Cadherin, OMIM: 192090) and more than 150 *CDH1* mutations have been reported in HDGC cases. Lynch Syndrome (LS) (OMIM:120435), or HNPCC (Hereditary Non Polyposis Colorectal Cancer), is the most prevalent inherited CRC susceptibility syndrome, responsible for up to 6% of all CRCs. It is characterized by point mutations and/or large rearrangements in DNA mismatch repair (MMR) genes resulting in a loss of MMR complex function and microsatellite instability (MSI). The MMR genes associated to this Syndrome are: MLH1 (OMIM:609310), MSH2 (OMIM:609309, MSH6 (OMIM:600678) and PMS2 (OMIM: 600259) [6,7].

In addition, germline deletions in the 3'end of *EPCAM* gene (OMIM: 613244), directly upstream to *MSH2*, have been found to underlie a small proportion of LS cases (< 5%) through methylation induced transcriptional silencing of *MSH2* [8].

LS has been subdivided into two subtypes LSI known as site-specific colonic cancer syndrome and LSII or extracolonic cancer syndrome (OMIM: 120435) [9]. The former subtype includes 2–3% of all CCRs and is characterized by a susceptibility to colorectal tumors in the absence of diffuse polyposis. The latest subtype is associated with a high risk of various tumors types such as Endometrium [10–12], Gastric [4, 13–15], Breast [16], Biliary tract, Urinary tract [17, 18], Small Bowel [9], Ovary[19–21], Brain and Skin Cancers [22–24].

In the context of LS II, MMR genes germ mutation carriers represent 50 to 80% lifetime risk to develop CCR, 40 to 60% risk of developing endometrial cancer (OMIM: 608089) and 13 to 19 % risk to develop gastric tumors [25]. Moreover, GC is, among LS carriers, far less common than colorectal at an average older age in western countries [13]. It's well established that MMR genes dysfunction is linked to LS susceptibility. Indeed, current data assume that MSH2, MLH1, and EpCAM mutation carriers have 13% of

cumulative risk incidence of developing gastric tumors at an age of 75 years; however MSH6 carriers have only 3% of the risk [26]. Even though GC is part of the LS tumors spectrum the risk of developing GC in LS families is unknown and surveillance strategies for GC in LS are still controversial [13]. To our best knowledge, very few studies have investigated the molecular basis and the genetic mutational profile underlying this form of GC. In Tunisia, some molecular and epidemiological based studies have been reported in colorectal cases in LS families [2, 27, 28]. However no study has investigated the molecular basis of GC cases in LS. Hence, it is of crucial importance to set up oncogenetic counseling based on specific genetic tests adapted to Tunisian population. This will help in early detection of individuals and families at high risk of developing LSII and will consequently reduce the mortality and morbidity due to the disease.

Our main objective in the current work was to identify the genetic mutational profile underlying the DGC occurring in a Tunisian family suspected with both HDGC and LSII by using classical sequencing and target NGS approaches. We performed a custom panel targeting 87 candidate genes recovering the four DNA repair systems (MMR, BER, NER and HR) known to be involved in DNA repair disorders. The proband diagnosed with a DGC was screened for a repair genes panel after HDGC syndrome exclusion.

Materials And Methods

Patient

This study was conducted according to the declaration of Helsinki and with the approval of the Institutional review board (IRB) of Institut Pasteur de Tunis. Four individuals, belonging to the same large Tunisian family suspected with LSII according to revised Bethesda Guidelines, were investigated after participants wrote informed consent (Fig. 1). The index case has been investigated first. She was a woman “JI-021” from a consanguineous marriage, fulfilling the first and fifth criteria of the Revised Bethesda Guidelines and the third criteria of IGCLC with no alcohol, smoking consumption habits, spicy foods with low consumption of meats. She was referred first for *CDH1* germline genetic testing to the Gastroenterology Department of Mohamed Tahar Maamouri Medical Hospital in Nabeul, Tunisia for HDGC suspicion since she was diagnosed with DG Adenocarcinoma (T3NxM1) at the antrum at an age of 52. This index case had an aggressive phenotype since diffuse gastric tumors are known to be an aggressive gastric tumor histotype. This patient had peritoneal carcinosis and she was therefore not operable and the only treatment she had is a palliative chemotherapy. This index case has got a healthy 23-years-old daughter, three unaffected brothers (47, 53 and 55 years old) and two affected sisters (29, 47 years old). Her youngest sister “JI-021-E”, suffered from primary sterility, developed successively ovarian and gastric carcinomas and was treated by hysterectomy and a total Gastrectomy and died at the same age. The other sister “JI-021-D” developed a BC and she died at the same age. The index case’s father “JI-021-H” suffered from Renal Failure and diabetes and died at the age of 65. In addition, this family seems to have a strong history of CCR from the paternal side (five cases) without clear information regarding their relationship with the index case. He has got a brother and a sister who died respectively at the age of 50 (Probably from CCR) “JI-021-J” and 32 years old from unknown cause and a third sister “JI-021-F”

who developed a CCR before the age of 50 and died at the age of 65. Her daughter “JI-021-G” also developed a CCR at the age of 50. Table 1 shows all the Clinico-pathological features of the index case.

Table 1
Description of clinicopathological features and family history of the index case.

Case	JI-021
Diagnosis Age/Sex	53/F
Family History	Colorectal / Ovarian / Breast/ Gastric Carcinomas
First degree relatives	
Second degree relatives	
IGCLC criteria *	Third
Revised Bethesda Guidelines +	First, fifth
HP	Yes
Lauren Classification	Diffuse
Location	Antrum
TNM	T3NxM1
Survival	Died after three years (56 years old)
* (1) Colorectal or uterine cancer diagnosed in a patient how is less than 50 years of age (2) Presence of synchronous, metachronous colorectal, or other HNPCC-associated tumors, regardless of age (3) Colorectal cancer with the MSI-H ** histology diagnosed in a patient who is less than 60 years of age. (4) Colorectal cancer diagnosed in one or more first-degree relatives with an HNPCC-related tumor, with one of the cancers being diagnosed under age 50 years. (5) Colorectal cancer diagnosed in two or more first- or second-degree relatives with HNPCC-related tumors, regardless of age.	

NB: Hereditary nonpolyposis colorectal cancer (HNPCC)-related tumors include colorectal, endometrial, stomach, ovarian, pancreas, ureter and renal pelvis, biliary tract, and brain (usually glioblastoma as seen in Turcot syndrome) tumors, sebaceous gland adenomas and keratoacanthomas in Muir-Torre syndrome, and carcinoma of the small bowel.

+ 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.

Sanger Sequencing

CDH1 and CTNNA1 coding regions Sanger sequencing

Referring to 2015 IGCLC [29] for the search of *CDH1* and *CTNNA1* [30] germ mutations, a complete sequencing of these two genes was carried out. Primers were designed using Primer Express [31]

(Additional File 1). Forward and reverse primers incorporated the extensions 18F tail (ACCGTTAGTTAGCGATTT) and 18R tail (CGGATAGCAAGCTCGT) respectively at their 5' ends. Then Generated Data were analyzed using SeqScape Version 3.0 and BioEdit Sequence Alignment Editor Version 7.2.

MSH2 Sanger Sequencing

We designed two pairs of primers (Forward primer sequence 5'AGAGGAGGAATTCTGATCAC3'), (Reverse primer sequence 5'CTGATTCTCCATTTCTGGC3') to amplify the following region: *MSH2*: exon 4 (NM_000251.2) (Additional File 1). PCR reactions were performed on genomic DNAs (gDNAs), following standard protocols, pursued by Sanger sequencing using an automated sequencer (ABI 3500; Applied Biosystems, Foster City, CA) using a cycle sequencing reaction kit (Big Dye Terminator kit, Applied Biosystems). Data were analyzed by BioEdit Sequence Alignment Editor Version 7.2.5.

Large CDH1 deletions/duplications analysis by multiplex ligation-dependent probe amplification (MLPA) assay

The index case's gDNA was tested for copy number changes using Multiplex Ligation Dependent Probe Amplification (MLPA). It was performed using the SALSA P083-D2 *CDH1* MLPA kit (MRC-Holland), following manufacturer's instructions. The kit contains 35 probes for the *CDH1* gene, one Upstream flanking probe and one downstream flanking probe. MLPA products were run on the ABI Prism 3730 xl Genetic Analyzer (Applied Biosystems) and analyzed with the Peak Scanner™ Software v1.0 (Applied Biosystems).

Next Generation Sequencing

Blood samples have been collected from the index case and her consent relatives and have been sampled in the gastroenterology department. gDNA was isolated from peripheral blood using Qiagen Kit and salting-out and it was used for the library preparation for NGS. HaloPlex assay incorporating molecular barcodes for high-sensitivity sequencing was used as a custom design (HaloPlexHS). Using SureDesign (Agilent Technologies Inc.), probes were generated to cover the exons and 15 bp of the surrounding intronic sequences of a total of 87 candidate genes known to be involved in DNA repair disorders as reported by a research team in our lab [32].

Amplicon libraries were prepared, from gDNA of "JI-021", using the HaloPlexHS PCR target enrichment system dedicated to Ion Torrent PGM according to the manufacturer's recommendations. Massively parallel sequencing was performed on an Ion Torrent PGM (Thermo Fisher Scientific). Raw data generated by the PGM sequencer were analyzed using the VarAft software version 2.5, which is freely available online (<http://varaft.eu/index.php>). Genome assembly and nucleotide coordinates were referenced to the GRCh37 version of the human genome.

Variants filtering strategy

We aimed herein to select pathogenic or potentially pathogenic variants associated with the development of LSII by removing neutral variants and sequencing errors. We prioritized rare functional variants (missense, nonsense, splice site variants, and indels) and excluded variants with a Minor Allele Frequency (MAF) > 0.01 in dbSNP137, and 138, in the Exome Variant Server (<http://evs.gs.washington.edu/EVS/>), 1000 Genomes Project (<http://www.1000genomes.org/>), or Exome Aggregation Consortium database (ExAC), Cambridge, MA (<http://exac.broadinstitute.org>) and the genome Aggregation Database (gnomAD) (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA398795>).

We also removed all synonymous and homozygous variants considering that LS is an autosomal dominant syndrome. All variants with a Depth < = 15 have been removed as well as variants classified as Benign and Likely Benign in ClinVar. Variants predicted with tolerated effect by more than two *in silico* prediction tools have also been removed.

Indeed, various *in silico* prediction tools have been used to assess the functional effect and pathogenicity of the selected variants 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 to 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 nonsynonymous or indel variants that are predicted to be functionally important. Mutation Taster (<http://www.mutationtaster.org/>) has been used to assess the impact of mutations on protein function and to look at effects on splicing sites, mRNA expression, MAPP-MMR (<http://mappmmr.blueankh.com/>) to accurate classification of missense variants in MMR genes, FATHMM (<http://fathmm.biocompute.org.uk/>) to predict the functional consequences of both coding variants (Non-Synonymous single nucleotide variants) and non coding variants and LRT. Variants not previously reported in healthy controls and classified as pathogenic were evaluated for sequencing depth and visually inspected using the Integrative Genomic Viewer (IGV) before validation by Sanger Sequencing.

Immunohistochemical study

The index case “JI-021” tumor gastric tissue was tested for E-cadherin, MSH2 and MLH1 protein expression profiles by IHC. Briefly, 3–4 µm tissue sections were obtained from FFPE gastric tissue and all immunohistochemistry steps were conducted as preconized in the Novolink MPolymer Detection Systems Kit (Biopole). Antigen retrieval was performed by incubation of tissue sections in a 10 mM Sodium Citrate buffer (pH 6.0) (RE7113, Leica Biosystems Newcastle Ltd) for 20 min in 95°C heated water.

E-Cadherin

Gastric tissue was incubated forward overnight with anti-E-Cadherin primary mouse monoclonal antibody (NCL-L-E-Cad, clone 36B5, Novocastra TM, Biopole diluted to 1/30) recognizing the Nt external domain of E-cadherin. The immunostaining reaction was visualized by adding the Diaminobenzidine DAB as a chromogenic substrate (Novolink TM, Biopole).

MSH2 and MLH1

Primary antibody anti-MLH1 (ES05) and anti-MSH2 (25D12) Leica Biosystems were diluted to 1:50 and were added and incubated for 60 min. The immunostaining reaction was visualized by adding the AEC as a chromogenic substrate (Novolink TM, Biopole).

Preparing the structures

The structure of two complexes

MSH2/MSH6 [33] and MSH2/MSH3 [34] were used in our computational study to analyze the effect of the mutation R243Q on MSH2. Missing segments and atoms were added using MODELLER [35]. We also used the structure of the N-terminus domain of the DNA mismatch repair protein MLH1 [36]. The mutation at position 243 of MSH2 in both complexes was performed *in silico* using FoldX [37].

Protein-protein docking

We conducted a protein-protein docking experiment using ZDOCK version 3.0 [38] using the structure of MLH1 as a ligand (the smallest partner) and the complex MSH2/MSH6 as a receptor (the largest partner). Input options were kept to their default values. We have, however, restrained the docking interface on MLH1 only to solve exposed residues of the large beta-sheet from the N-terminal domain. We generated 2000 complexes from which we took the best ten according to the Zdock scoring function. We then refined the complexes using sander from AMBER 19 [39] molecular dynamics package by running a restrained minimization consisting of 1000 steps of steepest descent (250 steps) and conjugate gradient (750 steps) algorithms. Restraints were applied to backbone atoms of the complex. A non-bonded cutoff of 12 Angstroms was applied for the refinement.

Electrostatic potential calculation

The electrostatic potential calculation [40] has been conducted using the Adaptive Poisson-Boltzmann Solver (APBS) software package. The PDB structure of a protein is first converted to the PQR format using the PDB2PQR server [41]. The PARSE force field was used to assign the partial charges to each atom and the ionization state of charged residues was assigned according to the calculation by PROPKA. The electrostatic potential calculation was effected using an implicit solvent model. The internal dielectric constants of the solute and the solvent are fixed to 2.0 and 78.45 respectively and the temperature is set to 298.15 K. The cutoff for non-bonded interactions was set to 15 Angstroms. Visualization of the electrostatic potential was generated using Chimera molecular viewer.

Stability Analysis

We estimated the free energy of folding between the wild type and the mutant structure ($\Delta\Delta G_{Wt-Mut}$) using DynaMut [42].

Results

CDH1 and CTNNA1 Sanger Sequencing

As a result, we have found no deleterious or potentially deleterious variants in the coding and flanking regions of both *CDH1* and *CTNNA1* genes. This finding suggests other candidate genes / variants predisposing to DGC.

Large CDH1 gene deletions/duplications analysis by Multiplex Ligation-Dependent Probe Amplification (MLPA) assay

Since heterozygous large genes rearrangements are hard to detect by conventional PCR-based sequencing of gDNA, we searched for possible rearrangements of the *CDH1* locus using the MLPA assay [43]. Results were processed by the Coffalyser software (MRC-Holland, Amsterdam, Holland). Our results showed no large rearrangements in the *CDH1* gene among this index case.

E-cadherin Immunohistochemistry expression profile

In order to assess the protein expression level of E-cadherin in the index case gastric tumor tissue we have performed an immunohistochemistry by using a monoclonal antibody directed against the Nt domain of E-cadherin. Our result showed a negative immunostaining expression in tumor cells compared to normal adjacent glandular cells as control. The adjacent normal gland showed a membranous normal E-cadherin expression pattern (Fig. 2A).

Selection of variants of interest detected by Targeted DNA repair genes panel

Besides IGCLC inclusion criteria, this index case also fulfilled Bethesda Guidelines inclusion criteria for LS (First, Fourth and Fifth criteria). Hence this case was a candidate for the screening of a panel of 87 candidate genes (see methods). Using SureDesign (Agilent Technologies Inc.), probes were generated to cover exons and 15 bp flanking sequences. The size of the final target region was 251.689 kpb with 33828 amplicons and the mean coverage was 99.74% of the target region. As a result, we have identified 33 variants (Additional File 2). Statistical distribution of variants is illustrated in (Fig. 3). Exonic variants represented 23 out of 33 variants. Among the 23 exonic variants 20 were novel.

Among the exonic variants, only the rare *MSH2* (rs63751455) variant corresponding to substitution c.728 G > A p.R243Q, classified in class 3 'VUS' in Clinvar, has been selected for Sanger sequencing validation and for structural model effect prediction analysis. Hence we tried to perform several *in silico* prediction analyses by using online prediction tools and databases (15 prediction tools) as well as structural modeling of MSH2 known complexes.

This mutation has been confirmed for the index case "JI-021" in an heterozygous state and not found in the index case's relatives (index case's daughter, brother and her niece). This variant has been forward searched in three other GC cases "JI-002", "JI-030" and "JI-036" belonging to unrelated families suspected with LS II and fulfilling the first criterion of Bethesda guidelines for the first case "JI-002" and

the fifth criterion for both “JI-030” and “JI-036” cases based on their family history of cancer. As a result, c.728 G > A has been found in none of the three screened index cases suspected with LSII. A search of potential linkage disequilibrium of this variant with known pathogenic variants has been performed and no information regarding such linkage disequilibrium is currently available from LDproxy Tool [44].

Based on structural analysis and DynaMut score on theoretical functional effect a novel identified variant on the *FANCD2* gene has been selected for *in silico* Prediction analysis as well.

MSH2 and MLH1 Immunohistochemistry staining in the index case tumor gastric tissue.

The immunohistochemistry analysis of MSH2 and MLH1 proteins in the index case gastric tissue has shown a positive diffuse nuclear immunostaining for MLH1 (Fig. 2B) and nuclear immunostaining of MSH2 in tumor cells and normal adjacent residual glandular cells with loss of MSH2 staining in some tumor cells (incomplete immunostaining) (Fig. 2C). Since no significant difference was noticed between mutated and wildtype models regarding the putative impact on the collective motion of the residues using MSH2/MSH6 and MSH2/MSH3 and MSH2 (monomeric protein) structures and based on MSH2/MLH1 effect, only MSH2 and MLH1 proteins have been investigated by IHC. Moreover we no longer have FFPE gastric tissue available for this case since it was a small biopsy.

Molecular Modeling of MSH2 mutation on the MSH2/MSH6/MLH1 complex

The mutation R243Q is located on the connector domain of MSH2 at a loop connecting two alpha-helices (Fig. 4A). Most of the atoms of the side chain of Arginine are highly packed against other residues forming the core of the domain while the polar guanidine group is exposed and forms a hydrogen bond with the D240 main chain. With such a configuration of R243, we thought first that the mutation could impact the conformational function of the connector domain to other parts of the protein. We then run a calculation of the normal modes for the Wild Type (WT) form (R243) and the mutant form (Q243) to investigate the putative impact on the collective motion of the residues using MSH2/MSH6 and MSH2/MSH3 and MSH2 (monomeric protein) structures, but no significant difference was noticed (data not shown). We then investigated the possibility that the connector domain is involved in protein-protein interaction with other partners. Indeed, there have been different arguments showing that the connector domain of MSH2 is involved in the interaction with a heterodimer of MLH1/PMS1[45]. We chose to run a protein-protein docking experiment to predict the ternary complex MSH2/MSH6/MLH1 (Fig. 4B). Of the 10 complexes with the best ZDOCK energy score, six show a binding mode where the connector domain interacts with MLH1 while in the docking solution with the best energy score, R243 do part of the interface.

Because the p.R243Q mutation involves a substitution of a positively charged residue with a neutral amino acid, we proceeded by calculating the electrostatic potential of MSH2 in both WT and mutant forms. At the site of the mutation, the electrostatic potential is majorly electronegative (Fig. 5A). We noticed the occurrence of some neutral surface patches and a slightly electropositive region at the exact

position of R243 in the WT form. In the mutant form, however, the electronegativity increases. We verified that the electrostatic potential at the site of the mutation is compatible with the interaction interface of MLH1 by calculating the electrostatic potential of the latter protein. Indeed, we found that the protein-protein interface is exclusively electropositive in MLH1 (Fig. 5B).

Molecular modeling of the FANCD2 variant

Among the other putative deleterious mutations, we have identified a novel p.A627S variant on FANCD2. FANCD2 forms a protein-protein semi-symmetrical complex with FANCI which is involved in the physiological mechanism of Fanconi anemia cancer. The substitution is located at the N-terminal end of an alpha helix of the FANCD2 NTD domain. The mutation p.A627S is located near a monoubiquitination site on FANCD2 K561 residue (Additional File 3). Ubiquitination can activate the Fanconi anemia DNA repair pathway [46]. This variant induces a significant destabilization of the folding energy calculated using DynaMut as well as a perturbation of the local flexibility of the structure [47]. The $\Delta\Delta G$ calculated by DynaMut shows an unfavorable energy of

-1kcal/mol. Moreover, four other tools including EnCoM, mCSM, SDM and DUET show also an unfavorable $\Delta\Delta G$ of -1.8, -2.9 and - 2.1 respectively. Therefore all the tools are in concordance about the destabilizing effect of the p.A627S variant.

Discussion

Currently, screening for serious diseases such as cancer represents one of the priorities of public health and therefore of medical research. Indeed, thanks to the development of highly throughput technologies, particularly in the field of genomics, it became easier to elucidate the molecular mechanisms underlying the development and progression of hereditary cancers. In addition such sequencing approaches contributed significantly to the identification of various genetic actionable variants allowing the development of molecular screening for early diagnosis of complex syndromic cancers particularly in under-investigated populations such as North African population.

GC the second most common extracolonic tumor in LS (HNPCC) [48, 49]. Although the risk of extracolonic tumors including gastric seems to be higher in *MSH2* compared to *MLH1* mutation carriers [50], the association between germline mutational profile and clinical phenotype is generally weak in LS [14]. Very few studies have performed molecular comprehensive investigation on patients with gastric tumors even if they are LS mutation carriers or not.

In the current study, we investigated a DGC patient “JI-021” belonging to a Tunisian family from North-East of Tunisia with suspicion of both HDGC and LSII (according to the tumor's cluster family history) shedding light on the molecular basis of GC in this region known to have relatively high proportion of digestive cancer syndromes (no epidemiological data available). Based on the International Gastric Cancer Linkage Consortium (IGCLC), this case meets the third criterion for *CDH1* germline screening. As a result, we found no pathogenic variant in both *CDH1* and *CTNNA1* genes coding and exons flanking

regions. In addition, no large rearrangements have been identified by MLPA in *CDH1* gene for this patient. This result is not surprising since the lack of *CDH1* pathogenic variant DGC patients meeting the IGCLC criteria has already been reported. Indeed, Jakubowska and colleagues have sequenced the entire coding region of *CDH1* gene in 86 Polish cancer patients from families fulfilling the criteria of HDGC and have found no deleterious mutations in this gene [51].

In fact, our study has covered the *CDH1* entire coding region (16 exons), splice junctions as well as large rearrangements analysis. Nevertheless, we need to investigate the non-coding regions particularly regulatory transcriptional regions and promotor to confirm the definitive HDGC exclusion in this case since its corresponding FFPE tumor tissue harbored a loss of E-cadherin expression in tumor cells comparing to normal adjacent gland (Fig. 2) suggesting a putative *CDH1* gene transcription defect at germinal or somatic level.

In addition to the IGCLC inclusion criteria, the index case “JI-021” also fulfilled Bethesda Guidelines inclusion criteria for LS (1th and 5th criteria). In the current study, our index case gDNA was screened via a custom panel targeting 87 DNA repair genes. Although several studies have used panels covering additional genes other than DNA repair as LS known associated genes or candidate genes, we aimed to focus on a panel covering the four DNA repair systems in order to identify new candidate DNA repair genes/variants predisposing to such cancer syndrome. As a result, we have identified the *MSH2* (c.728 G > A p.R243Q) variant in the index case “JI-021” but not in her asymptomatic relatives. According to our best knowledge, this is the first study investigating the genetic predisposition of GC case in the context of LSII in Tunisia and very few studies have reported such an investigation worldwide [13, 14].

Indeed, Capelle and colleagues have reported 2014 mutations identified in the MMR genes as part of LS in 236 Dutch families and among them GC was diagnosed in 32 subjects (1.6%), including 22 (69%) with family history of GC. The risk of developing GC was rated at 4.8% for carriers of the *MLH1* gene mutation and at 9% for carriers of the *MSH2* mutation. However, among 378 *MSH6* identified mutations, no carrier had GC [13].

Nevertheless with the emergence of precision oncology and development of germline panels, several studies have investigated LS susceptibility genes in individuals with cancers and individuals at risk to develop hereditary cancers. In 1990, a study was conducted in a series of 24 HNPCC families from Holland. In four families, CRC was the only type of cancer to occur suggesting a LSI type. Sixty-five extracolonic tumors were observed in the other 20 families. Carcinoma of the endometrium was present in 16 patients from 12 families; GC was present in 10 patients from five families besides other tumor types. They concluded that carcinoma of the endometrium, stomach, and urinary tracts were integral to the hereditary tumor spectrum of HNPCC [52]. These authors identified germline mutations in 11 (35.5%) families (six *MSH2*, five *MLH1*, including seven novel mutations), however only seven of which fulfilled the Amsterdam criteria. All these findings highlight that LS is histologically and genetically a complex disease with heterogeneous molecular background. Indeed, it seems that different families with the same LS inclusion criteria could have different mutational profiles in newly identified susceptibility genes not

initially known to be associated with LS. These identified mutations / variants appear to have low frequencies among screened cases and families. These reported results are in concordance with our finding since our *MSH2* c.728G > A identified mutation has been found only in one case among the four LSII suspected unrelated families. While several other works have investigated the molecular background of CCR in the context of LS worldwide [53–60], only three studies have been conducted in Tunisia in this field. Indeed, Moussa and colleagues have reported germline mutations in MMR genes in 11/31 families suspected for LS (six *MSH2* and five *MLH1* mutations). Among the six *MSH2* identified variants, our *MSH2* c.728G > A mutation has been found in only two probands (52 and 38 years old) belonging to two families sharing the fifth Bethesda guidelines selection criterion and the same mutational profile as well as a common newly identified *MSH2* c.1413dupA mutation. These authors reported that *MSH2* c.728G > A (p.R243Q) mutation was predicted as probably non-pathogenic since the substituted Arginine residue was relatively little conserved across phylogeny with weak physicochemical difference between Arginine and Glutamine. However, this claim does not consider the complexity of domain-domain cooperatively, the protein-protein interactions and the fact that Arg residues are positively charged and Gln is neutral. They have found this variant in two families, the first one with five cases of CCR, the second with two cases of CCR. Both families shared the fifth Bethesda Guidelines criterion and two mutations “*MSH2* c.728G > A” and “c.1413dupA”. Compared to our results, the proband “JI-021” carrying the same *MSH2* c.728G > A (p.R243Q) mutation shared the same age of GC onset (52) and the same Bethesda criteria (1st and 5th) with the two lastly reported Tunisian families. These homologies suggest that *MSH2* c.728G > A (p.R243Q) mutation might be a LS / LSII predisposing mutation in Tunisian population however supplementary molecular investigation among a larger cohort is needed to confirm this hypothesis. Intriguingly, Moussa and colleagues noted that the age of cancer onset as well as the tumors spectrum were similar in families with and without MMR germ mutations which is in contradiction with other studies conducted in other populations [2]. More recently, NGS of MMR genes, *POLE* and *POLD1* was performed, by Ben Sghaier and colleagues, to identify the genetic mechanisms underlying CCR in 24 Tunisian probands. As results, they identified in six cases five germline variants in *MLH1*, a somatic pathogenic variant in *MSH2* (c.2557G > T) and a germline variant in *MSH2* (c.1413 dupA) [27]. In addition, a team in our lab have identified in a LS family with discordant twins, an *MSH2* pathogenic mutation (c.1552C > T;p.Q518X) previously reported only in a Portuguese LS family and shared with all investigated family members suggesting that this variant is a causal mutation [28]. Furthermore, an Algerian recent study have reported the *MSH2* (c.728G > A) mutation in two cases (44 and 38 years old) with CCR belonging to the same family [61]. *MSH2* (c.728G > A) variant seems to be identified only in Mediterranean countries. These findings support the fact that this mutation might be an LSII rather than LSI predisposing mutations in Tunisian population and appears to be a shared in Mediterranean region. Indeed, previous work in our lab has reported that Tunisian population shares founder mutations with other North African and Middle Eastern populations for 43 inherited conditions [62]. Besides other mutations appeared to be specific to Tunisia and shared by other populations as the case for 11-β hydroxylase deficiency as well as breast and ovarian hereditary cancers [63–66]. Another common founder mutation responsible for cutaneous malignant melanoma shared between Jewish patients from Tunisia, Morocco, Spain and France has also been published [67]. More recently, the BC working group in

our lab has identified a novel *BRCA1* mutation appearing likely specific to North-East of Tunisia and seems to be associated with an aggressive phenotype. One of them seems to have a founder effect [68] emphasizing the particular genetic background of this region as well as Tunisian population. Such phenomenon could certainly be explained by all migratory waves that occurred with the colonial period resulting in the occurrence of new mutations with an important impact on genetic diversity of Tunisian population. All these data highlight the importance of founder and shared mutations in decision making tools for diagnosis and prevention of diseases in North Africa, Middle East and migrant populations living in Europe or Mediterranean region [62].

Moreover, our index case harboring the *MSH2* (c.728G > A) has an aggressive phenotype since Diffuse Gastric tumors are known to be an aggressive gastric tumor histotype. This patient had peritoneal carcinosis, she was therefore not operable and the only treatment she had is a palliative chemotherapy. The *MSH2* and *MLH1* IHC analysis in index case gastric tumor tissue has shown a positive nuclear immunostaining in tumor cells for *MLH1* and an incomplete nuclear immunostaining for *MSH2* protein (Fig. 2B and C). This result is in good agreement with *in silico* prediction analyses of *MSH2* variant since the predicted effect involves the *MSH2/MLH1* complex destabilization which has no effect on the protein tissue expression and detection. According to our best knowledge this study is one of the rare investigations of the MMR proteins expression profile in DGC tissue in a LSII suspected family. The MMR protein expression analysis by IHC is usually performed in CRC tissues according to international recommendation for LS diagnosis in routine. However very few studies have investigated MMR status in tumor tissues corresponding to other LS associated cancers such as GC. Indeed, it has been reported that MMR deficiency seems to be much more frequent in CRC tumors compared to gastric tumors belonging to LS families [69]. These authors have reported that among 45 LS confirmed families 31 cases had GC and only 4 gastric cases were MMR deficient versus 27 MMR proficient cases.

Shedding light on the molecular pathogenesis of GC in LS settings, gastric tumors have been included in the modified Amsterdam criteria II, however the clinical consensus criteria for Gastric tumors diagnosis in LS suspected cases / families remain still unclear and germ mutations search as well as MSI status are not yet clearly indicated for such tumors. Actually there is growing evidence that such tumors from MMR gene germ mutation carriers are part of LS tumor spectrum and relative risk of GC in LS mutation carriers is reported to be higher by 4–19 compared to general population in Western countries [48, 49, 70] at least by 2 fold in endemic areas in Asia [15] as our investigated region. Taking into account the gene-environment interactions our investigated family came from an endemic region of Tunisia with relatively high incidence of digestive cancers known with significant exposure of residents to pesticides and characterized by the spread of water with high levels of toxic chemicals, however no epidemiological or environmental studies have been carried out in this region. These findings highlight the crucial importance of early screening of individuals at high risk of developing such tumors particularly for this under-investigated region in Tunisia.

Our main predictive results showed that the interaction region between the *MSH2* gene and EXO1 may be lost due to this mutation p.R243Q. However, the genome annotation might provide only an inaccurate

level of evidence about the interaction of MSH2 with EXO1 that could be given only by direct mutagenesis binding assays or a high resolution crystal structure of MSH2/EXO1 complex. It has been showed that the interaction of EXO1 with MSH2 implicates an extended range of residues located on different domains but none of the tested mutations implicates the connector domain [71]. However, our docking results suggest that the binding surface of MSH2, respectively with MLH1 and EXO1 are close to each other as reported by [71]. The structural modelisation of this variant p.R243Q does not seem to have an effect on the folding of MSH2 neither on its conformational properties as we have not found any significant differences when we estimated the Gibbs free energy differences and the normal modes of the protein within the complex. Nevertheless, the mutation as part of the interface between MSH2 and MLH1 may be involved in the destabilization of the protein-protein interaction. The crystal structure, presented by [72], suggests the proximity of the connector domain to the MutL binding surface with MutS. Nevertheless, the structure of MutS deviates significantly from that of MSH2 to be able to infer any homology based conclusions about the contribution of R243, hence the utility of our protein-protein predicted complex.

The change in the electrostatic interaction surface of MSH2 upon mutation might result in different association equilibrium with MLH1. DNA repair mechanism is a multistep process that includes cooperatively between different proteins. Electrostatic forces that act at distances of 5–10 Å, can be affected by the p.R243Q mutation [73]. More recently, it has been shown that the electrostatic potential disruption between MLH1 and MSH2 can lead to the accumulation of DNA errors and an impact on the protein-protein complex [74]. The electrostatic properties of MLH1 seem to be highly regulated via acetylation/deacetylation by the Histone deacetylase 6 (HDAC6). Moreover, the results from [74] imply also that a dynamic of association/dissociation between MutL, MutS and HDAC6 is required for DNA repair. In this regard, electrostatic forces, disrupted by the p.R243Q, would play an important role and could result in a significant effect.

Beyond the *MSH2* mutation another putative deleterious mutation in the *FANCD2* gene (p.A627S) was identified. *FANCD2* is required for maintenance of chromosomal stability. It also promotes accurate and efficient pairing of homologs during meiosis. It is involved in the repair of DNA double-strand breaks, both by homologous recombination and single-strand annealing. It may participate in S phase and G2 phase checkpoint activation upon DNA damage playing a role in preventing breakage. This gene encodes the protein for complementation group D2. The A627 amino acid is located at the opposite helices facing the protein-protein interface with FANCI that also contains an ubiquitination site at residue K561. FANCD2 ubiquitination stabilizes a conformational change of the protein-protein complex [46] that is required to increase the affinity for the double stranded DNA [47]. The perturbation of the flexibility by p.A627S variant of FANCD2 may result in two effects. First, the formation of the complex FANCD2 and the ubiquitin-conjugating enzyme E2-UBE2T would have to overcome the energy barrier induced by p.A627S local destabilization since the variation happens at close distance to K561. The second effect may result in the impairment of the association between FANCD2 and FANCI induced by destabilization of the local interface near the site of the variant. Moreover, since monoubiquitination occurs on the evolutionarily

conserved lysine residues, mutations affecting FANCD2 at K561 residue seem to result in major molecular defects in response to DNA damage agents in the Fanconi Anemia cells [75, 76].

Our index case comes from a consanguineous marriage with accumulation of two potentially deleterious variants besides potential deficiency in the cell adhesion system suggesting a putative increasing risk of several cancers in this family. Unfortunately, we were not able to screen the index case's brother with CCR and her two affected sisters (breast and ovarian cancers) for this variant since they were dead. Susceptibility to cancer in offspring of consanguineous marriages has been largely studied. For countries such as Tunisia with common consanguineous marriage, the association between consanguinity and mortality due to cancers is highly important for public health programs [77–79].

Moreover, it has been reported that consanguinity and inbreeding plausibly led to the accumulation of population-specific founder pathogenic / or likely pathogenic sequence variants [80]. In fact, two reported studies have observed that inbreeding and ROH result in an increased risk of CCR [81] and developing leukaemia, lymphoma, colorectal and prostate cancer [82]. Supplementary genetic screenings of this mutation in larger cohorts in this region as in Tunisia are needed to confirm these hypotheses.

Conclusion

Our findings highlight the genetic heterogeneity of cancer predisposition in Tunisia and the importance of the use of target NGS to identify clinically actionable genetic variants for disease management. To our best knowledge, no study has investigated the molecular basis of GC in LS context neither in Tunisia nor in North Africa. Our findings suggest that *MSH2* c.728G > A (p.R243Q) rare mutation could represent a candidate marker for LS II screening. We suggested this variant be shared with Mediterranean countries. Nevertheless, genetic screening of a larger cohort of GC is needed to confirm its association with LSII syndrome in Tunisia. Our structural model prediction showed that this mutation is likely pathogenic and is involved in the MSH2-MLH1 complex stability suggesting the revision of its status from VUS to likely pathogenic. The novel *FANCD2* variant appears to have putative functional effects as well. The identification of such likely deleterious mutations will help in the setting up of specific clinical surveillance protocol for individuals at high risk of developing this disease in terms of yearly endoscopy and colonoscopy as well as a mammography biannually for women at risk.

Abbreviations

HDGC	Hereditary Diffuse Gastric Carcinoma
MLPA	Multiplex Ligation Probe Amplification
GC/DGC	Gastric Carcinoma / Diffuse Gastric Carcinoma
FIGC	Familial Intestinal Gastric Cancer
GAPPS	Proximal Polyposis Of the Stomach
LS	Lynch Syndrome
HNPCC	Hereditary Non Polyposis Colorectal Cancer
BC/LBC	Breast Cancer/Lobular Breast Cancer
MSI	Microsatellite Instability
OC	Ovarian Cancer
IGCLC	International Gastric Cancer Linkage Consortium
IRB	Institutional Review Board
gDNA	Genomic DNA
MAF	Minor Allele Frequency
sDNA	Somatic DNA
IGV	Integrative Genome Variant
HP	<i>Helicobacter Pylori</i>
PDB	Protein Data Bank
IHC	Immunohistochemistry
DAB	DiAminoBenzidine
AEC	AminoEthyl Carbazole
FFPE	Formalin Fixed Paraffin Embedded
APBS	Adaptive Poisson-Boltzmann Solver
WT	Wild Type
CCR	Colorectal Cancer
B/LB	Benign/Likely Benign
VUS	Variant of Uncertain Significance
NR	Not Reported
NI	Non Indicated

HDGC	Hereditary Diffuse Gastric Carcinoma
NGS	Next Generation Sequencing

Declarations

Ethics approval and consent to participate

Written informed consents were obtained from all participants. Ethical approval (Nub : 2017/6F/I/Gastric Cancer) according to the Declaration of Helsinki Principles was obtained from the biomedical ethics committee of Institut Pasteur de Tunis (OHRP registration number: IRB00005445, FWA00010074).

Consent for publication

We declare that we have obtained written informed consent for publication of anonymous clinical information from participants. Informed consent forms we used have been reviewed and approved by the biomedical ethics committee of Institut Pasteur de Tunis (IRB) (2017/6F/I/Gastric Cancer). No identifying images or other personal details were collected from participants for this study and no minors have been enrolled.

Availability of data and materials

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

Competing of interest

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

Funding

This work was supported by the Tunisian Ministry of Public Health, the Tunisian Ministry of Higher Education and Scientific Research (LR16IPT05) and MOBIDOC device funded by European Union in the frame of PASRI and EMORI program and administered by the National Agency of scientific research support (ANPR).

Authors' Contribution

MK led the study and the writing of the paper. **MK, SE, SA** and **SB** contributed to the experimental design and conception of the study. Sample and clinical data acquisition were performed by **MM, MMa, AK, HB, HT, LH, EC** and **MA**. **MK, JBAH** and **HO** contributed to data analysis and result interpretation. **MK, JBAH** and **HO** provided the Bioinformatics study analysis. **MK, JBAH, SL, AJG** and **AM** provided the technical

support for the experimental study. **MK, JBAH and HO** were Involved in the drafting of the manuscript. **SA** and **SB** provided the critical analysis of the paper.

All authors contributed to writing the manuscript read and agreed to the published version of the manuscript.

Acknowledgment

The authors are extremely grateful to the patients and their families whose participation made this work possible and to Dr. Etienne ROULEAU from Institut Gustave Roussy, Villejuif for supporting the Sanger Sequencing analysis of both *CDH1* and *CTNNA1* gene and MLPA assays.

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. doi:10.3322/caac.21492.
2. Moussa SA-B, Moussa A, Kourda N, Mezlini A, Abdelli N, Zerimech F, et al. Lynch syndrome in Tunisia: first description of clinical features and germline mutations. *Int J Colorectal Dis.* 2011;26:455–67.
3. Setia N, Clark JW, Duda DG, Hong TS, Kwak EL, Mullen JT, et al. Familial Gastric Cancers. *The Oncologist.* 2015;20:1365–77. doi:10.1634/theoncologist.2015-0205.
4. Vecchia CL, Negri E, Gentile A, Franceschi S. Family history and the risk of stomach and colorectal cancer. *Cancer.* 1992;70:50–5. doi:10.1002/1097-0142(19920701)70:1<50::AID-CNCR2820700109>3.0.CO;2-I.
5. 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.
6. Peltomäki P. Update on Lynch syndrome genomics. *Fam Cancer.* 2016;15:385–93.
7. Rahner N, Steinke V, Schlegelberger B, Olschwang S, Eisinger F, Hutter P. Clinical utility gene card for: Lynch syndrome (MLH1, MSH2, MSH6, PMS2). *Eur J Hum Genet.* 2010;18:1071–1071.
8. Lorans M, Dow E, Macrae FA, Winship IM, Buchanan DD. Update on Hereditary Colorectal Cancer: Improving the Clinical Utility of Multigene Panel Testing. *Clin Colorectal Cancer.* 2018;17:e293–305. doi:10.1016/j.clcc.2018.01.001.
9. Lynch HT, Smyrk TC, Lynch PM, Lanspa SJ, Boman BM, Ens J, et al. Adenocarcinoma of the small bowel in Lynch syndrome II. *Cancer.* 1989;64:2178–83. doi:10.1002/1097-0142(19891115)64:10<2178::AID-CNCR2820641033>3.0.CO;2-0.

10. Dillon JL, Gonzalez JL, DeMars L, Bloch KJ, Tafe LJ. Universal screening for Lynch syndrome in endometrial cancers: frequency of germline mutations and identification of patients with Lynch-like syndrome. *Hum Pathol*. 2017;70:121–8.
11. Hampel H, Frankel W, Panescu J, Lockman J, Sotamaa K, Fix D, et al. Screening for Lynch syndrome (hereditary nonpolyposis colorectal cancer) among endometrial cancer patients. *Cancer Res*. 2006;66:7810–7.
12. Stoffel E, Mukherjee B, Raymond VM, Tayob N, Kastrinos F, Sparr J, et al. Calculation of risk of colorectal and endometrial cancer among patients with Lynch syndrome. *Gastroenterology*. 2009;137:1621–7.
13. Capelle LG, Grieken NCTV, Lingsma HF, Steyerberg EW, Klokman WJ, Bruno MJ, et al. Risk and Epidemiological Time Trends of Gastric Cancer in Lynch Syndrome Carriers in The Netherlands. *Gastroenterology*. 2010;138:487–92. doi:10.1053/j.gastro.2009.10.051.
14. Gylling A, Abdel-Rahman WM, Juhola M, Nuorva K, Hautala E, Järvinen HJ, et al. Is gastric cancer part of the tumour spectrum of hereditary non-polyposis colorectal cancer? A molecular genetic study. *Gut*. 2007;56:926–33. doi:10.1136/gut.2006.114876.
15. Park YJ, Shin KH, Park JG. Risk of gastric cancer in hereditary nonpolyposis colorectal cancer in Korea. *Clin Cancer Res Off J Am Assoc Cancer Res*. 2000;6:2994–8.
16. Sheehan M, Heald B, Yanda C, Kelly ED, Grobmyer S, Eng C, et al. Investigating the Link between Lynch Syndrome and Breast Cancer. *Eur J Breast Health*. 2020;16:106–9.
17. Gylling AHS, Nieminen TT, Abdel-Rahman WM, Nuorva K, Juhola M, Joensuu EI, et al. Differential cancer predisposition in Lynch syndrome: insights from molecular analysis of brain and urinary tract tumors. *Carcinogenesis*. 2008;29:1351–9.
18. Rouprêt M, Yates DR, Comperat E, Cussenot O. Upper urinary tract urothelial cell carcinomas and other urological malignancies involved in the hereditary nonpolyposis colorectal cancer (lynch syndrome) tumor spectrum. *Eur Urol*. 2008;54:1226–36.
19. Lu KH, Daniels M. Endometrial and ovarian cancer in women with Lynch syndrome: update in screening and prevention. *Fam Cancer*. 2013;12:273–7.
20. Randall LM, Pothuri B. The genetic prediction of risk for gynecologic cancers. *Gynecol Oncol*. 2016;141:10–6.
21. Schmeler KM, Lynch HT, Chen L, Munsell MF, Soliman PT, Clark MB, et al. Prophylactic surgery to reduce the risk of gynecologic cancers in the Lynch syndrome. *N Engl J Med*. 2006;354:261–9.

22. Lynch HT, Lanspa SJ, Boman BM, Smyrk T, Watson P, Lynch JF, et al. Hereditary nonpolyposis colorectal cancer–Lynch syndromes I and II. *Gastroenterol Clin North Am*. 1988;17:679–712.
23. Lynch HT, Lynch PM, Lanspa SJ, Snyder CL, Lynch JF, Boland CR. Review of the Lynch syndrome: history, molecular genetics, screening, differential diagnosis, and medicolegal ramifications. *Clin Genet*. 2009;76:1–18.
24. Watson P, Vasen HF, Mecklin J-P, Bernstein I, Aarnio M, Järvinen HJ, et al. The risk of extra-colonic, extra-endometrial cancer in the Lynch syndrome. *Int J Cancer*. 2008;123:444–9.
25. Kanth P, Grimmer J, Champine M, Burt R, Samadder NJ. Hereditary Colorectal Polyposis and Cancer Syndromes: A Primer on Diagnosis and Management. *Am J Gastroenterol*. 2017;112:1509–25.
26. Møller P, Seppälä TT, Bernstein I, Holinski-Feder E, Sala P, Gareth Evans D, et al. Cancer risk and survival in path_MMR carriers by gene and gender up to 75 years of age: a report from the Prospective Lynch Syndrome Database. *Gut*. 2018;67:1306–16.
27. Ben Sghaier R, Jansen AML, Bdioui A, Van Wezel T, Ksaa M, Elgolli L, et al. Targeted next generation sequencing screening of Lynch syndrome in Tunisian population. *Fam Cancer*. 2019;18:343–8.
28. 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.
29. 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. doi:10.1001/jamaoncol.2014.168.
30. 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. doi:10.1016/S1470-2045(20)30219-9.
31. Singh A, Pandey GK. Primer Design Using Primer Express® for SYBR Green-Based Quantitative PCR. In: Basu C, editor. *PCR Primer Design*. New York, NY: Springer; 2015. p. 153–64. doi:10.1007/978-1-4939-2365-6_11.
32. Chikhaoui A, Elouej S, Nabouli I, Jones M, Lagarde A, Ben Rekaya M, et al. Identification of a ERCC5 c.2333T>C (L778P) Variant in Two Tunisian Siblings With Mild Xeroderma Pigmentosum Phenotype. *Front Genet*. 2019;10. doi:10.3389/fgene.2019.00111.
33. Warren JJ, Pohlhaus TJ, Changela A, Iyer RR, Modrich PL, Beese LS. Structure of the Human MutSa DNA Lesion Recognition Complex. *Mol Cell*. 2007;26:579–92. doi:10.1016/j.molcel.2007.04.018.
34. Gupta S, Gellert M, Yang W. Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. *Nat Struct Mol Biol*. 2012;19:72–8. doi:10.1038/nsmb.2175.

35. Šali A, Blundell TL. Comparative Protein Modelling by Satisfaction of Spatial Restraints. *J Mol Biol.* 1993;234:779–815. doi:10.1006/jmbi.1993.1626.
36. Wu H, Zeng H, Lam R, Tempel W, Kerr ID, Min J. Structure of the human MLH1 N-terminus: implications for predisposition to Lynch syndrome. *Acta Crystallogr Sect F Struct Biol Commun.* 2015;71:981–5.
37. Delgado J, Radusky LG, Cianferoni D, Serrano L. FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics.* 2019;35:4168–9. doi:10.1093/bioinformatics/btz184.
38. Pierce BG, Hourai Y, Weng Z. Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PloS One.* 2011;6:e24657.
39. D.A. Case, I.H. Ben-Shalom, S.R. Brozell, D.S. Cerutti, T.E. Cheatham, III, V.W.D. Cruzeiro, T.A. Darden, R.E. Duke, D. Ghoreishi, G. Giambasu, T. Giese, M.K. Gilson, H. Gohlke, A.W. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, R. Krasny, T. Kurtzman, T.S. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, V. Man, D.J. Mermelstein, K.M. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, A. Onufriev, F. Pan, R. Qi, D.R. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C.L. Simmerling, J. Smith, J. Swails, R.C. Walker, J. Wang, H. Wei, L. Wilson, R.M. Wolf, X. Wu, L. Xiao, Y. Xiong, D.M. York and P.A. Kollman (2019), AMBER 2019,
40. Vascon F, Gasparotto M, Giacomello M, Cendron L, Bergantino E, Filippini F, et al. Protein electrostatics: From computational and structural analysis to discovery of functional fingerprints and biotechnological design. *Comput Struct Biotechnol J.* 2020;18:1774–89. doi:10.1016/j.csbj.2020.06.029.
41. Dolinsky TJ, Nielsen JE, McCammon JA, Baker NA. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. *Nucleic Acids Res.* 2004;32 suppl_2:W665–7. doi:10.1093/nar/gkh381.
42. Rodriguez NJ, Reineke P, LaDuca H, Xicola RM, Llor X. 1066 - Phenotypic Expression of Cdh1 Germline Mutations Unselected for Hereditary Diffuse Gastric Cancer. *Gastroenterology.* 2018;154:S-204. doi:10.1016/S0016-5085(18)31081-3.
43. Schouten JP, McElgunn CJ, Waaijer R, Zwiijnenburg 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.
44. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics.* 2015;31:3555–7. doi:10.1093/bioinformatics/btv402.
45. Jorgensen NH, Byer DE, Gould AB. Aspiration pneumonitis. Prevention and treatment. *Minn Med.* 1989;72:517–9, 530.

46. Tan W, Murphy VJ, Charron A, Twest S van, Sharp M, Constantinou A, et al. Preparation and purification of mono-ubiquitinated proteins using Avi-tagged ubiquitin. *PLOS ONE*. 2020;15:e0229000. doi:10.1371/journal.pone.0229000.
47. Wang R, Wang S, Dhar A, Peralta C, Pavletich NP. DNA clamp function of the monoubiquitinated Fanconi anaemia ID complex. *Nature*. 2020;580:278–82.
48. Aarnio M, Sankila R, Pukkala E, Salovaara R, Aaltonen LA, de la Chapelle A, et al. Cancer risk in mutation carriers of DNA-mismatch-repair genes. *Int J Cancer*. 1999;81:214–8.
49. Goecke T, Schulmann K, Engel C, Holinski-Feder E, Pagenstecher C, Schackert HK, et al. Genotype-phenotype comparison of German MLH1 and MSH2 mutation carriers clinically affected with Lynch syndrome: a report by the German HNPCC Consortium. *J Clin Oncol Off J Am Soc Clin Oncol*. 2006;24:4285–92.
50. Vasen HF, Stormorken A, Menko FH, Nagengast FM, Kleibeuker JH, Griffioen G, et al. MSH2 mutation carriers are at higher risk of cancer than MLH1 mutation carriers: a study of hereditary nonpolyposis colorectal cancer families. *J Clin Oncol Off J Am Soc Clin Oncol*. 2001;19:4074–80.
51. Jakubowska A, Ławniczak M, Wojnarska B, Cybulski C, Huzarski T, Byrski T, et al. CDH1 gene mutations do not contribute in hereditary diffuse gastric cancer in Poland. *Fam Cancer*. 2010;9:605–8. doi:10.1007/s10689-010-9381-2.
52. Vasen HF, Offerhaus GJ, den Hartog Jager FC, Menko FH, Nagengast FM, Griffioen G, et al. The tumour spectrum in hereditary non-polyposis colorectal cancer: a study of 24 kindreds in the Netherlands. *Int J Cancer*. 1990;46:31–4.
53. Chen S, Wang W, Lee S, Nafa K, Lee J, Romans K, et al. Prediction of germline mutations and cancer risk in the Lynch syndrome. *Jama*. 2006;296:1479–87.
54. Drost M, Lützen A, Van Hees S, Ferreira D, Calléja F, Zonneveld JB, et al. Genetic screens to identify pathogenic gene variants in the common cancer predisposition Lynch syndrome. *Proc Natl Acad Sci*. 2013;110:9403–8.
55. Pande M, Wei C, Chen J, Amos CI, Lynch PM, Lu KH, et al. Cancer spectrum in DNA mismatch repair gene mutation carriers: results from a hospital based Lynch syndrome registry. *Fam Cancer*. 2012;11:441–7.
56. Pearlman R, Frankel WL, Swanson B, Zhao W, Yilmaz A, Miller K, et al. Prevalence and Spectrum of Germline Cancer Susceptibility Gene Mutations Among Patients With Early-Onset Colorectal Cancer. *JAMA Oncol*. 2017;3:464–71.
57. Pino MS, Chung DC. Application of molecular diagnostics for the detection of Lynch syndrome. *Expert Rev Mol Diagn*. 2010;10:651–65.

58. Tamura G. Alterations of tumor suppressor and tumor-related genes in the development and progression of gastric cancer. *World J Gastroenterol WJG*. 2006;12:192.
59. Watson P, Lynch HT. Cancer risk in mismatch repair gene mutation carriers. *Fam Cancer*. 2001;1:57–60.
60. Yurgelun MB, Kulke MH, Fuchs CS, Allen BA, Uno H, Hornick JL, et al. Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer. *J Clin Oncol*. 2017;35:1086–95. doi:10.1200/JCO.2016.71.0012.
61. Ziada-Bouchaar H, Sifi K, Filali T, Hammada T, Satta D, Abadi N. First description of mutational analysis of MLH1, MSH2 and MSH6 in Algerian families with suspected Lynch syndrome. *Fam Cancer*. 2017;16:57–66.
62. Romdhane L, Kefi R, Azaiez H, Halim NB, Dellagi K, Abdelhak S. Founder mutations in Tunisia: implications for diagnosis in North Africa and Middle East. *Orphanet J Rare Dis*. 2012;7:52.
63. Cherbal F, Bakour R, Adane S, Boualga K, Benais-Pont G, Maillet P. BRCA1 and BRCA2 germline mutations screening in Algerian breast/ovarian cancer families. *Dis Markers*. 2010;28:377–84.
64. Kharrat M, Trabelsi S, Chaabouni M, Maazoul F, Kraoua L, Jemaa LB, et al. Only two mutations detected in 15 Tunisian patients with 11 β -hydroxylase deficiency: the p.Q356X and the novel p.G379V. *Clin Genet*. 2010;78:398–401.
65. Troudi W, Uhrhammer N, Romdhane KB, Sibille C, Amor MB, El Khil HK, et al. Complete mutation screening and haplotype characterization of BRCA1 gene in Tunisian patients with familial breast cancer. *Cancer Biomark*. 2008;4:11–8.
66. Zhu Y-S, Cordero JJ, Can S, Cai L-Q, You X, Herrera C, et al. Mutations in CYP11B1 gene: phenotype-genotype correlations. *Am J Med Genet A*. 2003;122:193–200.
67. Yakobson E, Eisenberg S, Isacson R, Halle D, Levy-Lahad E, Catane R, et al. A single Mediterranean, possibly Jewish, origin for the Val59Gly CDKN2A mutation in four melanoma-prone families. *Eur J Hum Genet*. 2003;11:288–96.
68. 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.
69. Zeinalian M, Hashemzadeh-Chaleshtori M, Salehi R, Kazemi M, Emami MH. Tumor microsatellite instability and clinicopathologic features in Iranian colorectal cancer patients at risk for Lynch syndrome. *J Res Med Sci Off J Isfahan Univ Med Sci*. 2015;20:154–60. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4400710/>. Accessed 19 Mar 2021.

70. Vasen HF, Wijnen JT, Menko FH, Kleibeuker JH, Taal BG, Griffioen G, et al. Cancer risk in families with hereditary nonpolyposis colorectal cancer diagnosed by mutation analysis. *Gastroenterology*. 1996;110:1020–7. doi:10.1053/gast.1996.v110.pm8612988.
71. Goellner EM, Putnam CD, Graham WJ, Rahal CM, Li B-Z, Kolodner RD. Identification of Exo1-Msh2 interaction motifs in DNA mismatch repair and new Msh2-binding partners. *Nat Struct Mol Biol*. 2018;25:650–9.
72. Groothuizen FS, Winkler I, Cristóvão M, Fish A, Winterwerp HH, Reumer A, et al. MutS/MutL crystal structure reveals that the MutS sliding clamp loads MutL onto DNA. *Elife*. 2015;4:e06744.
73. Zhou H-X, Pang X. Electrostatic interactions in protein structure, folding, binding, and condensation. *Chem Rev*. 2018;118:1691–741.
74. Zhang M, Hu C, Moses N, Haakenson J, Xiang S, Quan D, et al. HDAC6 regulates DNA damage response via deacetylating MLH1. *J Biol Chem*. 2019;294:5813–26.
75. Romick-Rosendale LE, Lui VWY, Grandis JR, Wells SI. The Fanconi anemia pathway: Repairing the link between DNA damage and squamous cell carcinoma. *Mutat Res Mol Mech Mutagen*. 2013;743–744:78–88. doi:10.1016/j.mrfmmm.2013.01.001.
76. Nepal M, Che R, Ma C, Zhang J, Fei P. FANCD2 and DNA Damage. *Int J Mol Sci*. 2017;18:1804. doi:10.3390/ijms18081804.
77. El-Kheshen G, Saadat M. Prevalence of consanguineous marriages among Shi'a populations of Lebanon. *J Biosoc Sci*. 2013;45:675–82.
78. Rafiee L, Saadat M. Prevalence of consanguineous marriages among Iranian Georgians. *J Biosoc Sci*. 2011;43:47.
79. Saadat M, Vakili-Ghartavol R. Parental consanguinity and susceptibility to drug abuse among offspring, a case–control study. *Psychiatry Res*. 2010;180:57–9.
80. Laitman Y, Friebel TM, Yannoukakos D, Fostira F, Konstantopoulou I, Figlioli G, et al. The spectrum of BRCA1 and BRCA2 pathogenic sequence variants in Middle Eastern, North African, and South European countries. *Hum Mutat*. 2019;40:e1–23.
81. Rich JN. Cancer stem cells in radiation resistance. *Cancer Res*. 2007;67:8980–4.
82. Bener A, El Ayoubi HR, Chouchane L, Ali AI, Al-Kubaisi A, Al-Sulaiti H, et al. Impact of consanguinity on cancer in a highly endogamous population. *Asian Pac J Cancer Prev*. 2009;10:35–40.

Figures

Figure 1

The familial pedigree of the index case “JI-021” with family history of cancer.

Figure 2

Immunohistochemical expression patterns E-Cadherin, MLH1 and MSH2 in "JI-021" GC tissues A) Immunostaining of E-Cadherin in index case GC tissues showing a loss of E-cadherin in tumor cells compared to normal glandular adjacent cells showing positive membranous staining (x400). B) Nuclear positive immunostaining of MLH1 in index case GC tissue (x400). C) Uncomplete nuclear positive immunostaining of MSH2 in index case GC tissue (x200) original magnification.

Figure 3

Statistical distribution of variants identified after filtering strategy

Figure 4

Molecular modeling analysis of p.R243Q mutation. (A) Molecular model of the MSH2/MSH6 complex showing the location of the mutation on the connector domain of MSH2. (B) Ternary complex of MSH2/MSH6/MLH1 predicted by protein-protein docking.

Figure 5

Predicted effect of MSH2 p.R243Q mutation on the electrostatic properties of MSH2. Red and Blue colors correspond to potentials of electronegative and electropositive intensities respectively. The maximum intensity values are -5 and +5 kb T e-1. (A) Electrostatic potential of MSH2 showing the variation at the site of the mutation for the wild type and p.R243Q forms. (B) Electrostatic potential of MLH1 at the interaction surface with MSH2.

Supplementary Files

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

- [AdditionalFile1.xls](#)
- [AdditionalFile2.xls](#)
- [Additionalfile3.pdf](#)