**Role of germline aberrations affecting *CTNNA1*, *MAP3K6* and *MYD88* in gastric cancer susceptibility**

Robbert D.A. Weren1,\*, Rachel S. van der Post2,\*, Ingrid P. Vogelaar1, J. Han J.M. van Krieken2, Liesbeth Spruijt1, Jan Lubinski3, Anna Jakubowska3, Urszula Teodorczyk3, Cora M. Aalfs4, Liselotte P. van Hest5, Carla Oliveira6,7,8, Eveline J. Kamping1, Hans K. Schackert9, Guglielmina N. Ranzani10, Encarna B. Gómez García11, Frederik J. Hes12, Elke Holinski-Feder13, Maurizio Genuardi14, Margreet G.E.M. Ausems15, Rolf H. Sijmons16, Anja Wagner17, Lizet E. van der Kolk18, Annemieke Cats19, Inga Bjørnevoll20, Nicoline Hoogerbrugge1 and Marjolijn J.L. Ligtenberg1,2

1 Department of Human Genetics, Radboud university medical center, Nijmegen, the Netherlands

2 Department of Pathology, Radboud university medical center, Nijmegen, the Netherlands

3 Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland

4 Department of Clinical Genetics, Academic Medical Centre, Amsterdam, the Netherlands

5 Department of Clinical Genetics, VU University Medical Center, Amsterdam, the Netherlands

6 Expression Regulation in Cancer Group, Instituto de Investigação e Inovação em Saúde, Porto, Portugal (i3S)

7 Institute of Molecular Pathology and Immunology of the University of Porto (IPATIMUP), Porto, Portugal

8 Department of Pathology and Oncology, Faculty of Medicine, University of Porto, Al. Prof. Hernâni Monteiro, 4200-319 Porto, Portugal.

9 Department of Surgical Research, Universitätsklinikum Carl Gustav Carus, Technische Universität Dresden, Dresden, Germany

10 Department of Biology and Biotechnology, University of Pavia, Pavia, Italy

11 Department of Clinical Genetics, Maastricht University Medical Center, Maastricht, the Netherlands

12 Department of Clinical Genetics, Leiden University Medical Center, Leiden, the Netherlands

13 Medizinische Klinik und Poliklinik IV, Campus Innenstadt, Klinikum der Universität München, Germany

14 Institute of Genomic Medicine, Catholic University of the Sacred Heart, Rome, Italy

15 Department of Genetics, University Medical Center Utrecht, the Netherlands

16 Department of Genetics, University Medical Center Groningen, University of Groningen, the Netherlands

17 Department of Clinical Genetics, Erasmus University Medical Center, Rotterdam, the Netherlands

18 Family Cancer Clinic, The Netherlands Cancer Institute, Amsterdam, the Netherlands

19 Department of Gastrointestinal Oncology, The Netherlands Cancer Institute/ Antoni van Leeuwenhoek, Amsterdam, the Netherlands

20 Department of Medical Genetics, St. Olavs Hospital, Trondheim University Hospital

\*These authors contributed equally

|  |  |  |  |
| --- | --- | --- | --- |
| |  | | --- | | **Online Supporting Information** | | |  | | --- | | **Page** | |
| **Supplementary 1. Clinical criteria of included gastric cancer patients (n=299)** | 2 |
| **Supplementary 2. Materials and Methods** | 3 |
| **Supplementary 3. Average coverage of the open reading frame of *CDH1*, *CTNNA1*, *MAP3K6* and *MYD88* using smMIP-based targeted sequencing** | 4 |
| **Supplementary 4. Overview of known *CTNNA1* mutation-positive GC families** | 5 |

**Supplementary 1. Clinical criteria of included gastric cancer patients for which germline testing was successful (n=283)**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Category** | **Criteria\*** | **Number** | **Age (SD; Range)** | **Sex** |
| **HDGC** | **HDGC Criteria 2015:**  One DGC <40 years (n=50)  Two or more GC, one DGC<50 (n=53)  Three or more DGC at any age (n=9)  1 DGC & 1 LBC, 1<50 (n=2)  New criterion HDGC guidelines 2015, 2 GC at least 1 DGC, not fulfilling one of the criteria above (n=33) | **147** | **44 (15; 17-80)** | **M: 49**  **F: 98** |
| **FIGC** | **Criteria:**  One IGC <40 years (n=11)  Two or more GC, one IGC<50 (n=12)  Three or more IGC at any age (n=3) | **26** | **41 (11; 21-66)** | **M: 17**  **F: 9** |
| **FGC/EOGC** | **Familial GC or early-onset GC <40 years with mixed histology or (partly) unknown (not from all relatives were histology reports available):**  Index with DGC (n=4)  Index with MGC (n=6)  Index with IGC (n=30)  Histology unknown (n=36) | **76** | **52 (15; 28-87)** | **M: 42**  **F: 34** |
| **Other** | **One GC between age 40 – 50:**  Index with DGC (n=22)  Index with MGC (N=2)  Index with IGC (n=4)  Index with GC, histology unknown (n=6) | **34** | **42 (3; 40-49)** | **M: 14**  **F: 20** |

**\*Included are first/second degree relatives affected by gastric cancer; Abbreviations: DGC= diffuse gastric cancer; IGC= intestinal gastric cancer; MGC= mixed gastric cancer; GC= gastric cancer; LBC= lobular breast cancer; SD= standard deviation; M= male; F= female**

**Supplementary 2. Materials and Methods**

**smMIP-based targeted sequencing**

To determine the germline status of *CDH1*, *CTNNA1*, *MAP3K6* and *MYD88* in GC patients, we applied smMIP-based targeted sequencing on blood derived DNA of these patients (according to previously published methods[13-16]). Briefly, 81 smMIPs were designed to target both DNA strands of the protein encoding genomic regions of *CDH1* independently (i.e., double tiling). These smMIPs contained a stretch of eight random nucleotides to enable molecular tagging (smMIP sequences are available upon request). To sequence the open reading frame of *CTNNA1*, *MAP3K6* and *MYD88*, a total of 63, 78 and 16 smMIPs were designed, respectively (smMIP sequences are available upon request). In contrast to the *CDH1* smMIPs, these smMIPs contained a stretch of five random nucleotides for molecular barcoding and most genomic regions were only targeted by one smMIP. Following smMIP capture and endonuclease treatment (final volume 27µl), 20µl of the mixture was distributed equally across four PCR reactions. A unique barcoded reverse primer was used per sample to enable subsequent equimolar pooling of multiple libraries. Sequencing of these pooled libraries was performed using the Illumina NextSeq 500 system.

**Data analysis**

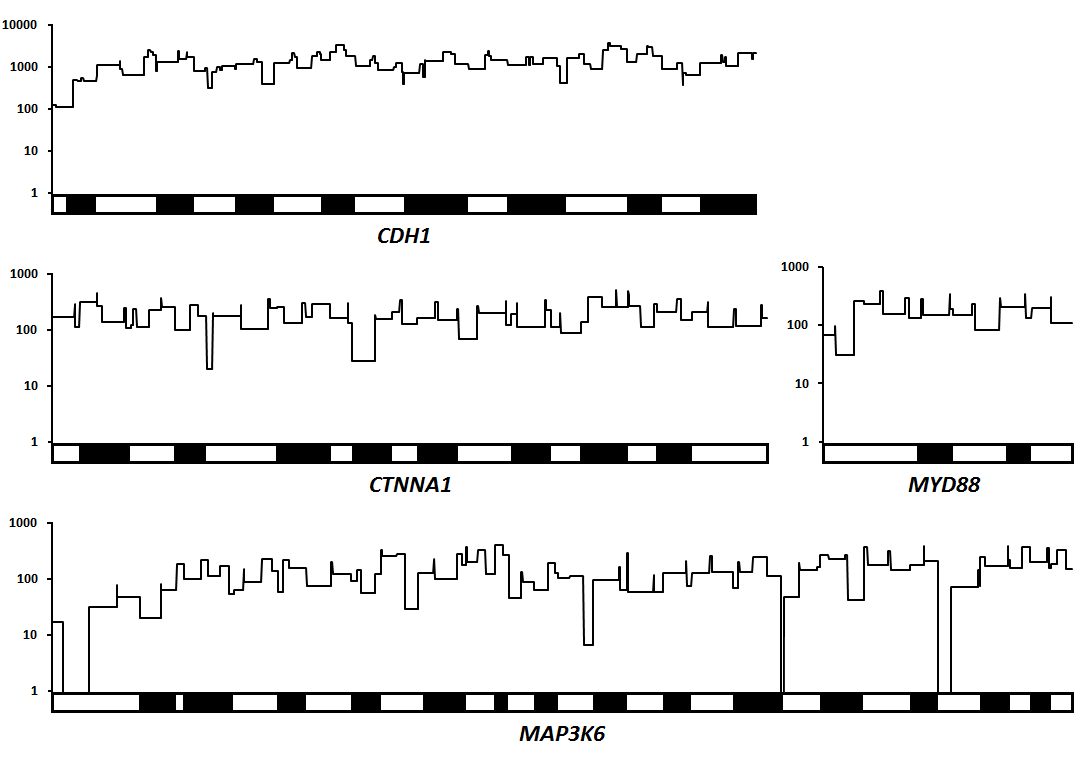
Prior to data analysis, fastq files were separated using the sample specific barcodes. Next, these fastq files were analyzed using the SeqNext software package version 4.3.0; JSI medical systems GmbH). Consensus reads were generated using the eight (*CDH1*) or five (*CTNNA1*, *MAP3K6*, *MYD88*) basepair molecular tags and variants were called if the corresponding variant was observed in ≥3 unique reads, ≥5% of all unique reads and ≥10 unique reads were observed at the corresponding genomic position. After manual exclusion of mapping artefacts and low confident variant calls (percentage variant reads ≤10), all non-synonymous variants in the protein coding regions and variants affecting the canonical splice sites were selected. The ExAC database, containing whole-exome sequencing data derived from 60,706 individuals, was used as a control dataset to establish the variant allele frequency of germline variants in *CDH1*, *CTNNA1*, *MAP3K6* and *MYD88* in the general population (Exome Aggregation Consortium (ExAC), Cambridge, MA (http://exac.broadinstitute.org) [01/04/2017 accessed]. [17] To predict the pathogenicity of encountered germline variants, *in silico* analyses were performed using the Alamut software package (Alamut Visual version 2.9 (Interactive Biosoftware, Rouen, France)). All reported germline variants in *CTNNA1*, *MAP3K6*, *MYD88* (Table 1) were validated using Sanger sequencing (primers available upon request).

**Statistics**

A two-tailed Chi-square test with Yates' correction was performed to determine the potential enrichment of germline variants affecting *CTNNA1*, *MAP3K6* and *MYD88* in GC patients compared to the general population (i.e., ExAC database). Patient 270A, who is a family member of a previously published *CTNNA1* family [9], was excluded from the statistical analysis which was applied to determine if strong loss-of-function germline variants in *CTNNA1* were enriched in our GC cohort compared to the general population (reducing the number of *CTNNA1* alleles in our cohort from 566 to 564).

**Immunohistochemistry for α-E-catenin**

To assess α-E-catenin protein expression, immunohistochemical staining was performed on 4-μm-thick formalin-fixed paraffin-embedded slides of gastric carcinomas. Tissue sections were deparaffinised in xylene and rehydrated with alcohol. Antigen retrieval was done by boiling in EDTA at 96°C for 30 min. Endogenous peroxidase activitity was blocked by 10 minutes incubation in 3% H2O2 in methanol. The rabbit monoclonal antibody against α-E-catenin (EP1793Y from Abcam) was diluted to 1:200 in normal antibody diluent (ImmunoLogic, Duiven, The Netherlands). Slides were labeled with primary anti- α-E-catenin antibody by 1 hour incubation at room temperature. Next, sections were incubated in Power Vision 1:2 plus poly-horseradisch peroxidase (HRP)-anti-mouse/rabbit/rat solution (ImmunoLogic, Duiven, The Netherlands) for 30 minutes. Visualization was done with DAB for 5 min. Nuclei were counterstained with hematoxylin. Slides were dehydrated, cleared in xylene, and mounted with micromount. Tissue samples of the gastric cancer specimen from patient 432A with *CTNNA1* variant p.Arg330fs were stained for α-E-catenin protein, as well as 105 anonymous control gastric adenocarcinomas (core biopsies on tissue microarrays, see also the study R.S. van der Post et al Virchows Arch (2014) 464:673-679). **Supplementary 3. Average coverage of the open reading frame of *CDH1*, *CTNNA1*, *MAP3K6* and *MYD88* using smMIP-based targeted sequencing**



*Black and white boxes represent alternating exons, including the canonical splice sites of the respective genes. Coverage of the open reading frame was determined based on the average number of reads obtained per smMIP in 107 (*CDH1*) and 283* (CTNNA1*,* MAP3K6 *and* MYD88*) germline DNA samples of gastric cancer patients, respectively. X-axis: alternating exons. Y-axis: average number of unique reads mapping to the corresponding genomic position.*

**Supplementary 4. Overview of known *CTNNA1* mutation-positive GC families**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Family** | **Amino acid change** | **Nucleotide change** | **Proband** | **Family history (at counseling)#** | **Other malignancies** | **Family origin** | **Signet-ring cell differentiation** | | **Original criteria by Caldas et al** | | **HDGC criteria 2010** | | **HDGC criteria 2015** | | **Remarks/ comments** | | **Reference paper(s)** |
| **1** | p.Arg27fs | c.80\_81del | DGC (62) mut\* | 1DGC (63) mut\*; 1GC (54); 1GC (71); 4x2GC (45, 51, 57, 72) | Brain tumor | Dutch | | Yes | | No | | No | | Yes | | Ten carriers (6 tested and 4 inferred). Endoscopic surveillance revealed invasive diffuse gastric cancer (pT1a signet ring cell mucosal foci) in four relatives at age 40, 41, 42 and 75. Three patients underwent gastrectomy and multiple mucosal SRCC foci were found, no typical *in situ* or pagetoid SRCC were reported. | [1, 2] |
| **2** | p.Asn71fs | c.211A>AT | DGC (22) mut\* | 1DGC (59) mut\*, 2BC (50) | NR | White/ Canadian | Yes with pools of mucin | | Yes | | Yes | | Yes | |  | | [3] |
| **3** | p.Arg129\* | c.385C>T | DGC (72) mut\* | 1DGC (52) | NR | Italian | Unknown | | No | | No | | Yes | |  | | [3] |
| **4** | p. Arg330fs | c.964\_988dup | DGC (40) mut\* | 2GC (56) | NR | Dutch | Yes | | Yes | | Yes | | Yes | |  | | Current paper |
| **5** | p.Asn443fs | c.1328dup | GC (30) mut\* | 1GC (48) | NR | Polish | Unknown | | No | | No | | No | |  | | Current paper |

*#number before ‘GC’ refers to the degree of family member (i.e. first- or second degree relative), number in brackets refers to age. Reference papers: [1] Majewski IJ, Kluijt I, Cats A, Scerri TS, de Jong D, Kluin RJ, Hansford S, Hogervorst FB, Bosma AJ, Hofland I, Winter M, Huntsman D, Jonkers J, Bahlo M, Bernards R. An alpha-E-catenin (CTNNA1) mutation in hereditary diffuse gastric cancer. J Pathol 2013;****229****(4):621-9. [2] van der Post RS, van Dieren J, Grelack A, Hoogerbrugge N, van der Kolk LE, Snaebjornsson P, Lansdorp-Vogelaar I, van Krieken JH, Bisseling TM, Cats A. Outcomes of screening gastroscopy in first-degree relatives of patients fulfilling hereditary diffuse gastric cancer criteria. Gastrointest Endosc 2017. [3] Hansford S, Kaurah P, Li-Chang H, Woo M, Senz J, Pinheiro H, Schrader KA, Schaeffer DF, Shumansky K, Zogopoulos G, Almeida Santos T, Claro I, Carvalho J, Nielsen C, Padilla S, Lum A, Talhouk A, Baker-Lange K, Richardson S, Lewis I, Lindor NM, Pennell E, MacMillan A, Fernandez B, Keller G, Lynch H, Shah SP, Guilford P, Gallinger S, Corso G, Roviello F, Caldas C, Oliveria C, Pharoah PDP, Huntsman DG. Hereditary Diffuse Gastric Cancer Syndrome: CDH1 Mutations and Beyond. JAMA Oncol 2015;1(1):23-32.*

Supplementary 5. A-E-catenin protein expression in gastric cancer controls

Panel figuur suppl 5A-H_def.tif

Immunohistochemistry for α-E-catenin of 105 anonymous control gastric adenocarcinomas showed strong membranous staining in 101 cases (96%). Sixty-two cases were of the intestinal-type of which 60 showed strong positive membranous staining (see example A-B, magnification 200x) and two showed no staining (one nuclear and one total negative, see example C-D, magnification 400x). Forty-two cases were of the diffuse-type of which 40 showed positive membranous staining (see example E-F, magnification 400x) and two showed no membranous staining: one showed nuclear staining (see example G-H, magnification 400x) and one showed no staining.