Pathology & Oncology Research (2019) 25:711–721
https://doi.org/10.1007/s12253-018-0547-9

ORIGINAL ARTICLE

Early Hereditary Diffuse Gastric Cancer (eHDGC) is Characterized by Subtle Genomic Instability and Active DNA Damage Response

Soroush Nasri[1,2] · Bostjan Humara[1] · Ahmad Anjomshoaa[1,3] · Nourodin Moradi[2] · Naghmeh Gholipour[2,4] · Sakineh Mashjoor[5] · Peng Zhang[6]

Received: 22 October 2017 / Accepted: 16 November 2018 / Published online: 13 December 2018
© Arányi Lajos Foundation 2018

Abstract

Diffuse gastric cancer (DGC) is one of the two primary types of stomach cancer. Carriers of germline mutations in the gene encoding E-cadherin are predisposed to DGC. The primary aim of the present study was to determine if genomic instability is an early event in DGC and how it may lead to disease progression. Chromosomal aberrations in early intramucosal hereditary diffuse gastric cancer (eHDGC) were assessed using array comparative genomic hybridization (array CGH). Notably, no aneuploidy or other large-scale chromosomal rearrangements were detected. Instead, all aberrations affected small regions (< 4.8 Mb) and were predominantly deletions. Analysis of DNA sequence patterns revealed that essentially all aberrations possessed the characteristics of common fragile sites. These results and the results of subsequent immunohistochemical examinations demonstrated that unlike advanced DGC, eHDGCs is characterized by low levels of genomic instability at fragile sites. Furthermore, they express an active DNA damage response, providing a molecular basis for the observed indolence of eHDGC. This finding is an important step to understanding the pathology underlying natural history of DGC and supports a revision of the current definition of eHDGC as a malignant disease.

Keywords Genomic instability · Diffuse gastric cancer · Array CGH · DNA fragility

Introduction

Gastric cancer is the fourth most common cancer [1], and the second most common cause of cancer associated death worldwide after lung cancer [2]. This cancer has two main histological subtypes: intestinal and diffuse [3]. One third of the diagnosed cases are diffuse [4]. While the incidence of intestinal gastric cancer has gradually declined in the last decades the incidence of diffuse gastric cancer (DGC) has remained steady or may even be on the rise [5]. DGC has the poor diagnosis (early stages can't be detected by gastroscopy) and prognosis (two third of cases are diagnosed with the unrespectable disease with 5–10% five-year survival [6]). Also, the biology of diffuse type of gastric cancer is largely unknown. One percent of all DGC cases are hereditary [7]. Hereditary diffuse gastric cancer (HDGC) is the autosomal dominant susceptibility for

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12253-018-0547-9) contains supplementary material, which is available to authorized users.

✉ Naghmeh Gholipour
ngholipour@nigeb.ac.ir

[1] Cancer Genetics Laboratory, Department of Biochemistry, University of Otago, P.O. Box: 56, Dunedin 9054, New Zealand

[2] Milad Center for Medical Genetics, P.O. Box: 48156-59176, Sari, Iran

[3] Department of Medical Genetics, Faculty of Medicine, Kerman University of Medical Sciences, P.O. Box: 6619-14155, Kerman, Iran

[4] Department of Medical Genetics, National Institute of Genetic Engineering and Biotechnology, Shahrak-e Pajoohesh, km 15, Tehran - Karaj Highway, P.O. Box: 14965-161, Tehran, Iran

[5] Department of Marine biology, Faculty of Marine Science and Technology, Hormozgan University, P.O. Box: 3995, Bandar Abbas, Iran

[6] Division of Biomedical Science and Biochemistry, Research School of Biology, ANU College of Medicine, Biology and Environment, Australian National University, P.O. Box: 2601, Canberra, Australia

DGC, largely associated with germline mutations in E-cadherin gene or *CDH1* which is a key cell-cell adhesion molecule and was first found in a NZ Maori family with the high rate of DGC [8]. Somatic mutations of *CDH1* are present in 50% of sporadic DGC cases and promoter methylation in 40–80%. Therefore the aetiology of HDGC seems to be comprehensively similar to many sporadic cases [9–11]. In New Zealand, *CDH1* germline mutation carriers are recommended to undergo prophylactic gastrectomy to prevent symptomatic disease [12]. Microscopic inspection of resected stomachs reveals multiple minutes, occult cancer foci (stage T1) which are predominantly composed of signet ring cells (SRCs) which are mitotically inactive and indolent. Also further bellow there are poorly differentiated cells (PDCs) which are the cells that invade muscularis mucosa and deeper layers in more advanced stages [13].

We have had access to the world largest collection of prophylactically removed stomachs with early HDGC. Our set of early HDGC samples were from patients all from a Maori tribe residing in Bay of Plenty, New Zealand, all diagnosed with an exactly same germline mutation in their *CDH1*.

The mechanism underlying the progression of indolent early (e) HDGC to invasive disease is unknown. Since E-cadherin loss is the defining feature of eHDGC foci [13] and lack of this protein can lead to disruption of the attachment, positioning, and segregation of mitotic chromosomes [13–16], we hypothesize that genomic instability is the driving force that shifts early stage foci to invasive disease.

Materials and Methods

Subjects

FFPE (formalin- fixed paraffin embedded) specimens from 12 members of one HDGC kindred were used in this study. All patients were carriers of the *CDH1* 1008G > T germline mutation and had undergone total gastrectomy following the endoscopic detection of DGC. Ethical approval was obtained from the Multi-Region Ethics Committee (reference number: MEC/06/05/053).

Section Preparation and Staining

4 μm section is prepared from a FFPE block. To be able to microdissect the pure SRC and PDC components of our eHDGC lesions, it was first necessary to select and optimize the staining conditions with consideration of not only cell morphology but also downstream applications. In the first set of experiments, different histological stains were evaluated for their suitability to identify SRCs and PDCs in eHDGC sections for laser capture microdissection (LCM). Toluidine Blue (TB), a popular nuclear stain used for LCM, readily marked nuclei within the gastric mucosa, however it did not provide a clear distinction between SRCs and PDCs. On the other hand, PAS (Periodic Acid Schiff, a mucin-binding stain), strongly stained mucin-rich SRCs, but was not as helpful in discriminating PDCs (little mucin, prominent nuclei) from other, non-neoplastic stromal cells. Cresyl Violet (CV), another nuclear stain, was superior to TB and it clearly marked both nuclei and (sub) cellular borders. We thus combined CV with PAS to visualise the prominent nuclei of neoplastic cells and their various degrees of mucin content (high in SRCs, low in PDCs).

Laser Capture Microdisection, DNA Extraction, and Multiplex PCR

SRCs and PDCs were carefully laser capture microdissected using Arcturus LM200 LCM microscope (Applied Biosystems,Foster City, California, United States) 7.5 μm laser beam was used to achieve maximum precision in microdissection of individual cells following re-examination of microdissected regions and material collected on cap (>90% neoplastic cells, and uncontaminated normal-appearing epithelia), DNA was extracted and subsequently were subjected to the multiplex PCR quality control exactly as it is defined in Nasri et al. [17]. As the DNA yield for one eHDGC focus was far below the recommended amount required for array comparative genomic hybridization (array CGH), we had to pool DNA from multiple foci. As a result, for each patient, consecutive FFPE sections from four eHDGC foci were microdissected and pooled to collect 30,000 SRCs and an equal number of cells from PDCs and adjacent, normal appearing gastric epithelium.

Agilent Oligonucleotide Array CGH

In our previous study [17] we showed Agilent oligonucleotide array CGH platform generates reliable genomic results on FFPE material. Agilent 44 K array CGH was used for genomic profiling of our HDGC DNA samples. Three array analyses were performed per patient; one for SRCs, one for PDCs and one for normal epithelial cells. Microarrays were scanned using Agilent SureScan high-resolution scanner (Agilent Technologies, Santa Clara, California, USA). Scanned images were analyzed using Feature Extraction 10.5 image analysis software (Agilent Technologies, Santa Clara, California, USA) without applying any background correction. QCmetrics were satisfactory. LOWESS or locally weighted scatter plot smoothing normalization [18] was used to adjust dye effect on each array. Subsequently data were log transformed to meet the requirements of the normality for ANOVA. Log 10 base values were converted into copy numbers. Regions of copy number variation were found using genomic segmentation method in Partek software (Partek

Incorporated, St. Louis, Missouri, USA). In genomic segmentation we kept the 10 markers default but increased the stringency by using 0.7 signal to noise and 0.0005 segmentation Copy number between 1.7–2.3 are considered normal, above that amplification and below that deletion.

Multiplex Ligation-Dependent Probe Amplification

MLPA oligonucleotide probes were designed according to the 'Designing synthetic MLPA probes' protocol v.10 (MRC-Holland, Amsterdam, Netherlands) (Table S1). We ensured there were no overlaps between the hybridizing sequences of different probes. Probes were designed to be from 88 to 168-nucleotide in length. There was a minimum 4 bp size difference between the probes. Our probes also met the MLPA recommendation of a maximum of 2 G/C in the final 5 bases at the 3′ end and a maximum of 3 G/C directly adjacent to the primer recognition sequence. The thermodynamic melting point (Tm) of each hybridization sequence was separately calculated using the 'RAW' program (http://www.mrcholland.com/WebForms/) to make sure all Tm values were above 70 °C as recommended by MRC-Holland. Our synthetic probes were manufactured by Invitrogen (Carlsbad, US) and supplied to us in a desalted purified form.

MLPA was performed as recommended by the manufacturers using SALSA MLPA P200-A1 kits (MRC-Holland, Amsterdam, Netherlands) and PCR product separated by capillary electrophoresis on an ABI 3730 Genetic Analyzer (Applied Biosystems, Foster City, California).

DNA Flexibility Study

TwistFlex software (http://margalit.huji.ac.il/TwistFlex/index.html), developed by the Life Sciences Institute of the Hebrew University of Jerusalem, was used to estimate the DNA fragility in our aberrant regions. In this study, the minimal number of peaks in a cluster was set to three and maximal distance between peaks in a cluster to 5000 bp. We also determined the A/T and AT-dinucleotide percentages of these regions.

Immunohistochemistry

We examined by immunohistochemistry the activation state of H2AX, ATM, ATR, CHK2 and CHK1 in eHDGC. Serial sections (5 μm) from HDGC-FFPE blocks were mounted on slides coated with aminopropyltriethoxysilone. Sections were stained using EnVision™ system (DakoCytomotion, Glostrup, Denmark) before being treated with 3.3′-diaminobenzidine (DAB) chromogen (Vector Laboratories, Burlingame, US). To each tissue section, 250 μl of normal goat serum (NGS) blocking solution (5% NGS in 0.02 M PBS containing 1% BSA) was added and spread evenly over the section. The following rabbit monoclonal antibodies were used for immunohistochemistry: ATM-pS1981 (no.YE070901R; Epitomics, Burlingame, US; diluted at 1:100), H2A.x-pS139 (no.YE080802r; Epitomics, Burlingame, US; diluted at 1:100), P-Chk2-T68 (no.2197P; Cell Signalling Technology Inc., Danvers, US; diluted at 1:200), ATR (C-20) (no.sc-21,848; Santa Cruz Biotechnology Inc., Santa Cruz, US; diluted at 1:200) and Chk1-pS345 (no.ab47318; Abcam, Cambridge, UK; diluted at 1:100). Then, 250 μl of anti-rabbit secondary antibodies conjugated to the labeled polymer-HRP were added to the sections followed by incubation in a moist chamber for 30 min. The antibody-antigen reaction was visualised using the DAB substrate kit (Vector Laboratories, Burlingame, US). Slides were rinsed in PBS to remove unbound secondary antibodies. The DAB substrate solution was prepared by mixing the supplied reagents according to the manufacturer's instructions. Slides were immersed in fresh haematoxylin (Sigma-Aldrich, St Louis, US) for 3–4 min and then washed thoroughly in water until the water ran clear.

Results

Using the CV-PAS double stain, SRCs were easily identifiable, whilst identification of PDCs became possible when considerable care was exercised. As a result, CV-PAS was the standard staining used in all LCM experiments. The black box on the bottom left is the magnified picture of the part of the lesion shown in a smaller box further up on the left. The black arrows indicate SRCs and red arrows show PDCs. SRCs are big and round indolent cells with cytoplasm full of mucin (mucin is stained strong purple by PAS) (Fig. 1).

All 36 samples produced the 100–300 bp fragment set in multiplex PCR quality control [17] and thus were of sufficient quality for array CGH. All the 36 arrays generated acceptable QC plots. The capacity of the Agilent platform to reliably detect genomic copy numbers from archived tissue has previously been validated by comparison with genomic profiles from fresh-frozen tissue [17].

A total of 78 different chromosomal aberrations were detected in SRCs and PDCs (Fig. 2). The size of aberrations ranged from 0.12–4.72 Mb and from 0.088–4.74 Mb in SRCs and PDCs, respectively. Using the same segmentation criteria as for the neoplastic cells, no aberrations were observed in neighbouring, normal appearing epithelia. Except for three amplifications in SRCs (11q23.3, 19q13.33, 20q11.22) and one in PDCs (12q13.3), the detected aberrations were all deletions.

Importantly, no aneuploidy, changes affecting whole chromosomal arms, or other macroscopic rearrangements were found. Deletion in 19q13.41 region, which was seen in five

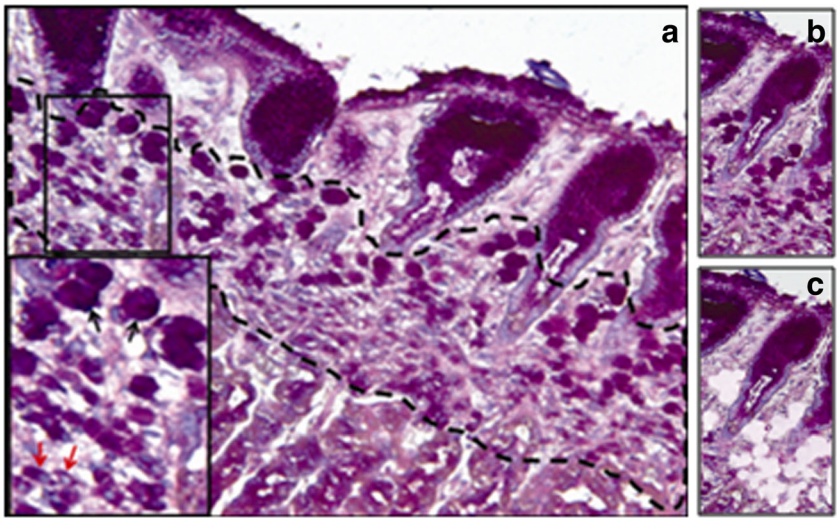

Fig. 1 Early HDGC lesions from a prophylactically removed stomach. **a** A medium size early HDGC lesion, encompassed by a dotted line. 4 μm section is prepared from a FFPE block subsequently double stained with PAS (Periodic Acid Schiff) and CV (Cresyl Violet). CV is a blue nuclear stain that gives a good cellular morphology visibility. The black box on the bottom left is the magnified picture of the part of the lesion shown in a smaller box further up on the left. The black arrows indicate Signet Ring Cells (SRCs) and red arrows show Poorly Differentiated Cells (PDCs). SRCs are big and round indolent cells with cytoplasm full of mucin (mucin is stained strong purple by PAS). **b** Lesion before microdissection. **c** Lesion after microdissection of cancer cells (SRCs and PDCs)

out of twelve patients (~ 42% frequency), has been the most common aberration across the patients. The chromosomal region 9q21.32 was imbalanced in four patients (~ 33% frequency) and 16p12.3 in only three (25% frequency), which made them second and third most frequent aberration. Both of the latter two aberrations were also deletions. These three

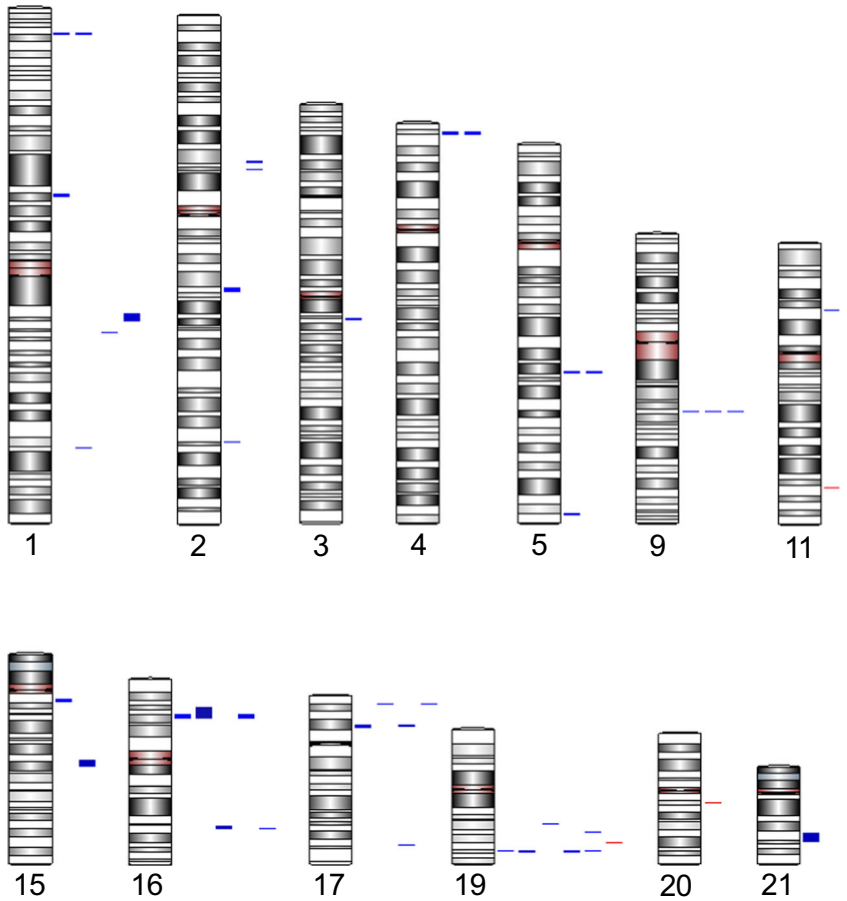

Fig. 2 Genomic imbalances detected by array CGH

aberrations were simultaneously present in only one eHDGC sample. To confirm the presence of detected chromosomal aberrations in eHDGC, MLPA was performed on DNA extracted from SRCs and PDCs.

Aberrations in 1p36, 4p16, 5q23, 9q21, 16p12, 17p12 and 17p13 each were detected in at least two patients by array CGH. Two genes representing each of these chromosomal regions were selected for MLPA validation.

The overall concordance rate between the MLPA and CGH results was 71.5% for SRCs and 72% for PDCs. Six out of thirteen discordant cases involved alterations that were detected by MLPA and not by CGH, probably due to a combination of lower resolution of CGH and occasionally false positive MLPA results.

Tables 1 (SRCs) and 2 (PDCs) show the gene changes as detected by MLPA and array CGH in each patient. In general, the MLPA results were in satisfactory agreement [19] with the CGH results, and the majority of the aberrations (33 out of 40) detected by array CGH was validated by MLPA (Table S2 and S3).

Next, aberrations observed in SRCs were compared with those detected in PDCs to determine whether any changes exist that are common across patients and discriminate SRCs from PDCs. Aberrations in chromosome 7 and 12 were only seen in PDCs and aberrations in chromosomes 11 only in SRCs. Aberrations in 12 chromosomes (1–5, 9, 15–17, 19–21) were seen in both cancer cell types. Fascinatingly, there was an extensive overall similarity between chromosomal imbalances observed in SRCs and PDCs. Only 10 out of 41 total aberrations seen in PDCs were not seen in SRCs, which means approximately 76% of all aberrations seen in PDCs were also seen in SRCs. Chromosomal aberrations that were present in PDCs but not SRCs and vice versa were uncommon and mostly specific for a given patient. Together with the above results, these findings are consistent with early stages of chromosomal instability where selection pressure is small.

Given the lack of a significant overlap of the detected aberrations across patients or between eHDGCs and sporadic DGC, the chromosomal aberrations were further analysed for loci that might point to commonalities among the aberrant regions. The three most frequent aberrations are home to 39 genes as shown in Table 4 (http://pubmatrix.grc.nia.nih.gov/). Of these genes, 22 were zinc finger proteins however this is mainly because the family of these proteins is clustered in 19q13.41 (Table 3).

To assess relationships and common themes among the 39 genes, 'Gene Ontology' and 'KEGG (Kyoto Encyclopedia of Genes and Genomes) Pathway' analyses (*gather.genome.duke.edu*) were performed (Table 4). These tools use a collection of online databases on genomic data, enzymatic pathways and signalling cascades to reveal molecular interactions within cells.

Aberrations in these regions cannot be justified by the function of the genes they cover. However majority of the aberrant regions we found have been reported in later stages of DGC but it is hard to relate them to the disease initiation. Rendering further investigation on the aberrant regions we noticed around 40% of these regions are in cytobands were common fragile sites were reported before. We thus investigated the fragility features of the aberrant regions detected in eHDGC.

We explored the fragility of these regions based on their nucleic acid composition using TwistFlex software (this software estimates the DNA fragility based on the DNA conformational free energy concept) [20]. This assessment showed almost all aberrant regions we found have the characteristics of regions with high fragility meaning high number of flexibility peaks more than one per 100 kb that is seen in non-fragile sites of genome) and high A/T ($> 61 \pm 3.6\%$) and AT-dinucleotide ($> 8 \pm 1\%$) content [21]. Apart from Mishmar's definition of non-fragile region we randomly chose a region in chromosome 1 where no fragile site was reported before. From 43,831,597 to 44,041,570 a region with 209,947 bp size 2 flexibility peaks were found by TwistFlex and A/T content was 51.14% and AT-dinucleotide content was 5.53% which matches the criteria of non-fragile DNA defined by Mishmar et al [21]. It does not seem to be any over-presentation of the fragile sites in chromosomal aberrations detected in previous studies on advanced DGC studies.

The fact that almost all aberrant regions detected in eHDGC had characteristics of fragile sites, indicates that eHDGC does not display chromosomal instability typically found in advanced disease, but may rather represent a precancerous and premalignant stage. The occurrence of DNA breaks at fragile sites, as any other break, will lead to the induction of classic DNA damage response pathways.

Sections for immunostaining were available for five of the 12 patients investigated in this study. In five out of the five samples (D, E, G, H, J), phosphorylated γ-H2AX, pS1981-ATM and pT68-CHK2 could be detected in eHDGC foci (Fig. 3). No activated H2AX, ATM and CHK2 were observed in adjacent, normal appearing epithelia, consistent with the lack of chromosomal aberrations in these regions. Further immunohistochemistry using antibodies against activated ATR and CHK1 did not reveal any positive signals in either eHDGC or adjacent epithelia. These results indicate that the DNA damage response is present and activated in the eHDGC foci of *CDH1* mutation carriers. Together with the finding of chromosomal instability specifically at sites with positive fragility features, the outcome of this project strongly suggests that eHDGC (i.e. hereditary T1a signet ring cell carcinoma) does not represent a malignancy as currently defined but rather is a precancerous stage with the potential to progress to invasive disease.

Table 1 Aberrations seen in SRCs and their flexibility characteristics

Chr	Cytoband	Start	End	del/amp	Sample	Number of flexibility peaks/b	% A/T	% AT- dn
chr1	1q21.1 - 1q21.2	144,744,997	148,125,625	del	F	123/3380628	75.70%	21.10%
chr1	1q21.3 - 1q22	153,200,208	153,319,557	del	I	13/119349	73.30%	19.00%
chr1	1p36.21	12,680,848	13,958,751	del	T	23/1277903	76.30%	22.30%
chr1	1q32.3	210,332,477	211,082,844	del	T	24/750367	76.60%	20.90%
chr1	1p36.21	12,680,848	13,958,751	del	W	23/1277903	76.30%	22.30%
chr1	1p22.3 - 1p22.2	87,329,427	88,817,562	del	W	47/1488135	76.70%	20.80%
chr2	2p14 - 2p13.3	70,041,244	70,952,581	del	I	29/911337	78.00%	20.90%
chr2	2p13.2 - 2p13.1	73,669,167	74,215,838	del	I	25/546671	80.10%	22.00%
chr2	2q14.3 - 2q21.1	128,833,713	130,725,074	del	T	59/1891361	75.90%	20.80%
chr2	2q33.1 - 2q33.2	203,260,314	203,886,203	del	T	20/625889	77.80%	21.00%
chr3	3q12.2 - 3q12.3	101,896,745	103,104,467	del	U	101/1207722	77.90%	21.70%
chr4	4p16.2	3,477,841	4,705,645	del	I	44/1227804	74.10%	21.30%
chr4	4p16.2	3,391,816	5,055,344	del	U	47/1663528	72.60%	20.80%
chr5	5q23.1	118,621,955	119,379,366	del	C	22/757411	75.60%	20.60%
chr5	5q23.1	118,621,955	119,379,366	del	U	22/757411	75.60%	20.60%
chr5	5q35.2 - 5q35.3	176,450,099	177,620,215	del	U	21/1170116	79.60%	23.50%
chr9	9q21.32	85,210,768	85,742,449	del	D	46/531681	76.60%	20.20%
chr9	9q21.32	85,210,768	85,742,449	del	F	46/531681	76.60%	20.20%
chr9	9q21.32	85,210,768	85,899,327	del	I	55/688559	76.30%	20.40%
chr11	11p13	31,913,126	32,821,625	del	T	38/908499	80.20%	25.00%
chr11	11q23.3	117,036,039	117,471,941	amp	T	12/435902	66.20%	16.00%
chr15	15q21.2 - 15q21.3	49,469,125	52,475,758	del	G	120/3006633	77.50%	21.10%
chr15	15q11.2	21,545,984	23,044,681	del	W	58/1498697	77.80%	21.90%
chr16	16q22.1 - 16q22.2	69,396,085	69,781,363	del	C	14/385278	76.30%	20.30%
chr16	16p12.3	17,181,622	18,968,830	del	E	73/1787208	77.80%	23.00%
chr16	16q22.1 - 16q22.3	68,363,971	69,925,944	del	J	36/1561973	75.70%	20.20%
chr16	16p13.12 - 16p12.3	14,254,275	18,968,830	del	R	162/4714555	78.10%	22.70%
chr16	16p12.3	17,181,622	18,873,757	del	U	73/1692135	77.80%	23.00%
chr17	17p13.1	6,856,942	7,115,875	del	I	3/258933	77.60%	20.30%
chr17	17p12 - 17p11.2	15,407,602	16,274,854	del	T	33/867252	76.40%	21.50%
chr17	17q25.1	71,177,114	71,398,190	del	T	1/221076	70.70%	17.30%
chr17	17p13.1	6,856,942	7,115,875	del	U	3/258933	77.60%	20.30%
chr17	17p12 - 17p11.2	15,407,602	16,819,916	del	W	46/1412314	76.70%	21.50%
chr19	19q13.33	54,496,739	54,949,356	amp	G	20/452617	75.90%	21.70%
chr19	19q13.32	50,094,762	50,216,724	del	I	2/121962	86.50%	26.80%
chr19	19q13.41	57,884,708	58,278,309	del	I	5/393601	74.30%	20.00%
chr19	19q13.41	57,731,842	58,474,028	del	J	13/742186	74.20%	21.40%
chr19	19q13.2	46,529,305	46,800,603	del	R	13/271298	75.10%	20.40%
chr19	19q13.41	57,765,384	58,724,670	del	T	19/959286	74.80%	21.30%
chr19	19q13.41	57,779,207	58,236,018	del	U	7/456811	76.60%	23.40%
chr20	20q11.22	33,204,027	33,372,584	amp	F	3/168557	76.40%	20.60%
chr21	21q22.11 - 21q22.12	32,410,102	36,478,834	del	I	176/4068732	75.80%	22.60%

Discussion

The discovery that *CDH1* germline mutations predispose to DGC has led to the description of HDGC. The introduction of prophylactic gastrectomy for *CDH1* mutation carriers - their only curative treatment - has made available tissue from asymptomatic individuals prone to malignancy. Histopathological mapping of HDGC gastrectomies revealed

Table 2 Aberrant regions detected in PDCs by CGH and their flexibility peaks

Chr	Cytoband	Start	End	del/amp	Sample	Number of flexibility peaks/bp	% A/T	% AT-dn
chr1	1q32.3	210,398,606	210,900,598	del	J	14/501992	77.90%	21.60%
chr1	1p35.1	32,912,340	33,151,711	del	R	10/239371	83.20%	23.90%
chr1	1p13.2	114,829,869	115,060,945	del	R	6/231076	75.00%	19.80%
chr1	1p36.21	12,680,848	13,958,751	del	T	23/1277903	76.30%	22.30%
chr1	1q32.3	210,332,477	211,218,800	del	T	24/886323	76.60%	20.90%
chr1	1p36.22 - 1p36.21	12,171,598	13,858,752	del	W	24/1687154	72.80%	21.90%
chr2	2p14 - 2p13.3	70,041,244	70,545,009	del	C	13/503765	75.20%	19.90%
chr2	2p14 - 2p13.3	70,041,244	70,545,009	del	H	13/503765	75.20%	19.90%
chr2	2p13.2 - 2p13.1	73,669,167	74,215,838	del	I	25/546671	80.10%	22.00%
chr2	2q14.3 - 2q21.1	128,833,713	130,725,074	del	T	59/1891361	75.90%	20.80%
chr3	3q12.2 - 3q12.3	101,896,745	103,104,467	del	U	101/1207722	77.90%	21.70%
chr4	4p16.2	3,477,841	4,866,124	del	I	57/1388283	75.00%	21.20%
chr5	5q23.1	118,621,955	119,379,366	del	C	22/757411	75.60%	20.60%
chr5	5q35.2 - 5q35.3	176,450,099	177,301,881	del	I	18/851782	81.20%	24.20%
chr5	5q15	94,889,694	95,398,231	del	J	22/508537	73.20%	18.30%
chr5	5q23.1	118,621,955	119,379,366	del	U	22/757411	75.60%	20.60%
chr5	5q35.2 - 5q35.3	176,450,099	177,669,329	del	U	22/1219230	79.20%	23.50%
chr7	7p13 - 7p12.3	45,966,501	48,954,613	del	W	154/2988112	75.80%	21.80%
chr9	9q31.1	104,439,232	106,328,898	del	G	165/1889666	78.90%	23.10%
chr9	9q21.32	85,210,768	85,792,652	del	I	52/581884	76.40%	20.40%
chr9	9q21.32	85,210,768	85,742,449	del	U	46/531681	76.60%	20.20%
chr12	12q13.3	56,099,262	56,187,394	amp	C	1/88132	73.30%	20.80%
chr15	15q21.2 - 15q21.3	49,469,125	52,656,037	del	G	126/3186912	77.50%	21.10%
chr15	15q15.3 - 15q21.1	42,393,924	43,152,796	del	I	18/758872	75.30%	20.60%
chr15	15q22.31	62,245,393	63,416,814	del	I	40/1171421	77.40%	21.50%
chr16	16q22.1 - 16q22.2	69,396,085	69,781,363	del	C	14/385278	76.30%	20.30%
chr16	16p13.12 - 16p12.3	14,254,275	18,990,655	del	H	162/4736380	78.10%	22.70%
chr16	16q22.1 - 16q22.3	69,396,085	69,820,355	del	J	17/424270	76.60%	20.50%
chr16	16p13.3	2,521,420	3,014,419	del	U	4/492999	79.10%	24.60%
chr16	16q12.2 - 16q13	54,418,154	55,854,746	del	W	54/1436592	75.80%	20.80%
chr17	17q21.31 - 17q21.32	40,810,446	42,325,613	del	I	33/1515167	77%	22.40%
chr17	17q25.1	71,120,626	71,398,190	del	T	3/277564	71.80%	19.50%
chr17	17p13.3	1,823,378	2,342,397	del	U	11/519019	71.20%	18.40%
chr17	17p13.3	1,823,378	2,342,397	del	W	11/519019	71.20%	18.40%
chr19	19p13.11	16,629,864	16,958,575	del	G	5/328711	76.60%	24.40%
chr19	19q13.41	57,765,384	58,458,191	del	G	11/692807	75%	21.70%
chr19	19q13.2	46,543,183	46,800,603	del	H	13/257420	75.10%	20.40%
chr19	19q13.41	57,884,708	58,297,126	del	I	5/412418	74.30%	20%
chr19	19q13.41	57,840,595	58,474,028	del	J	10/633433	73.60%	19.40%
chr19	19q13.2	46,543,183	46,800,603	del	R	13/257420	75.10%	20.40%

the identity of the earliest apparent manifestations of DGC prior to the onset of advanced disease, pathologically defined as TNM stage T1a intramucosal signet ring cell carcinoma (SRCC, referred to as eHDGC). Thus, the study of the initial stages of DGC has become possible.

In parallel, new functions of the E-cadherin protein encoded by *CDH1* have been described. Whilst initially being regarded as the key epithelial cell-cell adhesion molecule, it is now accepted that E-cadherin and the resulting intercellular adhesion not only confer simple mechanical stability to epithelial sheets, but also have profound effects on the polarity,

Table 3 Most common aberrations seen across patients and the genes they harbour

Aberrant chromosomal region	Genes
19q13.41	*ZNF547, ZNF548, ZNF17, ZNF749, VN1R1, VN1R107P, ZNF17, ZNF749, VN1R107P, ZNF772, ZNF419, ZNF773, MIR1274B, ZNF549, ZNF549, ZNF550, ZNF416, ZIK1, ZNF530, ZNF134, ZNF211, 2SCAN4, ZNF551, ZNF154, ZNF671, ZNF671, ZNF776*
9q21.32	*RASEF*
16p12.3	*XYLT1, RPL7P47, PKD1P4, PKD1P5, MIR3180–3, MIR3179–3, MOM02, ABCC6P1, RPSI5A, ARL6IP1, SMG1*

orientation, and integrity of individual cells. Of particular interest is the role E-cadherin has in the proper orientation of mitotic spindles to allow for a faithful segregation of sister chromatids into dividing cells of similar or divergent fate. If E-cadherin is lost or reduced in cells, misorientation of mitotic spindles occurs and can affect the differentiation program of the progeny [14].

Since in DGC, disease onset is triggered by downregulation of E-cadherin [22, 23], we hypothesised that a misalignment between mitotic spindles and the daughter-mother cell axis might exert some sort of mechanical tension on the chromosomes. Such stress in turn could immediately impact on chromosomal structure leading to aneuploidy or larger chromosomal aberrations [24]. Alternatively, damage might be subtler and manifest itself during subsequent replication, when the single strands of DNA are exposed. We therefore decided to perform a genome-wide analysis of eHDGC foci to assess the level and nature of genomic instability present in the earliest known stages of DGC.

In this study we used laser captured FFPE tissue material as the source of genomic DNA. As a result we had to deal with both low quantity and low quality of the DNA samples. We were also comparing three different cell types (two different cancer types and normal epithelium) in early HDGC lesions thus precision in mining out these cells was also important. To facilitate this procedure, we developed a staining method to distinguish these cells during LCM whilst concurrently optimising our tissue preparation, LCM method and DNA extraction method to increase the DNA yield. Numerous, array studies have been conducted in which FFPE derived DNA was used but the problem of DNA quality has never been carefully addressed [25, 26]. Likewise, several other array studies used whole genome amplification without assessing the faithfulness of the approach [27]. However, this study is the first of its kind where laser captured material from FFPE tissue was used as a source of genomic DNA with reasonable quality and quantity without the need for whole genome amplification prior to array analysis.

In this study, genomic profiles of SRCs, PDCs and normal epithelial cells from laser capture microdissected tissue of T1a HDGC lesions, were generated using Agilent 4 × 44 K array CGH and validated by MLPA. It was initially planned to use real-time qPCR to validate our genomic data at the expression level. However, RNA extracted from our samples was highly degraded and resulted in highly inconsistent amplification patterns. Real-time qPCR could also have been used to probe genomic copy numbers, but we opted for MLPA because it appears to have the most developed multiplexing capacities. In our hands, MLPA was a reliable, fast, cost effective and convenient validation method. It allowed for the simultaneous detection of multiple targets requiring as little as 50 ng of DNA, a great advantage when using FFPE tissue. In recent years the application of MLPA as a method of microarray data validation has grown in numbers [28–30]. This also includes use of MLPA kits in diagnostic laboratories [31, 32], which are traditionally very selective about their testing methods. Considering the limitations of starting material we were facing in this study, MLPA was the most sensible choice for the validation of the array CGH results. Chromosomal imbalances found in SRCs and PDCs, were extensively similar and there

Table 4 'Gene Ontology' and 'KEGG Pathway' analysis result of the most commonly aberrant genes

Gene Ontology	Genes	*P* value
Transcription, transcriptional regulation, nucleic acid metabolism	*ZNF134, ZNF154, ZNF17, ZNF211, ZNF419, ZNF530, ZNF547, ZNF548, ZNF549, ZNF550, ZNF551, ARL6IP, SMG1*	0.0001
Lipid phosphorylation and phosphoinositide phosphorylation, mRNA catabolism and nonsense-mediated decay	*SMG1*	0.0005
KEGG pathway		
Chondroitin / Heparan sulfate glycoprotein biosynthesis	*XYLT1*	0.0003

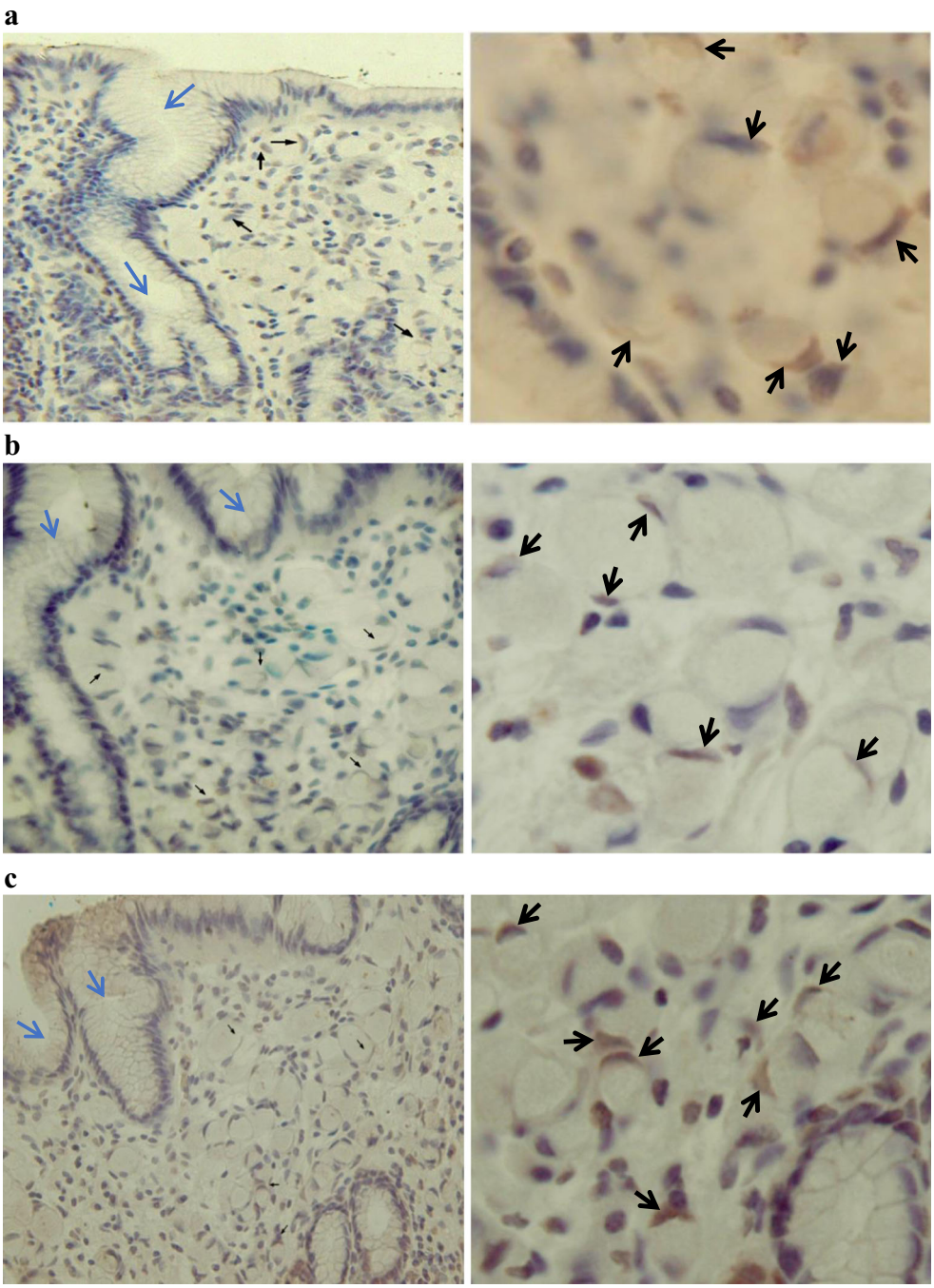

Fig. 3 Activation of DNA damage response in early HDGC lesions. IHC using (**a**) ATM phosphoS1981, **b** γH2AX and **c** CHK2 phosphoThr68 antibodies showed expression of these proteins in a number of cancer cells. Black arrows are pointing to the positively stained cells and blue arrows are pointing to the negatively stained normal cells

was no evidence of consistent acquisition of extra aberration among PDCs that could be related to disease progression. These imbalances were in the form of multiple small aberrations (mainly deletions) in the regions with high DNA fragility. No significant carcinogenesis related pathway was identified in the aberrant regions, which indicates how these regions are possibly selected based on their nucleic acid content rather than the functional importance of the genes they contain.

DNA damage usually activates cell cycle checkpoints. Checkpoint activation stops the cell cycle and gives the cell time to repair the damage before continuing to divide [33]. DNA damage checkpoints occur at the G1/S, G2/M boundaries and sometimes intra-S. ATM and ATR kinases are two major controllers of Checkpoint activation. ATM predominantly responds to DSBs while ATR is known to respond to SSBs such as stalled replication forks [34]. Phosphorylation of downstream targets in a signal transduction cascade by these kinases, ultimately lead to cell cycle arrest. Considering the fact that E-cadherin depletion can impede the replication machinery, together with the observation of multiple chromosomal imbalances in genomic regions with high fragility, has made us hypothesise that these aberrations are resulted from

the stalling of replication fork at these sites [35, 36]. To test this hypothesis we immunohistochemically examined eHDGC lesions for the evidence of DNA damage response (DDR). Our initial assumption was supported when we observed the DDR protein expression in eHDGC lesions however the family of the proteins we expected to see was not activated at all. As it was mention earlier stalled replication fork initially leads to DNA single strand breaks (SSBs), which primarily activates ATR family of DDR proteins. Surprisingly, our IHC examination showed no evidence of ATR family activation, while a clear expression of ATM family of DDR proteins, known to be activated in response to double strand breaks (DSBs) [37], was observed. Although it is seems that activation of ATM family protein is not the primary reaction to SSBs but it is widely accepted that if SSBs are left unrepaired, they could convert into DSBs and trigger ATM response [38]. It has previously been shown that E-cadherin is necessary for an appropriate and efficient DDR. It is possible that due to E-cadherin exhaustion in early HDGC cells the ATR dependent pathway either was not activated or its activation was limited and insufficient which was unable to successfully repair primary SSBs. Either way, the eHDGC cancer cells would be left with DSBs which activates the ATM-DDR pathway. Thus, the neoplastic cells could have had an initial phase of replication stress and correspondingly DNA damage in the form of SSBs, but we have captured them when they have already passed that phase. Alternatively, the loss of mitotic spindle appropriate orientation, caused by E-cadherin depletion [14, 15], might have exerted a specific tension or mechanical stress on chromatids, leading to direct DSBs of genomic segments with high fragility. These results also provide strong evidence for eHDGC being a precursor lesion rather than an invasive carcinoma, which is in contradiction with the current pathological definition of an eHDGC and the wide-held view of SRCC being highly malignant in general.

E-cadherin depletion in the proliferative mucous neck region possibly including committed progenitors and stem cells (located in the upper isthmus) hinders spindle orientation and thereby smooth chromosomal segregation during cell division and results in some level of genomic instability. The less differentiated cells in these lesions did not seem to have acquired further genomic instability (GIN), at least not to a detectable extent, through other mechanisms up to this level. Indeed the proliferative cells in these lesions will continue to select for the aberrations useful for disease progression as well as acquiring new aberrations helpful to the process of EMT and invasion. To explore these extra aberrations that might facilitate EMT and invasion, we suggest a similar comparative genomic study on dedifferentiated mesenchymal cells in later T1a and early T1b lesions. However extensive dissimilarities of PDCs and SRCs at protein level [13] might indicate a considerable difference in epigenetic level, as it has been shown that the epigenetic modifications, in accordance with tumour microenvironment, continue alongside with genetic changes in the process of carcinogenesis [39]. Studying eHDGC lesions in expression level is necessary to unravel such epigenetic mechanisms.

Conclusion

The availability of archived tissue from prophylactic gastrectomies of a large HDGC family provided us with a unique model to explore early stage genomic alterations in a background of rare homogeneity. Having a sample set from patients that are closely related, follow a similar life style, live in the same geographic area, and all harbour the identical *CDH1* germline mutation, is exceptional and motivated us to embark on this technically challenging and laborious project. Although it was not possible to successfully analyse all aspects that we aimed at (i.e. the genomic differences between SRCs and PDCs), the findings of this study have revealed novel and important insights into the genesis of diffuse gastric cancer.

The results of this project add new evidence that, together with existing knowledge, will help to establish eHDGC as a classic precursor lesion. Chromosomal aberrations strongly enriched at fragile sites, together with an intact and activated DNA damage response, are now considered as typical features of precancerous lesions before the acquisition of an invasive phenotype. We consistently observed both the instability at fragile sites and an active DNA damage response in all mutation carriers examined. No aberrations were observed that were common to the majority of patients, however, almost all aberrations were quite small deletions. These findings concur with what one would expect from very early neoplastic changes, namely an unselected state (i.e. random occurrence, little selection pressure) resulting from a common mechanism (i.e. breakage at fragile sites leading to small deletions). Certainly these observations do not provide a definite proof, but together with our previous findings, they strongly argue for eHDGC being a precursor to advanced HDGC.

Acknowledgements This research was supported by HS and JC Anderson Charitable Trust and the Health Research Council of New Zealand.

Compliance with Ethical Standards

Conflict of Interest The authors declare no conflict of interest.

References

1. Parkin DM, Bray F, Ferlay J, Pisani P (2005) Global cancer statistics, 2002. CA Cancer J Clin 55:74–108
2. Crew KD (2006) Neugut AI. Epidemiology of gastric cancer. World J Gastroenterol 12:354–362

3. Lauren P (1965) 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 64: 31–49
4. Ekstrom AM, Hansson LE, Signorello LB, Lindgren A, Bergstrom R, Nyren O (2000) Decreasing incidence of both major histologic subtypes of gastric adenocarcinoma-a population-based study in Sweden. Br J Cancer 83:391–396
5. Henson DE, Dittus C, Younes M, Nguyen H, Albores-Saavedra J (2004) Differential trends in the intestinal and diffuse types of gastric carcinoma in the United States, 1973-2000: increase in the signet ring cell type. Arch Pathol Lab Med 128:765–770
6. Kattan MW, Karpeh MS, Mazumdar M, Brennan MF (2003) Postoperative nomogram for disease-specific survival after an R0 resection for gastric carcinoma. J Clin Oncol 21:3647–3650
7. Stone J, Bevan S, Cunningham D, Hill A, Rahman N, Peto J, Marossy A, Houlston RS (1999) Low frequency of germline E-cadherin mutations in familial and nonfamilial gastric cancer. Br J Cancer 79:1935–1937
8. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, Taite H, Scoular R, Miller A, Reeve AE (1998) E-cadherin germline mutations in familial gastric cancer. Nature 392:402–405
9. Becker KF, Atkinson MJ, Reich U, Becker I, Nekarda H, Siewert JR, Höfler H (1994) E-cadherin gene mutations provide clues to diffuse type gastric carcinomas. Cancer Res 54:3845–3852
10. Tamura G, Yin J, Wang S, Fleisher AS, Zou T, Abraham JM, Kong D, Smolinski KN, Wilson KT, James SP, Silverberg SG, Nishizuka S, Terashima M, Motoyama T, Meltzer SJ (2000) E-cadherin gene promoter hypermethylation in primary human gastric carcinomas. J Natl Cancer Inst 92:569–573
11. Machado JC, Oliveira C, Carvalho R, Soares P, Berx G, Caldas C, Seruca R, Carneiro F, Sobrinho-Simões M (2001) E-cadherin gene (CDH1) promoter methylation as the second hit in sporadic diffuse gastric carcinoma. Oncogene 20:1525–1528
12. Blair V, Martin I, Shaw D, Winship I, Kerr D, Arnold J, Harawira P, McLeod M, Parry S, Charlton A, Findlay M, Cox B, Humar B, More H, Guilford P (2006) Hereditary diffuse gastric cancer: diagnosis and management. Clin Gastroenterol Hepatol 4:262–275
13. Humar B, Fukuzawa R, Blair V, Dunbier A, More H, Charlton A, , Kim WH, Reeve AE, Martin I, Guilford P (2007) Destabilized adhesion in the gastric proliferative zone and c-Src kinase activation mark the development of early diffuse gastric cancer. Cancer Res 67: 2480–2489
14. Le Borgne R, Bellaiche Y, Schweisguth F (2002) Drosophila E-cadherin regulates the orientation of asymmetric cell division in the sensory organ lineage. Curr Biol 12:95–104
15. Schluter MA, Pfarr CS, Pieczynski J, Whiteman EL, Hurd TW, Fan S, Liu CJ, Margolis B (2009) Trafficking of Crumbs3 during cytokinesis is crucial for lumen formation. Mol Biol Cell 20(22):4652–4663
16. Stehbens SJ, Akhmanova A, Yap AS (2009) Microtubules and cadherins: aneglected partnership. Front Biosci 14:3159–3167
17. Nasri S, Anjomshoaa A, Song S, Guilford P, McNoe L, Black M, Phillips V, Reeve A, Humar B (2010) Oligonucleotide array outperforms SNP array on formalin-fixed paraffin-embedded clinical samples. Cancer Genet Cytogenet 198(1):1–6
18. Cleveland WS (1979) Robust locally weighted regression and smoothing scatterplots. J Am Stat Assoc 74(368):829–836
19. van Dijk MC, Rombout PD, Boots-Sprenger SH, Straatman H, Bernsen MR, Ruiter DJ, Jeuken JW (2005) Multiplex ligation-dependent probe amplification for the detection of chromosomal gains and losses in formalin-fixed tissue. Diagn Mol Pathol 14(1):9–16
20. Sarai A, Mazur J, Nussinov R, Jernigan RL (1989) Sequence dependence of DNA conformational flexibility. Biochem 28(19): 7842–7849
21. Mishmar D, Rahat A, Scherer SW, Nyakatura G, Hinzmann B, Kohwi Y, Mandel-Gutfroind Y, Lee JR, Drescher B, Sas DE, Margalit H (1998) Molecular characterization of a common fragile site (FRA7H) on human chromosome 7 by the cloning of a simian virus 40 integration site. Proc Natl Acad Sci U S A 95:8141–8146
22. Humar B, Guilford P (2009) Hereditary diffuse gastric cancer: a manifestation of lost cell polarity. Cancer Sci 100(7):1151–1157
23. Mimata A, Fukamachi H, Eishi Y, Yuasa Y (2011) Loss of E-cadherin in mouse gastric epithelial cells induces signet ring-like cells, a possible precursor lesion of diffuse gastric cancer. Cancer Sci 102(5):942–950
24. Thoma CR, Toso A, Gutbrodt KL, Reggi SP, Frew IJ, Schraml P, Hergovich A, Moch H, Meraldi P, Krek W (2009) VHL loss causes spindle misorientation and chromosome instability. Nature Cell Biol 11(8):994–1001
25. Little SE, Vuononvirta R, Reis-Filho JS, Natrajan R, Iravani M, Fenwick K, Mackay A, Ashworth A, Pritchard-Jones K, Jones C (2006) Array CGH using whole genome amplification of fresh-frozen and formalin-fixed, paraffin-embedded tumor DNA. Genomics 87(2):298–306
26. Mc Sherry EA, Mc Goldrick A, Kay EW, Hopkins AM, Gallagher WM, Dervan PA (2007) Formalin-fixed paraffin-embedded clinical tissues show spurious copy number changes in array-CGH profiles. Clin Genet 72(5):441–447
27. Talseth-Palmer BA, Bowden NA, Hill A, Meldrum C, Scott RJ (2008) Whole genome amplification and its impact on CGH array profiles. BMC Res Notes 1(1):56
28. De Smith AJ, Tsalenko A, Sampas N, Scheffer A, Yamada NA, Tsang P, Ben-Dor A, Yakhini Z, Ellis RJ, Bruhn L, Laderman S (2007) Array CGH analysis of copy number variation identifies 1284 new genes variant in healthy white males: implications for association studies of complex diseases. Hum Mol Genet 16(23): 2783–2794
29. Brunet A, Armengol L, Heine D, Rosell J, García-Aragonés M, Gabau E, Estivill X, Guitart M (2009) BAC array CGH in patients with Velocardiofacial syndrome-like features reveals genomic aberrations on chromosome region 1q21. 1. BMC Med Genet 10(1):144
30. Ambros IM, Brunner B, Aigner G, Bedwell C, Beiske K, Bénard J, Bown N, Combaret V, Couturier J, Defferrari R, Gross N (2011) A multilocus technique for risk evaluation of patients with neuroblastoma. Clin Cancer Res 17(4):792–804
31. Shen Y, Wu BL (2009) Microarray-based genomic DNA profiling technologies in clinical molecular diagnostics. Clin Chem 55(4): 659–669
32. Hills A, Ahn JW, Donaghue C, Thomas H, Mann K, Ogilvie CM (2010) MLPA for confirmation of array CGH results and determination of inheritance. Mol Cytogen 3(1):19
33. Cook JG (2009) Replication licensing and the DNA damage checkpoint. Front Biosci (Landmark edition) 14:5013
34. Bakkenist CJ, Kastan MB (2003) DNA damage activates ATM through intermolecular autophosphorylation and dimer dissociation. Nature 421(6922):499–506
35. Abraham RT (2001) Cell cycle checkpoint signaling through the ATM and ATR kinases. Genes Dev 15(17):2177–2196
36. Kumagai A, Dunphy WG (2006) How cells activate ATR. Cell Cycle 5(12):1265–1268
37. Casper AM, Durkin SG, Arlt MF, Glover TW (2004) Chromosomal instability at common fragile sites in Seckel syndrome. Am J Hum Genet 75(4):654–660
38. Durkin SG, Glover TW (2007) Chromosome fragile sites. Annu Rev Genet 41:169–192
39. Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144(5):646–674