



CDH1 mutations in gastric cancer patients from northern Brazil identified by Next- Generation Sequencing (NGS)

Antonette El-Husny^{1,2}, Milene Raiol-Moraes^{1,2}, Marcos Amador^{1,2}, André M. Ribeiro-dos-Santos^{1,2}, André Montagnini³, Silvanira Barbosa⁴, Artur Silva^{2,4}, Paulo Assumpção^{2,5}, Geraldo Ishak^{2,5,6}, Sidney Santos^{1,2,5}, Pablo Pinto^{1,5}, Aline Cruz^{1,5} and Ândrea Ribeiro-dos-Santos^{1,2,5}

¹Laboratório de Genética Humana e Médica (LGHM), Instituto de Ciências Biológicas, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.

²Rede de Pesquisa em Genômica Populacional Humana, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.

³Instituto Sírio-Libanês de Ensino e Pesquisa, Hospital Sírio-Libanês, São Paulo, SP, Brazil.

⁴Laboratório de Polimorfismo de DNA - Instituto de Ciências Biológicas, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.

⁵Núcleo de Pesquisas em Oncologia, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.

⁶Hospital Universitário João de Barros Barreto, Universidade Federal do Pará (UFPA), Belém, PA, Brazil.

Abstract

Gastric cancer is considered to be the fifth highest incident tumor worldwide and the third leading cause of cancer deaths. Developing regions report a higher number of sporadic cases, but there are only a few local studies related to hereditary cases of gastric cancer in Brazil to confirm this fact. *CDH1* germline mutations have been described both in familial and sporadic cases, but there is only one recent molecular description of individuals from Brazil. In this study we performed Next Generation Sequencing (NGS) to assess *CDH1* germline mutations in individuals who match the clinical criteria for Hereditary Diffuse Gastric Cancer (HDGC), or who exhibit very early diagnosis of gastric cancer. Among five probands we detected *CDH1* germline mutations in two cases (40%). The mutation c.1023T > G was found in a HDGC family and the mutation c.1849G > A, which is nearly exclusive to African populations, was found in an early-onset case of gastric adenocarcinoma. The mutations described highlight the existence of gastric cancer cases caused by *CDH1* germline mutations in northern Brazil, although such information is frequently ignored due to the existence of a large number of environmental factors locally. Our report represent the first *CDH1* mutations in HDGC described from Brazil by an NGS platform.

Keywords: *CDH1*, germline mutations, HDGC, Gastric Cancer, NGS.

Received: November 30, 2014; Accepted: July 23, 2015.

Introduction

Gastric cancer was considered to be the fifth highest incident tumor in both genders worldwide in 2012 and the third leading cause of cancer deaths (8.8% of the total). Although East Asia represents the region of greatest gastric cancer mortality, high rates are also observed in both genders in Central and Eastern Europe and in Central and South America (Ferlay *et al.*, 2010). In Brazil, it was estimated that gastric cancer was the fourth most frequent in men and the fifth most common in women in 2014. In the northern region, gastric cancer is the second most common

in men and the third most common in women (INCA, 2014).

Corso *et al.* (2012) stated that developing regions of the world have a greater number of sporadic cases. Although only a few local studies have addressed this question, their results suggest that this statement is also applicable to Brazil. The gene related to the gastric cancer predisposition syndrome (Hereditary Diffuse Gastric Cancer - HDGC) is *CDH1* which encodes the E-cadherin protein, responsible for cell adhesion in non-neural epithelial cells, among other functions. Germline mutations of this gene as a cause of familial cases of diffuse gastric cancer were first described in the Maori tribe in New Zealand, in which the existence of a genetic syndrome predisposing to gastric cancer was first detected (Guilford *et al.*, 1998, 1999; Blair *et al.*, 2013).

Send correspondence to Ândrea Ribeiro-dos-Santos. Laboratório de Genética Humana e Médica, Instituto de Ciências Biológicas, Universidade Federal do Pará, Cidade Universitária Prof. José da Silveira Netto, 01, 66075-970 Belém, PA, Brazil. E-mail: akelyufpa@gmail.com or akely@ufpa.br.

To date, *CDHI* mutations have been described in various populations around the world. Noteworthy are the descriptions of mutations associated with carcinogenesis of different tumor types, such as breast cancer, prostate cancer, and gastric cancer (Richards *et al.*, 1999; Ikonen *et al.*, 2001; Masciari *et al.*, 2007; McVeigh *et al.*, 2014). Germline *CDHI* mutations have been described in approximately 30-40% of familial cases and in a smaller proportion of sporadic cases (Kaurah *et al.*, 2007; Garziera *et al.*, 2013;). Although several mutations have been detected in distinct families, no hotspot has been characterized. To date, there is only one molecular description of individuals from Brazil (Moreira-Nunes *et al.*, 2014).

Little has been discussed regarding the ancestral origin of pathogenic mutations described in cases of diffuse gastric cancer. However, a population approach is important when dealing with rare disorders like HDGC, because specific mutations observed in a population can guide the testing approach in other individuals of the same group.

Given the high incidence of gastric tumors in Brazil, specifically in the state of Pará, it is of great importance to focus the attention of clinicians and researchers on genetic factors potentially associated with gastric cancer in this population.

Subjects and Methods

Subjects

The study included six individuals of which only two were related (cases 1 and 3). Among the patients, four (cases 1, 3, 4 and 5) matched the clinical criteria for HDGC of the International Gastric Cancer Linkage Consortium (Fitzgerald *et al.*, 2010) and two (cases 2 and 6) had early onset (< 40 years) diffuse-type gastric cancer (Fitzgerald *et al.*, 2010; Kluijdt *et al.*, 2012). Five patients (cases 1 to 5) were from northern Brazil and one individual was from southeastern Brazil (case 6).

This study was approved by the Research Ethics Committee of Universidade Federal do Pará - Hospital João de Barros Barreto (protocol number 359.927), obeying the principles of the Declaration of Helsinki and Nuremberg Code. All individuals signed an Informed Consent form.

Genotyping analysis

DNA extraction was performed with the PureLink Genomics™ Mini Kit (Life Technologies, Foster City, CA, USA) according to the manufacturer's protocol. Amplification of the coding regions of the *CDHI* gene was performed by PCR, with a total of 20 amplicons per patient. The amplicons of the five index individuals (Table 1) were sequenced on an Ion Torrent PGM™ platform (Life Technologies).

Next Generation Sequencing (NGS)

Each of the 20 amplicons, including the promoter region of *CDHI* and its 16 exons (Table 1), were combined into a single equimolar pool with a total of 100 ng of DNA in a final volume of 35 µL.

The fragmentation of samples for 200 bp sequencing was performed with the ION shear plus reagent kit (Life Technologies) followed by purification with Agencourt™ AMPure Reagent XP™ (Agencourt Bioscience Corporation, Beverly, MA, USA). The connection of barcode adapters and repairs were made with the following kits: Ion Plus kit fragment library (Life Technologies) and Ion Xpress Barcode Adapters 1-16 kits (Life Technologies) in each one of the individual samples.

After purification, the samples passed through size selection in an e-Gel size selection 2% agarose gel, from which a band of 200 base pairs was retrieved. The library was then amplified, purified and assembled in the same pool concentration for emulsion PCR with the Ion PGM 200 template reagent kit (Life Technologies).

The sequencing reaction was performed with the ION sequencing reagent kit (Life Technologies). The resulting data was aligned to the reference genome hg19 (available at <http://genome.ucsc.edu>) and the mutations were identified using the GATK v.2.6. Toolkit. The variants were filtered by low quality calling (less than 50X depth and homopolymer runs) and analyzed using the Integrative Genomics Viewer software (IGV v.2.3) (Broad Institute; <https://www.broadinstitute.org/igv>).

Validation

After identifying the familial mutation (index individuals), first-degree relatives were screened for the specific mutation (Figure 1) by Sanger sequencing. Although being a limited analyses for clinical purposes, every missense mutation detected had its potential pathogenicity tested using Polyphen-2 (Adzhubei *et al.*, 2010) and SIFT (Ng and Henikoff, 2003) for predicting functional effects of human nsSNPs.

The main detected mutations were investigated by Sanger sequencing of 100 samples from the local population of Belém, PA, Brazil, for comparative purposes. Additionally, all populational data for mutations were checked using the 1,000 genomes project data (McVean *et al.*, 2012).

For each exon, the sequencing reaction was performed with 1 µL of purified PCR product of each exon, 0.5 µL of the reverse specific primer, 0.5 µL of Big Dye Terminator V3.1 Cycle Sequencing Kit (Life Technologies), and 3.0 µL of SaveMoney buffer to a final volume of 5 µL. The thermocycling reaction proceeded as follows: 95 °C for 2 min, followed by 35 cycles of 95 °C for 30 s, 60 °C for 20 s and 72 °C for 2 min.

Table 1 - Polymerase chain reaction primers used for amplification of *CDH1* promoter region and its 16 exons, their product size and flanking regions size.

Region	Forward and reverse primers	Product size (bp)	Flanking regions size (bp)
Promoter	5' GAGAACTCAGTAAAGGGGCTGA 3' 5' ACTAAGACCTGGGATCAGAAAGG 3'	853	5' - 108 -
Exon 1	5' CCATCTCCAAAACGAACAAAC 3' 5' GAACTTTCTTGGAAGAAGGGAAG 3'	752	- 3' - 113
Exon 2	5' CTAGTCTTGAGGGGGTGACT 3' 5' GTAAATCCAAGGGGTGTCGT 3'	486	5' - 236 3' - 135
Exon 3	5' GTAAATCCAAGGGGTGTCGT 3' 5' CAACCCCTACAACACAAAATCA 3'	421	5' - 80 3' - 117
Exon 4	5' TCAAACGTACTGCCCACA 3' 5' ATCCCAACTGGGCTTTTC 3'	347	5' - 117 3' - 86
Exon 5	5' TCTGTTTCTCTGGGAGGGATT 3' 5' TCAAGTTAAGCTCCTCATGTGTTTC 3'	383	5' - 111 3' - 106
Exon 6	5' GTCACCTCACTGGTTCTTTC 3' 5' CCGTAGGAAGGATCAGCTTAGT 3'	280	5' - 22 3' - 111
Exon 7	5' TTCTTTCTCCCTAGCACTTTG 3' 5' ACAACTGGCCTAGCAGGATT 3'	436	5' - 169 3' - 91
Exon 8	5' CTTGGTTGTGTCGATCTCTCTG 3' 5' GACCTTTCTTTGAAAACCTCT 3'	194	5' - 103 3' - 40
Exon 9	5' ATGATCGCTCAAATACACTCCA 3' 5' CTGCCAAAGCGAATCTACTTCT 3'	429	5' - 148 3' - 99
Exon 10	5' CATTGAAAGTCATGGCAGAAAC 3' 5' GCTGCAAGTCAGTTGAAAAATC 3'	420	5' - 142 3' - 33
Exon 11	5' GCTTAAGCCGTTTTCAGCTACA 3' 5' AACTCTCCCTCCAAAAGAAGG 3'	303	5' - 70 3' - 87
Exon 12	5' CTAGACTTGGTCTGGTGGAAAGG 3' 5' GGAAGCAAGTATCAATGGAAGG 3'	430	5' - 79 3' - 126
Exon 13	5' AAGCAGCTCTGCTCTTTCACT 3' 5' CTCTTTCCACATCAGCTAACC 3'	470	5' - 122 3' - 120
Exon 14	5' TCTGTGATAGCTGCTGCTTCTG 3' 5' AGCTGTTTCAAATGCCTACCTCT 3'	294	5' - 75 3' - 88
Exon 15	5' AAGGCATCATCCAACCATAATC 3' 5' TTTTGGACACAACCTCCTCTGA 3'	311	5' - 100 3' - 67
Exon 16.1	5' AAGTCTGGGTGCATTGTCGTA 3' 5' AGCTGACTTCTCCCTTCTTTT 3'	690	5' - 110 -
Exon 16.2	5' CAGCACCTGCAGATTTTCTTA 3' 5' CTAGTCAAGATGTGGCCAGACA 3'	840	- -
Exon 16.3	5' CAGTTGCTTTGCCCAAGATAG 3' 5' TAGCTTGAAGTCCGAAAAATC 3'	817	- -
Exon 16.4	5' GGTAGTGAGGATCTTGATTTGGA 3' 5' CCTCTTCTCCACGTTTGTACT 3'	398	- 3' - 90

After thermocycling, the product was prepared for sequencing in an ABI 3130 automatic sequencer (Life Technologies). The sequence information was interpreted by ABI Analysis SoftwareTM. The

electropherograms were analyzed using the ChromasPro1.49 software and compared with the reference sequence obtained from GenBank (NM_004360.3).

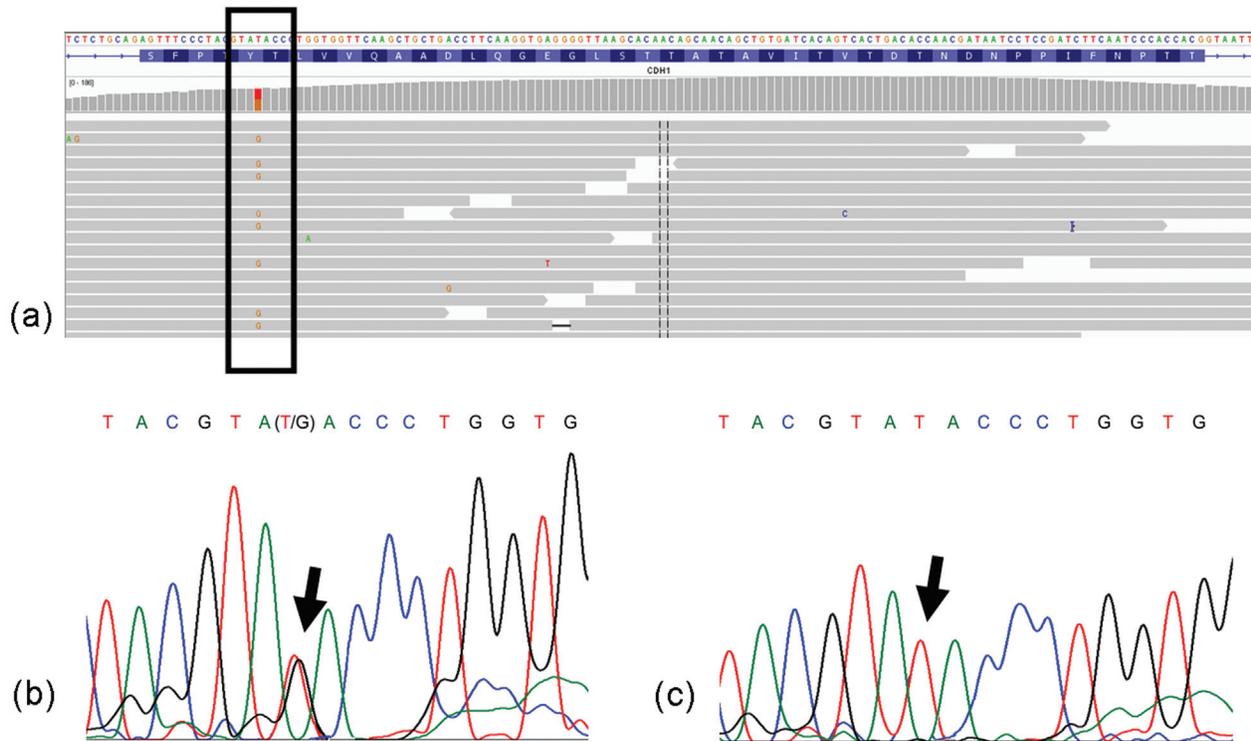


Figure 1 - Molecular analysis presenting exon 8 *CDH1* mutation, c.1023T > G. (a) Integrative Genomics Viewer - IGVTM software result; (b) Sanger sequencing analysis of a patient; (c) Sanger sequencing analysis of a control case.

Analysis of genetic ancestry

All five index cases had their genetic ancestry tests performed using 48 INDEL informative markers of ancestry (American Indians, Europeans and Africans), following the methodology previously described by Santos *et al.* (2010). Three multiplex PCR reactions were performed, each with 16 markers, followed by electrophoresis on an ABI-PRISM 3130 sequencer and analysis using GeneMapper ID v. 3.2 (Life Technologies). The individual proportions of European, African and Amerindian genetic ancestry were estimated using the STRUCTURE software v.2.3.3, assuming three parental populations (Europeans, Africans and Amerindians) and running with a 200,000 burn-in period and 200,000 Markov Chain Monte Carlo repetitions after burning.

Results

Results of the molecular analyses of the *CDH1* gene are summarized in Table 2. Among five probands, we detected *CDH1* germline mutations in two cases (40%). Case 1 exhibited a heterozygous *CDH1* exon 8 germline mutation c.1023T > G. This patient is member of a classic HDGC family. Case 3, who belongs to the same family, confirmed the presence of the c.1023T > G mutation.

Figure 1a presents the molecular analysis of the *CDH1* germline mutation c.1023T > G by Integrative Ge-

nomics Viewer (IGV v.2.5) software, which was validated by Sanger sequencing (index case 1; Figure 1b). In a sample of 100 individuals from the local population of Belém, PA (Brazil) no instances of the mutation were found (Figure 1c). This mutation was identified as familial by analyzing first-degree relatives of cases 1 and 3 by Sanger sequencing. The family pedigree with 46 individuals is presented in Figure 2. Ten of the family members were tested for the c.1023T > G mutation and all exhibited the heterozygous mutation.

The molecular analysis of index case 2 revealed the *CDH1* germline mutation exon 12 c.1849G > A in heterozygosis (Table 2), which was confirmed by Sanger sequencing. This patient exhibited early-onset gastric adenocarcinoma (by 28 years of age) without any other similar case in the family. Figure 3a presents the molecular analysis for the *CDH1* germline mutation c.1849G > A by IGV v.2.3. Similar to index case 1, this mutation was not identified in the sample from local population (Figure 3c).

Previous population studies performed by the 1000 Genomes Project (McVean *et al.*, 2012) did not detect the c.1023T > G mutation, but described the c.1849A mutation as almost exclusively African, with an allele frequency of 0.045 and overall database frequency (MAF) of 0.01.

NGS alignment presented two INDEL variations with good quality (more than 50X depth) in heterozygous state among all individuals: c.1649delG and c.2218delC located

Table 2 - Germline mutations and polymorphisms detected in CDH1 exons in hereditary diffuse gastric cancer in Brazil.

Case	Exon	CDH1 mutation	Type of mutation	Prediction	Ancestry panel		Previous References		Origin
					EUR	AME	AFR	Tumor type	
Case 1*	8	c.1023T > G / p.Tyr341*	Nonsense	Pathogenic	0,517	0,334	0,149	Gastric cancer (McVean et al., 2012)	Unavailable
	13	rs1801552	Silent	-	-	-	-	Populational studies (Adzhubai et al., 2010)	Unavailable
Case 2	12	c.2076T > C / p.Ala692=	Missense	Benign** (SIFT score 0.19) (PSIC score 0.04)	0,520	0,167	0,313	Endometrial cancer (Risinger et al., 2010)	Unavailable
		Sporadic diffuse gastric cancer (Guilford et al., 2010)						African-American	
	13	c.1849G > A / p.Ala617Thr	Silent	-	-	-	-	Early onset diffuse gastric cancer (Risinger et al., 1994)	African-American
		Populational studies (Adzhubai et al., 2010)						Unavailable	
Case 3*	8	c.1023T > G / p.Tyr341*	Nonsense	Pathogenic	0,391	0,385	0,224	Gastric cancer (McVean et al., 2012)	Unavailable
	13	rs1801552	Silent	-	-	-	-	Populational studies (Adzhubai et al., 2010)	Unavailable
Case 4	13	c.2076T > C / p.Ala692=	Silent	-	0,392	0,356	0,252	Populational studies (Adzhubai et al., 2010)	Unavailable
		Populational studies (Adzhubai et al., 2010)						Unavailable	
Case 5	13	c.2076T > C / p.Ala692=	Silent	-	0,511	0,399	0,090	Populational studies (Adzhubai et al., 2010)	Unavailable
		Populational studies (Adzhubai et al., 2010)						Unavailable	
Case 6	13	c.2634C > T / p.Gly878=	Silent	-	0,656	0,186	0,158	Populational studies (Adzhubai et al., 2010)	Unavailable
		c.2076T > C / p.Ala692=						Unavailable	

• Individuals from the same family, ** Pathogenicity prediction evaluated by SIFT (sorting intolerant from tolerant) and Polyphen2 (PSIC) softwares.

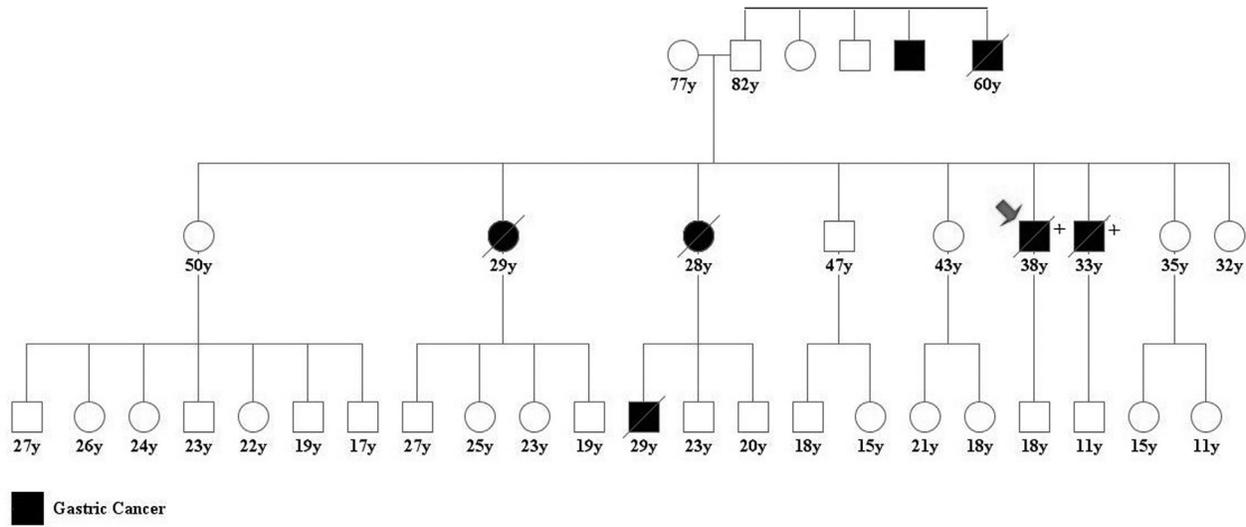


Figure 2 - Pedigree of the *HDGC* Northern Brazilian family described in this paper, red arrow showing the index case (case 1); (+) represents individuals with molecular analysis showing the *CDH1* c.1023T > G mutation.

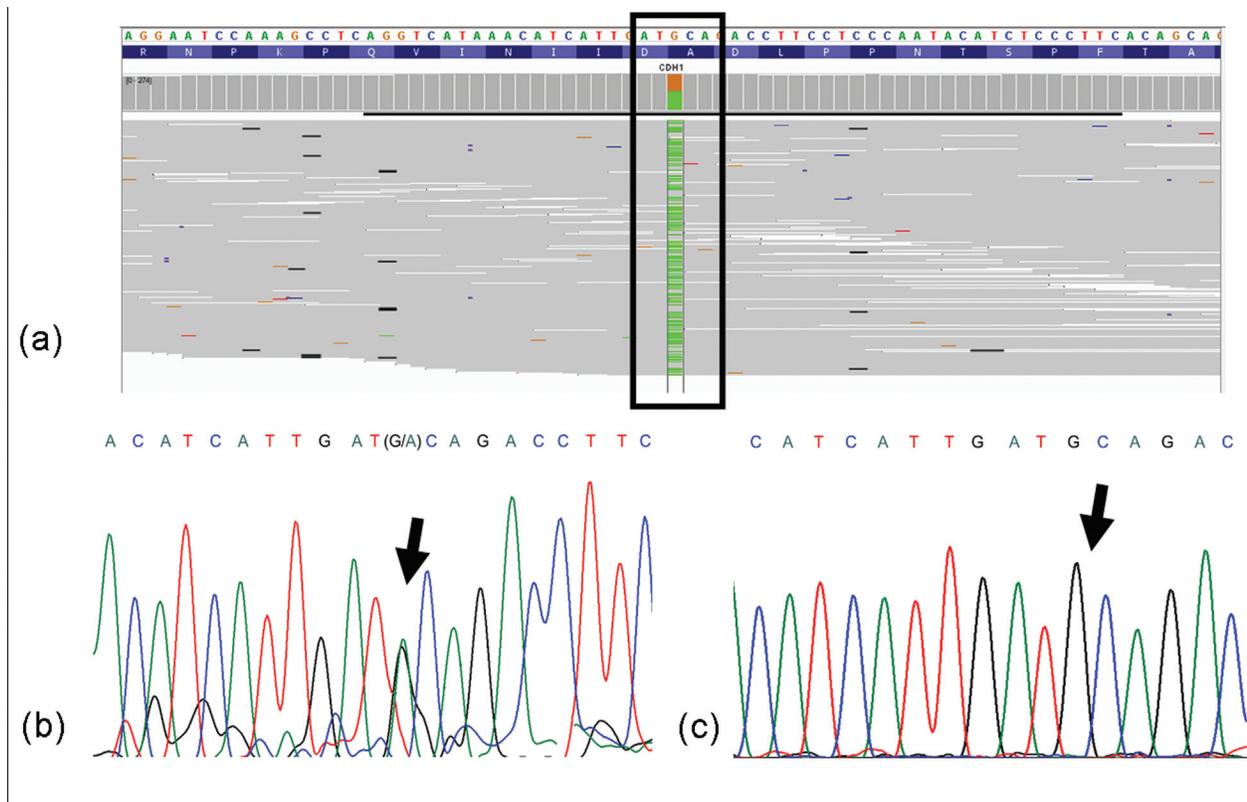


Figure 3 - Molecular analysis presenting exon 12 *CDH1* mutation, c.1849T > G. (a) Integrative Genomics Viewer - IGV™ software result; (b) Sanger sequencing analysis of the patient; (c) Sanger sequencing analysis of a control case.

in exon 11 and exon 14, respectively. Such kind of frame shift mutations would cause drastic damage for the final protein, and to us they seem to be incompatible with the biological and epidemiological background of gastric cancer.

These variants probably represent sequencing errors of the platform chemistry, since they were found in highly repetitive regions. Thus, we choose to validate all our findings by Sanger sequencing.

No pathogenic *CDHI* germline mutation was confirmed in the index cases 4, 5 and 6, and only polymorphisms were observed (Tables 2 and 3).

When analyzing the genetic ancestry contribution, index case 1 exhibited a 52% European, 33% Amerindian and 15% African contribution. Index case 2 exhibited a 52% European, 17% Amerindian and 31% African contribution. These results are presented in Table 2.

Besides the exonic alterations, NGS was also able to detect intronic modifications. The ones already described as polymorphisms and registered in NCBI dbSNP are shown in Table 3.

Discussion

The *CDHI* germline mutation of index case 1 (c.1023T > G) had been previously described in New Zealand patients, with only few clinical details available (Guilford *et al.*, 2010). It is responsible for the introduction of a premature stop codon at position 341 of the E-cadherin (p.Y341*) protein and is therefore pathogenic. This region encodes the second cadherin domain that is normally located in the extracellular portion of the protein and is essential for its juxtacellular adhesion function. Significant levels of mature protein cease to be translated in the presence of this mutation. Given the typ-

Table 3 - Known polymorphisms detected in *CDHI* introns and flanking regions.

CDHI region	Cases					
	Case 1*	Case2	Case 3*	Case 4	Case 5	Case 6
5' flanking region						
rs7194355 (C > A)	wt	A/A	wt	C/A	C/A	wt
rs35582463 (C > T)	C/T	wt	C/T	wt	wt	wt
rs33945903 (C > T)	C/T	wt	C/T	wt	wt	wt
rs5030625 (GA > G)	G/G	G/G	G/G	G/G	G/G	GA/G
rs3395334 (C > A)	C/A	wt	C/A	wt	wt	wt
rs16260 (C > A)	wt	A/A	A/A	C/A	C/A	wt
Intron 1						
rs3743674 (C > T)	T/T	T/T	*	T/T	T/T	wt
rs147838237 (C > CGCCCCAGCCCCGT)	hoz	hoz	*	hoz	hoz	wt
rs286579983 (T > C)	T/C	T/C	*	T/C	T/C	wt
rs12928281 (C > T)	C/T	C/T	*	C/T	C/T	C/C
Intron 6						
rs8059669 (A > C)	*	wt	A/C	wt	*	A/C
Intron 7						
rs34374107 (T > C)	wt	wt	wt	wt	wt	wt
Intron 9						
rs35423758 (C > T)	wt	wt	wt	C/T	wt	wt
rs339509003 (G > C)	wt	wt	wt	wt	wt	wt
Intron 12						
rs2276330 (T > C)	T/C	wt	wt	wt	T/C	wt
3'UTR						
rs1801026 (C > T)	C/T	wt	wt	wt	C/T	wt
rs8049282 (C > T)	C/T	wt	C/T	C/T	wt	wt
rs33956791 (C > T)	wt	wt	wt	wt	wt	wt
rs9282653 (G > A)	wt	wt	wt	wt	wt	wt
rs13689 (T > C)	T/C	wt	wt	wt	wt	wt
3' flanking region						
rs8045438 (A > G)	G/G	G/G	G/G	G/G	G/G	G/G
rs181705992 (T > A)	wt	wt	wt	wt	wt	wt
rs17690554 (C > G)	C/G	wt	wt	wt	wt	wt

wt, wildtype; *, no call; hoz, mutant homozygous

ical genetic admixture in Brazil, evidenced by studies of population ancestry (*e.g.* Santos *et al.*, 2010), it was not possible to associate ethnicity with the presence of the c.1023T > G mutation. However, the absence of this mutation in 100 individuals of the local population (Belém, PA, Brazil) and in the 1,000 genome project dataset (McVean *et al.*, 2012) puts in evidence that the mutation must be a rare mutation rather than a polymorphic variant.

Case 2 exhibited the c.1849G > A mutation previously detected by Risinger *et al.* (1994) in a tissue sample of endometrial cancer, and by Ascano *et al.* (2001) in patients with diffuse gastric cancer. Subsequently, Suriano *et al.* (2003) demonstrated the functional inactivation *in vitro* of c.1849G > A in cases of early onset gastric cancer like the present case. Supported by these studies, the pathogenicity of this mutation is described in the NCBI SNP database (<http://www.ncbi.nlm.nih.gov/projects/SNP>) under the code rs33935154.

However, analysis performed by Polyphen-2 (Adzhubei *et al.*, 2010) and SIFT (Ng and Henikoff, 2003) software suggest low pathogenicity for this mutation (PSIC score 0.04; SIFT score 0.19), since different species present substitutions in the same protein position (p.617). Population studies performed in the 1,000 genomes project (McVean *et al.*, 2012) described the overall frequency (MAF) of allele c.1849A as 0.01. In the same database, the presence of this mutation is almost exclusive of African populations, with an allele frequency increasing to 0.045.

This divergent information can be explained under at least four hypotheses: (i) the mutation has incomplete penetrance; (ii) the mutation is not truly pathogenic, given the high frequency in African populations, where there is no significant increase in the HDGC case number or diffuse type gastric cancer at a young age; although scientific data on this subject are minimal; (iii) the mutation is pathogenic for other population groups but not among Africans, given its local frequency of 4.5%; (iv) the mutation is only pathogenic in the presence of other genetic and/or epigenetic factors not yet studied. As there is no study testing this mutation in African patients with diffuse gastric cancer, all hypotheses above must be considered.

In case 2, it is likely that the mutation occurs as an effect of the patient's African ancestry contribution of 31%. Interestingly, Suriano *et al.* (2003) described the c.1849A mutation in two African-American-unrelated cases. The geographical origin of these two cases was not available.

Regarding the polymorphisms described in this paper, we highlight the exonic rs1801552 and the intronic: rs13689, rs16260 and rs17690554, already analyzed by Zhan *et al.* (2012) in a case-control study without any difference detected between genotypes in gastric cancer patients and control group.

Previous studies on rs16260 and also on rs1801026 have called attention to the possibility of being markers for

genetic susceptibility to cancer (Wang *et al.*, 2007; Li *et al.*, 2011). But conflicting results were shown when comparing different populations (Li *et al.*, 2012). Additional population studies of Brazilian subjects from different geographic regions of the country should be performed to find out whether these polymorphisms can provide useful susceptibility information in our country.

Recent studies of *CDHI* with NGS focused mostly on hereditary breast cancer (Castera *et al.*, 2014; Yang *et al.*, 2015). Dang *et al.* (2014) studied *CDHI* mutations in gastric cancer tissues with NGS, aiming to clarify their particular pathogenesis. All of these studies were performed on different sequencing platforms than the one used here, and concluded that NGS technology was an excellent method for their investigations.

Regarding the mutations c.1649delG and c.2218delC detected by NGS in all cases, it is unlikely that they represent low mosaicism with clinical significance, as they presented large read depth and calling quality in every case analyzed. Rather, they seem to be due to a sequencing error of a highly repetitive region of the genome. Although NGS techniques generate highly reliable data, they still can produce several miss calling due to specifics of their chemistry and software limitations. Thus, a critical analysis is necessary when interpreting variant reports, and suspicious results must be double checked by techniques such as Sanger sequencing or others.

Furthermore, although three of the six index cases presented no pathogenic mutation, the proportion of diffuse gastric cancer cases with mutations detected is similar to previously reports in the literature (40%; Kaurah *et al.*, 2007; Garziera *et al.*, 2013). Quite possibly, other genes or epigenetics factors may be the cause behind these undefined cases.

Conclusion

The mutations described in this paper demonstrate the existence of gastric cancer cases caused by *CDHI* germline mutations in an endemic region of gastric cancer (northern Brazil), and such information is frequently ignored due to the significant number of environmental factors present.

The presence of the c.1849G > A mutation, a mutation almost African-exclusive, demonstrates the importance of considering ancestry and ethnicity when studying genetic disorders.

These *CDHI* germline mutations (c.1849G > A; c.1023T > G) are the first described in association with HDGC and early onset gastric cancer from Brazil revealed by a Next-Generation Sequencing platform. Larger studies that examine the frequency of gastric cancer cases associated with an abnormal *E-cadherin* gene will be of great value to determine the true importance of this genetic factor for gastric cancer in this area.

Acknowledgments

The authors thank the patients from NPO (Núcleo de Pesquisas em Oncologia) and HUIBB/UFPA (Hospital Universitário João de Barros Barreto - Universidade Federal do Pará) for the participation of this study. Thank you Aline Cruz and Pablo Pinto for help. This work was part of the REDE DE PESQUISA EM GENÔMICA POPULACIONAL HUMANA (RPGPH - BioComputacional/CAPEs). Financial support was provided by PROPESP/UFPA-FADESP, MS/Decit, CNPq, FAPESPA, SESP and CAPEs. ASEH was supported by Programa de Bolsa Especial para Doutorado em Pesquisa Médica (PBE-DPM/CAPEs); SS and ARDS by CNPq/Produtividade grants.

References

- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS and Sunyaev SR (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7:248-249.
- Ascano JJ, Frierson H, Moskaluk CA, Harper JC, Roviello F, Jackson CE, El-Rifai W, Vindigni C, Tosi P and Powell, SM (2001) Inactivation of the E-cadherin gene in sporadic diffuse-type gastric cancer. *Mod Pathol* 14:942-949.
- Blair V, Kahokehr A and Sammour T (2013) Cancer in Māori: Lessons from prostate, colorectal and gastric cancer and progress in hereditary gastric cancer in New Zealand. *ANZ J Surg* 83:42-48.
- Castéra L, Krieger S, Rousselin A, Legros A, Baumann J-J, Bruet O, Brault B, Fouillet R, Goardon N, Letac O, *et al.* (2014) Next-generation sequencing for the diagnosis of hereditary breast and ovarian cancer using genomic capture targeting multiple candidate genes. *Eur J Hum Genet* 22:1305-1313.
- Corso G, Marrelli D, Pascale V, Vindigni C and Roviello, F (2012) Frequency of CDH1 germline mutations in gastric carcinoma coming from high- and low-risk areas: Meta-analysis and systematic review of the literature. *BMC Cancer* 12:e8.
- Dang Y, Wang YC and Huang QJ (2014) Microarray and next-generation sequencing to analyse gastric cancer. *Asian Pac J Cancer Prev* 15:8033-8039.
- Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Fitzgerald RC, Hardwick R, *et al.* (2010) Hereditary diffuse gastric cancer: Updated consensus guidelines for clinical management and directions for future research. *J Med Genet* 47:436-444.
- Garziera M, De Re V, Geremia S, Seruca R, Figueiredo J, Melo S, Simões-Correia J, Caggiari L, De Zorzi M, Canzonieri V, *et al.* (2013) A novel CDH1 germline missense mutation in a sporadic gastric cancer patient in north-east of Italy. *Clin Exp Med* 13:149-157.
- Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, Taite H, Scoular R, Miller A and Reeve AE (1998) E-cadherin germline mutations in familial gastric cancer. *Nature* 392:402-405.
- Guilford P, Hopkins JBW, Grady WM, Markowitz SD, Willis J, Lynch H, Rajput A, Wiesner GL, Lindor NM, Burgart LJ, *et al.* (1999) E-cadherin germline mutations define an inherited cancer syndrome dominated by diffuse gastric cancer. *Hum Mutat* 14:249-255.
- Guilford P, Humar B and Blair V (2010) Hereditary diffuse gastric cancer: Translation of CDH1 germline mutations into clinical practice. *Gastric Cancer* 13:1-10.
- Ikonen T, Matikainen M, Mononen N, Hyytinen E, Helin HJ and Tommola S (2001) Association of E-cadherin germ-line alterations with prostate cancer. *Clin Cancer Res* 7:3465-3471.
- Instituto Nacional de Câncer José Alencar Gomes da Silva (2014) Estimativa 2014: Incidência de Câncer no Brasil / Instituto Nacional de Câncer José Alencar Gomes da Silva, Coordenação de Prevenção e Vigilância. INCA, Rio de Janeiro, 126 pp.
- Kaurah P, MacMillan A and Boyd N (2007) Founder and recurrent CDH1 mutations in families with hereditary diffuse gastric cancer. *JAMA* 297:2360-2372.
- Kluijt I, Sijmons RH, Hoogerbrugge N, Plukker JT, de Jong D, van Krieken JH, van Hillegersberg R, Ligtenberg M, Bleiker E and Cats A (2012) Familial gastric cancer: Guidelines for diagnosis, treatment and periodic surveillance. *Familial Cancer* 11:363-369.
- Li Y, Tang Y, Zhou R, Sun D, Duan YA, Wang N, Chen ZF and Shen N (2011) Genetic polymorphisms in the 3'-untranslated region of the E-cadherin gene is associated with risk of different cancers. *Mol Carcinog* 50:857-862.
- Li YL, Tian Z, Zhang JB and Fu BY (2012) CDH1 promoter polymorphism and gastric cancer susceptibility. *Mol Biol Rep* 39:1283-1286.
- Masciari S, Larsson N, Senz J, Boyd N, Kaurah P, Kandel MJ, Harris LN, Pinheiro HC, Troussard A, Miron P, *et al.* (2007) Germline E-cadherin mutations in familial lobular breast cancer. *J Med Genet* 44:726-731.
- McVean GA, Donnelly P, Lunter G, Marchini JL, Myers S, Gupta-Hinch A, Iqbal Z, Mathieson I, Rimmer A, Xifara DK, *et al.* (2012) An integrated map of genetic variation from 1,092 human genomes. *Nature* 491:56-65.
- McVeigh T, Choi JK, Miller NM, Green AJ and Kerin MJ (2014) Lobular breast cancer in a CDH1 splice site mutation carrier: Case report and review of the literature. *Clin Breast Cancer* 14:47-51.
- Moreira-Nunes CA, Barros MB, Borges BN, Montenegro RC, Lamarão LM, Ribeiro HF, Bona AB, Assumpção PP, Rey JA, Pinto GR, *et al.* (2014) Genetic screening analysis of patients with hereditary diffuse gastric cancer from northern and northeastern Brazil. *Hered Cancer Clin Pract* 12:e18.
- Ng PC and Henikoff S (2003) SIFT-predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31:3812-3814.
- Richards FM, McKee SA, Rajpar MH, Cole TR, Evans DG, Jankowski JA, McKeown C, Sanders DS and Maher ER (1999) Germline E-cadherin gene (CDH1) mutations predispose to familial gastric cancer and colorectal cancer. *Hum Mol Genet* 8:607-610.
- Risinger J, Berchuck A, Kohler M and Boyd J (1994) Mutations of the E-cadherin gene in human gynecologic cancers. *Nat Genet* 7:98-102.
- Santos NPC, Ribeiro-Rodrigues E, Ribeiro-dos-Santos A, Pereira R, Gusmão L, Amorim A, Guerreiro JF, Zago MA, Matte C, Hutz MH, *et al.* (2010) Assessing individual interethnic admixture and population substructure using a 48-insertion-

deletion (INSEL) ancestry-informative marker (AIM) panel. *Hum Mutat* 31:184-190.

Suriano G, Oliveira C, Ferreira P, Machado J, Bordin M, De Wever O, Bruyneel EA, Moguelevsky N, Grehan N, Porter TR, *et al.* (2003) Identification of CDH1 germline missense mutations associated with functional inactivation of the E-cadherin protein in young gastric cancer probands. *Hum Mol Genet* 12:575-582.

Wang GY, Lu CQ, Zhang RM, Hu XH and Luo ZW (2007) The E-cadherin gene polymorphism -160C → A and cancer risk: A HuGe review and meta-analysis of 26 case-control studies. *Am J Epidemiol* 167:7-14.

Yang X, Wu J, Lu J, Liu G, Di G, Chen C, Hou Y, Sun M, Yang W, Xu X, *et al.* (2015) Identification of a comprehensive spectrum of genetic factors for hereditary breast cancer in a

Chinese population by next-generation sequencing. *PLoS One* 10:e0125571.

Zhan Z, Wu J, Zhang J, Yang Y, Ton S, Zhang C, Li J, Yang X and Dong W (2012) CDH1 gene polymorphisms, plasma CDH1 levels and risk of gastric cancer in a Chinese population. *Mol Biol Rep* 39:8107-8113.

Internet Resources

Forman D and Bray F (2012) Globocan v1.0 Cancer Incidence and Mortality Worldwide - IARC CancerBase No 11. International Agency for Research on Cancer-France, Lyon, <http://globocan.iarc.fr> (accessed Febr 10, 2014).

Associate Editor: Maria Isabel Achatz

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License (type CC-BY), which permits unrestricted use, distribution and reproduction in any medium, provided the original article is properly cited.