Identification of germline mutations in the cancer predisposing gene CDH1 in patients with orofacial clefts

Ingrid P. Vogelaar1,†, Joana Figueiredo5,6,‡, Iris A.L.M. van Rooij2, Joana Simões-Correia5,7, Rachel S. van der Post3, Soraia Melo5, Raquel Seruca5,6, Carine E.L. Carels4, Marjolijn J.L. Ligtenberg1,3,† and Nicoline Hoogerbrugge1,*,

1Department of Human Genetics, 2Department of Epidemiology, Biostatistics and HTA, 3Department of Pathology and 4Department of Orthodontics and Craniofacial Biology, Radboud University Nijmegen Medical Centre, PO Box 9101, Nijmegen 6500 HB, The Netherlands, 5IPATIMUP—Institute of Molecular Pathology and Immunology of the University of Porto, Porto 4200-465, Portugal, 6Faculty of Medicine, University of Porto, Porto 4200-319, Portugal and 7Centre of Ophthalmology and Vision Sciences—IBILI, Coimbra 3000-548, Portugal

Received October 2, 2012; Revised and Accepted November 21, 2012

Orofacial clefts (OFC) are among the most common birth defects worldwide. The etiology of non-syndromic OFC is still largely unknown. During embryonic development, the cell adhesion molecule E-cadherin, encoded by CDH1, is highly expressed in the median edge epithelium of the palate. Furthermore, in multiple families with CDH1 mutations, OFC cases are observed. To determine whether CDH1 is a causative gene for non-syndromic OFC and to assess whether CDH1 mutation screening in non-syndromic OFC patients enables identification of families at risk of cancer, direct sequencing of the full coding sequence of CDH1 was performed in a cohort of 81 children with non-syndromic OFC. Eleven children had heterozygous CDH1 sequence variants, 5 cases with 4 distinct missense mutations and 8 cases with 4 intronic variants. Using a combination of in silico predictions and in vitro functional assays, three missense mutations in four non-syndromic OFC patients were predicted to be damaging to E-cadherin protein function. The intronic variants including one tested in an in vitro assay appeared to be benign, showing no influence on splicing. Functionally relevant heterozygous CDH1 missense mutations were found in 4 out of 81 (5%) patients with non-syndromic OFC. This finding opens a new pathway to reveal the molecular basis of non-syndromic OFC. Cancer risk among carriers of these mutations needs to be defined.

INTRODUCTION

Orofacial clefts (OFC) are among the most common birth defects worldwide (1). They have a major clinical impact requiring surgical, dental, orthodontic, speech, hearing and psychologic treatments. Two types of OFC exist, cleft lip with or without cleft palate (CLP) and cleft palate only (CP) (1). Recently, advances have been made in the identification of causative genetic mutations underlying syndromic forms of OFC. The etiology of non-syndromic OFC, however, is still largely unknown (2). A meta-analysis of genome-wide association studies has confirmed and identified several loci that are implicated in the etiology of non-syndromic OFC, but the specific causal variants on these loci remain to be identified (3). Determining the relative risk of non-syndromic OFC based on genetic background will be useful for genetic counseling and development of future preventive measures.

The fusion of the palate is the result of complex cell–cell interactions that are controlled by structural and regulatory proteins (4). The CDH1 gene encodes the cell adhesion molecule E-cadherin, which is essential for the establishment of intercellular junction complexes and required for the adhesive
properties of epithelial cells (5). E-cadherin is highly expressed in the frontonasal prominence in the fourth and fifth week and in the lateral and medial nasal prominences in the sixth week of human embryogenesis, showing that it is expressed during the critical stages of lip and palate development (6).

Inactivating germline mutations in CDH1 are rare and tightly associated with hereditary diffuse gastric cancer (HDGC, OMIM: 137215), an autosomal dominant cancer-predisposition syndrome (7). HDGC is characterized by the presence of invasive diffuse signet ring cells and accounts for 1–3% of all gastric cancers (8). Carriers of a CDH1 mutation are counseled to have a high lifetime risk of developing diffuse gastric cancer (DGC) of up to 80% (9,10). In addition to the risk of DGC, female CDH1 mutation carriers are counseled to have a 60% lifetime risk of developing lobular breast cancer (LBC) (10). Due to the lack of effective surveillance programs, the absence of preneoplastic lesions, and the highly invasive behavior of this type of cancer, prophylactic gastrectomy is currently advised to adult carriers of mutations in CDH1 to eliminate the risk of lethality in this gastric cancer (GC) syndrome (11–14).

Contradictory results have been published on the risk of cancer in individuals with OFC (15–19). An overrepresentation of OFC patients was observed in CDH1 mutation carriers (6). OFC were reported in two HDGC families with germline CDH1 mutations, and in one family, CDH1 mutations cosegregated with DGC and OFC (6). The association between OFC and CDH1 was later supported by Kluijt et al. (20), who reported the occurrence of OFC in seven individuals from three HDGC families carrying a CDH1 mutation. Functionally relevant mutations in CDH1 have not yet been described in non-syndromic OFC patients, although non-coding polymorphisms at the CDH1 locus were found in such patients (21,22).

In conclusion, the clinical and genetic evidence suggests that CDH1 is a susceptibility gene for non-syndromic OFC. Thus, we decided to study, for the first time, the relevance of germline CDH1 mutations in a cohort of 81 children with non-syndromic OFC with unknown family history of HDGC.

RESULTS

CDH1 mutation analysis and in silico analyses

In the current study, 81 patients with non-syndromic OFC were screened for CDH1 mutations. Eleven patients presented heterozygous CDH1 sequence variants. Four distinct missense and four intronic variants were identified, which are shown in Tables 1 and 2, respectively.

All sequence variants were initially studied using in silico predictions. Afterwards, the ones with a putative impact on E-cadherin function (three missense and one intronic variant) were further analyzed using various in vitro functional assays.

The variant c.88C>A (p.Pro30Thr) was detected in two CLP patients. In silico predictions of the impact of the p.Pro30Thr variant varied between the different programs used. Although SIFT and AlignGVGD predicted this variant to be tolerated, PolyPhen predicted this variant to be possibly damaging to protein function. Structural modeling of the variant using FoldX predicted this variant to impact the total energy of the protein, inducing destabilization (ΔΔG = 1.35 kcal/mol).

---

**Table 1. In silico analysis of the identified CDH1 missense variants in our cohort of non-syndromic OFC patients**

<table>
<thead>
<tr>
<th>Missense variants</th>
<th>Grantham score</th>
<th>SIFT score</th>
<th>PolyPhen-2 score</th>
<th>PolyPhen-2 Grantham score</th>
<th>AlignGVGD score</th>
<th>FoldX prediction</th>
<th>Variant present in dbSNP</th>
<th>Frequency in ESP</th>
<th>Frequency in our in-house database</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.88C&gt;A (p.Pro30Thr)</td>
<td>38</td>
<td>Tolerated</td>
<td>Probably damaging 0.997</td>
<td>Class C0</td>
<td>ΔΔG = 1.35 kcal/mol</td>
<td>(HumVar) and 0.91</td>
<td>–</td>
<td>171</td>
<td>696</td>
</tr>
<tr>
<td>c.1108G&gt;T (p.Asp369 Tyr)</td>
<td>160</td>
<td>Deleterious</td>
<td>Probably damaging 0.999</td>
<td>Class C06</td>
<td>ΔΔG = −1.36 kcal/mol</td>
<td>(HumVar) and 0.91</td>
<td>–</td>
<td>0.02 kcal/mol</td>
<td>–</td>
</tr>
<tr>
<td>c.2413G&gt;A (p.Asp805Asn)</td>
<td>23</td>
<td>Deleterious</td>
<td>Probably damaging 0.996</td>
<td>Class C15</td>
<td>ΔΔG = 0.02 kcal/mol</td>
<td>(HumVar) and 0.91</td>
<td>–</td>
<td>0.04 kcal/mol</td>
<td>–</td>
</tr>
<tr>
<td>c.317A&gt;G (p.Lys113Glu)</td>
<td>56</td>
<td>Tolerated</td>
<td>Benign 0.008 (HumDiv) and 0.015 (HumVar)</td>
<td>Class C0</td>
<td>ΔΔG = 0.79 kcal/mol</td>
<td>(HumVar) and 0.91</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>c.317A&gt;G (p.Lys113Glu)</td>
<td>56</td>
<td>Tolerated</td>
<td>Benign 0.008 (HumDiv) and 0.015 (HumVar)</td>
<td>Class C0</td>
<td>ΔΔG = 0.79 kcal/mol</td>
<td>(HumVar) and 0.91</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Missense variants were predicted to have no influence on splicing.
Table 2. In silico analysis of the intronic variants with possible influence on splicing in our cohort of non-syndromic OFC patients

<table>
<thead>
<tr>
<th>Intronic variants identified (possible influence on splicing)</th>
<th>Frequency in in-house database</th>
<th>Frequency in ESP database</th>
<th>dbSNP Frequency in ESP database</th>
</tr>
</thead>
<tbody>
<tr>
<td>c.531+3A&gt;G (loss of splice donor site exon 4)</td>
<td>Not in in-house database</td>
<td>70%</td>
<td>72% (0.95)</td>
</tr>
<tr>
<td>c.532-18C&gt;T (stronger splice acceptor site exon 5)</td>
<td>Not in in-house database</td>
<td>99%</td>
<td>98% (0.97)</td>
</tr>
<tr>
<td>c.2439+10C&gt;T (loss of splice acceptor site exon 16)</td>
<td>Not in in-house database</td>
<td>99%</td>
<td>98% (0.97)</td>
</tr>
<tr>
<td>c.2440-6_2440-4delCTT (loss of splice acceptor site exon 16)</td>
<td>Not in in-house database</td>
<td>99%</td>
<td>98% (0.97)</td>
</tr>
<tr>
<td>p.Asp370Tyr and p.Asp805Asn</td>
<td>No splice site-prediction</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>p.Pro30Thr, p.Asp370Tyr and p.Asp805Asn affecting E-cadherin function</td>
<td>Not in in-house database</td>
<td>99%</td>
<td>98% (0.97)</td>
</tr>
</tbody>
</table>

Based on SIFT, PolyPhen and AlignGVGD predictions, the variant c.1108G>T (p.Asp370Tyr), found in one patient with CLP, is predicted to interfere with E-cadherin function. Structural modeling using FoldX suggests that this amino acid alteration leads to the establishment of an additional H-bond in the calcium coordination site between EC2-EC3, resulting in a decrease in the total energy ($\Delta G = -1.36$ kcal/mol) of the protein.

Although the variant c.2413G>A (p.Asp805Asn), identified in a patient with CP, seems to be structurally tolerated ($\Delta G = 0.02$ kcal/mol), the SIFT, PolyPhen and AlignGVGD algorithms predict a pathogenic effect of this variant.

The intronic variant c.531+3A>G might interfere with the recognition of the splice donor site of exon 4, although not all splice site-prediction programs are consistent.

For the variant c.337A>G, none of the programs predicted a functional impact. This was also the case for all the other intronic variants that were identified. Thus, functional assays were limited to the missense variants p.Pro30Thr, p.Asp805Asn and p.Asp370Tyr and the intronic variant c.531+3A>G.

**In vitro functional assays for identified missense variants**

To test the functional significance of the missense variants, Chinese Hamster Ovary (CHO) cells were transiently transfected with constructs encoding the wild-type (WT) protein and the different variants, as well as the empty vector (Mock). Transfected cells were analyzed for cell–cell adhesion competence, invasive capacity and E-cadherin expression levels and subcellular localization.

The slow aggregation assay showed that all three missense variants had an impact on cell–cell adhesion ability, but the type of aggregation showed distinct phenotypes (Fig. 1A). When compared with the aggregates formed by the WT E-cadherin-expressing cells, the variants p.Pro30Thr and p.Asp805Asn displayed smaller cellular aggregates, whereas the variant p.Asp370Tyr completely abrogated cell–cell adhesion and exhibited an isolated phenotype, similar to the phenotype of Mock-transfected cells.

The invasion assay revealed that, contrary to the cells expressing WT E-cadherin, the variant p.Asp370Tyr has lost the ability to suppress cell invasion through a Matrigel matrix ($P = 0.0078$) (Fig. 1B), indicating that this variant is more invasive than the WT-expressing cells.

Analyzing total E-cadherin expression by western blot, it seems that the variants p.Pro30Thr, p.Asp370Tyr and p.Asp805Asn do not interfere with the total amount of E-cadherin expressed in the cell (Fig. 1D). However, using immunostaining, abnormal E-cadherin localization was observed (Fig. 1C). Whereas WT E-cadherin was correctly located at the plasma membrane, the variants p.Pro30Thr, p.Asp370Tyr, and p.Asp805Asn presented reduced membranous E-cadherin expression.

Taken together, the *in vitro* results indicate that the variants p.Pro30Thr, p.Asp370Tyr and p.Asp805Asn affect E-cadherin protein function and its subcellular localization and thus can be considered as relevant mutations. The mutation p.Asp370Tyr shows the most dramatic impact on E-cadherin function.
In vitro functional assays for the c.531+3A>G intronic variant

To determine the possible effect on splicing, we transfected HeLa cells with a construct containing the CDH1 variant c.531+3A>G flanked by exons 3 and 5 of RHO. A cycloheximide control was used to suppress nonsense-mediated decay of a possible aberrant transcript. Twenty-four hours after transfection, RNA was isolated from transfected cells and a polymerase chain reaction (PCR) was performed with primers in the RHO exons. No indication of an altered splicing of CDH1 intron 4 for the construct with the c.531+3A>G variant was obtained.

DISCUSSION

To the best of our knowledge, this is the first report of a systematic analysis of CDH1 mutations in a consecutive series of non-syndromic OFC patients with an unknown family history of GC and/or LBC. We detected heterozygous CDH1 missense variants in 4 out of 81 (5%) patients. In vitro assays confirmed the functional impact of these variants, which can, therefore, be considered as relevant mutations. The data obtained in our study support the hypothesis that germline CDH1 mutations increase the risk of non-syndromic OFC.

The functional effect of missense variants is usually assessed using in silico prediction programs. Although in silico predictions are very useful and easy to perform, their potential is limited to amino acid conservation and the available structural information. In contrast to in silico predictions, in vitro assays allow to study the functional behavior of the cells with alterations in the protein in comparison to the WT form. Furthermore, immunocytochemistry analysis enables us to visualize the pattern of protein expression and localization and evaluate how the protein is affected in terms of trafficking and membrane expression.

The limitations of in silico predictions are underscored by the fact that two of the identified missense variants, p.Pro30Thr and p.Asp805Asn, for which in silico prediction programs were inconclusive, indeed seem to affect E-cadherin function and subcellular localization in our in vitro studies. For the p.Pro30Thr variant, the observed effect might be due to the localization of the variant in the prodomain. This may affect the processing to the mature form, possibly leading to degradation by endoplasmic reticulum-associated degradation, as described for other E-cadherin missense mutations (23,24).
Interestingly, p.Pro30Thr has been described before in a cohort of patients with LBC, but based only on \textit{in silico} analyses, the variant was considered non-causative (25). However, based on our functional analyses, we conclude that this variant affects E-cadherin function and relates to OFC. In light of this new information, its role in development of LBC might be reconsidered. The localization in the cytoplasmic tail of the variant p.Asp805Asn could impair the binding of other molecular partners, resulting in trafficking defects, abnormal cellular signaling or inefficient anchorage to the cytoskeleton, as described for other HDGC-associated missense mutations localized in proximity (26,27). The change of an Aspartic acid to a Tyrosine at amino acid 370 possibly leads to altered calcium affinity. This could result in the establishment of an aberrant conformation, justifying the loss of adhesive function observed for this variant in our aggregation assay.

In our cohort, two patients with CLP carry a p.Pro30Thr mutation. One of these patients self-reported a first-degree relative with CLP, the other patient reported a third-degree relative with OFC. Segregation analysis is needed to confirm whether affected relatives carry the \textit{CDH1} mutation. OFC were not reported in relatives of the patient carrying the p.Asp370Tyr variant. The family history of the patient carrying the p.Asp805Asn variant is unknown. These data suggest that the penetrance of developing OFC in subjects with a \textit{CDH1} mutation is incomplete. This is similar to what has been reported earlier in five \textit{CDH1} mutation-positive HDGC families with one or multiple cases of OFC (6,20).

The incomplete penetrance of OFC in \textit{CDH1} families suggests that apart from a genetic predisposition, environmental or other genetic factors may be involved in the occurrence of OFC in \textit{CDH1} mutation carriers. Several studies have shown that drinking alcohol (28,29), smoking (30,31) and a lack of folic acid (32,33) during pregnancy may increase the risk of OFC. The 11 chromosomal loci, that did not include \textit{CDH1}, identified in genome-wide association studies of non-syndromic OFC serve as good candidates for harboring genetic modifiers of the \textit{CDH1}-associated risk of OFC (3,34).

In HDGC families with \textit{CDH1} mutations, missense mutations occur in ~22% of the cases, and truncating mutations in about 78% (35). Lifetime risk of developing GC in individuals carrying a truncating germline \textit{CDH1} mutation has been estimated to be 80% by the age of 80, and the risk for women with a \textit{CDH1} mutation of developing LBC has been estimated to be 60% by the age of 80 (9,10). Systematic analysis of the risk of developing GC or LBC for missense mutations that affect E-cadherin function has not been described. Therefore, it is not known whether the mutation carriers that we identified in this non-syndromic OFC cohort are at increased risk of developing GC and LBC. The family history of \textit{CDH1}-associated cancer types of the mutation carriers identified in our non-syndromic OFC cohort is not yet available.

In conclusion, three functionally relevant missense mutations in \textit{CDH1} were identified in four patients with non-syndromic OFC. Before we can start implementing \textit{CDH1} mutation analysis in the standard diagnostic work-up of patients with non-syndromic OFC, a larger study is needed to confirm the findings of the current study. Moreover, calculation of cancer risks among carriers of functionally relevant missense \textit{CDH1} mutations is needed to assess whether prophylactic gastrectomy, which is currently the only option to prevent death from GC, should be considered in non-syndromic OFC patients with a missense mutation in \textit{CDH1}.

\section*{Materials and Methods}

\section*{Patients}

DNA was isolated from peripheral blood samples from 81 non-syndromic OFC patients from the \textit{Aetiologic} research into Genetic and Occupational/Environmental Risk Factors for Anomalies in Children (AGORA) study, that is a large Biobank at the Radboud University Nijmegen Medical Centre collecting questionnaires, phenotype data and DNA from patients diagnosed with congenital disorders or childhood cancers. Children in AGORA who had surgery for a cleft lip or a CLP were included. The majority of children in our cohort were of European descent and born between 1997 and 2011. Fifty-eight patients had CLP and 23 had an isolated CP.

All parents received questionnaires, with questions about demographic factors, family history of birth defects, cancer history of the parents, periconceptional information and pregnancy history. In addition, questions were asked about health status, prescribed medication, lifestyle and occupation pertaining to the 3 months before conception and during pregnancy.

The study was performed according to the rules of the Medical Ethics Committee of the Radboud University Nijmegen Medical Centre, registration number 2006/048. As \textit{CDH1} is a known tumor suppressor gene, the DNA samples were anonymized.

\section*{Mutation Analysis}

The full coding sequence of \textit{CDH1} (NM_004360.3), including splice junctions was amplified using polymerase chain reaction (primer sequences and PCR conditions are available on request) and screened for mutations using Big-Dye terminator sequencing BigDye Terminators (v 1.1) Applied Biosystems, USA. Analysis was performed on an ABI 3730 DNA Analyzer (Applied Biosystems). Subsequently, data were analyzed for variants in \textit{CDH1} using Sequence Pilot software (JSI Medical Systems, Germany).

\textit{In silico} predictions

Missense variants were analyzed using the Alamut 2.0 software package (Interactive Biosoftware, Rouen, France) incorporating SIFT (36), PolyPhen-2 (37) and Align GVGD (38). Furthermore, FoldX (http://foldx.crg.es/) was used to predict the structural impact of E-cadherin missense mutations, as described by Simões-Correia et al. (23). The previously established models of the different domains of E-cadherin (prodomain, extracellular and cytoplasmic domain) were used to calculate the total energies of the variants. The variants were generated with the Buildmodel command, and each variant was repeated in five runs. The energies are an automatic output in FoldX, and the native-state stability change between WT and mutant ($\Delta \Delta G = \Delta G_{\text{WT}} - \Delta G_{\text{Mut}}$) is also generated in a separate file, with the corresponding standard deviations and all the energetic penalties associated with each mutation. Mutations with $\Delta \Delta G > 0.8$ kcal/mol are considered destabilizing.
Intronic variants were analyzed using the Alamut 2.0 software package (Interactive Biosoftware), incorporating SpliceSiteFinder-like (39), MaxEntScan (40), NNSPLICE (41) and Human Splicing Finder (42).

To assess whether variants were present in individuals without OFC, we used the Exome Variant Server of the University of Washington (43) that contains CDHI sequencing data of 6500 individuals of European and African descent and our own in-house exome sequencing database (accessed July 2012), which contains data of over 1000 individuals, mainly of European descent. Finally, we used dbSNP (incorporated in the Alamut package, Build 135, accessed July 2012) to assess whether identified variants were recurrent.

**Plasmids construction**

Human E-cadherin cDNA was cloned in the pIRES2-EGFP vector according to the manufacturer’s instructions (Clontech, Takara Bio, Mountain View, CA, USA), and the variants c.88C>A (p.Pro30Thr), c.2413G>T (p.Asp370Tyr) and c.2413G>A (p.Asp805Asn) hE-cad were constructed by site-directed mutagenesis, following the protocol described by Wang and Wilkinson (44). The empty vector was used as control.

**Cell culture and transfection**

CHO cells (ATCC number: CCL-61, Barcelona, Spain) were grown at 37°C under 5% CO₂ humidified air, in α-minimum essential medium (MEM) (+) medium (Gibco, Invitrogen, Grand Island, NY, USA) supplemented with 10% fetal bovine serum (HyClone, Perbio, Cramlington, UK) and 1% penicillin/streptomycin (Gibco, Invitrogen, Grand Island, NY, USA). Cells were transiently transfected with the following vectors: empty vector (Mock), WT, p.Pro30Thr, p.Asp370Tyr and p.Asp805Asn hE-cad, using Lipofectamine 2000 (Invitrogen, Van Allen Way, Carlsbad, CA, USA), and the variants enriched with a protease inhibitor cocktail (Roche, Indianapolis, IN, USA) and a phosphatase inhibitor cocktail (Sigma). The proteins were quantified using a modified Bradford assay (Bio-Rad, Hercules, CA, USA). For analysis of total protein samples, 15 μg of proteins were eluted in sample buffer and loaded on a 7.5% SDS PAGE. The proteins were then electroblotted onto a Hybond ECL membrane (Amersham Biosciences, Buckinghamshire, UK). Membranes were blocked with 5% non-fat milk and 0.5% Tween-20 in PBS and immunoblotted with antibodies against E-cadherin (1:2000, Clone HEC1D Invitrogen, Camarillo, CA, USA) and α-tubulin (1:10,000, Sigma, St Louis, MO, USA). Sheep anti-mouse HRP-conjugated secondary antibody (Amersham Biosciences) was used, followed by ECL detection (Amersham Biosciences).

**Slow aggregation assay**

The functional significance of the E-cadherin missense mutations was assessed by a slow aggregation assay as described in Suriano et al. (45). A 96-well-plate was coated with 50 μl of agar solution (100 mg of Bacto-Agar in 15 ml of sterile PBS). Cells were detached with trypsin, resuspended in culture medium, and 2 × 10³ cells were seeded in each well. The plate was incubated at 37°C in a humidified atmosphere with 5% CO₂ for 24 h. Aggregation was evaluated using a Leica DM2000 microscope (Leica, Cambridge, UK).

**Immunofluorescence staining**

Cells were seeded on top of glass coverslips, and fixation was performed in ice-cold methanol for 20 min, followed by washing and blocking in 5% BSA in PBS for 30 min at room temperature. The mouse monoclonal E-cadherin antibody (BD Biosciences, Erembodegem, Belgium) was used at 1:300 dilution in PBS with 5% BSA and incubated at room temperature for 1 h. An Alexa Fluor 488 goat anti-mouse antibody (1:500, Invitrogen, Eugene, Oregon, USA) was applied as secondary antibody and incubated in the dark for 1 h. Coverslips were mounted on slides using Vectashield with DAPI (Vector Laboratories, Burlingame, CA, USA). Images were acquired on a Carl Zeiss Apotome Axiosvert 200M Fluorescence Microscope using 40× objective (Carl Zeiss, Jena, Germany). Images were taken with an Axioimc HRm camera and processed with the Zeiss Axion Vision 4.8 software.

**Western blotting**

Cells were lyzed in cold Catenin lysis buffer (1% Triton X-100, Sigma, St Louis, MO, USA), 1% Nonidet P-40 (Sigma) in PBS, and the variants were resuspended in culture medium, and 5 × 10³ cells were seeded in each chamber. The plate was incubated at 37°C in a humidified atmosphere with 5% CO₂ for 24 h. The non-invasive cells and matrigel from the upper side of the filters were removed with a prewet ‘cotton swab.’ The filters were washed in PBS, fixed in ice-cold methanol for 15 min and mounted in slides with Vectashield with 4′,6-diamidino-2-phenylindole (DAPI, Vector Laboratories, Burlingame, CA, USA). The total number of invasive nuclei was counted using a Leica DM2000 microscope (Leica, Cambridge, UK).

**Statistical analysis**

Two-tailed Student’s t-test was used to compare the number of invasive cells between the cell culture of the variants and that of the WT. In all analyses P < 0.05 was required for statistical significance.

**Functional analysis of intronic variants**

A plasmid containing the genomic region encompassing exons 3 to 5 of RHO inserted at the EcoRI/Sall sites in the pCI-NEO vector (46) was adapted to the Gateway cloning system and
used for in vivo splicing assays. The plasmid was digested with EcoNI and PfuMI (New England Biolabs, USA), resulting in the removal of exon 4 and part of the flanking intronic sequences and blunted using Large fragment Klenow DNA polymerase I (New England Biolabs), according to the manufacturer’s instructions. Subsequently, a blunt-end Gateway cloning cassette containing attR1 and attR2 sites and the pCI-neo vector were ligated using the Rapid DNA Ligation Kit (Roche, Germany) to generate pCI-NEO-RHO exon3,5:DEST.

PCR fragments containing exons 4 and 5 of CDH1 and surrounding intronic sequences were amplified from genomic DNA of the patient carrying the c.531+3A>G variant and a WT control using primers with attB sites. Using Gateway cloning technology (Invitrogen, USA), the attB-PCR product was cloned into pDONR201, sequence verified and subcloned in pCI-NEO-RHO exon3,5:DEST.

HeLa cells were transfected with 1 µg of plasmid DNA, using X-tremeGENE HP DNA Transfection Agent (Roche Applied Science, USA), according to the manufacturer’s protocol. After transfection, the cells were incubated at 37°C under 5% CO2 humidified air for 24 h, the last 4 h in the presence or absence of cycloheximide (100 µg/ml), and RNA isolation was performed using the RNAeasy RNA Mini Kit (Qiagen Benelux BV), according to the manufacturer’s protocol. RNA was reverse transcribed into cDNA according to standard procedures using random hexamer primers.

The RHO and CDH1 exons were amplified from the cDNA with forward primer 5′-ggaggtcaacaagctg-3′ and reverse primer 5′-gggtggagggaggagaggaggc-3′, which are located in RHO exon 3 and exon 5, respectively (PCR conditions available on request). The PCR product was sequenced as described above.

ACKNOWLEDGEMENTS

Authors would like to thank all the technicians from the Laboratory of Tumor Genetics and the technicians of the AGORA project, with special thanks to Saskia van de Velde-Visser. Furthermore, we would like to thank Lilian Vreede for practical assistance with the functional assays of the splicing variants. The Gateway-adapted pDEST vector was a kind gift from Erwin van Wyk. The Mock and Eccad pRES2-EGFP vectors were kindly provided by Dr Carla Oliveira and Renata Carriço.

Conflict of Interest statement. None declared.

FUNDING

Part of this work was supported by The Portuguese Foundation for Science and Technology (FCT) (PTDC/SAU-ONC/110294/2009).

REFERENCES


43. Exome Variant Server, NHLBI GO Exome Sequencing Project (ESP), Seattle, WA (URL: http://evs.gs.washington.edu/EVS/) [July 2012].

