

SUPPLEMENTARY MATERIALS AND METHODS:

Patient accrual:

327 LBC cases were identified through three different sources. The Breast Cancer Family Registry (Breast CFR) is an NCI-sponsored resource, which includes six population-based and clinic-based family registries and a collection of samples and data from more than 12,500 families with and without breast cancer [1]. Samples (n=168) were obtained from the Northern California, New York, Australia, Philadelphia, and Ontario sites of the Breast CFR. The Kathleen Cuninghame Foundation Consortium for Research into Familial Breast Cancer (kConFab) provided specimens (n=33) from families with a strong history of breast cancer, recruited from family cancer clinics in Australia and New Zealand [2]. The 126 remaining samples were collected through a Breast Cancer Research Foundation (BCRF)-funded Breast Cancer Genetics Consortium, a group of high-risk cancer clinics which included Dana-Farber Cancer Institute, Baylor College of Medicine, Beth Israel Deaconess Medical Center, Georgetown University, Massachusetts General Hospital, Memorial Sloan Kettering Cancer Center, Stanford University, the University of Chicago, the University of Pennsylvania, and the Hospital Vall d'Hebron, Barcelona, Spain. The inclusion criteria for the identification of the eligible LBC cases differed slightly between these groups. For cases from the Breast CFR and kConFab, eligibility for this study required a female case with documented invasive lobular or mixed (lobular and ductal) breast cancer, not known to carry germline *BRCA1* and *BRCA2* mutations, and either: (1) diagnosed before age 45 years, or (2) at any age but with two or more cases of breast cancer in first- or second-degree relatives. For cases from the BCRF-funded Breast Cancer Genetics Consortium, eligible women had a

diagnosis of invasive lobular or mixed (lobular and ductal) breast cancer, were not known to carry germline *BRCA1* and *BRCA2* mutations, and were either: (1) diagnosed before age 45 years, or (2) diagnosed at any age but with at least two or more cases of breast cancer in first- or second-degree relatives, and third-degree relatives in the paternal lineage, and with no reported family history of gastric cancer. The characteristics of the LBC cases screened for mutations in *CDH1* are summarized in Table 1. All cases had provided written informed consent and the study protocols were approved by the institutional review board at each participating center. DNA was extracted at the molecular laboratories for some of the collaborating centers using standard procedures (Qiamp DNA Blood Midi kit; Qiagen, Valencia, CA, USA) and anonymized genomic DNA samples were sent to the Centre for Translational and Applied Genomics (CTAG) at the British Columbia Cancer Agency (BCCA) where the analysis of *CDH1* was performed (Table 1).

Preparation of DNA and *CDH1* sequencing:

To accommodate the limited amount of DNA available, genomic DNA samples were subjected to whole-genome amplification using the GenomiPhi DNA amplification kit (GE Healthcare Bio-Sciences Inc., Quebec, Canada) as performed in our earlier study [3]. Exons and intron-exon boundary splice junctions of half the study cohort were amplified and screened for heterozygous base changes by the denaturing high pressure liquid chromatography (DHPLC) [3]. The primer sequences and conditions used have previously been described [4]. Exons displaying DHPLC changes consistent with a heterozygous variation were reamplified and PCR products were purified (Qiagen

MinElute; Qiagen, Mississauga, ON). Bidirectional sequencing was then performed (Big Dye Terminator V.3.1 Cycle Sequencing Kit; Applied Biosystems, Foster City, California, USA) and analysed (ABI Prism 310 Genetic Analyzer). Germline mutations in *CDHI* appear as heterozygous sequence changes, with the exception of large deletions, which are not detected by sequencing. Sequencing of all exons and intron-exon boundaries were carried out on the remaining samples using validated primer sets [5], at the Genome Sciences Centre on a service basis or in our laboratory.

Deletion Analysis:

Multiplex Ligation-dependant Probe Amplification (MLPA), previously used to identify large-scale deletions in *CDHI*, has been described previously [6]. MLPA could only be performed on 134 samples for which sufficient germline DNA was available.

Mutation Validation:

Basic Local Alignment Search Tool (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to compare sample sequences to the NCBI cDNA NM_004360.3 and the genomic DNA NG_008021.1, GI:190341080. Web-based software programs were used to look for predicted effects on splicing (Splice Site Prediction by Neural Network hosted by the Berkeley Drosophila Genome Project) and the predicted effects of amino acid changes on protein structure (Sorting Intolerant From Tolerant, SIFT, software version 2, Fred Hutchinson Cancer Research Center, Seattle, Wash).

Protein structure analysis:

Protein structure analysis and preparation of structural models was performed using the PyMol software package (DeLano Scientific). Structural co-ordinates were obtained from X-ray crystal structures of the ectodomain of C-cadherin (PDB code 1L3W) [7].

Functional Characterization:

Site directed mutagenesis:

To examine the effects of the missense mutations on protein localization, wild-type E-cadherin-WTpcDNA3.1 plasmids [8] were mutated using QuickChange site-directed mutagenesis kit (Stratagene, Cedar Creek, Texas) as per the manufacturer's instructions to create each of the novel non-synonymous variants we identified. The corresponding forward primer sequences are included following each variant: c.1223C>T, A408V, 5'-CCCCAATACCCCAGTGTGGGAGGCTGTAT-3'; c.8C>G, P3R, 5'-CTTACCATGGGCCGTTGGAGCCGCAGC-3'; c.88 C>A, P30T, 5'-GGAGCCCTGCCCACTGGCTTTGACGC-3'; c.1813A>G, 5'-ACTATATTCTTCTGTGAGGGGAATCCAAAGCCTCAGG -3'; c.1297G>A, D433N, 5'-CCACAAATCCAGTGAACAACAATGGCATTTTGAAAACAGCA-3'). The mutated plasmids were validated by DNA-sequencing.

Cell culture:

The immortalized ovarian surface epithelial cell-line, IOSE-80pc, that does not express E-cadherin [9] and the ovarian carcinoma cell line, OVCAR-3, that highly expresses E-cadherin [9], were maintained in a 1:1 mixture of MCDB 105 medium and Medium 199

(Sigma, St. Louis, MO, USA), supplemented with 5% fetal bovine serum (FBS) (Gibco BRL).

Transient transfections:

Aliquots of IOSE-80pc cells were transfected in parallel with the E-cadherin variants under investigation; c.1223C>T, A408V; c.8C>G, P3R; c.88 C>A, P30T; c.1813A>G, R695G; c.1297G>A, D433N and control plasmids. These controls included the empty vector LacZ as a control for the transfection procedure, wild-type (wt) E-cadherin and known loss of function E-cadherin mutants (c.1018A>G and c.2494G>A) that cause HDGC. Transient transfections were performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA) according to the manufacturer's guidelines.

Fluorescence microscopy:

E-cadherin staining was performed on subconfluent cell monolayers cultured on glass coverslips. Cells were washed once in media alone and then fixed in 4% paraformaldehyde (PFA) for 20 minutes at room temperature. Cells were washed with phosphate-buffered saline (PBS) before and after permeabilization with 0.2% Triton X-100 for 10min at room temperature. To decrease non-specific background, cells were incubated with serum-free protein block (Dako, Carpinteria, CA) for 30 minutes prior to incubation with a mouse monoclonal antibody [1:500] to the extracellular domain (HECD-1, 205601 Calbiochem) at room temperature. Primary antibody was detected using Alexa Fluor 488 goat anti-mouse IgG (Invitrogen, Carlsbad, CA) [1:500] for 60 minutes at room temperature. Cells were counterstained with DAPI [1:1000] prior to

mounting with glycerol. Slides were analyzed using the Axioplan 2, Zeiss (MetaSystems, Isis) camera VAC-30054 and pictures were obtained at 40x magnification.

References

[1] John EM, Hopper JL, Beck JC, Knight JA, Neuhausen SL, Senie RT, et al. The Breast Cancer Family Registry: an infrastructure for cooperative multinational, interdisciplinary and translational studies of the genetic epidemiology of breast cancer. *Breast Cancer Res* 2004;**6**(4):R375-89.

[2] Mann GJ, Thorne H, Balleine RL, Butow PN, Clarke CL, Edkins E, et al. Analysis of cancer risk and BRCA1 and BRCA2 mutation prevalence in the kConFab familial breast cancer resource. *Breast Cancer Res* 2006;**8**(1):R12.

[3] Masciari S, Larsson N, Senz J, Boyd N, Kaurah P, Kandel MJ, et al. Germline E-Cadherin mutations in familial lobular breast cancer. *J Med Genet* 2007;**44**(11):726-31.

[4] Suriano G, Yew S, Ferreira P, Senz J, Kaurah P, Ford JM, et al. Characterization of a recurrent germ line mutation of the E-cadherin gene: implications for genetic testing and clinical management. *Clin Cancer Res* 2005 Aug 1;**11**(15):5401-5409.

[5] Mullins FM, Dietz L, Lay M, Zehnder JL, Ford J, Chun N, et al. Identification of an intronic single nucleotide polymorphism leading to allele dropout during validation of a CDH1 sequencing assay: implications for designing polymerase chain reaction-based assays. *Genet Med* 2007 Nov;**9**(11):752-760.

[6] Oliveira C, Senz J, Kaurah P, Pinheiro H, Sanges R, Haegert A, et al. Germline CDH1 Deletions in Hereditary Diffuse Gastric Cancer Families. *Hum Mol Genet* 2009 May 1;**18**(9):1545-55.

[7] Boggon TJ, Murray J, Chappuis-Flament S, Wong E, Gumbiner BM, Shapiro L. C-cadherin ectodomain structure and implications for cell adhesion mechanisms. *Science* 2002 May 17;**296**(5571):1308-1313.

[8] Suriano G, Mulholland D, de Wever O, Ferreira P, Mateus AR, Bruyneel E, et al. The intracellular E-cadherin germline mutation V832 M lacks the ability to mediate cell-cell adhesion and to suppress invasion. *Oncogene* 2003 Aug 28;**22**(36):5716-5719.

[9] Wu C, Cipollone J, Maines-Bandiera S, Tan C, Karsan A, Auersperg N, et al. The morphogenic function of E-cadherin-mediated adherens junctions in epithelial ovarian carcinoma formation and progression. *Differentiation* 2008 Feb;**76**(2):193-205.