

Cadherin 26 is an alpha integrin-binding epithelial receptor regulated during allergic inflammation

Julie M. Caldwell, PhD¹, Margaret H. Collins, MD², Katherine A. Kemme¹, Joseph D. Sherrill, MD¹, Ting Wen, MD, PhD¹, Mark Rochman, PhD¹, Emily M. Stucke¹, Lissa Amin¹, Haitong Tai¹, Philip E. Putnam, MD³, Maximiliano J. Jiménez-Dalmaroni, PhD⁴, Mark R. Wormald, PhD⁵, Aleksey Porollo, PhD⁶, J. Pablo Abonia, MD¹, and ME Rothenberg, MD, PhD^{1*}

Affiliations:

¹Division of Allergy and Immunology, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, 3333 Burnet Ave., Cincinnati, OH 45229 USA.

²Division of Pathology and Laboratory Medicine, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, 3333 Burnet Ave., Cincinnati, OH 45229 USA.

³Division of Gastroenterology, Hepatology and Nutrition, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, 3333 Burnet Ave., Cincinnati, OH 45229 USA.

⁴Department of Biological Chemistry, John Innes Centre, The Sainsbury Laboratory, Norwich Research Park, Norwich, NR4 7UH, United Kingdom.

⁵The Glycobiology Institute, Department of Biochemistry, University of Oxford, Oxford, United Kingdom.

⁶Center for Autoimmune Genomics and Etiology, Division of Biomedical Informatics, Cincinnati Children's Hospital Medical Center, University of Cincinnati College of Medicine, 3333 Burnet Ave., Cincinnati, OH 45229 USA.

*To whom correspondence should be addressed: Marc Rothenberg, M.D., Ph.D., Cincinnati Children's Hospital Medical Center, Division of Allergy and Immunology, 3333 Burnet Ave. ML7028, Cincinnati, Ohio 45229 USA. Phone: 513.802.0257; Fax: 513.636.3310; E-mail: Rothenberg@cchmc.org.

Disclosure: M.E.R. is a consultant for Receptos, Pulm-One and NKT Therapeutics, and has an equity interest in Immune Pharmaceuticals, Pulm-One and NKT Therapeutics and received royalties for reslizumab (Teva Pharmaceuticals). M.H.C. is a consultant for Meritage Pharma, Novartis, Receptos, Regeneron, and Aptalis. J.M.C. and M.E.R. are co-inventors of a patent application, owned by Cincinnati Children's Hospital, concerning some of the data in this paper.

Abstract

Cadherins mediate diverse processes critical in inflammation including cell adhesion, migration, and differentiation. Herein, we report that the uncharacterized cadherin 26 (CDH26) is highly expressed by epithelial cells in human allergic gastrointestinal tissue. *In vitro*, CDH26 promotes calcium-dependent cellular adhesion of cells lacking endogenous cadherins by a mechanism involving homotypic binding and interaction with catenin family members (alpha, beta, and p120), as assessed by biochemical assays. Additionally, CDH26 enhances cellular adhesion to recombinant integrin $\alpha 4\beta 7$ *in vitro*; conversely, recombinant CDH26 binds αE and $\alpha 4$ integrins in biochemical and cellular functional assays, respectively. Interestingly, CDH26-Fc inhibits activation of human CD4⁺ T cells *in vitro* including secretion of IL-2. Taken together, we have identified a novel functional cadherin regulated during allergic responses with unique immunomodulatory properties, as it binds $\alpha 4$ and αE integrins and regulates leukocyte adhesion and activation, and may thus represent a novel checkpoint for immune regulation and therapy via CDH26-Fc.

Introduction

Cadherins, a family of transmembrane cell surface glycoproteins, mediate calcium-dependent cell adhesion and exhibit a tightly regulated tissue-specific expression pattern. Cadherins expressed predominantly within epithelia such as E-cadherin (CDH1) and P-cadherin (CDH3)¹ contribute to the maintenance of skin and mucosal barriers, regulating access of pathogens and allergens to underlying tissue and immunocytes. Modulation of cadherin expression and function has been associated with a number of diseases such as cancer, in which metastatic tumorigenesis is often associated with a switch in cadherin expression such as CDH1 to CDH2 (N-cadherin)², as well as diverse inflammatory diseases, including but not limited to asthma, eczema, chronic rhinosinusitis, inflammatory bowel disease, and rheumatoid arthritis³⁻⁷. Focusing on epithelial CDHs, reduced CDH1 occurs in gastroesophageal reflux disease, asthma, and eczema and has been shown to contribute to loss of epithelial integrity, impairment of barrier function, and production of pro-inflammatory cytokines^{6,8-10}. A substantial advance in the cadherin field has been the finding that CDH1 binds lymphocyte integrin $\alpha E\beta 7$ and regulates the activation and localization of epidermal and intestinal intraepithelial lymphocytes¹¹⁻¹⁴. Despite these observations regarding CDH1, the involvement of other cadherins in the regulation of immunologic processes mediated by the mucosal epithelium such as their binding to integrins has not been described.

Allergic disorders are characterized by a Th2 immune response that involves the accumulation of distinct subsets of activated leukocytes at the affected site. In particular, in addition to eosinophils, increased numbers of CD4⁺ T cells and type 2 innate lymphoid cells have been observed at sites of allergic gastrointestinal (GI) inflammation¹⁵⁻¹⁸. Furthermore,

highly elevated levels of Th2 cytokines, including IL-13, suggest activation of these cells at sites of allergic GI inflammation^{19,20}. In general, it is known that leukocytes, in part under the influence of soluble mediators including cytokines and chemokines, localize from the blood to tissue in a multi-step process involving the coordinate expression and activation of leukocyte-expressed selectins and integrins and their counter-receptors on activated endothelium²¹. However, less is known about how leukocyte-expressed receptors interact with epithelial ligands to influence cellular localization and activation, particularly in the context of allergic inflammation.

We identified a previously uncharacterized cadherin-like molecule, CDH26, that was markedly overexpressed in human gastrointestinal tissue with active eosinophilic inflammation²⁰, yet no studies of this molecule have been reported, although its transcript appears to be upregulated in epithelial cells under Th2-associated conditions²²⁻²⁵. Herein, we demonstrate that CDH26 is a unique functional cadherin, in that it (1) has an epithelial cell-restricted expression pattern that is particularly prominent following gene induction during allergic inflammation; (2) is an $\alpha 4$ and αE integrin receptor; (3) has the capacity to modulate leukocyte adhesion and activation; and (4) has immunomodulatory function that can be exploited via a CDH26-Fc fusion protein, which has immunosuppressive activity.

Results

CDH26 is overexpressed during pathologic allergic inflammation

Gastric tissue of control patients and patients with an allergic gastroenteropathy, eosinophilic gastritis (EG), was subjected to global transcript analysis²⁰. Additional information about these patients can be found in the Supplementary Methods and Supplementary Table S2. We identified

that the most upregulated transcript that passed the criteria $P < 0.01$ and 2-fold filter was the uncharacterized cadherin family member cadherin 26 (*CDH26*) (12.3-fold, $P < 0.005$)²⁰. We verified by real time polymerase chain reaction (RT-PCR) analysis that the *CDH26* mRNA level was highly increased in the gastric tissue of EG patients within the same cohort subjected to microarray (15.3-fold, $n = 5$ EG vs. $n = 5$ control, $P = 0.03$; Supplementary Figure S1a) as well as in additional EG patient gastric tissue (35.6 fold, $n = 10$ EG vs. $n = 10$ control, $P < 0.0001$; Supplementary Figure S1b). Comparison of the genes differentially regulated in EG²⁰ and in eosinophilic esophagitis (EoE)²⁶ revealed that *CDH26* was the only cadherin family member that exhibited a significant change in gene expression in both allergic disorders. Indeed, as previously observed¹⁹, *CDH26* mRNA expression was significantly increased (115 fold) in the esophageal tissue of patients with active EoE compared to control patients (Figure 1a). Only *CDH1* (E-cadherin) and *CDH26* exhibited raw signals indicative of the transcripts being substantially expressed (i.e. at least one sample for that probe exhibited raw signal ≥ 400) in the gastric tissue of patients with active EG (Figure 1b). Raw signals for cadherin transcripts in esophageal tissue from EoE patients only showed such values for *CDH1*, *CDH3* (P-cadherin), and *CDH26* (Figure 1c). As a control, no significant change in *CDH26* or *CCL26* (eotaxin-3) expression was observed in gastric tissue of patients with *H. pylori* gastritis compared to control patients, although as a control *C3* transcript was elevated in the *H. pylori* cohort as previously reported (Figure 1d and ^{27,28}). Notably, a microarray study of gastric antrum tissue of patients with *H. pylori* gastritis did not identify *CDH26* as being upregulated compared to normal tissue²⁷. As such, *CDH26* is a unique cadherin in terms of its expression level and regulation in two distinct allergic states.

CDH26 protein expression is increased in inflamed allergic gastrointestinal tissue and is localized to epithelial cells

Immunohistochemical staining for CDH26 showed localization nearly exclusively in the surface and gland epithelial cells in gastric tissue of patients with EG (Figure 2a and b). The mean peak number of CDH26-positive cells was 241 cells/400X high-power field (HPF) (mean \pm SEM, 104.2 ± 40.8 ; $P = 0.0476$, Mann Whitney test) in EG patients compared with no expression above background in controls. Western blot revealed an increased level of gastric CDH26 (4.9 fold) in EG patients compared to control samples (Figure 2c). Patients with active EoE had high levels of esophageal CDH26 protein expression compared to control patients. In control esophageal biopsies, the staining was confined to epithelial cells near the surface, but in active EoE the staining was both more intense and prevalent and included both surface epithelial cells as well as epithelial cells in the expanded basal layer (Figure 2d). By western blot analysis, esophageal tissue of EoE patients showed 3.4-fold increased CDH26 protein levels compared to control tissue (Figure 2e).

CDH26 is a functional cadherin

CDH26 exhibits sequence homology to the cadherin family of proteins, with five extracellular cadherin repeats (EC)²⁹ in the putative extracellular portion of the protein, a predicted transmembrane domain, and a C-terminal cytoplasmic region (Figure 3a). To determine if CDH26 localized to the cell membrane of esophageal epithelial cells, affinity isolation of biotinylated surface proteins was performed. Proteins present on the surface of cells were labeled with biotin, the cells were lysed and proteins solubilized, and biotinylated proteins were subjected to affinity isolation using streptavidin beads. Proteins bound to the streptavidin beads

were subsequently solubilized in loading buffer and subjected to SDS-PAGE. Western blot for CDH26 indicated that it was present at the cell surface in esophageal epithelial cells that express high levels of CDH26 (Figure 3b), because it was pulled down only from cells overexpressing CDH26 that underwent surface biotinylation (Figure 3b, lane 8).

CDH26 contains five asparagine residues in its extracellular domain located within the consensus sequence for N-linked glycosylation (N81, N85, N171, N177, N462). To test whether CDH26 is modified by N-linked glycosylation, CDH26 was immunoprecipitated and then treated with peptide: *N*-glycosidase F (PNGase F) that was either active or heat-inactivated as a negative control. The treated, immunoprecipitated proteins were then subjected to SDS-PAGE and western blot analysis for CDH26. Immunoprecipitated CDH26 treated with peptide: *N*-glycosidase F (PNGase F) (Figure 3c, lane 5), but not heat-inactivated PNGase F (Figure 3c, lane 6), exhibited increased mobility compared to CDH26 from total cell lysates, indicating that the protein is modified by N-linked glycosylation under baseline conditions.

CDH26 contains a tryptophan residue at position 2 of the most N-terminal EC domain known to be critical for dimerization *in trans* of type I and type II cadherins³⁰. Therefore, we tested whether CDH26 molecules interact in a homotypic manner. When co-expressed, myc-tagged CDH26 co-immunoprecipitated with HA-tagged CDH26 (Figure 3d, lane 7), and the reciprocal immunoprecipitation confirmed that CDH26-HA co-immunoprecipitated with CDH26-MYC (Figure 3d, lane 8).

We next tested whether CDH26 interacted with beta-catenin, which binds other cadherin molecules to link them indirectly to the actin cytoskeleton³⁰⁻³². The region of CDH1 known to interact with beta-catenin exhibited 68% similarity to the same region of CDH26 (^{31,32}, Supplementary Figure S2 a, b, and d). When immunoprecipitation of HA-tagged beta-catenin

was performed, CDH26 was also detected in the precipitates (Figure 3e, lanes 7 and 9), indicating that beta-catenin and CDH26 exist in the same complex.

Beta-catenin interacts with alpha-catenin to link cadherin molecules indirectly to the actin cytoskeleton and thus support cell adhesion³⁰. We tested whether alpha-catenin exists in the same complex as CDH26 and found that CDH26 co-immunoprecipitated with alpha-catenin (Figure 3f, lane 6). As a positive control, beta-catenin was also observed to co-immunoprecipitate with alpha-catenin (Figure 3f, lane 6).

p120-catenin binds the juxtamembrane domain (JMD) of the cytoplasmic portion of cadherin molecules and has been shown to function in maintenance of cadherin stability and localization to the cell surface³⁰. The primary amino acid sequence of the JMD of CDH1 was notably homologous (48%) to that in CDH26 (Supplementary Figure S2 a, c, and d). We therefore tested whether CDH26 and p120-catenin could exist in the same protein complex and found that p120-catenin and CDH26 co-immunoprecipitated (Figure 3g, lane 9).

We tested whether CDH26 could promote calcium-dependent cellular adhesion. To do this, L929 cells, which lack endogenous cadherins, were used in an aggregation assay. L929 cells that were transduced with a CDH26 expression construct and thus expressed high levels of CDH26 showed a high degree of aggregation only in the presence of calcium (Figure 3h, column 4 vs. column 3), whereas cells transduced with a control expression construct aggregated neither in the presence nor the absence of calcium (Figure 3h, columns 1 and 2).

CDH26 binds $\alpha 4$ and αE integrins

Heterotypic binding of cadherins and integrins has been reported^{12,33}. There are a number of proteins known to bind integrins, including CDH1 as well as ICAM-1, MAdCAM-1, and

fibronectin, for which 3D structures have been resolved and integrin binding sites have been localized³⁴. To map these sites to CDH26, we generated a 3D model of CDH26 structure using homology modeling with the CDH1 structure (PDB ID 3Q2V) as a primary template. The pairwise structure alignment of the model (blue) with other resolved structures (gray) indicates putative integrin binding sites in CDH26 (Figure 4a). As can be seen, known integrin binding sites are located in unstructured loops of the extracellular domain and are negatively charged. D42 of MAdCAM-1 and E34 of ICAM-1 overlap with D98, E99, and E102 in CDH26. D1495 of fibronectin corresponds to E138 of CDH26. The presence of these negatively charged residues in unstructured loops on the surface suggests that CDH26 would bind integrins.

We tested whether L929 cells expressing high levels of CDH26 exhibited adhesion to $\alpha 4\beta 7$ compared to control cells. Indeed, L929 cells transduced with a CDH26 expression vector showed increased adhesion to $\alpha 4\beta 7$ -coated wells compared to control cells, in contrast to BSA-coated wells, which did not support adhesion of either cell type (Figure 4b). Visualization of cellular binding revealed a marked increase of CDH26-expressing cells binding $\alpha 4\beta 7$ (Figure 4c). To further prove that CDH26 could directly bind $\alpha 4\beta 7$, a solid phase adhesion assay was performed in which recombinant CDH26-hIgG1-Fc (referred to as CDH26-Fc hereafter; see Supplementary Figure S3 and Supplementary Methods for details of isolation) was added to wells that were coated with recombinant $\alpha 4\beta 7$. CDH26-Fc bound specifically to $\alpha 4\beta 7$ and not BSA, while negative control IgG did not bind either $\alpha 4\beta 7$ or BSA (Figure 4d).

To further substantiate and address the specificity and molecular requirements for CDH26/ $\alpha 4\beta 7$ interaction, overexpression and co-immunoprecipitation studies were carried out using HEK 293T cells. When HA-tagged CDH26 (CDH26-HA) and integrin $\alpha 4$ were overexpressed in HEK 293T cells, integrin $\alpha 4$ co-immunoprecipitated with CDH26-HA (Figure

4e), and the reciprocal was observed as CDH26-HA was found to co-immunoprecipitate with integrin $\alpha 4$ (Figure 4f). In addition to integrin $\alpha 4$, V5-tagged integrin αE (ITGAE-V5) was observed to co-immunoprecipitate with CDH26-HA (Figure 4g). Integrin $\alpha 4$ did not co-immunoprecipitate with CDH1, although positive control integrin αE co-immunoprecipitated with CDH1 (Figure 4h).

We next tested whether $\alpha 4$ bound to integrin $\beta 1$, could mediate interaction with CDH26 by observing whether cells that express integrin $\alpha 4\beta 1$ adhered to recombinant CDH26-Fc. Jurkat cells, which express integrin $\alpha 4\beta 1$ but not $\alpha 4\beta 7^{35}$, adhered to CDH26-Fc to a significantly greater degree than they adhered to IgG control antibody (Figure 4i). Pre-incubation of Jurkat cells with anti-integrin $\alpha 4$ antibody (HP2/1), but not an equivalent amount of control mIgG1, blocked their binding to CDH26-Fc in a dose-dependent manner (Figure 4i). Jurkat cells stimulated with antibodies that activate integrin $\beta 1$ (clone TS2/16) adhered to CDH26-Fc in an integrin $\alpha 4$ -dependent manner to a greater degree than they adhered to CDH1-Fc (Figure 4j).

CDH26 modulates CD4⁺ T cell activation

Because other $\alpha 4\beta 1$ ligands including VCAM-1 have been shown to co-stimulate CD4⁺ T cell activation³⁶⁻³⁸, we sought to test the hypothesis that the putative integrin $\alpha 4$ ligand CDH26 had this property. After human peripheral blood CD4⁺ T cells isolated by negative selection (detailed in Supplementary Methods) were incubated for 48 h in the presence of plate-bound anti-CD3 antibodies (clone OKT3) to induce suboptimal stimulation of the T cell receptor (TCR), an increased percentage of the cells expressed CD25 at the cell surface as assessed by flow cytometry analysis (detailed in Supplementary Methods). However, the presence of CDH26-Fc, but not IgG control, inhibited the increase in percentage of cells expressing CD25 in a dose-

dependent manner (Figure 5a; [Supplementary Figure S4a](#)). Likewise, production of the cytokine IL-2 was inhibited by CDH26-Fc in a dose-dependent manner in cells subjected to TCR stimulation (Figure 5b). A similar inhibition of IL-4 secretion and CD69 and CD154 surface expression were also observed (data not shown). We next tested whether this effect was specific to CDH26-Fc or whether other cadherin-Fc proteins mediated inhibition of T cell activation. Both CDH26-Fc and CDH1-Fc attenuated the increase of CD25 surface expression following TCR stimulation of CD4⁺ T cells (Figure 5c; [Supplementary Figure S4b](#)). The inhibition of CD4⁺ T cell activation by both cadherin-Fc constructs was also apparent at the level of cytokine secretion; cadherin-Fc constructs inhibited the secretion of IL-2 in response to TCR stimulation (Figure 5d).

Discussion

Herein we have elucidated the properties of CDH26, identifying it as a functional cadherin with unique features in that it is the only cadherin family member significantly upregulated in human allergic gastrointestinal tissue, where it localizes to epithelial cells in the inflamed esophagus and stomach. CDH26 mediates calcium-dependent cell adhesion, dimerizes/multimerizes, and interacts with alpha-, beta-, and p120-catenins. CDH26 also has the ability to impact leukocyte migration and adhesion. Moreover, we present biochemical, molecular, and functional evidence that CDH26-integrin interactions impact cellular adhesion; specifically, integrins $\alpha 4$ and αE co-immunoprecipitate with CDH26, recombinant CDH26-Fc binds recombinant $\alpha 4\beta 7$, CDH26-expressing cells adhere to integrin $\alpha 4\beta 7$, and Jurkat cells adhere to recombinant CDH26-Fc in a manner dependent on integrin $\alpha 4$ (see Figure 6 and Supplementary Table S1 for summary and comparison to CDH1). Besides uncovering a novel role for this molecule, we present evidence

that it can be exploited to generate a potential therapeutic as CDH26-Fc is an immunosuppressive molecule, at least *in vitro*. Taken together, we have identified a novel functional cadherin regulated during allergic inflammation and determined that it binds α integrins and has immunomodulatory properties.

Given our observation that CDH26 interacts with not only integrin $\alpha 4$ but also integrin αE , CDH26 might have a function similar to E-cadherin to regulate localization or activation of leukocytes during allergic responses via interacting with leukocyte integrins. Epithelial CDH26 may impact the localization or activation status of diverse $\alpha 4^+$ and/or αE^+ cells (e.g. $CD4^+$ T cells, eosinophils, mast cells) within the epithelium in the context of allergic inflammation. In particular, these subsets of cells are known to be increased in the esophageal epithelium of EoE patients compared to control patients^{17,26,39}; furthermore, numerous intraepithelial eosinophils are observed in EoE and EG but not in normal esophageal or gastric tissue. This altered intraepithelial localization of several subsets of cells correlates with the fact that CDH26 appears to be a highly inducible molecule in epithelial cells that is present at only low levels under homeostatic conditions.

We identified several lines of evidence that CDH26 interacts with alpha integrins and facilitates cellular binding to $\alpha 4$ -containing integrins, including both integrin $\alpha 4\beta 7$ and $\alpha 4\beta 1$. Biochemical assays suggested that the extracellular portion of CDH26 could directly bind the $\alpha 4$ -containing integrin $\alpha 4\beta 7$. Furthermore, we observed that cells expressing high levels of CDH26 adhered to recombinant integrin, and in the reciprocal situation, recombinant CDH26 was sufficient to mediate adhesion of Jurkat T cells in a manner dependent on integrin $\alpha 4$. Such interactions are consistent with the structural properties of CDH26, which include the presence of solvent-exposed acidic residues that could be critical to facilitate the CDH26/integrin

interaction. We speculate that in addition to mediating adhesion, interaction of epithelial-expressed CDH26 with leukocyte integrins could initiate intracellular signaling in both epithelial cells and leukocytes. This could impact diverse processes such as alteration of gene expression, regulation of barrier function, or production of mediators by either cell type.

It has been shown that integrin $\alpha 4$ ligands including fibronectin, VCAM-1, and MAdCAM-1 costimulate CD4⁺ T cells activated by suboptimal TCR activation³⁶⁻³⁸. Our data show that CDH26-Fc inhibited CD4⁺ T cell activation mediated by suboptimal TCR engagement. We additionally observed that CDH1-Fc had similar effects. In fact, although CDH1 has previously been shown to co-stimulate CD4⁺ T cell activation⁴⁰, it has additionally been shown to be an inhibitory molecule in several settings; for example, ligation of E-cadherin expressed by dendritic epidermal T cells with CDH1 expressed by epidermal keratinocytes inhibits DETC IFN- γ production, TNF- α production, and degranulation in response to TCR stimulation¹⁴. Moreover, CDH1 has been shown to be a counterreceptor for killer cell lectin-like receptor G1 (KLRG1), which is expressed by NK cells, memory T cells, and type 2 innate lymphoid cells. Engagement of KLRG1 by E-cadherin promotes inhibitory effects in the KLRG1-expressing cell including inhibition of NK cell cytotoxicity, inhibition of antigen-induced proliferation and induction of cytolytic activity of CD8⁺ T cells, and inhibition of type 2 cytokine production of type 2 innate lymphoid cells⁴¹⁻⁴³. We speculate that CDH26 may act as a counterregulatory molecule either through engagement of integrins or other unknown counterreceptors to dampen the Th2-associated inflammatory responses. Although molecules upregulated during disease might be assumed to be involved in promoting disease, upregulation of CDH26 could be part of the mechanism by which resolution of inflammation occurs to promote the return of the tissue to homeostasis. Alternatively, it is possible that CDH26

engagement may inhibit particular subsets of CD4⁺ T cells; for example, regulatory T cells, which are known to be increased in EoE and EG^{16,20,44}, or other more select CD4⁺ T cell subsets such as Th1 or Th2 cells. Depending on the subset of T cells inhibited, the molecule could serve either to dampen or accelerate Th2-associated inflammatory responses. In summary, CDH26 is now the second cadherin (besides CDH1) that has been shown to bind α -integrins, extending the paradigm of cadherin/integrin interactions and leading to the unexpected finding that CDH26-Fc (as well as CDH1-Fc) has T cell immunosuppressive activity, providing a potential novel therapeutic strategy.

Methods

Methods related to patient studies, expression constructs, cell isolation, FACS analysis, ELISA, and bioinformatics analysis can be found in the Supplementary Material.

Quantitation of transcript levels

Microarray analysis

For microarray analysis, biopsy samples collected during the index endoscopy were stored in RNAlater until subjected to RNA isolation using the miRNeasy kit (Qiagen) per the manufacturer's instructions. RNA labeling and hybridization to the GeneChip Human Genome U133 Plus 2.0 Array (Affymetrix) was performed as reported²⁶. RNA labeling, hybridization, and generation of expression data were performed by the Gene Expression Microarray Core at CCHMC. In some cases, the raw signal intensity of particular probe sets was reported. Transcripts for which all individual samples exhibited a raw signal of ≤ 100 were considered to be not expressed. Transcripts with raw signal ≥ 400 were considered to be substantially expressed.

Quantitative PCR

Total RNA (100 ng-1 μ g) isolated from biopsy specimens using the miRNeasy kit (Qiagen) or from cells using Trizol (Invitrogen) was used to synthesize cDNA using Superscript II Reverse Transcriptase (Invitrogen). Real-time (RT)-PCR was performed using the iQ5 system (Biorad) and SYBR green mix (BioRad). The value obtained for each primer set (Supplementary Table S3) was normalized to the *GAPDH* value for the corresponding sample.

Protein and cell detection in tissue

Histopathology

Biopsies for histologic evaluation were fixed in 10% formalin, routinely processed, and embedded in paraffin. Sections (5 μm) were stained with H&E or with specific antibodies. Immunohistochemical stains were performed using Ventana Benchmark XT automated immunostainer. Antigen retrieval (EDTA, 30 min) was followed by staining with anti-CDH26 (1:50, Sigma-Aldrich) or anti-*Helicobacter pylori* (Ventana Medical Systems, Inc.) antibodies. For quantitative microscopy, multiple levels of biopsies were surveyed, and the areas containing the greatest concentration of eosinophils or immunopositive cells were identified and enumerated at 400X (0.3 mm^2) to generate a peak count per biopsy. Quantitative evaluations were performed in well-oriented areas when feasible.

Cell culture and manipulation

Culture of cell lines and cytokine treatment

Human esophageal squamous epithelial cancer cell line TE-7 was kindly provided by Dr. Hainault (IARC, Lyon, France). These cells were maintained in RPMI medium (Invitrogen) supplemented with 5% FBS (Atlanta Biologicals) and 1% penicillin/streptomycin (Invitrogen). HEK 293T cells and L929 cells were grown in DMEM medium (Invitrogen) supplemented with 10% FBS (Atlanta Biologicals) and 1% penicillin/streptomycin (Invitrogen). Jurkat cells were cultured in RPMI medium (Invitrogen) supplemented with 10% FBS (Atlanta Biologicals) and 1% penicillin/streptomycin (Invitrogen).

Lentivirus production and transduction of HEK 293T, TE-7, and L929 cells

For pMIRNA1 constructs, lentivirus production was carried out by the CCHMC Viral Vector Core. HEK 293T, TE-7, and L929 cells were transduced by incubating lentivirus with the cells for 24 h in the presence of 5 µg/ml polybrene. Media were then changed, and after 24 h medium containing 2 µg/ml puromycin was added. After selection for 48 h in puromycin, cells were dispersed and plated to single cells in 96 well plates to obtain clones derived from single cells. A second round of dispersing, plating to single cells, and picking single colonies was done. CDH26 expression was verified by western blot analysis.

Protein methods

Protein extracts and immunoprecipitation (IP)

Cell lysates were prepared from HEK 293T cells generally as described. Cells (approximately 2×10^6) were washed one time with PBS and incubated in IP buffer (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 2 mM EDTA, 1 mM dithiothreitol, 1% Nonidet P-40 [NP-40]; or 10 mM imidazole, 100 mM NaCl, 1 mM MgCl₂, 5 mM EDTA, 1% Triton X-100, pH 7.4) containing 1X cOmplete protease inhibitor cocktail (Roche) for 10 min on ice. Cells were scraped from the plate and rotated at 4°C for 10 min. Lysates were cleared by centrifugation at 20,000 x g at 4°C for 10 min. An equal amount of protein was added to total 500 µl of IP buffer plus protease inhibitors (Roche). Antibodies (2 µg α-HA (Covance), α-myc (Covance), α-p120 (BD Transduction Laboratories), α-ITGA4 (Cell Signaling Technology, Inc.), mouse IgG1 control (AbD Serotec), or normal rabbit IgG control (R & D Systems)) were added and rotated overnight at 4°C. Subsequently, 20 µl of protein A/G agarose beads (Santa Cruz Biotechnology, Inc.) were added per sample. After 2 h of rotation (4°C), beads were washed 5 times in IP buffer containing

protease inhibitors. 2X Laemmli buffer was added to the immunoprecipitates or total cell lysates saved prior to IP (input).

Biopsy protein extracts

Distal esophagus or gastric antrum biopsy specimens were transferred into 100 µl of IP buffer (50 mM Tris-HCl [pH 7.4], 150 mM NaCl, 2 mM EDTA, 1 mM dithiothreitol, 1% Nonidet P-40 [NP-40], 1X protease inhibitors (Roche)) and sonicated. Lysates were cleared by centrifugation (20,000 x g, 4°C, 10 min). Alternatively, protein was isolated from the organic fraction remaining after RNA isolation from biopsy specimens using the miRNeasy kit (Qiagen). DNA was precipitated by the addition of 0.3 volumes of ethanol followed by a 2,000 x g spin. Protein was precipitated from the supernatant by addition of 3 volumes of acetone, pelleted by centrifugation (20,000 x g, 10 min, 4°C), dried, and solubilized in Laemmli buffer (2X).

Biotinylation of cell surface proteins

Adherent cells were washed with ice-cold biotinylation buffer (100 mM HEPES, 50 mM NaCl, pH 8.0) twice; cold biotinylation buffer plus sulfo-NHS-LC-biotin (0.9259 mg/ml) (Thermo Scientific) was then added (30 min on ice). Buffer was removed, cells were washed 3 times with ice-cold PBS with 100 mM glycine, and protein was then extracted as described above using IP buffer with protease inhibitors. Cell lysates were incubated with streptavidin-agarose beads (Sigma-Aldrich) for 2 h at 4°C. Beads were washed 5 times with cold IP buffer containing 1 mM PMSF followed by addition of 2X Laemmli buffer.

SDS-PAGE and Western blot analysis

Total protein, inputs, or immunoprecipitates were loaded onto either 4-12% NuPage Tris-bis gels (Invitrogen) and electrophoresed for 1.5 h at 150 V or 4-12% BOLT gels (Invitrogen) and electrophoresed for 50 min at 165V. Proteins were then transferred to nitrocellulose membranes. Primary antibodies were diluted in TBS/0.1% Tween 20 containing 5% milk or Odyssey blocking buffer (LI-COR Biosciences) + 0.2% Tween 20: rabbit anti-CDH26 (Sigma-Aldrich), 1:500; rabbit anti-beta-catenin (Cell Signaling Technology, Inc.), 1:1000; mouse anti-p120 (BD Transduction Laboratories), 1:1000; mouse anti-HA (Covance), 1:1000; rabbit anti-HA (Santa Cruz Biotechnology, Inc.), 1:200; mouse anti-Myc (Cell Signaling Technology, Inc.), 1:1000; mouse anti-beta-actin (Sigma-Aldrich), 1:1000; rabbit anti-V5 (Bethyl Laboratories), 1:1000; rabbit anti-ITGA4 (Cell Signaling Technology, Inc.), 1:1000; rabbit anti-E-cadherin (Cell Signaling Technology, Inc.), 1:1000. HRP-conjugated secondary antibodies were incubated with the membranes in TBS/0.1% Tween 20 containing 5% milk: anti-rabbit HRP, 1:10,000 (Cell Signaling Technology, Inc.); anti-mouse HRP, 1:10,000 (Cell Signaling Technology, Inc.). Blots were developed using ECL Plus reagent (GE Healthcare). Densitometry measurements were performed using Multi Gauge V3.0 (Fugifilm). Alternatively, secondary antibodies conjugated to infrared fluorophores (anti-rabbit IRDye 800CW or anti-mouse IRDye 680RD; 1:15,000) (LI-COR Biosciences) were used, and blotting was carried out in a similar manner except that antibodies were diluted in Odyssey blocking buffer (LI-COR Biosciences) + 0.2% Tween 20 and infrared signal was visualized and quantified using the Odyssey[®] CLx Infrared Imaging System and ImageStudio software (LI-COR Biosciences).

Solid phase adhesion assay

Recombinant human integrin $\alpha 4\beta 7$ (R&D Systems) was diluted to the indicated concentration in buffer (150 mM NaCl, 20 mM HEPES); 100 μ l per well was added to Costar half-well ELISA plates (Corning Inc.) and incubated overnight at 4°C. The following day, wells were washed and blocked with 5% BSA in buffer (150 mM NaCl, 20 mM HEPES) overnight at 4°C. The following day, the wells were washed, and IgG1 or CDH26-hIgG1-Fc diluted in assay buffer (150 mM NaCl, 20 mM HEPES, 1 mM CaCl₂, 1 mM MgCl₂, 1 mM MnCl₂) containing 5% BSA was added to wells for 60 min at 37°C. Wells were washed 3x, and detection antibody (biotinylated anti-human IgG1 (Vector Laboratories, Inc.); 0.5 μ g/ml in assay buffer with 5% BSA) was added for 2 h at room temperature. Wells were washed 3x with assay buffer with 5% BSA and streptavidin-HRP was added (1:200 in assay buffer with 5% BSA; R & D Systems). Wells were washed 3x with assay buffer with 5% BSA, and a 1:1 dilution of TMB substrate (BD Biosciences) was added. The reaction was stopped by the addition of 2N H₂SO₄. Absorbance at 450 nm and 900 nm was measured using a plate reader (BioTek).

Functional Assays

Aggregation assay

L929 cells were treated with DMEM containing 0.1% trypsin at a final concentration of 5 mM CaCl₂ (30 min, 37°C). Cells were washed once with DMEM containing 10% FBS and then twice with 1X HBSS containing 1% FBS. Cells were counted, and 2 x 10⁶ cells were aliquoted into 1.5-ml tubes (two tubes per cell type). Cells were spun down and resuspended in HEPES-buffered magnesium-free saline (HMF; 10 mM HEPES in saline) that either contained or lacked 1 mM CaCl₂. The initial particle number was counted, and then tubes were rotated at 37°C for 30

min. The final particle number was then counted. The aggregation index was expressed as [(initial particle number – final particle number)/initial particle number].

Cell adhesion to recombinant integrin assay

Recombinant human $\alpha 4\beta 7$ (R&D Systems) was diluted to the indicated concentration in buffer (150 mM NaCl, 20 mM HEPES), and 100 μ l per well was added to half-well Costar ELISA plates (Corning) and incubated overnight at 4°C. The following day, wells were washed and coated with 5% BSA in buffer (150 mM NaCl, 20 mM HEPES) (3 h, 37°C). L929 cell clones transduced with the indicated construct and grown to confluency were then dispersed and resuspended in assay buffer (150 mM NaCl, 20 mM HEPES, 1 mM CaCl₂, 1 mM MgCl₂, 1 mM MnCl₂). For each cell type, 50,000 cells were added per well. The plate was spun at 10 x g for 1 min and subsequently incubated at 37°C for 1 h. Wells were washed by gravity one time by inverting the plate in a large beaker of assay buffer for 10 min. Two additional washes were performed by pipetting 100 μ l of assay buffer into the wells 4 times per wash. Fluorescence was measured prior to the washes and after each wash using a plate reader (ex/em 485/20 nm / 528/20 nm) (BioTek). The results are expressed as the percent fluorescence remaining (fluorescence after the last wash/initial fluorescence) for each well.

Cell adhesion assay

CDH26-hIgG1-Fc, CDH1-Fc (R&D Systems), and/or hIgG1 (Southern Biotech) was diluted in buffer (150 mM NaCl, 20 mM HEPES) to appropriate concentrations so that the indicated amount of protein was added to Costar half-well ELISA plates in 50 μ l per well (Corning, Inc.) and incubated overnight at 4°C. The next day, Jurkat cells (1×10^6 /ml) were incubated in HEPES

medium (132 mM NaCl, 6 mM KCl, 1 mM CaCl₂, 1 mM MgSO₄, 1.2 mM KH₂PO₄, 20 mM HEPES, 5.5 mM glucose, 0.5% BSA) plus 4 µg/ml calcein-AM (Sigma-Aldrich) for 1 h at 37°C. Cells were washed 3x in HEPES medium. In the indicated cases, cells were incubated with 1.4 µg of either anti-integrin β1-activating antibodies (clone TS2/16, Santa Cruz Biotechnology, Inc.) or an equivalent amount of mIgG1 (Southern Biotech). In the indicated cases, labeled cells were then pre-incubated with the indicated amount of either mIgG1 (AbD Serotec), anti-integrin α4 (HP2/1; AbD Serotec), or anti-CD32 (Stem Cell Technologies, Inc.) antibodies in assay buffer for non-TS2/16-treated cells (150 mM NaCl, 10 mM HEPES, 1 mM CaCl₂, 1 mM MgCl₂, 1 mM MnCl₂) or HEPES medium for TS2/16-treated cells for 15 min at 4°C prior to their addition to wells (50,000 cells/well). Plates were spun at 10 x g for 1 min and then incubated at 37°C for 45 min. Initial fluorescence per well was then measured using a plate reader ((ex/em 360/40 nm / 460/40 nm); BioTek). Wells were washed by gravity one time by inverting the plate in a large beaker of assay buffer for 10 min. Two additional washes were performed by pipetting 100 µl of assay buffer into the wells 4 times per wash. Final fluorescence was then measured.

CD4⁺ T cell activation assay

The indicated amounts of anti-CD3 (clone OKT3, eBioscience) antibodies, CDH26-hIgG1-Fc, CDH1-Fc (R&D Systems), and/or hIgG1 control antibody (Southern Biotech) were suspended in coating buffer (20 mM HEPES, 150 mM NaCl), added to wells of a 96 well cell culture plate, and incubated overnight at 4°C. The following day, isolated human peripheral blood CD4⁺ T cells (150,000 per well) were added to the wells in RPMI supplemented with 10% FBS, 1% penicillin/streptomycin, and 200 mM glutamine. Cells were incubated for 48 h. Cells and supernatants were then collected for analysis by flow cytometry and ELISA, respectively.

Statistics

Data are expressed as mean \pm SEM or median \pm interquartile range. Statistical significance was determined using the unpaired *t* test (2 groups, normal distribution, equal variance), Mann-Whitney test (2 groups, nonparametric), 1-way ANOVA followed by the Tukey post-test (>2 groups), or the Kruskal-Wallis test followed by Dunn's multiple comparison test (>2 groups, nonparametric) with Prism 5.0 software (GraphPad Software, Inc.).

Acknowledgments: We thank the CCED and clinical research coordinators for facilitating patient sample and data collection; Michael Eby for EGID database maintenance, queries, and retrieval of patient information; the Cincinnati Digestive Health Center Integrative Morphology Core for tissue processing, sectioning, histology, and immunohistochemical staining; Betsy DiPasquale for assistance with immunohistochemical stains; Dr. Andrew Herr for advice regarding protein expression and purification; Dr. Bruce Bochner for critical review of the manuscript; and physicians who collected biopsy samples including Drs. James Franciosi, Kathleen Campbell, Michael Farrell, Ajay Kaul, and Gitit Tomer. This work was supported in part by NIH U19 AI070235, NIH R01 DK076893, the PHS Grant P30 DK0789392, American Heart Association 11POST440046, the Sunshine Charitable Foundation and its supporters, Denise A. Bunning and David G. Bunning, the Buckeye Foundation, and the Campaign Urging Research for Eosinophilic Diseases (CURED) Foundation.

Author contributions: J.M.C. and M.H.C. designed and performed experiments, analyzed data, and wrote the paper; K.A.K. performed tissue sample processing and performed experiments; J.D.S., T.W., M.R., L.A., and H.T. designed and performed experiments and analyzed data; M.J.D., M.R.W., and A.P. performed protein structure and bioinformatics analyses and gave conceptual advice; E.M.S. performed tissue sample processing and assisted in acquiring clinical information; P.E.P. collected patient tissue; J.P.A. acquired clinical information and gave conceptual advice; M.E.R. designed the study, analyzed data, and wrote the paper.

Disclosure: M.E.R. is a consultant for Receptos, Pulm-One and NKT Therapeutics, and has an equity interest in Immune Pharmaceuticals, Pulm-One and NKT Therapeutics and royalties in reslizumab (Teva Pharmaceuticals). M.H.C. is a consultant for Meritage Pharma, Novartis, Receptos, Regeneron, and Aptalis. J.M.C. and M.E.R. are co-inventors of patents and/or applications, owned by Cincinnati Children's Hospital, concerning some of the data in this paper.

References:

1. Shimoyama Y, Hirohashi S, Hirano S, Noguchi M, Shimosato Y, Takeichi M *et al.* Cadherin cell-adhesion molecules in human epithelial tissues and carcinomas. *Cancer Res* 1989; **49**(8): 2128-2133.
2. Wheelock MJ, Shintani Y, Maeda M, Fukumoto Y, Johnson KR. Cadherin switching. *J Cell Sci* 2008; **121**(Pt 6): 727-735.
3. Gassler N, Rohr C, Schneider A, Kartenbeck J, Bach A, Obermuller N *et al.* Inflammatory bowel disease is associated with changes of enterocytic junctions. *Am J Physiol Gastrointest Liver Physiol* 2001; **281**(1): G216-228.
4. Heijink IH, Nawijn MC, Hackett TL. Airway epithelial barrier function regulates the pathogenesis of allergic asthma. *Clin Exp Allergy* 2014; **44**(5): 620-630.
5. Hupin C, Gohy S, Bouzin C, Lecocq M, Polette M, Pilette C. Features of mesenchymal transition in the airway epithelium from chronic rhinosinusitis. *Allergy* 2014; **69**(11): 1540-1549.
6. Trautmann A, Altnauer F, Akdis M, Simon HU, Disch R, Brocker EB *et al.* The differential fate of cadherins during T-cell-induced keratinocyte apoptosis leads to spongiosis in eczematous dermatitis. *J Invest Dermatol* 2001; **117**(4): 927-934.
7. Valencia X, Higgins JM, Kiener HP, Lee DM, Podrebarac TA, Dascher CC *et al.* Cadherin-11 provides specific cellular adhesion between fibroblast-like synoviocytes. *J Exp Med* 2004; **200**(12): 1673-1679.
8. Trautmann A, Kruger K, Akdis M, Muller-Wening D, Akkaya A, Brocker EB *et al.* Apoptosis and loss of adhesion of bronchial epithelial cells in asthma. *Int Arch Allergy Immunol* 2005; **138**(2): 142-150.
9. de Boer WI, Sharma HS, Baelemans SM, Hoogsteden HC, Lambrecht BN, Braunstahl GJ. Altered expression of epithelial junctional proteins in atopic asthma: possible role in inflammation. *Can J Physiol Pharmacol* 2008; **86**(3): 105-112.
10. Heijink IH, Kies PM, Kauffman HF, Postma DS, van Oosterhout AJ, Vellenga E. Down-regulation of E-cadherin in human bronchial epithelial cells leads to epidermal growth factor receptor-dependent Th2 cell-promoting activity. *J Immunol* 2007; **178**(12): 7678-7685.
11. Cepek KL, Parker CM, Madara JL, Brenner MB. Integrin alpha E beta 7 mediates adhesion of T lymphocytes to epithelial cells. *J Immunol* 1993; **150**(8 Pt 1): 3459-3470.
12. Cepek KL, Shaw SK, Parker CM, Russell GJ, Morrow JS, Rimm DL *et al.* Adhesion between epithelial cells and T lymphocytes mediated by E-cadherin and the alpha E beta 7 integrin. *Nature* 1994; **372**(6502): 190-193.
13. Schon MP, Arya A, Murphy EA, Adams CM, Strauch UG, Agace WW *et al.* Mucosal T lymphocyte numbers are selectively reduced in integrin alpha E (CD103)-deficient mice. *J Immunol* 1999; **162**(11): 6641-6649.
14. Uchida Y, Kawai K, Ibusuki A, Kanekura T. Role for E-cadherin as an inhibitory receptor on epidermal gammadelta T cells. *J Immunol* 2011; **186**(12): 6945-6954.
15. Teitelbaum JE, Fox VL, Twarog FJ, Nurko S, Antonioli D, Gleich G *et al.* Eosinophilic esophagitis in children: immunopathological analysis and response to fluticasone propionate. *Gastroenterology* 2002; **122**(5): 1216-1225.

16. Tantibhaedhyangkul U, Tatevian N, Gilger MA, Major AM, Davis CM. Increased esophageal regulatory T cells and eosinophil characteristics in children with eosinophilic esophagitis and gastroesophageal reflux disease. *Annals of clinical and laboratory science* 2009; **39**(2): 99-107.
17. Lucendo AJ, Navarro M, Comas C, Pascual JM, Burgos E, Santamaria L *et al.* Immunophenotypic characterization and quantification of the epithelial inflammatory infiltrate in eosinophilic esophagitis through stereology: an analysis of the cellular mechanisms of the disease and the immunologic capacity of the esophagus. *Am J Surg Pathol* 2007; **31**(4): 598-606.
18. Doherty TA, Baum R, Newbury RO, Yang T, Dohil R, Aquino M *et al.* Group 2 innate lymphocytes (ILC2) are enriched in active eosinophilic esophagitis. *J Allergy Clin Immunol* 2015; **136**(3): 792-794 e793.
19. Blanchard C, Mingler MK, Vicario M, Abonia JP, Wu YY, Lu TX *et al.* IL-13 involvement in eosinophilic esophagitis: transcriptome analysis and reversibility with glucocorticoids. *J Allergy Clin Immunol* 2007; **120**(6): 1292-1300.
20. Caldwell JM, Collins MH, Stucke EM, Putnam PE, Franciosi JP, Kushner JP *et al.* Histologic eosinophilic gastritis is a systemic disorder associated with blood and extragastric eosinophilia, TH2 immunity, and a unique gastric transcriptome. *J Allergy Clin Immunol* 2014; **134**(5): 1114-1124.
21. Bochner BS. Road signs guiding leukocytes along the inflammation superhighway. *J Allergy Clin Immunol* 2000; **106**(5): 817-828.
22. Woodruff PG, Boushey HA, Dolganov GM, Barker CS, Yang YH, Donnelly S *et al.* Genome-wide profiling identifies epithelial cell genes associated with asthma and with treatment response to corticosteroids. *Proc Natl Acad Sci U S A* 2007; **104**(40): 15858-15863.
23. Shum BO, Mackay CR, Gorgun CZ, Frost MJ, Kumar RK, Hotamisligil GS *et al.* The adipocyte fatty acid-binding protein aP2 is required in allergic airway inflammation. *J Clin Invest* 2006; **116**(8): 2183-2192.
24. Zhen G, Park SW, Nguyenvu LT, Rodriguez MW, Barbeau R, Paquet AC *et al.* IL-13 and epidermal growth factor receptor have critical but distinct roles in epithelial cell mucin production. *Am J Respir Cell Mol Biol* 2007; **36**(2): 244-253.
25. Li RW, Gasbarre LC. A temporal shift in regulatory networks and pathways in the bovine small intestine during *Cooperia oncophora* infection. *Int J Parasitol* 2009; **39**(7): 813-824.
26. Blanchard C, Wang N, Stringer KF, Mishra A, Fulkerson PC, Abonia JP *et al.* Eotaxin-3 and a uniquely conserved gene-expression profile in eosinophilic esophagitis. *J Clin Invest* 2006; **116**(2): 536-547.
27. Ikuse T, Ohtsuka Y, Kudo T, Hosoi K, Obayashi N, Jimbo K *et al.* Microarray analysis of gastric mucosa among children with *Helicobacter pylori* infection. *Pediatr Int* 2012; **54**(3): 319-324.
28. Wen S, Felley CP, Bouzourene H, Reimers M, Michetti P, Pan-Hammarstrom Q. Inflammatory gene profiles in gastric mucosa during *Helicobacter pylori* infection in humans. *J Immunol* 2004; **172**(4): 2595-2606.
29. Truong K, Ikura M. The cadherin superfamily database. *J Struct Funct Genomics* 2002; **2**(3): 135-143.
30. Shapiro L, Weis WI. Structure and biochemistry of cadherins and catenins. *Cold Spring Harb Perspect Biol* 2009; **1**(3): a003053.
31. Stappert J, Kemler R. A short core region of E-cadherin is essential for catenin binding and is highly phosphorylated. *Cell Adhes Commun* 1994; **2**(4): 319-327.

32. Jou TS, Stewart DB, Stappert J, Nelson WJ, Marrs JA. Genetic and biochemical dissection of protein linkages in the cadherin-catenin complex. *Proc Natl Acad Sci U S A* 1995; **92**(11): 5067-5071.
33. Whittard JD, Craig SE, Mould AP, Koch A, Pertz O, Engel J *et al.* E-cadherin is a ligand for integrin α 2 β 1. *Matrix Biol* 2002; **21**(6): 525-532.
34. Taraszka KS, Higgins JM, Tan K, Mandelbrot DA, Wang JH, Brenner MB. Molecular basis for leukocyte integrin α (E) β (7) adhesion to epithelial (E)-cadherin. *J Exp Med* 2000; **191**(9): 1555-1567.
35. Seminario MC, Sterbinsky SA, Bochner BS. Beta 1 integrin-dependent binding of Jurkat cells to fibronectin is regulated by a serine-threonine phosphatase. *Journal of leukocyte biology* 1998; **64**(6): 753-758.
36. Lehnert K, Print CG, Yang Y, Krissansen GW. MAdCAM-1 costimulates T cell proliferation exclusively through integrin α 4 β 7, whereas VCAM-1 and CS-1 peptide use α 4 β 1: evidence for "remote" costimulation and induction of hyperresponsiveness to B7 molecules. *Eur J Immunol* 1998; **28**(11): 3605-3615.
37. van Seventer GA, Newman W, Shimizu Y, Nutman TB, Tanaka Y, Horgan KJ *et al.* Analysis of T cell stimulation by superantigen plus major histocompatibility complex class II molecules or by CD3 monoclonal antibody: costimulation by purified adhesion ligands VCAM-1, ICAM-1, but not ELAM-1. *J Exp Med* 1991; **174**(4): 901-913.
38. Shimizu Y, van Seventer GA, Horgan KJ, Shaw S. Costimulation of proliferative responses of resting CD4⁺ T cells by the interaction of VLA-4 and VLA-5 with fibronectin or VLA-6 with laminin. *J Immunol* 1990; **145**(1): 59-67.
39. Abonia JP, Blanchard C, Butz BB, Rainey HF, Collins MH, Stringer K *et al.* Involvement of mast cells in eosinophilic esophagitis. *J Allergy Clin Immunol* 2010; **126**(1): 140-149.
40. Berg RW, Yang Y, Lehnert K, Krissansen GW. Mouse M290 is the functional homologue of the human mucosal lymphocyte integrin HML-1: antagonism between the integrin ligands E-cadherin and RGD tripeptide. *Immunol Cell Biol* 1999; **77**(4): 337-344.
41. Grundemann C, Bauer M, Schweier O, von Oppen N, Lassing U, Saudan P *et al.* Cutting edge: identification of E-cadherin as a ligand for the murine killer cell lectin-like receptor G1. *J Immunol* 2006; **176**(3): 1311-1315.
42. Salimi M, Barlow JL, Saunders SP, Xue L, Gutowska-Owsiak D, Wang X *et al.* A role for IL-25 and IL-33-driven type-2 innate lymphoid cells in atopic dermatitis. *J Exp Med* 2013; **210**(13): 2939-2950.
43. Ito M, Maruyama T, Saito N, Koganei S, Yamamoto K, Matsumoto N. Killer cell lectin-like receptor G1 binds three members of the classical cadherin family to inhibit NK cell cytotoxicity. *J Exp Med* 2006; **203**(2): 289-295.
44. Fuentebella J, Patel A, Nguyen T, Sanjanwala B, Berquist W, Kerner JA *et al.* Increased number of regulatory T cells in children with eosinophilic esophagitis. *J Pediatr Gastroenterol Nutr* 2010; **51**(3): 283-289.

Figure Legends

Figure 1. *CDH26* expression in allergic tissue. **(a)** Relative esophageal tissue *CDH26* transcript levels were determined ($n = 66$ control [CTL], $n = 77$ eosinophilic esophagitis [EoE] patients). **(b)** and **(c)**. The mean raw expression value for each cadherin probe in which any patient sample exhibited a signal intensity greater than 100 is graphed for **(b)** the gastric tissue of 5 CTL and 5 eosinophilic gastritis (EG) patients²⁰ or for **(c)** the esophageal tissue of 14 CTL and 18 EoE patients characterized previously²⁶. For each cadherin, only the probe with the highest raw signal is shown. **(d)** Relative *CDH26*, *CCL26* (eotaxin-3), and *C3* transcript levels from the gastric antrum tissue of CTL patients ($n = 5$), EG patients ($n = 5$), and *H. pylori* gastritis patients ($n = 3$) were determined. For **(a)** and **(d)**, data were analyzed by Mann-Whitney test.

Figure 2. *CDH26* expression and localization in epithelial cells in allergic gastrointestinal tissue. **(a)** Representative control (CTL) and eosinophilic gastritis (EG) patient gastric biopsy specimen (200X) stained with anti-*CDH26* antibody or control IgG. **(b)** CTL and EG biopsy specimen (800X) stained with anti-*CDH26* antibody. **(c)** Left: Gastric antrum protein lysates were subjected to SDS-PAGE and western blot analysis. Right: The ratio of *CDH26* to beta-actin signal was graphed. **(d)** Immunohistochemical staining for *CDH26* was performed on esophageal biopsy specimens ($n = 7$ CTL, $n = 3$ EoE). Representative CTL and EoE biopsies are shown (200X, 800X inset). **(e)** Left: Esophageal protein lysates ($n = 4$ CTL, $n = 4$ EoE) were subjected to SDS-PAGE and western blot analysis. Right: The ratio of *CDH26* to beta-actin signal was graphed. For **(c)** and **(e)**, data were analyzed by unpaired t test. For **(a)**, **(b)**, and **(d)**, the location of the lumen is denoted to facilitate interpretation of the orientation of the tissue section.

Figure 3. Biochemical and functional properties of CDH26. **(a)** Schematic representing human CDH26 domain structure. S: signal peptide, EC1-EC5: extracellular cadherin repeat 1-5, T: transmembrane domain, JMD: juxtamembrane domain, C: CBD, catenin binding domain. **(b)** Surface biotinylation of TE-7 cells. Cell surface proteins were labeled with biotin and pulled down with streptavidin beads. Total cell lysates (input) and proteins bound to the streptavidin beads were subjected to SDS-PAGE and western blot analysis. Predicted CDH26 molecular weight: 92.4 kDa. **(c)** Immunoprecipitates from transiently transfected HEK 293T cells were treated with either peptide: *N*-glycosidase F (PNGase F) (+) or heat-inactivated PNGase F (-). Inputs (1/10 of amount used for IP) and treated immunoprecipitates were subjected to SDS-PAGE and western blot analysis. Each blot shown is representative of three independent experiments. **(d-g)** Immunoprecipitates from transiently transfected HEK 293T cells and inputs (1/10 of amount used for IP) were subjected to SDS-PAGE and western blot analysis. Each blot shown is representative of 3 independent experiments. **(h)** Transduced L929 cells were dispersed, incubated in buffer either containing or lacking 1 mM CaCl₂, and assessed for the degree of aggregation. Data show 1 experiment representative of 3 and were analyzed by one-way ANOVA followed by Tukey post-test.

Figure 4. Binding of CDH26 to integrins. **(a)** Pairwise structure alignment of CDH26 with known integrin ligands. CDH26 structure was modeled (blue) and aligned to the resolved structures of MAdCAM-1 (PDB ID 1BQS), ICAM-1 (PDB ID 1IC1), CDH1 (PDB ID 1EDH), and fibronectin (PDB ID 1FNF) (gray). For each pair, integrin binding amino acids and the corresponding CDH26 residues are labeled (arrows) and rendered using a stick representation.

(b) Transduced L929 cell clones were dispersed and added to wells coated with either BSA or recombinant $\alpha 4\beta 7$. The percentage of adherent cells remaining after wells were washed is shown. The graph represents seven experiments combined that each involved separate control and CDH26-overexpressing clones. (c) Pictures of Giemsa-stained wells from b were taken (magnification = 4X), with 1 control and 1 CDH26-overexpressing clone shown. (d) Wells were coated with or without recombinant $\alpha 4\beta 7$ and then blocked with BSA, followed by addition of either hIgG1 or CDH26-hIgG1-Fc (CDH26-Fc). Bound antibody or fusion protein was then detected and expressed as $A_{450nm}-A_{900nm}$. Each condition was performed in triplicate. This graph shows 1 experiment representative of 3. (e-h) Inputs (1/10 of amount used for IP) and immunoprecipitates from transiently transfected HEK 293T cells were subjected to SDS-PAGE and western blot analysis. Each blot shown is representative of 3 independent experiments. (i-j) Fluorescently labeled Jurkat cells (untreated [i] or incubated with TS2/16 integrin $\beta 1$ -activating antibodies [j]) pre-incubated with the indicated amount of either control mIgG1, anti-integrin $\alpha 4$ (HP2/1), or anti-CD32 antibodies were added to wells coated with control hIgG1, CDH26-hIgG1-Fc (CDH26-Fc), or CDH1-hIgG1-Fc (CDH1-Fc), as indicated. The graph indicates the percentage of fluorescence remaining after wells were washed. For (b), (d), (i), and (j), data were analyzed by one-way ANOVA followed by Tukey post-test.

Figure 5. Effect of CDH26-Fc and CDH1-Fc on CD4⁺ T cell activation. Human peripheral blood CD4⁺ T cells were isolated and cultured for 48 h in wells coated with the indicated amounts of proteins (IgG, anti-CD3, CDH26-Fc, and/or CDH1-Fc). Cells were stained for flow cytometry analysis to detect CD4 and CD25, and supernatants were analyzed to detect IL-2 levels by ELISA. For (a) and (c), the percent of live CD4⁺ cells that are CD25⁺ are shown, and for (b) and

(d), the amount of IL-2 detected in the supernatant is shown. The dotted lines represent the detection limit for the ELISA. Data were analyzed by one-way ANOVA followed by Tukey post-test. Data are results from one subject representative of those from five individual subjects for (a) and (b), and from one subject representative of those from four individual subjects for (c) and (d).

Figure 6. Model of CDH26 expression and function in allergic inflammation. CDH26 is expressed by GI epithelial cells in allergic GI inflammatory conditions. CDH26 dimerizes, interacts with beta-, alpha-, and p120-catenins, and mediates calcium-dependent cell adhesion. CDH26 additionally interacts with integrin $\alpha 4$ (ITGA4) and integrin αE (ITGAE), which may impact leukocyte migration, localization, or activation status in allergic tissue.

Figure 1

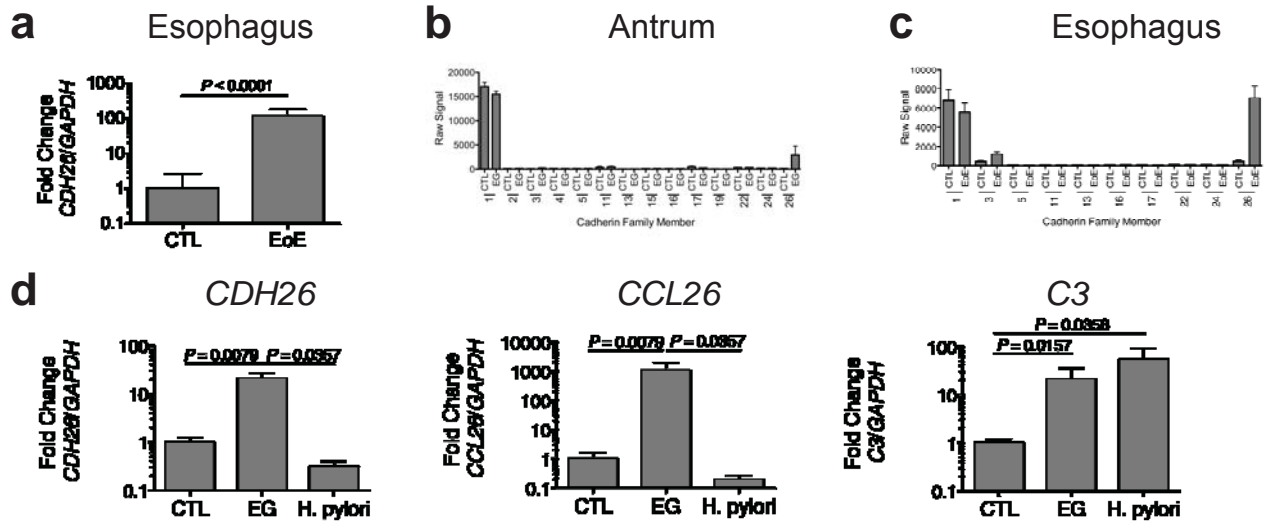


Figure 2

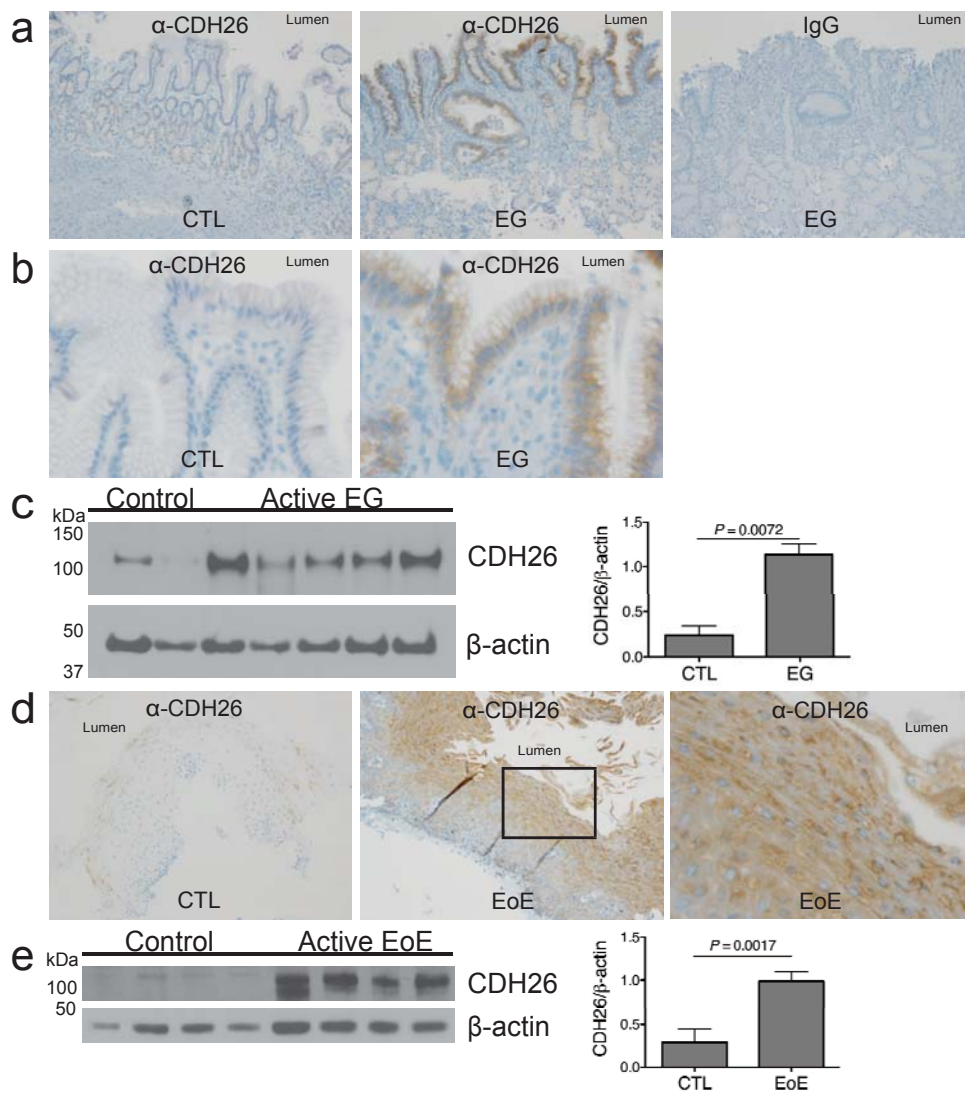


Figure 3

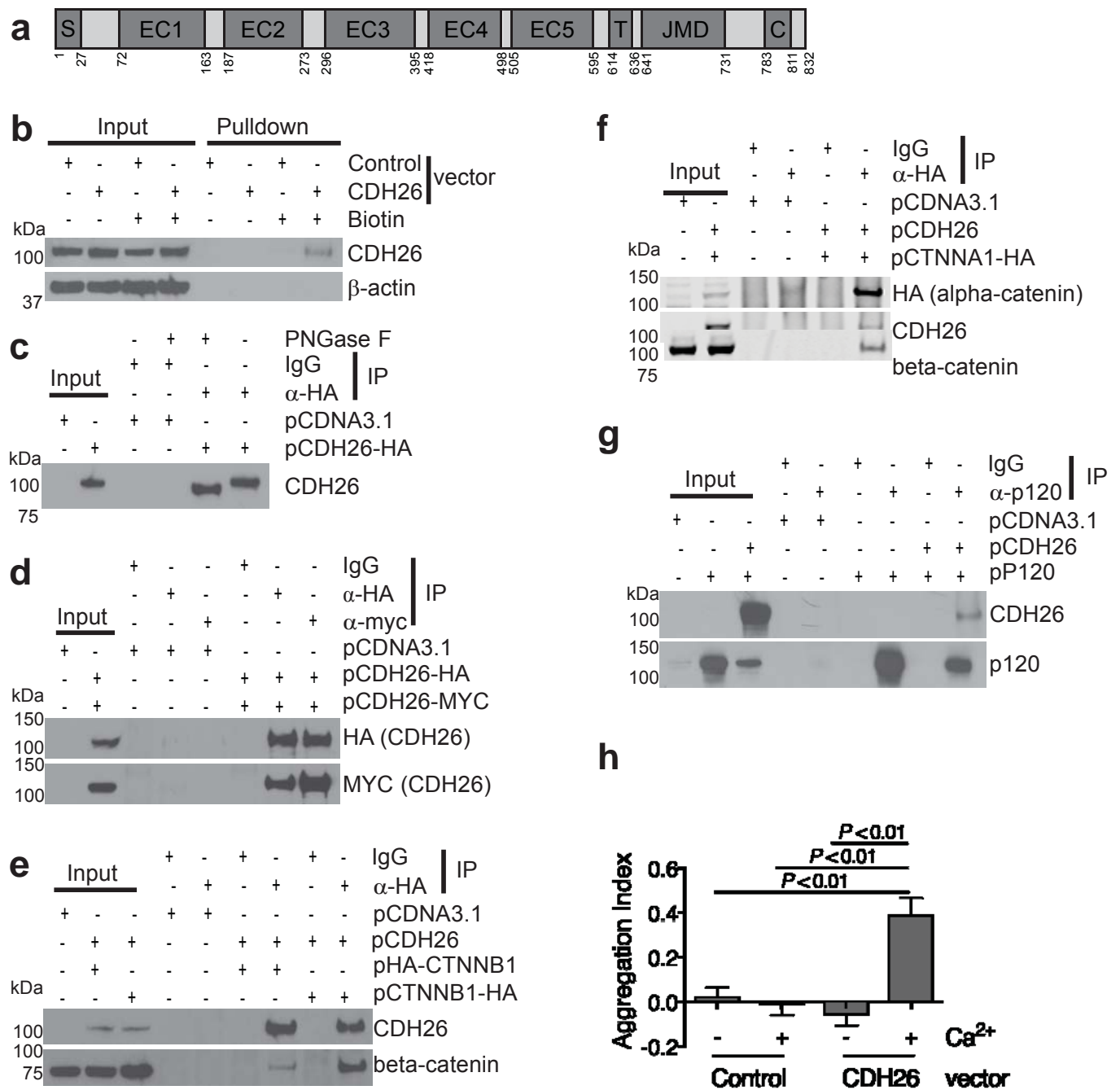


Figure 4

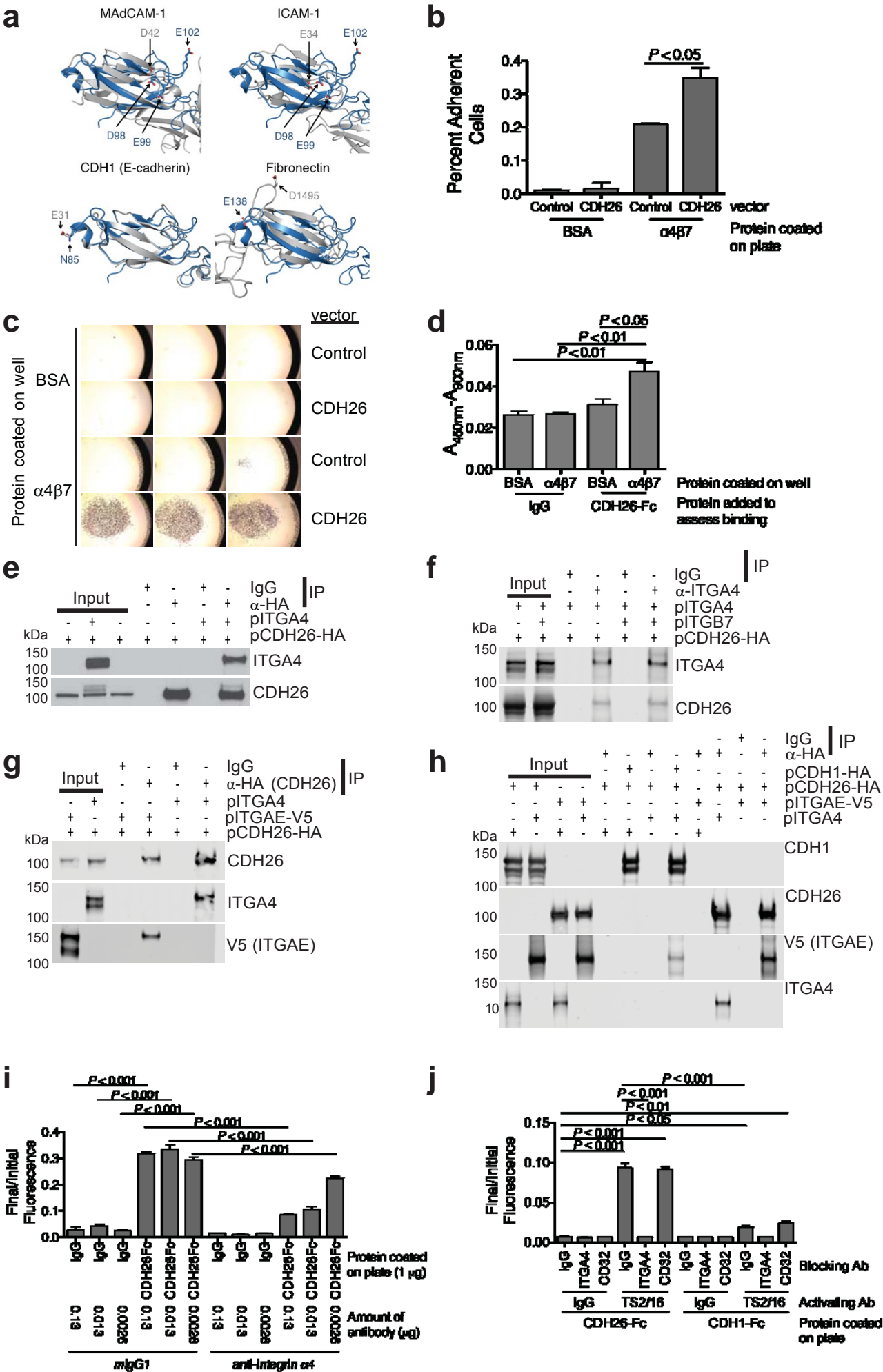


Figure 5

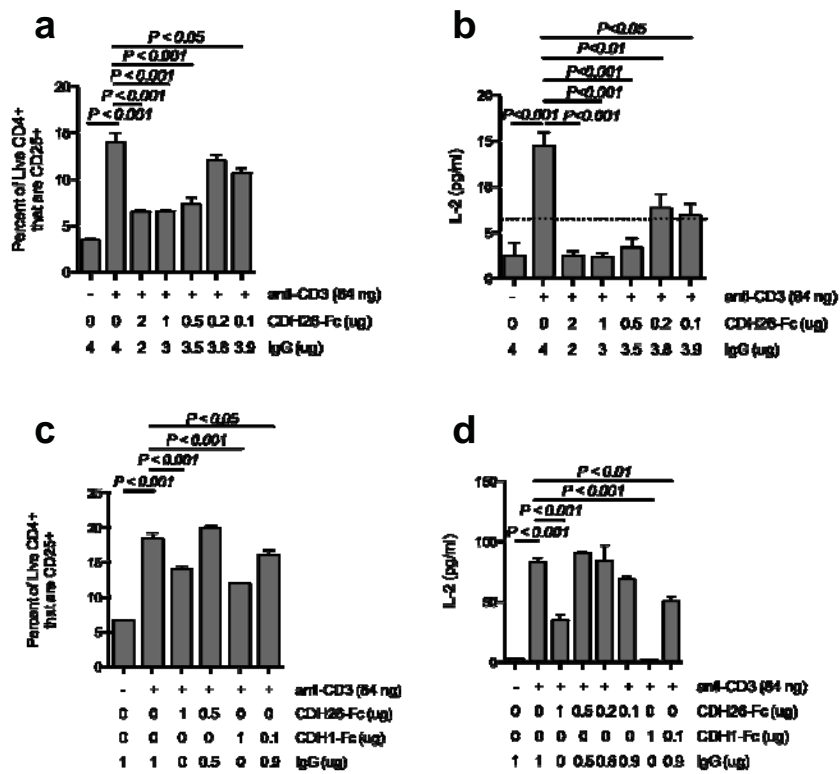
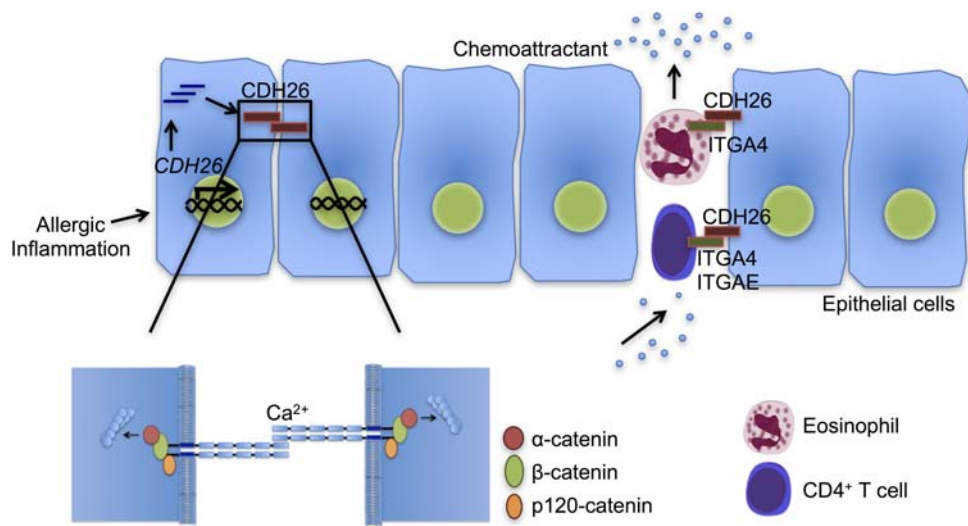


Figure 6



Supplementary Material

Cadherin 26 is an alpha integrin-binding epithelial receptor regulated during allergic inflammation

JM Caldwell¹, MH Collins², KA Kemme¹, JD Sherrill¹, T Wen¹, M Rochman¹, EM Stucke¹, L Amin¹, H Tai¹, PE Putnam³, MJ Jiménez-Dalmaroni⁴, MR Wormald⁵, A Porollo⁶, JP Abonia¹, and ME Rothenberg^{1*}

Supplementary Methods

Patient studies

Entry criteria

Patients in this study included those who had signs and symptoms consistent with upper gastrointestinal (GI) tract disease and who underwent endoscopy and had biopsies taken from various GI segments. The endoscopic procedure at which samples for histopathology, microarray, and PCR analyses were obtained is designated the index endoscopy. Gastric samples were obtained from the antrum or body, and esophageal samples were obtained from the distal esophagus. Duodenum samples were also acquired during all endoscopies, and colon samples were obtained in some cases. This study was performed with the approval of the Cincinnati Children's Hospital Medical Center (CCHMC) IRB. Informed consent was obtained from patients or their legal guardians to donate tissue samples for research and to have their clinical information entered into the Cincinnati Center for Eosinophilic Disorders (CCED) Eosinophilic Gastrointestinal Disorder (EGID) Database.

Definition of active EGID and control patients

Patients were defined to have active eosinophilic gastritis (EG) if they had gastric biopsies that met the criteria for histological EG (i.e. ≥ 127 eosinophils/mm² [≥ 30 peak eosinophils/mm²] in five separate 400X high-power fields [HPF])¹ and that also showed architectural abnormalities, including excessively branched and/or coiled glands. Similar criteria were used to diagnose eosinophilic duodenitis (ED), eosinophilic jejunitis (EJ), and eosinophilic colitis (EC), where increased numbers of eosinophils were defined as at least twice the normal level of eosinophils reported previously². Eosinophilic esophagitis (EoE) was diagnosed if at least 15

eosinophils/HPF were present in esophageal biopsies in patients on proton pump inhibitor (PPI) therapy³. Control patients included patients who had undergone endoscopy, had no history of any EGID, had normal pathology results for all GI tract segments surveyed during the index endoscopy, were not taking any swallowed or oral glucocorticoid treatments in the period immediately prior to the index endoscopy, and also had gastric samples collected for research purposes during the index endoscopy. After identifying all patients who met these control criteria from the EGID database, the ones most closely matching the EG patient cohort were selected based on the criteria of 1: sex and 2: age. Clinical information was obtained from the EGID Database and review of medical records. Patient information is listed in Supplementary Table S2.

Expression Constructs

pCDNA3.1(-) was obtained from Promega. pLX304-hITGAE-V5 (ID HsCD00436970) was obtained from DNASU Plasmid Repository (The Biodesign Institute, Arizona State University). Expression plasmids were constructed by PCR amplification of the relevant open reading frame using primers listed in Supplementary Table S4. The following primers were used: pCDH26-HA: 5170 and 4242, pCDH26-MYC: 5170 and 4241, pCDH26: 5170 and 4105, pHA-CTNNB1: 4590 and 4367, pCTNNB1-HA: 4366 and 4593, pCTNNA1-HA: 4370 and 4701, pCTNND1: 4468 and 4469, pITGA4: 6074 and 6075, pITGB7: 6078 and 6079, pCDH1: 5427 and 5140. PCR products were then ligated into the following restriction sites of pCDNA3.1(-): pCDH26-HA: EcoRI/KpnI, pCDH26-MYC: EcoRI/KpnI, pCDH26: EcoRI/NotI, pHA-CTNNB1: XbaI/KpnI, pCTNNB1-HA: XbaI/KpnI, pCTNNA1-HA: EcoRI/KpnI, pCTNND1: EcoRI/KpnI, pITGA4: XhoI/NotI, pITGB7, NotI/EcoRI, pCDH1: NotI/HindIII. pMIRNA1-puro-control has

been described previously and is referred to as “pMIRNA1-control” henceforth. pMIRNA1-CDH26 was constructed by introducing the CDH26 open reading frame into the EcoRI and NotI sites of pMIRNA1-control. pFUSE-CDH26-hIgG1-Fc was constructed by first amplifying nucleotides encoding the extracellular domain of CDH26 using primers 5067 and 5108 with pCDH26 as template. These primers introduced a region of nucleotide sequence complementary to the 5' portion of hIgG1-Fc at the 3' end of the PCR product. This PCR product was isolated and used as a primer along with primer 5109 in a second PCR reaction using pFUSE-hIgG1-Fc1 plasmid (Invivogen) as template. The resulting PCR product was then digested with EcoRI and BsrGI restriction enzymes and ligated into the corresponding sites of pFUSE-hIgG1-Fc1 plasmid. This construct and pCDH26 were then digested with EcoRI and Sall, and the 1329 base pair restriction fragment from pCDH26 was ligated into the corresponding restriction sites of the pFUSE-based EcoRI/Sall-digested construct. Primers 5254 and 7491 were then used to amplify the CDH26-Fc open reading frame, and it was introduced into the NotI restriction site of pMIRNA1-control to make pMIRNA1-CDH26-hIgG1-Fc.

Production of CDH26-hIgG1-Fc

HEK 293T cells (20 x 150 mm dishes) stably transduced with pMIRNA1-CDH26-hIgG1-Fc (described in Constructs) were grown in DMEM containing 10% ultra-low IgG FBS (Invitrogen). After 8 days, the supernatants from all dishes were combined. The supernatants were filtered (0.2 μ m) and the resulting volume was diluted 1:1 in binding buffer (20 mM HEPES, 50 mM NaCl, 1 mM CaCl₂, pH 7.4) and then passed over a protein G column (GE Healthcare) that had been pre-washed with binding buffer. The column was then washed with binding buffer, and fusion protein was eluted in fractions using 100 mM glycine, pH 3.0. The

eluates were immediately neutralized by the addition of 1 M Tris HCl, pH 8.0. Fractions were assayed by SDS-PAGE and Coomassie stain as well as western blot analysis to identify the fractions that contained the fusion protein. The pooled five fractions containing the most fusion protein were desalted using G20 columns (GE Healthcare) and binding buffer. Control hIgG1 antibody (low endotoxin, azide-free preparation) (Southern Biotech) was also desalted in a similar manner. Quantitation of LPS in the isolated proteins using the Limulus amoebocyte lysate (LAL) chromogenic endotoxin quantitation kit (Thermo Scientific) showed that the CDH26-Fc preparations had endotoxin concentrations of less than 0.2 endotoxin units (EU)/ml, which was lower than that detected in the control hIgG1 antibody (Supplementary Figure S5).

Isolation of CD4⁺ T cells from human peripheral blood

CD4⁺ T cells were isolated from the peripheral blood mononuclear cells (PBMCs) derived from normal human donors using the Dynabeads Untouched Human CD4 T cells Kit (Life Technologies) according to the manufacturer's protocol. Briefly, heparinized blood was diluted with at least 3 parts 1X DPBS + 2 mM EDTA and layered onto Ficoll-Paque Plus (GE Healthcare) and centrifuged for 10 min at 1000 x g. PBMCs were collected, and 3 washes with 1X DPBS + 2 mM EDTA followed by centrifugation (10 min at 300 x g twice; 10 min at 200 x g once) were done to eliminate platelets. PBMCs were incubated with the antibody cocktail provided in the kit, washed Dynabeads were added, and tubes were placed in a magnet to separate the cell-bound beads from the non-bead-bound cells. Non-bead-bound cells (i.e., CD4⁺ T cells) were collected and either used immediately in functional assays or subjected to FACS analysis to determine purity and viability. Cells were routinely >91% CD4⁺ and >98% viable.

FACS analysis

Human peripheral blood mononuclear cells or human peripheral blood CD4⁺ T cells were washed with 1X DPBS and then incubated in 100 μ l 1X DPBS containing 1 μ l of Zombie Violet dye (BioLegend) for 15 minutes. Cells were then washed in FACS buffer and incubated for 30 min at 4°C in the dark in FACS buffer containing primary antibodies or an equivalent amount of the appropriate isotype control antibodies (anti-CD4-FITC [clone RPA-T4, Biolegend 300506], FITC isotype [clone MOPC-21, BioLegend 400109], anti-CD25-AF647 [clone BC96, BioLegend 302617], AF647 isotype [clone HTK888, BioLegend 400924]. Cells were washed in FACS buffer, fixed in 2% paraformaldehyde in FACS buffer for 10 min at room temperature, washed again with FACS buffer, and resuspended in FACS buffer for analysis. For all cell types, flow cytometry analysis was performed using the FACS Canto (BD Biosciences) and analysis was done using FlowJo software (TreeStar). Unstained cells and fluorescence minus one (FMO) controls were used to set gates for each fluorophore, and compensation was performed using data obtained from eCompUltra beads (eBioscience) that were stained with antibodies conjugated to the fluorophores relevant to each experiment.

ELISA

ELISA to detect human IL-2 was performed generally according to the manufacturer's instructions (eBioscience). Briefly, wells of half-area 96 well plates (Costar) were coated overnight (4°C) with buffer containing capture antibody. Plates were then washed with wash buffer (1X DPBS + 0.05% Tween-20) three times. Blocking buffer was added to each well and plates were incubated at room temperature for 1 h. Plates were then washed three times with wash buffer. Samples or standards (50 μ l) were added to wells and incubated at room

temperature for 2 h. Plates were washed with wash buffer three times, followed by addition of buffer containing detection antibody (50 μ l per well). Plates were incubated at room temperature for 1 h, washed 3 times with wash buffer, followed by addition of buffer containing streptavidin-HRP. Plates were incubated at room temperature for 30 min and then washed with wash buffer 5 times. TMB substrate solution (BD Biosciences) was then added. The reaction was stopped by addition of 2N H₂SO₄. Absorbance at 450 nm and 900 nm was measured using a plate reader (BioTek).

Bioinformatics analysis

Protein homology and structural prediction analyses

Protein domain prediction was done by submitting the primary amino acid sequence of human CDH26 (Uniprot accession Q8IXH8-4, CDH26 isoform 3) to SMART (Simple Modular Architecture Research Tool) protein analysis program^{4,5} and was also compared to the juxtamembrane domain and catenin binding domain of other cadherin molecules. Multiple sequence alignment was done using Constraint-based Multiple Protein Alignment Tool (COBALT)⁶, and the multiple sequence alignments generated were input into Clustal Omega⁷⁻⁹ for visualization. A 3D model of CDH26 structure was obtained using the Phyre2 server¹⁰. Resolved 3D structures of CDH1, MAdCAM-1, fibronectin, and ICAM-1 were obtained from Protein Databank (PDB IDs: 1EDH, 1BQS, 1FNF, and 1IC1, respectively). Protein structure alignment was done using the DaliLite server¹¹. Protein visualization was performed using PyMOL (The PyMOL Molecular Graphics System, Version 0.99rc6, DeLano Scientific, LLC). For the CDH26 model shown in Supplementary Figure S2d, molecular modeling was performed on a Silicon Graphics Fuel workstation using Insight II and Discover software (Accelrys Inc.,

San Diego, USA). Figures were produced using PyMOL. Protein structures used for modeling were obtained from the Protein Databank database¹². The structure of the ectodomain of glycosylated human CDH26 was based on the crystal structures of *Xenopus* C-cadherin ectodomain¹³. The intracellular domain was modeled as fully disordered, based on the SPINE-D disorder prediction¹⁴, which only predicted two very short regions not to be disordered. Biantennary complex glycans were added to each possible *N*-glycosylation site in the ectodomain (asparagines: 81, 85, 171, 175, 462). *N*-linked glycan structures were generated using the database of glycosidic linkage conformations¹⁵ and *in vacuo* energy minimisation to relieve unfavorable steric interactions. The Asn-GlcNAc linkage conformations were based on the observed range of crystallographic values¹⁶, the torsion angles around the Asn C α -C β and C β -C γ bonds then being adjusted to eliminate unfavorable steric interactions between the glycans and the protein surface. The structure of human CDH26 in complex with p120 and beta-catenin was based on the crystal structures of the complex between human p120 and E-cadherin¹⁷ and the complex between mouse beta-catenin and E-cadherin¹⁸.

Accession codes

Gastric tissue microarray data are deposited in Gene Expression Omnibus (GEO) (Accession GSE54043; <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54043>).

Supplementary References

1. Lwin T, Melton SD, Genta RM. Eosinophilic gastritis: histopathological characterization and quantification of the normal gastric eosinophil content. *Mod Pathol* 2011; **24**(4): 556-563.
2. DeBrosse CW, Case JW, Putnam PE, Collins MH, Rothenberg ME. Quantity and distribution of eosinophils in the gastrointestinal tract of children. *Pediatric and developmental pathology : the official journal of the Society for Pediatric Pathology and the Paediatric Pathology Society* 2006; **9**(3): 210-218.
3. Liacouras CA, Furuta GT, Hirano I, Atkins D, Attwood SE, Bonis PA *et al.* Eosinophilic esophagitis: updated consensus recommendations for children and adults. *J Allergy Clin Immunol* 2011; **128**(1): 3-20 e26; quiz 21-22.
4. Letunic I, Doerks T, Bork P. SMART 7: recent updates to the protein domain annotation resource. *Nucleic acids research* 2012; **40**(Database issue): D302-305.
5. Schultz J, Milpetz F, Bork P, Ponting CP. SMART, a simple modular architecture research tool: identification of signaling domains. *Proc Natl Acad Sci U S A* 1998; **95**(11): 5857-5864.
6. Papadopoulos JS, Agarwala R. COBALT: constraint-based alignment tool for multiple protein sequences. *Bioinformatics* 2007; **23**(9): 1073-1079.
7. Li W, Cowley A, Uludag M, Gur T, McWilliam H, Squizzato S *et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic acids research* 2015; **43**(W1): W580-584.
8. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J *et al.* A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic acids research* 2010; **38**(Web Server issue): W695-699.
9. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W *et al.* Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular systems biology* 2011; **7**: 539.
10. Kelley LA, Sternberg MJ. Protein structure prediction on the Web: a case study using the Phyre server. *Nature protocols* 2009; **4**(3): 363-371.
11. Hasegawa H, Holm L. Advances and pitfalls of protein structural alignment. *Current opinion in structural biology* 2009; **19**(3): 341-348.
12. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, Weissig H *et al.* The Protein Data Bank. *Nucleic acids research* 2000; **28**(1): 235-242.
13. 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; **296**(5571): 1308-1313.
14. Zhang T, Faraggi E, Xue B, Dunker AK, Uversky VN, Zhou Y. SPINE-D: accurate prediction of short and long disordered regions by a single neural-network based method. *J Biomol Struct Dyn* 2012; **29**(4): 799-813.
15. Wormald MR, Petrescu AJ, Pao YL, Glithero A, Elliott T, Dwek RA. Conformational studies of oligosaccharides and glycopeptides: complementarity of NMR, X-ray crystallography, and molecular modelling. *Chem Rev* 2002; **102**(2): 371-386.
16. Petrescu AJ, Milac AL, Petrescu SM, Dwek RA, Wormald MR. Statistical analysis of the protein environment of N-glycosylation sites: implications for occupancy, structure, and folding. *Glycobiology* 2004; **14**(2): 103-114.

17. Ishiyama N, Lee SH, Liu S, Li GY, Smith MJ, Reichardt LF *et al.* Dynamic and static interactions between p120 catenin and E-cadherin regulate the stability of cell-cell adhesion. *Cell* 2010; **141**(1): 117-128.
18. Huber AH, Weis WI. The structure of the beta-catenin/E-cadherin complex and the molecular basis of diverse ligand recognition by beta-catenin. *Cell* 2001; **105**(3): 391-402.
19. Blanchard C, Wang N, Stringer KF, Mishra A, Fulkerson PC, Abonia JP *et al.* Eotaxin-3 and a uniquely conserved gene-expression profile in eosinophilic esophagitis. *J Clin Invest* 2006; **116**(2): 536-547.
20. Blanchard C, Mingler MK, Vicario M, Abonia JP, Wu YY, Lu TX *et al.* IL-13 involvement in eosinophilic esophagitis: transcriptome analysis and reversibility with glucocorticoids. *J Allergy Clin Immunol* 2007; **120**(6): 1292-1300.
21. Sugihara T, Kobori A, Imaeda H, Tsujikawa T, Amagase K, Takeuchi K *et al.* The increased mucosal mRNA expressions of complement C3 and interleukin-17 in inflammatory bowel disease. *Clin Exp Immunol* 2010; **160**(3): 386-393.

Supplementary Figure Legends

Supplementary Figure S1. *CDH26* gene expression in gastric tissue. **(a)** *CDH26* and *GAPDH* transcript levels from the same patient samples subjected to microarray analysis (n = 5 control patients [CTL], n = 5 patients with active EG [EG]) were quantified by real-time (RT-)PCR. *CDH26* levels were normalized to *GAPDH* levels for each sample and are presented as fold-change relative to control. **(b)** Relative *CDH26* levels from a replication cohort of patients (n = 10 CTL, n = 10 EG) were determined by RT-PCR. Data are presented as fold-change relative to control. For **(a)** and **(b)**, data were analyzed using the Mann Whitney test.

Supplementary Figure S2. Similarity of human *CDH26* protein domains to those of other cadherin molecules and structural modeling of human *CDH26*. **(a)** *CDH26* primary amino acid sequence was aligned with that of other cadherin molecules by multiple sequence alignment using COBALT, and alignments were visualized using Clustal Omega. Residues and domains of interest are marked with arrows or underlined, and the associated numbers denote the following features: (1) W56, conserved W residue at position 2 of EC1; (2) putative N-glycosylation sites N81, N85, N171, N177, N462; (3) residues or motifs shown to be important for integrin binding in *CDH1* or other adhesion molecules; (4) transmembrane domain, (5) the juxtamembrane domain; (6) the catenin binding domain. **(b)** Top: The catenin binding domain (CBD/CM2) of *CDH1* was aligned to the cytoplasmic domain of other cadherins. Bottom: The residues aligning to the *CDH1* CBD are shown. **(c)** Top: The juxtamembrane domain (JMD) of *CDH1* was aligned to the cytoplasmic domain of other cadherins. Bottom: The residues corresponding to the *CDH1* JMD are shown. **(d)** Three-dimensional structure model of human *CDH26* in complex with p120

and beta-catenin. Human CDH26 is depicted in red, its N-linked glycosylation in blue, calcium in yellow, p120 in green, and beta-catenin in cyan. The predicted disordered regions are indicated as pointed lines.

Supplementary Figure S3. Properties of CDH26-hIgG1-Fc. **(a)** Schematic representation of CDH26-hIgG1-Fc fusion protein, including the nucleotide sequence encoding the junction between the CDH26 extracellular domain and the Fc portion of human IgG1 (Fc) compared to the corresponding sequence encoding the wild-type protein (WT). EC1-EC5: extracellular cadherin repeat 1-5; CH2, CH3: heavy chain constant domains. **(b)** CDH26-hIgG1-Fc fusion protein isolated by protein G affinity chromatography was subjected to western blot analysis to detect human IgG to verify the presence of the IgG1-Fc portion of the molecule and also CDH26 to detect the presence of the extracellular domain of CDH26 in the chimeric molecule. Lower arrow, CDH26-hIgG1-Fc monomer; Upper arrow, CDH26-hIgG1-Fc dimer.

Supplementary Figure S4. Analysis of surface expression of CD25 on live CD4⁺ T cells. **(a)** Representative flow cytometry plots for the data shown in Figure 5a are shown. CD25 expression of live (Zombie Violet negative), CD4⁺ cells is shown. The percentage of events in the upper right quadrant of this graph represent data plotted in Figure 5a. **(b)** Representative flow cytometry plots for the data shown in Figure 5c are shown. CD25 expression of live (Zombie Violet negative), CD4⁺ cells is shown. The percentage of events in the upper right quadrant of this graph represent data plotted in Figure 5c.

Supplementary Figure S5. Detection of endotoxin content of CDH26-hIgG1-Fc preparations. Two CDH26-hIgG1-Fc (CDH26-Fc) preparations (prepared on two separate days) and human IgG1 (low endotoxin, azide-free [LEAF] preparation) used as a control for this construct were subjected to analysis using the Limulus amoebocyte lysate (LAL) assay, and the concentration of endotoxin (Endotoxin Units [EU]/ml) for each preparation is shown.

Supplementary Table S1. Comparison of characteristics of CDH1 (E-cadherin) and CDH26

	CDH1	CDH26
CDH type	C-1 subfamily, Type I	C-1 subfamily, solitary
Expression (baseline)	High	Low
Expression (allergic tissue)	High	Moderate
Inducible (allergic tissue)	No	Yes
Catenin binding	Yes	Yes
αE binding	Yes	Yes
α4 binding	No	Yes

Supplementary Table S2. Index endoscopy information

Patient Number	Gastric DX ^A	Age (yr) ^B	Sex	Indication ^C	Endoscopic Findings ^D	DX ^E	Prior EGID DX ^F	Duration of EG (yr) ^G	Medication at IE ^H	Diet at IE ^I	Asthma ^J	AC ^J	AR ^J	HF ^J	U/A ^J	Eczema ^J	Drug Allergy ^J	Food Allergy ^J	E Allergy ^J	SPT ^K	Clo ^L	Eosinophils ^M
1	CTL	1	F	Vomiting, siblings with EoE	Normal	NDA	None	None	PPI	U	N	N	N	N	N	N	N	Y	Y	NT	Neg	17, 5, 9, 3, 3
2	CTL	15	M	Rectal bleeding	Normal	NDA	None	None	Miralax	Ad lib	N	N	N	N	N	N	N	N	N	NT	Neg	2, 3, 0, 0, 0
3	CTL	1	F	Daily spit ups	Normal	NDA	None	None	PPI, hydrocortisone cream	Ad lib	N	N	N	N	Y	N	N	U	U	NT	U	8, 6, 4, 5, 9
4	CTL	14	F	Abdominal pain	Normal	NDA	None	None	PPI	Ad lib	Y	N	N	N	N	N	N	Y	Y	NT	Neg	9, 12, 6, 15, 12
5	CTL	14	F	Abdominal pain	Normal	NDA	None	None	Bactroban	Ad lib	N	N	Y	Y	N	N	N	Y	Y	NT	Neg	3, 5, 4, 0, 0
6	CTL	17	M	Abdominal pain	Normal	NDA	None	None	PPI	Ad lib	N	N	N	U	N	N	N	N	N	NT	Neg	6, 2, 1, 1, 2
7	CTL	4	M	Vomiting, gagging/choking	Normal	NDA	None	None	H2RA	Self-restricted	N	N	N	N	N	N	N	NT	NT	NT	Neg	11, 8, 6, 5, 5
8	CTL	9	M	Reflux, abdominal pain	Normal	NDA	None	None	PPI	Elim	N	N	N	N	N	N	Y	NT	NT	NT	Neg	0, 9, 1, 1, 0
9	CTL	5	M	Regurgitation	Erythema in distal esophagus	NDA	None	None	PPI	Ad lib	N	N	Y	N	N	N	N	NT	NT	NT	Neg	4, 1, 1, 1, 0
10	CTL	4	M	Dyspepsia	Normal	NDA	None	None	None	Ad lib	N	N	N	N	Y	N	N	NT	NT	NT	Neg	6, 7, 12, 5, 0
11	CTL	13	F	Abdominal pain	Normal	NDA	None	None	None	Ad lib	N	N	Y	Y	Y	N	Y	NT	NT	NT	Neg	11, 8, 8, 7, 1
12	CTL	13	M	Vomiting	Normal	NDA	None	None	None	Ad lib	N	Y	Y	Y	N	N	N	N	N	NT	Neg	0, 0, 0, 3, 2
13	CTL	2	M	FTT	Normal	NDA	None	None	None	Ad lib	N	N	N	Y	Y	Y	N	N	N	U	U	6, 6, 7, 6, 7
14	CTL	8	F	Abdominal pain	Normal	NDA	None	None	H2RA, PPI	Ad lib	N	N	N	N	N	N	N	NT	NT	NT	Neg	11, 15, 9, 10, 16
15	CTL	11	F	Abdominal pain	Normal	NDA	None	None	INHGC, PPI, B	Ad lib	N	N	N	N	N	N	N	N	Y	NT	Neg	13, 25, 4, 9, 37
16	EG	3	F	Duodenal mass	Antrum and jejunum: nodules	EG, EIT	EG, EIT	0.2	Ferrous sulfate	Ad lib	U	U	U	U	U	U	U	Y	U	Fo, E	Neg	488, 546, 493, 398, 252
17	EG	13	M	medication change	Fundus and body: nodules; Antrum: normal	EG	EG	1.4	M1A, iron acid, PPI	Ad lib	N	N	N	N	N	N	N	N	N	Neg	U	114, 137, 101, 102, 97

18	EG	20	F	Abdominal pain	Esophagus: thickened muocsa; Antrum: erythema	EG, EoE	EoE	None	INHGC, SWGC, PPI	Elim	Y	Y	Y	Y	Y	Y	Y	Y	Y	Fo, E	Neg	270, 242, 141, 92, 162
19	EG	15	F	Diarrhea	Esophagus: white plaques; Antrum: nodules	EG, EoE	EG	0.6	Iron, vitamins, probiotics	Elim	N	Y	Y	N	N	N	Y	Y	Y	Fo, E	Neg	664, 449, 275, 308, 242
20	EG	12	F	Medication change	Esophagus: white patches; Stomach: peristomal nodules	EG, EoE	EG, EoE, EJ	5.3	SWGC, 6 MP, growth hormone, lamisil	Ad lib	U	N	N	N	N	U	U	Y	Y	Neg	Neg	183, 171, 107, 85, 109
21	EG	14	M	Vomiting, dysphagia	Esophagus: thickened and furrowed; body/antrum: punctate erythematous areas with suggestion of ulceration	EoE, EG	EoE, EG	2.1	SWGC, INHGC, H1RA, LTRA	Elim	Y	Y	Y	Y	Y	Y	N	Y	Y	Fo, E	Neg	258, 195, 268, 134, 163
22	EG	12	M	Abdominal discomfort, nausea, poor growth	Esophagus: diffusely furrowed, thickened and friable; antrum/body: small nodules; duodenum: nodularity	EG	EoE, EG	0.8	PPI, LTRA	Ad lib	N	N	Y	N	N	Y	Y	Y	N	Neg	Neg	155, 176, 92, 125, 95
23	EG	11	M	Vomiting	Esophagus: mild furrows; antrum: nodularity, erythema, friability	EoE, EG	EoE, EG	4.0	PPI	Elim	N	Y	Y	N	N	Y	N	Y	Y	Neg	Neg	183, 263, 201, 236, 201
24	EG	30	M	U	U	EoE, EG	EoE	None	PPI, SWGC	Elemental + food	N	Y	Y	Y	Y	N	N	Y	Y	Neg	U	91, 116, 59, 172, 87
25	EG	6	M	Surveillance	Prepyloric stomach: 3-4 small sessile polyps; duodenum: studded with small sessile polyps, some pedunculated, some lobulated, some ulcerated	EG	EG, ED	1.5	None	Ad lib	N	N	N	N	N	N	N	N	N	Neg	Not done	68, 50, 56, 49, 64
26	EG	11	F	Hx EoE/EG	Antrum, fundus: 2-3 polyps	EoE, EG	EoE, EG	0.9	PPI, SWGC, LTRA, B	Elim	Y	N	Y	Y	Y	N	Y	Y	Y	Neg	Not done	162, 179, 122, 101, 58
27	EG	23	M	Hx EoE/EG; surveillance following wheat elimination	Antrum: nodularity, erythema	EG	EoE, EG, ED	10	SWGC, H1RA	Elim	Y	Y	Y	Y	Y	Y	Y	Y	Y	Fo, E	Not done	231, 93, 115, 342, 163

28	EG	14	M	Hx of EG, abdominal pain, edema	Antrum: marked nodularity and erythema; distal duodenum and proximal jejunum: multiple superficial areas of erythema	EG, ED, EJ	EG	0.6	PPI, NGC, H1RA	Ad lib	N	Y	Y	Y	N	N	N	Y	Y	Fo, E	Neg	280, 264, 316, 360, 261
29	EG	7	F	Hx of EG, possible eosinophilic enteropathy	Antrum: erythema, large prominent lymphonodules; esophagus: distal erythema	EG	EG	4.0	INHGC, OGC, B	Elim	Y	N	Y	Y	Y	Y	N	Y	Y	NT	Neg	154, 146, 161, 162, 168
30	EG	12	F	EoE, food allergies	Proximal esophagus: edema, exudate; mid esophagus: exudate; fundus of stomach: loss of vascular pattern, nodularity	EG	EoE	0.8	SWGC, NGC, H1RA, PPI	Elim	N	Y	Y	Y	Y	Y	N	Y	Y	Fo, E	Not done	329, 305, 327, 445, 149

^AGastric diagnosis at index endoscopy

^BAge at time of index endoscopy

^CIndication for index endoscopy

^DMacroscopic findings observed during the index endoscopy

^EDiagnosis derived from the index endoscopy

^FAny EGID diagnosis that was assigned prior to the index endoscopy

^GTime elapsed between the initial diagnosis of EG and the index endoscopy

^HMedications that the patient was prescribed during the time period immediately prior to the index endoscopy

^IDiet that the patient was prescribed during the time period immediately prior to the index endoscopy

^JIf known, it is noted whether a patient has a history of the indicated condition. This information is self-reported by the patient or his/her parent and verified when possible by chart review. Unknown indicates that the medical record was reviewed but it could not be determined whether the patient had the indicated condition. Not Tested indicates that the medical record specified that the patient had never undergone testing for the indicated condition.

^KResults of skin prick tests performed on the date closest to the index endoscopy, when available. Unknown indicates that the medical record was reviewed but it could not be determined whether the patient had ever undergone skin prick testing. Not Tested indicates that the medical record specified that the patient had never undergone skin prick testing.

^LResults of the campylobacter-like organism test/rapid urease test for the index endoscopy are shown. Unknown indicates that the medical record was reviewed but it could not be determined whether the patient had the Clo test at the index endoscopy. Not done indicates that the Clo test was not performed at the index endoscopy.

^MThe areas of the gastric biopsy with the most dense eosinophil infiltrate were identified, and the number of eosinophils in five separate 400X high-power fields (HPF) in these areas were counted. The counts from the 5 HPFs are shown.

Abbreviations: 6 MP, 6 mercaptopurine; AC, allergic conjunctivitis; Ad lib, ad libitum; AR, allergic rhinitis; B, beta-agonist; Clo, campylobacter-like organism test/rapid urease test; CTL, control; Dx, diagnosis; EG, eosinophilic gastritis; EGID, eosinophilic gastrointestinal disorder; EJ, eosinophilic jejunitis; Elim, elimination diet; E, environmental; EoE, eosinophilic esophagitis; F, female; Fo, food; FTT, failure to thrive; H1RA, H1 receptor antagonist; HF, hay fever; Hx, history; IE, index endoscopy; INHGC, inhaled glucocorticoid; LTRA, leukotriene receptor antagonist; M, male; Med, medication; mo, months; MTX, methotrexate; N, No; Neg, negative; NDA, no diagnostic abnormality; NGC, nasal glucocorticoid; NT, not tested; OGC, oral glucocorticoid; PPI, proton pump inhibitor; SWGC, swallowed glucocorticoid; U, unknown; U/A, urticaria/angioedema; Y, yes.

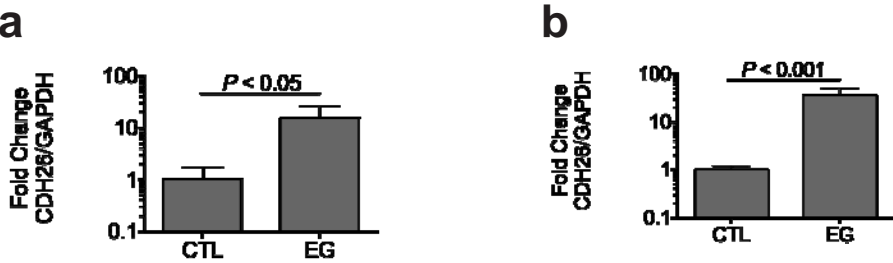
Supplementary Table S3. Primers used for RT-PCR

Transcript	Forward primer (5' to 3')	Reverse primer (5' to 3')	Reference
<i>GAPDH</i>	TGGAAATCCCATCACCATCT	GTCTTCTGGGTGGCAGTGAT	¹⁹
<i>CDH26</i>	TGCTTTTCTGTTGCGATGCT	CTTGCCATAACCCCAGCTC	This Study
<i>IL13</i>	ACAGCCCTCAGGGAGCTCAT	TCAGGTTGATGCTCCATACCAT	²⁰
<i>C3</i>	GCTGAAGCACCTCATTGTGA	CTGGGTGTACCCCTTCTTGA	²¹
<i>CCL26</i>	AACTCCGAAACAATTGTGACTCAGCTG	GTA ACTCTGGGAGGAAACACCCTCTCC	²⁰

Supplementary Table S4. Primers used to generate expression constructs

Primer	Primer Sequence (5' to 3')
Designation	
4105	GGGGTACCTTAGGAAGGAACACCTGACT
4241	GGGGTACCTTACAGGTCCTCCTCGCTGATCAGCTTCTGCTCGGAAGGAACACCTGACT
4242	GGGGTACCTTAGGCGTAGTCGGGCACGTCGTAGGGGTAGGAAGGAACACCTGACT
4366	GCTCTAGACACCATGGCTACTCAAGCTGATTTG
4367	GGGGTACCTTACAGGTCAGTATCAAACC
4370	GGAATTCACCATGACTGCTGTCCATGCAGG
4468	GGAATTCACCATGGACGACTCAGAGGTGG
4469	GGGGTACCCTAAATCTTCTGCATGGAGG
4590	GCTCTAGACACCATGTACCCCTACGACGTGCCCCGACTACGCCGCTACTCAAGCTGATTTG
4593	GGGGTACCTTAGGCGTAGTCGGGCACGTCGTAGGGGTACAGGTCAGTATCAAACC
4701	GGGGTACCTTAGGCGTAGTCGGGCACGTCGTAGGGGTAGATGCTGTCCATAGCTTTG
5067	GTGGGCATGTGTGAGTTTGTCTCTGCAAGCTCCACACATGT
5108	GGAATTCGAGAGGCTCGTCTTCTGTGAG
5109	GGGGTACCTGACCTGGTTCTTGGTCATC
5140	CCCAAGCTTCAACGTGATTTCTGCATTTC
5170	GGAATTCACCATGGCCATGAGATCCGGGAGG
5254	ATAAGAATGCGGCCGCACCATGGCCATGAGATCCGGGAGG
7491	ATAAGAATGCGGCCGCTCATTTACCCGGAGACAGGGAGA
5427	ATAAGAATGCGGCCGCCACCATGGGCCCTTGGAGC
6074	CCGCTCGAGCACCATGGCTTGGGAAGCGAGG
6075	ATAAGAATGCGGCCGCTTAATCATCATTGCTTTTACTG
6078	ATAAGAATGCGGCCGCCACCATGGTGGCTTTGCCAATGGT
6079	GGAATTCTCAGAGAGTGGGACTGTCTG

Supplementary Figure S1



Supplementary Figure S2

a

CDH26	-----
CDH2	---MCRIAGALRTLLPLLAALLQASVEASGEIALCKTGFPEDVYSAVLSKDVH-EGQPLL
CDH4	---MTAGAGVLL-LLLSLSGALRAHNEDLTTRETCKAGFSEDDYTALISQNIL-EGEKLL
CDH18	MKITSTSCICPVLVC-LCF-----VQRCYGTAAHSSIK-----
CDH6	-----MRTYRYFLLLFVVGQPYPTLSTPLSKRTSGFPAK-----
CDH12	-----MLTRNCLSLLLWVLFDGLLT-PLQPQPQQLAT-----
CDH11	MKENYCLQA-----ALVCLGMLCHSHAFA----PERRGHLRP-----
CDH20	MWTSGRMSNAKNWLG-LGMSLYFWGLMDLTTTVLSDTPTPQG-----
CDH9	-----MRTYHYIPLFIWTYMFHTVDTILLQEKPN SYLSS-----
CDH5	-----MQRLMMLLATSGACLGLLAVAAVAAAGANP-----
CDH8	MPERLAEMLLDLWTPLIILWITLPPCIYMAPMNQSQVLMMSG-----
CDH13	MQPR-----TPLVLCVLLSQVLLLTSAEDLDCTPGFQQKVFHINQPAEFI-EDQSIL
CDH7	-----M-KLGKVEFCHFLQLIALFLCFSGMSQA-----
CDH1	MGPWSRSLSALLLLLQVSSWL-----CQEPEPCHPGFDAESYTFVPRRHLEGRVLG
CDH24	MWG-----LVRLLLAWLGGWGC MGR LA-----A-----
CDH3	-MGLPRGPLASLLLLQVCWLQ-----CAASEPCRAVFREAEVTLEAGGAEQEPGQALG
CDH15	-----MDAAFLLVLG
CDH10	-----MTIHQFLLLFLFWVCLPHFCSPEIMFRRTVPVQQR-----
CDH22	MRPRPEGRGLRAGVALSPALLLLLLLPPPTLLGRLWAAGTP-----
CDH26	-----
CDH2	NVKFSNCGKRKVQYESSEPADFKVDEDMVYAVRSFPLSSEHAKFLIYAQDKETQEKWQ
CDH4	QVKFSSCVGTKGTQYETNS-MDFKVGADGTVFATRELQVPSEQVAFTVTAWDSQTAEKWD
CDH18	-----
CDH6	-----
CDH12	-----
CDH11	-----
CDH20	-----
CDH9	-----
CDH5	-----
CDH8	-----
CDH13	NLTFSDCKGNDKLRYEVSS-PYFKVNSDGGLVALRNITAVG--KTLFVHARTPHAEDMAE
CDH7	-----
CDH1	RVNFEDCTGRQRTAYFSLD-TRFKVGTGCVITVKRPLRFHNPQIHFLVYAWDSTYRKF--
CDH24	-----
CDH3	KV-FMGCPGQEPALFSTDN-DDFTVRNGETVQERRSLKERNPLKIFPS-----
CDH15	LLAQSLCCLSLGVPGWRRPT-----
CDH10	-----
CDH22	-----

Supplementary Figure S2

a

		(1)
		(3) ↓
CDH26	-MAMRSGRHPSLLLLLVLLLWLLQVSIIDSVQQETDDLTKQT--KEKIYQPLRRSKRRWV	
CDH2	VAVKLSLKPTLTTEESVKESAEVEEI-----VFPRQ--FSKHSGLHQRQKRDWV	
CDH4	AVVRLLLVAQTSSPHSGHKPQKGKKVVALDPSPPPKDTLLPWP--QHQNANGLRRRKRDWV	
CDH18	-----VMRNQTKHIEGETEVHHRPKRGWV	
CDH6	-----KRALEL--SGNSKNELNRSKRSWM	
CDH12	-----EPRENVIHLPGQRSHFQRVKRGWV	
CDH11	-----SFHGHHE-KGKEGQVLQRSKRGWV	
CDH20	-----ELEALLS-D--KPQSHQRTKRSWV	
CDH9	-----KKIAGL--TKDDGKMLRRTKRGWM	
CDH5	-----AQRD--THSLLPTHRRQKRDWI	
CDH8	-----P-LELNS-LGEEQRILNRSKRGWV	
CDH13	LV-----IVGGKDIQGSGLQDIFKFAR-----TSPVPRQKRSIV	
CDH7	-----ELSRERS-KPYFQSGRSRTKRSWV	
CDH1	-STKVTLNTVGHHRPPPHQASVSGIQAELLT-----FPNSSPGLRRQKRDWV	
CDH24	-----PARAWAGSREHPGPALLRTRRSWV	
CDH3	-----KRILRRHKRDWV	
CDH15	-----TLYPWR----RAPALSRVRRWV	
CDH10	-----ILSSRV--PRSDGKILHRQKRGWM	
CDH22	-----SPSAPGA-RQDGALGAGRVKRGWV	
		* : *
	(2,3) (2) (3)	
CDH26	<u>ITTTLELEEDPGPFKLI</u> GELFNNMSY-NMSLMYLISGPGVDEYPEIGLESLEDHENGRI	
CDH2	IPPINLPENSRGPFQQLVRIIRSDRDKNLSLRYSVTGPGADQPPT-GIFIIN-PISGQL	
CDH4	IPPINVPENSRGPFQQLVRIIRSDKDN-DIPIRYSITGVGADQPPM-EVFSID-SMSGRM	
CDH18	WNQFFVLEEHMGPDQPQYVGKLHSNSDKGDGSKYIILTGEAG-----TIFIID-DTTGDI	
CDH6	WNQFFVLEETGSDYQYVGKLHSDQDRGDGSLKYILSGDGAG-----DLFIIN-ENTGDI	
CDH12	WNQFFVLEEYVGSEPPQYVGKLHSDLDKGEETVKYITLSGDGAG-----TVFTID-ETTGDI	
CDH11	WNQFFVIEEYTGPDVPLVGRHSDIDSGDGNIKYILSGEGAG-----TIFVID-DKSGNI	
CDH20	WNQFFVLEEYTGTDPLYVGKLHSDMDRGDSIKYILSGEGAG-----IVFTID-DTTGDI	
CDH9	WNQFFVLEEYTGTDQYVGKLHSDQDKGDGNLKYIILTGDGAG-----SLFVID-ENTGDI	
CDH5	WNQMHIDEENKNTSLPHHVGIKSSVSRK--NAKYLLKGEYVG-----KVFRVD-AETGDV	
CDH8	WNQMFVLEEFSGPEPILVGRHSDLDPGSKIKYILSGDGAG-----TIFQIN-DVTGDI	
CDH13	VSPILIPENQRQPFPRDVGKVV-DSDR-PERSKFRLTGKGVDPQEPK-GIFRIN-ENTGSV	
CDH7	WNQFFVLEEYMGSDPLYVGKLHSDVDKGDGSIKYILSGEGAS-----SIFIID-ENTGDI	
CDH1	IPPISCPENEKGPFKPNLVQIKSNKDK-EGKVFYSITGQADTPPV-GVFIIE-RETGWL	
CDH24	WNQFFVIEEYAGPEPVLIGKLHSDVDRGEGRTKYLLTGEAG-----TVFVID-EATGNI	
CDH3	VAPISVPENGKGPFPQRLNQLKSNKDR-DTKIFYISITGPGADSPPE-GVFAVE-KETGWL	
CDH15	IPPISVSENHNR-LPYPLVQIKSDKQQ-LGSVIYSIQPGVDEEPR-GVFSID-KFTGKV	
CDH10	WNQFFVLEEYTGSDYQYVGKLHSDQDKGDGSLKYILSGDGAG-----TLFIID-EKTGDI	
CDH22	WNQFFVVEEYTGTEPLYVGKIHSDEGDGAIKYTISGEGAG-----TIFLID-ELTGDI	
	: * : : . : . : : * . : * : . * :	

Supplementary Figure S2



Supplementary Figure S2

a

CDH26	CLDYETA--PQFTLLIRARDCGEP--SLSSTTTVHVDVQEGNNHRPAFTQENYKVQIPE
CDH2	-LDREKV--QQYTLIIQATDMEGNPTYGLSNTATAVITVTDVNDNPPEFTAMTFYGEVPE
CDH4	-LDREKV--QQYTVIVQATDMEGNLNYGLSNTATAIITVTDVNDNPPEFTASTFAGEVPE
CDH18	-MDREAR--EHYSVVIQAKDMAGQV--GGLSGSTTVNITLTDVNDNPPRFPQKHQYQLYVPE
CDH6	-MDREN--EQYQVVIQAKDMGGQM--GGLSGSTTVNITLTDVNDNPPRFPQSTYQFKTPE
CDH12	-MDREVK--EQYQVLIQAKDMGGQL--GGLAGTTIVNITLTDVNDNPPRFPKSIHFLKVPE
CDH11	-MDREAK--EEYHVVIQAKDMGGHM--GGLSGTTKVTITLTDVNDNPPKFPQSVYQMSVSE
CDH20	-MDREAK--EYVEVIIQAKDMGGQL--GGLAGTTTVNITLSDVNDNPPRFPQKHQYQMSVLE
CDH9	-MSREN--EQYQVVIQAKDMGGQM--GGLSGSTTVNITLTDVNNNPPRFPQSTYQFNSPE
CDH5	-LDREKQ--ARYEIVVEARDAQGLR--GDSGTATVVLVTLQDINDNFPFFTQTKYTFVVPPE
CDH8	-MDREAK--EEYLVVIQAKDMGGHS--GGLSGTTTLTVTLTDVNDNPPKFAQSLYHFSVPE
CDH13	LLDRETLENPKYELII EAQDMAGLD-VGLTGTATATIMIDDKNDHSPKFTKKEFQATVEE
CDH7	-MDREAK--DQYLLVIQAKDMVGQN--GGLSGTTSVTVTLTDVNDNPPRFPRRSYQYNVPE
CDH1	-LDRESF--PTYTLVVQAADLQGE--GLSTTATAVITVTDNDNPPIFNPPTYKQGVPE
CDH24	-MDRETQ--EEFLVVIQAKDMGGHM--GGLSGSTTVTVTLSDVNDNPPKFPQSLYQFSVVE
CDH3	-LDREKV--PEYTLTIQATDMGD--GSTTTAVAVVEILDANDNAPMFDPPQKYEAHVPE
CDH15	-LDREVV--AVYNLTQVADMSGD--GLTATASAIITLDDINDNAPEFTRDEFFMEIAIE
CDH10	-MNREN--EQYQVVIQAKDMGGQM--GGLSGSTTVNITLTDVNDNPPRFPQNTIHLRVLE
CDH22	-LDRESQ--ERYEVVIQATDMAGQL--GGLSGSTTVTIVTDVNDNPPRFPQKMYQFSIQE
	: . * : : : . . * . : : : : : * : . * * *
CDH26	GR-ASQGVLRLLVQDRDSPFTSAWRAKFNILH--GNEEGHFDISTDPETNEGILNVIKPL
CDH2	NR-VDIIIVANLTVTDKQPHTPAWNNAVYRISG--GDPTGRFAIQTDPNNSNDGLVTVVKPI
CDH4	NR-VETVVANLTVMDRDQPHSPNWNNAVYRIIS--GDPSGHFSVRTDPVTNEGMMVTVVKAV
CDH18	SAQVGSVAVGIKANDADTG--SNADMTYSIIN--GDGMGIFSISTDKETREGILSLKKPL
CDH6	SSPPGTPIGRIKASDADVG--ENAEIEYSITD--GEGLDMFVDVITDQETQEGIIITVKKLL
CDH12	SSPIGSAIGRIRAVDPDFG--QNAEIEYNIVP--GDGGNLFDIVTDEDTQEGVIKLLKKPL
CDH11	AAVPGEEVGRVKAKDPDIG--ENGLVTYNIVD--GDGMESFEITTDYETQEGVIKLLKKPV
CDH20	SAPISSTVGRVFAKDLDEG--INAEMKYTIVD--GDGADAFDISTDPNFQVGIITVKKPL
CDH9	SVPLGTHLGRIKANDPDVG--ENAEMEYSIAE--GDGADMFDVITDKDTQEGIIITVKQNL
CDH5	DTRVGTSGSLFVEDPDEP--QNRMTKYSILR--GDYQDAFTIETNPAHNEGIIKPMKPL
CDH8	DVVLGTAIGRVKANDQDIG--ENASQSYDIID--GDGTALFEITSDAQADGIIRLRKPL
CDH13	GA-VG-VIVNLTVEDKDDPTTGAWRAAYTIIN--GNPGQSFEIHTNPQTNEGMLSVVKPL
CDH7	SLPVASVVARIKAADADIG--ANAEMEYKIVD--GDGLGIFKISVDKETQEGIIITIKEL
CDH1	NE-ANVVITTLKVTADAPNTPAWEAVYTIIN--DDGGQFVVTTNPVNNDGILKTAKGL
CDH24	TAGPGTLVGRRLRAQDPDLG--DNALMAYSILD--GEGSEAFSISTDLQGRDGLLTVRKPL
CDH3	NA-VGHEVQRLTVTDLDAPNSPAWRATYILIMG--GDDGDHFTITTHPESNQGILTTRKGL
CDH15	AV-SGVDVGRLEVEDRDLPGSPNWWARFTILE--GDPDGQFTIRTDPKTNEGVLSIVKAL
CDH10	SSPVGTAIGSVKATDADTG--KNAEVEYRIID--GDGDMFDIVTEKDTQEGIIITVKKPL
CDH22	SAPIGTAVGRVKAEDSDVG--ENTDMTYHLKDESSSGGDVFKVTTSDTQEAIIVVQKRL
	: : . * * : : . * : . . : : : :

Supplementary Figure S2

a

CDH26	DYETRPAQSLIIVVENEERLVFCERGLQPPRKAASATVSVQVTDANDPPAFHPQSFIIV
CDH2	DFETNRMFVLTVAAENQVPLAK-----GIQHPPQSTATVSVTVIDVNENPYFAPNPKII
CDH4	DYELNRAFMLTVMVSNQAPLAS-----GIQMSFQSTAGVTISIMDINEAPYFPSNHKLI
CDH18	NYEKKSSTLNLIEGANTHLDFR-----FSLGLGPFKDATMLKIIIGDVDEPPLFSMPSYLM
CDH6	DFEKKKVYTLKVEASNYPVEPR-----FLYLGPFKDSATVRIIVEDVDEPPVFSKLAYIL
CDH12	DFETKKAYTFKVEASNLHLDHR-----FHSAGPFKDTATVKISVLDVDEPPVFSKPLYTM
CDH11	DFETKRAYSLKVEAANVHIDPK-----FISNGPFKDTVTVKISVEDADEPPMFLAPSYIH
CDH20	SFESKKSSTLNLKVEGANPHLEMR-----FLNLGPFQDTTTHISVEDVDEPPVFEPGFYFV
CDH9	DFENQMLYTLRVDASNTHDPDR-----FLHLGPFKDTAVVKISVEDIDEPPVFTKVSYLI
CDH5	DYEYIQQYSFIVEATDPTIDLR-----YMS-PPAGNRAQVIINITDVDEPPIFQQPFYHF
CDH8	DFETKKSSTLNLKVEAANVHIDPR-----FSGRGPFKDTATVKIIVVEDADEPPVFSSPTYLL
CDH13	DYEISAFHTLLIKVENEDPLVP-----DVSYPGSSSTATVHITVLDVNEGPFVFPDPMV
CDH7	DFEAKTSYTLRIEAANKDADPR-----FLSLGPFSDTTTVKIIIVEDVDEPPVFSSPLYPM
CDH1	DFEAKQQYILHVAVTNVVPFEV-----SLTTSTATVTVLDVLDVNEAPIFVPPEKRV
CDH24	DFESQRSYSFRVEATNTLIDPA-----YLRRGPFKDVASVRVAVQDAPEPPAFTQAAYHL
CDH3	DFEAKNQHTLYVEVTNEAPFVL-----KLPTSTATIVVHVEDVNEAPVFVPPSKVV
CDH15	DYESCEHYELKVSQNEAPLQA-----AALRAERGQAKVRVHVQDTNEPPVFQENPLRT
CDH10	DYESRRLYTLKVEAENTHVDPR-----FYYLGPFKDTTIVKISIEDVDEPPVFSRSSYLF
CDH22	DFESQPVHTVILEALNKFVDPR-----FADLGTFRDQAIVRVAVTDVDEPPEFRPPSGLL
	. : * . : :
CDH26	NKEEGARPGTLLGTFNAMDPDS----QIRYELVH--DPANWVSVDKNSGVVITVEPIDRE
CDH2	RQEEGLHAGTMLTTFTAQDPDRYMQQNIRYTKLS--DPANWLKIDPVNGQITTIIVLDRE
CDH4	RLEEGVPPGTVLTTFSAVDPDRFMQQAVRYSKLS--DPASWLHINATNGQITTAIVLDRE
CDH18	EVYENAKIGTVVGTVLAQDPDS--TNSLVRYFINYNVEDDRFFNIDANTGTIRTTKVLDR
CDH6	QIREDAQINTTIGSVTAQDPDA--ARNPVKYSVDRHTDMDRIFNIDSGNGSIFTSKLLDRE
CDH12	EVYEDTPVGTIIGAVTAQDLDV--GSSAVRYFIDWKS DGDSYFTIDGNEG TIATNELLDR
CDH11	EVQENAAAGTVVGRVHAKDPDA--ANSPIRYSIDRHTDLDRFFTINPEDGFIKTTKPLDR
CDH20	EVPEDVAIGTTIQIISAKDPDV--TNNSIRYSIDRSSDPGRFFYVDITTGALMTARPLDR
CDH9	EVDEDVKEGSIIGQVTAYDPDA--RNNLIKYSVDRHTDMDRIFGIHSENGSIFTLKDRE
CDH5	QLKEN--QKKPLIGTVLAMDPDA--ARHSIGYSIRRTSDKGQFFRVT--KKGDIYNEKELDR
CDH8	EVHENAALNSVIGQVTARDPDI--TSSPIRFSIDRHTDLERQFNINADDGKITLATPLDR
CDH13	TRQEDLSVGSVLLTVNATDPDSLQHQTI RYSVYK--DPAGWLNINPINGTVDTTAVLDRE
CDH7	EVSEATQVGNIIGTVAAHDPDS--SNSPVRYSIDRNTDLERYFNIDANSGVITTAKSLDR
CDH1	EVSEDFGVGQEITSYTAQEPDTFMEQKITRYIRIWR--DTANWLEINPDTGAISTRAELDR
CDH24	TVPENKAPGTLVGQISAADLDS--PASPIRYSILPHSDPERCFSIQPEEGTIHTAAPLDR
CDH3	EVQEGIPTGEPVCVYTAEDPDK--ENQKISYRILR--DPAGWLAMDPDSGQVTAVGTLDRE
CDH15	SLAEGAPPGLTVATFSARDPDTEQLQRLSYSKDY--DPEDWLQVDAATGRIQTQHVLSPA
CDH10	EVHEDIEVGTTIIGTVMARDPDS--ISSPIRFSIDRHTDLDRIFNIHSGNGSLYTSKPLDR
CDH22	EVQEDAQVGSVLGVVTARDPDA--ANRPVRYAIDRESDLQIFDIDADTGAIVTGKGLDR
	* : * : * : : . : * : :

Supplementary Figure S2

a

(2)

CDH26	SP-HVNNSFYVIIIIHAVDDG-----FP
CDH2	SP-NVKNNIYNATFLASDNG-----IP
CDH4	SL-YTKNNVYEATFLAADNG-----IP
CDH18	ET-----PWYNITVTASEID-----NP
CDH6	TL-----LWHNITVIATEIN-----NP
CDH12	ST-----AQYNFSIIASKVS-----NP
CDH11	ET-----AWLNITVFAAEIH-----NR
CDH20	EF-----SWHNITVLAMEMN-----NP
CDH9	SS-----PWHNITVTATEIN-----NP
CDH5	VY-----PWYNLTVEAKELD-----STGTPT
CDH8	LS-----VWHNITIIATEIR-----NH
CDH13	SP-FVDNSVYTALFLAIDSG-----NP
CDH7	TN-----AIHNITVLAMESQ-----NP
CDH1	DFEHVKNSTYTALIIATDNG-----SP
CDH24	AR-----AWHNLTVLATELGWSWGPGRGWVPLLVAEWSAPAAPPQRSPVGSAGVGPQDSS
CDH3	DEQFVRNNIYEVVMLAMDNG-----SP
CDH15	SP-FLKGGWYRAIVLAQDDA-----SQ
CDH10	LS-----QWHNLTVIAAEIN-----NP
CDH22	TA-----GWHNITVLAMEAD-----NH
	. * .
CDH26	PQTATGTLMLFLSDINDNVPTLRPRSRMEVCESA-----VHEPLHIEAEDPDLEPFSD
CDH2	PMSGTGTQLQIYLLDINDNAPQVLPQE--AETCETP-----DPNSINITALDYDIDPNAG
CDH4	PASGTGTQLQIYLLIDINDNAPELLPKE--AQICEKP-----NLNAINITAADADVDPNIG
CDH18	DLLSHVTVGIRVLDVNDNPPELAREYD-IIVCENSK----PGQVIHTISATDKDDFANGP
CDH6	KQSSRVPLYIKVLDVNDNAPEFAEFYE-TFVCEKAK----ADQLIQTILHAVDKDDPYSGH
CDH12	LLTSKVNILINVLDVNEFPPEISVPYE-TAVCENAK----PGQIIQIVSAADRDLSPAGQ
CDH11	HQEAKVPVAIRVLDVNDNAPKFAAPYE-GFICESDQTKPLSNQPIVTISADDKDDTANGP
CDH20	SQVGSVPVTIKVLDVNDNAPEFPRFYE-AFVCENAK----AGQLIQTVSAVDQDDPRNGQ
CDH9	KQSSHIPVFIIRILDINDHAPEFAMYIE-TFVCENAK----PGQLIQTVSVMDDKDDPPRGH
CDH5	GKESIVQVHIEVLDENDNAPEFAKPYQ-PKVCENAV----HGQLVLQISAIKDKITPRNV
CDH8	SQISRPVPAIKVLDVNDNAPEFASEYE-AFLCENGK----PGQVIQTVSAMDKDDPKNGH
CDH13	PATGTGTLLITLEDVNDNAPFIYPTV--AEVCDDAK----NLSVVILGASDKDLHPNTD
CDH7	SQVGRGYVAITILDINDNAPEFAMDYE-TTVCENAQ----PGQVIQKISAVDKDEPSNGH
CDH1	VATGTGTLLLLILSDVNDNAPIPEPT--IFFCERN-----PKPQVINIIDADLPNTS
CDH24	AQASRVQVAIQTLDENDNAPQLAEPYD-TFVCDSAA----PGQLIQVIRALDRDEVGNSS
CDH3	PTTGTGTLLLLTLIDVNDHGPVPEPRQ--ITICNQS-----PVRQVLNITDKDLSPHTS
CDH15	PRTATGTLSIEILEVNDHAPVLAPPPP-GSLCSEPH-----QGPGLLLGATDEDLPPHGA
CDH10	KETTRVAVFVRILDVNDNAPQFAVFYD-TFVCENAR----PGQLIQTISAVDKDDPLGGQ
CDH22	AQLSRASLRIRILDVNDNPPELATPYE-AAVCEDAK----PGQLIQTISVVDREDEPQGGH
	: : : * : * * . : * *

Supplementary Figure S2

a

CDH26	PFTFELDNTWGN AEDTWKLG RNWQSV ELLTL--RSL--PRGNYLVPLFIGDKQG--LSQ
CDH2	PFAFDLPLSPVTIKRNWTITRLNGDFAQLNLK-IKFL--EAGIYEVPIIITDSGNPPKSN
CDH4	PYVFELPFVPAAVRKNWTITRLNGDYAQLSLR-ILYL--EAGMYDVPIIIVTDSGNPPLSN
CDH18	RFNFFLDERLP-VNPNFTLKDNDNTASILTRRRRFSRTVQDVYYLPIMISDGGIPSLSS
CDH6	QFSFSLAPEAA-SGSNFTIQDNKDNTAGILTRKNGYNRHEMSTYLLPVVISDNDYPVQSS
CDH12	QFSFRLSPEAA-IKPNFTVRDFRNNNTAGIETRRNGYSRRQQELYFLPVVIEDSSYPVQSS
CDH11	RFIFSLPPEII-HNPNFTVRDNRDNTAGVYARRGGFSRQKQDLYLLPIVISDGGIPPMSS
CDH20	HFYYS LAPEAA-NNPNFTIRDNDNTARILTRRSGFRQQEQSVFHLPIILIADSGQPVLS
CDH9	KFFFEPVPEFT-LNPNFTIVDNKDNTAGIMTRKDGYSRNKMSTYLLPILIFDNDYPIQSS
CDH5	KFKFILNT----ENNFTLTDNDNTANITVKYGQFDREHTKVHFLPVVISDNGMPSRTG
CDH8	YFLYSLLPEMV-NNPNFTIKKNEDNSLSILAKHNGFN RQKQEVYLLPIIISDSGNPPLSS
CDH13	PFKFEIHKQAV-PDKVWKISKINNTHALVSL--QNL--NKANYNLPIMVTD SGKPMTN
CDH7	QFYFSLTTDAT-NNHNFS LKDNKDNTASILTRRNGFRQQEQSVYYLP I FIVDSGSPSLSS
CDH1	PFTAELTHGA---SANWTIQYNDPTQESIILKPKMAL--EVGDYKINLKLMDNQN--KDQ
CDH24	HVSFQGPLG---PDANFTVQDNRDGSASLLLP SRPAPPRH-APYLVPIELWDWGQPALSS
CDH3	PFQAQLTDDS---DIYWTAEVNEEGDTVV-LSLKKFL--KQDTYDVHLSLSDHGN--KEQ
CDH15	PFHFQLSPRLPELGRNWSLSQVNVSHARLRPR--HQV--PEGLHRLSLLLRDSGQPQQR
CDH10	KFFFSLAA----VNPNTVQDNEDNTARILTRKNGFN RHEISTYLLPVVISDNDYPIQSS
CDH22	RFYFRLVPEAP-SNP HFSLLDIQDNTAAVHTQHVGFN RQEQDVFFLPILVVD SGPPPTLSS

: . : . : : *

(4)

CDH26	KQTVHVRICPCASGLT---CVELA-DAE---VGLHVGALFPVCAAFVALAVALFLLR CY
CDH2	ISILRVKVCQCDSNGD---CTD VDRIVG---AGLGTGAI IAILL---CIIILLILVLMF
CDH4	TSIIKVKVPCDDNGD---CTTIGAVAA---AGLGTGAIVAILI---CIIILLTMVLLF
CDH18	SSTLTIRVCACERDGRVRTCHAE AFLSS---AGLSTGALIAILL---CVLILLAIVVLF
CDH6	TGTVTVRVCACDHHGNMQSCHAEAL IHP---TGLSTGALVAILL---CIVILLVTVVLF
CDH12	TNTMTIRVCRCDSDGTILSCNVEAIFLP---VGLSTGALIAILL---CIVILLAIVVLY
CDH11	TNTLTIKVCGCDVNGALLSCNAEAYILN---AGLSTGALIAILA---CIVILLVIVVLF
CDH20	TGTLTIQVCSCDDDGHVMSCSPEAYMLP---VSLSRGALIAILA---CIFVLLVLVLLI
CDH9	TGTLTIRVCACDNQGNMQSCTAEALILS---AGLSTGALVAILL---CVLILLILVVLF
CDH5	TSTLTAVCKCNEQGEFTFCEDMAAQVG-----VSIQAVVAILL---CILTITVITLLI
CDH8	TSTLTIRVCGCSNDGVVQSCNVEAYVLP---IGLSMGALIAILA---CIILLLVIVVLF
CDH13	ITDLRVQVCSCRNSKV---DCNA-----AGALRFSLPSV---LL---LSLFSLACL---
CDH7	TNTLTIRVCDADGVAQTCNAEAYVLP---AGLSTGALIAILA---CVLTLLVLILLI
CDH1	VTTLEVSVCDCEGAAG---VCRKAQPVE---AGLQIPAILGILG---GILALLILILL
CDH24	TATVTVSVCRCPDGSVASCWPEAHL SA---AGLSTGALLAIIT---CVGALLALVVLF
CDH3	LTVIRATVCDCHGHE---TCP-----GPWKGGFILPVLG---AVLALLFLLLV
CDH15	EQPLNVTVCRCGKDG V---CLPGAAALLAGGTGLSLGALVIVLA---SALLLVLVLLV
CDH10	TGTLTIRVCACDSQGNMQSCSAEALLLP---AGLSTGALIAILL---CIIILLVIVVLF
CDH22	TGTLTIRICGCDSSGTIQSCNTTAFVMA---ASLSPGALIALLV---CVLILVLVLLI

: : * * :

Supplementary Figure S2

a

(5)

CDH26	<u>FVLEPKR-----HGCSV-SNDEGHQTLVMYNAESKGTSAQTWSDVEGQRPALLICTAA</u>
CDH2	VVWMKRDKERQAKQLLID-PEDDVRDNILKYDEEGGGEEDQD-YDLSQLQQPDTVEPDA
CDH4	VMWMKRREKERHTKQLLID-PEDDVRDNILKYDEEGGGEEDQD-YDLSQLQQPEAMGHVP
CDH18	ITLRRSK-----KEPLIIS-E-EDVRENVVTYDDEGGGEEDTEAFDITALRNPSA-----
CDH6	AALRRQR-----KKEPLII-SKEDIRDNIIVSYNDEGGGEEDTQAFDIGTLRNPEA-----
CDH12	VALRRQK-----KKDTLMT-SKEDIRDNIHYDDEGGGEEDTQAFDIGALRNPKV-----
CDH11	VTLRRQK-----KEPLIVF-EEEDVRENIITYDDEGGGEEDTEAFDIATLQNPDG-----
CDH20	LSMRRHR-----KQPYIID-DEENIHENIVRYDDEGGGEEDTEAFDIAAMWNPRE-----
CDH9	AALKRQR-----KKEPLII-SKDDVRDNIVTYNDEGGGEEDTQAFDIGTLRNPEA-----
CDH5	FLRRRLR-----KQARAHGKSVPEIHEQLVTYDEEGGEMDTTSYDVSVLNSVRR-----
CDH8	VTLRRHK-----NEPLIIK-DEEDVRENIIRYDDEGGGEEDTEAFDIATLQNPDG-----
CDH13	-----
CDH7	VTMRRRK-----KEPLIFD-EERDIRENIVRYDDEGGGEEDTEAFDMAALRNLNV-----
CDH1	LLFLRRR---AVVKEPLLP-PEDDTRDNVYYYDEEGGGEEDQD-FDLSQLHRGLD-----
CDH24	VALRRQK-----QEALMVL-EEEDVRENIITYDDEGGGEEDTEAFDITALQNPDG-----
CDH3	LLLVRKK---RKIKEPLLL-PEDDTRDNVFIYDEEGGGEEDQD-YDITQLHRGLE-----
CDH15	ALRARFW-KQSRGKGLLHG-PQDDLDRDNLNYDEQGGGEEDQDAYDISQLRHPTA-LSLP
CDH10	AALKRQR-----KKEPLIL-SKEDIRDNIIVSYNDEGGGEEDTQAFDIGTLRNPA-----
CDH22	LTLRRHH-----KSHLSSD-EDEDMRDNIKYNDEGGGEQDTEAYDMSALRSLYDFGELK

CDH26	<u>AGPTQGVKDLLEEVPPSAASQSAQARCALGSWGKPFEPKSVKNIHSTPAYPDATMHRQL</u>
CDH2	IKPVG-----IRRM-----DERP-IHAEPQYPVRS-----
CDH4	SKAPG-----VRRV-----DERP-VGAEPQYPIRPM-----
CDH18	-----AEELKYRR-----DIRPEVKLTTPRHQTSST-----
CDH6	-----IEDNKLRR-----DIVPEALFLPRRTP-TA-----
CDH12	-----IEENKIRR-----DIKPSLCLPRQRP-PM-----
CDH11	-----INGFIPRK-----DIKPEYQYMPRPGLRPA-----
CDH20	-----AQAGAAPKTRQ-----DMLPEIESLSRYVPQTC-----
CDH9	-----REDSKLRR-----DVMPETIFQIRRTV-PL-----
CDH5	-----GGAKPPRPALDARP-----SLYAQVQKPPRHAPGAH-----
CDH8	-----INGFLPRK-----DIKPDQLQFMPRQGLAPV-----
CDH13	-----
CDH7	-----IRDTKTRR-----DVTPEIQFLSRPAFKSI-----
CDH1	ARPEV-----TRN-----DVAPTLMSVPRYLPRPA-----
CDH24	-----AAPPAGPPARR-----DVLPRARVSRQPR--P-----
CDH3	ARPEV-----VLRN-----DVAPTIIPTPMYRPRPA-----
CDH15	LGPPP-----LRRD-----APQGRLLHPQP--PRVL-----
CDH10	-----IEEKKLRR-----DIIPETLFIPTPRTP-TA-----
CDH22	GGDGGGSAGGGAGGGSGGGAGSPPQA-----HLPSEHSLPQGPPSPE-----

Supplementary Figure S2

a

(6)

CDH26	LAPVEGRMAETLNQKLHVANVLEDDPGY-LPHVYSEEGECGGAPSLSSLASL---EQEL
CDH2	-APHPGDIGDFINEGLKAADNDPTAPPYDSL LV FDYEGSGSTAGSLSSLN SS ---SGGEQ
CDH4	-VPHPGDIGDFINEGLRAADNDPTAPPYDSL LV FDYEGSGSTAGSVSSLN SS ---SSGDQ
CDH18	--LESIDVQEFIKQRLAEADLDPSVPPYDSLQTYAYEGQRSEAGSISSLD SAT ---TQSDQ
CDH6	--RDNTDVRDFINQRLKENDTDPTAPPYDSLATYAYEGTGSVADSLSSLE SVT ---TDADQ
CDH12	--EDNTDIRDFIHQRLQENDVDPTAPPYDSLATYAYEGSGSVAESLSSID SLT ---TEADQ
CDH11	--PNSVDVDDFINTRIQEADNDPTAPPYDSIQIYGYEGRGSVAGSLSSLE SAT ---TDSDL
CDH20	--AVNSTVHSYVLAKLYEADMDLWAPPFDSLQTYMFEGDGSVAGSLSSLQ SAT ---SDSEQ
CDH9	--WENIDVQDFIHRRLKENDADPSAPPYDSLATYAYEGNDSIADSLSSLE SLT ---ADCNQ
CDH5	--GGPGEMAAMIEVKKDEADHDGDGPPYDTLHIYGYEGSE IAESLSS LGTD S ---SDSDV
CDH8	--PNGVDVDEFINVRLHEADNDPTAPPYDSIQIYGYEGRGSVAGSLSSLE STT ---SDSDQ
CDH13	-----
CDH7	--PDNVIFREFIWERLKEADVDPGAPPYDSLQTYAFEGNGSVAESLSSLD SIS ---SNSDQ
CDH1	---NPDEIGNFIDENLKAADTDPTAPPYDSL LV FDYEGSGSEASLSSLN SSE ---SDKDQ
CDH24	--PGPADVAQLLALRLREADEDPGVPPYDSVQVYGYEGRGSSCGSLSSLGSGSEAGGAPG
CDH3	---NPDEIGNFIIENLKAANTDPTAPPYDTLLVFDYEGSGSDAASLSSLT S SA---SDQDQ
CDH15	-PTSPLDIADFINDGLEAADS DP SVPPYDTALIYDYEGDGSVAGTLSSIL SSQ ---GDEDQ
CDH10	--PDNTDVRDFINERLKEHDLPTAPPYDSLATYAYEGNDSIAESLSSLE SGT ---TEGDQ
CDH22	--PDFSVFRDFISRKVALADGDLSPVPPYDAFQTYAFEGADSPAASLSSLHSGS---SGSEQ
CDH26	<u>QPDLLDSL</u> GSKATPFEEIYSESGVPS-----
CDH2	DYDYLNDWGPRFKKLADMYGGGDD-----
CDH4	DYDYLNDWGPRFKKLADMYGGGEED-----
CDH18	DYHYLGDWGPEFKKLAELYGEIESERTT-----
CDH6	DYDYLSDWGPRFKKLADMYGGVDS DKDS -----
CDH12	DYDYLTDWGPRFKVLADMFGEEESYNPDKVT-----
CDH11	DYDYLQNWGPRFKKLADLYGSKDTFDDDS-----
CDH20	SFDFLTDWGPRFRKLAELYGASEGPAPLW-----
CDH9	DYDYLSDWGPRFKKLADMYGGDDSDRD-----
CDH5	DYDFLNDWGPRFKMLAELYGSDPREELLY-----
CDH8	NFDYLSDWGPRFKRLGELYSVGESDKET-----
CDH13	-----
CDH7	NYDYLSDWGPRFKRLADMYGTGQESLYS-----
CDH1	DYDYLNEWGNRFKKLADMYGGGEDD-----
CDH24	PAEPLDDWGPLFRTLAELYGAKEPPAP-----
CDH3	DYDYLNEWGSRFKKLADMYGGGEDD-----
CDH15	DYDYLRDWGPRFARLADMYGHPCGLEYGARWDHQAREGLSPGALLPRHRGRTA
CDH10	NYDYLREWGPRFNKLAEMYGGGESDKDS-----
CDH22	DFAYLSSWGPRFRPLAALYAGHRGDDEAQAS-----

Supplementary Figure S2

b

CDH4	-VFDYEGSGSTAGSVSSSLNSSS--SG-DQDYDYL
CDH2	-VFDYEGSGSTAGSLSSSLNSSS--SGGEQDYDYL
CDH1	-VFDYEGSGSEAASLSSSLNSSE--SDKDQDYDYL
CDH3	-VFDYEGSGSDAASLSSSLTSSA--SDQDQDYDYL
CDH15	-IYDYEGDGSVAGTLSSILSSQ--GDEDQDYDYL
CDH11	QIYGYEGRGSVAGSLSSLESAT--TDSLDYDYL
CDH8	QIYGYEGRGSVAGSLSSLESTT--SDSDQNFDYL
CDH6	ATYAYEGTGSVADSLSSLESVT--TDADQDYDYL
CDH12	ATYAYEGSGSVAESLSSIDSLT--TEADQDYDYL
CDH10	ATYAYEGNDSIAESLSSLESST--TEGDQNYDYL
CDH9	ATYAYEGNDSIADSLSSLESST--ADCNQDYDYL
CDH18	QTYAYEGQRSEAGSISSSLDSAT--TQSDQDYHYL
CDH7	QTYAFEGTGSLAGSLSSLESAT--SDQDESYDYL
CDH19	QTYAFEGTGSLAGSLSSLESAT--SDQDESYDYL
CDH20	QTYMFEGDGSVAGSLSSSLQSAT--SDSEQSFDFL
CDH22	-TYAFEGADSPAASLSSSLHSGS--SGSEQDFAYL
CDH5	-IYGYEGSESIAESLSSSLGTDS--SDSDVDYDFL
CDH24	QVYGYEGRGSSCGSLSSSLGSGSEAGGAPGPAEPL
CDH26	-VYSEEGECGGAPSLSSSLASLE---QELQPDLL
	: ** . . : : * : *

Alignment of CDH26 and CDH1 only:

CDH26	VYSEEGECGGAPSLSSSLASLEQELQPDLL--
CDH1	VFDYEGSGSEAASLSSSLNSSES SDKDQDYDYL
	*:. **. . * ***** * *.: : *

Supplementary Figure S2

C

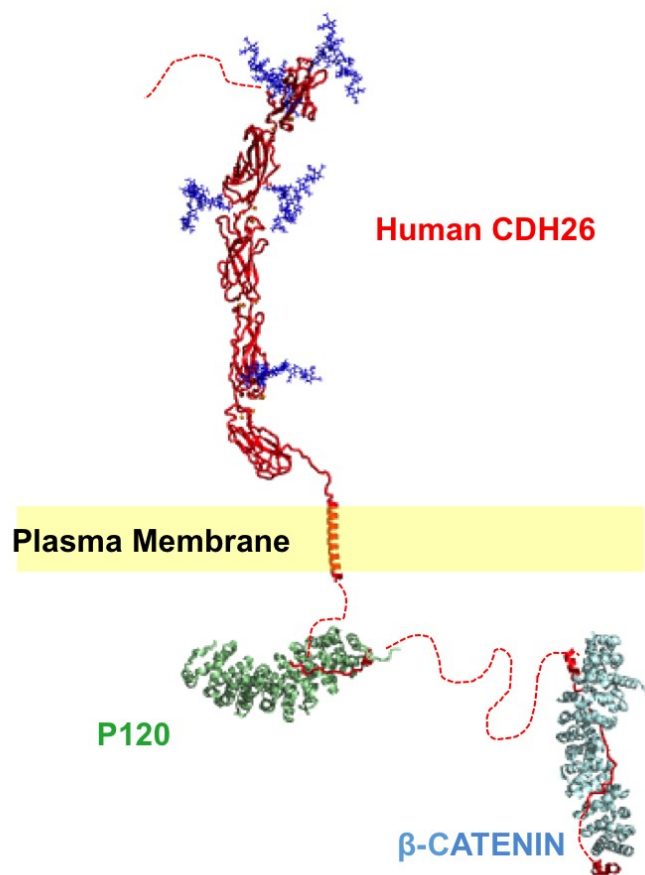
CDH6	R----
CDH10	R----
CDH9	R----
CDH12	K----KKDTLMTSKE--DIRDNVIHYDDEGGGEEDTQAFDIGALR-----N
CDH18	-----KKEPLIISEE--DVRENVVTYDDEGGGEEDTEAFDITALR-----N
CDH11	-----KKEPLIVFEE--EDVRENIITYDDEGGGEEDTEAFDIATLQ-----N
CDH8	-----KNEPLIIKDD--EDVRENIIRYDDEGGGEEDTEAFDIATLQ-----N
CDH24	-----KQEALMVLEE--EDVRENIITYDDEGGGEEDTEAFDITALQ-----N
CDH7	-----RKQILFPEKS--EDFRENIFQYDDEGGGEEDTEAFDIAELR-----S
CDH19	-----RKQILFPEKS--EDFRENIFQYDDEGGGEEDTEAFDIAELR-----S
CDH20	-----RKQPYIIDDE--ENIHENIVRYDDEGGGEEDTEAFDIAAMW-----N
CDH4	REKERHTKQLLIDPED--DVRDNILKYDEEGGGEEDQD-YDLSQLQ-----Q
CDH2	RDKERQAKQLLIDPED--DVRDNILKYDEEGGGEEDQD-YDLSQLQ-----Q
CDH15	W-KQSRGKGLLHGPQD--DLRDNVLNYDEQGGGEEDQDAYDISQLR-----H
CDH1	R---AVVKEPLLPPED--DTRDNVYYYDEEGGGEEDQD-FDLSQLH-----R
CDH3	K---RKIKEPLLLPED--DTRDNVFYYGEEGGGEEDQD-YDITQLH-----R
CDH22	-----HKSHLSSDED--EDMRDNVIKYNDEGGGEQDTEAYDMSALR-----SLYDFG--
CDH5	L-----RKQARAHGKSVP EIH EQLVTYDEEGGGMDDTTSYDVSVLN-----S
CDH26	L---EPKRHGCSVSND--EGHQTLVMYNAESKGTSAQTWSDVEGQRPALLICTAAAGPTQ
	. . . : : : * . : . * *

CDH6	-----PEAIEDN-----KLRRDI
CDH10	-----PAAIEEK-----KLRRDI
CDH9	-----PEAREDS-----KLRRDV
CDH12	-----PKVIEEN-----KIRREDI
CDH18	-----PSAAEEL-----KYRRDI
CDH11	-----PDGINGF-----IPRKDI
CDH8	-----PDGINGF-----LPRKDI
CDH24	-----PDGAAPPAPGP--PARRDV
CDH7	-----STIMRER-----KTRKTT
CDH19	-----STIMRER-----KTRKTT
CDH20	-----PREAQAG---AAPKTRQDM
CDH4	-----PEAMGHVPSKAP-----GVRRVDE
CDH2	-----PDTVEPDAIKPV-----GIRRMDE
CDH15	-----PTALSLPLGPPP-----LRRDAP
CDH1	-----GLDARPE-----VTRNDV
CDH3	-----GLEARPEV-----VLRNDV
CDH22	-----ELKGGDGGGSAGGGAGGSGGGAGSPQAH
CDH5	-----VRRGGAKPPR--PALDARPSL
CDH26	G-----VKDLEEVPPSAASQSAQARCAL

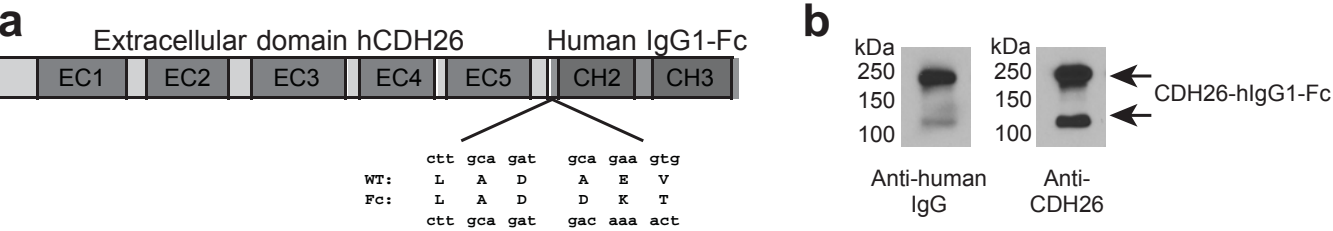
Alignment of CDH26 and CDH1 only:

CDH26	LEPKRHGCSVSND EGHQTLVMYNAESKGTSAQTWSDVEGQRPALLICTAAAGPTQGVKDL
CDH1	RAVVKEPLLPPEDDTRDNVYYYDEEGGGEEDQDFDLSQLHR-----GLDARPEVTRNDV
	: . : * : : : * : * . * : . : * * * :

d

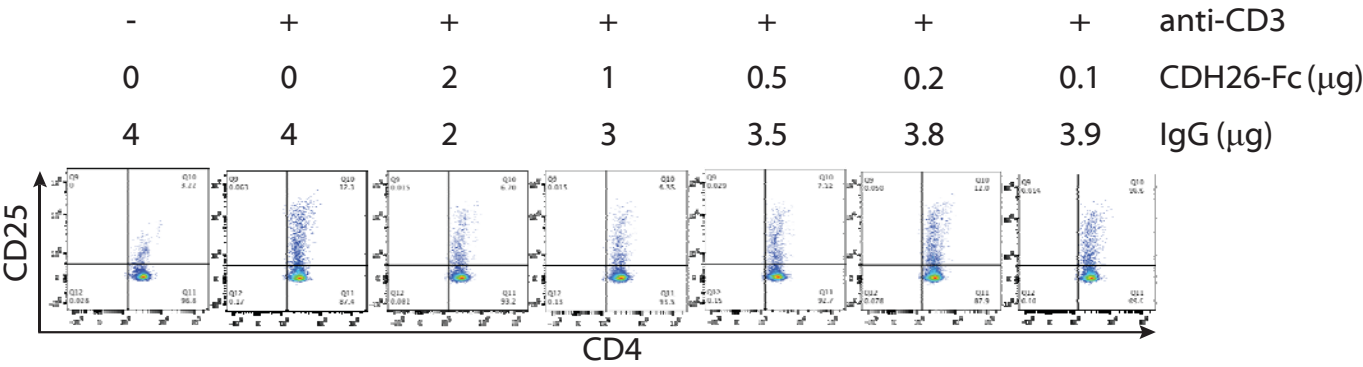


Supplementary Figure S3

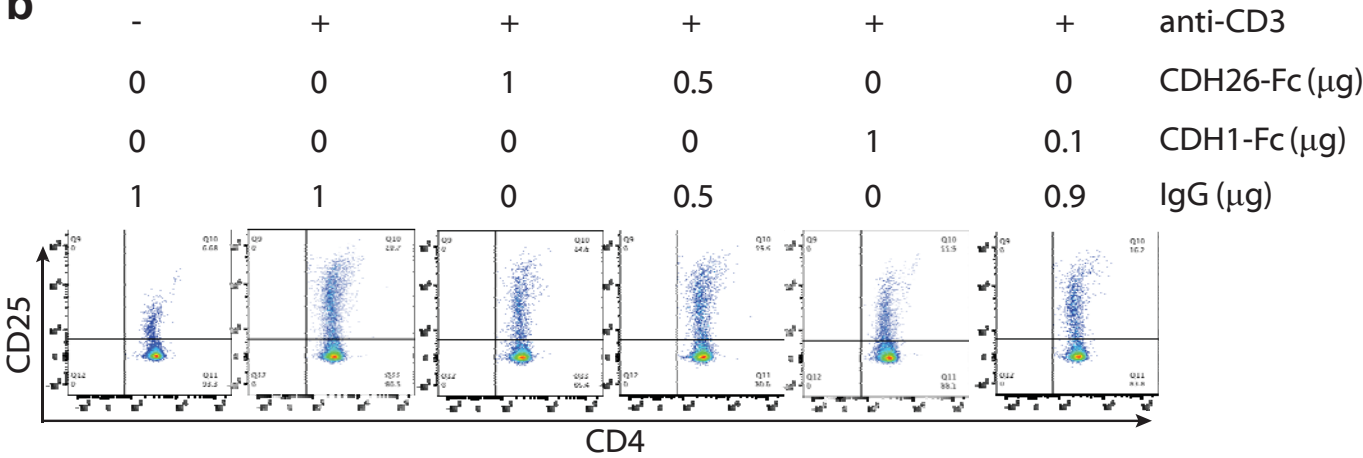


Supplementary Figure S4

a



b



Supplementary Figure S5

