E-cadherin interactome complexity and robustness resolved by quantitative proteomics
Zhenhuan Guo, Lisa J. Neilson, Hang Zhong, Paul S. Murray, Sara Zanivan and Ronen Zaidel-Bar (December 2, 2014)
Science Signaling 7 (354), rs7. [doi: 10.1126/scisignal.2005473]

The following resources related to this article are available online at http://stke.sciencemag.org.
This information is current as of December 2, 2014.

Article Tools
Visit the online version of this article to access the personalization and article tools:
http://stke.sciencemag.org/content/7/354/rs7

Supplemental Materials
"Supplementary Materials"
http://stke.sciencemag.org/content/suppl/2014/11/26/7.354.rs7.DC1.html

Related Content
The editors suggest related resources on Science's sites:
http://stke.sciencemag.org/content/sigtrans/2/89/pe58.full.html
http://stke.sciencemag.org/content/sigtrans/7/312/ra15.full.html
http://stke.sciencemag.org/content/sigtrans/2/87/ra51.full.html
http://stke.sciencemag.org/content/sigtrans/6/265/ra14.full.html
http://stke.sciencemag.org/content/sigtrans/4/177/ra41.full.html

References
This article cites 61 articles, 31 of which you can access for free at:
http://stke.sciencemag.org/content/7/354/rs7#BIBL

Glossary
Look up definitions for abbreviations and terms found in this article:
http://stke.sciencemag.org/cgi/glossarylookup

Permissions
Obtain information about reproducing this article:
http://www.sciencemag.org/about/permissions.dtl
E-cadherin interactome complexity and robustness resolved by quantitative proteomics

Zhenhuan Guo, Lisa J. Neilson, Hang Zhong, Paul S. Murray, Sara Zanivan, Ronen Zaidel-Bar

E-cadherin–mediated cell-cell adhesion and signaling plays an essential role in development and maintenance of healthy epithelial tissues. Adhesiveness mediated by E-cadherin is conferred by its extracellular cadherin domains and is regulated by an assembly of intracellular adaptors and enzymes associated with its cytoplasmic tail. We used proximity biotinylation and quantitative proteomics to identify 561 proteins in the vicinity of the cytoplasmic tail of E-cadherin. In addition, we used proteomics to identify proteins associated with E-cadherin–containing adhesion plaques from a cell-glass interface, which enabled the assignment of cellular localization to putative E-cadherin–interacting proteins. Moreover, by tagging identified proteins with GFP (green fluorescent protein), we determined their subcellular localization of 83 putative E-cadherin–proximal proteins and identified 24 proteins that were previously uncharacterized as part of adherens junctions. We constructed and characterized a comprehensive E-cadherin interaction network of 79 published and 394 previously uncharacterized proteins using a structure-informed database of protein-protein interactions. Finally, we found that calcium chelation, which disrupts the interaction of the extracellular E-cadherin domains, did not disrupt most intracellular protein interactions with E-cadherin, suggesting that the E-cadherin intracellular interactome is predominantly independent of cell-cell adhesion.

INTRODUCTION

Epithelial (E)-cadherin (CDH1) is required for cell-cell adhesion in embryonic and adult epithelia (1). E-cadherin is essential for development of all animals, and decreased abundance of E-cadherin is associated with cancer metastasis (2–4). E-cadherin receptors cluster and organize into larger dynamic structures termed adherens junctions (2, 5). The extracellular domains of E-cadherin are sufficient to mediate both trans and cis clustering and promote cell-cell adhesion (6). However, interaction of the cytoplasmic tail of E-cadherin with F-actin is crucial for the stability of E-cadherin clusters (7). Moreover, the coupling of E-cadherin with the actomyosin cytoskeleton, which enables force generation, facilitates the integration of individual cell shape changes into dynamic tissue rearrangements, such as those that occur during wound healing or morphogenesis (8). Proteins, such as catenins (9), mediate the physical linkage of E-cadherin with the actin cytoskeleton (10). Moreover, an array of signaling enzymes regulates the interactions among cadherins, catenins, and the cytoskeleton through posttranslational modifications (11). Regulatory proteins localized on the cytoplasmic side of cell-cell junctions control actin dynamics and endocytosis (12, 13). Together, this ensemble of structural and regulatory proteins and their interactions is called the cadherin adhesome (“cadhesome”) (14). The dynamics of cell-cell adhesion are an emergent property of protein interactions within the cadhesome (15). Therefore, elucidating the components and interactions within the cadhesome is an essential step in understanding how cell adhesion is regulated in health and misregulated in disease.
provide a comprehensive analysis of the molecular composition, dependence on cell-cell contact, and localization of components of the E-cadherin interactome. This analysis validated many published components of the cadhesome, identified many previously uncharacterized components of the cadhesome, and enabled predictions regarding the organization and regulation of E-cadherin adhesions. We found that many intracellular proteins interacted with E-cadherin independent of trans-ligation of cadherins (cell-cell adhesion).

RESULTS

Isolation of E-cadherin–proximal proteins with E-cad-BirA*

To identify proteins associated with E-cadherin, we used BioID, a technique that enables spatially restricted protein biotinylation by a promiscuous mutant of the bacterial biotin ligase BirA (BirA*). In the presence of biotin, BirA* and BirA*-fusion proteins generate a local cloud of biotinoyl-5′-adenosine monophosphate (bioAMP), the active form of biotin that reacts with primary amines (31), and thereby promote the biotinylation of neighboring proteins representing both direct and indirect protein interactions (32–34). We fused BirA* to the C terminus of E-cadherin (E-cad-BirA*) and generated a stable line of MKN28 human gastric adenocarcinoma cells (MKN28-E-cad-BirA* cells). E-cad-BirA* colocalized at adherens junctions with endogenous E-cadherin in MKN28-E-cad-BirA* cells, and these junctions were indistinguishable from those in parental MKN28 cells with regard to the organization of E-cadherin, catenins, IQGAP, and F-actin (Fig. 1A). To determine where biotinylation occurred in MKN28-E-cad-BirA* cells, we labeled cells with streptavidin conjugated to the fluorophore Alexa 647. In parental MKN28 cells incubated with biotin, streptavidin labeling was weak and diffuse, whereas in MKN28-E-cad-BirA* cells incubated with biotin, streptavidin robustly and specifically labeled cell-cell junctions (Fig. 1B). Without the addition of biotin, we detected little to no streptavidin signal in MKN28-E-cad-BirA* cells incubated with biotin, streptavidin robustly and specifically labeled cell-cell junctions (Fig. 1B). Without the addition of biotin, we detected little to no streptavidin signal in MKN28-E-cad-BirA* cells incubated with biotin, streptavidin robustly and specifically labeled cell-cell junctions (Fig. 1B), demonstrating the specificity of the labeling as well as the minimal amount of endogenous biotinylated proteins. Double labeling demonstrated overlap between biotinylated proteins and E-cadherin at cell-cell junctions in MKN28-E-cad-BirA* cells incubated with biotin (Fig. 1C). Moreover, Western blot analysis with streptavidin–horseradish peroxidase (HRP) confirmed that stable expression of E-cad-BirA* increased the abundance of biotinylated proteins in MKN28 cells incubated with biotin (Fig. 1D).

To identify proteins associated with E-cadherin, we isolated the biotinylated proteins using streptavidin affinity chromatography from lysates of intact monolayers of MKN28-E-cad-BirA* cells incubated with or without the calcium chelator EGTA. In addition, we performed streptavidin affinity chromatography on lysates of parental MKN28 cells as a negative control. Quadruplicate samples enriched for biotinylated proteins from each condition were trypsinized, analyzed by high-resolution MS, and quantified using an advanced label-free quantification (LFQ) algorithm, MaxLFQ (35). We omitted proteins identified in parental MKN28 cells from the list of those identified in MKN28-E-cad-BirA* cells (table S1). We quantified 560 proteins that were represented by at least two peptides in three of four replicates of MKN28-E-cad-BirA* cells incubated with or without the calcium chelator EGTA (table S1). Rab10 was not statistically enriched in MKN28-E-cad-BirA* cells compared to parental MKN28 cells but was added to this list. We estimated the abundance of each protein using the iBAQ sum, which is the sum of peptide intensities for a given protein normalized to the number of theoretical tryptic peptides (as estimation of the protein size) (36). The iBAQ sums of the proteins identified across the eight replicates spanned five orders of magnitude (Fig. 2A), demonstrating a wide range of quantifiable protein abundances, and did not correlate with the number of lysine residues per protein (fig. S1), excluding the possibility of a bias toward

![Fig. 1. Characterization of MKN28-E-cad-BirA* cells.](image-url)
proteins with abundant primary amines. Proteins were ranked according to their combined iBAQ values in control and EGTA conditions, and the 35 most abundant proteins (excluding ribosomal proteins) accounted for more than 65% of the total mass of all the quantified proteins. Most of these proteins were either components of the literature-based cadhesome or paralogs of such components (14) (Table 1). However, literature-based cadhesome components were also present among the less abundant proteins identified (Fig. 2A), suggesting that there was only a weak correlation between the iBAQ sum of a protein and its likelihood of being a cadhesome component. Altogether, 82 of the 561 quantified proteins were members of the literature-based cadhesome (14) (table S1). On the basis of literature searches, we identified an additional 28 proteins as having interactions with E-cadherin or catenins or as components of tight junctions, desmosomes, or integrin adhesions (table S1).

Characterization of the E-cad-BirA* interactome

We annotated the function of the 561 quantified proteins using UniProt (37), Entrez Gene (38), and the primary literature and manually classified each protein into one of 20 functional categories (table S1). The largest group of proteins (163 proteins) in the E-cadherin interactome consisted of adaptor proteins: 46 of these bind actin, 7 bind microtubules, and 25 bind the plasma membrane (Fig. 2B). Other highly represented groups were transmembrane proteins (including adhesion receptors), guanosine triphosphatase (GTPase) regulators, kinases and phosphatases, actin dynamics regulators, and cytoskeleton structural and motor proteins. Together, these six functional categories accounted for more than 50% of the 561 quantified proteins and resembled the functional composition of the literature-based cadhesome (14) (table S1). The remaining proteins were categorized as being involved in transcription, translation, trafficking, proteolysis and metabolism, or unknown functions. Analysis of the iBAQ sums of the proteins in each category showed that 60% of the mass of the identified proteins...
Chemical cross-linking enabled the isolation of integrin-associated protein interactions, we isolated and characterized E-cadherin junctions and proteins that interact with E-cadherin at other cellular locations. To differentiate between proteins that interact with E-cadherin at adherens junctions containing adhesion plaques, which supports cell-glass E-cadherin junctions (39–41). MKN28 cells were rounded when plated on noncoated glass, but spread out when plated on E-cadherin–coated glass (fig. S2A). Plating MKN28 cells in the presence of an antibody against the extracellular domain of E-cadherin inhibited spreading on E-cadherin–coated glass (fig. S2A). Plating MKN28 cells on E-cadherin–coated glass, but not E-cadherin–coated glass, induced the formation of focal adhesion, as indicated by immunolabeling for paxillin (fig. S2B). To visualize E-cadherin–containing adhesion plaques, we labeled MKN28 cells plated on E-cadherin–coated glass with antibodies against E-cadherin and β-catenin as well as phalloidin to visualize F-actin. Although there were numerous E-cadherin and β-catenin double-positive foci throughout the cell, only foci located near the periphery of the cell colocalized with F-actin (Fig. 3A). Moreover, after chemical cross-linking, cell lysis, and aggressive washing with hydrodynamic pressure, only the F-actin–containing adhesion plaques, which contains foci located near the periphery of the cell colocalized with F-actin (Fig. 3A).

### Table 1. The 35 most abundant proteins in the E-cad-BirA* interactome

<table>
<thead>
<tr>
<th>Mass relative to CDH1</th>
<th>Gene names</th>
<th>Protein names</th>
<th>Functional category</th>
<th>Published component of cadhesome</th>
</tr>
</thead>
<tbody>
<tr>
<td>6.48</td>
<td>CTNNA1</td>
<td>β-E-catenin</td>
<td>Actin binding adaptor</td>
<td>1</td>
</tr>
<tr>
<td>5.77</td>
<td>JUP</td>
<td>Junction plakoglobin</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>5.71</td>
<td>AHNAK</td>
<td>Neuroblast differentiation–associated protein</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>1.74</td>
<td>CTNND1</td>
<td>Catenin; p120-catenin</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>1.73</td>
<td>FLNA</td>
<td>Filamin-A</td>
<td>Actin dynamics regulator</td>
<td>2</td>
</tr>
<tr>
<td>1.23</td>
<td>CTNNB1</td>
<td>β-Catenin</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>1.00</td>
<td>CDH1</td>
<td>E-cadherin; Cadherin-1</td>
<td>Adhesion receptor</td>
<td>1</td>
</tr>
<tr>
<td>0.78</td>
<td>FAM129B</td>
<td>Niban-like protein 1</td>
<td>Membrane binding adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.71</td>
<td>ANXA2</td>
<td>Annexin A2</td>
<td>Membrane binding adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.43</td>
<td>CRKL</td>
<td>Crk-like protein</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.43</td>
<td>SKT</td>
<td>Sickle tail protein homolog</td>
<td>Unknown</td>
<td>No</td>
</tr>
<tr>
<td>0.40</td>
<td>ERBB2IP</td>
<td>ErbB2 interacting protein</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.34</td>
<td>SCRIB</td>
<td>Protein scribble homolog</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.26</td>
<td>CTTN</td>
<td>Src substrate cortacin</td>
<td>Actin dynamics regulator</td>
<td>1</td>
</tr>
<tr>
<td>0.23</td>
<td>EGFR</td>
<td>Epidermal growth factor receptor</td>
<td>Receptor tyrosine kinase</td>
<td>1</td>
</tr>
<tr>
<td>0.22</td>
<td>DLG1</td>
<td>Disks large homolog 1</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.22</td>
<td>KRT18</td>
<td>Keratin, type I cytoskeletal 18</td>
<td>Cytoskeleton</td>
<td>No</td>
</tr>
<tr>
<td>0.21</td>
<td>TJP1</td>
<td>Tight junction protein ZO-1</td>
<td>Actin binding adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.21</td>
<td>CTNNA2</td>
<td>α-N-catenin</td>
<td>Actin binding adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.19</td>
<td>MKL2</td>
<td>MKL myocardin-like protein 2</td>
<td>Transcription factor</td>
<td>No</td>
</tr>
<tr>
<td>0.15</td>
<td>ANXA1</td>
<td>Annexin A1</td>
<td>Membrane binding adaptor</td>
<td>2</td>
</tr>
<tr>
<td>0.15</td>
<td>TNKS1BP1</td>
<td>Tankyrase-1-binding protein</td>
<td>Adaptor</td>
<td>No</td>
</tr>
<tr>
<td>0.15</td>
<td>TAGLN2</td>
<td>Transgelin-2</td>
<td>Actin binding adaptor</td>
<td>No</td>
</tr>
<tr>
<td>0.14</td>
<td>NDRG1</td>
<td>N-myc downstream regulated 1</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.14</td>
<td>EPB41</td>
<td>Protein 4.1</td>
<td>Membrane binding adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.12</td>
<td>SNAP29</td>
<td>Synaptosomal-associated protein 29</td>
<td>Vescicle docking or fusion</td>
<td>No</td>
</tr>
<tr>
<td>0.12</td>
<td>EFHD2</td>
<td>EF-hand domain-containing protein D2</td>
<td>Unknown</td>
<td>No</td>
</tr>
<tr>
<td>0.11</td>
<td>SLC3A2</td>
<td>Solute carrier family 3, member 2</td>
<td>Transmembrane transporter</td>
<td>3</td>
</tr>
<tr>
<td>0.12</td>
<td>NUMB</td>
<td>Protein numb homolog</td>
<td>Adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.16</td>
<td>VAMP3</td>
<td>Vescicle-associated membrane protein 3</td>
<td>Vescicle docking or fusion</td>
<td>No</td>
</tr>
<tr>
<td>0.12</td>
<td>PDLIM1</td>
<td>PDZ and LIM domain protein 1</td>
<td>Actin binding adaptor</td>
<td>No</td>
</tr>
<tr>
<td>0.09</td>
<td>MARK2</td>
<td>Serine and threonine kinase MARK2</td>
<td>Protein kinase</td>
<td>No</td>
</tr>
<tr>
<td>0.09</td>
<td>ARHGAP32</td>
<td>Rho GTPase–activating protein 32</td>
<td>Rho GAP</td>
<td>1</td>
</tr>
<tr>
<td>0.08</td>
<td>MLLT4</td>
<td>Afdin</td>
<td>Actin binding adaptor</td>
<td>1</td>
</tr>
<tr>
<td>0.07</td>
<td>S100A11</td>
<td>Protein S100-A11</td>
<td>Adaptor</td>
<td>2</td>
</tr>
</tbody>
</table>

is made up of adaptors, 25% of translation-related proteins, and 15% of all the other functional categories (Fig. 2C). We determined the number of proteins for each functional category that were identified in our MS analysis but previously uncharacterized as interacting with E-cadherin, those identified in our MS analysis that overlapped with the literature-based cadhesome, and those that were present in the literature-based cadhesome but absent from our MS analysis (Fig. 2D). This analysis showed that some literature-based cadhesome components, mostly consisting of adaptors, GTPase regulators, and kinases or phosphatases, were not detected by the E-cadherin proximity biotinylation method, but that many proteins that were identified by the E-cadherin proximity biotinylation method and previously uncharacterized as components of the cadhesome were involved in functions known to be associated with the cadhesome, such as adaptors and actin dynamics regulators.

**Isolation and characterization of proteins associated with E-cadherin adhesion plaques**

To differentiate between proteins that interact with E-cadherin at adherens junctions and proteins that interact with E-cadherin at other cellular locations, we isolated and characterized E-cadherin–containing adhesion plaques. Chemical cross-linking enabled the isolation of integrin-associated protein complexes from cells adhering to fibronectin-coated plates or beads (28, 29). We adapted this method for the isolation of E-cadherin adhesion plaques by coating glass coverslips with recombinant E-cadherin extracellular domains, which supports cell-glass E-cadherin junctions (39–41). MKN28 cells were rounded when plated on noncoated glass, but spread out when plated on E-cadherin–coated glass (fig. S2A). Plating MKN28 cells in the presence of an antibody against the extracellular domain of E-cadherin inhibited spreading on E-cadherin–coated glass (fig. S2A). Plating MKN28 cells on E-cadherin–coated glass, but not E-cadherin–coated glass, induced the formation of focal adhesion, as indicated by immunolabeling for paxillin (fig. S2B). To visualize E-cadherin–containing adhesion plaques, we labeled MKN28 cells plated on E-cadherin–coated glass with antibodies against E-cadherin and β-catenin as well as phalloidin to visualize F-actin. Although there were numerous E-cadherin and β-catenin double-positive foci throughout the cell, only foci located near the periphery of the cell colocalized with F-actin (Fig. 3A). Moreover, after chemical cross-linking, cell lysis, and aggressive washing with hydrodynamic pressure, only the F-actin–containing adhesion plaques remained (Fig. 3A).

We isolated the chemically cross-linked E-cadherin– and F-actin–containing adhesion plaques for analysis by high-resolution MS in three replicate experiments and identified a total of 1980 proteins, which were...
plotted by functional categories. Among these were 70 members of the literature-based cadhesome (14), including the core components E-cadherin, p120-catenin, β-catenin, α-catenin, vinculin, and eplin, as well as many other proteins from diverse functional groups. We found that 275 proteins overlapped between the proteins identified in MKN28-E-cad-BirA* cells and those isolated from adhesion plaques from E-cadherin–coated glass (table S1).

To determine the putative functions of the proteins that overlapped between these data sets, we assessed the functional categories for these proteins as described above. The annotated molecular functions of the overlapping data set were similar to that of the full data set derived from MKN28-E-cad-BirA* cells (Fig. 3B). Although proteins in the largest subgroup in the overlapping data set were annotated as adaptors, the second and third largest subgroups of proteins were annotated as being involved in trafficking or translation-related proteins (Fig. 3B), suggesting that interactions between E-cadherin and proteins involved in these processes may take place at the plasma membrane.

Identification of E-cadherin interactome components that localize at adherens junctions

We predicted that E-cad-BirA* could biotinylate vicinal proteins regardless of whether it localized at cell junctions or elsewhere in the cell (for example, in vesicles). Therefore, we examined the subcellular localization of proteins identified by MS in MKN28-E-cad-BirA* cells. We created C-terminal EGFP fusions of 107 candidate E-cadherin interacting proteins corresponding to 18 of 20 functional categories. We could not obtain complementary DNA (cDNA) for proteins from the remaining two categories. We transiently transfected plasmids encoding GFP-fusion proteins into parental MKN28 cells. For 83 fusion proteins with observable GFP, we determined their subcellular localization by confocal microscopy (Fig. 4A and table S3): 26 of the proteins localized to cell-cell junctions, 42 to the cytoplasm, and 15 to other specific cellular areas, most often in or around the nucleus or in vesicles. Overall, 31% of proteins showed junctional localization (Fig. 4B). However, none of the proteins annotated as being involved in transcription, translation, or metabolism showed junctional localization, whereas 55% of actin dynamics regulators and 50% of adaptors showed junctional localization (Fig. 4B). This suggested that many proteins may interact with E-cadherin at cellular locations other than cell junctions. However, because the EGFP tag could interfere with protein-protein or other molecular interactions required for protein localization, our observations do not preclude the possibility that the endogenous version of these proteins may have junctional localization. We found that 26 EGFP-fusion proteins that localized to cell-cell junctions colocalized with endogenous E-cadherin (fig. S3), suggesting that these proteins interact with E-cadherin at adherens junctions.

Construction of an E-cadherin interactome using known and predicted protein-protein interactions

To map published and predicted interactions among the proteins identified in the MKN28-E-cad-BirA* cells (hereafter referred to as the “E-cad-BirA* interactome”), we used PrePPI, a structure-informed database of human protein-protein interactions (42, 43). This analysis identified 2035 interactions among 394 of the 556 biotinylated proteins that were present in the PrePPI database (Fig. 5A and table S4). The 162 proteins for which no interactions were identified consisted primarily of metabolic enzymes, ribosomal and trafficking proteins, and proteins with no annotated function (table S4). As a negative control, we used PrePPI to map a network of interactions between the set of 81 literature-based cadhesome proteins that were also identified in MKN28-E-cad-BirA* cells and a set of 475 proteins randomly selected from the human proteome. This analysis identified 681 interactions among 226 proteins and 187 interactions between the literature-based cadhesome proteins and randomly selected proteins (table S4). In contrast, there were 726 interactions between proteins in the literature-based cadhesome and proteins in the E-cad-BirA* data set that were previously uncharacterized with respect to E-cadherin (table S4). Likewise, there were 978 interactions...
among previously uncharacterized E-cadherin–interacting proteins, compared to 167 interactions among randomly selected proteins (table S4).

We analyzed the topology of the E-cad-BirA* interactome. Using the network analyzer (44) in Cytoscape (45), we found that the topology of the E-cad-BirA* interactome followed the power law typical of other biological networks (46) (fig. S4A), and determined the clustering coefficient, average number of neighbors, and characteristic path length to be 0.29, 10.33, and 3.14, respectively.

We used the E-cad-BirA* interactome to categorize proteins according to the network distance from E-cadherin. We plotted the iBAQ sum of a protein as a function of network distance and found that the median iBAQ sum for first-degree interactors of E-cadherin was threefold higher than that for second-degree interactors (fig. S4B). There was no difference between second- and third-degree or second- and fourth-degree interactors.

To visualize the E-cad-BirA* interactome as a function of connectivity, we arranged the 394 proteins into concentric circles according to the number of interactions per protein and color-coded proteins according to whether they were previously uncharacterized or published components of the literature-based cadhesome (Fig. 5A). Forty-one percent of the 64 proteins in the three innermost circles (>20 interactions) were published cadhesome components, whereas only 12% of the 254 proteins in the outermost circle (1 to 9 interactions) were published cadhesome components. The core of the E-cad-BirA* interactome (>20 interactions) contained 38 proteins that were not previously published to interact with E-cadherin (Fig. 5A). These proteins, which are likely to be bona fide components of the cadhesome, were primarily annotated as membrane-bound adaptors, cytoskeleton-bound adaptors, or cytoskeletal proteins (Fig. 5A).

To focus on the structural elements of adherens junctions within the E-cad-BirA* interactome, we visualized proteins annotated as adhesion receptors, membrane binding adaptors, actin dynamics regulators, cytoskeleton and motor proteins, and adaptors, according to the network distance from E-cadherin (Fig. 5B). This analysis revealed a large number of previously uncharacterized proteins that may serve to link the plasma membrane at adherens junctions with the actin cytoskeleton. Several of the proteins identified in MKN28-E-cad-BirA* cells were annotated with functions not previously considered integral to the function of the cadhesome, including trafficking, transcription, and translation. We examined the connectivity between proteins within these functional groups and published cadhesome components using the E-cad-BirA* interactome (fig. S5). We found 369 interactions among 91 proteins annotated as adaptors that were not previously published to interact with E-cadherin and 79 published cadhesome components. Moreover, we found that 79 published cadhesome components participated in 49 interactions with previously uncharacterized E-cadherin–interacting proteins annotated with trafficking, Golgi, or endoplasmic reticulum functions; in 36 interactions with those annotated with DNA, transcription and RNA, or translation functions; and in 23 interactions with those annotated as metabolic enzymes (fig. S5).

Response of E-cad-BirA* interactome to reduction of free calcium
To assess the extent to which the E-cad-BirA* interactome depends on the formation of cadherin-mediated cell-cell junctions, we investigated the
Fig. 5. E-cad-BirA* protein interaction network. (A) Diagram of 2035 protein-protein interactions [probability ≥ 50%, derived from the PrePPI database (42)] for 394 proteins identified in MKN28-E-Cad-BirA* cells by MS. Proteins are grouped in concentric circles on the basis of the number of interactions. The circles correspond to 1 to 9, 10 to 19, 20 to 29, 30 to 39, or 40 to 80 from outside to inside. Proteins in the literature-based cadhesome (14) are depicted in blue, except for E-cadherin (CDH1), which is depicted in yellow; proteins not previously characterized with respect to E-cadherin are depicted in green, and proteins with junctional localization of GFP-fusion proteins (Fig. 4) are depicted in green with a blue border. (B) Diagram of a subnetwork of the network in (A) containing only adhesion receptors, actin dynamics regulators, cytoskeleton and motor proteins, and adaptors, arranged by network distance relative to E-cadherin.
differences between biotinylated proteins from MKN28-E-cad-BirA* cells incubated with or without EGTA. Incubation with EGTA led to the dissociation of cell-cell adhesions in MKN28 cells, which detached from their neighbors and adopted a rounded morphology, but remained attached to the substrate even after 24 hours of incubation with EGTA (Fig. 6A and fig. S6). MKN28 cells were returned to calcium-rich medium after 24 hours of incubation with EGTA, and the cells spread out and reformed cell-cell junctions (fig. S6), indicating that they were still viable. We compared the LFQ intensity of individual proteins from MKN28-E-cad-BirA* cells incubated with or without EGTA and found that most (60%) of the proteins showed less than a twofold change (Fig. 6B and table S1). Proteins with the highest abundance, including E-cadherin and catenins, showed the smallest difference between conditions (table S1).

The independence of the E-cad-BirA* interactome from the formation of cell-cell junction was unexpected, and thus, to corroborate this finding, we used immunofluorescence to examine the colocalization of published cadhesome components with E-cadherin in the presence of EGTA. Similar to MKN28 cells grown in control medium (Fig. 1), in MKN28 cells incubated with EGTA for 24 hours, β-catenin, α-catenin, γ-catenin, and IQGAP colocalized with E-cadherin at the periphery of the cells (Fig. 6C). E-cadherin undergoes endocytosis in the absence of cell-cell contacts (23, 24). To address the possibility that calcium chelation interfered with endocytosis of E-cadherin resulting in junctional components remaining at the membrane, we examined the colocalization of published cadhesome components with E-cadherin in sparsely seeded MKN28 cells that had not yet formed cell-cell junctions. This analysis revealed that the colocalization of E-cadherin with β-catenin, α-catenin, γ-catenin, or IQGAP at the plasma membrane was similar to that in cells incubated with EGTA (fig. S7), suggesting that the interaction of E-cadherin with cadhesome components is likely independent of cell-cell junctions.

Biotinylated proteins that showed greater than a twofold change in abundance between MKN28-E-cad-BirA* cells incubated with or without EGTA (Fig. 6B) could represent proteins important for E-cadherin-associated signaling. To better understand the nature of these proteins, we analyzed the distribution of protein domains among biotinylated proteins using Pfam (47) (table S5). We found that proteins from MKN28-E-cad-BirA* cells incubated in the absence of EGTA were enriched (defined as the ratio of the number of proteins greater than 1.5) for ankyrin and armadillo repeats, as well as PAK (p21-activated protein kinase)-binding, guanylate kinase, SH3, LIM, and PDZ domains. In contrast, proteins from MKN28-E-cad-BirA* cells incubated in the presence of EGTA were enriched for spectrin, calponin homology, and FERM domains (table S5). Thus, different protein domains may confer different abilities to associate with E-cadherin in the presence or absence of cell-cell junctions.

To experimentally test this prediction, we used live cell imaging of MKN28 cells expressing GFP fused to the zyxin family protein TRIP6, which contains three LIM domains and was more abundant among biotinylated proteins isolated from MKN28-E-cad-BirA* cells incubated without EGTA (table S1). In contrast to E-cadherin-GFP, which localized to the cell cortex in the presence or absence of EGTA, TRIP6-GFP was present at cell-cell junctions before exposure to EGTA and rapidly dissociated from the cell cortex in the presence of EGTA (Fig. 6D).

**DISCUSSION**

We used BioID (34) to identify E-cadherin vicinal proteins in MKN28 cells. A similar study using BioID in MDCK cells identified 303 proteins associated with E-cadherin (26), 114 of which overlapped with the 561 proteins identified here (table S1). The differences in the identified proteins could be due to differences in cell types or the specific experimental conditions. For example, whereas our experiments used parental cells as a negative control, Van Itallie et al. (26) used cells expressing cytoplasmic BirA*. Nevertheless, the observed overlap between these studies supports the conclusion that the BioID approach using E-cad-BirA* is a robust and reliable method to identify proteins associated with E-cadherin. About 50% of proteins identified using BioID are present within 20 to 30 nm of BirA* (34). Moreover, a BioID study focused on ZO-1 interacting proteins found that BirA* fused to either the N or C terminus of ZO-1 yielded partially
overlapping sets of biotinylated proteins (48), suggesting that the radius of biotinylation is on the order of the size of ZO-1.

The E-cad-BirA* BioID approach used here successfully identified 82 of 173 literature-based cadhesome proteins. The remaining 91 published cadhesome components that were not biotinylated by E-cad-BirA* were mostly cytoplasmic proteins, primarily including kinases, phosphatases, and GTPases. We used a brief permeabilization step before lysis of cells that likely introduced a bias toward identification of insoluble proteins. Furthermore, the detection of biotinylated proteins in this assay likely depended not only on proximity to E-cadherin but also on the abundance and turnover rate of individual proteins. Moreover, some cadhesome components may not be present in MKN28 cells, or the subcellular localization of these proteins could depend on specific culture conditions.

In addition to the identification of previously uncharacterized E-cadherin interacting proteins with functional annotations similar to those of published cadhesome components, we identified proteins with functions that are not published to be integral to the cadhesome. For example, we found proteins involved in transcription and translation. Some cadhesome components, such as β-catenin, function both as structural components of cell-cell junctions and as transcriptional coactivators in the nucleus (49). Similarly, mRNA and translation machinery can localize to cell-cell junctions (50). Moreover, the cytoplasmic tail of E-cadherin can be cleaved and translocate to the nucleus (51). Whereas our analysis of the localization of GFP-fusion proteins showed that many of the candidate E-cadherin interacting proteins annotated with transcriptional or translational functions did not localize predominantly to cell-cell junctions, the MS analysis of isolated E-cadherin adhesion plaques indicated that at least some of these proteins are localized at the plasma membrane in proximity to E-cadherin.

E-cadherin protein interactions that occur in cellular locations other than cell-cell junctions may have effects on cell adhesion. A genome-wide RNA interference (RNAi) screen in Drosophila S2 cells discovered more than 800 genes encoding proteins required for cadherin-mediated cell adhesion, and only a small fraction of these localized to cell-cell junctions (52). Fifty-five proteins overlapped between those from this RNAi screen and the E-cad-BirA* interactome described here. Among the overlapping proteins, 31 are involved in trafficking, translation, or transcription (table S1). Moreover, 48 proteins, including the transmembrane receptor SCARBL1, the Rho GTPTase–activating proteins ARHGAP1 and ARHGAP21, and the actin binding adaptors ANLN, CORO1B, and LASP1, are not reported to interact with E-cadherin but were present in our proteomics analysis and the functional genomic screen and, thus, are promising candidates for further investigation.

We found that the composition of the E-cad-BirA* interactome was largely independent of cell-cell adhesion, suggesting that the difference between adherens junctions and nonadhesive E-cadherin clusters is not in the composition of high abundance plaque proteins, but possibly in the structural organization or posttranslational modifications of these proteins. In contrast, we observed that biotinylated proteins with low iBAQ enrichments showed larger fold-change differences between cells incubated with and without EGTA. The low abundance of these proteins in the biotinylated fraction suggests that either the overall abundance of the protein was low or the protein weakly interacted with E-cadherin. In contrast to stable protein interactions, such as those between receptors and adaptors or scaffolding proteins, transient protein interactions, such as those between kinases and substrates, often have signaling and regulatory roles.

Among the differently biotinylated, low-abundance proteins in the E-cad-BirA* interactome, there was an enrichment for specific protein domains, suggesting specific regulatory functions. For example, calcium chelation reduced the E-cad-BirA*–dependent biotinylation of the LIM domain–containing protein TRIP6 and caused relocation of TRIP6-GFP away from cell-cell junctions. Similar to zyxin, TRIP6 is a component of focal adhesions that promotes cell migration and also functions as a transcription coactivator in the nucleus (53). TRIP6 interacts with the cell-cell junction–associated protein MAGI-1 (membrane-associated guanylate kinase WW and PDZ domain–containing protein 1) (53), but its function at adherens junctions is not known. The association between TRIP6 and E-cadherin in an adhesion-dependent manner suggests that TRIP6 may be involved in signaling downstream of E-cadherin trans-ligation.

**MATERIALS AND METHODS**

**Plasmids and cell culture**

The E-cad-BirA* fusion construct was created by replacing GFP in E-cadherin–GFP (5) (a gift from J. Nelson, Stanford University) with BirA* (34) (a gift from B. Burke, Institute of Molecular Biology, Agency for Science, Technology and Research, Singapore) using Xho I and Mfe I restriction sites. For establishing a stable cell line, E-cad-BirA* was cloned into pJTI Fast DEST vector (Life Technologies) with a cytomegalovirus promoter. pJTI-E-cad-BirA* was stably integrated into MKN28 cells, using Jump-In Fast Gateway Kit (Life Technologies). Colonies expressing E-cad-BirA* were selected with hygromycin B (200 µg/ml) and screened by streptavidin staining.

For GFP fusion proteins, entry vectors containing open reading frames (gift from T. C. Cornvik, Nanyang Technical University, Singapore) were recombined with pcDNA-6.2-C-EmGFP-DEST (Life Technologies), using Gateway cloning (Life Technologies). The resulting cDNA-GFP fusion constructs were transfected into MKN28 cells, using Lipofectamine 2000 (Life Technologies).

**Immunolabeling**

Cells were fixed with paraformaldehyde (4%) in phosphate-buffered saline (PBS) for 30 min and permeabilized with 0.1% Triton X-100 (0.1% in PBS) for 20 min. Primary antibodies used were E-cadherin (BD Biosciences, 610181, or Sigma, U3254), β-catenin (Sigma, C7082), α-catenin (Cell Signaling, 3134), γ-catenin (Abcam, 15153), and IQGAP1 (Santa Cruz Biotechnology, 10792). Reagents used were streptavidin–Alexa 647 (Life Technologies, S32357) and TRITC (tetramethyl rhodamine isothiocyanate)–phalloidin (Sigma, 77418). Species-specific secondary antibodies used were rat (Life Technologies, A10777), mouse (Life Technologies, A31571, or Abnova, PAB 10733), and rabbit (Abcam, 6799 or ab6717).

**Microscopy**

Confocal images were acquired using a Nikon Eclipse Ti inverted microscope with a spinning-disk confocal head (CSU-X1, Yokogawa Corporation), a laser launch unit (iLas2, Roper Scientific), an electron-multiplying charge-coupled device camera (Evolve Rapid-Cal, Photometrics), 60× or 100× Plan Apo objectives (Nikon), and MetaMorph software (Molecular Devices).

**Purification of biotinylated proteins**

MKN28 or MKN28-E-cad-BirA* cells (140-mm dish) were grown to confluence in biotin-free medium (Dulbecco's modified Eagle's medium + 10% dialyzed fetal bovine serum) and then switched to medium containing biotin (0.2 mg/liter) with or without EGTA (1 mM) for 24 hours. Cells were permeabilized with ice-cold Tween 20 (1% in PBS) for 1 min, washed twice with PBS, lysed in 3 ml of lysis buffer [50 mM tris (pH 7.4), 500 mM NaCl, 0.4% SDS, 5 mM EDTA, 1 mM dithiothreitol (DTT), and protease inhibitor (Roche)], and sonicated as described previously (34). Supernatants were cleared by centrifugation, and biotinylated proteins were enriched with 300 µl of streptavidin beads (MyOne Streptavidin C1, Life
Isolation of E-cadherin–specific adhesion plaques

Coverslips were not coated or coated with recombinant human E-cadherin–Fc chimera (30 μg/ml, R&D Systems) or fibronectin (10 μg/ml, Sigma-Aldrich) for 2 hours at room temperature and rinsed with PBS with calcium and magnesium (Life Technologies). MKN28 cells were trypsinized and seeded on coverslips in RPMI without serum. An antibody against E-cadherin extracellular domain (DEMCA-1, 1:100; Sigma, U3254) was added to the medium where indicated. Cells were incubated for 6 hours to recover and spread. For MS analysis, cells were seeded at high density (>6 × 10^4 cells/cm^2) in triplicate on round coverslips (50-mm diameter, custom-made by Paul Marienfeld GmbH & Co. KG), and adhesion plaques were isolated according to a modification of a published protocol (29). Briefly, cells were rinsed with PBS with calcium and magnesium (PBS+[Ca^2+/Mg^2+], Life Technologies, 14080-055), cross-linked with DSP [1,4-bis(3-(2-pyridyldithio)propionamido)butane, 50 μM; Sigma-Aldrich] in PBS+[Ca^2+/Mg^2+] for 15 min, and rinsed with tris (50 mM) for 15 min to quench cross-linkers. Cross-linked cells were lysed for 30 min with ice-cold lysis buffer [25 mM tris-HCl (pH 7.5), 150 mM NaCl, 1% Triton X-100, 0.2% SDS, 0.5% sodium deoxycholic acid, and protease inhibitors (Roche)] and washed extensively with a water jet (Oral-B Oxyjet + 1000). Cross-linked protein residues were eluted from coverslips at 60°C for 30 min with an elution buffer [25 mM tris-HCl (pH 7.5), 10 mM NaCl, 0.1% SDS, 100 mM DTT], precipitated with 4 volumes of cold acetone and Glycoblue (1:500, Life Technologies), reconstituted in Laemmli buffer, and boiled for 5 min.

MS and bioinformatic analysis

Biotinylated proteins or proteins isolated from adhesion plaques were separated on 4 to 12% gradient NuPAGE Novex bis-tris gel (Invitrogen) and digested in-gel with trypsin (Sequencing Grade Modified, Promega, V511C) (54), and the peptides were concentrated and desalted on StageTips (55). Peptides were analyzed by EASY-nLC system (Thermo Fisher Scientific) coupled on line to a linear trap quadrupole (LTQ)-Orbitrap Elite through a nanoelectrospray ion source (Thermo Fisher Scientific). Peptides were resolved in a 20-cm fused silica emitter (Thermo Fisher Scientific) packed in-house with reversed-phase ReproSil-Pur C18-AQ, 1.5-μm resin (Dr. Maisch GmbH) and eluted with a flow rate of 200 nl/min from 5% to 60% solvent (80% acetonitrile, 0.5% acetic acid in water) over 110 min. The full-scan MS spectra [from 300 to 1650 mass/charge ratio (m/z)] were acquired with a resolution of 120,000 at m/z 400 and target value of 1,000,000 charges in the Orbitrap. The top 10 most intense ions were sequentially isolated for fragmentation using high-energy collision dissociation (HCD) at the MS^3 target value of 40,000 charges and recorded in the Orbitrap with a resolution of 15,000 at m/z 400. All data were acquired with Xcalibur software (Thermo Fisher Scientific).

The MS RAW files were processed with MaxQuant software (56) version 1.3.8.2 and searched with Andromeda search engine (57) against the human UniProt database (37) (release-2013 05, 88,847 entries). To search parent mass and fragment ions, an initial mass deviation of 4 and 20 ppm, respectively, was required. Trypsin enzyme specificity with a maximum of two missed cleavages and peptides more than seven amino acids were allowed. Carbamidomethylation (Cys) was set as fixed modification. Oxidation (Met) and N-acetylation were considered as variable modifications. For identification of proteins and peptides, we required a maximum of 1% FDR. Scores were calculated in MaxQuant as described previously (56).

Supplementary materials

www.sciencesignaling.org/cgi/content/full/7/354/rs7/DC1
Fig. S1. Relationship between the iBAQ sum and the number of lysines.
Fig. S2. E-cadherin–specific adhesive interaction between cells and E-cadherin–coated glass.
Fig. S3. Colocalization of GFP-fusion proteins with E-cadherin.
Fig. S4. Network topology of E-cad-BirA* interactome.
Fig. S5. E-cad-BirA* interactome graphed according to protein functional annotation.
Fig. S6. EGTA-induced loss of cell adhesion and recovery.
Fig. S7. Cortical localization of cadhesome proteins in isolated cells.
Table S1. MS results of E-cadherin
Table S2. MS results of E-cadherin adhesion plaques from cells adhering to E-cadherin
Table S3. Protein interactions within E-cad-BirA* interactome predicted by PrePPI.
Table S4. Protein interactions within E-cad-BirA* interactome predicted by PrePPI.
Table S5. Protein domain analysis of the E-cad-BirA* interactome.

References and Notes


RESEARCH RESOURCE


Acknowledgments: We thank B. Burke (IMB, A-STAR, Singapore) for the BirA* plasmid, T. Carl Cornvik (Nanyang Technological University, Singapore) for access to the ORFeome and MGC libraries, A. Hedley (Cancer Research UK Beatson Institute) for informatics assistance, and J. Nelson (Stanford University, USA) for E-cadherin::GFP plasmid. Funding: This work was supported by the National Research Foundation Singapore under its NRF fellowship (NRF-RF2009-RF001-074) awarded to R.Z.-B., and Cancer Research UK. Author contributions: Z.G. designed and performed all the experiments and analyzed the data; L.J.N. performed MS; H.Z. performed localization experiments; P.S.M. performed bioinformatic analysis; S.Z. and R.Z.-B. conceived the work, designed experiments, analyzed the data, supervised the work, and wrote the paper. Competing interests: The authors declare that they have no competing interests. Data and materials availability: The MS files have been deposited to the ProteomeXchange Consortium (http://proteomecentral.proteomexchange.org/cgi/GetDataset) through the PRIDE partner repository. Data can be accessed using Accession: PXD000448; Username: reviewer92280@ebi.ac.uk; Password: 5iRRT6pA.

Submitted 12 May 2014
Accepted 12 November 2014
Final Publication 2 December 2014
10.1126/scisignal.2005473