Title
Cadherin evolution and the origin of animals

Author
Abedin, Monika

Publication Date
2010

Peer reviewed|Thesis/dissertation

Cadherin evolution and the origin of animals

By

Monika Abedin

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Molecular and Cell Biology

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Assistant Professor Nicole King, Chair
Associate Professor David Bilder
Associate Professor Steven E. Brenner
Professor Matthew D. Welch

Spring 2010

Cadherin evolution and the origin of animals

By

Monika Abedin

Abstract

Cadherin evolution and the origin of animals

by

Monika Abedin

Doctor of Philosophy in Molecular and Cell Biology

University of California, Berkeley

Professor Nicole King, Chair

The question of how animals evolved from a unicellular ancestor has challenged evolutionary biologists for decades. Because cell adhesion and signaling are required for multicellularity, understanding how these cellular processes evolved will provide key insights into the origin of animals. A critical finding is that choanoflagellates, the closest living unicellular relatives of animals, express members of the cadherin superfamily. Cadherins are pivotal for animal cell adhesion and signaling and were previously thought to be unique to animals, making them crucial to understanding the evolutionary origin and transition to multicellularity. Importantly, the presence of cadherins in choanoflagellates allows a consideration of their ancestral function in the unicellular progenitor of animals. To gain insight into the ancestral structure and function of cadherins, I reconstructed the domain content of cadherins from the last common ancestor of choanoflagellates and metazoans. Conservation of diverse protein domains in the choanoflagellate *Monosiga brevicollis* and metazoan cadherins suggests that ancestral cadherins served both signaling and adhesive functions. I find that two *M. brevicollis* cadherin containing an ancestral domain combination MBCDH1 and a close paralog, MBCDH2, localize to the actin-filled microvilli of the feeding collar. Interestingly, the protein abundance of these cadherins changes in response to bacterial food availability. *In vitro* studies, as well as experiments performed in a heterologous cell culture system, suggest that MBCDH1 does not mediate homophilic adhesion in the context of these experiments like some metazoan cadherins and thus may play a role in cell signaling. Taken together, these data suggest that cadherins may mediate interactions with the extracellular environment, including the recognition and capture of bacterial prey. I hypothesize that metazoan cadherins were co-opted to mediate cell-cell interactions from ancestral proteins that interpreted and responded to extracellular cues.

Table of Contents

Tables and Figures

Chapter 1: Origins and evolution of eukaryotic cell adhesion

SUMMARY

The multiple independent transitions to multicellularity that have occurred on the tree of life help explain differences in the morphology and cell physiology of macroscopic organisms. A prerequisite for the origin of multicellularity in each lineage was the evolution of stable cell adhesion. Reconstructing the evolution of cell junction proteins in diverse animal species and their unicellular relatives emphasizes the importance of both co-option and innovation during the evolution of multicellularity. The graded complexity of volvocine algae reveals their transition to multicellularity through the conversion of cell wall material into extracellular matrix. Furthermore, comparisons between animals and *Dictyostelium* reveal the extent to which the biology of their unicellular ancestors influenced their adhesive mechanisms. Understanding the unicellular ancestry of cell adhesion helps illuminate the basic cell biology of multicellular development in modern organisms.

INTRODUCTION

The structural integrity of multicellular organisms depends upon the establishment and maintenance of stable cell adhesion. In fact, multicellular life would not have been possible without the *de novo* evolution of mechanisms for attaching ancestrally solitary cells together. Multicellular organisms have evolved many times throughout the history of eukaryotes, each time from a distinct unicellular ancestor (1, 2). The independent origins of multicellularity can be viewed as repeated experiments that illuminate the role of historical contingency in the evolution of new morphologies (3). Comparing the biology of diverse unicellular and multicellular organisms reveals the genome content and cell biology of long-extinct unicellular ancestors and may help elucidate the cellular, molecular, and evolutionary foundations of modern multicellularity and development.

Separated at birth: multicellular lineages and their unicellular relatives

If the now extinct unicellular progenitors of multicellular organisms lived today, their study might reveal the molecular changes that accompanied the evolution of cell adhesion. Instead, one must look to extant unicellular organisms to identify genes that were potentially important for the evolution of their multicellular sisters. To this end, a clear understanding of the phylogenetic relationships between diverse unicellular and multicellular eukaryotes is critical.

An important pattern emerges when unicellular and multicellular lineages are mapped onto the tree of life. Multicellularity has arisen many times during life's history and the relic is that many multicellular lineages have unicellular relatives with similar cell biology. Comparisons of three pairs of lineages within this phylogenetic framework shed light on the diverse evolutionary history of cell adhesion mechanisms (Figure 1.1). Insights into genetic and cell biological changes that occurred during the origin and evolution of animal multicellularity have emerged through the analysis of diverse animal and their closes living relatives, the choanoflagellates (Figure 1.2) (4, 5). Choanoflagellates are unicellular and colony-forming microbial eukaryotes that inhabit both marine and freshwater environments throughout the globe. Second, comparisons made across multicellular lineages, between animals and dictyostelids (slime

molds) (Figure 1.1), identify common themes in cell adhesion mechanisms that evolved independently. Last, multicellular transitions in a class of green algae, the Volvocaceans, are illuminated by unicellular *Chlamydomonas* (Figure 1.1) as well as specific taxa within the Volvovaceans that represent transitional forms in multicellular evolution (Figure 1.4) (6). Together, these data provide a window into the events that took place during the evolution and diversification of multicellular forms.

Co-option and Innovation During Animal Cell Junction Evolution

One feature that distinguishes animals from all other multicellular lineages is the epithelium, a single layer of tightly packed, polarized cells held together by a distinct set of stable cell junctions (7). Each junction serves a different function within the epithelium and contains a unique collection of proteins. Together, these proteins provide adhesive and structural support to protect against mechanical stress, barrier function to block particles from crossing the epithelial layer, and communicating function allowing small molecules to pass between neighboring epithelial cells. Variations in the diversity and organization of junctions, as well as differences in junctional protein composition, contribute to the development and differentiation of discrete tissue types and perhaps played a role in the evolution of novel animal forms. By studying the phylogenetic distribution of junctions and their associated proteins, we can reconstruct the evolutionary history of cell junctions and illuminate the connection between epithelia and body plan evolution in animals. Here we analyze the distribution of epithelial junctions based on available ultrastructural, genomic and functional analyses (Figure 1.2).

The clear requirement for cell adhesion in multicellular organisms suggests that the metazoan ancestor must have possessed, at a minimum, rudimentary cellular connections. If so, what form did those connections take? Evidence from the most basal extant animals has shed light on this question. Adherens junctions, which physically tether adjacent cells to one another (8), are found in rudimentary form in sponge epithelia (9). From studies in later-branching lineages, we know that the molecules providing adherens junctions with their adhesive function are the classical cadherins. The presence of classical cadherins in sponges (10, 11) indicates that the molecular foundations for adherens juctions were in place in the last common ancestor of all animals (Figure 1.2). Future experiments to test whether cadherins localize to sponge junctional structures and provide adhesive function would provide important insights into the earliest stages of adherens junction evolution.

In addition to adhesion, epithelial tissues exhibit barrier properties that prevent passage of particles such as ions and other solutes across the epithelial cell layer. In invertebrates and some mammals, septate junctions provide occluding function and evidence suggests these junctions may extend back to sponges (12-17), raising the possibility that the barrier capacity of epithelia also evolved early in animal history. Addressing this prospect will require further characterization of putative septate junctions in sponges. More generally, sponge epithelia offer the opportunity to study the most basic functions of epithelia, those that were likely in place in the last common ancestor of animals.

The arrival of new, more complex animal morphologies was coincident with the appearance of pannexin-based gap junctions (characterized by the presence of pannexin proteins) that afford epithelial cells the ability to communicate through the passage of small molecules between cells

(18). The earliest branching organisms with pannexin-based gap junctions are cnidarians such as sea anemones and *Hydra*, which also have clearly identifiable adherens and septate junctions (19, 20). The emergence of gap junctions in organisms with tissue-level organization suggests that the evolution of intercellular communication in epithelia may have contributed to the elaboration of eumetazoan body plans during the Cambrian radiation.

The distribution of adherens, septate and pannexin-based gap junctions suggests that core functions of epithelia were in place before the evolution of Bilateria. Yet, not all bilaterian epithelia are alike. Chordates evolved new junctions that have surprisingly similar functions to those that already existed. Septate junctions appear to have been replaced in later-branching animals by tight junctions (21-23) that, until recently, were thought to have independent evolutionary origins. Although most of the molecules that comprise septate junctions are not shared with tight junctions, recent studies indicate that one primary component of tight junctions – claudins – serve barrier functions in septate junctions of *Drosophila* (fruit fly) epithelia (24-26). This indicates that tight and septate junctions may share a common ancestry or, alternatively, that the use of claudin-family members in *Drosophila* septate junctions and the tight junctions of later-branching lineages is convergent.

Chordates also evolved a second class of gap junction, made up of connexins rather than pannexins. Interestingly, unlike tight junctions, connexin and pannexin-based gap junctions may work alongside each other, as they are often expressed in the same tissues (20). Despite functional similarities, pannexin and connexin protein families have no recognizable sequence similarity, indicating that they are not evolutionarily related or that they have diverged beyond recognition.

The final major elaboration of epithelial junction diversity came in a new form of adhesive junction called the desmosome. Desmosomes, which are restricted to vertebrates, support epithelia against mechanical stress and, like adherens junctions, are composed of a specific type of cadherin (8). The absence of desmosomes in other animals implies that they might serve a function specific to vertebrate biology or simply that their emergence was coincident with the evolution vertebrates. The evolution of new junctions in later branching lineages may also have contributed to the evolution of highly specialized tissues (e.g. skeletal muscle, lung epithelium, mammary glandular tissue).

The characterization of junctions in animal provides important insights into animal epithelial evolution, but how can we understand the deepest roots of epithelia? An instructive example can be found in the study of cadherins in the closest known relatives of animals, the choanoflagellates (5, 27). Cadherins were long thought to be an animal innovation not found in other eukaryotes. However, genomic analysis of the choanoflagellate *Monosiga brevicollis* has uncovered an unexpected diversity of cadherins, on a par with some animals, despite its apparently simple, single-celled morphology (11). The presence of cadherins in a unicellular species of choanoflagellate, as well as the localization of two *M. brevicollis* cadherins to the choanoflagellate feeding structure, suggests that the ancestral functions of cadherins may have been for sensing and responding to extracellular cues. Furthermore, comparisons of cadherin protein domain composition in *M. brevicollis* and diverse animal cadherins suggest a link between ancestral cadherins and intracellular signal transduction, such as tyrosine kinase and

hedgehog signaling. Only after the divergence of the choanoflagellate and animal lineages were cadherins co-opted in animals to form stable junctions between epithelial cells.

The phylogenetic distribution of epithelial junctions across animals suggests that many of the major epithelial junctions evolved early in animal evolution. The epithelial cells of the metazoan ancestor likely possessed adhesive function and possibly barrier capabilities, with junction-mediated communication emerging soon after. As animals evolved, so too did their epithelia, acquiring new proteins and eventually new junctional types to fine-tune their epithelia and allow the development of more specialized tissues. In some cases the evolutionary path of junctional proteins appears to be one of convergence as is the case with pannexin and connexin that compose gap junctions and, possibly, septate and tight junctions. In other cases, such as that of the cadherins, the molecules related to junctional proteins evolved before the junction itself, suggesting that co-option to new functions may have been a driving force for animal evolution. Sponges, apparently lacking some of the well-defined junctions present in later-branching lineages may represent the ancestral state of epithelia and further study of sponge epithelia will yield valuable insights into basic epithelial biology. It is also possible that heretofore undiscovered cell junction types may exist in the less well-studied early-branching lineages. Continued research into the molecular foundations of cell biology in choanoflagellates will bring to light critical steps in animal evolution and a more complete picture of the origin and evolution of animal cell junctions will emerge.

Parallels in *Dictyostelium* and Animal Adhesion

Multicellular organisms typically develop in one of two ways, through division without cell separation or through cell aggregation (28). The first mode of multicellular development is exemplified by organisms like plants, animals and fungi while the second mode, a less common strategy among eukaryotes, is nicely illustrated by the dictyostelid slime molds. Dictyostelids exist as unicellular amoeboid cells that, when deprived of nutrients, initiate a morphogenetic program ultimately leading to the formation of a multicellular fruiting body that facilitates spore dispersal. Like animals, dictyostelids lack a cell wall, allowing their cells to change position within the multicellular organism. The absence of a cell wall also allows them to make dynamic changes in their adhesive properties during development through the regulated expression of diverse adhesion molecules. The parallels between animal and dictyostelid cell biology as well as the development of dictyostelid fruiting bodies through cell aggregation predict that dictyostelids can yield valuable insights into the evolution of cell adhesion.

To draw comparisons between animals and dictyostelids, an understanding of basic dictyostelid development is necessary. The best-studied dictyostelid, *Dictyostelium discoideum*, has a defined lifecycle consisting of both unicellular and multicellular states (Figure 1.3). In the vegetative state, *D. discoideum* remains unicellular; if nutrients become low and starvation ensues, the amoeboid unicells aggregate to form a motile multicellular slug. Once a nutrient-rich environment is found, the slug will develop into a fruiting body that consists of a long stalk bearing a spore-filled sac at its tip. The spatial and temporal regulation of diverse adhesion molecules occurs throughout development and correlates with dramatic morphological changes (Figure 1.3).

Cell adhesion mechanisms employed during *D. discoideum* morphogenesis have striking similarities with animal cell adhesion. Many cell-cell and cell-matrix adhesion genes have been characterized in *D. discoideum* and appear to be tightly regulated, becoming activated at specific stages of development (Figure 1.3). The calcium-dependent adhesion molecule DdCAD-1 provides one example that highlights the overlap between animal and *D. discoideum* adhesion. DdCAD-1 functions early in *D. discoideum* development to initiate cell contacts and has been proposed to play a role similar to that of animal cadherins, which are important for establishing stable cell junctions and also require calcium to function (29-31). In both animals and *D. discoideum*, expression of specific adhesion genes is frequently limited to a subset of cells within the multicellular organism and this spatial regulation is critical for proper development (32-34). During the slug stage, after cell contacts have been established, DdCAD-1 localizes to the periphery of the slug while other adhesion molecules are expressed in the interior (35). Furthermore, deletion of the DdCAD-1gene leads to defects in morphogenesis, delayed development and aberrant cell sorting indicating that, like many animal adhesion molecules, DdCAD-1 has important developmental functions (33, 36-38). In animals, the phosphorylation state of an adhesion protein often has dramatic affects on its adhesive properties (39). Similarly, ras-dependent dephosphorylation of DdCAD-1 increases DdCAD-1-mediated cohesion of cells (40). Despite the apparent connections between *D. discoideum* and animal cell adhesion during development, sequence comparisons suggest that *D. discoideum* adhesion molecules have few, if any, homologs in animals. Thus, it is not likely that their adhesion proteins are evolutionarily related. Instead, the commonalities between these two multicellular lineages may be the product of convergent evolution due to shared aspects of their cell biology.

Other animal-like features of *D. discoideum* cell adhesion not to be overlooked are structures resembling adherens junctions in the fruiting body (41). Not only do these structures have the appearance of adherens junctions in electron micrographs, but, like animal adherens junctions, they are actin-rich. Furthermore, a protein called aardvark localizes to the junctions and, independent from its junctional context, loss of aardvark alters gene expression suggesting that it has a cell signaling function. Aardvark belongs to the protein superfamily that also contains the animal protein β-catenin, a cytoplasmic component of animal adherens junctions that organizes the junction-associated actin cytoskeleton and has additional roles in cell signaling (42). The discovery of adherens-like junctions in the mature fruiting body provides yet another example of an adhesive feature common to both animals and *D. discoideum.*

Orchestration of *D. discoideum* development requires regulated cell adhesion and the mature organism possesses stable cell junctions associated with the actin cytoskeleten. The molecular basis for this adhesion comes from a diverse set of proteins expressed throughout morphogenesis, each with a specific function within the organisms. The adhesive properties of *D. discoideum* are remarkably similar to those of animals yet most *D. discoideum* adhesion molecules have little sequence similarity to animal proteins. This suggests that their cell adhesion molecules have distinct evolutionary origins due to the unique unicellular ancestors from which they evolved (Figure 1.1). Although the genomic foundation of dictyostelids is different from that of animals, the cell biological characteristics that they share, such as the absence of a cell wall and locomotive capabilities, may have lead to the convergence of adhesive mechanisms. An alternative explanation is that the cell adhesion proteins in animals and dictyostelids are homologous but have evolved beyond detection at the molecular level and have been lost in the

unicellular and multicellular organisms that branch between these two lineages. Continued investigation into the molecular underpinnings *D. discoideum* adhesion will expand our understanding of dictyostelid multicellularity and its relationship to animal biology.

Despite the apparent cell biological similarities between animals and dictyostelid there are also striking differences. In contrast to most animal species, dictyostelids display a relatively simple morphology. Why is it that animals are so much more complex than dictyostelids? The morphological disparity could stem from differences in the nature of multicellularity between the two lineages. The lifecycle of dictyostelids has both a unicellular and a multicellular stage, while animals are strictly multicellular. Perhaps maintaining the capacity to switch between unicellular and differentiated multicellular forms restricts the level of complexity that can be reached by the dictyostelids. Another possibility is that while animal cells within an organism are clonal, the cells that comprise a dictyostelid fruiting body can differ genetically. Building a multicellular organism requires intercellular cooperation, with some cells devoted to specific somatic functions while others are afforded the opportunity to transmit their genes. When cells are genetically different, this situation invites loss of cooperation and cheating among cells that could explain why organisms with aggregation-based development are less morphologicaly complex than those with clonal development (43). The differing levels of complexity between animals and dictyostelids suggests that, while certain cell biological features can influence key properties of multicellularity, like cell adhesion, many other factors impact the elaboration and evolution of multicellular morphology.

The Volvocine Algae: From Cell Walls to Extracellular Matrix

Many of the multicellular lineages present on earth today evolved from unicellular ancestors with rigid cell walls. Indeed, the cells of land plants, volvocine algae, fungi, and red and brown algae (Figure 1.1) are all bound by a cell wall. For these cells, adhesion is a passive process: physical connections are established as new cells form and the resulting attachments between cells are stabilized and maintained throughout life. This type of multicellular development, in which cells divide and remain linked by their shared cell wall, has important implications for the developing organism as the cells cannot reposition themselves after cytokinesis. Understanding the origins of cell wall-based connections may provide insights into the changes that facilitated the evolution of multicellularity.

Despite the many independent origins of multicellularity, the intermediate forms that reveal important steps in multicellular transitions have been largely wiped away by evolution and extinction. An exception to this is the volvocine algae, with diverse species reflecting everything from unicellular to colonial to multicellular morphology. The graded complexity within the volvocine algae provides a superb opportunity to explore the evolution of cell adhesion in organisms with cell walls (Figure 1.4).

The phylogenetic relationships among the volvocine algae provide a robust framework for examining the genomic and cell biological changes that occurred during their evolution (Figure 1.4). The group contains organisms ranging from the unicellular flagellate *Chlamydomonas reinhardtii* to the fully differentiated multicelluar *Volvox carteri* (*Volvox*), in which somatic cells and reproductive cells adhere to form a hollow sphere. Molecular phylogenetic studies indicate that *Volvox* and its close relatives (family Volvocaceae) form a monophyletic group that shares a

recent common ancestor with *Chlamydomonas* (44). Branching between *Volvox* and *Chlamydomonas* are *Gonium pectorale* (*Gonium*) and *Pandorina morum* (*Pandorina*), whose morphologies appear to represent intermediate forms, suggesting that multicellular *Volvox* may have evolved though progressive increases in size and complexity (Figure 1.4).

Characterizations of cell wall and extracellular matrix (ECM) components in the volvocine algae indicate that the conversion of cell wall structures to ECM and the subsequent expansion of this ECM was critical during the transition to multicellularity (45). The cell wall of unicellular *Chlamydomonas* can be broken into two main parts: the highly organized outer cell wall or 'tripartite layer', named after its striped appearance in electron micrographs, and the less structured inner cell wall (Figure 1.4) (46). The tripartite layer is found in all volvocine algae; in early-branching lineages it surrounds individual cells, whereas in later-branching lineages it surrounds entire colonies or entire multicellular organisms. In colonial *Gonium* the cell wall resembles that of *Chlamydomonas,* with each cell surrounded by tripartite layer yet connected by cytoplasmic bridges that are stabilized by specialized cell wall structures, potentially representing an early stage in the evolution of the cell wall into ECM (45). Unlike in Gonium, the tripartite layer in *Pandorina* and *Volvox* encases the whole multicellular organism and is important for holding the cells together once cytoplasmic connections have broken down (47).

The inner cell wall has also changed throughout volvocine evolution, appearing as an amorphous region below the outer cell wall of *Chlamydomonas* and growing into a voluminous ECM in which cells are embedded in *Pandorina* and *Volvox*. The dramatic change in inner cell wall volume has been so great that each *Volvox* cell has approximately 10,000 times as much ECM as a *Chlamydomonas* cell does (45). The morphological differences in cell wall organization among the volvocine algae may reflect its evolution from a simple protective layer into a scaffold that provides both protection and structural support for cells within colonial and multicellular contexts.

Molecular studies corroborate morphological evidence that the cell wall was transformed into a voluminous ECM during volvocine evolution and that co-option and diversification of cell wall protein families to ECM may have played an important part in the transition to multicellularity. Both the inner and outer cell wall of *Chlamydonomas* is comprised primarily of hydroxyproline-rich glycoproteins (HRGPs). Functional equivalency of the *Chlamydomonas* outer cell wall and the tripartite layer of colonial and multicellular volvocine species has been demonstrated experimentally (48) and the homology of cell wall proteins has been asserted due to high sequence similarity. For instance, homologs of the *Chlamydomonas* outer cell wall HRGP, GP2, have been identified in *Volvox* and *Pandorina* and shown to localize to the tripartite layer on (49). Additionally, the protein ISG (inversion specific glycoprotein), critical for ECM assembly during *Volvox* development (50), has a close relative in *Chlamydomonas* called VSP-3 that is found in the outer cell wall (51).

The HRGP constituents of the inner wall have been modified, diversified and specialized over time to serve as ECM in the later branching algae. Pherphorins are a class of HRGP found throughout the ECM of *Pandorina* and *Volvox* (52); pherphorin homologs have also been identified in *Gonium* and *Chlamydomonas* (53). The expression of some pherphorins changes in response to wounding and is also modulated by pheromones that trigger sexual development in

Volvox, indicating that pherphorins are a developmentally regulated class of proteins. Phylogenetic analyses of pherphorins suggest that the unicellular volvocine ancestor possessed a relatively diverse array of pherphorin proteins and the protein family has further diversified in *Volvox*, potentially adopting new developmental functions. Together, these data indicate that the early cell-wall-associated HRGPs of the basal volvocine algae evolved into a class of proteins with ECM function in colonial and multicellular lineages. Furthermore, comparison of HRGPs from diverse species suggests that gene duplication, divergence and domain shuffling played an important role in volvocine evolution (45, 51, 53).

The volvocine algae represent a well-characterized example of how a unicellular organism can evolve mechanisms of cellular attachment through the co-option of cell wall proteins to ECM. The expansion and diversification of the HRGPs correlates with the appearance of ECM and homologs of *Chlamydomonas* cell wall-associated HRGPs localize to ECM-rich structures in colonial and multicellular species, suggesting that this protein family played an important role in the evolution of multicellularity. The transformation of cell wall to ECM was a critical step in the evolution of multicellularity in Volvocaceae, yet it represents one of many ways that cells can evolve stable attachments. Investigation of cell adhesion machinery in other multicellular cell wall-bound organisms indicates that despite similarities in cell biology, adhesion mechanisms can differ greatly from one lineage to the next. For example, cells of land plants, which share a closer ancestry with the volvocine algae than most other multicellular lineages (Figure 1.1), adhere using pectin polysaccharides found throughout the cell wall and enriched in the middle lamellae and tricellular junctions (54). Fungi, whose cell wall composition differs greatly from plants, use entirely different molecules to maintain intercellular connections (55). These differences in cell adhesion reflect the unique evolutionary history of each multicellular lineage and suggest how the biology and genomic content of the unicellular ancestor exerts strong influences on the evolution of multicellular forms.

Contingency and chance in the evolution of cell adhesion

What explains the existence of dramatically different modes of cell adhesion in each of the different multicellular lineages? Part of the answer derives from the disparate cell biology and genome composition of the different unicellular progenitors of each lineage. For example, we can infer that the unicellular progenitor of animals was a heterotrophic flagellate that lacked a cell wall. Animal cells also lack cell walls, permitting them to adhere dynamically and reorganize into complex tissues and organs during development. In contrast, the last common ancestor of *Chlamydomonas* and *Volvox* was encased in a cell wall, a feature that provides structural integrity but prohibits cell rearrangement. The genome contents of the distinct progenitors of animals and *Volvox* may have also predisposed the two lineages to certain forms of multicellularity. Through comparative genomics and cell biology one can identify those genes that represented preadaptations for cell adhesion in each lineage (56). HRGP cell wall components in the unicellular progenitor were co-opted for use in the ECM of *Volvox*, just as cadherins in the unicellular progenitor of animals were co-opted to function as adhesion receptors in epithelia. An outstanding question is whether these molecules and those used in other lineages (e.g. *Dictyostelium*) were especially suited for mediating cell-cell interactions, or whether their recruitment to function in cell adhesion was primarily through chance (3, 57).

Understanding the unicellular ancestry of cell adhesion mechanisms will reveal fundamental aspects of cell biology in multicellular organisms and provide insights into the development and evolution of morphologically complex eukaryotes. Progress on this front will require expanded comparisons of genomes from diverse multicellular organisms paired with their closest single-celled relatives. With advances in genome sequencing technology, coupled with improvements in genetic manipulation in non-model organisms, we can elucidate critical genomic and functional changes that launched eukaryotic transitions to multicellularity. The approach proposed, while most directly relevant to the origin of cell adhesion, holds the promise of further illuminating the roles of historical contingency, biological constraint, and chance during evolution.

FIGURES

Figure 1.1. Diverse multicellular eukaryotes and their closest unicellular or colonial relatives. The phylogenetic relationships among select unicellular, colonial and multicellular eukaryotic lineages indicate that multicellularity evolved multiple times. Some lineages are strictly multicellular (filled circle) and some are unicellular or display simple/undifferentiated colonies (open circle), while others have a mix of unicellular or colonial and multicellular forms (half filled circle).

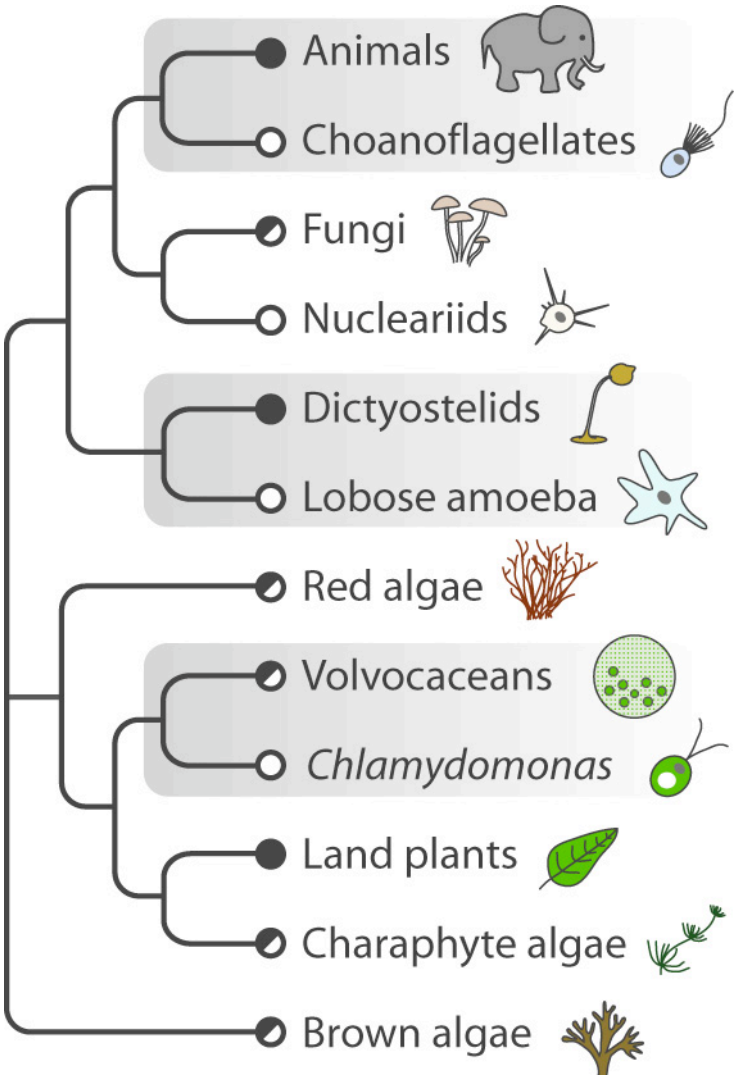
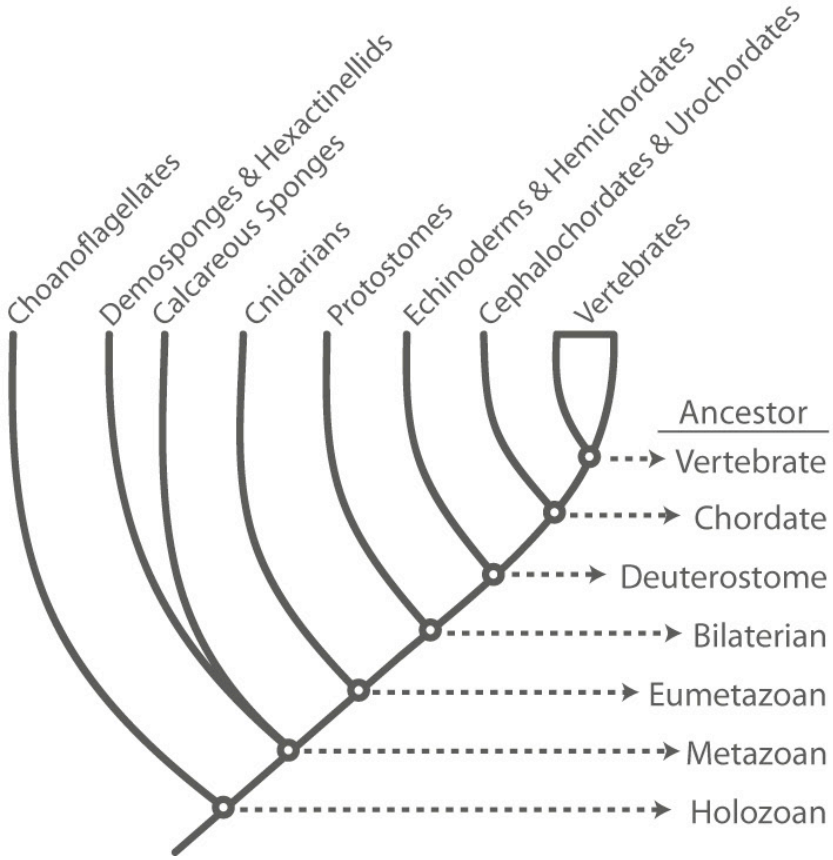


Figure 1.2. Phylogenetic distribution of epithelial junctions based on ultrastructural, genomic and functional data. Animals possess epithelial junctions with three main functions: adhesion, barrier and communication. Adherens, septate and pannexin-based gap junctions appeared relatively early while desmosomes, tight and connexin-based gap junctions emerged later in animal evolution. Presence (filled box) or absence (open box) in the genome of diagnostic genes ("Gene"), junctional ultrastructure ("Morph") and experimental support in extant lineages (e.g. stable cell adhesion, barrier function, or protein localization; "Expt"). For the inferred last common ancestors of different lineages, it is possible to reconstruct the nature of their cell junctions based on genomic ("Gene"), ultrastructural ("Morph") and experimental ("Expt") data from living organisms. A filled box in each column indicates: "Gene", diagnostic adhesion/junctional genes are present; "Morph", junction has been identified by electron microscopy; "Expt", adhesion/junctional proteins have been shown to localize to the junctional structure and/or are required for its functions. Numbered boxes: **1**, reports of desmosomes observed in non-vertebrates by electron microscopy are controversial (7); **2**, septate junctions have been identified only in mammals (12); **3**, evidence for septate junctional proteins in sponges is based solely on BLAST analysis of EST data (15); **4,** tight and connexin-based gap junctions are found in Urochordates but not Cephalochordates (18, 21-23); **5**, a single example of tight junctions has been reported in the spider central nervous system (58); **6**, pannexins are present in *Hydra* but absent from the *Nematostella* genome suggesting that they are not found in all cnidarians (20); **n/d**, not determined. References for figure data: (7-18, 20, 22, 59-63)



	Adhesion									Barrier						Communication					
	Adherens			Desmo-somal			Septate			Tight			Pannexin Gap			Connexin Gap					
Ancestor	Gene	Morph	Expt	Gene	Morph	Expt	Gene	Morph	Expt	Gene	Morph	Expt	Gene	Morph	Expt	Gene	Morph	Expt	Gene	Morph	Expt
Vertebrate				1				2	2												
Chordate										4	4	n/d				4	4	n/d			
Deuterostome																					
Bilaterian											5										
Eumetazoan			n/d				n/d		n/d				6		n/d						
Metazoan			n/d				3		n/d												
Holozoan																					

Figure 1.3. Cell adhesion proteins are dynamically expressed during *D. discoideum* development. Upon starvation, a complex developmental program is initiated involving temporal regulation of diverse cell-cell and cell- substrate adhesion protein. Colored bars indicate gene expression. The cell-cell adhesion protein DdCAD-1 is expressed when aggregations initiates and is followed by gp80 and then gp150. PsA expression is turned on at the mound stage and persists until fruiting body development. The cell-substrate adhesion protein SadA is activated during vegetative growth while SibA is constitutively expressed. Adapted from (34).

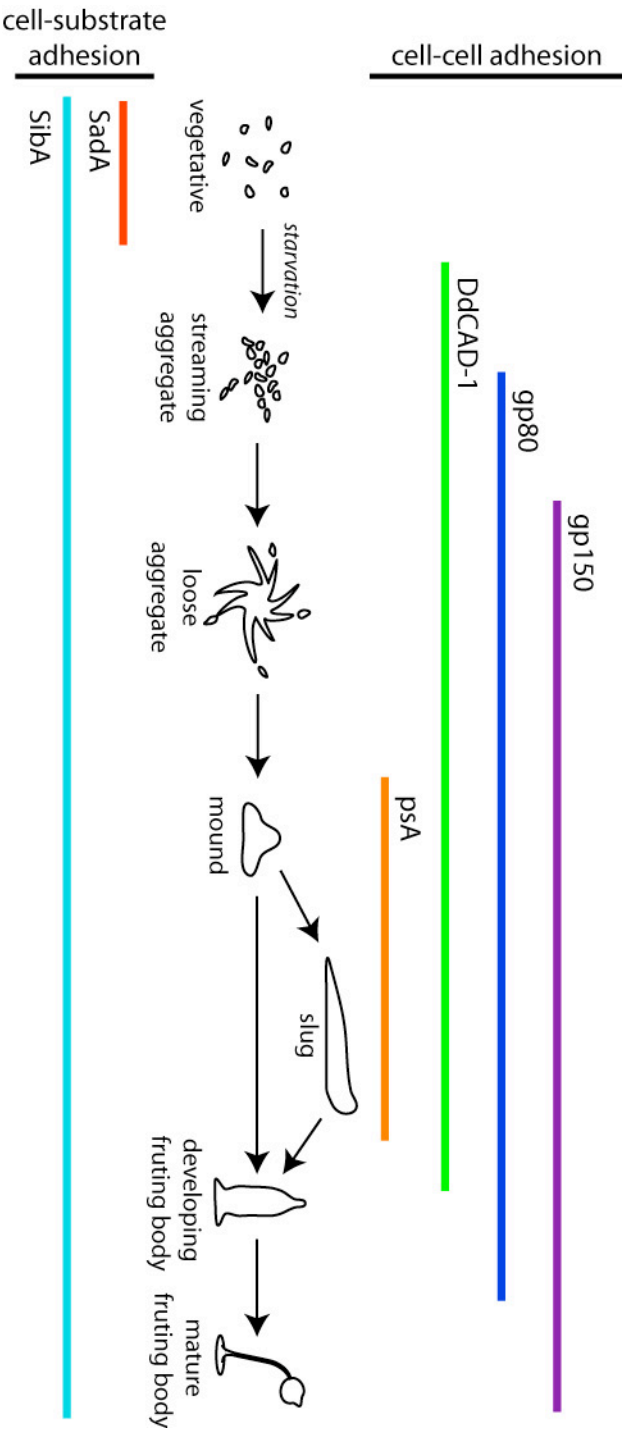
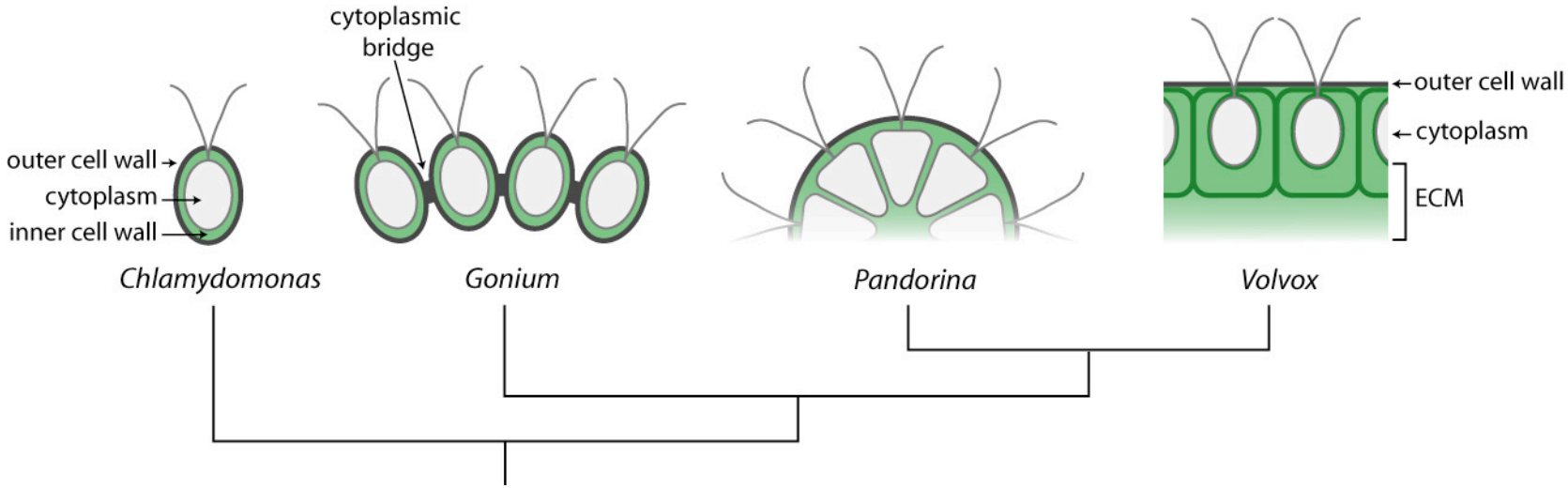

cell-substrate adhesion
cell-cell adhesion
SibA
SadA
DdCAD-1
gp80
gp150
psA
vegetative
starvation
streaming aggregate
loose aggregate
mound
slug
developing fruting body
mature fruting body

Figure 1.4. Evolution of cell wall into ECM in the volvocine algae. *Chlamydomonas,* possesses an HRGP rich cell wall consisting of an outer cell wall (dark grey) with a characteristic tripartite structure (3 layers) and a relatively amorphous inner cell wall (green) surrounding the cell membrane and cytoplasm (light grey). Colonial *Gonium* has a similar cell wall and utilizes cell wall components to stabilize cytoplasmic bridges between cells. In *Pandorina*, the outer cell wall encases the entire colony and the cells are embedded in ECM comprised of inner cell wall HRGP homologs. Multicellular *Volvox* has an extremely voluminous ECM derived from diverse HRGP family members that have specialized functions. Adapted from (45)


cytoplasmic bridge
outer cell wall
cytoplasm
inner cell wall
outer cell wall
cytoplasm
ECM
Chlamydomonas
Gonium
Pandorina
Volvox

Chapter 2: The premetazoan ancestry of cadherins

SUMMARY

Cadherin-mediated cell adhesion and signaling is essential for metazoan development (64, 65) and yet is absent from all other multicellular lineages. Because the evolution of metazoans from a single-celled ancestor required novel cell adhesion and signaling mechanisms (66-68), the discovery of cadherins in single-celled choanoflagellates, the closest known relatives of metazoans, potentially links cadherin evolution to metazoan origins. I find a surprising abundance of cadherin genes in the choanoflagellate *Monosiga brevicollis*, rivaling the number found in morphologically complex metazoans. Conservation of diverse protein domains in *M. brevicollis* and metazoan cadherins, including the Hedgehog N-terminal peptide and the SH2, Laminin G, Epidermal Growth Factor and Immunoglobulin domains, suggests that ancestral cadherins served both signaling and adhesive functions. The finding of cadherins in the strictly single-celled *M. brevicollis* raises questions about their functions in choanoflagellates and in the last common ancestor of metazoans. I show that two *M. brevicollis* cadherins containing features conserved from ancestral cadherins localize to the actin-filled microvilli of the feeding collar, suggesting that they may mediate interactions with the extracellular environment, including the recognition and capture of bacterial prey.

INTRODUCTION

The evolution of animals (metazoans) from their single-celled ancestors required genomic innovations that allowed cells to adhere and communicate (66-68). Of particular interest is the evolutionary history of cadherins, critical mediators of metazoan cell adhesion and signaling. Cadherins are expressed in diverse embryonic and adult tissues and provide the structural basis for vital developmental processes including tissue morphogenesis and maintenance, cell sorting and cell polarization. Despite the pivotal role filled by cadherins in the establishment and mediation of metazoan multicellularity they are apparently lacking from all non-metazoan multicellular organisms. In fact, the only non-metazoan taxon in which cadherins have been found are choanoflagellates, the sister group of metazoans (27, 69).

Since their discovery as cell adhesion proteins over 25 years ago, more than 100 cadherins family members have been identified in animals (70, 71). Most cadherins possess a region containing multiple extracellular cadherin repeats (ECs), a transmembrane domain, and an intracellular region with protein interaction domains. The number of ECs and the structure of the cytoplasmic region vary widely among different types of cadherins and some family members have additional protein domains that add to the range of cadherin architectures.

The diversity of cadherin structures is reflective of their diverse functions. The role of cadherins in epithelial biology, in which they physically tether epithelial cells together, is perhaps the best characterized. Evidence from both vertebrates and invertebrates has shown that homophilic adhesion between classic cadherins on neighboring cells forms adherens junctions. These junctions provide structural support to epithelial tissue (72, 73) and are important for establishing and maintaining apico-basal polarity (74-76). Classic cadherins also have critical functions in the nervous system. For example, N-cadherin localizes to synaptic junctions in the mouse brain

(77) and loss of N-cadherin function in developing flies and chicks leads to severe neural patterning defects (78, 79). More broadly, many classic cadherin and most protocadherins (which display weak adhesive properties) are expressed in specific patterns within the brain where they are involved in neurite outgrowth, axon guidance and synapse formation (64, 80, 81). Earlier in development, before neural specification, cadherins control the compaction of cells within the morula (82), orchestrate cellular rearrangements that occur during gastrulation (83) and, later in development, are involved in the formation of tissue boundaries (65, 84). Furthermore, cadherins mediate cell sorting that occurs throughout embryogenesis and is critical for morphogenesis [(65, 85, 86) and discussed further in Chapter 4].

The roles of cadherins described thus far are based, at least in part, on their adhesive properties (87). Another important role of cadherins is to activate and transduce intracellular signals that coordinate developmental processes. The atypical cadherins Fat and Dachsous (sometimes classified as protocadherins) are critical components of the Hippo signaling pathway, which regulates tissue growth, organ size and apoptosis (88). Based on studies in *Drosophila*, a working model has emerged in which Dachsous (ligand) acts downstream of morphogens to activate Fat (receptor) and initiate intracellular Hippo signal transduction causing restricted growth (89-95). Fat and Dachsous also mediate PCP signaling although, some details remain to be uncovered. A proposed model based on current data suggests that a gradient of Dachsous stimulates Fat to regulate the organization of cells within the plane of an epithelium (96-99). In a pathway thought to be separate from Fat and Dachsous, the seven-pass transmembrane cadherin Flamingo (also known as Starry Night) also functions in PCP signaling, possibly through homophilic interactions with Flamingo on adjacent cells (100-102). Recent studies suggest that the roles of these cadherins in signaling are at least partially conserved in mammals (98, 103, 104).

The fundamental roles of cadherins in animal multicellularity suggests that the study of choanoflagellate cadherins may illuminate the transition from single-celled organisms to multicellular animals. Choanoflagellates are unicellular and colony-forming microbial eukaryotes that inhabit both marine and freshwater environments and are ubiquitously distributed across the globe. As filter-feeding bacteriovores, choanoflagellates are important players in the ecological food web, providing a bridge between trophic levels by facilitating carbon flow to organisms higher on the global food chain. Their cellular architecture consists of an ovoid cell body and a collar of actin-filled microvilli that surrounds a single apical flagellum used to swim and capture bacterial prey. The observation that both the cell morphology and feeding strategy of choanoflagellates are nearly indistinguishable from those of feeding cells (choanocytes) in sponges (105, 106) first raised the possibility that choanoflagellates and animals are closely related. Indeed, phylogenetic analyses of nuclear and mitochondrial genes from diverse animals, choanoflagellates and other unicellular eukaryotes (107-110), as well as comparative studies of mitochondrial genome structure and sequence (111, 112) have revealed that choanoflagellates are among the closest relatives of metazoans. Furthermore, phylogenetic studies using large amounts of sequence data demonstrate that choanoflagellates are not derived from sponges or more recently branching metazoan phyla (5, 113).

Unlike sponges and other metazoans, most choanoflagellates exist primarily in a unicellular state. Thus, the common ancestor of choanoflagellates and metazoans was likely unicellular or capable of forming simple colonies. The unicellular life-history stage shared by all

choanoflagellates and their phylogenetic position has inspired their development as an experimental model for reconstructing features of the ancestor of metazoans. The completion of a whole genome sequence from the unicellular choanoflagellate *Monosiga brevicollis* (5) offers the first opportunity to consider the unicellular ancestry of cadherins and their potential contributions to metazoan origins. Here I show that *M. brevicollis* possesses a diverse array of cadherins, comparable to the abundance of cadherins in some animal species. Protein domain composition shared by *M. brevicollis* and animal cadherins suggest that ancestral cadherins may have had signaling and adhesive function. Colocalization of two *M. brevicollis* cadherins, MBCDH1 and MBCDH2, with polymerized actin raises the possibility that they associate with the actin-cytoskeleton like many animal cadherins. Additionally, localization of MBCDH1 and MBCDH2 to the feeding collar suggests a potential role for choanoflagellate cadherins in sensing environmental stimuli such as bacterial prey.

MATERIALS AND METHODS

Identification and domain annotation of cadherins

To identify cadherins in *M. brevicollis* I used the advanced search tool at the Joint Genome Institute (JGI) *Monosiga brevicollis v1.0* genome browser (114) to search for protein models that contain extracellular cadherin (EC) repeats. Both the key word "cadherin" and the EC repeat InterPro code, IPR002126, were used as search terms. To ensure that no EC repeat-containing proteins were missed due to incomplete gene predictions, I also searched the *M. brevicollis* genome with representative metazoan cadherin protein sequences using tblastn (115). EC repeat-containing proteins in the *N. vectensis* (116) and *C. intestinalis* (117) genomes, which also have JGI genome browsers, were identified in a similar manner to *M. brevicollis*. To search for EC repeats in the *M. musculus* genome, I used the "Genes and Markers Query Form" on the Mouse Genome Informatics webpage (118) and entered IPR002126 in the "Protein domains" field. Over 98% of the proteins identified were previously annotated as cadherins. To identify *D. melanogaster* cadherins I referred to previously published data that cataloged the name and domain structure of all the cadherins in the *D. melanogaster* genome (119). To obtain the complete sequence of each cadherin I searched FlyBase (119) using the protein name. To determine if the EC repeat-containing genes in the *M. brevicollis* genome were expressed, I utilized the EST data available on the Joint Genome Institute (JGI) *Monosiga brevicollis v1.0* genome browser. I interpreted EST support for any region of an EC-containing gene as evidence of expression (Table S3.1).

After identifying all potential cadherin proteins in the five genomes listed above, I analyzed each protein with SMART(120) and Pfam (121) to confirm the presence of ECs and to determine the overall domain structure. When the number of EC repeats predicted by the two programs differed, I cited the EC repeat architecture provided by SMART. I considered all other domains present if they were predicted by both SMART and Pfam. If only SMART or Pfam had the capacity to identify a specific domain, I considered it present if it was predicted by that annotation tool.

Choanoflagellate culture conditions

I cultured *M. brevicollis* at 25°C in natural seawater infused with cereal grass (5 g/ L) in 100 mm x 20 mm polystyrene dishes (Falcon). *Enterobacter aerogenes* or *Flavobacter sp.* were co-cultured with *M. brevicollis* and a grain of rice was added to each culture.

MBCDH1 antibody production

I generated rabbit anti-MBCDH1 polyclonal antibodies by PCR amplifying the region of MBCDH1 encoding EC1-EC3 (Fig. S2.3) from *M. brevicollis* cDNA using the following primers: forward CCGGAATTCTACACCTTCAACGTGACTGAGG and reverse CCGCTCGAGTTAGATGGTCTCAGGCTGCTCAAT. The forward primer contained an *EcoRI* restriction site (underlined) and the reverse primer contained a *XhoI* restriction site (underlined) as well as a stop codon. I cloned the synthesized DNA into pGEX-6P-1 (Amersham Biosciences, Piscataway, NJ) and pMAL-c2X (New England BioLabs, Ipswich, MA) expression vectors using the engineered restriction-sites to generate two fusion proteins: the MBCDH1 peptide N-terminally tagged with glutathione-S-transferase (GST) and the same peptide tagged N-terminally with maltose binding protein (MBP).

I transformed *Escherichia coli* BL21 cells to express each construct and purified the fusion proteins using affinity chromatography. Antibodies were raised in rabbits against GST-MBCDH1 antigen (Covance Research Products Inc., Princeton, NJ) and I affinity purified anti-MBCDH1 antibodies from the rabbit's serum using MBP-MBCDH1 fusion protein.

Immunofluorescence microscopy

I used indirect immunofluorescence to examine cadherin localization in *M. brevicollis* cells.

Staining after cell permeabilization

I fixed *M. brevicollis* cells that had grown to a density between 10^6 and 10^7 cells/mL by adding formaldehyde to a final concentration of 4%. I then applied approximately 0.7 mL of the fixed culture to poly-L-lysine coated coverslips and incubated for 30 minutes. After gently washing the coverslips 4 times with PEM (100 mM PIPES, pH 6.9, 1 mM EGTA, 0.1 mM MgSO4) I blocked and permeabilized the cells for 30 minutes with blocker (PEM/1% BSA/0.3% TritonX-100) and subsequently replaced the blocker with anti-MBCDH1 and E7 anti-β-tubulin primary antibodies (Developmental Studies Hybridoma Bank, Iowa City, IA). After incubating the cells with the antibodies for 1 hr, I washed the coverslips 4 times with blocker, applied Alexa Fluor 488 goat anti-rabbit IgG (H+L) and Alexa Fluor 350 goat anti-mouse IgG (H+L) (Molecular Probes, Carlsbad, CA) secondary antibodies and incubated for 1 hr in the dark, subsequently washing 4 times with PEM. To visualize F-actin, I incubated the cells with 6 U/ mL rhodamine phalloidin (Molecular Probes, Carlsbad, CA) diluted in PEM for 15 minutes in the dark. I then washed the coverslips 3 times with PEM and mounted them onto slides using 10 μl ProLong Gold antifade reagent (Molecular Probes, Carlsbad, CA). All steps were performed at room temperature and all antibodies were diluted in blocker.

Staining before cell permeabilization

I grew *M. brevicollis* cells to a density between 10^6 and 10^7 cells/ mL, pelleted the cells by spinning 1 mL at 500 xg for 5 minutes and resuspended them in PEM. I then added anti-MBCDH1 or both anti-MBCDH1 and E7 anti-β-tubulin primary antibodies to the cells and incubated for 30 minutes. After pelleting and washing the cells once with PEM, I fixed them in

4% formaldehyde in PEM for 30 minutes. After fixation, I pelleted the cells, resuspended them in PEM and applied them to poly-L-lysine coated coverslips for 30 minutes. I washed the cells 4 times and blocked and permeabilized for 30 minutes in blocker. I performed the incubations with the secondary antibodies and rhodamine phalloidin staining as described above.

Image collection
I captured all images using a Leica DMI6000 B inverted compound microscope and Leica DFC350 FX camera. I collected between 25 and 50 optical sections of each field along the Z-axis at 100X magnification using oil immersion. I deconvolved the Z stacks using Image-Pro AMS 3D blind deconvolution software and flattened two to five sections from roughly the center of the stack to produce the final images.

RESULTS

Given the absence of overt cell adhesion in *M. brevicollis,* one might expect choanoflagellates to have significantly fewer and less diverse cadherin genes than metazoans, perhaps reflective of the ancestral condition. In contrast, I find that both the absolute and relative abundances of *M. brevicollis* cadherins are comparable to those of diverse metazoan genomes (Table 2.1). I identified a total of 23 putative cadherin genes (representing 0.25% of the gene catalog), over half (13/23) of which are expressed during logarithmic growth in culture [based on EST sequencing (114)]. To establish benchmarks for metazoan cadherin gene abundance I performed exhaustive searches for cadherins in genomes of four representative metazoan species: *Nematostella vectensis* (Phylum Cnidaria), *Drosophila melanogaster* (Phylum Arthropoda), *Ciona intestinalis* (Phylum Chordata) and *Mus musculus* (Phylum Chordata). The total number of cadherin genes in metazoan genomes ranges from 17 in *D. melanogaster* to 127 in *M. musculus* (Table 2.1), representing between 0.12% and 0.39% of the gene catalog. The unexpected abundance of cadherins in *M. brevicollis* as compared with metazoans suggests that they serve a similar diversity of functions despite a lesser degree of morphological complexity.

M. brevicollis cadherins differ from most metazoan cadherins in that they tend to contain a greater number and diversity of protein domains. For example, *M. brevicollis* cadherins on average have more EC repeats, the domain that defines cadherins, per protein than do metazoan cadherins (Table 2.1). Unlike it's more recently branching metazoan relatives, cadherins from *N. vectensis* also contain a relative enrichment of EC repeats (Table 2.1), suggesting that choanoflagellate and cnidarian cadherins more closely reflect the ancestral condition. The comparatively limited diversity in EC repeat number of cadherins from bilaterally symmetric metazoans may reflect a shift in function from long-range interactions to short-range adhesion over the course of metazoan evolution.

Protein domains linked to EC repeats in *M. brevicollis* and metazoans were likely present in cadherins from their common ancestor. I find that cadherins containing epidermal growth factor (EGF), Laminin G (LamG) and transmembrane domains are present in *M. brevicollis* and metazoans as diverse as sponges, sea urchins and humans (MBCDH21; Fig. 2.1A and B), suggesting that the physical linkage of these domains with EC repeats is ubiquitous and predates the origin of metazoans. In bilaterians, cadherins possessing these four linked domains are classified as Fat cadherins. Among Fat cadherins, the serial arrangement and number of EGF

and LamG domains varies, despite the fact that they all likely share a common heritage (122). Likewise, the position of EGF and LamG domains in *M. brevicollis* MBCDH21 differs from metazoan Fat cadherins. While the specific role of EGF and LamG domains in the context of cadherins is unclear, the conserved linkage of these domains to EC repeats among metazoan phyla, and presumably in their last common ancestor, suggests a conservation of core function that warrants further investigation.

M. brevicollis, the sponge *Amphimedon queenslandica*, and *N. vectensis* (123) share cadherins with EC repeats linked to the Src Homology 2 (SH2), Hedgehog N-terminal peptide (N-hh), Immunoglobulin (Ig), von Willebrand type A domains and Ig domains, suggesting that cadherins containing these domains were ancestral to metazoans. The presence of SH2 and N-hh domains in cadherins from choanoflagellate and cnidarian lineages is particularly interesting because it hints at ancient linkages between cadherin function and intercellular signaling. Because SH2 domains bind sites of tyrosine phosphorylation, choanoflagellate and cnidarian cadherins containing cytoplasmic SH2 domains (MBCDH1 and NvHedgling; Fig. 2.1A and C) could connect extracellular cues to intracellular processes such as cell cycle regulation and cellular metabolism. The presence of a protein tyrosine phosphatase domain in two *M. brevicollis* cadherin (MBCDH21 and MBCDH7, Fig. 1 A and B and S1) provides further evidence of a connection between choanoflagellate cadherins and tyrosine kinase signaling.

Likewise, the presence of the N-hh domain at the amino termini of cadherins from choanoflagellates and cnidarians (MBCDH11 and NvHedgling; Fig. 1A and C) suggests an ancestral connection between cadherin function and developmentally important hedgehog signaling components. *M. brevicollis* and a second choanoflagellate species, *Monosiga ovata*, also contain the HINT domain characteristic of the C-terminus of metazoan hedgehog proteins (5, 124). Thus, although the choanoflagellate genome lacks a canonical hedgehog gene, it encodes each of the essential protein modules suggesting that protein domain shuffling contributed to the evolution of metazoan hedgehogs (122). Interestingly, although the *M. brevicollis* genome lacks a hedgehog gene, it does contain a clear homolog of Patched (5), the receptor that binds hedgehog.

The linkage of EC repeats to SH2 and N-hh domains in *M. brevicollis* and *N. vectensis* is absent from bilaterians, suggesting that this ancient protein architecture was lost relatively early in metazoan evolution. In contrast, the connection between metazoan cadherins and β-catenin, an important player in the Wnt pathway (42, 125, 126), likely evolved after the divergence of choanoflagellate and metazoan lineages, and represents a metazoan innovation. Metazoan classical cadherins contain a highly conserved classical cadherin cytoplasmic domain (CCD) that has binding sites for β-catenin, a regulator of cadherin-mediated adhesion (127). *N. vectensis* possesses 5 CCDs (Table 2.1, Fig. S2.2), indicating that this domain evolved before the origin of Bilatieria. Indeed, recent EST and genome sequence data from a sponge, *Amphimedon queenslandica,* reveal that cadherin CCDs may have more ancient roots, dating back to the earliest ancestors of metazoans. Onur *et al*. (2007) (10) identified an *A. queenslandica* cadherin that, upon further analysis, we find contains a cadherin CCD domain (Fig 2.1B). Its domain structure suggests orthology with *D. melanogaster* N-cadherin (78), whereas BLAST (115) analysis indicates that it shares the greatest percent sequence identity with Fat cadherins. In contrast to metazoans, the *M. brevicollis* genome lacks both β-catenin and the CCD (Table 2.1),

indicating that this domain evolved after the origin of Metazoa or was lost within the choanoflagellate lineage.

M. brevicollis leads a unicellular lifestyle and is not known to form stable cell-cell contacts. Therefore, the biological processes mediated by choanoflagellate cadherins remain enigmatic. The subcellular localization of MBCDH1 and MBCDH2, two nearly identical cadherins with a domain content resembling that of inferred ancestral cadherins [including six EC repeats, a von Willebrand type D domain and an EGF domain in the extracellular region, and a single intracellular SH2 domain (Fig. 2.1.C)], may provide insight into potential functions of choanoflagellate cadherins. Antibodies raised against an extracellular portion of MBCDH1 with 95% sequence identity to MBCDH2 recognize a single protein band of the correct predicted size (~192 kDa) and this signal is competed away following pre-incubation of the antibody with the soluble form of the MBCDH1 epitope, indicating that the antibody is specific (Fig. S2.4).

MBCDH1 localizes to four regions of the choanoflagellate cell: the apical collar of actin-filled microvilli, the basal pole of the cell, an unidentified structure at the apical end of the cell, and puncta within the cell body (Fig. 2.2 B-K). The subcellular distribution of extracellular MBCDH1 relative to intracellular MBCDH1 is revealed by probing cells with the antibody either before or after cell membrane permeablization. Extracellular MBCDH1 colocalizes with polymerized actin, most strikingly in the apical collar and to a lesser extent at a focus near the basal end of the cell (Fig. 2.2 G, H, J). In metazoans, classical cadherins interact indirectly with the underlying actin cytoskeleton to promote proper microfilament assembly and organization (128-130). Although further biochemical investigation is necessary, the colocalization of actin filaments and MBCDH1 in *M. brevicollis* raises the possibility that the connection between cadherins and actin filaments predates the diversification of choanoflagellates and metazoans. Metazoan SH2 domain containing proteins such as Nck are known to interact with actin-binding proteins through phosphorylated tyrosines (131, 132) suggesting that the cytoplasmic tail of MBCDH1 could interact indirectly with the underlying actin cytoskeleton.

The intracellular localization of MBCDH1 also reveals parallels with metazoan cadherins. The punctate MBCDH1 localization and its enrichment at the apical end of the cell body (Fig. 2.2 B) are likely to represent vesicles and the Golgi apparatus, respectively, based on their intracellular location and by reference to known ultrastructural landmarks in choanoflagellates (133). This would be consistent with patterns of intracellular trafficking of metazoan cadherins during the establishment of cell contacts (134-136).

DISCUSSION

On the basis of shared domain content of choanoflagellate and metazoan cadherins (Fig. 2.1), I infer that cadherins in the last common ancestor of choanoflagellates and metazoans possessed, among others, SH2, N-hh, LamG, EGF, Ig and transmembrane domains. The functions of these domains suggest that some choanoflagellate cadherins may mediate intracellular signaling. Specifically, the SH2 domain has the potential to interact with targets of tyrosine kinase phosphorylation. Interestingly, the receptor tyrosine kinase pathway is one of the few metazoan-type signaling networks that is well represented in choanoflagellates (5, 27, 137, 138). Moreover, in metazoans, cytoplasmic and receptor tyrosine kinases are commonly associated

with adherens junctions, sites of abundant tyrosine phosphorylation (139-141). Thus, I propose that the interaction between cadherins and the tyrosine kinase signaling pathway evolved before the origin of metazoans.

In metazoan epithelial cells, recruitment of β-catenin facilitates interactions between classical cadherins and the actin cytoskeleton, an association that is essential for the establishment and maintenance of cell shape and cell polarity (142, 143). The finding that *M. brevicollis* MBCDH1 and MBCDH2 colocalize with microfilaments comes as a surprise considering that it does not have the CCD domain characteristic of metazoan classical cadherins and given that the *M. brevicollis* genome lacks a β-catenin ortholog. If it is determined that *M. brevicollis* cadherins associate with the local actin cytoskeleton, this interaction may contribute to the striking cell morphology and polarity that choanoflagellates exhibit.

The abundance, diversity and high proportion of expressed cadherins in the *M. brevicollis* genome indicate that these proteins play important roles in choanoflagellate biology and raise questions about their function in a unicellular organism. Like the tip-link interactions between human cadherin 23 and protocadherin 15 that bundle microvilli-like stereocilia in the inner ear (144), choanoflagellate cadherins may stabilize the shape and integrity of the feeding collar by mediating interactions among neighboring microvilli. In addition, choanoflagellate cadherins may contribute to transient cell-cell interactions, such as cell recognition, or mediate adherence to a substrate. Finally, in metazoans, E-cadherins and Flamingo cadherins are exploited as extracellular tethers by pathogenic bacteria during host cell invasion (145-147), demonstrating that cadherins have the capacity to interact with bacteria. It is possible that choanoflagellate cadherins fill an equivalent role in binding bacterial prey for recognition or capture, functions consistent with the enrichment of MBCDH1 on the feeding collar (Fig. 2.2). If ancient cadherins bound to bacteria in the unicellular progenitor of choanoflagellates and metazoans, cadherin-mediated cell adhesion in metazoans may reflect the co-option of a class of proteins whose earliest function was to interpret and respond to cues from the extracellular milieu. Indeed, the transition to multicellularity likely rests on the co-option of diverse transmembrane and secreted proteins to new functions in intercellular signaling and adhesion. Using the *M. brevicollis* model system to distinguish between these hypotheses will provide novel insights into the ancestral function of this important protein family and its potential contribution to the origin of metazoans.

FIGURES AND TABLES

Table 2.1. Cadherin abundance and EC repeat content in *M. brevicollis* and diverse eukaryotes. *Atha, Arabidopsis thaliana; Ddis, Dictyostelium discoideum; Scer, Saccharomyces cerevisiae; Mbre, Monosiga brevicollis; Nvec, Nematostella vectensis; Dmel, Drosophila melanogaster; Cint, Ciona intestinalis; Mmus, Mus musculus.*

	Plant	Slime mold	Fungus	Choanoflagellate			Chordata	
						Bilateria		
					METAZOA			
	Atha	*Ddis*	*Scer*	*Mbre*	*Nvec*	*Dmel*	*Cint*	*Mmus*
Genes/ genome	27,273	13,607	6,609	9,196	18,000	13,601	14,182	32,661
Cadherins/ genome	0	0	0	23	46	17	32	127
Normalized Cadherin abundance*	0	0	0	0.25%	0.26%	0.12%	0.23%	0.39%
EC repeats/ cadherin (avg)	N/A	N/A	N/A	14.7	11	12.2	6.2	5.2

*Percent EC repeat-encoding genes in draft gene catalog.

Figure 1.1. Phylogenetic reconstruction of ancestral cadherin domain composition. (**A**) Venn diagram analysis of domains linked to EC repeats in *M. brevicollis*, *N. vectensis* and *M. musculus* cadherins. (**B**) Representative relationships of Fat-related cadherins from *M. brevicollis* and diverse metazoan taxa. The cladogram depicts currently accepted relationships among metazoan phyla (148, 149)**.** Green boxes highlight clusters of EGF and LamG domains and EC repeats are shown in blue. (**C**) Protein domains shared by *M. brevicollis* MBCDH1, MBCDH10 and MBCDH11, the sponge cadherin AmqHedgling and the cnidarian cadherin NvHedgling are absent from *M. musculus* cadherins. Blue boxes contain SH2 domains, yellow boxes contain Ig domains and red boxes contain N-hh and VWA domains. (**D**) Symbols used in B and C. LCA: last common ancestor. LamN and LamG domains in MBCDH21 are below the SMART e-value threshold but above the Pfam threshold (120, 121). See Fig. S1 for the complete domain structure of MBCDH21, Table S2 for protein identifiers and species names and Table S3 for domain abbreviations.

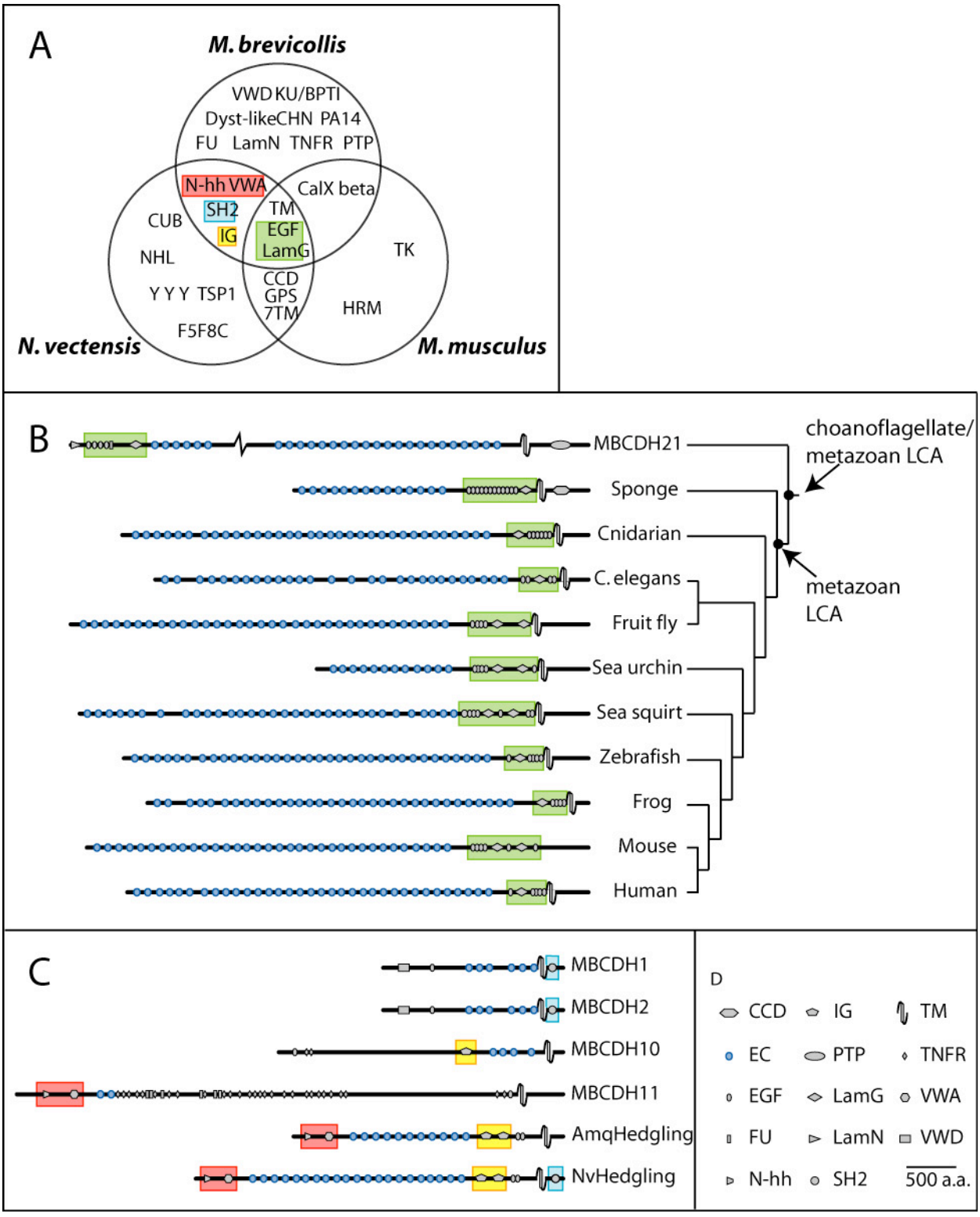
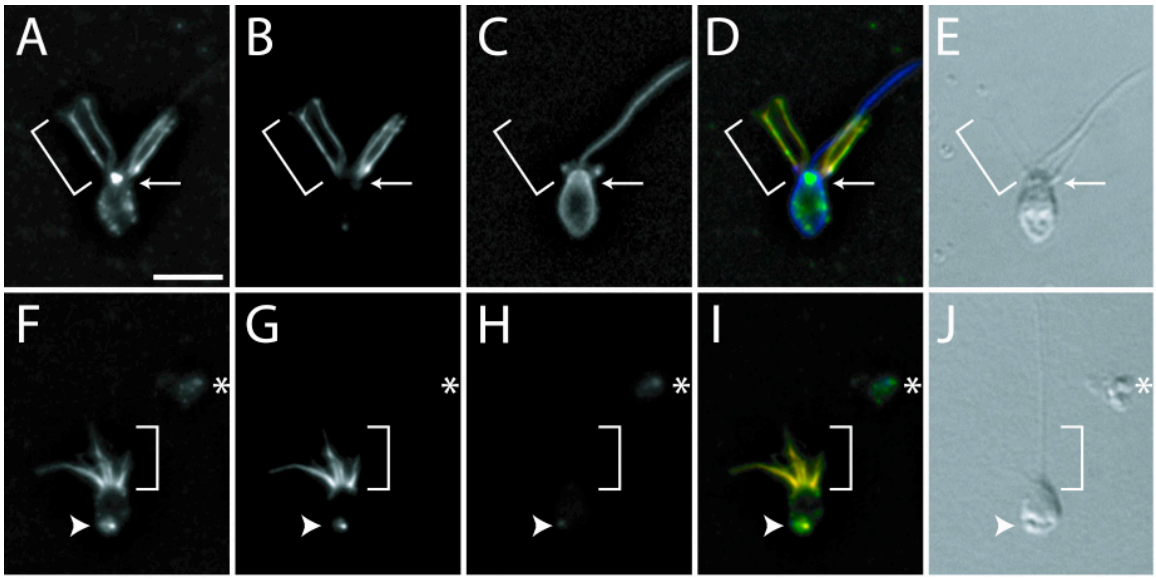

A
M. brevicollis
VWD KU/BPTI
Dyst-likeCHN PA14
FU LamN TNFR PTP
N-hh VWA
CalX beta
SH2
TM
CUB
IG
EGF
LamG
NHL
TK
CCD
GPS
7TM
Y Y Y TSP1
HRM
F5F8C
N. vectensis
M. musculus
B
MBCDH21
Sponge
Cnidarian
C. elegans
Fruit fly
Sea urchin
Sea squirt
Zebrafish
Frog
Mouse
Human
choanoflagellate/
metazoan LCA
metazoan
LCA
C
MBCDH1
MBCDH2
MBCDH10
MBCDH11
AmqHedgling
NvHedgling
D
CCD
IG
TM
EC
PTP
TNFR
EGF
LamG
VWA
FU
LamN
VWD
N-hh
SH2
500 a.a.

Figure 2.2. Localization of MBCDH1 and MBCDH2 to the actin-filled microvilli and cell body of *M. brevicollis*. Subcellular localization of MBCDH1 and MBCDH2 (**A, F**), compared to polymerized actin stained with rhodamine-phalloidin (**B, G**), or antibodies against β-tubulin (**C, H**). Cells were exposed to antibodies against MBCDH1 after (**A-E**) or before (**F-J**) permeabilization. Overlay of MBCDH1 and MBCDH2 (green), actin (red) and β-tubulin (blue) reveals colocalization of MBCDH1 and MBCDH2 with actin (yellow) on the collar and at the basal pole (**D, I**). Differential interference contrast microscopy shows cell morphology (**E, J**). Brackets: collar of microvilli, arrowhead: apical organelle, *: cluster of autofluorescent bacterial detritus. Scale bar: 5 microns.



SUPPLAMENTAL FIGURES AND TABLES

Table S2.1. *M. brevicollis* cadherin expression during logarithmic growth.

Gene	Transcribed?
MBCDH1	Yes
MBCDH2	Yes
MBCDH3	Yes
MBCDH4	No
MBCDH5	Yes
MBCDH6	No
MBCDH7	Yes
MBCDH8	Yes
MBCDH9	Yes
MBCDH10	Yes
MBCDH11	Yes
MBCDH12	No
MBCDH13	No
MBCDH14	Yes
MBCDH15	Yes
MBCDH16	No
MBCDH17	No
MBCDH18	No
MBCDH19	No
MBCDH20	Yes
MBCDH21	No
MBCDH22	Yes
MBCDH23	No

Table S2.2. Cadherin protein identifiers for Figure 2.1.

Protein	Species	Accession Number
Panel B		
C. elegans	*Caenorhabditis elegans*	NP_497917
Cnidarian	*Nematostella vectensis*	XP_001633830
Frog	*Xenopus tropicalis*	* (150)
Fruit fly	*Drosophila melanogaster*	NP_477497
Human	*Homo sapien*	NP_005236
MBCDH21	*Monosiga brevicollis*	* (114)
Mouse	*Mus musculus*	ABB88946
Sea squirt	*Ciona intestinalis*	* (117)
Sea urchin	*Strongylocentrotus purpuratus*	XP_785601
Sponge	*Amphimedon queenslandica*	* (10)
Zebrafish	*Danio rerio*	NP_998132
Panel C		
MBCDH1	*Monosiga brevicollis*	* (150)
MBCDH10	*Monosiga brevicollis*	* (150)
MBCDH11	*Monosiga brevicollis*	* (150)
AmqHedgling	*Amphimedon queenslandica*	* (123)
NvHedgling	*Nematostella vectensis*	* * (123)

*See Note S2 for mRNA and protein sequences in FASTA format.

Table S2.3. Domain abbreviations for Figure 2.1.

Abbreviation	Domain
7TM	seven-pass transmembrane domain
CalXbeta	calcium exchange beta
CCD	classical cadherin cytoplasmic domain
CHN	cohesion domain
CUB	complement C1r/C1s, Uegf, Bmp1
NHL	NCL-1, HT2A and LIN-41
Dyst-like	dystroglycan-like cadherin-type
EGF	epidermal growth factor
F5F8C	F5/F8 type C
FU	furin-like
GPS	G-protein-coupled receptor proteolytic site domain
HRM	hormone receptor domain
IG	immunoglobulin domain
PTP	protein tyrosine phosphatase
KU/BPTI	Kunitz/ Bovine pancreatic trypsin inhibitor
LamG	Laminin G
LamN	Laminin N-terminal
N-hh	hedgehog N-terminal peptide
PA14	PA14 domain
SH2	src homology type 2
TK	tyrosine kinase
TM	transmembrane
TNFR	tumor necrosis factor receptor
TSP1	Thrombospondin type 1 repeats
VWA	von Willebrand type A
VWD	von Willebrand type D
Y Y Y	Y Y Y domain

Table S2.4. Protein identifiers for *M. brevicollis* cadherins in Figure S2.1.

Protein*	Scaffold number and position
MBCDH1	scaffold 3:2082322-2088244
MBCDH2	scaffold 24:59573-66261
MBCDH3	scaffold 16:685938-691335
MBCDH4	scaffold 17:320979-326244
MBCDH5	scaffold 17:673239-690723
MBCDH6	scaffold 18:478917-482854
MBCDH7	scaffold 2:2797908-2819316
MBCDH8	scaffold 18:434817-475101
MBCDH9	scaffold 2:2884545-2896706
MBCDH10	scaffold 20:676972-685860
MBCDH11	scaffold 28:283225-288861
MBCDH12	scaffold 28:283225-288861
MBCDH13	scaffold 28:513136-535020
MBCDH14	scaffold 3:670125-676643
MBCDH15	scaffold 30:223040-230280
MBCDH16	scaffold 31:130217-148856
MBCDH17	scaffold 12:1046678-1061067
MBCDH18	scaffold 36:114574-145715
MBCDH19	scaffold 36:166826-182718
MBCDH20	scaffold 40:32342-44590
MBCDH21	scaffold 48:60640-91459
MBCDH22	scaffold 54:79290-93779
MBCDH23	scaffold 9:1252738-1266382

Table S2.5. Protein identifiers for *N. vectensis* cadherins in Figure S2.2.

Scaffold number and position	Scaffold accession number
scaffold 77:798400-816589	DS469583
scaffold 100:445997-482773	DS469606
scaffold 27:101436-116117	DS469533
scaffold 193:332198-350335	DS469699
scaffold 71:633314-642612	DS469577
scaffold 416:43332-58684	DS469922
scaffold 14475:461-786	DS479215
scaffold 4224:534-8006	DS473521
scaffold 236:331404-343949	DS469742
scaffold 23:25977-55968	DS469529
scaffold 212:49834-114843	DS469718
scaffold 270:155593-158039	DS469776
scaffold 12488:115-1286	DS478634
scaffold 21:857978-903769	DS469527
scaffold 48:93827-108841	DS469554
scaffold 85:312269-324474	DS469591
scaffold 157:282-12640	DS469663
scaffold 236:344187-353204	DS469742
scaffold 270:170452-172182	DS469776
scaffold 270:159645-167352	DS469776
scaffold 2646:2417-3403	DS472047
scaffold 6181:413-6601	DS475139
scaffold 77:757328-786001	DS469583
scaffold 77:843092-896167	DS469583
scaffold 212:116301-139891	DS469718
scaffold 416:106486-127211	DS469922
scaffold 557:85479-94927	DS470063
scaffold 851:13-19643	DS470357
scaffold 2546:395-6927	DS471948
scaffold 2837:370-1336	DS472225
scaffold 9093:64-889	DS477286
scaffold 14604:438-743	DS479256
scaffold 7:818956-823802	DS469513
scaffold 21:828149-846115	DS469527
scaffold 77:817743-833487	DS469583
scaffold 131:198546-214026	DS469637
scaffold 557:14863-50158	DS470063
scaffold 416:10481-17409	DS469922
scaffold 21:909299-912326	DS469527
scaffold 86:481944-489013	DS469592
scaffold 21:767716-804426	DS469527
scaffold 104:381241-412272	DS469610
scaffold 5:2420416-2446071	DS469511
scaffold 67:468550-490914	DS469573
scaffold 318:249865-260879	DS469824
scaffold 9:2048365-2076242	DS469515
scaffold 280:47108-49824	DS469786

Figure S2.1. Complete set of *M. brevicollis* cadherins. Schematic diagram of the domain organization of all 23 *M. brevicollis* cadherins shows diversity in size, domain content and domain arrangement. For clarity, signal peptides were omitted from protein diagrams. Laminin N-terminal and Laminin G domains in MBCDH21 are below the SMART e-value threshold but above the Pfam threshold (120, 121).

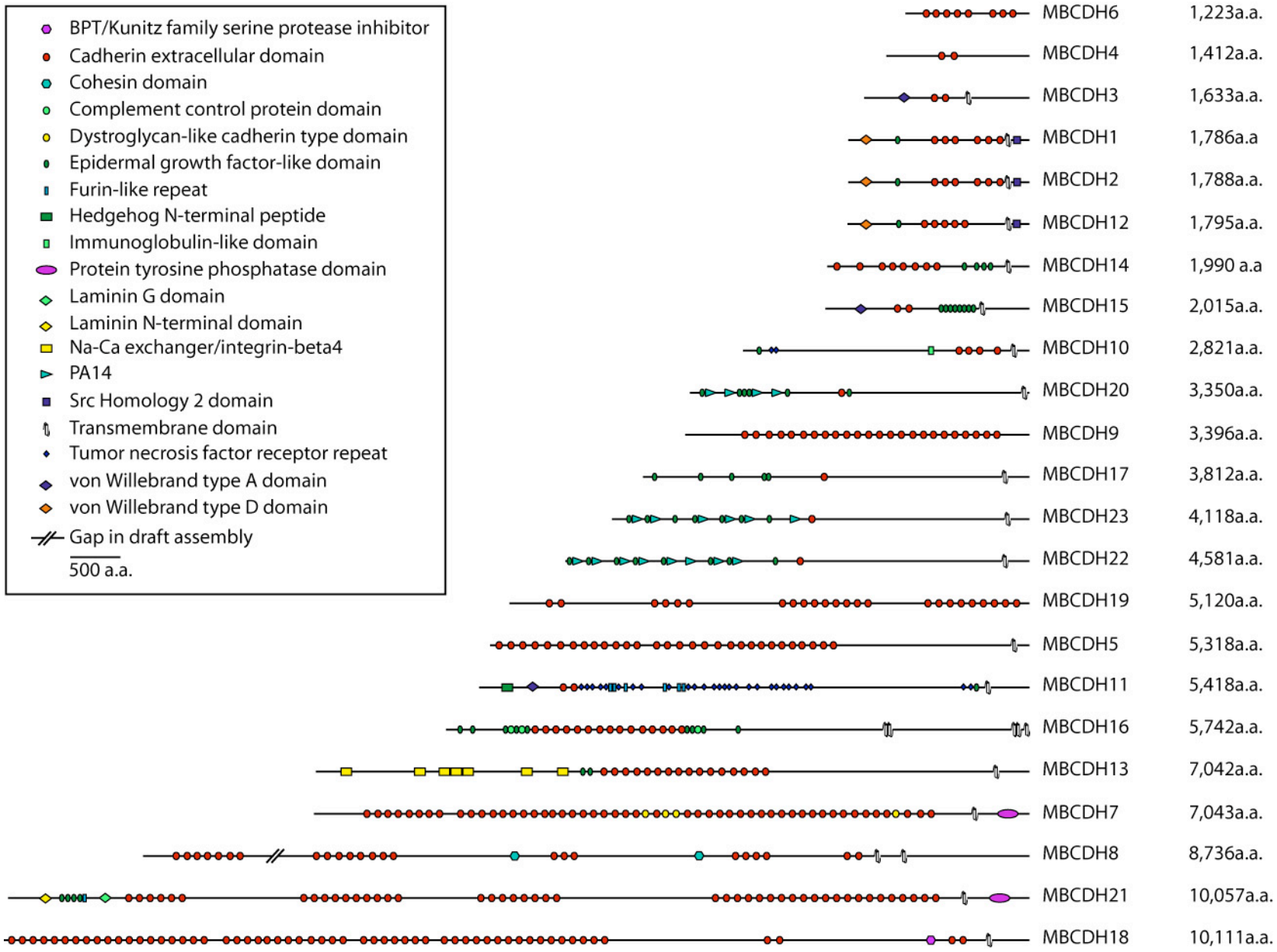

BPT/Kunitz family serine protease inhibitor
Cadherin extracellular domain
Cohesin domain
Complement control protein domain
Dystroglycan-like cadherin type domain
Epidermal growth factor-like domain
Furin-like repeat
Hedgehog N-terminal peptide
Immunoglobulin-like domain
Protein tyrosine phosphatase domain
Laminin G domain
Laminin N-terminal domain
Na-Ca exchanger/integrin-beta4
PA14
Src Homology 2 domain
Transmembrane domain
Tumor necrosis factor receptor repeat
von Willebrand type A domain
von Willebrand type D domain
Gap in draft assembly
500 a.a.
MBCDH6 1,223a.a.
MBCDH4 1,412a.a.
MBCDH3 1,633a.a.
MBCDH1 1,786a.a
MBCDH2 1,788a.a.
MBCDH12 1,795a.a.
MBCDH14 1,990 a.a
MBCDH15 2,015a.a.
MBCDH10 2,821a.a.
MBCDH20 3,350a.a.
MBCDH9 3,396a.a.
MBCDH17 3,812a.a.
MBCDH23 4,118a.a.
MBCDH22 4,581a.a.
MBCDH19 5,120a.a.
MBCDH5 5,318a.a.
MBCDH11 5,418a.a.
MBCDH16 5,742a.a.
MBCDH13 7,042a.a.
MBCDH7 7,043a.a.
MBCDH8 8,736a.a.
MBCDH21 10,057a.a.
MBCDH18 10,111a.a.

Figure S2.2. Domain structure of *N. vectensis* cadherins. *N. vectensis* possesses orthologs of previously identified cadherins (e.g. Fat-like, EGF LAG seven-pass G-type receptor-like, protocadherin-like) as well as proteins with unique domain arrangements not found in the other metazoans sampled here. Note: predicted proteins may be truncated due to gaps in draft genome assembly.

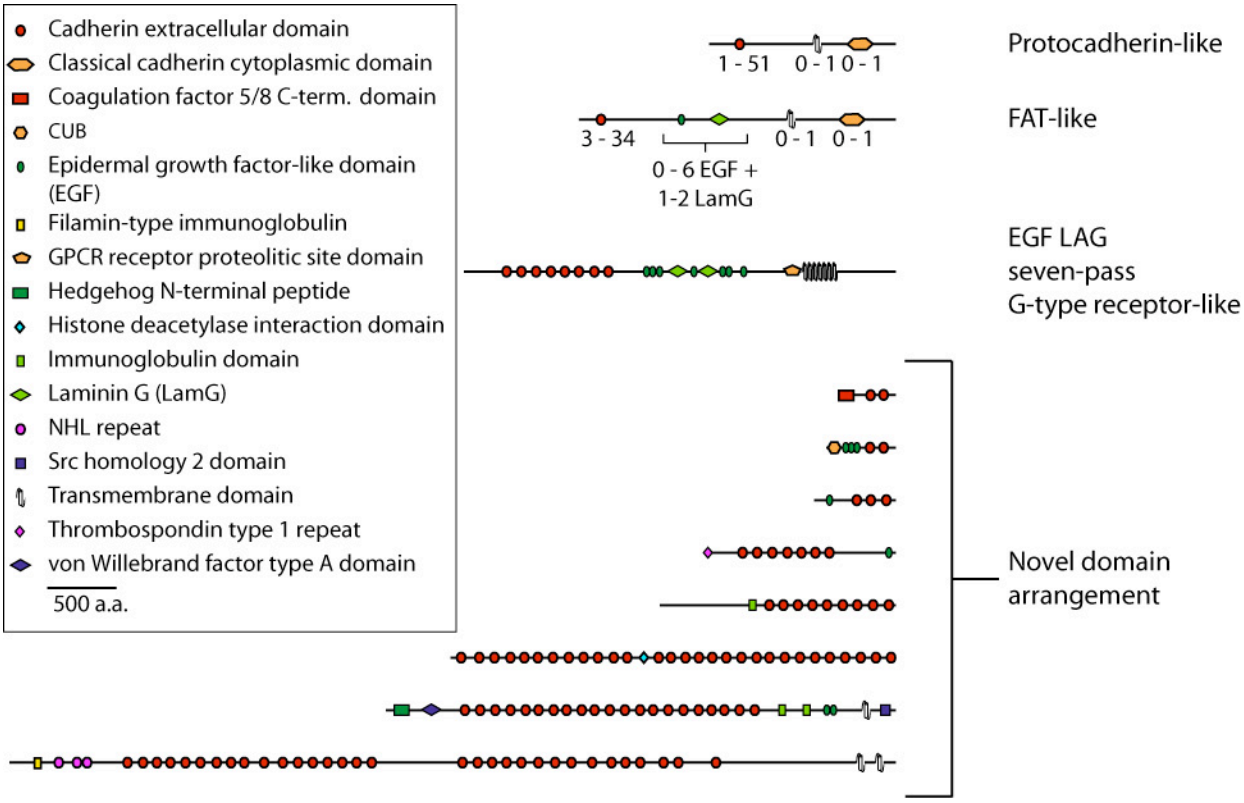


Figure S2.3. MBCDH1 and MBCDH2 alignment. MBCDH1 and MBCDH2, two proteins with identical protein domain composition, are encoded by distinct loci and are 93% identical across their lengths. The positions of predicted domains and the region of the protein against which antibodies were generated are highlighted. Alignments were generated using ClustalW. VWA: von Willebrand type A domain, EGF: epidermal growth factor domain, EC: extracellular cadherin repeat, TM: transmembrane domain, SH2: src homology 2 domain, Ab Epitope: antibody epitope.

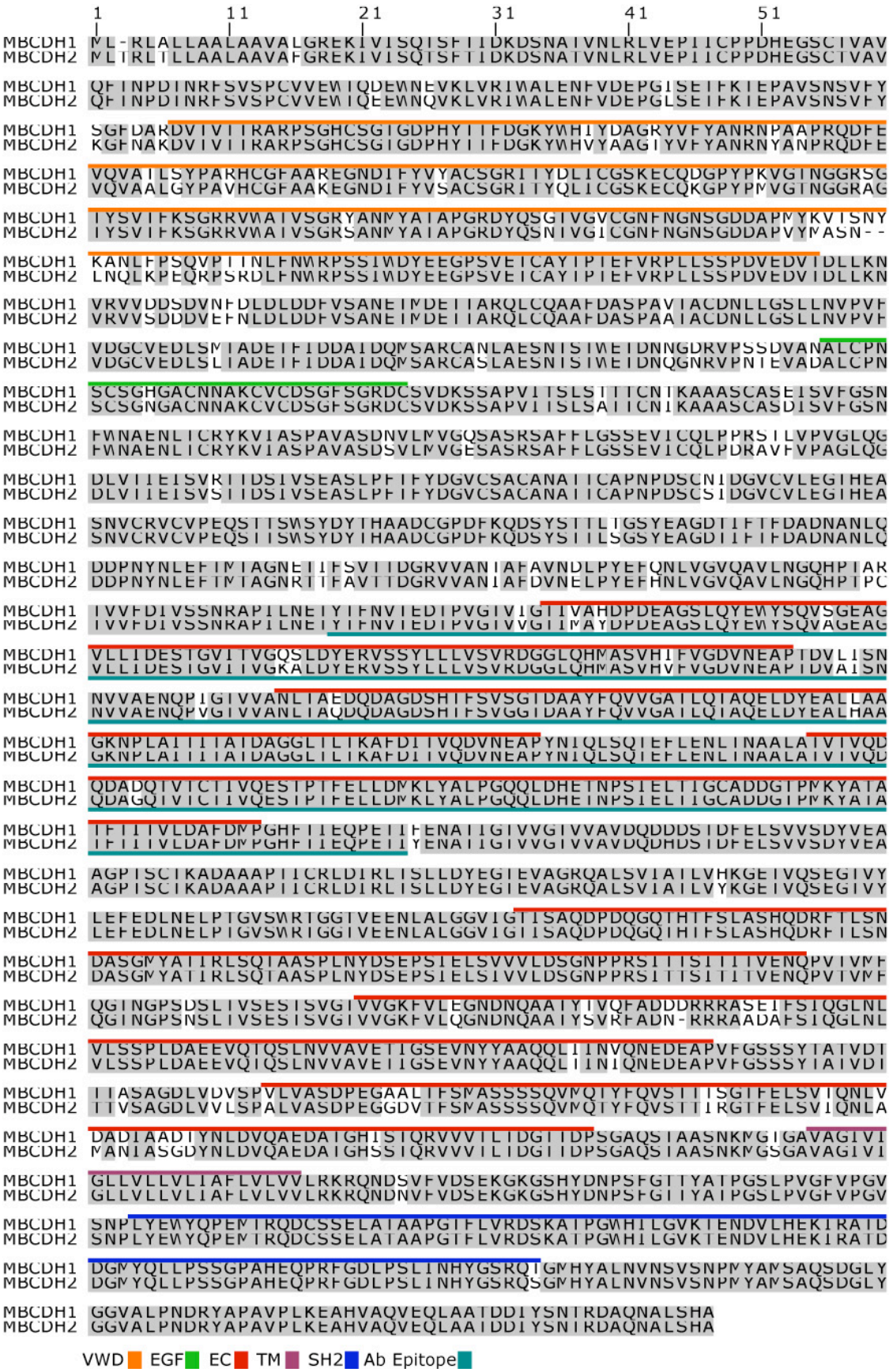
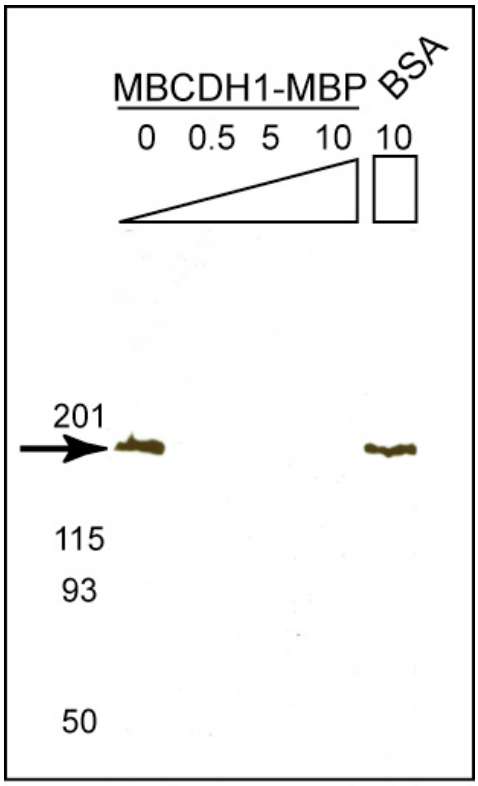

1
11
21
31
41
51
MBCDH1
MBCDH2
VWD EGF EC TM SH2 Ab Epitope

Figure S2.4. Specificity of MBCDH1 antibodies. *M. brevicollis* cell lysates probed with antibodies against MBCDH1 in the absence and presence of competitor demonstrating that in the absence of competitor, the antibody recognizes a single band of approximately 192 kDa (arrow). Pre-incubation of the antibody with 0.5, 5 or 10 μg of maltose binding protein-tagged MBCDH1 eliminated detection of the protein, indicating that the competitor bound to and saturated the available antibody.



Note S2.1. Similarity of MBCDH1 and MBCDH2.
MBCDH1 and MBCDH2 are two closely related *M. brevicollis* cadherins encoded by distinct loci. Both genes are expressed, based on EST data (Table S1) and RT-PCR using gene specific primers for MBCDH1 (TCAGTTCGCCGATGACGAT and CGTTGTGGTGTCGACGGTA) and MBCDH2 (GCGGTTTGCCGACAACCGC and ACTGTCGTAGTGTCCACAGTG). The predicted proteins are 93% identical and the region of MBCDH1 used to generate the antibodies is 95% identical to MBCDH2. The localization patterns in Fig.1 may reflect non-overlapping distributions of the two proteins, although their high degree of sequence identity suggests that they are likely to co-localize.

Chapter 3: Dynamic relationship between MBCDH1/2 protein levels and bacterial prey availability

SUMMARY

Choanoflagellate are proficient bacteriovores and have the potential to shed light on the evolution of animal-bacteria interactions like those that occur during bacterial infection. Some pathogenic bacteria bind to animal cadherins during host cell infection demonstrating that bacteria and animal cadherins have the capacity to interact. Moreover, localization of two *M. brevicollis* cadherins, MBCDH1 and MBCDH2 (MBCDH1/2), to the feeding collar suggests that they may be involved in sensing or capturing bacterial prey. To characterize the interaction between MBCDH1/2 and bacteria, I measured the effect of varying bacterial abundance and biodiversity on MBCDH1/2 protein levels. I discovered that long-term cultures of *M. brevicollis* fed either on *Flavobacter* sp. or *Enterobacter aerogenes* have differing levels of MBCDH1/2 protein depending on the bacterial species. In addition, rapid changes in bacterial prey availability cause a drop in MBCDH1/2 levels, suggesting that bacterial species identity and abundance affect MBCDH1/2 protein stability, turnover or translation. The connection between choanoflagellate cadherins and bacterial prey raises the possibility that cadherins in the unicellular ancestor of animals were attuned to sense changes in the extracellular environment.

INTRODUCTION

Interactions between animals and bacteria are wide-spread and have profound impacts on the development and physiology of animals. Some animal-bacteria relationships are symbiotic, such as those between gut-colonizing bacteria and intestinal epithelial cells (151). Others are host-pathogen in nature, such as those between disease-causing bacteria and animal host cells or cells of the innate immune system (152, 153). Understanding how these interactions evolved has important implications for our knowledge of animal biology and disease.

The close relationship between choanoflagellates and animals coupled with the intimate associations of choanoflagellates with their prey bacteria provide an opportunity to learn about the evolution of animal-bacterial interactions. Choanoflagellates are filter-feeding bacteriovores. To capture bacteria, they undulate their flagellum creating flow fields that draw water through the feeding collar, in turn trapping bacteria on the outer surface (154). Once in contact with the collar, bacteria are phagocytosed and transported to the food vacuole. How bacteria are recognized and captured by choanoflagellates is not understood. Indeed, little is known about the receptors used by any unicellular flagellates to capture bacteria from the water column (155).

In contrast to unicellular eukaryotes, much is understood about the ways in which animal and bacterial cells interact. Cells involved in vertebrate innate immunity typically use C-type lectins and Toll-like receptors to recognize and capture invading bacteria (156). Genome analyses indicate that *M. brevicollis* lacks Toll-like receptors and has only two membrane-bound C-type lectins (5), suggesting that choanoflagellates may use other classes of proteins to bind to bacteria. In animals, pathogenic bacteria are known to bind cadherins during infection. For example, epithelial E-cadherins (discussed briefly in Chapter 2) are used by the bacterial pathogen *Listeria monocytogenes* as an entry-point into the animal host cell. *Listeria monocytogenes* binds to E-

cadherin on the surface of an epithelial cell stimulating bacterial engulfment and ultimately leading to infection (145, 146, 157). In addition, E-cadherin and (putatively) Flamingo cadherin are receptors exploited by *Streptococcus pneumoniae* (147, 158) to gain entry into animal cells. These host-pathogen interactions demonstrate the capability of cadherins to bind to bacteria.

The fact that two *M. brevicollis* cadherins, MBCDH1 and MBCDH2 (MBCDH1/2), localize to the feeding collar [Chapter 2 and (11)] raises the possibility that choanoflagellate cadherins, like some of their animal homologs, interact with bacteria. One could imagine that the diversity of cadherins in choanoflagellates reflects the ability of different cadherins to bind specific classes of bacteria. Because disrupting expression of cadherins in *M. brevicollis* is not yet possible, I investigated the relationship between choanoflagellates and their bacterial prey by altering the concentration and diversity of bacteria in *M. brevicollis* cultures and measuring MBCDH1/2 protein levels in response to these changes. I found that MBCDH1/2 protein abundance differs between choanoflagellate cultures fed continuously with different species of bacteria. I also discovered that rapid changes in the bacterial environment created by removing or enriching bacterial prey cause a decrease in MBCDH1/2 after two days. Although the detailed interactions between MBCDH1/2 and bacteria have yet to be elucidated, the dynamic changes in MBCDH1/2 protein levels in response to bacteria suggest that choanoflagellate cadherins, and possibly those in the ancestor of choanoflagellates and animals, may be important for sensing and responding to environmental cues and were later co-opted by animals for intercellular interactions.

MATERIALS AND METHODS

Choanoflagellate culture conditions

I cultured *M. brevicollis* at 25°C in Artificial Seawater (32.9 g/L Tropic Marine) infused with cereal grass (5 g/ L) in a tissue culture dish. *Enterobacter aerogenes* or *Flavobacter sp.* were co-cultured with *M. brevicollis*. To maintain cultures, I transferred 1-3 ml of culture (containing both *M. brevicollis* and bacterial cells) to a 100 mm x 20 mm polystyrene dish adding fresh culture medium to obtain a final volume of 15 ml. The *M. brevicollis* + *Flavobacter sp.* monoxenic culture line was generated in August, 2004 and the *M. brevicollis* + *E. aerogenes* monoxenic culture line was created in July, 2005 by Rick Zuzow. (confirm dates)

Choanoflagellate-free culture conditions for bacteria

To grow *E. aerogenes* I picked a single bacterial colony into 10 ml of Artificial Seawater infused with cereal grass and incubated for approximately 24 hours at 25°C. To culture *Flavobacter sp.* I picked a single colony into 10 ml of Artificial Seawater infused with cereal grass and incubated for 5 days at 25°C.

Supplementing choanoflagellate cultures with bacteria

For experiments where bacteria were added to *M. brevicollis* cultures (Figure 3.2 A), I plated 1.7×10^6 *M. brevicollis* cells/well in a 6-well plate after spinning down the cells and resuspending them in 2.8 ml fresh cultured medium. I pelleted 200 µl of *E. aerogenes* and *Flavobacter sp.* (16,000*xg* for 1 minute) cultured as described above. Pellet size between the two samples was equal indicating that approximately equivalent amounts of bacteria were used. After resuspending bacteria in 200 µl of fresh medium I added them to the well containing *M.*

brevicollis culture. The cultures were incubated at 25°C for 48 hrs before harvesting cells for immunoblotting.

Immunoblotting

I prepared cell lysates by resuspending cell pellets (containing both *M. brevicollis* and the bacteria present in the cultures) in SDS-PAGE loading buffer (4% SDS, 100 mM Tris-HCL, 20% glycerol, 25 ng/ml Bromophenol Blue). To remove insoluble protein, I spun samples at 16,000*xg* for 3 minutes and collected the supernatant. After adding 100 mM 2-Mercaptoethanol to each sample, the proteins were separated by denaturing SDS-PAGE and subsequently transferred to a PVDF membrane. I blocked the membrane for 1hr at room temperature with PBS, 0.1% Tween-20, 5% Milk. After blocking, I probed with antib-MBCDH1/2 or E7 anti-β-tubulin primary antibodies (Developmental Studies Hybridoma Bank) and then Horseradish Peroxidase-conjugated secondary (Jackson ImmunoResearch Laboratories) antibodies diluted in blocking solution. After treating blots with SuperSignal Chemiluminescent Substrate (Pierce) I detected antigen by exposure to film.

Semi-quantitative RT PCR

I isolated total RNA from choanoflagellate culture lines using the Qiagen RNeasy Midi Kit (Qiagen) and subsequently isolated mRNA using the Qiagen mRNA Isolation Kit. For making cDNA and PCR amplification, I used the Superscript III RT PRC Kit (Invitrogen). To amplify MBCDH1 I used the following primers: forward TCAGTTCGCCGATGACGAT and reverse CGTTGTGGTGTCGACGGTA. To amplify MBCDH2 I used the following primers: forward GCGGTTTGCCGACAACCGC and reverse ACTGTCGTAGTGTCCACAGTG. To amplify GAPDH I used the following primers: forward TGGCTATCAAGGTTGGCATCAA and reverse CCTGAACGTTGGGGTTGCTC.

RESULTS

Based on the observation that MBCDH1/2 localizes to the feeding collar [Figure 2.2 and (11)], I hypothesized that these cadherins might interact with or respond to bacterial prey. To investigate this possibility, I compared MBCDH1/2 protein levels in cell lysates from *M. brevicollis* cultures grown continuously with one of two distinct prey bacteria: *Flavobacter sp*. and *E. aerogenes*. These cultures were generated approximately 5 years ago and have been mantained as stable long-term cultures. *Monosiga brevicollis* grown on *E. aerogenes* (hereafter, *Mbrev-Ea*) had significantly higher levels of MBCDH1/2 protein than those cultured on *Flavobacter. sp*. (hereafter, *Mbrev-Fs*) (Figure 3.1 A). To determine if this difference in protein abundance was a result of transcriptional regulation, I measured the levels of MBCDH1/2 mRNA by semi-quantitative RT-PCR (Figure 3.1 B, data corroborated by real-time PCR; Brock Roberts, pers. comm.). In contrast to the protein, MBCDH1/2 mRNA levels did not differ significantly between the two culture lines, suggesting that variation in the protein abundance may reflect changes in the rate of protein turnover or regulation at the level of translation.

If *M. brevicollis* cadherins are involved in sensing or capturing bacterial prey, it is possible that different cadherins respond to different bacterial species and this might explain the disparity in MBCDH1/2 protein levels between *Mbrev-Ea* and *Mbrev-Fs*. Given this hypothesis, I next sought to determine if the enrichment of MBCDH1/2 protein in *Mbrev-Ea* was species specific.

I therefore asked: does *E. aerogenes* directly stimulate accumulation of MBCDH1/2 protein or, conversely, does *Flavobacter sp.* directly reduce the level of MBCDH1/2? To address these questions, I performed a mixing experiment in which *Flavobacter sp*. was added to *Mbrev-Ea* and reciprocally, *E. aerogenes* was added to *Mbrev-Fs.* After two days of growth, I discovered that addition of a fresh aliquot of either species of bacteria to either *Mbrev-Ea* or *Mbrev-Fs* causes MBCDH1/2 levels to decrease (Figure 3.2 A). Therefore, it is unlikely that rapid changes in MBCDH1/2 levels is dependent on *E. aerogenes or Flavobacter sp.* species identity

An alternative explanation for the rapid decrease in MBCDH1/2 levels is that it is related to nutritional availability. Interestingly, bacteria in *Mbrev-Ea* and *Mbrev-Fs* form biofilm (possibly in response to choanoflagellate predation); these bacterial complexes are too large for the choanoflagellates to consume and their presence in choanoflagellate cultures may lead to a reduction in available nutrients. In contrast, bacteria added to the cultures in the mixing experiment described above are primarily single cells and are easily consumed by *M. brevicollis.* Therefore, it is possible that the change in MBCDH1/2 abundance upon addition of planktonic bacteria into *M. brevicollis* cultures is a response to increased food availability. To test this hypothesis, I starved the *M. brevicollis* culture lines by transferring the cells into unenriched seawater, which halts bacterial growth, thereby depriving the choanoflagellates of food. If an increase in food availability causes a decrease in MBCDH1/2, I predicted that removing bacteria would have the opposite affect of raising MBCDH1/2 protein levels. In contrast, I found that starvation has a similar affect to adding bacteria, causing MBCDH1/2 levels to drop (Figure 3.2 B and data not shown). This indicates that the level of MBCDH1/2 is not linearly related to the presence of prey bacteria.

DISCUSSION

The results of the *M. brevicollis* feeding and starvation experiments raise questions about the relationship between MBCDH1/2 and bacteria. Despite the fact that two long-term culture lines with distinct bacterial prey species have differing levels of MBCDH1/2, addition of the second species of bacteria did not equalize MBCDH1/2 protein abundance (Figure 3.1 and 3.2). In fact, addition of either bacterial species to either culture line caused a decrease in MBCDH1/2 protein, although the total amount of MBCDH1/2 was always greater in *M. brevicollis* cells from the culture line fed *E. aerogenes* than that of the culture line fed *Flavobacter sp*. This suggests that the modulation in MBCDH1/2 levels between the two cultures results from a different signal than that which triggers a decrease in MBCDH1/2 when I add or remove bacteria from the cultures.

Why has long-term culturing of *M. brevicollis* with a specific bacterial species changed the level MBCDH1/2? It is possible that *E. aerogenes* stabilizes MBCDH1/2 at the membrane, allowing the protein to accumulate on the cell surface. This would explain why the differences in protein levels do not reflect a difference in transcript abundance between the two culture lines. Alternatively, the bacteria could be indirectly affecting MBCDH1/2 by activating intracellular signaling that leads to changes in the rate of cadherin degradation or translation. Ancillary data supporting this hypothesis includes the presence of a tyrosine kinase Src homolog in *M. brevicollis* (159), which regulates E-cadherin degradation through phosphorylation in animals (160, 161). Another possibility is that *Flavobacter sp.* causes a decrease in MBCDH1/2 by

stimulating its cleavage or removal from the cell surface. *Helicobacter pylori*, a bacterial pathogen that causes gastritis and increases the risk of gastric cancer in animals, induces cleavage of E-cadherin on the surface of epithelial cells. The metalloprotease ADAM10, expressed by the animal host cell, is thought to mediate this cleavage (162-164). *Monosiga brevicollis* has ADAM-related proteins (114) and it is possible that, like *H. pylori* in animals, *Flavobacter sp.* is stimulating ADAM-mediated removal of the extracellular portion of MBCDH1/2 from the surface of *M. brevicollis*. If this is the case, the anti-MBCDH1 antibodies used in the experiments described here, which bind to the extracellular domain of MBCDH1/2, would only recognize intracellular or uncleaved protein. This would appear as a decrease in protein level in the cells from *Mbrev-Fs* cultures.

Although the studies of mixed cultures of *M. brevicollis* and bacteria did not reveal a clear function for MBCDH1/2, they did yield the discovery that increased food availability causes a drop in MBCDH1/2 protein (Figure 3.2). This suggests a link between MBCDH1/2 and the growth or nutrient state of *M. brevicollis*. Similarly, starvation of *M. brevicollis* also caused a decrease in MBCDH1/2 indicating that the relationship between MBCDH1/2 and bacterial prey is complex and warrants further investigation. As antibodies to additional *M. brevicollis* cadherins become available, it will be important to determine if they too respond to bacterial prey.

Bacteria have a profound impact on the physiology and development of animals (165-167). *M. brevicollis* represents an unprecedented opportunity to study the relationship between choanoflagellates and their bacterial prey and provides a system for analyzing the effect of diverse bacterial species on eukaryotic biology. Evidence suggesting that MBCDH1/2 respond to bacteria through changes in protein abundance provides a potential first step in uncovering the function or regulation of cadherins in *M. brevicollis* and is consistent with the notion that cadherins in the unicellular ancestor of animals were important for sensing or responding to the extracellular environment.

FIGURES

Figure 3.1. Sensitivity of MBCDH1 and MBCDH2 protein levels to bacterial prey species. *M. brevicollis* cells grown on *E. aerogenes* (*Ea*) have higher levels of MBCDH1/2 protein than *M. brevicollis* continuously fed *Flavobacter sp.* (*Fs*) (A). Cell lysates probed with anti-MBCDH1/2 antibodies by Western blot show the abundance of MBCDH1/2 relative to the loading control, β-tubulin (A). In contrast to MBCDH1/2 protein, MBCDH1 and MBCDH2 transcript levels analyzed by semi-quantitative RT-PCR (compared to GAPDH loading control) do not differ significantly between the two *M. brevicollis* culture conditions (B).

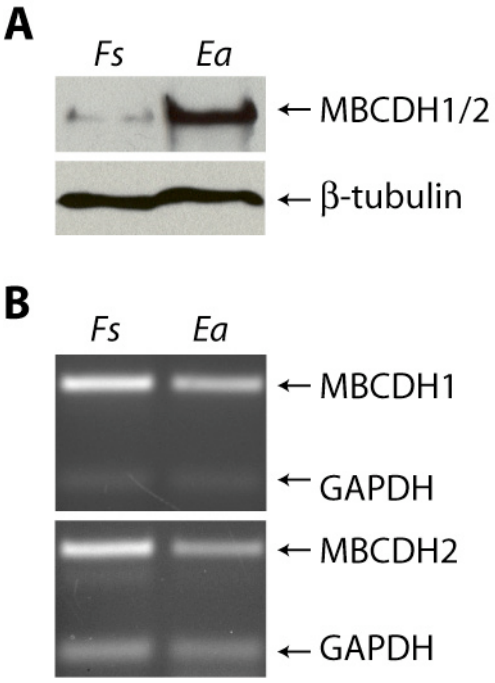
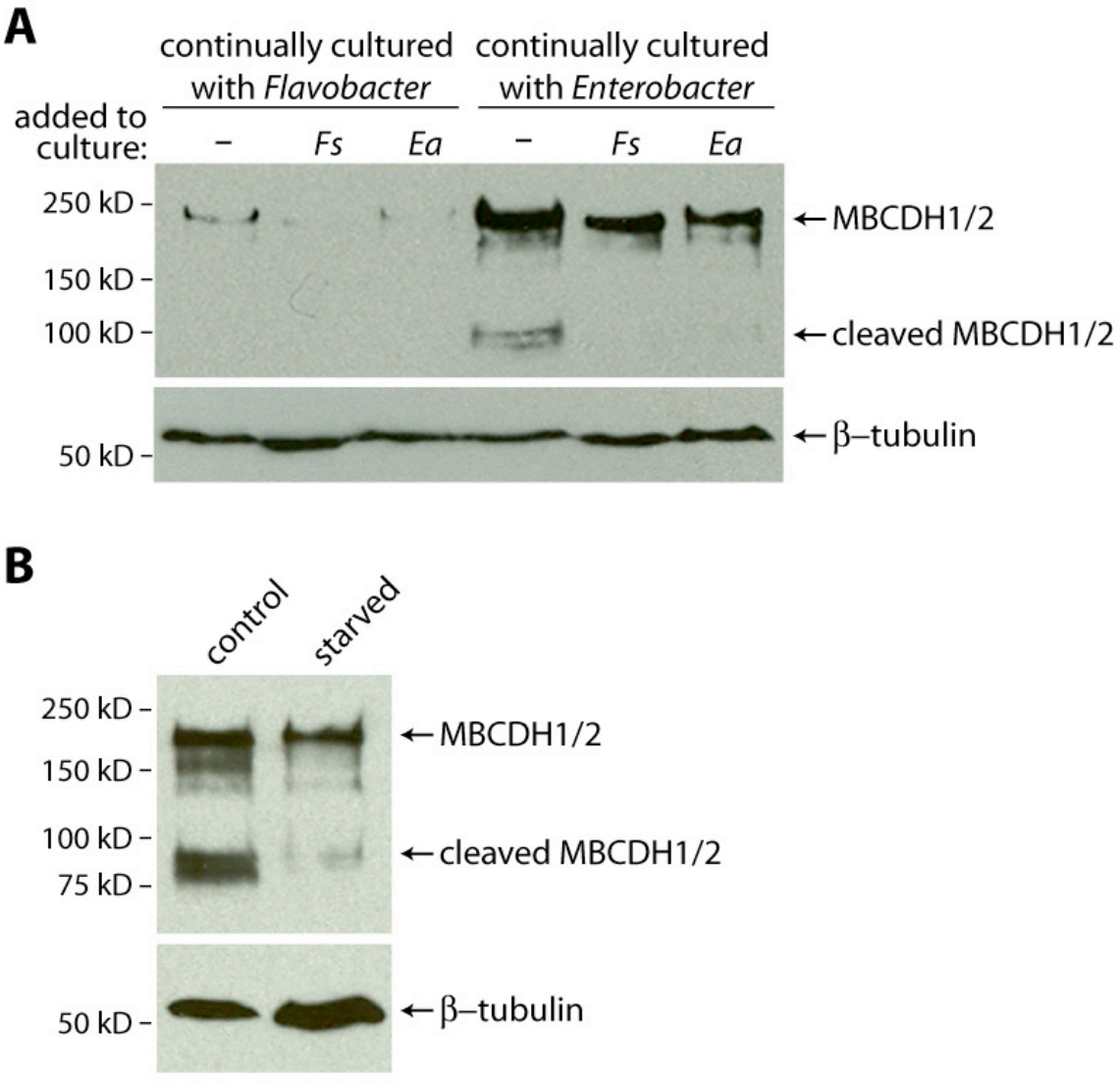


Figure 3.2. Altering bacterial food availability causes a change in MBCDH1 and MBCDH2 protein levels. (A) MBCDH1/2 levels decrease when bacteria are added to *M. brevicollis* culture lines. *E. aerogenes* (*Ea*) and *Flavobacter sp*. (*Fs*) bacteria cultured in the absence of choanoflagellates were added to *M. brevicollis* cultures that had been continuously grown in the presence of *Flavobacter sp*. or *E. aerogenes*. Upon addition of bacteria, cultures were incubated for 48 hours before harvesting cells. (B) Starvation of *M. brevicollis* co-cultured with *E. aerogenes*, achieved by transferring cells into artificial seawater for 48 hours, also caused a decrease in MBCDH1/2 levels. For both (A) and (B), lysates from the cultured cells were probed with anti-MBCDH1/2 antibodies to detect abundance of MBCDH1/2 protein. Additional, lower molecular weight bands are presumed cleavage products of MBCDH1/2. β-tubulin served as a loading control.



Chapter 4: MBCDH1 does not mediate homophilic adhesion *in vitro* or in a heterologous system

SUMMARY

Homophilic adhesion mediated by classical cadherins represents one of the best-understood functions of animal cadherins and has important roles in development and disease. *Trans* interactions between individual ECs on neighboring cells are responsible for the specificity and adhesive force of cadherins and binding of Ca^{2+} between each EC is required for homophilic adhesion. To test for homophilic adhesion activity by a choanoflagellate cadherin, I performed binding assays with the *M. brevicollis* cadherin MBCDH1 *in vitro* and in a heterologous mammalian cell culture system. In the context of these experiments, I found that MBCDH1 does not interact homophilically. One interpretation of these results supported by insights from choanoflagellate biology and characterization of non-classical animal cadherins is that cadherin-mediated homophilic adhesion evolved after the divergence of animals and choanoflagellates. Alternatively, uncharacterized post-translational modifications of MBCDH1 may be required to facilitate homophilic adhesion *in vivo* and in laboratory assays.

INTRODUCTION

The evolution of adhesion molecules was a critical step in animal evolution, allowing for the evolution of coordinated cell adhesion that is indispensable during development. Differential adhesion between cells contributes to the complex morphogenic changes that occur during embryogenesis. For example, alterations in cell adhesion during neurulation of chick embryos (168) and mesoderm formation in the *Drosophila* gastrula (169) are critical for proper development. Moreover, loss of cell adhesion is required for the epithelial-mesenchymal transitions that occur throughout ontogeny and has important implications during cancer progression (170, 171). Because of the central importance of cell adhesion in development and disease, the discovery of the first cadherin adhesion protein over thirty years ago (70, 172) launched an ongoing effort to understand the molecular and biochemical underpinnings of cadherin-based cell adhesion.

The founding member of the cadherin superfamily is a classical cadherin and, until recently, cadherin research has focused primarily on this subclass (64, 71). The structure of most classical cadherins consists of 5 tandemly arrayed ECs, a transmembrane domain and a CCD (discussed in Chapter 2). Between each EC is a Ca^{2+} binding pocket. Investigation into the functional role of Ca^{2+} revealed its requirement for cadherins to maintain their stiff, rod-like structure (72) explaining why cadherin-mediated adhesion is dependent on Ca^{2+}.

Early studies of cadherins showed that mixed cell populations expressing different classical cadherins segregated based on cadherin subtype (173). This sorting function is mediated by homophilic interactions between cadherins on neighboring cells (174) and the N-terminal-most EC domain is primarily responsible for binding specificity (175). Although binding specificity lies in the first EC, other ECs also contribute to cell adhesion. Testing the binding affinity of truncated forms of E-cadherin suggests that at least the second EC is required and similar experiments with C-cadherin indicate that multiple ECs are important for adhesion (176, 177).

Nevertheless, it is unclear whether non-N-terminal EC domains play a direct role in adhesion or are more important for protein structure (178). Interestingly, analysis of cadherin crystal structure and sequence alignment data suggests that the molecular mechanism of protein interactions between classical cadherins (exchange of N-terminal β-strands at the adhesive interface) differs from that of non-classical cadherins (179). Indeed, a growing body of evidence indicates that cadherins perform diverse functions not limited to homophilic adhesion (64, 98, 180)

The well-defined adhesive function of animal cadherins raises the possibility that cadherins in choanoflagellates can also mediate adhesion. Although a clear role for adhesion in the strictly unicellular *M. brevicollis* is not immediately obvious, it is possible that cadherins play a structural function in maintaining aspects of choanoflagellate cell architecture. For example, MBCDH1 localizes to the microvilli of the feeding collar, which it may stabilize through homophilic interactions among MBCDH1 molecules on adjacent microvilli [Chapter 2 and (11)]. Alternatively, choanoflagellates might use cadherins to mediate adhesion during intercellular interactions not yet documented in the laboratory, e.g. mating type recognition, sex or competition for a substrate attachment site (181). In this chapter, I investigate the homophilic binding capacity of one choanoflagellate cadherin, MBCDH1. Contrary to the animal classical cadherins, I have found that MBCDH1 does not mediate homophilic adhesion *in vitro* or in a mammalian cell adhesion assay.

MATERIALS AND METHODS

Cadherin-Fc constructs

Cloning

I generated MBCDH1-Fc fusions by PCR amplifying extracellular MBCDH1 from *M. brevicollis* cDNA using the following primers: forward CCGGAATTCGCCATGCTGCGCTTGGCTTTGCTCGCCG and reverse CCGGGATCCCGGCACCAGTGCCCATCTTGTTG. The forward primer contained an *EcoRI* restriction site (underlined) and the reverse primer contained a *BamHI* restriction site (underlined). I PCR amplified both the monomeric and dimeric forms of human IgG1 off plasmids gifted by the Nelson Lab (Sanford) using the following primers: forward CCGGGATCCTACTCGAGCTCGACAAAACTCAC and reverse CCGTCTAGATTACCCCGGAGACAGGGAGAG. The forward primer contained a *BamH1* restriction site (underlined) and the reverse primer contained an *XbaI* restriction site (underlined). I first sub-cloned the gene fragments into pCR2.1 using the TOPO TA Cloning Kit (Invitrogen), and subsequently cloned them into the mammalian expression vector pCDNA3.1/myc-His A (Invitrogen) using engineered restriction-cut sites to generate MBCDH1-Fc fusions. The E-cadherin-Fc monomer and dimer expression constructs (in pCDNA3.1A) were a gift from the Nelson Lab.

Tissue culture, transfection and selection

I maintained HEK 293 cells (ATCC # CRL-1573) in High Glucose Dulbecco's modified Eagles's medium (DMEM) with 10% fetal bovine serum and Penicillin and streptomycin (Invitrogen). For transfection, I plated cells into 24-well dish in DMEM such that the following day, cells were approximately 50% confluent. I transfected 500 ng of plasmid purified from *E.*

coli DH5-alpha cells and diluted in Optimem (Invitrogen) into HEK 293 cells using Lipofectamine LTX (Invitrogen) according to the manufactures protocol. The cells were cultured for approximately two weeks in the presence of 1.3 mg/ml G418 (Fisher Scientific) to select for stable transfectants.

Full-length and GFP-tagged cadherin constructs

Cloning

To clone full-length MBCDH1, I PCR amplified off *M. brevicollis* cDNA using the following primers: forward CGGGGATCCACCATGCTGCGCTTGGCTTTGCT and reverse GGCGAATTCTTAGGCATGGCTCAGCGC. The forward primer contained a *BamH1* restriction site (underlined) and the reverse primer contained an *EcoRI* restriction site (underlined). I PCR amplified full-length MBCDH1 for the GFP fusion construct from *M. brevicollis* cDNA using the same forward primer mentioned above and the following reverse primer: GCCTCCGGA**GCCGCCGCC**GGCATGGCTCAGCGCGTTCT containing a *BspEI* restriction site (underlined) and a linker region (bold) coding for three glycines. Canine E-cadherin was amplified off a plasmid supplied by the Nelson Lab using the following primers: forward CGAGCGGCCGCACCATGGGCCCTCGGTACGGCGG and reverse GGGGAATTCTAGTCGTCCTCGCCACCT. The forward primer contained a *Not1* restriction site (underlined) and the reverse primer contained an *EcoRI* restriction site (underlined). For the GFP fusion construct, I PCR amplified E-cadherin using the same forward primer mentioned above and the following reverse primer: GCCTCCGGA**GCCGCCGCC**GTCGTCCTCGCCACCT containing a *BspEI* restriction site (underlined) and a linker region (bold) coding for three glycines. GFP was PCR amplified off pmaxGFP (Amaxa Biosystems) using the following primers: forward CGGTCCGGA**GGCGGC**ATGGTGAGCAAGGGCGAGG and reverse GCCGAATTCTTACTTGTACAGCTCGTCCA. The forward primer contained a *BspEI* restriction site (underlined) and a linker region (bold) coding for 2 glycines and the reverse primer contained an *EcoRI* restriction site (underlined). I cloned all fragments directly into the retroviral expression vector pQCXIN (Clontech) except MBCDH1 with *BamHI* and *EcoRI* restriction sites, which I first subcloned into pCR2.1 using the TOPO-TA Cloning Kit (Invitrogen). The resulting cadherin-GFP fusion proteins contain a linker region between the cadherin C-terminus and the N-terminus of GFP that consists of the following amino acids: Gly-Gly-Gly-Ser-Gly-Gly.

Tissue culture, viral-mediated transduction and selection

I cultured L cells (ATCC # CRL-2648) in High Glucose DMEM with 10% fetal bovine serum and Penicillin and streptomycin. For virus production, I plated viral packaging cell line (GP2-293T - gift from Barton Lab, UC Berkeley) at 7.5 x 10^6 cells per 10 cm tissue culture dish and grew overnight. I co-transfected the cells with 15ug of expression plasmid DNA and 8ug of VSVG plasmid (gift from Barton Lab - constructed by Nolan Lab, Stanford) as described in section above. The next day, I transferred the packaging cells to 32 ^{o}C for 24 hours. Virus was collected from cell supernatant by centrifugation (50,000*xg*, 95 min, 4^{o}C), resuspended in culture medium + PolyBrene (Invitrogen) and applied to L cells at 50% confluency (plated the day before). After incubating the cells at 32 ^{o}C for 24 hours, I transferred them to 37 ^{o}C. Two days after transduction, the cells were cultured for approximately 1.5 weeks under 1 mg/ml G418 selection. To isolate cells expressing high levels of cadherins, I sorted the cells by GFP

expression levels (for cadherin-GFP fusion proteins) or by immunofluorescence detection (for cadherins without GFP tag) of protein using the MoFlo XDP cell sorter (Beckman Coulter). The cells with the highest fluorescence intensity (top 2-4%) were collected in batch. Between 3,000 and 15,000 cells/cell line were obtained and expanded for hanging drop assays.

Immunofluorescent staining and microscopy

Staining for cell sorting

I immunostained live L cells by spinning them at 700*xg* for 10 minutes at 4 ^{o}C, resuspended them in PBS, 5% FBS and added either anti-MBCDH1/2 primary antibodies or DECMA monoclonal primary antibodies (recognizes E-cadherin) and incubated on ice for 30 minutes. I then washed the cells with DPBS, 5% FBS and applied Alexa Fluor 488 goat anti-rabbit IgG (H+L) secondary antibodies (Molecular Probes) to the MBCDH1 expressing cells and Alexa Fluor 488 goat anti-rat IgG (H+L) secondary antibodies to the cells expressing E-cadherin. I incubated the cells on ice for 30 minutes, washed with DPBS, 5% FBS and resuspended them in DPBS, 3% FBS.

Staining after cell permeabilization for microscopy

The day before staining, I plated L cells in a glass bottom dish such that they would be approximately 40% confluent the next day. To fix the cells, I washed 2 times with DPBS (Invitrogen) and applied 3% formaldehyde diluted in DPBS to the cells in the dish. After incubating for 20 minutes, I removed the formaldehyde, washed the cells 3 times with DPBS, applied the primary antibodies (either anti-MBCDH1/2 or DECMA) diluted in blocking solution (DPBS, 0.1% Triton X-100, 5% BSA) and incubated for 1 hour. I then washed the cells 3 times and applied the Alexa Fluor 568 goat anti-rabbit IgG (H+L) to cells treated with anti-MBCDH1/2 and Alexa Fluor 568 goat anti-rat IgG (H+L) to cells treated with DECMA. Both secondary antibodies were diluted in blocking solution. After incubating for 1 hour, I washed the cells 4 times with DPBS and stored the cells in DPBS for visualization.

Staining before cell permeabilization for microscopy

I plated L cells as described above in glass bottom dishes and washed 2 times with DPBS before applying primary anti-MBCDH1/2 or DECMA antibodies diluted in DBPS + 5% FBS to cells. After incubating for 30 minutes at 37 ^{o}C, I washed the cells 3 times with DBPS and fixed and stained with secondary antibodies as described above.

Immunoblotting

I prepared cell lysates by resuspending cell pellets in SDS-PAGE loading buffer (4% SDS, 100mM Tris-HCL, 20% glycerol, 25ng/ml Bromophenol Blue, 100mM 2-Mercaptoethanol). Proteins were separated by denaturing SDS-PAGE and subsequently transferred to a PVDF membrane. I blocked the membrane for 1hr at room temperature with PBS, 0.1% Tween-20, 5% Milk. After blocking, I probed with primary and then Horseradish Peroxidase-conjugated secondary antibodies diluted in blocking solution. After treating blots with SuperSignal Chemiluminescent Substrate (Peirce) I detected antigen by exposure to film.

Bead assays

Cadherin-Fc constructs were purified as described previously (182) using Protein-A agarose beads (Millipore) and stored in 10mM Hepes, 50mM Nacl, pH 7.2. For bead assays, I

conjugated cadherin-Fc proteins to Protein-A-coated microbeads (Bangs Laboratories, 0.86um diameter), transferred the beads into Artificial Seawater (Tropic Marine) and sonicated with a jewelry sonicator for 1 minute to break up bead clumps. Beads were incubated in 0.5 ml tube for 1 hour at room temperature and subsequently visualize at 40X magnification and brightfield microscopy.

Hanging drop assays

For hanging drop assays performed in DMEM, I trypsinized the L cells and diluted to a density of 200 cells/µl. I then applied 20µl of cells (4,000 total) onto the lid of a petridish. After quickly flipping over the lid, I placed it on the bottom half of the dish, which contained a few ml of PBS to prevent evaporation. The cells were incubated overnight at 37 ^{o}C. The next day, I transferred the droplets to a slide using a p1000 pipette tip with the end cut off and placed a coverslip over the drop of cells. The cells were imaged at 10X magnification using phase microscopy.

For hanging drop assays performed in Artificial Seawater, I lifted cells off the substrate using 5mM EDTA in DPBS (without calcium and magnesium) and diluted them in Artificial Seawater + 5% FBS, 25mM Hepes, 1x L-glutamine, 1x sodium pyruvate to a concentration of 5000 cells/µl. I applied 20µl of cells (100,000 total) to the lid of a 33mm dish containing water in the bottom to prevent evaporation. After incubating for 2 hours at room temperature, I triturated the drop of cells 5 times with a p200 pipette tip and transferred them to a slide for visualization. The cells were imaged at 10X magnification using DIC microscopy.

Image collection

I captured all images using a Leica DMI6000 B inverted compound microscope and Leica DFC350 FX camera.

RESULTS

To determine if MBCDH1 can mediate adhesion through homophilic binding I used two different assays previously demonstrated to detect homophilic adhesion between diverse classical cadherins. In the first I tested the ability of recombinant MBCDH1 to stimulate bead aggregation (177) and in the second I determined if MBCDH1 could induce adhesion of mouse fibroblasts (173, 174, 183).

Bead Assay

The extracellular domain of animal classical cadherins ectopically expressed and purified from mammalian cells is sufficient for homophilic adhesion *in vitro* (177, 184). To test if MBCDH1 can mediate homophilic adhesion in a biochemical assay I generated fusion proteins consisting of the extracellular domain of the cadherin fused at its C-terminus to the Fc region of human IGg1 (Figure 4.1). There is evidence, albeit controversial (73), suggesting that *cis* interactions between cadherins on the surface of the same cell are important for adhesive function and that cadherins permitted to form *cis* dimers have greater adhesive activity than the monomeric forms in bead aggregation assays (185, 186). To account for the potential importance of *cis* interactions for MBCDH1 function, I constructed dimeric and monomeric forms of MBCDH1-Fc and E-cadherin-Fc (Figure 4.1 A), which served as a positive control. The secreted cadherin-Fc dimer proteins form a homodimer through disulfide bridges in the hinge region of the IgG

fragment. In the monomer form, two cysteine residues were mutated to serines to prevent dimerization via disulfide bonds. The fusion proteins were expressed and purified from HEK 293 cells. Analysis by SDS-PAGE indicated that Fc dimers formed properly and Fc monomers did not form disulfide-based dimers (Figure 4.1 B).

To test for homophilic adhesion, I coupled the purified cadherin-Fc proteins to protein A-coated beads, which bind Fc, and incubated the beads in artificial seawater in the presence and absence of Ca^{2+}. As expected, both monomeric and dimeric E-cadherin-Fc-coated beads formed large aggregates in the presence of calcium but did not aggregate when calcium was chelated (Figure 4.2 E-H). In contrast, MBCDH1-Fc-coated beads did not aggregate when Ca^{2+} was present (Figure 4.2 A-D), indicating that the extracellular domain of recombinant MBCDH1 is not sufficient for homophilic adhesion *in vitro*.

L cell assay

It is possible that the transmembrane domain and cytoplasmic tail of MBCDH1 are required for homophilic adhesion, highlighting a need to test the binding capacity of full-length MBCDH1. Ectopic expression of full-length E-cadherin in mouse fibroblast cells (L cells), which do not normally adhere, can induce cell aggregates (175, 183). To test the ability of MBCDH1 to direct adhesion between animal cells I stably expressed full-length MBCDH1 in L cells. To facilitate visualization of cadherin in live L cells, I also expressed MBCDH1 fused at its C-terminus to GFP. E-cadherin and E-cadherin-GFP again served as positive controls. To select for L cells expressing high levels of cadherin, I sorted cells by flow cytometry, collecting the 4% of cells with the highest protein levels based on GFP or antibody fluorescence intensity. Western blots with anti-MBCDH1/2 and anti-E-cadherin antibodies reveal that the cadherins are expressed (Figure 4.3).

To determine if the cadherins reached the surface of the L cells, I visualized the subcellular localization of the proteins. Fluorescence microscopy showed that MBCDH1-GFP and E-cadherin-GFP are enriched at the membrane (Figure 4.2 A and C) and IMF using antibodies against MBCDH1 and E-cadherin demonstrate that GFP is properly labeling cadherins (Figure S4.2 A and B). IMF of fixed, permeablized L cells expressing MBCDH1 and E-cadherin shows that they are enriched at the membrane like the GFP fusion proteins. To confirm that the proteins are indeed at the plasma membrane, I applied MBCDH1/2 and E-cadherin antibodies to live cells, a condition in which only proteins displayed on the surface become labeled (Figure 4.2 B and D; Figure S4.2 A and B). Brightly staining cadherins indicate that a large fraction of the protein pool was present on the surface of the cell. Lastly, I tested the specificity of the cadherin antibodies by applying E-cadherin antibodies to L cells expressing MBCDH1 and MBCDH1/2 antibodies to E-cadherin-expressing L cells and saw no fluorescence (Figure S4.1 C).

Adhesion of L cells to the surface of the culture dish can make it difficult to determine definitively if the cells are actively adhering to each other, or are simply close neighbors. To prevent substrate adhesion while forcing cells to make contacts with each other, I performed hanging drop assays. Briefly, cells were suspended in a drop of liquid to maximize cell-cell interactions and minimize cell-substrate interactions (Figure 4.5 A) and subsequently assayed for aggregation by microscopy. L cells expressing MBCDH1 or MBCDH1-GFP did not adhere while E-cadherin and E-cadherin-GFP L cells formed large aggregates (Figure 4.5 and S4.2),

corroborating the data from the bead assays that recombinant MBCDH1 is not sufficient for homophilic adhesion.

Relative sizes and potential modifications of MBCDH1

Due to glycosylation and N-terminal cleavage, the molecular weights of mature animal classical cadherins that have passed through the Golgi apparatus are typically higher than predictions made based on their amino acid sequences (187). Indeed, the molecular weights of both the E-cadherin-Fc and MBCDH1-Fc proteins are greater than predicted (Figure 4.1 B) suggesting that posttranslational modifications occurred in the HEK 293 cells. In contrast, the size of endogenous MBCDH1 detected in choanoflagellate lysates is similar to its predicted molecular weight (Figure 4.3), suggesting that it undergoes fewer post-translational modifications *in vivo* than when expressed in HEK 293 cells. In addition to processing by glycosylation, cleavage of the E-cadherin N-terminal pre sequence in the Golgi is required for proper protein function (188). A second, larger species of E-cadherin-Fc monomer was present in purified sample likely representing unprocessed protein still containing the pre sequence.

Like the cadherin-Fc fusion proteins, the predicted molecular weight of full-length cadherins expressed in L cells differs from the actual size of the proteins. Amino acid sequence indicates that MBCH1 is expected to be 192 kDa. Comparisons between endogenous *M. brevicollis* MBCDH1/2 and MBCDH1 from L cell lysates show that the molecular weight of endogenous MBCDH1/2 is less than MBCDH1 in L cells. This suggests that MBCDH1 may be more heavily glycosylated or modified in L cells or that cleavage of MBCDH1 occurs in *M. brevicollis*.

DISCUSSION

The fact that MBCDH1 does not appear to mediate homophilic adhesion comes as no surprise. Despite our relatively detailed understanding of homophilic adhesion among the classical cadherins, they are only a small class of proteins within the much larger cadherin superfamily and it is clear that many other family members do not interact in this way. As mentioned in chapter 2, tip-link-associated cadherin 23 and protocadherin 15 interact heterophilically (144) to connect adjacent microvilli on hair cells (stereocilia) in the inner ear. *M. brevicollis* possesses a diversity of cadherins (Figure S2.1), any of which could be interacting heterophilically with MBCDH1 to connect the microvilli that make up the feeding collar, a function reminiscent of stereocilia tip-linkages. Binding between the atypical cadherins Fat and Daschous stimulates planar cell polarity signaling in *Drosophila* and mice, which is critical for regulating tissue growth and patterning during development (96, 98, 103). Likewise, heterophilic binding between cadherins may trigger intracellular signaling in choanoflagellates in response to transient cell-cell contact or adhesion during conjugation. In support of a cell signaling function for choanoflagellate cadherins, the conserved SH2 domain on the cytoplasmic tail of MBCDH1 hints at a connection between cadherins and tyrosine kinase signaling.

Interactions between cadherin and non-cadherin proteins have also been documented in animals. The protocadherin PAPC and the membrane protein Frizzled, which together regulate the separation of ectodermal and mesodermal cells during *Xenopus* gastrulation, have been shown to bind through extracellular interactions in a biochemical assay (189). In addition, contact between cadherins and the receptor tyrosine kinases Epidermal Growth Factor Receptor and

Fibroblast Growth Factor Receptor can activate or maintain downstream MAPK signaling (180, 190-192) again highlighting the link between cadherins and tyrosine kinase signaling. The remarkable abundance of choanoflagellate tyrosine kinase signaling machinery (193) and the prevalence of related domains in *M. brevicollis* cadherins lends support for the idea that ancient cadherins intersected directly with tyrosine kinase signaling pathways.

In addition, it is possible that MBCDH1 proteins do interact through homophilic binding but the assays used here failed to detect binding. Postranslational processing is often critical for protein function and alteration of cadherin glycosylation has been shown to affect cadherin-based cell adhesion (187, 194). If MBCDH1 was not properly modified in the mammalian cells in which it was expressed, this could block homophilic binding. Indeed, the size difference between endogenous MBCDH1 and ectopically expressed MBCDH1 in L cells (Figure 3.3) suggests that ectopic postranslational modification occurred in the L cells.

Characterizations of cadherins in early-branching animal lineages [Figure S2.2 nematostella cadherins, (60, 71, 195)] indicate that the well-studied classical cadherin family that mediates homophilic adhesion was not elaborated until the appearance of the vertebrates. The abundance of cadherins lacking homophilic adhesive function in animals and the incredible diversity of cadherin protein architectures in both animals and choanoflagellates suggest that cadherin-mediated homophilic adhesion evolved after the divergence of choanoflagellates and animals and may have arisen concomitantly with or after multicellularity.

FIGURES AND TABLES

Figure 4.1. Purified Cadherin-Fc chimeric fusion proteins. (A) Schematic representation of cadherin-Fc protein constructs. Chimeric fusions were constructed of either canine E-cadherin or MBCDH1 extracellular domain cDNA (from start codon to transmembrane domain) linked to the Fc domain of human immunoglobulin (IgG1) heavy chain monomer or dimer cDNA. The mature secreted cadherin-Fc dimer protein forms a homodimer through disulfide bridges in the hinge region of the IgG fragment. In the monomer form, two cysteine residues are mutated to serines to prevent dimerization via disulfide bonds. (B) Secreted recombinant proteins were purified from the supernatant of HEK 293 cells by incubation with protein A-coated beads and analyzed by SDS-PAGE in the absence of a reducing agent with subsequent Coomassie staining. Lane 1, E-cadherin Fc-Monomer; Lane 2, E-cadherin-Fc dimer; Lane 3, MBCDH1-Fc monomer; Lane 4, MBCDH1-Fc dimer. Predicted molecular weight based on amino acid sequence is indicated in parentheses. Note that predicted and actual molecular weights differ. The presence of both the pre and pro species of E-cadherin-Fc-monomer indicates that both processed and unprocessed protein is secreted.

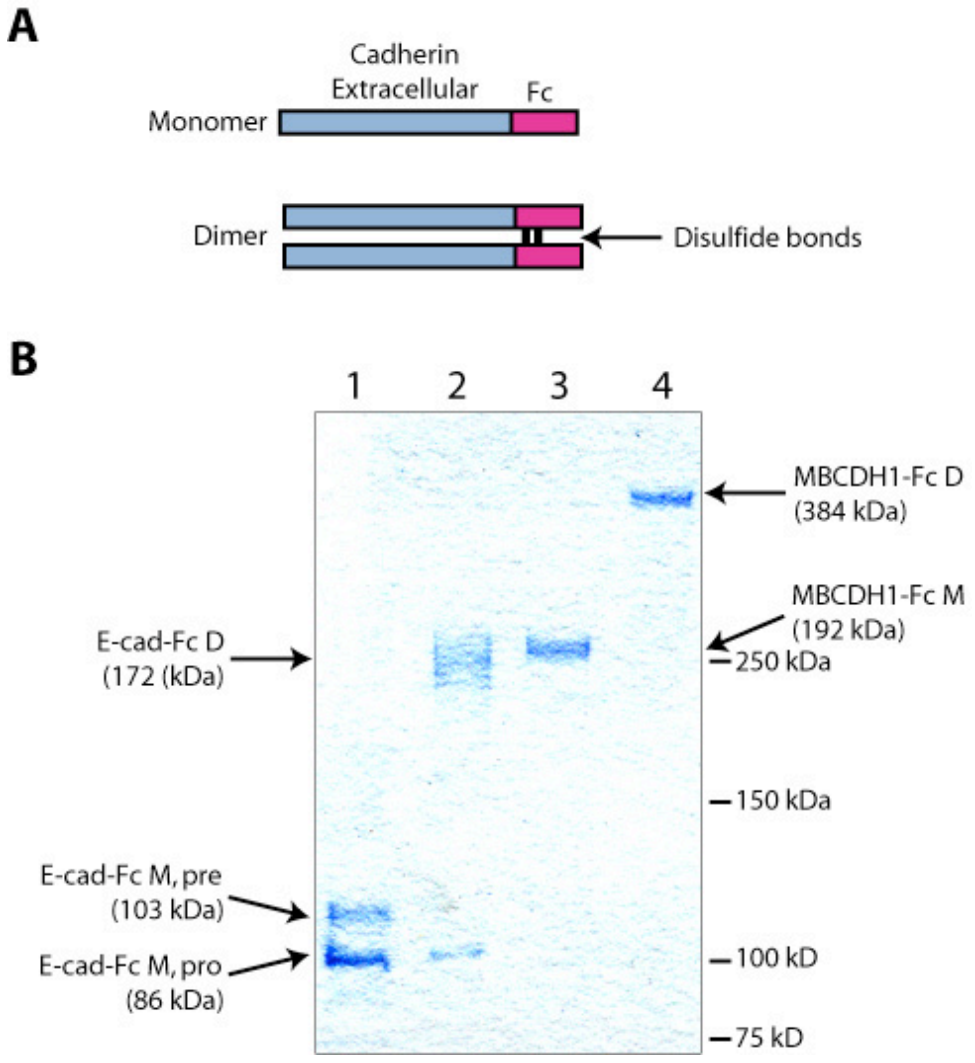


Figure 4.2. MBCDH1 does not mediate homophilic adhesion *in vitro.* Homophilic binding ability was assayed by monitoring the aggregation of beads coated with MBCDH1-Fc monomers (A & B), MBCDH1-Fc dimers (C & D), E-cadherin-Fc monomers (E & F) or E-cadherin-Fc dimers (G & H). MBCDH1-Fc-coated beads incubated in artificial seawater containing an assumed Ca^{2+} concentration of approximately 7 mM in the absence (A, C) or presence of 10 mM EGTA, to chelate Ca^{2+}, (B, D) did not aggregate. In contrast, beads coated with E-cadherin-Fc proteins formed large aggregates in the presence of free Ca2+ (E, G) but not when incubated with 10 mM EGTA (F, H), indicating that homophilic adhesion is calcium dependent. Scale bar: 20 micron.

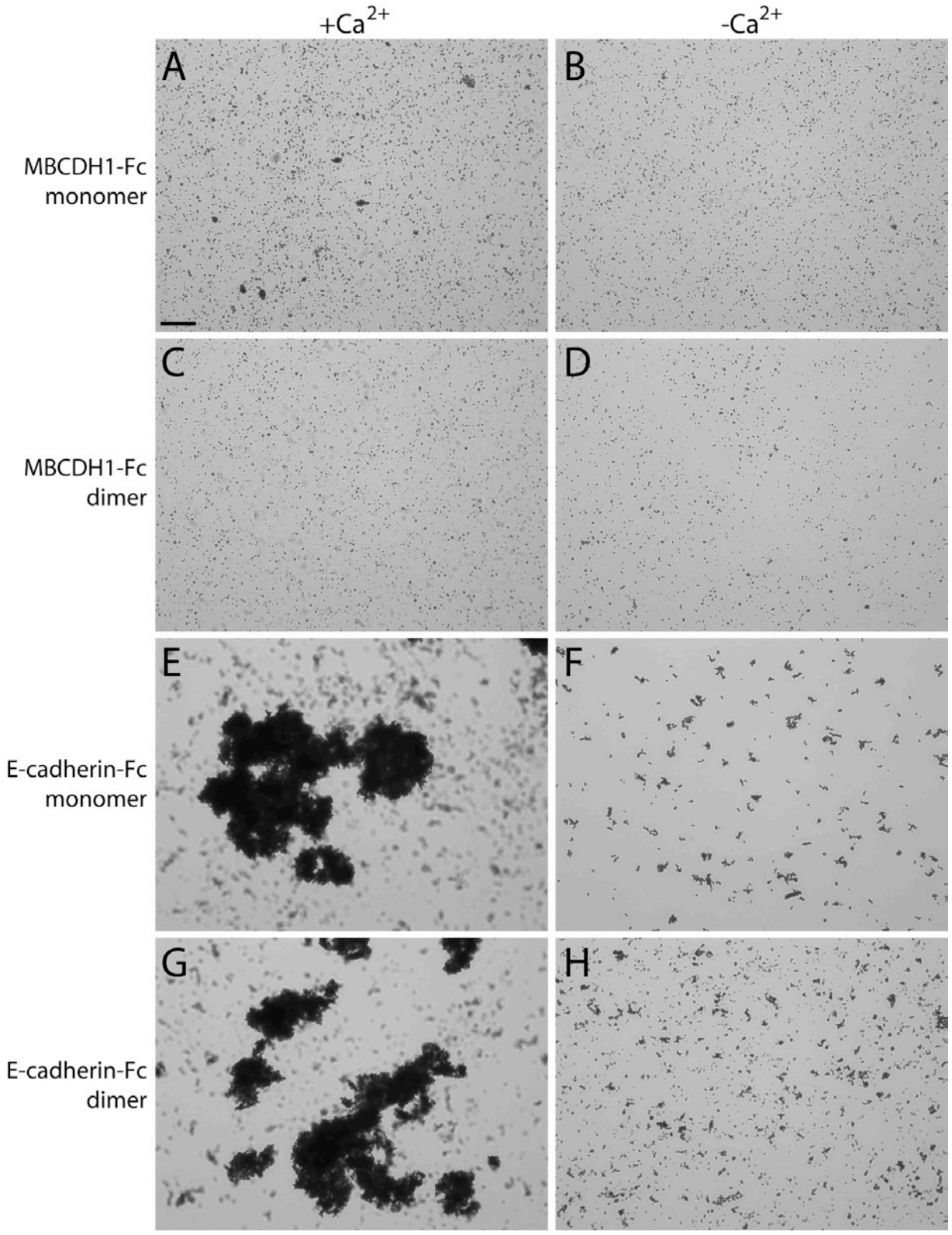

+Ca2+
-Ca2+
A
B
MBCDH1-Fc
monomer
C
D
MBCDH1-Fc
dimer
E
F
E-cadherin-Fc
monomer
G
H
E-cadherin-Fc
dimer

Figure 4.3. L cells express cadherin transgenes. Expression of cadherins was detected by immunoblotting lysate from cell lines stably transduced with the indicated cadherin gene construct. Lysates were probed with either anti-MBCDH1/2 antibodies (A) or anti-E-cadherin antibodes (B). Untransduced L cells show no expression of MBCDH1, but MBCDH1 and MBCDH1-GFP transduced cell lines exhibit prominent expression of the respective transgene. Comparisons lysates from *M. brevicollis* and MBCDH1-expressing L cells indicate that endogenous MBCDH1/2 has a lower molecular weight than MBCDH1 and MBCDH1-GFP expressed in L cells (A). E-cadherin and E-cadherin-GFP are expressed in transduced cell lines and not in untransduced cells (B). Predicted molecular weight based on amino acid sequence: MBCDH1, 192 kDa; MBCDH1-GFP, 219 kDa; E-cadherin, 97 kDa; E-cadherin-GFP, 124 kDa. Posttranslational glycosylation of E-cadherin represents 20% of its actual molecular weight (187).

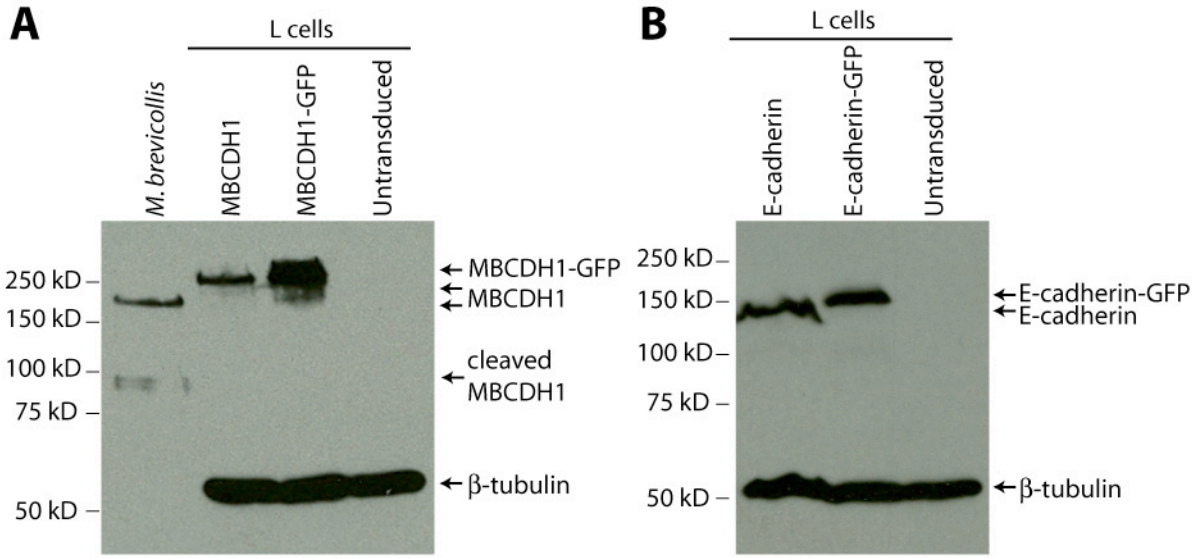


Figure 4.4. Cadherins ectopically expressed in L cells localize to the plasma membrane. Fluorescent detection of GFP shows that E-cadherin-GFP (A) and MBCDH1-GFP (C) are enriched at the cell cortex. Live L cells expressing E-cadherin (B, 'live') or MBCDH1 (D, 'live') probed with indicated antibody to label only cell-surface proteins demonstrate that recombinant cadherins reach the plasma membrane. Immunofluorescent staining of cadherins in fixed L cells (B and D, 'fixed') shows a similar staining pattern, indicating that a high proportion of the total protein is displayed on the surface of the cell. Scale bar: 50 micron, inset scale bar: 20 microns.

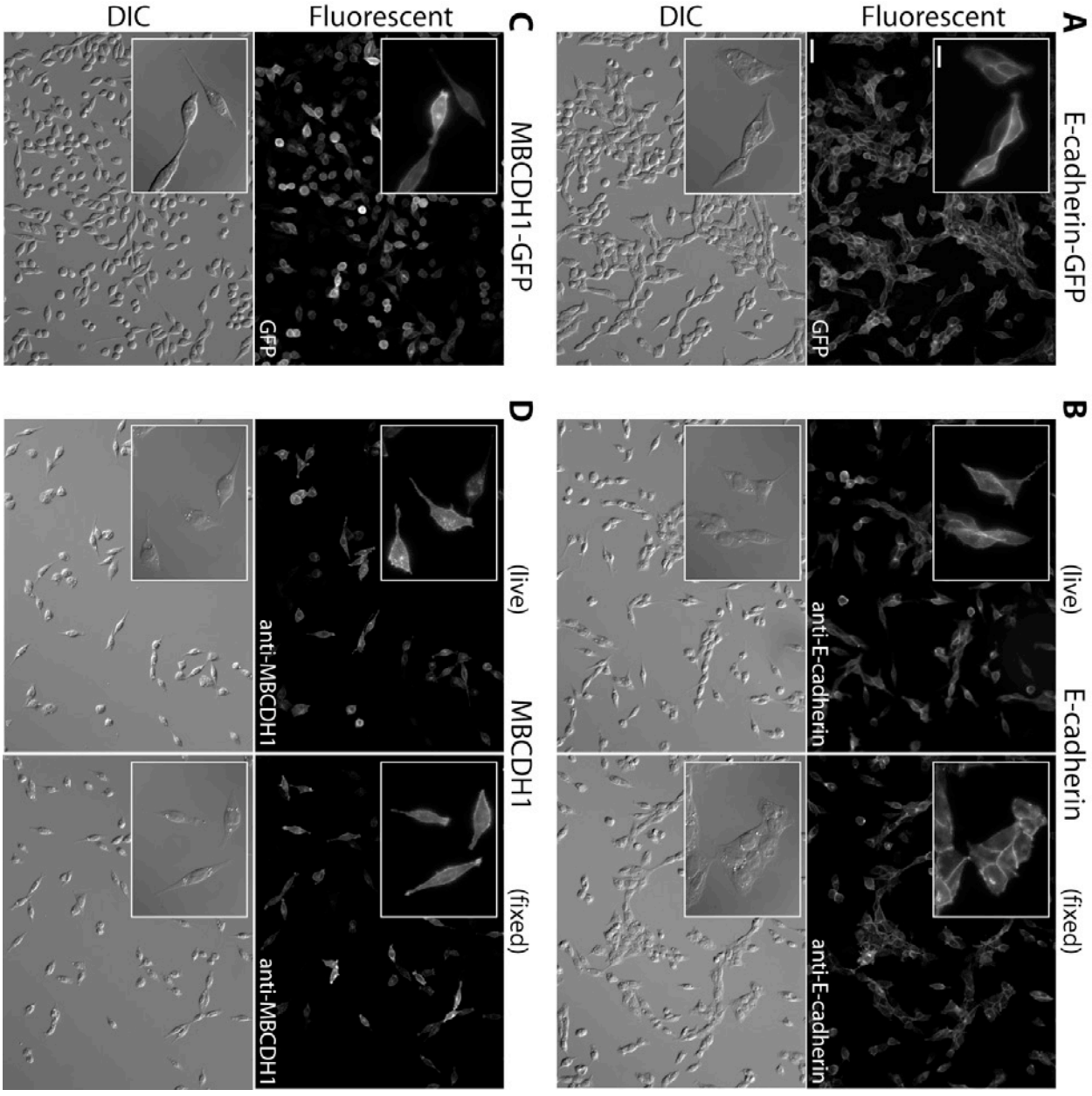
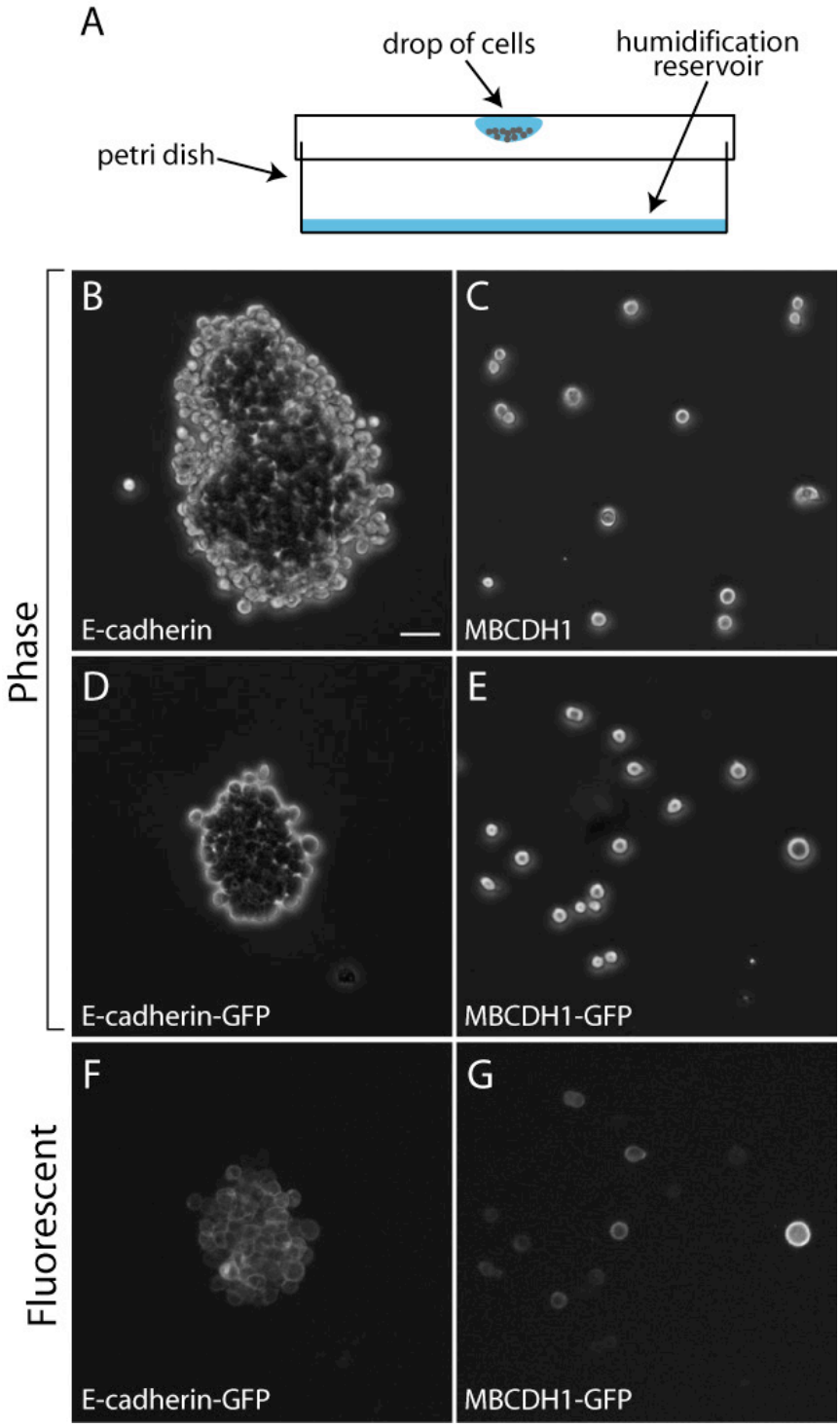


Figure 4.5. L cells expressing MBCDH1 do not form intercellular connections. Adhesion between cells was tested using a hanging drop assay in which cells were suspended from the lid of a petridish in a drop of liquid to maximize intercellular contact (A). Control L cells expressing E-cadherin (B) or E-cadherin-GFP (D and F) formed aggregates when incubated in a drop of culture medium for approximately 16 hours. In contrast, L cells expressing MBCDH1 (C) or MBCDH1-GFP (E and G) did not adhere. Scale bar: 10 microns.



SUPPLAMENTAL FIGURES AND TABLES

Figure S4.1. Membrane localization of cadherin-GFP fusion proteins confirmed by indirect immunofluorescence. E-cadherin-GFP (A) and MBCDH1-GFP (B) L cells immunostained before permeabilization demonstrate that a significant fraction of expressed fusion proteins reaches the surface of the cell. Merged images of GFP fluorescent (green) and IMF (red) show colocalization (yellow) of cadherin at the plasma membrane marked by indicated antibodies and total cadherin labeled with GFP. E-cadherin L cells probed with anti-MBCDH1/2 antibodies and MBCDH1 L cells probed with anti-E-cadherin antibodies (C) show that both antibodies are specific. Scale bar: 50 micron, inset scale bar: 20 microns

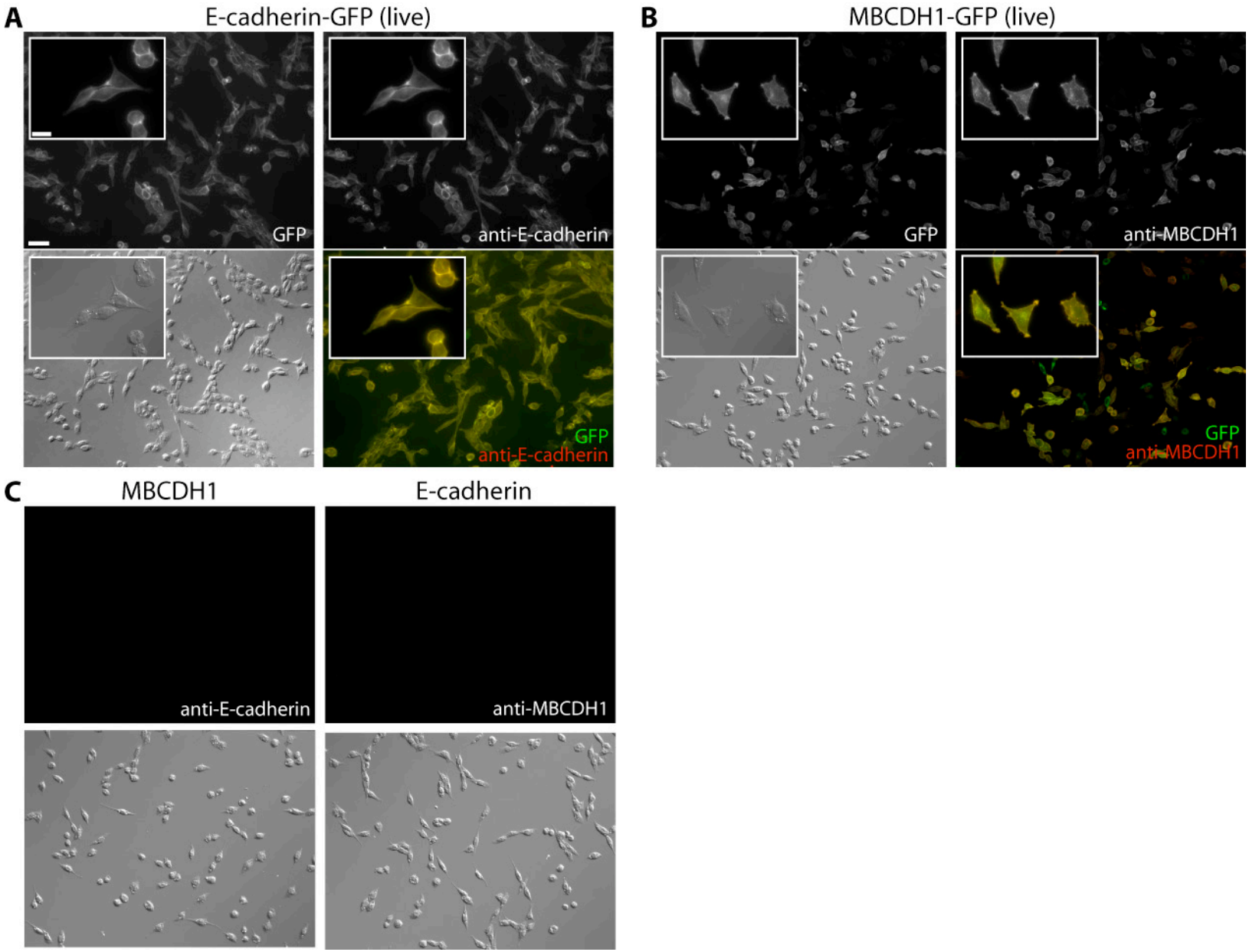
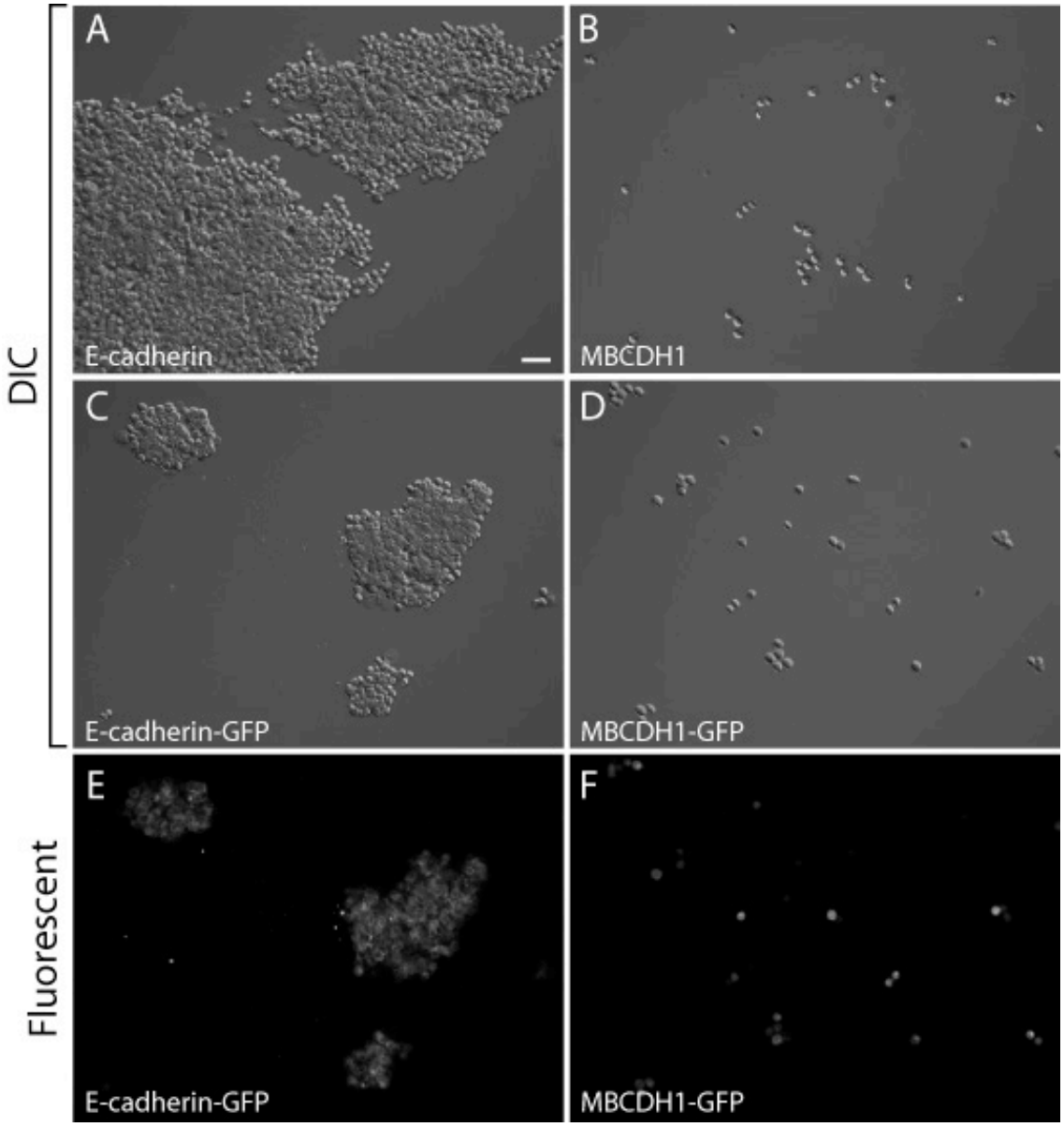

A
E-cadherin-GFP (live)
GFP
anti-E-cadherin
GFP
anti-E-cadherin
B
MBCDH1-GFP (live)
GFP
anti-MBCDH1
GFP
anti-MBCDH1
C
MBCDH1
E-cadherin
anti-E-cadherin
anti-MBCDH1

Figure S4.2. Hanging drop assay in seawater recapitulates results in DMEM. Adhesion between cells after sorting for clones expressing high levels of cadherins was tested using a hanging drop assay. Consistent with experiments performed in DMEM before cell sorting, control L cells expressing E-cadherin (B) or E-cadherin-GFP (D and F) formed aggregates when incubated in a drop of artificial seawater for 2 hours. In contrast, L cells expressing MBCDH1 (C) or MBCDH1-GFP (E and G) did not adhere. Scale bar: 20 microns.



Appendix: Identifying MBCDH1 and MBCDH2 cytoplasmic binding partners

In animals, homo- and heterophilic cadherin interactions elicit specific intracellular responses mediated by multi-protein complexes that form on cadherin cytoplasmic tails (42, 96, 196-198). Identifying intracellular interactors of MBCDH1/2 with defined cellular functions will provide significant insights into the functional role of cadherins in *M. brevicollis* and could potentially identify shared functions between animal and choanoflagellate cadherins. Discovery of previously uncharacterized binding partner may unveil novel functions for choanoflagellate cadherins. To identify MBCDH1/2 interacting partners, I used two techniques, Pull-down and Immunoprecipitation (IP). The experimental outline (Figure *A*1) and the progress that I made in optimizing these techniques are described below.

Experimental overview
The basic method to identify protein-protein interactions (Figure *A*1) involves a bait protein (Pull-down) or antibody (IP) that is used to affinity purify binding partners from a cellular pool of proteins. Binding partners are then eluted from the bait or antibody and detected by SDS-PAGE. Subsequently, Mass Spectrometry (MS) is used to determine protein identity.

Pull-down versus Immunoprecipitation
There are advantages and disadvantages to both Pull-downs and IPs. It is generally thought that Pull-downs have a reduced incidence of contaminating, non-specific proteins in the final sample because target proteins must actively bind to the bait protein. A disadvantage is that interactors can be titrated away from the bait if they are tightly associated with endogenous proteins, making it more difficult to successfully obtain target proteins. In addition, many protein-protein interactions rely on modifications such as phosphorylation, which are likely to be missing from the recombinant bait. IP experiments typically display more background or non-specific proteins in the final sample and thus have a higher risk of falsely identifying an interacting protein. An advantage is that antibody precipitation of endogenous proteins increases the likelihood that binding partners will be purified from cell extracts. A second advantage of IPs is that optimization is somewhat easier because the presence of your protein of interest, in this case MBCDH1/2, in eluted samples can be used as a metric for experimental success.

I first chose to use Pull-downs to identify MBCDH1 interacting proteins. After much optimization with little success, I switched to the IP method. Although my IP experiments did not ultimately lead to identification of MBCDH1/2 binding partners, I made significant experimental progress (Figure *A*4) and believe that with a few more modifications, the IP will succeed.

Step 1: Grow large volumes of choanoflagellates to high density
A major challenge of Pull-downs and IPs is obtaining enough cellular material to detect binding partners by SDS-PAGE and Mass Spectrometry (MS). To grow large volumes of choanoflagellates, I passaged cells every day for approximately 3 days to obtain cultures that were in logarithmic growth and inoculated multiple flasks with the rapidly dividing *M. brevicollis* cultures. The flasks were incubated for approximately 24 hours with gentle agitation (100rpm) before harvesting the cells.

A second hurtle is maximizing choanoflagellate growth while minimizing the bacteria that are co-cultured with choanoflagellates to provide a food source. The bacteria contain high quantities of proteases that complicate subsequent steps in the experiment. After experimenting with differential centrifugation to separate the bacteria from the choanoflagellates, which was inefficient, and filtering, which was extremely labor intensive and caused loss of a significant fraction of cells, I discovered that supplementing cultures with *E. aerogenes* was a good method for obtaining cultures with low levels of bacteria and higher choanoflagellate density (Figure *A*2). Not only do the cells grow to high density because of the added bacteria, but after 24 hours of growth, the bacteria have been almost completely consumed by the choanoflagellates, leaving very few in the water column. Although it is unclear exactly why this method reduces the bacteria, I hypothesize that the *E. aerogenes* added to the culture is easily consumed by the choanoflagellates, because it not yet formed biofilms like bacteria already present in the cultures. This allows the choanoflagellates to grow and rapidly graze down the bacteria so that few are left when the cultures are harvested.

The addition of *E. aerogenes* also decreases levels of MBCDH1/2 (Figure 3.1), so there is a trade-off between achieving dense, bacteria free cultures, and obtaining cells with higher levels of MBCDH1/2. Immunoblotting shows that more MBCDH1/2 is ultimately obtained from cultures when *E. aerogenes* is added (Figure *A*2), likely due to higher *M. breviocollis* cell densties.

Step 2: Prepare whole-cell lysate

The two main techniques for cell lysis are physical and chemical disruption of cell membranes. Because the presence of detergent in samples can negatively affect protein identification by MS, I first attempted to physically lyse the cells. Homogenization using a Dounce Homogenizer, freeze- thaw, and osmotic shock coupled with sheering (by passing cells through a Z-shaped needle) were successful to some degree, but typically left a significant number of cells unlysed. The importance of obtaining large quantities of lysate led me to use the non-ionic detergent NP-40 to lyse the choanoflagellates (Figure *A*2), which could later be removed from the eluted sample by trichloroacetic acid precipitation. In addition to lysis, NP-40 completely solublized the membranes allowing for efficient capture of MBCDH1/2 on antibody-coated beads during the IPs.

Step 3: Conjugate bait to beads

The bait protein is used to affinity purify binding partners from the cell lysates. For the Pull-downs, I purified recombinant, GST-tagged MBCDH1 cytoplasmic tail (GST-MBCDH1cyto) from *E. coli* to use as my bait (199). For the IPs, I used anti-MBCDH1/2 antibodies. I conjugated the bait to a support matrix made of beads that allow for recovery of the bait and any interactors after Step 5. To avoid co-elution of my bait during Step 6, I covalently linked the bait to the support matrix.

Pull-down

I performed initial Pull-down experiments by coupling GST-MBCDH1cyto to glutathione-coated agarose beads (GE Healthcare) and discovered that the abundant bait protein eluted from beads (Step 6) was potentially masking the presence of other proteins. To address this issue, I tried covalently cross-linking GST-MBCDH1cyto to the glutathione beads using the potent

crosslinking agent Glutaraldehyde. This method worked well but is known to mask protein epitopes and could thus affect efficiency of the pull-down. After experimenting with several types of beads, I settled upon an activated sepharose matrix (CH-Sepharose 4B, Sigma) that specifically and irreversibly couples to amine groups present in the bait protein.

Immunoprecipitation
To generate anti-MBCDH1/2-coated beads I used magnetic microspheres conjugated with either protein A and anti-rabbit IgG antibodies (Bangs Laboratories and Invitrogen), which work equally well. Both proteins bind specifically to the Fc region of rabbit IgG such that the antigen-binding portion of the anti-MBCDH1/2 antibodies face outward. To crosslink the beads, I used Dimethylpimelimidate (Peirce), which links amine groups in the bait antibody to the proteins on the beads. I chose to use magnetic beads because they can be rapidly cleared from the water-column, making washing steps much faster and easier.

Optional Step: Pre-clear the lysate
Pre-clearing the lysate reduces the background caused by proteins that bind non-specifically to the protein coated beads.

Pull-down
To remove any proteins from the *M. brevicollis* lysate that interact with GST, I incubated the lysate with GST-coated beads for one hour before moving on to Step 5.

Immunoprecipitation
I pre-cleared the lysate as described above using beads coated with antibodies from the pre-innoculation serum of the rabbit in which the anti-MBCDH1/2 that were generated. This serum was collected before injecting the MBCDH1 antigen into the rabbit and should not have antibodies specific to MBCDH1/2.

Step 4: Incubate beads with whole-cell lysate
In this step, I incubated the beads from step 3 with the lysate prepared in step 2. Initially, I tried performing the bead-lysate binding at room temperature to more closely mimic the *M. brevicollis* cellular environment. However, protease activity was too high and proteins in the sample degraded during the hour incubation at room temperature, despite the fact that I added large quantities of protease inhibitors (Phenylmethylsulfonyl fluoride, Thermo Scientific; Complete Mini, Roche). Because proteases are less active at lower temperatures, I decided to perform the incubation at 4 ^{o}C.

Step 6: Elute protein bound to beads
The basic strategy that I used for elution was to chemically denature the proteins bound to the beads. I initially used an acid solution (100mM Glycine, pH 2.5) but later found that elution by ionic detergent (SDS in SDS-PAGE sample buffer) allowed me to obtain a much more concentrated samples. Interestingly, I discovered that, when using magnetic beads, boiling the beads in SDS-PAGE sample buffer to further denature proteins and reduce disulphide bonds, 100mM 2-Mercaptoethanol before loading onto and a polyacrylamide gel (step 7) caused high levels of background that appeared as a smear throughout the gel lane. Removing the beads

from the sample before adding the reducing agent and boiling abolished the background (Figure *A*3).

Step 7: Identify protein(s) by SDS-PAGE and Mass Spectrometry
After elution of proteins from the bait, I analyzed the samples by SDS-PAGE followed by silver stain to detect low abundance proteins.

Experimental versus Mock
To ensure that proteins present in my eluted sample were specific binding partners of MBCDH1/2, I performed mock experiments along side the actual experiments. For the Pull-downs, I used GST-coated beads as the mock sample, because the MBCDH1 bait protein was conjugated to GST. In this way, I could identify proteins that specifically bound to MBCDH1, as opposed to those that bound to the GST tag. For the IPs, I carried out mock experiments using beads coated with antibodies from rabbit pre-immune serum. This distinguished between proteins that bound non-specifically to the rabbit antibodies, and the target proteins bound to MBCDH1/2. I considered bands on the polyacrylamide gel that were present in both the mock and experimental to be non-specific, and those only in the experimental, to be putative binding partners.

I tried two different types of MS to identify putative MBCDH1 interactors: one-dimensional liquid chromatography MS/MS to sequence a single protein species from a gel band, and Multidimensional Protein Identification Technology for identifying proteins in a complex mixture. I was not able to definitively identify proteins using either method, which I attribute primarily to low protein concentration in my samples.

Current state and future directions
Based on my efforts with both Pull-downs and IPs, I achieved greater success with IPs. I isolated quantities of MBCDH1/2 that are visible on a silver stained polyacrylamide gel (Figure *A*4). This suggests that any associated binding partners may also be present in the sample. However, the non-specific proteins (in both mock and experimental samples) make it difficult to identify putative binding partners. Thus, reducing background is will be critical for achieving a successful MBCDH1/2 IP. A useful method for background reduction is addition of the anionic detergent, Sodium Dodecyl Sulfate (SDS) to the lysate. The negatively charged SDS molecules block proteins from sticking to one another. If SDS is added to lysis buffer, the cells must be reversibly cross-linked before lysis to maintain associations between MBCDH1/2 and its binding partners, which could potentially be disrupted by the SDS. I have optimized cross-linking using Dithiobis (succinimidyl propionate) (Pierce) and confirmed that the anti-MBCDH1/2 antibodies are functional in the presence of 0.1% SDS.

A second important aspect of the IP that has not been completely achieved is obtaining adequate amounts of cellular material. As mentioned, I can isolate visible quantities of MBCDH1/2 using cell lysate from 8 L of culture (Figure *A*4). Ideally, one would hope to visualize interacting proteins by Coomassie stain, which would indicate that the protein is abundant and would increase the likelihood of determining protein identity by MS. Doubling the volume of cell culture used for the experiment will most likely accomplish this goal.

The progress I have made on MBCH1/2 IPs is promising. Reducing background and increasing cell culture volume will potentially lead to identification of proteins that interact with the cytoplasmic tail of MBCDH1/2. These data will shed light on the functions of choanoflagellate cadherins and provide new directions for future research efforts.

FIGURES

Figure *A*1. Flowchart depicting experimental procedure.

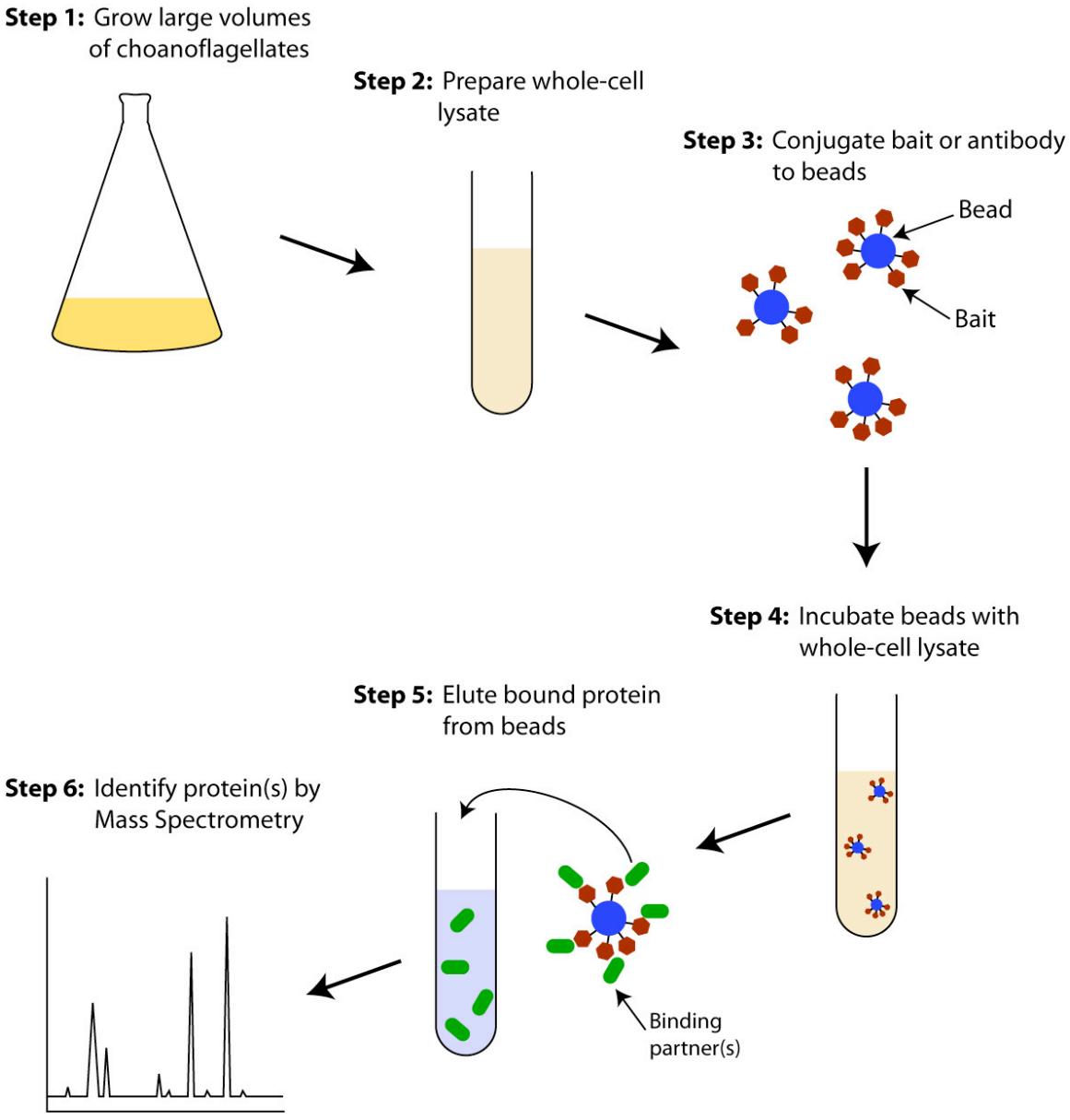


Figure *A*2. Method of *M. brevicollis* growth and lysis affects MBCDH1/2 abundance in whole-cell extracts. Addition of *E. aerogenes* (+*E.a.*) to *M. brevicollis* cultures (lanes 3 and 4 versus 1 and 2) and NP-40 detergent during cell lysis (lanes 2 and 4 compared to 1 and 3) resulted in higher levels of MBCDH1/2 protein in cell extracts. Immunoblot of lysate obtained from *M. brevicollis,* cultured with out added *E. aerogenes,* followed by filtering through 8 μm filter to remove bacterial biofilms (lanes 1 and 2) shows that addition of NP-40 during lysis (lane 1 compared to 2) releases MBCDH1/2 into cell extracts. When *E.a.* was added to *M. brevicollis* cultures 24 hours before harvesting cells, more *M. brevicollis* and less bacteria were obtained, leading to higher levels of MBCDH1/2 in cell lysate (lanes 3 and 4). Again, addition of NP-40 increased release of MBCDH1/2 into cell lysate (lanes 3 vs. 4).

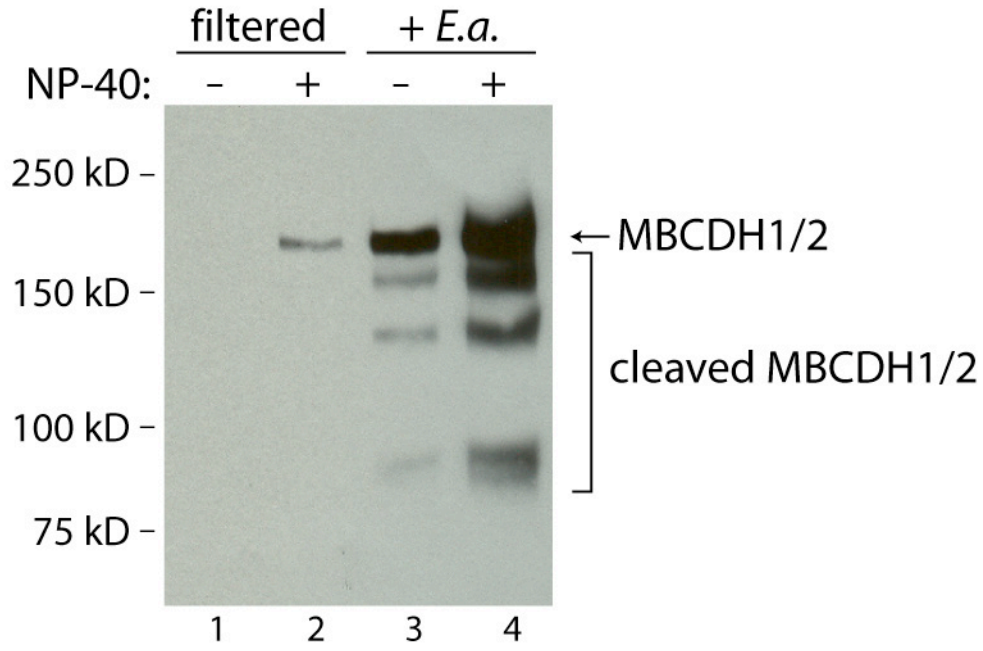


Figure *A*3. Boiling beads causes high levels of background on SDS-PAGE polyacrylamide gel. Magnetic beads coated with anti-MBCDH1/2 antibodies covalently linked to anti-rabbit IgG antibodies were suspended in SDS-PAGE sample buffer. The sample was split in two and the beads were removed from one of the samples (- Beads) and left in the other sample (+ Beads). 100mM 2-Mercaptoethanol was added to the samples before boiling for 10 minutes. Samples were analyzed by silver stain after SDS-PAGE.

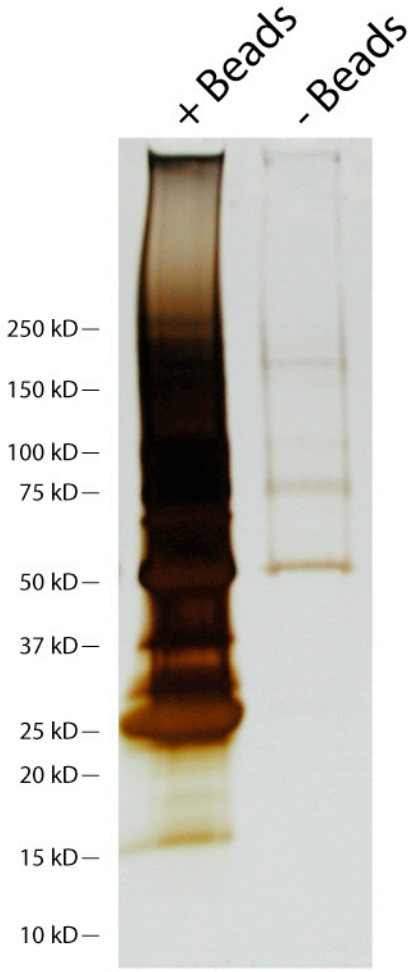
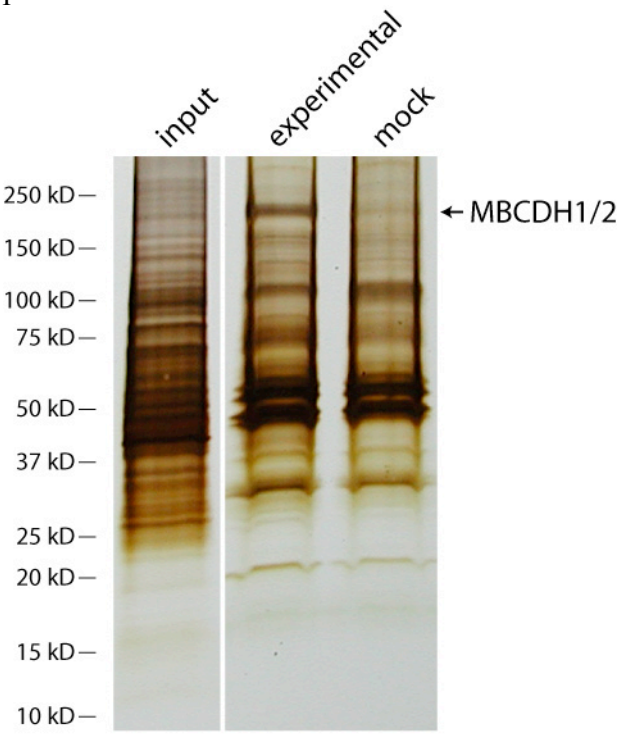


Figure *A*4. Immunoprecipitation extracts MBCDH1/2 from *M. brevicollis* whole-cell lysate. Whole-cell lysates prepared by lysing *M. brevicollis* with buffer containing NP-40 (input), were incubated with anti-MBCDH1/2 antibodies (experimental) or antibodies from rabbit pre-immune serum (mock). Protein was eluted from beads by incubation in SDS-PAGE sample buffer and protein content was analyzed by SDS-PAGE followed by silver stain. A brightly staining band that corresponds to MBCDH1/2 (192 kDa) is present in the experimental, but not the mock sample, indicating that the anti-MBCDH1/2 antibodies specifically precipitated MBCDH1/2 protein.



Bibliography

1. King, N., (2004). The unicellular ancestry of animal development. *Dev. Cell*, 7, 313-25.
2. Baldauf, S.L., (2003). The deep roots of eukaryotes. *Science*, 300, 1703-6.
3. Carroll, S.B., (2001). Chance and necessity: the evolution of morphological complexity and diversity. *Nature*, 409, 1102-9.
4. Medina, M., Collins, A.G., Taylor, J.W., Valentine, J.W., Lipps, J.H., Amaral-Zettler, L., and Sogin, M.L., (2003). Phylogeny of Opisthokonta and the evolution of multicellularity and complexity in Fungi and Metazoa. *Int. J. Astrobiol.*, 2, 203-11.
5. King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., Fairclough, S., Hellsten, U., Isogai, Y., Letunic, I., Marr, M., Pincus, D., Putnam, N., Rokas, A., Wright, K.J., Zuzow, R., Dirks, W., Good, M., Goodstein, D., Lemons, D., Li, W., Lyons, J.B., Morris, A., Nichols, S., Richter, D.J., Salamov, A., Sequencing, J.G., Bork, P., Lim, W.A., Manning, G., Miller, W.T., Mcginnis, W., Shapiro, H., Tjian, R., Grigoriev, I.V., and Rokhsar, D., (2008). The genome of the choanoflagellate Monosiga brevicollis and the origin of metazoans. *Nature*, 451, 783-8.
6. Nozaki, H., Misawa, K., Kajita, T., Kato, M., Nohara, S., and Watanabe, M.M., (2000). Origin and evolution of the colonial volvocales (Chlorophyceae) as inferred from multiple, chloroplast gene sequences. *Mol. Phylogenet. Evol.*, 17, 256-68.
7. Tyler, S., (2003). Epithelium—The Primary Building Block for Metazoan Complexity. *Integr. Comp. Biol.*, 43, 55-63.
8. Pokutta, S. and Weis, W.I., (2007). Structure and mechanism of cadherins and catenins in cell-cell contacts. *Annu. Rev. Cell Dev. Biol.*, 23, 237-61.
9. Boury-Esnault, N., Ereskovsky, A., Bezac, C., Tokin, D., (2003). Larval development in the Homoscleromorpha (Porifera, Demospongiae) *Invertebr. Biol.*, 122, 187-202.
10. Onur Sakarya, K.A.A., Maja Adamska, Marcin Adamski, I-Fan Wang, Bruce Tidor, Bernard M. Degnan, Todd H. Oakley, Kenneth S. Kosik, (2007). A Post-Synaptic Scaffold at the Origin of the Animal Kingdom. *PLoS ONE*, 1-9.
11. Abedin, M. and King, N., (2008). The premetazoan ancestry of cadherins. *Science*, 319, 946-8.
12. Spiegel, I. and Peles, E., (2002). Cellular junctions of myelinated nerves (Review). *Mol. Membr. Biol.*, 19, 95-101.
13. Bing-Yu, M., Xiao-Yang, S., Hong-Wei, Z., Shi-Cui, Z., Xian-Han, W., (1997). Morphological and functional studies on the epidermal cells of amphioxus (Branchiostoma belcheri tsingtauense) at different developmental stages. *Chin. J. Oceanol. Limnol.*, 15, 236-41.
14. Banerjee, S., Sousa, A.D., and Bhat, M.A., (2006). Organization and function of septate junctions: an evolutionary perspective. *Cell. Biochem. Biophys.*, 46, 65-77.
15. Nichols, S.A., Dirks, W., Pearse, J.S., and King, N., (2006). Early evolution of animal cell signaling and adhesion genes. *Proc. Natl. Acad. Sci. U. S. A.*, 103, 12451-6.
16. Ledger, P.W., (1975). Septate junctions in the calcareous sponge Sycon ciliatum. *Tissue Cell*, 7, 13-8.
17. Leys, S.P., Nichols, S.A., Adams, E.D.M., (2009). Epithelia and integration in sponges. *Integr. Comp. Biol.*, 49, 167-77.
18. Shestopalov, V.I. and Panchin, Y., (2008). Pannexins and gap junction protein diversity. *Cell. Mol. Life Sci.*, 65, 376-94.

19. Westfall, J.A., Kinnamon, J.C., and Sims, D.E., (1980). Neuro-epitheliomuscular cell and neuro-neuronal gap junctions in Hydra. *J. Neurocytol.*, 9, 725-32.
20. Litvin, O., Tiunova, A., Connell-Alberts, Y., Panchin, Y., and Baranova, A., (2006). What is hidden in the pannexin treasure trove: the sneak peek and the guesswork. *J. Cell Mol. Med.*, 10, 613-34.
21. Sasakura, Y., Shoguchi, E., Takatori, N., Wada, S., Meinertzhagen, I.A., Satou, Y., and Satoh, N., (2003). A genomewide survey of developmentally relevant genes in Ciona intestinalis. X. Genes for cell junctions and extracellular matrix. *Dev. Genes Evol.*, 213, 303-13.
22. Lane, N.J., Dallai, R., Burighel, P., and Martinucci, G.B., (1986). Tight and gap junctions in the intestinal tract of tunicates (Urochordata): a freeze-fracture study. *J. Cell Sci.*, 84, 1-17.
23. Lane, N.J., Dallai, R., Martinucci, G.B., and Burighel, P., (1987). Cell junctions in amphioxus (Cephalochordata): a thin section and freeze-fracture study. *Tissue Cell*, 19, 399-411.
24. Anderson, J.M., Van Itallie, C.M., and Fanning, A.S., (2004). Setting up a selective barrier at the apical junction complex. *Curr. Opin. Cell Biol.*, 16, 140-5.
25. Wu, V.M., Schulte, J., Hirschi, A., Tepass, U., and Beitel, G.J., (2004). Sinuous is a Drosophila claudin required for septate junction organization and epithelial tube size control. *J. Cell Biol.*, 164, 313-23.
26. Behr, M., Riedel, D., and Schuh, R., (2003). The claudin-like megatrachea is essential in septate junctions for the epithelial barrier function in Drosophila. *Dev. Cell*, 5, 611-20.
27. King, N., Hittinger, C.T., and Carroll, S.B., (2003). Evolution of key cell signaling and adhesion protein families predates animal origins. *Science*, 301, 361-3.
28. Bonner, J.T., (1997). The Origins of Multicellularity. *Int. Biol.*, 27-36.
29. Nelson, W.J., (2008). Regulation of cell-cell adhesion by the cadherin-catenin complex. *Biochem. Soc. Trans.*, 36, 149-55.
30. Brar, S.K. and Siu, C.H., (1993). Characterization of the cell adhesion molecule gp24 in Dictyostelium discoideum. Mediation of cell-cell adhesion via a Ca(2+)-dependent mechanism. *J. Biol. Chem.*, 268, 24902-9.
31. Sesaki, H. and Siu, C.H., (1996). Novel redistribution of the Ca(2+)-dependent cell adhesion molecule DdCAD-1 during development of Dictyostelium discoideum. *Dev. Biol.*, 177, 504-16.
32. Gumbiner, B.M., (1996). Cell adhesion: the molecular basis of tissue architecture and morphogenesis. *Cell*, 84, 345-57.
33. Halbleib, J.M. and Nelson, W.J., (2006). Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. *Genes Dev.*, 20, 3199-214.
34. Siu, C.H., Harris, T.J., Wang, J., and Wong, E., (2004). Regulation of cell-cell adhesion during Dictyostelium development. *Semin. Cell Dev. Biol.*, 15, 633-41.
35. Wang, J., Hou, L., Awrey, D., Loomis, W.F., Firtel, R.A., and Siu, C.H., (2000). The membrane glycoprotein gp150 is encoded by the lagC gene and mediates cell-cell adhesion by heterophilic binding during Dictyostelium development. *Dev. Biol.*, 227, 734-45.
36. Wong, E., Yang, C., Wang, J., Fuller, D., Loomis, W.F., and Siu, C.H., (2002). Disruption of the gene encoding the cell adhesion molecule DdCAD-1 leads to aberrant

cell sorting and cell-type proportioning during Dictyostelium development. *Development*, 129, 3839-50.
37. Nicol, A., Rappel, W., Levine, H., and Loomis, W.F., (1999). Cell-sorting in aggregates of Dictyostelium discoideum. *J. Cell Sci.*, 112 (Pt 22), 3923-9.
38. Steinberg, M.S., (2007). Differential adhesion in morphogenesis: a modern view. *Curr. Opin. Genet. Dev.*, 17, 281-6.
39. Sallee, J.L., Wittchen, E.S., and Burridge, K., (2006). Regulation of cell adhesion by protein-tyrosine phosphatases: II. Cell-cell adhesion. *J. Biol. Chem.*, 281, 16189-92.
40. Secko, D.M., Siu, C.H., Spiegelman, G.B., and Weeks, G., (2006). An activated Ras protein alters cell adhesion by dephosphorylating Dictyostelium DdCAD-1. *Microbiology*, 152, 1497-505.
41. Grimson, M.J., Coates, J.C., Reynolds, J.P., Shipman, M., Blanton, R.L., and Harwood, A.J., (2000). Adherens junctions and beta-catenin-mediated cell signalling in a non-metazoan organism. *Nature*, 408, 727-31.
42. Nelson, W.J. and Nusse, R., (2004). Convergence of Wnt, Beta-catenin and cadherin pathways. *Science*, 303, 1483-7.
43. Grosberg, R.K., Strathmann, R.R., (2007). The Evolution of Multicellularity: A Minor Major Transition? . *Annu. Rev. Ecol. Evol. Syst.* , 38, 621-54.
44. Nakada, T., Misawa, K., and Nozaki, H., (2008). Molecular systematics of Volvocales (Chlorophyceae, Chlorophyta) based on exhaustive 18S rRNA phylogenetic analyses. *Mol. Phylogenet. Evol.*, 48, 281-91.
45. Kirk, D.L., (2005). A twelve-step program for evolving multicellularity and a division of labor. *Bioessays*, 27, 299-310.
46. Woessner, J.P., Goodenough, U.W., (1994). Volvocine cell walls and their constituent glycoproteins: an evolutionary perspective. *Protoplasma*, 1-14.
47. Fulton, A.B., (1978). Colonial development in Pandorina morum. II. Colony morphogenesis and formation of the extracellular matrix. *Dev. Biol.*, 64, 236-51.
48. Adair, W.S., Steinmetz, S.A., Mattson, D.M., Goodenough, U.W., and Heuser, J.E., (1987). Nucleated assembly of Chlamydomonas and Volvox cell walls. *J. Cell Biol.*, 105, 2373-82.
49. Adair, W.S., Appel, H., (1989). Identification of a highly conserved hydroxyproline-rich glycoprotein in the cell wall of Chlamydomonas reinhardtii and two other Volvocales. *Planta*, 179, 381-6.
50. Hallmann, A. and Kirk, D.L., (2000). The developmentally regulated ECM glycoprotein ISG plays an essential role in organizing the ECM and orienting the cells of Volvox. *J. Cell Sci.*, 113 Pt 24, 4605-17.
51. Woessner, J.P., Molendijk, A.J., Van Egmond, P., Klis, F.M., Goodenough, U.W., and Haring, M.A., (1994). Domain conservation in several volvocalean cell wall proteins. *Plant Mol. Biol.*, 26, 947-60.
52. Hallmann, A., (2003). Extracellular matrix and sex-inducing pheromone in Volvox. *Int. Rev. Cytol.*, 227, 131-82.
53. Hallmann, A., (2006). The pherophorins: common, versatile building blocks in the evolution of extracellular matrix architecture in Volvocales. *Plant. J.*, 45, 292-307.
54. Willats, W.G., Orfila, C., Limberg, G., Buchholt, H.C., Van Alebeek, G.J., Voragen, A.G., Marcus, S.E., Christensen, T.M., Mikkelsen, J.D., Murray, B.S., and Knox, J.P., (2001). Modulation of the degree and pattern of methyl-esterification of pectic

homogalacturonan in plant cell walls. Implications for pectin methyl esterase action, matrix properties, and cell adhesion. *J. Biol. Chem.*, 276, 19404-13.
55. Klis, F.M., Ram, A.F.J., Degroot, P.W.J., (2007). A Molecular and Genomic View of the Fungal Cell Wall. . *Biology of the Fungal Cell, 2nd Edition, The Mycota VIII*, R.J. Howard and N.A.R. Gow (Eds.), 97-112.
56. Marshall, C.R. and Valentine, J.W., (2009). The Importance of Preadapted Genomes in the Origin of the Animal Bodyplans and the Cambrian Explosion. *Evolution*, 64, 1189-201.
57. Monod, J., (1971). *Chance and Necessity,*, Vintage Books
58. Lane, N.J. and Chandler, H.J., (1980). Definitive evidence for the existence of tight junctions in invertebrates. *J. Cell Biol.*, 86, 765-74.
59. Fritzenwanker, J.H., Genikhovich, G., Kraus, Y., and Technau, U., (2007). Early development and axis specification in the sea anemone Nematostella vectensis. *Dev. Biol.*, 310, 264-79.
60. Magie, C.R. and Martindale, M.Q., (2008). Cell-cell adhesion in the cnidaria: insights into the evolution of tissue morphogenesis. *Biol Bull*, 214, 218-32.
61. Nollet, F., Kools, P., Roy, F., (2000). Pylogenetic analysis of the cadherin superfamily allows identification of six major subfamilies besides several solitary members. *J. Mol. Biol.*, 229, 551-72.
62. Mitic, L.L., Van Itallie, C.M., and Anderson, J.M., (2000). Molecular physiology and pathophysiology of tight junctions I. Tight junction structure and function: lessons from mutant animals and proteins. *Am J Physiol Gastrointest Liver Physiol*, 279, G250-4.
63. Cruciani, V. and Mikalsen, S.O., (2006). The vertebrate connexin family. *Cell Mol Life Sci*, 63, 1125-40.
64. Halbleib, J.M. and Nelson, W.J., (2006). Cadherins in development: cell adhesion, sorting, and tissue morphogenesis. *Genes Dev*, 20, 3199-214.
65. Gumbiner, B.M., (2005). Regulation of cadherin-mediated adhesion in morphogenesis. *Nat Rev Mol Cell Biol*, 6, 622-34.
66. King, N., (2004). The unicellular ancestry of animal development. *Dev Cell*, 7, 313-25.
67. Tyler, S., (2003). Epithelium-the primary building blocks for metazoan complexity. *Integr Comp Biol*, 43, 55-63.
68. Brooke, N.M. and Holland, P.W., (2003). The evolution of multicellularity and early animal genomes. *Curr Opin Genet Dev*, 13, 599-603.
69. Steenkamp, E.T., Wright, J., and Baldauf, S.L., (2006). The protistan origins of animals and fungi. *Mol Biol Evol*, 23, 93-106.
70. Yoshida-Noro, C. and Takeichi, M., (1982). Teratocarcinoma cell adhesion: identification of a cell-surface protein involved in calcium-dependent cell aggregation. *Cell*, 28, 217-24.
71. Hulpiau, P. and Van Roy, F., (2009). Molecular evolution of the cadherin superfamily. *Int J Biochem Cell Biol*, 41, 349-69.
72. Pokutta, S., Herrenknecht, K., Kemler, R., and Engel, J., (1994). Conformational changes of the recombinant extracellular domain of E-cadherin upon calcium binding. *Eur J Biochem*, 223, 1019-26.
73. Perez, T.D. and Nelson, W.J., (2003). Cadherin adhesion: mechanism and molecular interactions. In: Handbook of Experimental Pharmacology. *Cell Adhesion (Ed. J. Behrens, W. J Nelson)*.

74. Bilder, D., Schober, M., and Perrimon, N., (2003). Integrated activity of PDZ protein complexes regulates epithelial polarity. *Nat Cell Biol*, 5, 53-8.
75. Tanentzapf, G. and Tepass, U., (2003). Interactions between the crumbs, lethal giant larvae and bazooka pathways in epithelial polarization. *Nat Cell Biol*, 5, 46-52.
76. Nelson, W.J., (2003). Adaptation of core mechanisms to generate cell polarity. *Nature*, 422, 766-74.
77. Uchida, N., Honjo, Y., Johnson, K.R., Wheelock, M.J., and Takeichi, M., (1996). The catenin/cadherin adhesion system is localized in synaptic junctions bordering transmitter release zones. *J Cell Biol*, 135, 767-79.
78. Iwai, Y., Usui, T., Hirano, S., Steward, R., Takeichi, M., and Uemura, T., (1997). Axon patterning requires DN-cadherin, a novel neuronal adhesion receptor, in the Drosophila embryonic CNS. *Neuron*, 19, 77-89.
79. Ganzler-Odenthal, S.I. and Redies, C., (1998). Blocking N-cadherin function disrupts the epithelial structure of differentiating neural tissue in the embryonic chicken brain. *J Neurosci*, 18, 5415-25.
80. Yagi, T., (2008). Clustered protocadherin family. *Dev Growth Differ*, 50 Suppl 1, S131-40.
81. Frank, M. and Kemler, R., (2002). Protocadherins. *Curr Opin Cell Biol*, 14, 557-62.
82. Yoshida-Noro, C., Suzuki, N, Takeichi, M., (1984). Molecular nature of the calcium-dependent cell-cell adhesion system in mouse teratocarcinoma and embryonic cells studied with a monoclonal antibody. *Dev. Biol.*, 101, 19-27.
83. Zhong, Y., Brieher, W.M., and Gumbiner, B.M., (1999). Analysis of C-cadherin regulation during tissue morphogenesis with an activating antibody. *J Cell Biol*, 144, 351-9.
84. Kim, S.H., Jen, W.C., De Robertis, E.M., and Kintner, C., (2000). The protocadherin PAPC establishes segmental boundaries during somitogenesis in xenopus embryos. *Curr Biol*, 10, 821-30.
85. Godt, D. and Tepass, U., (1998). Drosophila oocyte localization is mediated by differential cadherin-based adhesion. *Nature*, 395, 387-91.
86. Price, S.R., De Marco Garcia, N.V., Ranscht, B., and Jessell, T.M., (2002). Regulation of motor neuron pool sorting by differential expression of type II cadherins. *Cell*, 109, 205-16.
87. Borghi, N. and James Nelson, W., (2009). Intercellular adhesion in morphogenesis: molecular and biophysical considerations. *Curr Top Dev Biol*, 89, 1-32.
88. Zhang, L., Yue, T., and Jiang, J., (2009). Hippo signaling pathway and organ size control. *Fly (Austin)*, 3, 68-73.
89. Matakatsu, H. and Blair, S.S., (2006). Separating the adhesive and signaling functions of the Fat and Dachsous protocadherins. *Development*, 133, 2315-24.
90. Bennett, F.C. and Harvey, K.F., (2006). Fat cadherin modulates organ size in Drosophila via the Salvador/Warts/Hippo signaling pathway. *Curr Biol*, 16, 2101-10.
91. Silva, E., Tsatskis, Y., Gardano, L., Tapon, N., and Mcneill, H., (2006). The tumor-suppressor gene fat controls tissue growth upstream of expanded in the hippo signaling pathway. *Curr Biol*, 16, 2081-9.
92. Cho, E., Feng, Y., Rauskolb, C., Maitra, S., Fehon, R., and Irvine, K.D., (2006). Delineation of a Fat tumor suppressor pathway. *Nat Genet*, 38, 1142-50.

93. Willecke, M., Hamaratoglu, F., Kango-Singh, M., Udan, R., Chen, C.L., Tao, C., Zhang, X., and Halder, G., (2006). The fat cadherin acts through the hippo tumor-suppressor pathway to regulate tissue size. *Curr Biol*, 16, 2090-100.
94. Willecke, M., Hamaratoglu, F., Sansores-Garcia, L., Tao, C., and Halder, G., (2008). Boundaries of Dachsous Cadherin activity modulate the Hippo signaling pathway to induce cell proliferation. *Proc Natl Acad Sci U S A*, 105, 14897-902.
95. Rogulja, D., Rauskolb, C., and Irvine, K.D., (2008). Morphogen control of wing growth through the Fat signaling pathway. *Dev Cell*, 15, 309-21.
96. Matakatsu, H. and Blair, S.S., (2004). Interactions between Fat and Dachsous and the regulation of planar cell polarity in the Drosophila wing. *Development*, 131, 3785-94.
97. Simons, M. and Mlodzik, M., (2008). Planar cell polarity signaling: from fly development to human disease. *Annu Rev Genet*, 42, 517-40.
98. Sopko, R. and Mcneill, H., (2009). The skinny on Fat: an enormous cadherin that regulates cell adhesion, tissue growth, and planar cell polarity. *Curr Opin Cell Biol*, 21, 717-23.
99. Simon, M.A., (2004). Planar cell polarity in the Drosophila eye is directed by graded Four-jointed and Dachsous expression. *Development*, 131, 6175-84.
100. Usui, T., Shima, Y., Shimada, Y., Hirano, S., Burgess, R.W., Schwarz, T.L., Takeichi, M., and Uemura, T., (1999). Flamingo, a seven-pass transmembrane cadherin, regulates planar cell polarity under the control of Frizzled. *Cell*, 98, 585-95.
101. Lawrence, P.A., Struhl, G., and Casal, J., (2007). Planar cell polarity: one or two pathways? *Nat Rev Genet*, 8, 555-63.
102. Casal, J., Lawrence, P.A., and Struhl, G., (2006). Two separate molecular systems, Dachsous/Fat and Starry night/Frizzled, act independently to confer planar cell polarity. *Development*, 133, 4561-72.
103. Saburi, S., Hester, I., Fischer, E., Pontoglio, M., Eremina, V., Gessler, M., Quaggin, S.E., Harrison, R., Mount, R., and Mcneill, H., (2008). Loss of Fat4 disrupts PCP signaling and oriented cell division and leads to cystic kidney disease. *Nat Genet*, 40, 1010-5.
104. Curtin, J.A., Quint, E., Tsipouri, V., Arkell, R.M., Cattanach, B., Copp, A.J., Henderson, D.J., Spurr, N., Stanier, P., Fisher, E.M., Nolan, P.M., Steel, K.P., Brown, S.D., Gray, I.C., and Murdoch, J.N., (2003). Mutation of Celsr1 disrupts planar polarity of inner ear hair cells and causes severe neural tube defects in the mouse. *Curr Biol*, 13, 1129-33.
105. Leadbeater, B.S. and Kelly, M., (2001). Evolution of animals - choanoflagellates and sponges. *Water and Atmosphere*, 9, 9-11.
106. James-Clark, H., (1868). On the spongiae ciliatae as infusoria flagellata; or observations on the structure, animality, and relationships of Leucosolenia botrydoides, Bowerbank. *Ann Mag Natl Hist*, 1, 133-42; 88-215; 50-64.
107. Snell, E.A., Furlong, R.F., and Holland, P.W., (2001). Hsp70 sequences indicate that choanoflagellates are closely related to animals. *Curr Biol*, 11, 967-70.
108. Lang, B.F., O'kelly, C., Nerad, T., Gray, M.W., and Burger, G., (2002). The closest unicellular relatives of animals. *Curr Biol*, 12, 1773-8.
109. Carr, M., Leadbeater, B.S., Hassan, R., Nelson, M., and Baldauf, S.L., (2008). Molecular phylogeny of choanoflagellates, the sister group to Metazoa. *Proc Natl Acad Sci U S A*, 105, 16641-6.
110. Ruiz-Trillo, I., Roger, A.J., Burger, G., Gray, M.W., and Lang, B.F., (2008). A phylogenomic investigation into the origin of metazoa. *Mol Biol Evol*, 25, 664-72.

111. Burger, G., Forget, L., Zhu, Y., Gray, M.W., and Lang, B.F., (2003). Unique mitochondrial genome architecture in unicellular relatives of animals. *Proc Natl Acad Sci U S A*, 100, 892-7.
112. Lavrov, D.V., Forget, L., Kelly, M., Lang, B.F., (2005). Mitochondrial genomes of two demosponges provide insights into an early stage of animal evolution. *Mol. Biol. Evol*, 22, 1231-9.
113. Rokas, A., Kruger, D., and Carroll, S.B., (2005). Animal evolution and the molecular signature of radiations compressed in time. *Science*, 310, 1933-8.
114. (2006). *JGI M. brevicollis v1.0 Home. http://genome.jgi-psf.org/Monbr1/Monbr1.home.html*.
115. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J., (1990). Basic local alignment search tool. *J Mol Biol*, 215, 403-10.
116. (2006). *JGI N. vectensis v1.0 Home. http://genome.jgi-psf.org/Nemve1/Nemve1.home.html*.
117. (2005). *JGI C. intestinalis v2.0 Home. http://genome.jgi-psf.org/Cioin2/Cioin2.home.html*.
118. (2007). *MBI_3.54 - Mouse Genome Informatics. http://www.informatics.jax.org/*.
119. (2007). *FlyBase Homepage. http://flybase.bio.indiana.edu/*.
120. Letunic, I., Copley, R.R., Pils, B., Pinkert, S., Schultz, J., and Bork, P., (2006). SMART 5: domains in the context of genomes and networks. *Nucleic Acids Res*, 34, D257-60.
121. Bateman, A., Coin, L., Durbin, R., Finn, R.D., Hollich, V., Griffiths-Jones, S., Khanna, A., Marshall, M., Moxon, S., Sonnhammer, E.L., Studholme, D.J., Yeats, C., and Eddy, S.R., (2004). The Pfam protein families database. *Nucleic Acids Res*, 32, D138-41.
122. Patthy, L., (1999). Genome evolution and the evolution of exon-shuffling--a review. *Gene*, 238, 103-14.
123. Adamska, M., Matus, D.Q., Adamski, M., Green, K., Rokhsar, D.S., Martindale, M.Q., and Degnan, B.M., (2007). The evolutionary origin of hedgehog proteins. *Curr Biol*, 17, R836-7.
124. Snell, E.A., Brooke, N.M., Taylor, W.R., Casane, D., Philippe, H., and Holland, P.W., (2006). An unusual choanoflagellate protein released by Hedgehog autocatalytic processing. *Proc Biol Sci*, 273, 401-7.
125. Gumbiner, B.M., (2000). Regulation of cadherin adhesive activity. *J Cell Biol*, 148, 399-404.
126. Jamora, C. and Fuchs, E., (2002). Intercellular adhesion, signalling and the cytoskeleton. *Nat Cell Biol*, 4, E101-8.
127. Kemler, R., (1993). From cadherins to catenins: cytoplasmic protein interactions and regulation of cell adhesion. *Trends Genet*, 9, 317-21.
128. Aberle, H., Butz, S., Stappert, J., Weissig, H., Kemler, R., and Hoschuetzky, H., (1994). Assembly of the cadherin-catenin complex in vitro with recombinant proteins. *J Cell Sci*, 107, 3655-63.
129. Adams, C.L., Nelson, W.J., Smith, S.J., (1996). Quantitative analysis of cadherin-catenin-actin reorganization during development of cell-cell adhesion. *J Cell Biol*, 135, 1899-911.
130. Drees, F., Pokutta, S., Yamada, S., Nelson, W.J., and Weis, W.I., (2005). Alpha-catenin is a molecular switch that binds E-cadherin-beta-catenin and regulates actin-filament assembly. *Cell*, 123, 903-15.

131. Rivera, G.M., Antoku, S., Gelkop, S., Shin, N.Y., Hanks, S.K., Pawson, T., and Mayer, B.J., (2006). Requirement of Nck adaptors for actin dynamics and cell migration stimulated by platelet-derived growth factor B. *Proc Natl Acad Sci U S A*, 103, 9536-41.
132. Jones, N., Blasutig, I.M., Eremina, V., Ruston, J.M., Bladt, F., Li, H., Huang, H., Larose, L., Li, S.S., Takano, T., Quaggin, S.E., and Pawson, T., (2006). Nck adaptor proteins link nephrin to the actin cytoskeleton of kidney podocytes. *Nature*, 440, 818-23.
133. Hibbert, D.J., (1975). Observations on the ultrastructure of the choanoflagellate *Codosiga botrytis* (Ehr.) Saville-Kent with special reference to the flagellar apparatus. *J Cell Sci*, 17, 191-219.
134. Wollner, D.A. and Nelson, W.J., (1992). Establishing and maintaining epithelial cell polarity. Roles of protein sorting, delivery and retention. *J Cell Sci*, 102, 185-90.
135. Wollner, D.A., Krzeminski, K.A., and Nelson, W.J., (1992). Remodeling the cell surface distribution of membrane proteins during the development of epithelial cell polarity. *J Cell Biol*, 116, 889-99.
136. Bryant, D.M. and Stow, J.L., (2004). The ins and outs of E-cadherin trafficking. *Trends Cell Biol*, 14, 427-34.
137. Segawa, Y., Suga, H., Iwabe, N., Oneyama, C., Akagi, T., Miyata, T., and Okada, M., (2006). Functional development of Src tyrosine kinases during evolution from a unicellular ancestor to multicellular animals. *Proc Natl Acad Sci U S A*, 103, 12021-6.
138. Manning, G., Young, S.L., Miller, W.T., and Zhai, Y., (2007). Diversity and convergence in the tyrosine kinase network of a unicellular choanoflagellate. *(in review)*.
139. Takata, K. and Singer, S.J., (1988). Phosphotyrosine-modified proteins are concentrated at the membranes of epithelial and endothelial cells during tissue development in chick embryos. *J Cell Biol*, 106, 1757-64.
140. Tsukita, S., Oishi, K., Akiyama, T., Yamanashi, Y., Yamamoto, T., and Tsukita, S., (1991). Specific proto-oncogenic tyrosine kinases of src family are enriched in cell-to-cell adherens junctions where the level of tyrosine phosphorylation is elevated. *J Cell Biol*, 113, 867-79.
141. Lilien, J. and Balsamo, J., (2005). The regulation of cadherin-mediated adhesion by tyrosine phosphorylation/dephosphorylation of beta-catenin. *Curr Opin Cell Biol*, 17, 459-65.
142. Perez-Moreno, M., Jamora, C., and Fuchs, E., (2003). Sticky business: orchestrating cellular signals at adherens junctions. *Cell*, 112, 535-48.
143. Carthew, R.W., (2005). Adhesion proteins and the control of cell shape. *Curr Opin Genet Dev*, 15, 358-63.
144. Kazmierczak, P., Sakaguchi, H., Tokita, J., Wilson-Kubalek, E.M., Milligan, R.A., Muller, U., and Kachar, B., (2007). Cadherin 23 and protocadherin 15 interact to form tip-link filaments in sensory hair cells. *Nature*, 449, 87-91.
145. Mengaud, J., Ohayon, H., Gounon, P., Mege, R.M., and Cossart, P., (1996). E-cadherin is the receptor for internalin, a surface protein required for entry of L. monocytogenes into epithelial cells. *Cell*, 84, 923-32.
146. Boyle, E.C. and Finlay, B.B., (2003). Bacterial pathogenesis: exploiting cellular adherence. *Curr Opin Cell Biol*, 15, 633-9.
147. Blau, K., Portnoi, M., Shagan, M., Kaganovich, A., Rom, S., Kafka, D., Chalifa Caspi, V., Porgador, A., Givon-Lavi, N., Gershoni, J.M., Dagan, R., and Mizrachi Nebenzahl,

Y., (2007). Flamingo cadherin: a putative host receptor for Streptococcus pneumoniae. *J Infect Dis*, 195, 1828-37.
148. Telford, M.J., (2006). Animal phylogeny. *Curr Biol*, 16, R981-5.
149. (2007). *Tree of Life Web Project. http://www.tolweb.org/tree/*.
150. (2005). *JGI X. tropicalis v4.1 Home http://genome.jgi-psf.org/Xentr4/Xentr4.home.html*.
151. Pai, R. and Kang, G., (2008). Microbes in the gut: a digestable account of host-symbiont interactions. *Indian J Med Res*, 128, 587-94.
152. Pluddemann, A., Mukhopadhyay, S., and Gordon, S., (2006). The interaction of macrophage receptors with bacterial ligands. *Expert Rev Mol Med*, 8, 1-25.
153. Hauck, C.R., Agerer, F., Muenzner, P., and Schmitter, T., (2006). Cellular adhesion molecules as targets for bacterial infection. *Eur J Cell Biol*, 85, 235-42.
154. M. E. Pettitta, B.A.A.O., J. R. Blakeb and B. S.C. Leadbeater, (2002). The hydrodynamics of filter feeding in choanoflagellates *European Journal of Protistology*, 38, 313-32.
155. Bozzaro, S., Bucci, C., and Steinert, M., (2008). Phagocytosis and host-pathogen interactions in Dictyostelium with a look at macrophages. *Int Rev Cell Mol Biol*, 271, 253-300.
156. Lloyd, D.H., Viac, J., Werling, D., Reme, C.A., and Gatto, H., (2007). Role of sugars in surface microbe-host interactions and immune reaction modulation. *Vet Dermatol*, 18, 197-204.
157. Ireton, K., (2007). Entry of the bacterial pathogen Listeria monocytogenes into mammalian cells. *Cell Microbiol*, 9, 1365-75.
158. Anderton, J.M., Rajam, G., Romero-Steiner, S., Summer, S., Kowalczyk, A.P., Carlone, G.M., Sampson, J.S., and Ades, E.W., (2007). E-cadherin is a receptor for the common protein pneumococcal surface adhesin A (PsaA) of Streptococcus pneumoniae. *Microb Pathog*, 42, 225-36.
159. Li, W., Young, S.L., King, N., and Miller, W.T., (2008). Signaling properties of a non-metazoan Src kinase and the evolutionary history of Src negative regulation. *J Biol Chem*, 283, 15491-501.
160. Fujita, Y., Krause, G., Scheffner, M., Zechner, D., Leddy, H.E., Behrens, J., Sommer, T., and Birchmeier, W., (2002). Hakai, a c-Cbl-like protein, ubiquitinates and induces endocytosis of the E-cadherin complex. *Nat Cell Biol*, 4, 222-31.
161. Behrens, J., Vakaet, L., Friis, R., Winterhager, E., Van Roy, F., Mareel, M.M., and Birchmeier, W., (1993). Loss of epithelial differentiation and gain of invasiveness correlates with tyrosine phosphorylation of the E-cadherin/beta-catenin complex in cells transformed with a temperature-sensitive v-SRC gene. *J Cell Biol*, 120, 757-66.
162. Schirrmeister, W., Gnad, T., Wex, T., Higashiyama, S., Wolke, C., Naumann, M., and Lendeckel, U., (2009). Ectodomain shedding of E-cadherin and c-Met is induced by Helicobacter pylori infection. *Exp Cell Res*, 315, 3500-8.
163. Maretzky, T., Reiss, K., Ludwig, A., Buchholz, J., Scholz, F., Proksch, E., De Strooper, B., Hartmann, D., and Saftig, P., (2005). ADAM10 mediates E-cadherin shedding and regulates epithelial cell-cell adhesion, migration, and beta-catenin translocation. *Proc Natl Acad Sci U S A*, 102, 9182-7.
164. Ye, W., Held, M., Lagergren, J., Engstrand, L., Blot, W.J., Mclaughlin, J.K., and Nyren, O., (2004). Helicobacter pylori infection and gastric atrophy: risk of adenocarcinoma and

squamous-cell carcinoma of the esophagus and adenocarcinoma of the gastric cardia. *J Natl Cancer Inst*, 96, 388-96.
165. Bates, J.M., Mittge, E., Kuhlman, J., Baden, K.N., Cheesman, S.E., and Guillemin, K., (2006). Distinct signals from the microbiota promote different aspects of zebrafish gut differentiation. *Dev Biol*, 297, 374-86.
166. Vrieze, A., Holleman, F., Zoetendal, E.G., De Vos, W.M., Hoekstra, J.B., and Nieuwdorp, M., The environment within: how gut microbiota may influence metabolism and body composition. *Diabetologia*, 53, 606-13.
167. Lee, P.N., Mcfall-Ngai, M.J., Callaerts, P., and De Couet, H.G., (2009). The Hawaiian bobtail squid (Euprymna scolopes): a model to study the molecular basis of eukaryote-prokaryote mutualism and the development and evolution of morphological novelties in cephalopods. *Cold Spring Harb Protoc*, 2009, pdb emo135.
168. Hatta, K., Takagi, S., Fujisawa, H., and Takeichi, M., (1987). Spatial and temporal expression pattern of N-cadherin cell adhesion molecules correlated with morphogenetic processes of chicken embryos. *Dev Biol*, 120, 215-27.
169. Oda, H., Tsukita, S., and Takeichi, M., (1998). Dynamic behavior of the cadherin-based cell-cell adhesion system during Drosophila gastrulation. *Dev Biol*, 203, 435-50.
170. Thiery, J.P., Acloque, H., Huang, R.Y., and Nieto, M.A., (2009). Epithelial-mesenchymal transitions in development and disease. *Cell*, 139, 871-90.
171. Yilmaz, M. and Christofori, G., (2009). EMT, the cytoskeleton, and cancer cell invasion. *Cancer Metastasis Rev*, 28, 15-33.
172. Takeichi, M., (1977). Functional correlation between cell adhesive properties and some cell surface proteins. *J Cell Biol*, 75, 464-74.
173. Nose, A., Nagafuchi, A., and Takeichi, M., (1988). Expressed recombinant cadherins mediate cell sorting in model systems. *Cell*, 54, 993-1001.
174. Katsamba, P., Carroll, K., Ahlsen, G., Bahna, F., Vendome, J., Posy, S., Rajebhosale, M., Price, S., Jessell, T.M., Ben-Shaul, A., Shapiro, L., and Honig, B.H., (2009). Linking molecular affinity and cellular specificity in cadherin-mediated adhesion. *Proc Natl Acad Sci U S A*, 106, 11594-9.
175. Nose, A., Tsuji, K., and Takeichi, M., (1990). Localization of specificity determining sites in cadherin cell adhesion molecules. *Cell*, 61, 147-55.
176. Perret, E., Benoliel, A.M., Nassoy, P., Pierres, A., Delmas, V., Thiery, J.P., Bongrand, P., and Feracci, H., (2002). Fast dissociation kinetics between individual E-cadherin fragments revealed by flow chamber analysis. *Embo J*, 21, 2537-46.
177. Chappuis-Flament, S., Wong, E., Hicks, L.D., Kay, C.M., and Gumbiner, B.M., (2001). Multiple cadherin extracellular repeats mediate homophilic binding and adhesion. *J Cell Biol*, 154, 231-43.
178. Perez, T.D., Nelson, W.J., (2005). Cadherin adhesion: mechanisms and molecular interactions. *Handbook of Experimental Pharmacology*.
179. Posy, S., Shapiro, L., and Honig, B., (2008). Sequence and structural determinants of strand swapping in cadherin domains: do all cadherins bind through the same adhesive interface? *J Mol Biol*, 378, 954-68.
180. Stemmler, M.P., (2008). Cadherins in development and cancer. *Mol Biosyst*, 4, 835-50.
181. Carr, M., Leadbeater, B.S., and Baldauf, S.L., Conserved meiotic genes point to sex in the choanoflagellates. *J Eukaryot Microbiol*, 57, 56-62.

182. Drees, F., Reilein, A., and Nelson, W.J., (2005). Cell-adhesion assays: fabrication of an E-cadherin substratum and isolation of lateral and Basal membrane patches. *Methods Mol Biol*, 294, 303-20.
183. Nagafuchi, A., Shirayoshi, Y., Okazaki, K., Yasuda, K., and Takeichi, M., (1987). Transformation of cell adhesion properties by exogenously introduced E-cadherin cDNA. *Nature*, 329, 341-3.
184. Lambert, M., Padilla, F., and Mege, R.M., (2000). Immobilized dimers of N-cadherin-Fc chimera mimic cadherin-mediated cell contact formation: contribution of both outside-in and inside-out signals. *J Cell Sci*, 113 (Pt 12), 2207-19.
185. Brieher, W.M., Yap, A.S., and Gumbiner, B.M., (1996). Lateral dimerization is required for the homophilic binding activity of C-cadherin. *J Cell Biol*, 135, 487-96.
186. Takeda, H., Shimoyama, Y., Nagafuchi, A., and Hirohashi, S., (1999). E-cadherin functions as a cis-dimer at the cell-cell adhesive interface in vivo. *Nat Struct Biol*, 6, 310-2.
187. Liwosz, A., Lei, T., and Kukuruzinska, M.A., (2006). N-glycosylation affects the molecular organization and stability of E-cadherin junctions. *J Biol Chem*, 281, 23138-49.
188. Ozawa, M. and Kemler, R., (1990). Correct proteolytic cleavage is required for the cell adhesive function of uvomorulin. *J Cell Biol*, 111, 1645-50.
189. Medina, A., Swain, R.K., Kuerner, K.M., and Steinbeisser, H., (2004). Xenopus paraxial protocadherin has signaling functions and is involved in tissue separation. *Embo J*, 23, 3249-58.
190. Pece, S. and Gutkind, J.S., (2000). Signaling from E-cadherins to the MAPK pathway by the recruitment and activation of epidermal growth factor receptors upon cell-cell contact formation. *J Biol Chem*, 275, 41227-33.
191. Williams, E.J., Furness, J., Walsh, F.S., and Doherty, P., (1994). Activation of the FGF receptor underlies neurite outgrowth stimulated by L1, N-CAM, and N-cadherin. *Neuron*, 13, 583-94.
192. Sanchez-Heras, E., Howell, F.V., Williams, G., and Doherty, P., (2006). The fibroblast growth factor receptor acid box is essential for interactions with N-cadherin and all of the major isoforms of neural cell adhesion molecule. *J Biol Chem*, 281, 35208-16.
193. Manning, G., Young, S.L., Miller, W.T., and Zhai, Y., (2008). The protist, Monosiga brevicollis, has a tyrosine kinase signaling network more elaborate and diverse than found in any known metazoan. *Proc Natl Acad Sci U S A*, 105, 9674-9.
194. Pinho, S.S., Reis, C.A., Paredes, J., Magalhaes, A.M., Ferreira, A.C., Figueiredo, J., Xiaogang, W., Carneiro, F., Gartner, F., and Seruca, R., (2009). The role of N-acetylglucosaminyltransferase III and V in the post-transcriptional modifications of E-cadherin. *Hum Mol Genet*, 18, 2599-608.
195. Nichols, S.A., Dirks, W., Pearse, J.S., and King, N., (2006). Early evolution of animal cell signaling and adhesion genes. *Proc Natl Acad Sci U S A*, 103, 12451-6.
196. Perez, T.D., Tamada, M., Sheetz, M.P., and Nelson, W.J., (2008). Immediate-early signaling induced by E-cadherin engagement and adhesion. *J Biol Chem*, 283, 5014-22.
197. Jaiswal, M., Agrawal, N., and Sinha, P., (2006). Fat and Wingless signaling oppositely regulate epithelial cell-cell adhesion and distal wing development in Drosophila. *Development*, 133, 925-35.

198. Wheelock, M.J. and Johnson, K.R., (2003). Cadherin-mediated cellular signaling. *Curr Opin Cell Biol*, 15, 509-14.
199. Brymora, A., Valova, V.A., Robinson, P.J., (2003). Protein-Protein Interactions Identified by Pull-Down Experiments and Mass Spectrometry. *Current Protocols in Cell Biology*.