Structural molecular components of Septate Junctions in cnidarians point to the origin of epithelial junctions in Eukaryotes

Philippe Ganot¹, Didier Zoccola¹, Eric Tambutté¹, Christian R. Voolstra², Manuel Aranda², Denis Allemand¹, Sylvie Tambutté¹

¹ Centre Scientifique de Monaco, Marine Biology Department, MC98000, Monaco.
² Red Sea Research Center, KAUST, Thuwal, Saudi Arabia.

Correspondence: Philippe Ganot, Centre Scientifique de Monaco, Marine Biology Department, 8 Quai Antoine Premier, MC98000, Monaco
Tel: +377 97 77 44 76; Fax: +377 97 77 44 01; E-mail: pganot@centrescientifique.mc

Keywords: Epitheliozoa, Claudin, Neurexin, Contactin, Neuroglian, Coracle, MAGUK, Na+/K+ ATPase transporter, DSCAM, Nbl4, para-cellular pathway, permselectivity, corals, ctenophores, poriferans, Monosiga, Capsaspora.

Running title: SJs in Cnidarians
Abstract

Septate junctions (SJs) insure barrier properties and control paracellular diffusion of solutes across epithelia in invertebrates. However, the origin and evolution of their molecular constituents in Metazoa has not been firmly established. Here, we investigated the genomes of early branching metazoan representatives to reconstruct the phylogeny of the molecular components of SJs. Although Claudins and SJ cytoplasmic adaptor components appeared successively throughout metazoan evolution, the structural components of SJs arose at the time of Placozoa/Cnidaria/Bilateria radiation. We also show that in the scleractinian coral *Stylophora pistillata*, the structural SJ component Neurexin IV (StpNrXIV) co-localizes with the cortical actin network at the apical border of the cells, at the place of SJs. We propose a model for SJ components in Cnidaria. Moreover, our study reveals an unanticipated diversity of SJ structural component variants in cnidarians. This diversity correlates with gene specific expression in calcifying and non-calcifying tissues, suggesting specific paracellular pathways across the cell layers of these diploblastic animals.
Introduction

A unifying characteristic of metazoans evolution has been the building of joined layers of cells that form a physical barrier between the environment and the inner body, or between different compartments within the body. Desmosomes and adherens junctions insure mechanical binding between cells forming the epithelia, whereas occluding junctions seal and control paracellular transport across the epithelial layer. Two structurally different types of occluding junctions have been characterized, the Tight Junction (TJ) and the Septate Junction (SJ) (Banerjee, Sousa, et al. 2006; Magie and Martindale 2008).

Multiple studies have investigated ultrastructure and molecular composition of TJs and SJs in bilaterians. TJs appear restricted to chordates and form circular strands around the apical cell border joining together two adjacent plasma membranes (Shen, et al. 2011). In protostomes, SJs are the predominant occluding junctions typically arranged in a spiral manner around the cell lateral border forming large macromolecular complexes that span the extra-cellular space between two neighboring cells. In transmission electron microscope (TEM) cross-sections images, SJs display characteristic electron-dense ladder-like structures of 10-20 nm width called septa (Tepass, et al. 2001). SJs are also found in mammals, at the nodes of Ranviers where they form the paranodal junction between axons and myelinated glial cells (Hortsch and Margolis 2003; Poliak and Peles 2003; Nans, et al. 2011). In non-bilaterians, cell-cell junctions have been structurally investigated using different electron microscope techniques in the diverse phyla composing early branching metazoans, i.e. Cnidaria, Placozoa, Porifera, and Ctenophora. In cnidarians, both medusozoans and anthozoans possess belt junctions referred to as SJs that form a belt around the apical circumference of the cell, although the “Hydra type” (Hydrozoa) and the “Anthozoan type” (Actiniaria) of SJs were shown to slightly differ structurally (Wood 1959; Filshie and Flower 1977; Green and Flower 1980). In Placozoa, very little is known. In the ventral epithelium of Trichoplax adhaerens (Trichoplax), apical belt desmosomes with proximal “periodic connection” of intercellular material joining two adjacent cells have been noted (Ruthmann, et al. 1986). Although this is reminiscent of the SJ ladder structure, it awaits further clarification. Porifera encompass four distinct taxonomic classes, (Philippe, et al. 2009; Sperling, et al. 2009; Erwin, et al. 2011) with differences in junction depending on the class. No clear SJ were described in the more distant Hexactinellida or Demospongiae (Leys, et al. 2009). In Homoscleromorph, the presence of SJ is uncertain as the presence of septa is unclear (Leys, et al. 2009; Gazave, et al. 2010). In fact, the only clear report of SJ in Porifera was made in calcareous by Ledger (1975). In this study, the authors could show an electron dense ladder.
between the spicule-secreting sclerocytes Sycon ciliatum, although electron dense junctions with no visible septa were described in Sycon coactum (Eerkes-Medrano and Leys 2006). Ctenophores share a very similar unique junctional structure, where epithelial cells are linked with distinctive belt junction (see supplementary Figure S8). The junctional membranes are 2-3 nm apart but they do not fuse nor are they linked by septa (Hernandez-Nicaise, et al. 1989; Hernandez-Nicaise 1991).

At the molecular level, most functional studies rely on mammalian and insect model organisms. Components of TJs and SJs can be sub-divided into inter-cellular structural and cytosolic scaffolding/polarity proteins (see Table 1). While the structural transmembrane proteins mediate cell-cell adhesion, the cytosolic junction plaque contains various types of proteins that link the junction transmembrane proteins to the underlying cytoskeleton.

The structural components of Drosophila SJs and human axo-glial SJs consist of a core complex of three cell-adhesion molecules (Hortsch and Margolis 2003). These are Neurexin IV (NrxIV), Contactin (Cont), and Neuroglian (Nrg) in Drosophila corresponding respectively to Caspr, Contactin, and Neurofascin in mammals. Loss of any one of these proteins in either Drosophila or Mouse disrupts SJ formation and function, and cellular trafficking of these proteins to SJs is interdependent (Baumgartner, et al. 1996; Boyle, et al. 2001; Genova and Fehon 2003; Faivre-Sarrailh, et al. 2004; Banerjee, Pillai, et al. 2006; Bonnon, et al. 2007; Thaxton, et al. 2010; Tiklova, et al. 2010; Banerjee, et al. 2011). In both, Drosophila and mammalian SJs, NrxIV/Caspr associates with Cont/Contactin in cis and with Nrg/Neurofascin in trans (Hortsch and Margolis 2003). Further trans-membrane proteins have been characterized to be required for SJ formation and/or function in Drosophila. Three Claudin-like proteins, Megatrachea, Sinuous and Kune, are essential for SJ (see below) (Behr, et al. 2003; Wu, et al. 2004; Nelson, et al. 2010). The Na+K+ ATPase subunits alpha and beta are also necessary for SJs in a pump-independent function (Genova and Fehon 2003; Paul, et al. 2003; Paul, et al. 2007; Krupinski and Beitel 2009). Other proteins have been characterized as part of SJs, including Lachesin, Fasciclin III, Macroglobulin complement-related ((Woods, et al. 1997; Llimargas, et al. 2004; Narasimha, et al. 2008; Batz, et al. 2014), see also (Hall, et al. 2014) for a recent exhaustive listing).

Structural components of vertebrate TJs are members of the tetraspan family, i.e Claudins, Occludin, and Tricellulin, and members of the immunoglobulin superfamily, i.e. JAM1-3 (junctional adhesion molecules), ESAM (endothelial cell-selective adhesion molecule), and CAR (coxsackie- and adenovirus receptor) (Furuse 2010). Claudins are the focus of an abundant literature as they probably represent the principle TJ barrier-forming
proteins in vertebrates (Van Itallie and Anderson 2006; Shen, et al. 2011). The human Claudins-1-27, here referred to as Claudins sensu stricto (Claudins s.s.), represent a family of small (20–27 kDa) highly specialized proteins predicted to have four transmembrane helices (tetraspan), with two extracellular loops (EL1 and EL2). EL1 is characterized by a conserved W-GLW-C amino acid (AA) motif. Claudins s.s. have been shown to interact with each other in cis on the plasma membrane and in trans forming kissing complexes in the paracellular space (Krause, et al. 2008). Moreover, several Claudin-like proteins sharing the very same tetraspan topology and W-GLW-C-C signature motif have been identified, although their function is much less understood. These further include human Claudin Like1-2, lens fiber membrane intrinsic protein isoform 2 (Lim2), epithelial membrane proteins (EMP1-3), and peripheral myelin protein 22 (PMP22), for which adhesive or barrier-forming properties have been described (Van Itallie and Anderson 2006; Günzel and Alan 2013). Additionally, more distantly related Claudin-like proteins with a similar structure, but with a less conserved signature motif, have been identified. Among these are human Lipoma HMGIC fusion partner (LHFP1-4) proteins, Clarins3, and uncharacterized protein C16orf52 (CPO52), for which a function remains to be fully clarified (Huang, et al. 2004; Geng, et al. 2009). In invertebrates, Claudin-like proteins with a similar tetraspan topology are also found, although their relation to the vertebrate Claudins is unclear. In Drosophila, of the eight Claudin-like proteins referenced in the genome, only three were shown to be required for the formation and barrier function of SJs. These are Megatrachea, Sinuous, and Kune, all of which carry the W-GLW-C-C signature. Furthermore, deletion of any of these genes impairs SJ formation (Behr, et al. 2003; Wu, et al. 2004; Furuse and Tsukita 2006; Nelson, et al. 2010). Megatrachea co-immunoprecipitates with NrxIV, Nrg, and Cont, among others (Jaspers, et al. 2012), showing that at least some of the invertebrate Claudin-like proteins participate in SJs. Nonetheless, despite the identification of many Claudin-like proteins in diverse invertebrate phyla, their phylogeny, association with SJs, and/or function remains to be clarified.

Although SJs and TJs show striking differences in their respective structural components, the cytosolic adaptor proteins responsible for their assembly and maintenance at the plasma membrane appear to share, in part, similar machineries. MAGUK proteins are evolutionary conserved scaffolding proteins that create and maintain multi-molecular complexes, such as adherens and occluding junctions, at distinct subcellular sites like the cytoplasmic surface of the plasma membrane for instance (Ikenouchi, et al. 2007; de Mendoza, et al. 2010). In Drosophila, members of the MAGUK protein superfamily,
including Discs Large (Dlg), Zona Occludens (ZO), Varicose (Vari), and Stardust (Std) are necessary for epithelial polarity and scaffolding of SJs (Woods, et al. 1996; Bachmann, et al. 2001; Jung, et al. 2006; Moyer and Jacobs 2008). Similarly, the mammalian MAGUK proteins Discs Large (Dlg), Zona Occludens (ZO), MPP5 (Pals1), and MPP7 are also part of the cytosolic adaptors involved in TJ formation (Roh, et al. 2003; Van Itallie and Anderson 2006; Stucke, et al. 2007; Fanning, et al. 2012; Su, et al. 2012). ZO1 and ZO2 have been shown to bind several Claudins s.s. (Van Itallie and Anderson 2006). Likewise, the cytosolic FERM domain protein Coracle (Cora) binds to the cytoplasmic domain of NrxIV in Drosophila SJs, and its homologs in vertebrates (Band 4.1) participate in TJ formation (Fehon, et al. 1994; Lamb, et al. 1998; Ward, et al. 1998; Mattagajasingh, et al. 2000; Denisenko-Nehrbass, et al. 2003; Jensen and Westerfield 2004; Laprise, et al. 2009; Xia and Liang 2012). Finally, the Na\(^{+}/K\)^{+} ATPase alpha- and beta-subunits (ATPalpha and Nervana2) have repeatedly been associated with both TJs and SJs although the specific function of this transporter in junctions is unclear (Paul, et al. 2003; Rajasekaran, et al. 2005; Laprise, et al. 2009; Vagin, et al. 2012).

Outside Bilateria, several studies have identified members of the structural and scaffolding SJ molecular components in early branching Metazoa. However, these data appear contradictory and incomplete (Chapman, et al. 2010; Fahey and Degnan 2010; Leys and Riesgo 2012). For example, Claudins appeared in Porifera (Leys and Riesgo 2012), in Cnidaria (Chapman, et al. 2010), or in Bilateria (Fahey and Degnan 2010). One homolog to NrxIV and Cont was present in both Porifera and Cnidaria (Chapman, et al. 2010) or absent in Porifera and noted as “aberrant” in Cnidaria (Fahey and Degnan 2010). In a more recent analysis by Suga and coworkers (Suga, et al. 2013), NrxIV was present in Trichoplax, cnidarians and bilaterians. In contrast Cont was specific to bilaterians, whereas Riesgo et al (2014) identified a Cont homolog in poriferans. Homology criteria may have been misaddressed and/or intermediate evolutionary precursor may have been missed. Thus, despite SJs having been structurally characterized already three decades ago, their gene complement, respective diversification, and evolution in early branching metazoans remains elusive. In other words, how and when body compartmentalization has arisen in Metazoa is still a controversial question.

In order to gain molecular insight into cnidarian SJs, we initiated the characterization of their principal molecular components in three different cnidarian representatives (i.e. the scleractinian coral Stylophora pistillata, the actiniarian Nematostella vectensis, and the medusozoan Hydra magnipapillata). We monitored expression and localization of key SJ
proteins in the coral *S. pistillata*, a tractable species for calcification studies. After having defined the principal members of SJs in Cnidaria, we extended our genomic search to the other representatives of the early branching metazoans (i.e. the placozoan *Trichoplax adherens*, the homoscleromoph *Oscarella carmella*, the demosponge *Amphimedon queenslandica*, and the ctenophore *Mneniopsis leidyi*) as well as the unicellular organisms suspected to be at the origin of the metazoan lineage (i.e. the choanoflagellate *Monosiga brevicollis* and the filasterean *Capsaspora owczarzaki*). Our analysis also includes evolutionary related gene families (not necessarily functionally related) to apprehend SJ gene evolution. The present study aims at providing a comprehensive repertoire of the components involved in sealing epithelia of early metazoans as well as to reconstruct the stepwise evolution of SJs in invertebrates that preceded the formation of TJs in chordates.
Results


Early branching metazoans encode SJ, but not TJ, components

Transcriptome and genome data mining was based on BLAST (bilaterian query sequences against non-bilaterian databases) and reciprocal BLAST (non-bilaterian candidate sequences against bilaterian databases) approaches (Supplementary Figure S3). The search was performed in an iterative manner, first targeting cnidarians and then extended to include the other phyla of interest. In addition to homology approaches, based on reverse BLAST against human and *Drosophila*, we used domain composition (SMART) and phylogenetic trees (PhyML and Bayesian) to identify and name homologs of known occluding junction components (our terminology followed the *Drosophila* nomenclature). Table 1 summarizes the presence/absence of homologs across non-bilaterians. None of the TJ structural components characteristic of chordates were found in non-bilaterians. However, all SJs components that we searched for were present in cnidarians, often in multiple copies. In the other phyla, the range of SJ component homologs was variable with a correlating trend of fewer homologs/copies and organismal simplicity.
i) Claudins

Human Claudin 1-27 (Claudins s.s.) homologs were not found. However, iterative search with bilaterian Claudin-like sequences identified a variable number of Claudin-like homologs in the different representative species of early branching metazoans as well as protists (Table 1). Profile based search against the PFAM database confirmed that all belonged to the PMP22_Claudin (PF00822), Claudin_2 (PF13903), or L_HGMIC_fpl (PF10242) domain family, except for 3 of the Claudins identified in Oscarella (OcaClau4,5,8) (Supplementary Table S1). Transmembrane domain prediction confirmed that all sequences were tetraspan proteins (data not shown) with a larger EL1 (50.8 AA +/- 16.5) than EL2 (19.8 AA +/- 8.0) (Supplementary Table S1). The Claudin signature motif within EL1 appeared slightly modified (i.e. W-G[LVI][WFYL]-C-C), except for a few cases. Bayesian and Maximum Likelihood methods gave incongruent albeit comparable phylogenetic trees, i.e. several well supported groups could be outlined using both methods (Figure 1, Supplementary Figure S5a). Use of an alternative alignment method (MUSCLE) prior to phylogenetic analyses supported the same groups (Supplementary Figure S5a). Group Ia contains anthozoan Claudins AS1,2 with human Claudin domain-containing protein 2 (HsClauL2) and lens fiber membrane intrinsic protein isoform 2 (HsLIM2), and group Ib contains Hydra Claudins 4,7,8,10,12,14,15 with human epithelial membrane protein 1-3 (HsEMP1-3) and peripheral myelin protein 22 (HsPMP22). Of note, the Stylophora, Acropora, and Nematostella Claudin AS1 and AS2 proteins corresponded to two splice variants conserved in anthozoans which vary in their first exon, and consequently in their first ~80 AA. This gave rise to two Claudins differing in their EL1. Group II corresponds to homologs of the human TMP211 and LHFP family of which some members have been involved in ear hair cell formation (Xiong, et al. 2012). This large group contains Claudin members from Drosophila (CG3770, CG12026), cnidarians (anthozoan Clau3-6, Hydra Clau4,5), homoscleromorph (OcaClau1,2), and the placozoan and ctenophore Claudins (TriClau and MleClau1-4, respectively). Group III (Drosophila CG14182, anthozoan Clau8-9, Oscarella Clau8, Monosiga Clau8, Capsaspora Clau1), IV (anthozoan Clau2, Oscarella Clau3,4) and V (anthozoan Clau7, Oscarella Clau5) comprise homologs of the human CPO52 (uncharacterized protein C16orf52), Clarin3 and TMP127, respectively, for which a function has not yet been determined. Other Claudins sequences (for example MonoClauA,B,C or amphiClau) could not be reliably positioned on the tree, inferring that the Claudin primary sequences have considerably diverged during evolution. The human Claudins s.s. have been recently subdivided into 5 subgroups (Gunzel and Fromm 2012). Two members from each
subgroup were randomly selected (HsCLDN1,2,3,8,11,12,16,18a,21,23) as representatives of
the human Claudins s.s. and included in our phylogenetic analysis. These human TJ specific
Claudins clustered as a single outgroup. Interestingly, the three Drosophila Claudin-like
proteins Megatrachea, Sinuous, and Kune, for which functional characterizations are
available, also clustered outside our 5 groups.

\textit{ii) Neurexins}

The Drosophila Neurexin IV (NrxIV) and human Caspr family of proteins are closely
related extracellular ligands with parallel domain architecture (i.e. LamG, EGF and FBG).
One clear homolog of NrxIV/Caspr was found in cnidarians (NRX1), displaying the same
domain architecture, except for a missing NH2term FA58C domain (Figure 2A and
supplementary Figure S5b). In Stylophora NRX1, domain homology search using SMART
revealed a Band 4.1 binding motif. The presence of this motif indicates that StpNRX1
potentially binds to the putative Cora/Band 4.1 homolog as known from bilaterians. NRX1
was found to be duplicated in Nematostella (NvNRX1-2), A. digitifera (AdiNRX1-2), and
Hydra (HydNRX1-3). In addition, several extra copies for cnidarian NRX (StpNRX2-5,
AdiNRX3-5, NvNRX3-6, HydNRX4-5) were found, with missing domains and/or long
intracellular portions in comparison to the \textit{bona fides} NRX1 homologs. Within the
phylogenetic tree, the position of these supernumerary homologs within non-bilaterians NRX
suggests duplication within the cnidarian lineage. StpNRX2 did not cluster with any
anthozoa homolog, suggesting that it may be either specific to Stylophora or the robust clade
of scleractinian corals, since it is found in Acropora (complex clade).

In Trichoplax we identified 5 potential NrxIV homologs, representing placozoa
specific duplications, showing variable domains composition, except for TriNRX1 which
harbor canonical Nrx domain composition. No NrxIV/Caspr homologs were found in the
remaining analyzed phyla. However, homologs of the more distant gene family, human
Neurexin 1-3 (HsNeu1-3) and Drosophila Neurexin 1 (DmNeu1) were present. Neurexins are
synaptic cell adhesion molecules in bilaterian composed of alternating LamG and EGF
domains (Bang and Owczarek 2013), similar to the NrxIV/Caspr, but without the FBG
domain (Figure 2A). One clear Neurexin1 homolog (see Supplementary Table S1 for reverse
BLAST hits) was found in Oscarella (but not in Amphimedon, supplementary Figure S6) and
in Mnemiopsis, with the same domain architecture as in bilaterians. Of note, omission of the
bilaterian Neurexin protein sequences placed the OcaNeu and MleNeu sequences at the base
of the cnidarian/placozoa NRX phylogeny (data not shown). Several conserved Neurexin
homologs were also found in cnidarians, although here, they were substantially shorter: 2
LamG and 1 EGF domain, instead of 6 LamG and 3 EGF domains in the canonical form. Moreover, one putative Neurexin1 homolog was found in *Capsaspora* (also identified in (Suga, et al. 2013)), which is composed of 6 LamG domains (Figure 2A) and positions in between the Neu and Nrx families in the phylogenetic tree (see Radial representations in Supplementary Figure S4), potentially representing the metazoan Neu-Nrx ancestor.

**iii) NRG, CONT, DSCAM**

The *Drosophila* Nrg and Cont, and the human Neurofascin and Contactin have closely related domain structures, i.e. succession of Ig domains followed by FN3 domains. However Nrg/Neurofascin has a Cterm transmembrane (TM) domain spanning the plasma membrane, whereas Cont/Contactin is attached to the membrane via a GPI anchor. Both Nrg and Cont have homologs in cnidarians displaying similar domain architecture as well as TM and GPI anchor attachment, respectively (Figure 2B and Supplementary Figure S5b). In comparison to *Nematostella*, the scleractinians *Stylophora* and *Acropora* have additional NRG copies (StpNRG2, AdiNRG2,3). Their position within the phylogenetic tree indicates that they represent scleractinian specific duplications. In *Trichoplax*, 4 NRG and 1 CONT homologs were also found, likewise with placozoan-specific duplications (Figure 2B). Note that differentiation between the *Trichoplax* NRGs and CONT was solely based on the TM/GPI anchor prediction. In *Oscarella*, 1 potential NRG homolog (OcaNRGCAM) could be found. However, based on the reverse BLAST hit approach, this protein could either be a NRG or a DSCAM homolog. Indeed, NRG/CONT share a very similar domain composition with other Ig/FN3 domain adhesion molecules such as Hemicentin, and DSCAM (Down Syndrome Cell Adhesion Molecule), the latter being the closest relative of NRG/CONT. DSCAM are extracellular ligands capable of homophilic associations and heterophilic interactions involved in neural wiring in bilaterian as well as innate immunity in protostomes (Schmucker and Chen 2009). We thus undertook a characterization of DSCAM proteins in the different early branching metazoan to estimate the evolutionary convergence of the NRG/CONT and DSCAM families. DSCAM homologs were identified in *Mnemiopsis* (1), *Oscarella* (2), *Trichoplax* (1), *Hydra* (2) and anthozoans (2), but neither in *Amphimedon* nor in protists (Figure 2B). With respect to domain architecture, the cnidarian DSCAM1 resembled the bilaterian DSCAMs, whereas the cnidarian DSCAM2, TriDSCAM, OcaDSCAM and MleDSCAM showed higher similarity to the NRG/CONT architecture, despite being closer to DSCAM at the sequence level. In line with these finding, the OcaNRGCAM protein represents an evolutionary intermediate between the two families (see supplementary Figure S4).
iv) MAGUK

Members of the MAGUK super family share a central PDZ-SH3-GuKc domains module. The various MAGUK members essentially differ by the addition of other domains, commonly PDZ and L27 (Funke, et al. 2005). The phylogenetic analysis of MAGUK members across early branching metazoans was based on the central module sequences (Figure 3A and supplementary Figure S5c). This analysis complements the previous analysis by de Mendoza et al. (2010). Both Capsaspora and Monosiga possess a MPP and Dlg ancestor that gave rise to the phylogenetic diversity of the metazoan MAGUK family. We show that MPP2-7 is split in 2 distinct groups, MPP2,6 (Varicose) with extended members in all early branching metazoans, and MPP3,4,7 (Mena3) which is restricted to bilaterians and cnidarians. MPP5 (Stardust) appears to have several related members (MPPb) in poriferans, ctenophores and placozoan. However, we could not ascribe a bona fides Stardust homolog to ctenophores. The ZO family is present in all early branching metazoans, except Amphimedon.

v) Coracles

Coracle, Yurt, and Nbl4 are structurally related FERM-FA domains proteins (Tepass 2009). Cnidarians possess a clear Coracle homolog (CORA) and 2 additional Coracle variants that mainly differ by their COOH terminal moiety (Figure 3B and supplementary Figure S5c). Trichoplax, Oscarella, Amphimedon and Mnemiopsis also harbor Coracle-like proteins, structurally closer to the cnidarian Coracle variants than the canonical one. A Yurt homolog is found in anthozoans, Trichoplax and Oscarella whereas the Nbl4 (Human 4.1-Like) appears to have emerged at the time of Cnidarian/bilaterian radiation. Of note, OcaYurt clusters with Nbl4 protein sequences in the Bayesian tree and Yurt protein sequences in the Maximum Likelihood tree (Figure 3B and supplementary Figure S5c), and may therefore represent an ancestor of Yurt-Nbl4 families. OcaYurt was ascribed as Oscarella Yurt homolog based on BLAST results (Supplementary Table S1).

vi) Phylogenetic conclusive remarks

Taken together, cnidarians and placozoans appear to share the complete SJ complement. Several gene duplications were observed in cnidarians (NRX, NRG, CORA), some of which are likely specific to reef building corals. Dichotomy between Hydra and the anthozoans was apparent in the gene phylogeny (e.g. Claudin-like, NRX, NRG), which indicates class specific diversification with possible subsequent divergence in SJ structures. In contrast, genes encoding for the structural components of SJs, i.e. NrxIV, Nrg and Cont, are absent in the other early branching metazoan phyla analyzed here (Table 1), although members of the scaffolding and polarity genes of SJs (MAGUK, Cora, Na+/K+ ATPase
exchanger, Supplementary Figure S5d) are present (despite noticeable losses in *Amphimedon*). The absence of structural SJ proteins suggests that intercellular junctions in these taxa are structurally different from those found in cnidarians and bilaterians.

*The diversified SJ components in anthozoans show distinct and tissue-specific gene expression.*

Electron microscope investigation of *Stylophora* across the different tissue layers clearly hallmarks the presence of SJs between the apical border of every ectodermal and endodermal cell (Supplementary Figure S2) (Tambutté, et al. 2007). They are 0.2 to 1 μm long, depending on the section, and display a characteristic ladder structure. On micrographs where the two tissue layers are visible, SJs of the endoderm layer appear to show higher electron density than those of the ectoderm layer. As the SJ complement in cnidarians appears to have diversified, we next asked what the relative expression of different SJ components was and whether differential expression between the oral (non-calcifying) and aboral (calcifying) tissues could be observed in the adult coral. We developed a protocol to micro-dissect the oral discs from the coral colony using the anesthetic drug MS222 and micro-scissors (see Material and Methods). *Stylophora* total RNA and proteins were extracted from a colony fragment (oral and aboral) or from the oral disc (oral only) and expression was quantified by real-time PCR and western blotting for the genes described in Figure 4. qPCR expression estimates were normalized arbitrarily to StpNRX1=1 as relative expression of the SJ components was our primary focus and because NRX is a core-component of SJs in bilaterians. StpNRX3, 4 and 5 showed relatively low expression in contrast to StpNRX2 (0.53-fold to NRX1) (Figure 4a). The two StpNRG copies were expressed at strikingly different levels (StpNRG1=0.34-fold, StpNRG2=3.1-fold). Unexpectedly, StpCONT was weakly expressed (0.047, see discussion). Claudin-like mRNAs were all expressed, although at relatively low levels in comparison to StpNRX1. The SJ adaptor component StpCORA1 was expressed at a similar level as StpNRX1 and the different variants StpCORA2-4 and StpYURT were also expressed, strongly arguing that these conserved anthozoan genes represent functional rather than pseudo-genes. When assessing tissue specificity, three SJ genes, namely StpNRX2, StpClaud3, and StpClaud6, were strongly down-regulated in the oral disc as opposed to the total colony (oral and aboral) tissue, suggesting that these were mainly expressed in the calcifying aboral tissue, similar to the TFZPD9 calicoblast control (Figure 4b). Although to a lesser extent, StpNRG2 as well showed preferential expression in the aboral tissue. Reversely, StpNRX3, StpCONT, and StpClaudAS2 showed high expression
in the oral disc, albeit displaying a lower colony-wide expression. In order to estimate the relative expression between the endodermal and ectodermal tissue layers, we took advantage of the large size of the sea anemone \textit{Anemonia viridis} tentacles (oral tissue), where the endodermal and ectodermal layer can be manually separated. A partial \textit{A. viridis} cDNA database is available (Sabourault, et al. 2009) and incomplete sequences corresponding to SJ components homologs could be identified. Measurement of their relative tissue expression show predominant tissue specificity for one gene among those tested, namely the duplicated copy of NRX1 (AvNRX1b) (Supplementary Figure S7).

We generated antibodies against the StpNRX1 and StpClaud3 proteins. These antibodies were specific as little to no cross-reactivity could be observed in Western blots (Figure 4c). Similar to the Actin control, StpNRX1 was equally expressed in both the oral disc and the total colony fractions and was present in the blot as a single band <150KDa. This ascertained our qPCR results, namely, that StpNRX1 represents a central component in most SJs. Conversely, StpClaud3 was mostly absent from the oral discs fraction but present in the total colony. This Claudin-like protein is thus likely to be mainly expressed in the aboral calcifying tissues of the coral. In conclusion, anthozoan specific gene diversification is accompanied by differential tissue expression, suggesting the presence of multiple SJ architectures and functions in the different cell layers comprising these diploblastic animals.

\textit{Stylophora NRX1 is glycosylated and co-localizes with F-actin}

StpNRX1 has a predicted molecular weight of 126.5 KDa, which is in disagreement with the molecular weight of 141 KDa determined by Western blotting (Figure 4c). In humans, the Caspr1 protein is glycosylated (Bonnon, et al. 2003); we therefore examined whether StpNRX1 also exhibits post-translational N-linked glycosylation that contributes to the difference between the apparent and predicted molecular weight. Total protein extract was treated with and without PGNaseF (which specifically cleaves between asparagine and N-acetylglucosamines), and Western blotting showed a shift from 141 KDa to 128 KDa of StpNRX1 after PGNase treatment (Figure 5f). StpNRX1 is conclusively N-glycosylated similar to Caspr1 in human. We next addressed the cellular localization of StpNRX1 and StpClaud3 in adult \textit{Stylophora}. An immuno-localization protocol was therefore established. Coral fragments were fixed, decalcified and cut into parts for investigation of the aboral calcifying tissues (Figure 5a). Labeling performed on the basal discs (see Supplementary Figure S1) with phalloidin identified the F-actin network framing every cell (Figure 5b). This cortical F-actin is supposedly adjacent to SJs as SJs are linked to the cytoskeleton in
bilaterians, and anthozoan SJs display similar protein composition to bilaterian SJs. Immuno-localization with phalloidin and anti-StpNRX1 showed overlapping signals for most of the F-actin network (Figure 5c). Optical sectioning sagittal to the epidermal (calicoderm) tissue layer showed that NRX1 and F-actin overlapped, albeit partially, on the apical face of the cell layer (Figure 5d). In order to eliminate optical interference between the Alexa-conjugated secondary antibody and potential endogenous autofluorescence (e.g., GFP), the rabbit anti-NRX1 was detected simultaneously with anti-rabbit-Alexa488 and anti-rabbit-Alexa405. Both channels showed identical labeling in the calicoderm layer (figure 5d1-2). Thus, StpNRX1 could be co-localized with, or very close to, the F-actin network at the apical border of the calicoderm layer, strongly supporting that StpNRX1 is a core component of SJs in *Stylophora*. Immuno-labeling of StpClaud3 showed a different pattern. First, labeling was restricted to groups of calicodermal cells along the basal disc. In such groups, although labeling juxtaposed the F-actin labeling, the overlap between Claud3 and F-actin was only partial. In some cases, StpClaud3 encircled two or more cells (Figure 5e). Such a pattern rather suggests that StpClaud3 has a supra-cellular function within the calicoderm layer for yet-to-define specialized cells.

*Model of cnidarian SJ as a blueprint of bilaterian SJ*

Several lines of concordant evidences led us to propose a model for cnidarian SJs (Figure 6), as inferred from bilaterian SJs (Laval, et al. 2008; Shimoda and Watanabe 2009): i) protein sequence and domains conservation of the different SJ components, which suggest common functionality; ii) congruent phylogeny of bilaterian and cnidarian SJ components, which suggest evolutionary conserved function; iii) localization at the apical border of the cells for StpNRX1; iv) conserved N-linked glycosylation between StpNRX1 and its mammalian counterpart Caspr1; v) similar SJ ultrastructure in insects and anthozoans on electron micrographs. The tripartite NRX-NRG-CONT complex forms the structural base linking two adjacent cells. Coracle and Yurt proteins serve as intracellular scaffolds, possibly attaching the intracellular part of the structural components. Members of the MAGUK superfamily also serve in scaffolding and cellular polarity. Na\(^+\)/K\(^+\) transporters in SJs have been verified in various species. Our StpClaud3 labeling data substantiates Claudin-like association with SJs, although the expression of this particular Claudin appears restricted to specific cell types. Limitations of the above model include the low mRNA expression of CONT, the absence of the diverse conserved variants of NRX, and the lack of evidence for the presence of Claudin-like proteins as core-components of cnidarian SJs. However, the
model presented here accounts for both medusozoans and anthozoans, two cnidarian clades that diverged probably more than 540 million years ago (Chapman, et al. 2010): besides ultrastructural variation recognized in electron micrographs, the SJ components of the two clades are comparable and SJs should therefore be considered as structurally similar and evolutionary related.
Discussion

Data mining of representatives from the early branching metazoans using known molecular components of bilaterian occluding junctions (TJs and SJs) has conclusively identified SJs as the sole type of occluding junctions present in Cnidaria and Placozoa, thereby asserting previous electron microscope investigations on these phyla. Although the core components of SJ have not been definitively defined, Nrx, Nrg, Cont and Claudins are likely to represent the structural core components and thus their expression in early metazoan lineages is meaningful in determining the evolution of this occluding junction. In cnidarians, the SJ gene repertoire is diversified, with differential tissue expression for variants of the structural SJ components NRX and NRG, which suggests an unexpected complexity of SJs in these diplobastic animals. Although epithelium sealing properties have been documented in poriferans (Adams, et al. 2010), lack of SJ structural homologs in the poriferan Demospongiae and Homoscleromorpha as well as in Ctenophora indicates that SJ arose in metazoans before the placozoan/cnidarian/bilaterian radiation.

Epitheliozoans as defined by acquisition of septate junctions

The molecular phylogeny of the principal occluding junction components across the metazoan lineages (restricted to representative organisms with complete genomes) allows reconstructing a scenario of stepwise evolution for sealing epithelia, i.e. the emergence of body compartments (Figure 7). However, the phylogeny of early branches is not settled (Philippe, et al. 2011; Nosenko, et al. 2013). The tree presented in Figure 7 follows minimal gene loss across metazoan evolution of the SJ complement. Ctenophores are positioned at the base of the metazoan lineage, according to current studies (Ryan, et al. 2013; Moroz, et al. 2014); demosponges and homoscleromorphes are separated according to (Sperling, et al. 2009; Erwin, et al. 2011) although consensus on the mono vs paraphyly debate of poriferans has not been reached (Worheide, et al. 2012). In the protists Capsaspora and Monosiga, we identified the Na+/K+ ATPase exchanger (the Beta subunit appeared with Monosiga), MAGUK ancestors (Dlg and MPP), and Claudin-like members, which prove that these were already present in the metazoan ancestor lineage. The Na+/K+ ATPase transporter is an integral part of occluding junctions (Krupinski and Beitel 2009). Although this exchanger is required for SJ formation in insects, its function in SJs is pump-independent (Genova and Fehon 2003; Paul, et al. 2007). Interestingly, the beta subunit of the Na+/K+ ATPase transporter has been shown to create molecular bridges between two adjacent cells (Vagin, et al. 2012). This moonlighting function of the Na+/K+ transporter might have represented a
potential building block for the further development of occluding junction in epithelia (Krupinski and Beitel 2009). With multicellular animals, components of the cytosolic adaptor plaque appeared successively. Homologs of the MAGUK members Varicose (MP2,6), ZO and an ancestral form of MPP 3,4,5,7, as well as the FERM protein Coracle, arose in ctenophores. These represent cytosolic components involved in cellular polarization (and junction scaffolding) in bilaterians. Stardust (MPP5) appeared with demosponges but ZO was absent. In homoscleromorphs, Yurt as well as a putative NRG ancestor (intermediate between DSCAM and NRG) were identified. However, it is only with placozoans and cnidarians that the structural components of SJs, i.e. NRX, NRG and CONT, emerged, hereby pointing to the origin of SJs in metazoans. In bilaterians, SJs were kept as the principal type of occludin junctions in protostomes, whereas vertebrates within the deuterostome lineage evolved a specialized Claudin family (here referred as Claudins s.s.) and other structural proteins (JAM, Marvels…) that permitted a novel type of junction, the Tight Junction.

Epitheliozoa, which includes the Bilateria, Cnidaria, and Placozoa, was originally proposed to characterize animals with true epithelia defined as cell layers held together by belt desmosomes (Ax 1996; Dohrmann and Worheide 2013). Our present study extends the characteristics of the Epitheliozoa as animals with epithelia sealed by occluding junctions (TJs and SJs). Importantly, the lack of structural SJ components in poriferans was not assessed in calcareous sponges in this study, as no calcareous genome is available hitherto. However, Ledger described potential SJs in the calcareous sponge *Sycon ciliatum* using TEM experiments (Ledger 1975). Hence, genomic exploration of calcareous sponges is required before a complete picture of SJ evolution can be drawn.

Are structural SJs components derived from neuronal junctions?

Poriferans and placozoan do not have recognized neurons contrary to ctenophores and cnidarians which have well defined neurons and nerves (Moroz 2012). However candidate neurosecretory cells have been found in both poriferans (flask cells (Renard, et al. 2009)) and *Trichoplax* (fiber cells (Smith, et al. 2014)). Further, a set of protosynaptic genes have been identified in poriferans (Sakarya, et al. 2007; Conaco, et al. 2012). Thus, irrespective of whether or not Ctenophora represents the basal Metazoa, the genetic origin of the neural system starts with animal multicellularity. Central to the organization of the bilaterian neuronal network is the Neurexin-Neurolignin interaction (Bang and Owczarek 2013). Neurexins are found at the synaptic membranes and bind to Neurolignin on the opposite synaptic membrane across the 20 nm wide synaptic cleft (Sudhof 2008; Chen, et al. 2010). On
the cytosolic side, Neurexin binds to the MAGUK proteins Dlg (PSD-95) and CASK and the FERM protein 4.1/Coracle (Hata, et al. 1996; Biederer and Sudhof 2001; Chen, et al. 2005; Chen and Featherstone 2011). Thus, in addition to structural similarities (Bellen, et al. 1998), the synaptic Neurexin (Neu) and the Septate junction Neurexin (Nrx) share common cytosolic partners. Although the function of the Neurexin 1 homologs in homoscleromorphs (Nichols, et al. 2006) and ctenophores (Moroz, et al. 2014) is not known, Neu arose before Nrx, possibly originating from a Capsaspora ancestor (Figure 7 and Supplementary Figure S4). A possible scenario implies, the primary addition of EGF domains to the Capsaspora ancestor in the first multicellular animals, and subsequently, after a duplication event, the rearrangement of central LamG-EGF domains by an FBG domain, that could have led to a neofunctionalization of the Neu protein to the Nrx protein. Another molecular actor necessary to neural wiring in bilaterians and structurally related to NRG is the DSCAM cell adhesion molecule. DSCAM controls repulsion/attraction between two neurons via extracellular homophilic recognition, (Schmucker and Chen 2009). Besides, DSCAM is part of the innate immune response in arthropods via heterophilic binding to different pathogen molecules (Cerenius and Soderhall 2013). Our investigation of non-bilaterian genomes shows that DSCAM is an evolutionary conserved molecule with homologs in ctenophores, poriferans, placozoan and cnidarians (Figure 2 and 7), although localization and function are unknown. However, among the 2 DSCAM copies present in Homoscleromorpha, one (OcaDSCAM) clusters with the DSCAM family whereas the other clusters with the NRG family, likely pointing to the emergence of the NRG family. Thus, NRG could represent an evolution of DSCAM, a molecule linked to neuronal development and immunity.

**Structural SJ components appeared and diversified in Cnidarians**

The SJ components of cnidarians and bilaterians are very similar at the protein level and therefore, a common model for the SJ structure can be inferred from the characterization of SJs both in insects and mammalian paranodal junctions (Charles, et al. 2002; Bonnon, et al. 2003; Faivre-Sarrailh, et al. 2004) (Figure 6). In particular, in *Stylophora*, StpNRX1 localizes at the apical border of each cell (Figure 5), which is in strict correlation with the position of SJs observed in TEM images (Supplementary Figure S2). StpNRX1 also co-localizes with the F-actin network (Figure 5), strongly supporting a model where SJs are attached to the cytoskeleton via cytoplasmic adaptor proteins. Finally, in human, Caspr proteins associate with contactin during their biosynthesis, resulting in the expression of high-mannose glycoforms of the two proteins at the cell surface (Bonnon, et al. 2003). In this study,
StpNRX1 has been shown to be N-Glycosylated as demonstrated by the apparent molecular weight reduction after PGnaseF treatment on Western blot (Figure 5). StpNRX1 thus performs as the faithful homolog of Drosophila NrxIV and human Caspr1. Moreover, several other copies of NRX1 (StpNRX2-5) are also present in Stylophora (Figure 2). These supernumerary anthozoan specific copies display similar domain architectures in their NH2 terminus but differ in their COOH terminus (all have a transmembrane signature). These variations in domain architectures, conserved among anthozoans, may reflect functional diversification, in conjunction with specific tissue expression. StpNRX2 (a Stylophora specific NRX), which accounts for about half of StpNRX1 in the total fraction, is mostly expressed in the aboral tissue (Figure 4). Conversely, StpNRX3 is dominantly expressed in the oral tissue. This suggests that the aboral (calcifying) and oral (polyp) tissues harbor a different set of structural components for their SJs. Indeed, one of the scleractinian specific NRG copies shows preferential expression in the oral tissue. Although the cellular localization and function of these additional NRX and NRG homologs remains to be addressed, SJs with different composition may reflect structural differences and possibly result in different paracellular properties between the different tissues layers or specialized cell types. Along the same line of evidence, in the tentacle of Anemonia viridis, one of the two copies of NRX1 (AvNRX1b) is found mostly expressed in the endoderm, unlike AvNRX1a which is equally expressed in both tissue layers (Supplementary Figure S7). Thus, at least in A. viridis, the structural composition of SJs between the ectoderm and the endoderm appears to differ. Such discrepancy in NRX composition may be the cause of the differences observed in TEM images of sea anemone SJs between the endoderm with double septum and the ectoderm with single septum (Green and Flower 1980). Patchwork expression of different SJ components in tissues/layers is substantiated by the differential expression of the additional copies encoding for the cytoplasmic adaptor CORA in Corals. One surprising result in our expression analysis was the very low mRNA expression level of StpCONT as compared to StpNRX1 and StpNRG1,2, as these form a trimolecular complex in bilaterians. This result was confirmed by other independent RNAseq approaches (data not shown). In addition, data mining of other cnidarian EST databases (including Nematostella and Hydra) also showed that putative ESTs homologous to CONT were scarce. Although thought provoking, our data raise the possibility that CONT is not part of the tri-molecular core-complex that structures all SJs in cnidarians. Alternatively, the turnover of the CONT mRNA and protein may be very slow and therefore present in low copy numbers. Also, CONT may be required for specific
developmental stages or cell types. Hence, as the CONT mRNA is indeed expressed, we included CONT in our cnidarian SJ model.

**Conserved Claudin-like proteins**

Claudins were first identified in vertebrate TJs and so far 27 members have been identified in vertebrates (Gunzel and Fromm 2012). Beside these TJ specific Claudins (Claudin s.s.), other Claudin-like proteins have been identified based on sequence and tetraspan structure similarities both in human and invertebrates. In *Drosophila*, three Claudin-like proteins were functionally associated with SJs. However, Claudin phylogeny is unclear as these proteins loosely cluster in highly divergent clades (Simske and Hardin 2011). The addition of Claudin-like proteins from early branching metazoans to the Claudin repertoire highlights several clusters of evolutionarily conserved Claudin-like members. Claudins s.s. are specific to TJs, which correlates with the fact that they form an outgroup to the other Claudin-like sequences. Group Ia, Ib, II, and IV encompass human Claudin-like proteins (LIM2, PMP22 EMP1-3, LHFPL1-4, Clarin-3) proposed to have cell-cell interaction properties (reviewed in (Van Itallie and Anderson 2004; Simske and Hardin 2011; Cosgrove and Zalocchi 2013). For example, LIM2 (Group Ia) and PMP22 (Group Ib) have been shown to associate with TJ constituents and to display barrier properties (Notterpek, et al. 2001; Grey, et al. 2003; Roux, et al. 2005), whereas a LHFP member (also called TMHS, group II) was associated with hair-cell anchoring independently of TJs (Xiong, et al. 2012). What is the role of analogous Claudin-like in invertebrates? The *Stylophora* StpClau3 (group II) clearly localizes at the cell-cell border of specific cells (Figure 5), in agreement with specialized cell interaction properties. In *Hydra*, the Claudin HydClau1 also localizes to the apical junctional complexes (Bert Hobmayer, University of Innsbruck, personal communication). However, outside the three *Drosophila* Claudin-like proteins associated with SJ formation, it would be premature to involve any other invertebrate Claudin-like proteins with a particular function in SJs. A junctional interaction in *trans* between two cells is highly improbable as the distance separating adjacent plasma membranes is too large to allow kissing complexes in SJs. A function in regulating the paracellular transport across an epithelium has only been described for the TJ specific Claudin (Van Itallie and Anderson 2006). The ancient and diversified Claudin repertoire may well represent diverse conserved functions, as part of macromolecular complexes associated with the plasma membrane. Further biochemical characterization will be needed to clarify the apparent discrepancies between the Claudin phylogeny presented here and the function of Claudins inferred from vertebrates. Indeed, Claudin-like proteins are
present in the unicellular *Capsaspora* and *Monosiga* suggesting that tetraspan proteins had ancestral functions besides promoting cell-cell interaction. Claudin-like Group III appears to contain the most evolutionary conserved Claudin clade with Claudin-like members found in vertebrates, *Drosophila*, cnidarians, poriferans, *Monosiga* and *Capsaspora*; yet functional data is not available for any of them.

*Functional implications*

Occluding Junctions govern paracellular transport across epithelia. In invertebrates, SJs control this paracellular pathway, as shown in insects using conductance experiments on epithelia and by dextran injection after gene knock down (Pannabecker, et al. 1993; Lamb, et al. 1998; Banerjee, Sousa, et al. 2006). Although classified as “leaky epithelium”, as compared to the vertebrates’ “tight epithelia”, epithelia in insects are nevertheless able to show barrier properties comparable to vertebrates. For example, in female mosquitoes, Malpighian tubes maintain very high [K+] gradients and allow rapid paracellular transport of Cl− across the Malpighian epithelium after blood meals to maintain homeostasis (Beyenbach and Piermarini 2011). In cnidarians, the epithelial layer also show different permselective properties to Ca2+, Na+, and Cl− (Bénazet-Tambutté, et al. 1996), suggesting that SJs potentially control ion exchange across cnidarian tissue layers. In reef building corals, the oral and aboral tissues have specialized roles in the process of biomineralization, and the transport of ions from the surrounding sea water to the site of calcification is central to the understanding of how the calcium carbonate skeleton is formed (Tambutté, et al. 2007; Allemand, et al. 2011). Although the transcellular pathway is part of this ion transport (for recent reviews see (Allemand, et al. 2011; Tambutté, et al. 2011), experiments have raised the possibility that paracellular transport might also be involved (Tambutté, et al. 2012). Since small ions such as calcium (molecular radius 1.8 Å) should also, in principle, be able to pass via the paracellular pathway (Tambutté, et al. 2011; Tambutté, et al. 2012). However, in chordate epithelia, TJs not only regulate the flow of molecules based on the size, but also based on the charge of the molecule/atom. Although in chordates it is generally accepted that Claudins (Claudins s.s.) define the TJ permselective properties, almost nothing is known about the mechanisms that govern the flow of molecules through SJs. In other words, the respective roles of the Claudin-like, NRX, NRG and CONT proteins (or other molecules) in regulating the paracellular transport still remain to be characterized, especially in regards to tissue specific permselectivity. Further experiments using heterologous expression of SJ
components in conjunction with electrophysiological measurements will help to better understand the role of these molecules in the permselective passage of ions. In addition to shedding light onto the coral calcification process, determining the permeability/permselectivity of SJs is also of major importance in the environmental context of ocean acidification. Previous studies have shown that the decrease in pH in the oceans, due to the increase in atmospheric CO$_2$ and its dissolution into seawater, negatively affects coral calcification (Andersson and Gledhill 2013). One parameter that might explain this effect, among others, is the degree to which the site of calcification is isolated from seawater. It has been proposed that the sensitivity of corals to ocean acidification could readily be explained if the paracellular route is the major supply of ions for calcification (Erez, et al. 2011). Different studies have suggested a protective role of tissue layers against skeletal dissolution (Ries, et al. 2009; Rodolfo-Metalpa, et al. 2011). However none of them has examined the potential role of SJs in such a protection because no molecular data on junctions were hitherto available. The results presented here lay the foundations for future studies that will allow to monitor differential expression of genes involved in the formation of SJs and to determine whether they play a role in the resistance to ocean acidification.
Material and Methods

Model organisms: Cnidaria comprise two major classes, Medusozoa (including Hydrozoa) and Anthozoa (including Hexacorallia). Actiniaria and Scleractinia constitute two major subclasses of Hexacorallia. Commonly, Actiniaria are represented by sea anemones such as Nematostella vectensis (Nematostella) and Scleractinia are represented by reef building corals, such as Stylophora pistillata and Acropora digitifera (Kayal, et al. 2013), which are colonial polyps and have a specialized calcifying tissue layer (calicoderm, Supplementary Figure S1 and S2).

Sequences: all human and Drosophila melanogaster reference protein sequences listed in Supplementary Table S1 were retrieved from NCBI. Early branching metazoan sequences were retrieved from the databases listed in Supplementary Table S2. The Stylophora pistillata sequences were deduced from transcriptome and/or genome assemblies (C.R.V. and M.A.). Note that some Hydra and A. digitifera protein families were omitted in our phylogenetic analysis due to inconsistent sequences assemblies (gaps, Ns, misassemblies) of some members.

Softwares and strategy used: BLAST (2.2.22) genome/transcriptome analysis was run locally, at NCBI and JGI depending on organisms’ database. An online version of MAFFT (mafft.cbrc.jp/alignment/server/) was used with strategy L-INS-i default parameters. Prottest (v2.4), PhyML (v3.0), MrBayes (v3.2.1) and FigTree (v1.3.1) (Huelsenbeck and Ronquist 2001; Abascal, et al. 2005; Rambaut and Drummond 2009; Guindon, et al. 2010) were run locally. PFAM (pfam.sanger.ac.uk/) (and SMART (http://smart.embl-heidelberg.deSMART/set_mode.cgi) were used to predict protein domains. Transmembrane domains were predicted at www.cbs.dtu.dk/services/TMHMM/. GPI anchors were predicted using the webtools “FragAnchor” (http://navet.ics.hawaii.edu/~fraganchor/NNHMM/NNHMM.html), PredGPI (http://gpcr2.biocomp.unibo.it/gpipe/pred.htm) and “GPI-anchored Protein Prediction” (http://bolero.bi.a.u-tokyo.ac.jp:8201/GPI-Predictor/), which gave similar results.

Blast searches were run on the cnidarians databases to identify putative homologs using Drosophila and Human protein sequences of the molecular SJ and TJ constituents (Supplementary Figure S3). All newly identified protein sequences were added to the previous pool of query sequences for iterative BLAST searches to identify novel potential homologs. Selection criteria were based on reciprocal BLAST against the Human and...
Drosophila RefSeq databases as well as domain homology. Once identified in cnidarians, searches for homologs were further carried out in Trichoplax, then in Amphimedon and finally in Monosiga, always using the entire pool of identified proteins as bait. When homologs were missing in Amphimedon, BLAST searches were carried out against the whole NCBI sponge database before deemed absent. For each family of SJ components identified, protein sequences were aligned using MAFFT. Alignments were trimmed to the largest conserved part of the proteins (Supplementary File S1) and then subjected to phylogenetic analyses. Best substitution matrices and parameters were calculated using Prottest before running PhyML. Alternatively, Bayesian analyses using MrBayes were run with default specification until convergence reached standard deviation below 0.01, except for Claudins, which were stopped after 7 million MCMC generations. Resulting trees were visualized using FigTree.

Oral disc dissection: fragments of the same Stylophora pistillata colony were used for both RNA and protein extractions. Fragments were set to rest in a glass petri dish filled with sea water until polyps were extended. 0.4% stock solution of Tricaine mesylate (MS-222, Sigma) dissolved in sea water was added into the petri dish to a final concentration of 0.04% and left to rest under dimmed light for 15 minutes. Subsequently, oral discs (the apparent portion of the polyp, Supplementary Figure S1) were cut from the colony under binocular using 5 mm blades micro-dissection scissors (Vannas). Batches of 10-15 oral discs were collected and transferred into Trizol® or TNE solutions (see below). Dissections were stopped after a maximum of 45 minutes of MS-222 incubation to elude any potential secondary effect of the drug.

RNA extraction. Freshly dissected oral discs were put into Trizol® and homogenized for 1 minute using an electrical potter. Alternatively, entire fragments of colony were cryo-crushed (Spex sampleprep® 6770) and the resulting powder was dissolved in Trizol®. RNA extraction was carried out using a standard protocol (Moya, et al. 2008). Extracted RNAs were treated with RNase-free DNaseI (Roche) and precipitate with NaAcetate/EtOH. Concentrations were determined by spectrophotometry using an Epoch Microplate Spectrophotometer (BioTek).

Real-time qPCR, cDNAs were synthesized using the Superscript®III kit (Invitrogen). qPCR runs were performed on an ABi 7300 using “EXPRESS SYBR® GreenER™ qPCR Supermix with Premixed ROX” (Lifetechnologies) for PCR amplification. Experimental procedures were performed as in (Moya, et al. 2008). Data were either relative to stpNRX1 [dCt =
(Ct\text{gene\_of\_interest} - Ct\text{NRX1})_{\text{Total}}) for the expression in the whole colony (Total), or normalized to stpNRX1 and relative to Total [ddCt = (Ct\text{gene\_of\_interest} - Ct\text{NRX1})_{\text{Oral}} - (Ct\text{gene\_of\_interest} - Ct\text{NRX1})_{\text{Total}}] for the expression in the oral disc (Oral) as compared to Total. Fold expressions were further extrapolated using the $2^{\text{ddCt}}$ and $2^{\text{ddCt}}$ function.

Protein extraction protocol for enriched membrane proteins (all step on ice). Freshly dissected *Stylophora* oral discs were kept in cold TNE buffer [100 mM Tris pH=7.2; 100 mM NaCl; 5 mM EDTA; 1X Protein Inhibitor Cocktail (Sigma)]. Alternatively, total tissue was extracted from the skeleton in cold TNE using the air-pick (i.e. pressurized air through a pipet tip) method on coral fragments. Batches of 1 ml crude extract were then dilacerated by repeatedly passing them through a 21" syringe gauge until homogeneity was reached. Extracts were centrifuged at 1000 g for 1 minute at 4°C. The supernatant (S1) was collected and kept on ice. The pellet was resuspended in TNE, passed through a syringe and centrifuged again. The supernatant (S2) was pooled with S1 and triton X-100 was added to a 1% final concentration. The resulting intermediate extract was incubated for 1hr at 4°C on a rotating wheel and then centrifuged at 15,000 g for 15 minutes at 4°C. The resulting supernatant was the final total extract, now cleared of zooxanthellae but enriched in membrane proteins. Protein concentrations were measured by comparison to a BSA standard curve. For the N-linked glycosylation analysis, 250 µl of *Stylophora* total extract was divided into two reaction tubes; half was incubated with 2 µl 2-mercaptoethanol and 8 µl PGNaseF (Roche) for 3H at 37 °C and the other half was used as control.

Custom made antibodies (Eurogentec). Two antibodies were produced in rabbit using synthetic peptides, one against the peptide KTNPYDPTSGRTDDD (AA 1057-1073) corresponding to the beginning of the extracellular part of StpNRX1 and the other one against the peptide GRMASHGYNNQDTTTL (AA 220-236) corresponding to the COOH terminus of the StpClaud3. For each antibody, ten rabbits were initially screened for non-cross-reactivity with *Stylophora* proteins and two were selected for the Speedy program. Each selected antibody was affinity purified before use.

Western blotting. Equal amounts of protein extracts were loaded onto 6% (NRX1), gradient 4-15% (actin) and 15% (Claud3) TGX precast gel (Bio-Rad). After electrophoresis, gels were transferred onto PVDF membrane and blotted using SNAP i.d.® with anti-Stp\_NRX1 (1:200), anti-StpClaud3 (1:200), anti-actin (mouse A4700 Sigma) (1:500) primary antibodies, and HRP-coupled goat anti-rabbit (Sigma) (1:2000) or HRP-coupled goat anti-mouse (Sigma)
(1:2000) secondary antibody. ECL was conducted using Amersham ECL detection reagents (GE Life Sciences). Imaging was carried out on a Fusion Fx7 (Peqlab).

**Immunolocalization.** One microcolony of *Stylophora* grown on a slide (Venn, et al. 2011) was fixed in 25 ml chilled artificial-sea-water / parafomaldehyde (PAF) fixation buffer [425 mM NaCl; 9 mM KCl; 9.3 mM CaCl2; 25.5 mM MgSO4; 23 mM MgCl2; 2 mM NaHCO3; 100mM HEPES pH=7.9; 4.5% PAF] for 2H on ice. The microcolony was transferred into a 50 ml Falcon tube and decalcified in 50 ml [100 mM HEPES pH=7.9; 500 mM NaCl; 250 mM EDTA pH=8.0; 0.4% PAF] (renewed after 48h) at 4°C until dissolution of the skeleton (3-5 days). The remaining soft tissue was transferred into a small petri dish containing 10 ml decalcifying buffer/1X PBS (50/50 v/v). Here, basal discs were cut with 5 mm blades micro-dissection scissors under a binocular. Basal discs were collected in 1X PBS and rinsed 3 times for 5 minutes. Samples were blocked in [1X PBS; 0.05% Tween_20 (PBST); 2% BSA; 2% donkey serum; 0.1% Triton_X100] for 2H at 4°C. Samples were incubated in antibody solution [PBST; 1% BSA] with either anti-Stp_NRX1 or anti-StpClaud3 (dilution 1:25) for 2 days at 4°C, then rinse 3 times 10 minutes in [PBST; 0.1% BSA] and further incubated for 1 day at 4°C in antibody solution supplemented with 10 μl Phalloidin-Alexa568 and anti-rabbit-Alexa488 and anti-rabbit-Alexa405 (dilution 1:200 each). Finally, samples were rinsed 3 times for 5 minutes in PBST, mounted in ProLong® Gold antifade reagent (Molecular Probes) and left for 24h in darkness.

**Imaging.** Confocal imaging was performed using a Leica SP5 and the LAS AF lite software. For imaging, each channel was acquired sequentially to ascertain lack of cross-emission. Merging was achieved using the LAS AF lite tools option. Light microscope images were acquired using a Leica Macroscope Z16 APO. Sample preparation and electron micrographs obtained with a JEOL transmission microscope were described in (Tambutté, et al. 2007). Image contrast and brightness were adjusted with the Photoshop levels tool.

**Aknowledgments:** Thanks are due to Natacha Segonds and Nathalie Techer for technical assistance and Dominique Desgré for coral maintenance. We are very Grateful to M.L. Hernandez-Nicaise for discussion and photograph on intercellular junction in Ctenophora.
This work was supported by the Centre Scientifique de Monaco research program, funded by the Government of the Principality of Monaco. This project was partially funded by KAUST baseline funds to CRV and MA.
**Figure Legends**

Table 1: Molecular components of occluding junctions in representative eukaryotes.

List of the major components of Tight Junctions (TJs) and Septate Junctions (SJs) in Human and *Drosophila*, and the respective protein homologs in *S. pistillata*, *N. vectensis*, *H. magnipapillata* (Cnidaria), *T. adherens* (Placozoa), *O. carmela* (Porifera, Homoscleromorpha), *A. queenslandica* (Porifera, Demospongiae), *M. leidyi* (Ctenophora), *M. brevicollis* (Choanoflagellata), and *C. owczarzaki* (Filasterea). Numbers in brackets refer to the number of homologs found. “Not found” means absence of homologs whereas “nd” stands for “non-determined” due to limitations in the assembly of the reference sequence (see Material and methods). Claudins were arguably separated into two subgroups: Claudin s.s. (Claudin sensu stricto) refers to the vertebrate Claudins (1-27) that are unique to the vertebrate TJs, whereas Claudin-like proteins encompass other members of the tetraspan family that are similar to Claudins in structure and also belong to the PFAM families PF00822, PF13903, and PF10242 (see Table S1).

**Figure 1: Phylogenetic trees of Claudins**


**Figure 2: Phylogenetic trees of NRX, NRG-CONT**

Unrooted Bayesian phylogenetic tree (rectangular) of holozoan Neurexins (A) and NRG, CONT and DSCAM (B). The different taxa are color shaded. The domain arrangements of each protein are schematized on the right hand side of the trees. Nrx and correspond to Caspr/NrxIV and Neurexin1-3 homologs, respectively; NRG, CONT and DSCAM correspond Neurofascin, Contactin and Down Syndrome Cell Adhesion Molecule homologs, respectively. Hs: human; Dm: *Drosophila*, Adi: *A. digitifera*; Stp: *S. pistillata*; Nv: *N. vectensis*; Hyd: *Hydra*; Tri: *Trichoplax*; Oca: *Oscarella*; Amphi: *Amphimedon*; Mle: *Mnemiopsis*; Pba: *Pleurobrachia bachei*; Cow: *Capsaspora*.

**Figure 3: Phylogenetic trees of MAGUKs and 4.1/Coracle-Yurt**
Unrooted Bayesian phylogenetic tree (rectangular) of holozoan MAGUKs (A) and 4.1/Coracle-Yurt (B). The domain arrangements of each protein are schematized on the right hand side of the trees. Vari, mena3, Stard, and Cora abbreviate Varicose, menage a trois, Stardust and Coracle respectively. Hs: human; Dm: Drosophila, Adi: A. digitifera; Stp: S. pistillata; Nv: N. vectensis; Tri: Trichoplax; Oca: Oscarella; Amphi: Amphimedon; Mle: Mnemiopsis; Mon: Monosiga; Cow: Capsaspora.

Figure 4: Expression of the principal molecular components of Septate Junctions (SJ) in S. pistillata.

a-b) Real-time PCR expression analysis of the SJ components NRX, NRG-CONT, Claudins, Cora-Yurt, and the calicoblast specific three-fingered protein TFZPD9. The S. pistillata three finger domain protein TFZPD9 is the nearest homolog of the Pocillopora damicornis Pdcyst-rich protein specifically expressed in the calicoderm (Vidal-Dupiol, et al. 2009) and served as a tissue positive control. Expressions were measured in (a) total tissues (whole coral fragment), relative to NRX1, or in (b) oral disc versus total tissues, normalized to NRX1 and relative to total. c) Western blotting of the total and oral discs protein extracts with anti-Stp_NRX1, anti-StpClaud3 and anti-actin. Note that the pseudo-band around 75KDa (*) in NRX1 corresponds to a compression of middle to low MW non-specific bands apparent in this 6% polyacrylamide gel that are barely detectable on gradient 4-15% gels.

Figure 5: Immunolocalization of Stp_NRX1 and StpClaud3 and N-linked glycosylation of Stp_NRX1

a) Bright-field microscope image of the calcifying tissue including basal discs, after decalcification. b) Phalloidin-Alexa 568 staining of a basal disc showing the F-actin network framing each cell (Z-stack projection). c1-3) Portion of a basal disc showing the near complete superimposition of the NRX1 (c1) and the F-actin (c2) labeling (Z-stack projection). d1-4) Optical cross-section through the calicoblastic epithelium showing the co-localization of NRX1 (d1,2) with F-actin (d3). e1-4) Detail of a group of cells where Claud3 labeling is visible. Claud3 (e1,2) juxtaposes with F-actin (e3) although the overlap is partial (e4) (Z-stack projection). Scale bars are in μm. f) Western blot of S. pistillata total protein extract after PGNaseF treatment with anti-Stp_NRX1. After removal of N-linked glycosylation, the original 141 KDa NRX1 has a reduced molecular weight of 125 KDa (as estimated with Fusion software tools).
The relative arrangement of the different SJ components in cnidarians is inferred from the models of SJs in *Drosophila* and human paranodes. We distinguished the structural components NRX, NRG and CONT from the cytosolic adaptor protein families CORA/YURT and MAGUK shown to be involved in SJ scaffolding in bilaterians. Although Claudins are involved the permselectivity of Tight Junctions, a similar role in SJs has not yet been demonstrated.

**Figure 7: Proposed evolution of occluding junctions.**

Model of the evolutionary emergence of occluding junctions based on the presence/absence of the different occluding junction (TJ+SJ) components in diverse representative taxa of Holozoa. See text for details.
Supplementary Information

Figure S1: Coral anatomy.

Figure S2: Transmission electron microscopy of the oral and aboral tissue.

Figure S3: Analysis pipeline

Figure S4: Radial tree representation of the NRX, NRG-CONT, MAGUK, and CORA-YURT

Figure S5a-d: PhyML and/or MrBayes Trees of a) Claudins; b) NRX, NRG-CONT; c) MAGUKs and CORA-YURT; d) Na+/K+ ATPase transporter alpha and beta

Figure S6: Domain structure of the best NRX, NRG-CONT homologs in demosponges.

Figure S7: Relative Ectoderm-Endoderm expression of several SJ components in Anemonea viridis

Figure S8: The ctenophore apical belt junction

File S1: Alignments (Nexus) used to compute Trees relative to SJ components:
Nexus_align_seq.txt

Table S1: Sequences used in this analysis: Occluding_Junction_sequences.xls

Table S2: Web site location of the different databases used for the phylogeny: databases.xls

Table S3: Primer used for PCR amplification: Primers.xls
References


