


RESEARCH ARTICLE

Open Access



# Evolution of mechanisms controlling epithelial morphogenesis across animals: new insights from dissociation-reaggregation experiments in the sponge *Oscarella lobularis*

Amélie Vernale<sup>1,2</sup>, Maria Mandela Prünster<sup>2,3</sup>, Fabio Marchianò<sup>3</sup>, Henry Debost<sup>2</sup>, Nicolas Brouilly<sup>2</sup>, Caroline Rocher<sup>1</sup>, Dominique Massey-Harroche<sup>2</sup>, Emmanuelle Renard<sup>1,2†</sup>, André Le Bivic<sup>2†</sup>, Bianca H. Habermann<sup>2,3\*†</sup> and Carole Borchiellini<sup>1\*†</sup> 

## Abstract

**Background:** The ancestral presence of epithelia in Metazoa is no longer debated. Porifera seem to be one of the best candidates to be the sister group to all other Metazoa. This makes them a key taxon to explore cell-adhesion evolution on animals. For this reason, several transcriptomic, genomic, histological, physiological and biochemical studies focused on sponge epithelia. Nevertheless, the complete and precise protein composition of cell–cell junctions and mechanisms that regulate epithelial morphogenetic processes still remain at the center of attention.

**Results:** To get insights into the early evolution of epithelial morphogenesis, we focused on morphogenic characteristics of the homoscleromorph sponge *Oscarella lobularis*. Homoscleromorpha are a sponge class with a typical basement membrane and adhaerens-like junctions unknown in other sponge classes. We took advantage of the dynamic context provided by cell dissociation-reaggregation experiments to explore morphogenetic processes in epithelial cells in a non-bilaterian lineage by combining fluorescent and electron microscopy observations and RNA sequencing approaches at key time-points of the dissociation and reaggregation processes.

**Conclusions:** Our results show that part of the molecular toolkit involved in the loss and restoration of epithelial features such as cell–cell and cell–matrix adhesion is conserved between Homoscleromorpha and Bilateria, suggesting their common role in the last common ancestor of animals. In addition, sponge-specific genes are differently expressed during the dissociation and reaggregation processes, calling for future functional characterization of these genes.

**Keywords:** Adhaerens junction, Basement membrane, Evolution, Cytoskeleton remodeling, Transcriptome, Porifera

\*Correspondence: bianca.habermann@univ-amu.fr; carole.borchiellini@imbe.fr

<sup>†</sup>Emmanuelle Renard, André Le Bivic, Bianca H. Habermann and Carole Borchiellini—Co PIs

<sup>1</sup> Aix Marseille Univ, CNRS, IRD, IMBE UMR 7263, Avignon Université, Institut Méditerranéen de Biodiversité et d'Ecologie Marine et Continentale, Station Marine d'Endoume, Marseille, France

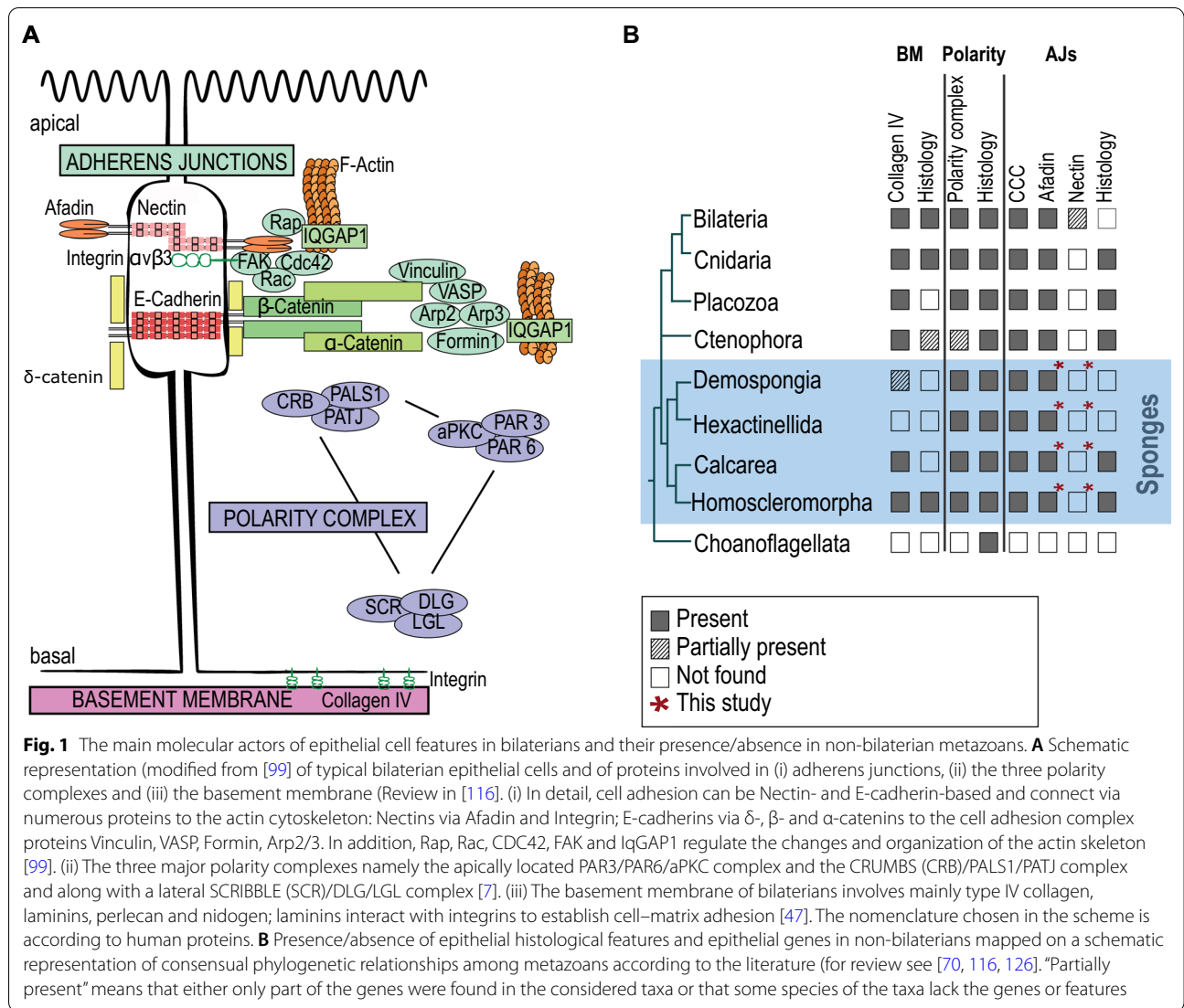
<sup>2</sup> Aix Marseille Univ, CNRS, UMR 7288, Developmental Biology Institute of Marseille Luminy (IBDM), Marseille, France

Full list of author information is available at the end of the article

## Background

The presence of epithelial tissues is a fundamental feature of all Metazoa. They are the first visible sign of cellular differentiation appearing as early as the blastula stage, and are a starting point for building animal bodies during development. The acquisition of the epithelium represents a pivotal evolutionary innovation at the origin of the establishment of a multicellularity in the last





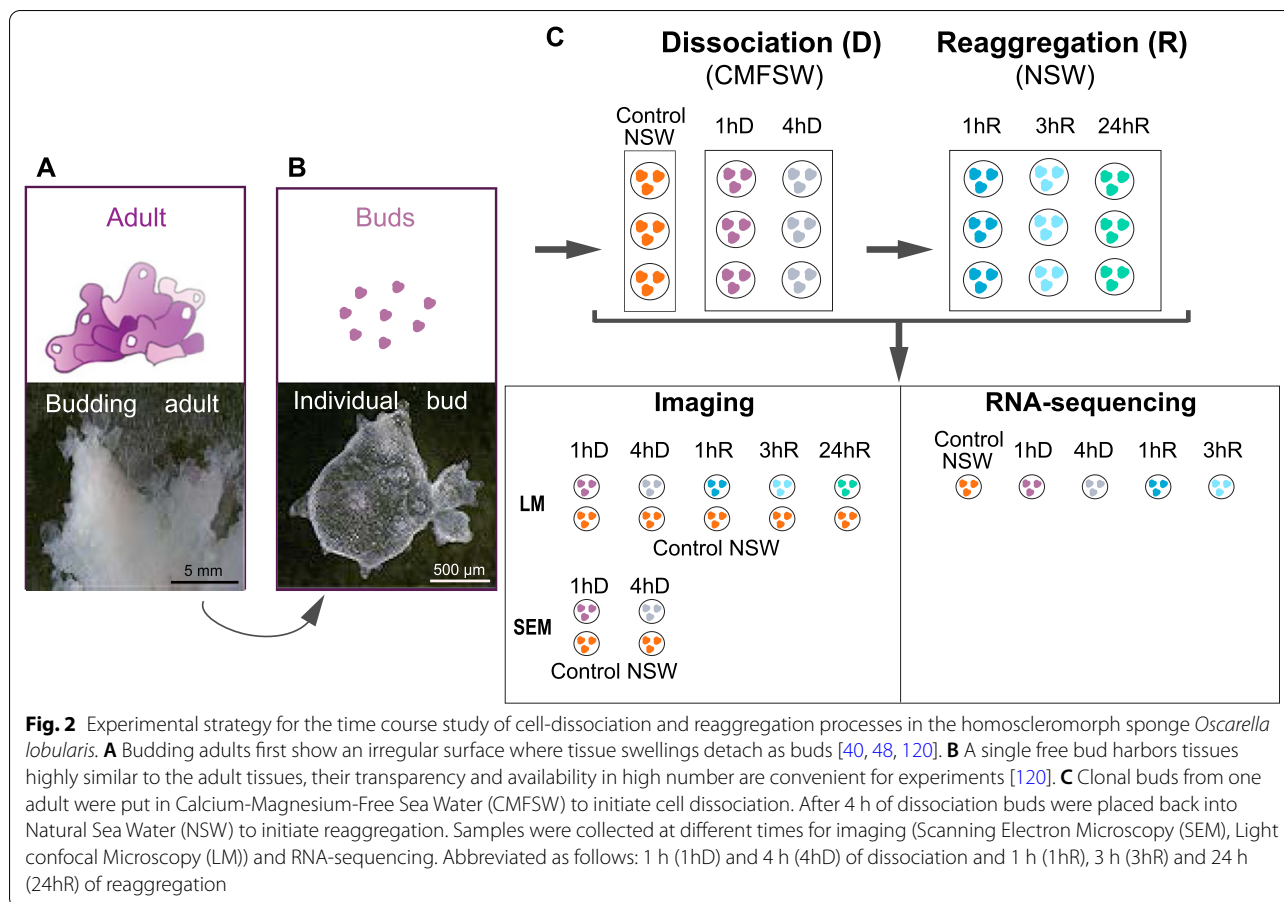
metazoan common ancestor (LMCA) [12, 22, 84, 85, 96, 112, 116, 135].

According to bilaterian features, an epithelium is usually defined as a layer of cells (i) showing coordinated polarity controlled by three polarity complexes (Crumbs, Scribble and Par); (ii) connected by cell–cell junctions (adhesive, communicating and sealing junctions) and (iii) attached by their basal pole to a basement membrane (containing type IV collagen) via cell–matrix junctions involving Integrins [135] (Fig. 1A). Such a typical bilaterian organization of the epithelium, fitting all three criteria is however, not commonly present in non-bilateria lineages except Cnidaria (reviewed in [116]) (Fig. 1B). It is therefore necessary to characterize and compare epithelial features of the three remaining non-bilateria phyla (Porifera, Ctenophora and Placozoa). This will allow us to

understand how and when the cnidarian-bilaterian epithelial features emerged.

While phylogenetic relationships at the base of the animal tree have been heavily debated for over ten years, Porifera (sponges) are frequently considered as one of the most basal branches in the metazoan tree [26, 43, 46, 49, 70, 74, 88, 89, 126, 131]. Porifera therefore constitute an important group in which to trace the emergence of epithelia.

The phylum Porifera is divided into four distinct classes (Calcarea, Demospongiae, Hexactinellida and Homoscleromorpha) each harboring distinct epithelial features according to histological observations [38, 41, 47, 84, 85] (Fig. 1B). Among them, only homoscleromorph sponges (Homoscleromorpha) possess tissues that meet many of the criteria established for bilaterian epithelia.



Indeed, the three other sponge classes (Demospongiae, Hexactinellida and Calcarea) lack a basement membrane and are usually considered to be devoid of adherens-like junctions according to histological features [41, 84, 85, 111, 135]; even though recent studies may challenge this statement [128].

Despite these histological differences, comparative genomic and transcriptomic analyses provided evidence that all four sponge classes possess a whole set of genes encoding core proteins needed to build bilaterian epithelia (catenins, cadherins, polarity complexes) [11, 44, 50, 72, 110, 119]. According to this discrepancy between histological and genetic data (reviewed in [117]) and despite recent efforts to understand the differences by modern immunodetection and proteomic approaches [97, 98, 128], it is partly unknown which protein interactions are involved in sponge cell–cell junctions and which molecular mechanisms are involved in epithelial morphogenetic processes in sponges [58]. Moreover, experimental and functional approaches are currently lacking to establish links between candidate genes of epithelial function and their effective purpose in the cell.

To explore the molecular mechanisms involved in epithelial morphogenesis, we focused on the homoscleromorph sponge *Oscarella lobularis* and took advantage of the dynamic context offered by cell dissociation-reaggregation experiments [24, 34, 52, 66, 68, 78, 79, 81, 91, 120, 137]. We carried out cell dissociation-reaggregation experiments in buds (resulting from asexual reproduction), taking advantage of the high number of replicates and the transparency of tissue which enable fluorescent imaging [120]. The observation of the dissociation and the reaggregation processes at different time-points (by confocal imaging and electronic microscopy) allowed us to monitor the time-scale of the processes and determine time-points of key events (Fig. 2). Even though numerous processes might be affected by such a treatment, we chose to focus our analyses on epithelia and more specifically on the disassembly and restoration of the internal epithelium, the choanoderm. This tissue is composed of choanocytes, a highly specialized cell type organized in hollow spheres called choanocyte chambers. This tissue has epithelial properties: the presence of a basement membrane composed of type IV collagen and highly polarized cells with cell–cell junctions histologically

(See figure on next page.)

**Fig. 3** Structural changes in the choanoderm during dissociation. **A–C''** Control conditions in natural sea water (1 h NSW): **(A)** Schematic view of a choanocyte chamber organized in a round shape with choanocyte surrounded by an apical microvilli (Mi) rich in actin and a flagellum (Fl) composed by  $\alpha$ -tubulin. PhaE: PhaE lectin staining, Nu: Nucleus **(A')** Confocal microscopic view (LM) of a choanocyte chamber, the classical round shape of the choanocyte chamber is visible (dotted orange line). The choanocyte (orange line) is recognizable with its typical collar of apical microvilli (blue arrow). The type IV collagen staining (in magenta) is lining the basal pole of choanocyte chambers (pink arrow). **B** LM focused view on the actin-rich cellular junction between choanocytes (yellow arrow). **C** Scanning Electron Microscopic (SEM) view of a choanocyte chamber with collar of apical microvilli (cyan arrow), and a basement membrane lining the choanocyte chamber (pink arrow). **C'** Focus on the cell–cell junction (yellow arrow), basal lamina (pink arrow) and apical microvilli (cyan arrow). **C''** Choanocyte with its complete apical microvilli collar (cyan arrow) and complete basal lamina (pink arrow). **D–F'''** Choanocyte chambers observed one hour after incubation in calcium-magnesium-free sea water (1hD CMFSW): **D** The general architecture of the choanocyte chamber is still observable in LM (dotted orange line), immunostaining of type IV collagen is not visible anymore (dotted pink arrow), apical microvilli are disintegrated (dotted cyan arrow). **E** Focus (LM) on a choanocyte with its disintegrated apical microvilli (dotted cyan arrow) and numerous basal actin protrusions (red arrow). **F** SEM view of a choanocyte chamber. Choanocytes present vacuoles (black arrow), disintegrated apical microvilli (dotted cyan arrow) and protrusions (red arrow). The basal lamina is detached from the choanocytes (dotted pink arrow). **F'** Focus (SEM) on the disintegration of the apical microvilli (cyan dotted arrow) and on the loss of cell–cell contact (dotted yellow arrow). **F''** Focus on basal actin protrusions at the basal pole of a choanocyte (SEM) (red arrow). **F'''** Focus on two choanocytes with a clear detachment from the basal lamina (dotted pink arrow). **G–I''** Choanocyte chambers observed four hours after incubation in calcium-magnesium-free sea water (4hD CMFSW): **G** LM view of a destructured choanocyte chamber (dotted orange line), immunostaining of type IV collagen is not visible and the general architecture of the choanocyte chamber is no more recognizable (yellow dotted line). **H** (LM) Focus on a choanocyte with numerous actin basal protrusions (red arrow). **I** SEM view of choanocyte showing actin protrusions (red arrow) and the whole destructuring of the basal lamina (dotted pink arrow). **I'** (SEM) focus on the actin protrusions (red arrow). **I''** The basal lamina is completely lost at this time-point (SEM) (dotted pink arrow). For all confocal (LM) pictures: Staining in grey: Phalloidin, Green: PhaE lectin, Blue: DAPI, Magenta: type IV collagen. Additional picture of control condition in NSW is available in the Additional file 2: Figure S2B

similar to adhaerens junctions [15–17, 38, 39, 41, 54, 84, 85, 120]. In parallel, we performed high-throughput RNA sequencing at selected time-points (Fig. 2) and looked at temporal expression dynamics to identify genes involved in epithelial patterning (loss of adhesion and polarity, and re-epithelization) using a de novo assembled and annotated *Oscarella lobularis* transcriptome. This unprecedented approach to study epithelial morphogenesis in a sponge provides a large set of new data and insights on the evolution of the metazoan toolkit involved in epithelial morphogenesis. In addition, the same dataset we provide could be explored by others to study other processes.

## Results

### Loss of epithelial features in the choanoderm during dissociation

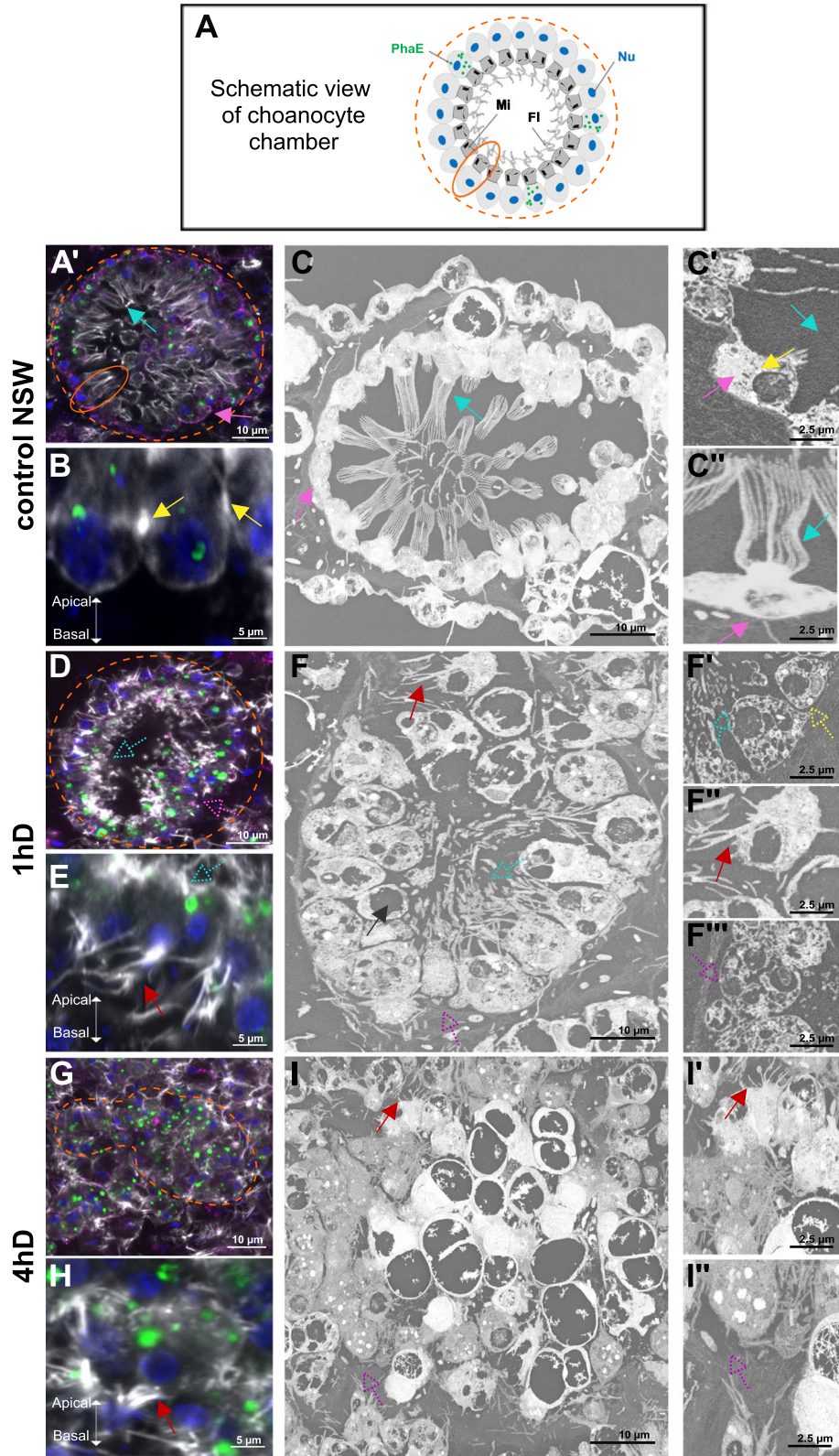
Live buds of *Oscarella lobularis* were first incubated with fluorescent labeled lectin PhaE (*Phaseolus vulgaris Erythroagglutinin*) to specifically stain choanocytes [14, 120] (Additional file 1: Figure S1). Then, cell dissociation was induced by depletion of divalent cations in calcium-magnesium-free artificial sea water (CMFSW). As a control, we used buds incubated in natural sea water (NSW) in culture plates for the same duration. Choanoderm features were observed by both confocal microscopy imaging (after type IV collagen or acetylated tubulin immunostaining, phalloidin and DAPI counter-staining) and Scanning Electron Microscopy (SEM-3D).

The general architecture of the choanocyte chamber is organized as a round structure (Fig. 3A) and for all the incubation times of the control (NSW), the choanocyte

chambers looked like as described before in this species [16, 39, 120]: they appeared as a rounded and organized structures (Fig. 3A', C). Choanocytes had a canonical shape and were highly polarized. Their larger basal side rested on a basal lamina and they had an actin-rich collar of microvilli surrounding a flagellum at their thinner apical side (Fig. 3A', C, C', C''). Cell–cell contacts could be identified as an actin-rich ring around each cell (Fig. 3B). 3D reconstructions of SEM images showed that cell–cell contacts on the lateral side were almost non-existent below the actin ring, indicating that the tissue cohesion mainly relied on these thin lateral adhaerens-like junctions and on a large area of latero-basal cell–matrix contacts (Fig. 3C', C'').

After incubating the buds for either 20 min, 1, 2, 3 or 4 h in CMFSW, we identified two main steps in the dissociation process.

After 1 h of incubation in CMFSW (1hD), the choanocyte chambers could still be identified and were forming cohesive structures (Fig. 3D, F). However, the shape of choanocytes was strongly affected compared to the control condition; they adopted a round shape and the collar of actin microvilli at their apical side lost its organization. The microvilli displayed some degree of fragmentation (Fig. 3D, E, F, F'; Additional file 2: Figure S2A–1hD). At this stage, 99.4% of the choanocyte chambers were affected ( $n = 178$ ; Additional file 3: Table S1A). In addition, large vacuoles were present in choanocytes (Fig. 3F). In parallel, F-actin containing protrusions were much more numerous at the basal pole of choanocytes (Fig. 3E, F, F'') and cell adhesion, including both cell–cell



**Fig. 3** (See legend on previous page.)

contacts (Fig. 3F') as well as cell-basal lamina contacts, (Fig. 3F'') were lost. This disorganization of the basement membrane was also estimated by the decrease of type IV collagen labeling (Fig. 3D) surrounding these structures.

After four hours of calcium/magnesium deprivation (4hD), the choanocyte chambers were disorganized and only remained recognizable due to PhaE staining. They formed compact irregular pseudo-spheres, the internal cavity and the former monolayered epithelial organization of the choanoderm were no longer present in any of the observed chambers (Fig. 3G, I; Additional file 2: figure S2A-4hD and B-4hD for the controls). At this stage, 100% of the choanocyte chambers were unstructured (n=216; Additional file 3: Table S1B). The apical collar of microvilli was completely absent (Fig. 3G, H). F-actin protrusions were still present at the basal pole of these cells (Figs. 3H, I, I'). The flagellum (anti-acetylated tubulin immunostaining) was still present despite the absence of a clearly polarized cell-body (Additional file 4: Figure S3A). Moreover, type IV collagen staining was lost (Fig. 3G) and the basement membrane was not visible anymore (Fig. 3I, I').

All phenotypes observed were reproducible between different individuals as well as between the majority of choanocyte chambers of the same bud (see images at lower magnification provided in Additional file 2: Figure S2A; statistical analyses are provided in Additional file 3: Table S1).

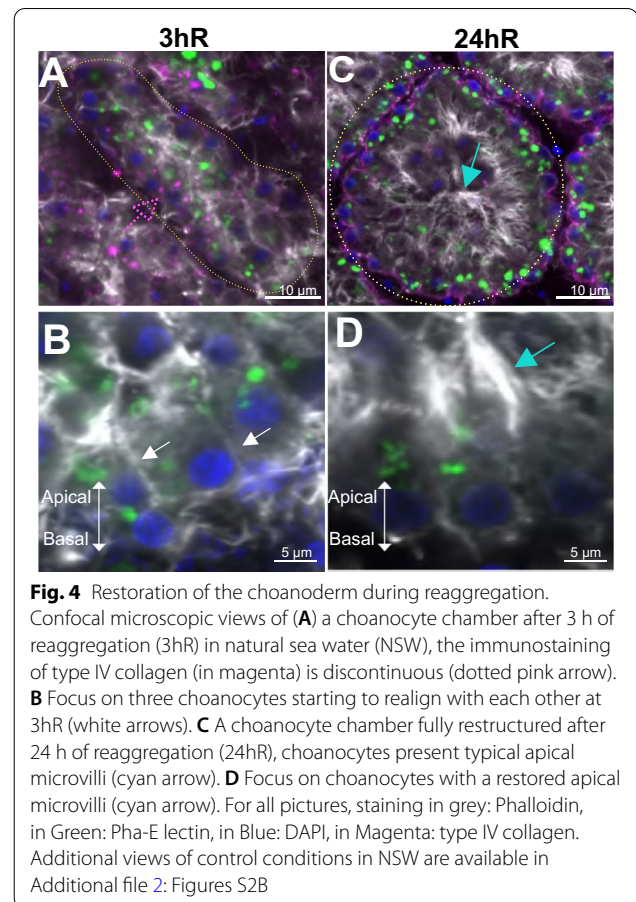
In conclusion, after four hours of calcium-magnesium deprivation, the main epithelial features of the choanoderm were lost, defined by the loss of the basement membrane, of cell–cell junctions, of cell–matrix contacts and loss of a clear cell polarity.

#### Reaggregation, reorganization and re-epithelialization of the choanoderm

Buds dissociated previously by incubation for 4 h in CMFSW were put back into NSW. By monitoring the reaggregation process at different time-points (30 min, 1, 2, 3, 5, 6, 24 h), we were able to define three main steps.

After 1 h incubation in NSW (1hR) (Additional file 2: Figure S2C; Control in S2B-1hR), PhaE-stained cells (previous choanocytes) started to realign, although the structure of the chamber was not recovered at this step compared to the control.

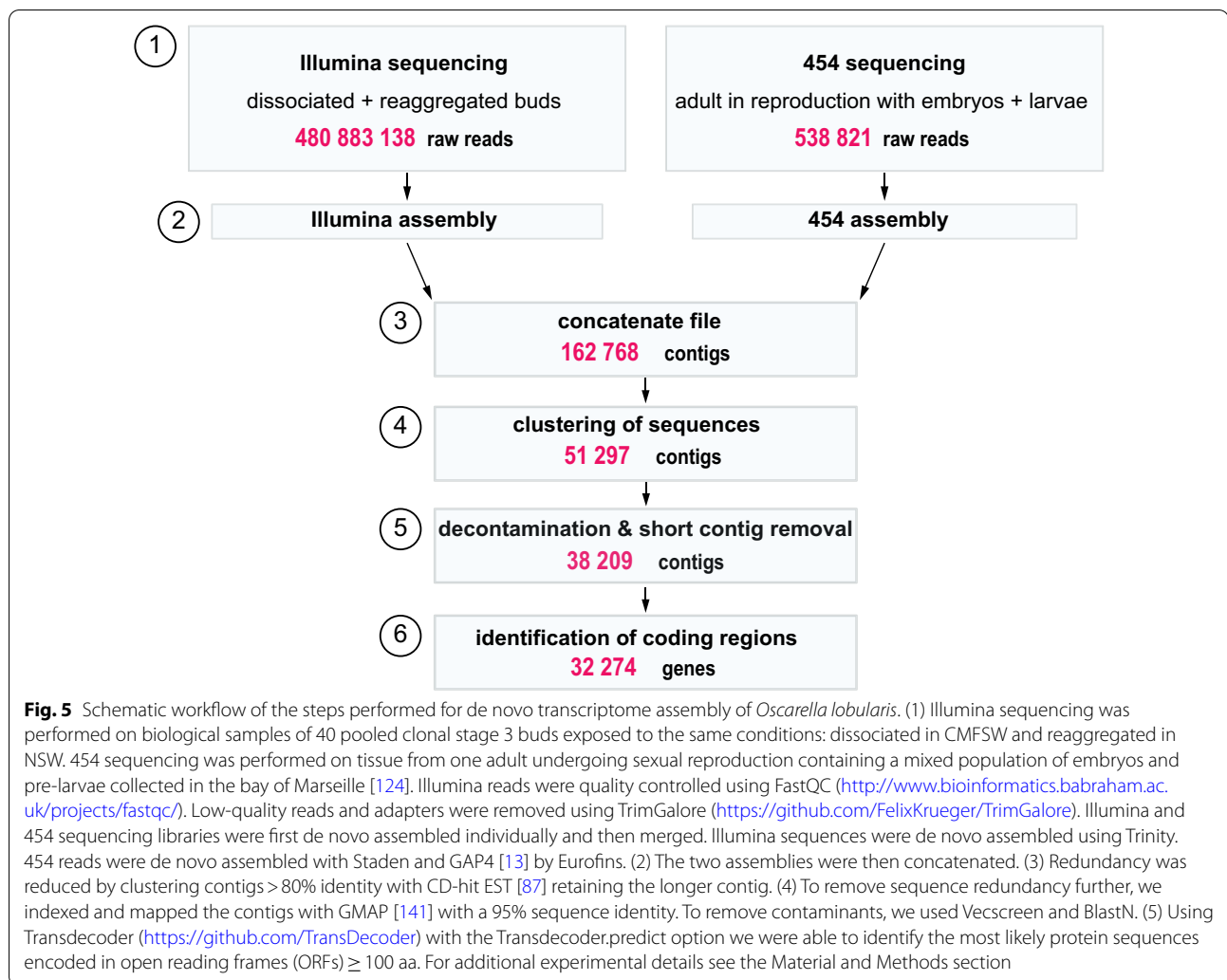
After 3 h in NSW (3hR) (Fig. 4A, B; Control in Additional file 2: Figure S2B-3hR) an accumulation of F-actin detectable in the center of PhaE-stained cells indicated not only the reformation of a small central cavity but also the reformation of a microvillar collar: cell polarity was already re-established at that time-point and the cells began to reorganize into an monostratified cell layer (Fig. 4A, B). Choanocytes re-established contacts with



each other and with the basement membrane that began to recover. Indeed, a discontinuous, spot-like type IV collagen immunostaining could again be observed around these structures. Thus, at that time-point, the main epithelial features were being recovered: basement membrane, cell–matrix and cell–cell adhesion, as well as cell polarity. The cells were not only reaggregated and reorganized (compared to 1hR) but underwent a re-epithelialization process.

After 24 h in NSW (24hR), 96.8% of choanocyte chambers recovered their initial structure (n=94, Additional file 3: Table S1C) and choanocytes regained their classical morphology (Fig. 4C, D; Control in Additional file 2: Figure S2B–24hR). The basement membrane was recovered with a continuous type IV collagen line surrounding the chambers, as well as a clear reformation of the apical microvilli.

In summary, the reaggregation of choanocytes resulted in a complete recovery of the normal morphology of the choanoderm, demonstrating the reversibility of the dissociation process.



### Characterization of the transcriptome of *Oscarella lobularis*

An earlier attempt to obtain the genome sequence of the *O. lobularis* genome resulted only in a partial, highly fragmented genome due to a high number of repetitive sequences and a large number of contaminant reads from bacteria and archaea [11]. The estimated gene content of this partial resource was  $\sim 17,885$  protein-coding genes, which is low compared to other sponges [42, 72]. We therefore decided to sequence and de novo assemble the transcriptome of *O. lobularis*, attempting to obtain a near-complete transcriptome of this sponge species. To this end, two types of sequencing libraries were merged: one library was prepared from a reproductive adult with embryos and larvae previously sequenced by 454 sequencing [124], the second library was obtained from dissociating/reaggreating buds and from buds in NSW and sequenced by Illumina sequencing in the course of the present work. Sequencing of the two libraries resulted in 538,821 reads from 454 sequencing

and 480,883,138 Illumina reads, respectively (Fig. 5). Low quality reads with a Phred quality score  $< 20$  were trimmed and removed from both libraries. All reads were submitted to the Short Read Archive at NCBI (Bioproject PRJNA659410).

De novo assembly of the two different libraries was done separately (see Material and Methods section and Fig. 5), the concatenated dataset resulted in 162,768 contigs. After removing redundant contigs, as well as possible contaminations (Archaea, Bacteria, Viridiplantae, Fungi and Dinoflagellata) the final *O. lobularis* transcriptome contained 32,274 contigs. The average contig length was 958 bp, with the smallest contig having a length of 255 bp and the largest 47,628 bp. The mean GC content based on all contigs was 48.2% (Fig. 6A).

We functionally annotated the transcriptome using groups of orthologs, gene ontology terms (GO), as well as conserved domains (CDs). We found an orthologous group for 18,129 contigs (56%), could assign a GO term

(See figure on next page.)

**Fig. 6** Characteristics of the de novo transcriptome assembly of *O. lobularis*. **A** Contig length distribution and GC content of the *O. lobularis* transcriptome from mixed stages and types of sequencing. Statistics was done after de novo assembly, concatenation, decontamination, and length cleaning steps. The smallest contig is 255 bp long and the longest contig has 47 628 bp coming from a single long ORF encoding a Mucin ortholog. The average contig length is 958 bp. The mean percentage of GC content is 48.2%. **B** Annotated coding regions either with eggNOG [65] (orthologs) (18,129 sequences), GO Terms (11,526), and Conserved Domains (CDs) [93] (28,440 sequences) (Additional file 5: Table S2). For 1793 contigs no described CD or ortholog could be found, corresponding to 5.5% of the identified coding regions. **C** The ten most frequent Conserved Domains (CDs) in the *O. lobularis* transcriptome. Among the top ten CDs are kinases, nucleotide binding domains, domains that are involved in membrane transport or membrane binding as well as those inferring changes in cell shape. The protein kinase domain was most abundant and was present in 4057 coding regions. Occurrences were counted once per coding region. **D** Percentage of orthologs, *O. lobularis* has in common with other species. 29% orthologs are shared with human, 25.2% with *Trichoplax adhaerens*. Shared orthologs with other sponges vary from 24 to 57%. 56.6% of contigs from *O. lobularis* have a clear ortholog in its close relative, *O. pearsei* (Additional file 6: Table S3). **E** DOGMA analysis [29] to evaluate the completeness of different metazoan transcriptomes (Additional file 8: Table S5). For *O. lobularis*, we found 72.2% of expected CD arrangements. This is within the range of other sponge transcriptomes

to 11,526 contigs (35.7%) and predicted a CD for 28,440 contigs (88.1%) (Fig. 6B, Additional file 5: Table S2A–E). Protein kinases were the most abundant CD identified, occurring in 4057 contigs. Among the other top CDs were membrane-associated domains, domains involved in membrane transport or membrane binding as well as those inferring changes in cell shape (Fig. 6C). We further investigated with Orthofinder [37] the percentage of orthologs shared between *O. lobularis* transcripts from our assembly and several other metazoans (Fig. 6D, Additional file 6: Table S3). Shared orthologs with other sponges ranged between 24.4 and 56.6%, whereby the highest percentage of orthologs was found for the closest relative of *O. lobularis*, *Oscarella pearsei*. Between 20 and 33% of orthologs were shared with non-sponge animal species, with 29% of orthologs found in *H. sapiens*. We estimated the completeness of this transcriptome to be rather high, compared to other sponge transcriptomes (Additional file 7: Table S4), as our transcriptome assembly had a BUSCO [130] score of 82%. To further support the completeness of the *O. lobularis* transcriptome, we looked at the percentage of conserved domain arrangements using DOGMA [29] (Fig. 6E). We found that 72.2% of expected CD arrangements were present in the *O. lobularis* transcriptome. This number is similar to the other sponge transcriptomes tested, which had a maximum of 81.65% expected CD arrangements found in *Amphimedon queenslandica* (Demospongiae) and a minimum of 60% found in *Aphrocallistes vastus* (Hexactinellida) (Additional file 8: Table S5).

To conclude, the *O. lobularis* transcriptome presented here shows a high degree of completeness, with most of the contigs being assigned to an orthologous group, identifying a conserved domain or a gene ontology term. Furthermore, BUSCO and DOGMA results support the comprehensiveness of our resource. The transcriptome is available in the Bioproject PRJNA659410 on NCBI.

### Identification of differentially expressed genes during the cell-dissociation-reaggregation process

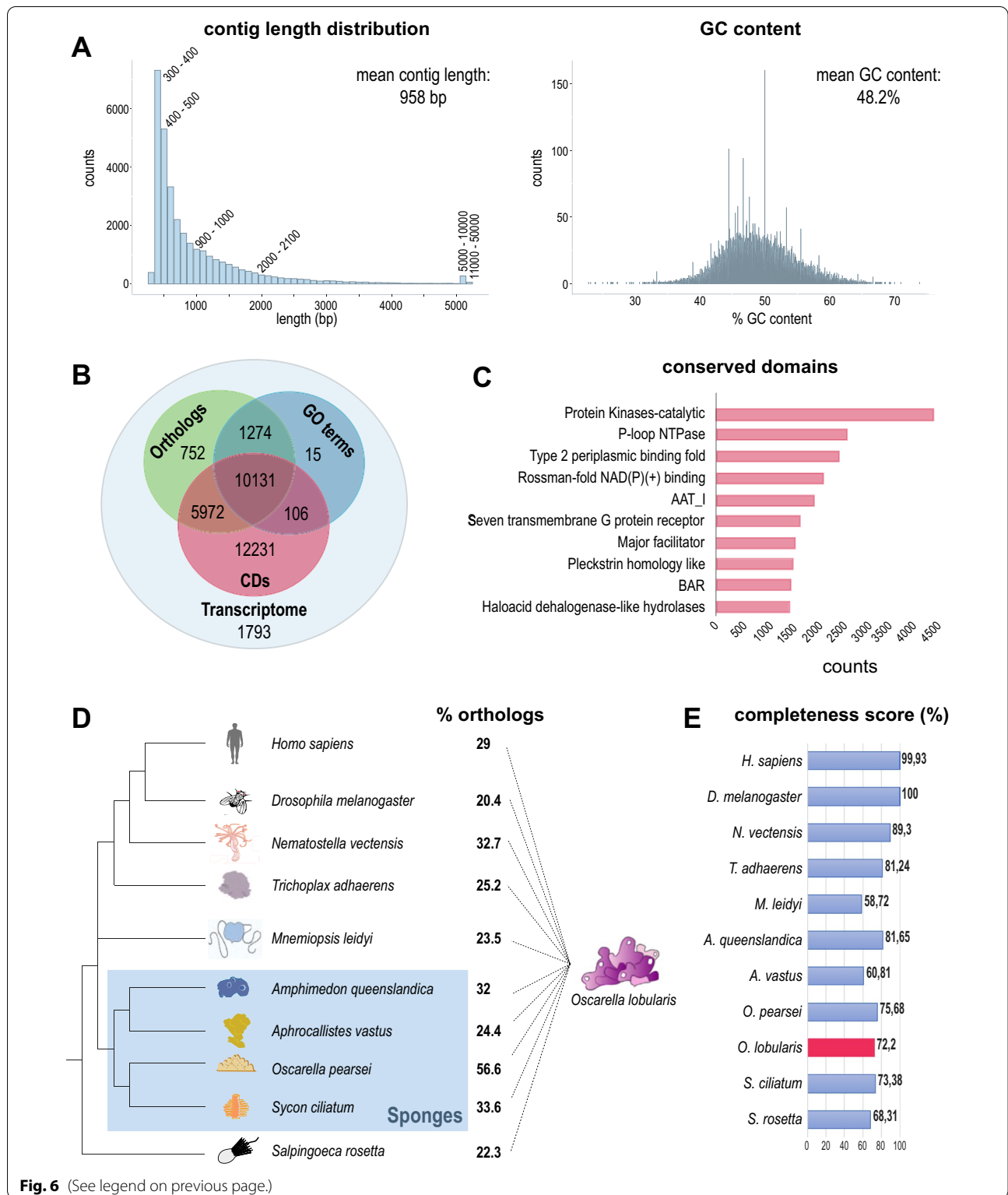
To identify putative molecular players involved in epithelium organization, we performed time-series RNA-sequencing at selected key time-points of the dissociation-reaggregation process: Two time-points during the dissociation, 1 h (1hD) and 4 h (4hD); and two time-points during the reaggregation/re-epithelialization process, 1 h (1hR) and 3 h (3hR) (Fig. 2). After trimming of raw reads, we mapped the reads to the de novo assembled *O. lobularis* transcriptome using Kallisto [18] and Tximport [134]. Principal component analysis (PCA) of normalized read counts showed that the different time-point replicates cluster together, except for the first time-point of reaggregation (1hR), which was dispersed between 4hD and 3hR clusters, suggesting a variability in the process timing between pooled individuals (Fig. 7A).

### Number and evolutionary emergence of genes involved in dissociation and reaggregation

We looked for differentially regulated genes between time-points by pair-wise comparisons using DESeq2 [90]. We found 5903 genes significantly differentially expressed (adjusted p-value < 0.05 and a log<sub>2</sub>|FC| of at least 1.5) between each time-point of dissociation and the NSW control and between each time-point of reaggregation and 4hD. The number of exclusively differentially expressed genes during dissociation (2510 genes) was comparable to the number of exclusively differentially expressed genes during reaggregation (2713 genes). 11.5% of differentially expressed genes (680 genes) were shared in both processes, dissociation and reaggregation (Fig. 7B, see Additional file 9: Table S6A–C for the lists of differentially expressed genes of the processes under study).

To evaluate the proportion of sponge-specific, pre-metazoan or metazoan genes differentially expressed during the dissociation-reaggregation process, we





performed a phylostratigraphy analysis according to Sogabe et al. [133]. A phylostratigraphy analysis is based on sequence similarity with genes present in other

organisms with a defined phylogenetic distance [30]. *O. lobularis* genes were classified in three categories of evolutionary emergence: (1) before or (2) after the

(See figure on next page.)

**Fig. 7** Gene expression changes in *O. lobularis* buds during dissociation and reagggregation. **A** Principal component analysis (PCA) plot of the 5 different conditions sequenced during dissociation and reagggregation. Dissociation time-points are circled in blue, those of reagggregation in pink. Replicates cluster well except for the samples of the first hour of reagggregation (1hR) which show larger dispersion, and which are located between time-points of four hours of dissociation (4hD) and three hours of reagggregation (3hR), indicating different reagggregation states of the sequenced buds. **B** A total of 5903 differentially expressed genes (DEGs) were found during dissociation and reagggregation with a log<sub>2</sub> fold change (log<sub>2</sub>[FC]) of at least 1.5 (Additional file 9: Table S6). Pairwise comparisons were done as follows: for dissociation (1hD/CT, 4hD/CT); for reagggregation (1hR/4hD, 3hR/4hD); genes in common for both processes are shown as overlapping (680 genes). Gene ontology enrichment analysis of genes differentially expressed during dissociation shows that genes relevant for the process under study are enriched, including those involved in cell adhesion and its regulation, substrate dependent cell migration and actin bundle assembly were enriched (Additional file 11: Table S8). Enriched GO terms for shared DEGs of both conditions included cell migration, cell adhesion, regulation of collagen metabolic processes as well as integrin-mediated signaling pathway. During reagggregation, relevant GO terms associated with cell adhesion, regulation of integrin and actin filament-based process were enriched. **C** Phylostratigraphy analyses on the whole transcriptome and during dissociation and reagggregation provide the evolutionary age of differentially regulated genes during the dissociation and reagggregation process, as well as in the entire transcriptome of mixed stages. On the entire transcriptome 16.6% (5389) of the genes are *O. lobularis* specific (see Additional file 10: Table S7). During dissociation and reagggregation, there is a higher percentage of sponge-specific DEGs as compared to the entire transcriptome. The percentage of metazoan DEGs is higher during dissociation than during reagggregation. Analysis was done after Sogabe and collaborators in [133]. Species used are listed in Additional file 10: Table S7A. **D** Heatmap showing the hundred most differentially regulated genes during dissociation and reagggregation (1hD/CT, 4hD/CT, 1hR/4hD, 3hR/4hD). For genes upregulated during dissociation or reagggregation, we performed enrichment analysis for GO terms (with TopGO [4]) (Additional file 11: Table S8B). Among the top enriched GO terms, we found regulation of extracellular matrix disassembly for dissociation and establishment and maintenance of an epithelium for the reagggregation

divergence between choanoflagellates and metazoans (pre-metazoan and metazoan genes, respectively); and (3) after the divergence between the phylum Porifera and the other metazoan phyla (sponge-specific genes) (Additional file 10: Table S7A, B). We first looked at the overall distribution of the phylogenetic category of conserved genes in the entire transcriptome and found that most of the genes (48%) were sponge-specific and a slightly higher percentage of the remaining genes predated the metazoan emergence (30%), leaving 22% metazoan genes (Fig. 7C; Additional file 10: Table S7C). The proportion of genes in each phylogenetic category was slightly shifted in both dissociation and reagggregation: during both processes, we found slightly more sponge-specific genes than present in the transcriptome (52.3% in dissociation and 59.6% in reagggregation, respectively). During both processes the percentage of Pre-metazoan genes decreased and the percentage of metazoan genes slightly increased. We found a higher percentage of metazoan genes (25% and 22% in dissociation and reagggregation, respectively) than pre-metazoan genes (22.7% and 18.4% in dissociation and reagggregation, respectively) (Fig. 7C; Additional file 10: Table S7C).

#### Identification of the molecular toolkit regulated during dissociation and reagggregation of a sponge

Among the 5903 differentially expressed genes, we next looked at enriched gene ontology (GO) terms associated with genes involved only in dissociation (2510 genes) or only in reagggregation (2713 genes), as well as those genes shared between both processes (680 genes) (Fig. 7B; the complete list of enriched GO terms is available in Additional file 11: Table S8A). In order to investigate, whether

epithelial organization shares molecular components with bilaterians we focused on genes related to cell–cell contacts, cell adhesion and migration.

Genes involved in both dissociation and reagggregation were enriched for GO terms related to cell-migration, control of cell projection and cell adhesion (including integrin-dependent mechanisms) and control of collagen metabolism (Fig. 7B). For the significant transcripts associated with these GO biological processes, we identified the encoded genes by identifying orthologous groups with eggNOG [65], as well as by performing BlastP searches [5] (Additional file 12: Table S9A). We observed that most of these genes encoded proteins involved in cell surface receptor tyrosine kinase or phosphatase activities.

During dissociation, mainly genes involved in regulation of cell adhesion, cell migration, actin bundle assembly were enriched, and even more genes involved in extracellular matrix organization. This latter category included in particular extracellular matrix-structuring proteins such as Collagens, Microfibril-Associated Glycoprotein (MAGP), Laminin or Papilin and Integrin receptors (Additional file 12: Table S9A).

Genes involved in reagggregation were enriched for GO terms related to cell adhesion, actin filament-based processes and in the regulation of integrin complex activation (Fig. 7B). The latter GO category included genes such as Protocadherin Fat 4, Cadherin 23-like or Afadin (Additional file 12: Table S9A).

We finally looked at the 100 most differentially expressed genes over all four key steps of dissociation-reagggregation (Fig. 7D). We found different sets of genes specific either for dissociation or for reagggregation.



**Fig. 7** (See legend on previous page.)

(See figure on next page.)

**Fig. 8** Gene expression dynamics during dissociation and reagggregation. **A** Fuzzy clustering [51] of temporal expression profiles during dissociation and reagggregation. We chose 15 clusters to represent our data, determined by hierarchical clustering of normalized read counts of the five conditions using hclust in R (Additional file 20: Figure S7). We found subgroups of profiles with a decrease in expression until 4hR (clusters 12,15), early peaks and a flat curve during reagggregation (clusters 3,8,5,6,10), peaks at 1hR (clusters 4,14), decreasing expression during dissociation and increasing expression in reagggregation (clusters 1,11,13), strong ascending trends (clusters 10,5) and a peak at 4hD (clusters 7, 2, 9). Cluster cores have a membership value > 0.7 (pink) (Additional file 13: Table S10). Other membership values are shown as follows: membership value between 0.5 and 0.7 (blue), membership value < 0.5 (green). **B** Selected GO terms corresponding to Biological Processes found to be enriched for each core cluster as determined by topGO [4] (Additional file 11: Table S8C). Colored boxes indicate significant enrichment in a given term (adj p-value < 0.05). The dendrogram at the top was calculated by hierarchical clustering using hclust with Euclidian distance of the normalized read counts of genes qualifying as cluster core (membership value > 0.7). **C** Expression profile of genes involved in epithelia organization and cell adhesion in bilaterians (see also Fig. 1A). Expression of genes of the three polarity complexes decrease during dissociation and increase upon reagggregation. The expression of the gene coding for the basement membrane collagen was rather stable during the process, while the expression of integrin  $\beta 2$  increased at 4hD and became stable at 1hR. Genes involved in adherens junctions globally show a decrease in expression during dissociation and an increase around 1hR. Expression of the cytoskeleton component actin increased upon dissociation and decreased when reagggregation started. Finally, four of the Aggregation Factors (AFs) candidates (AF 1, 3, 4, 5) were downregulated during dissociation and upregulated during reagggregation (in particular AF3). In contrast, AF2 was downregulated during the entire process. NSW stands for natural condition which is the control condition, 1hD and 4hD correspond to 1 h and 4 h of dissociation (in CMFSW) respectively, 1hR and 3hR stand for 1 h and 3 h of reagggregation (in NSW) respectively. See Additional file 15: Table S11 for Standardized TPM and log2 Fold Change values

We used those genes for GO-term enrichment analysis (Additional file 11: Table S8B). During the dissociation process, in agreement with our previous results, the most differentially expressed genes were mainly associated with the GO category “Extracellular matrix disassembly”. For the reagggregation process, we found GO-terms related to “Regulation of integrin activation”. In addition, we noticed that transcripts belonging to two orthologous groups of C-type lectins (Additional file 12: Table S9B) were strongly increased during the reagggregation process.

In summary, we found biological processes related to cellular adhesion, cell migration, integrin signaling and actin assembly as well as processes important for organizing the extracellular matrix enriched in the differentially expressed genes during the *O. lobularis* dissociation- reagggregation processes.

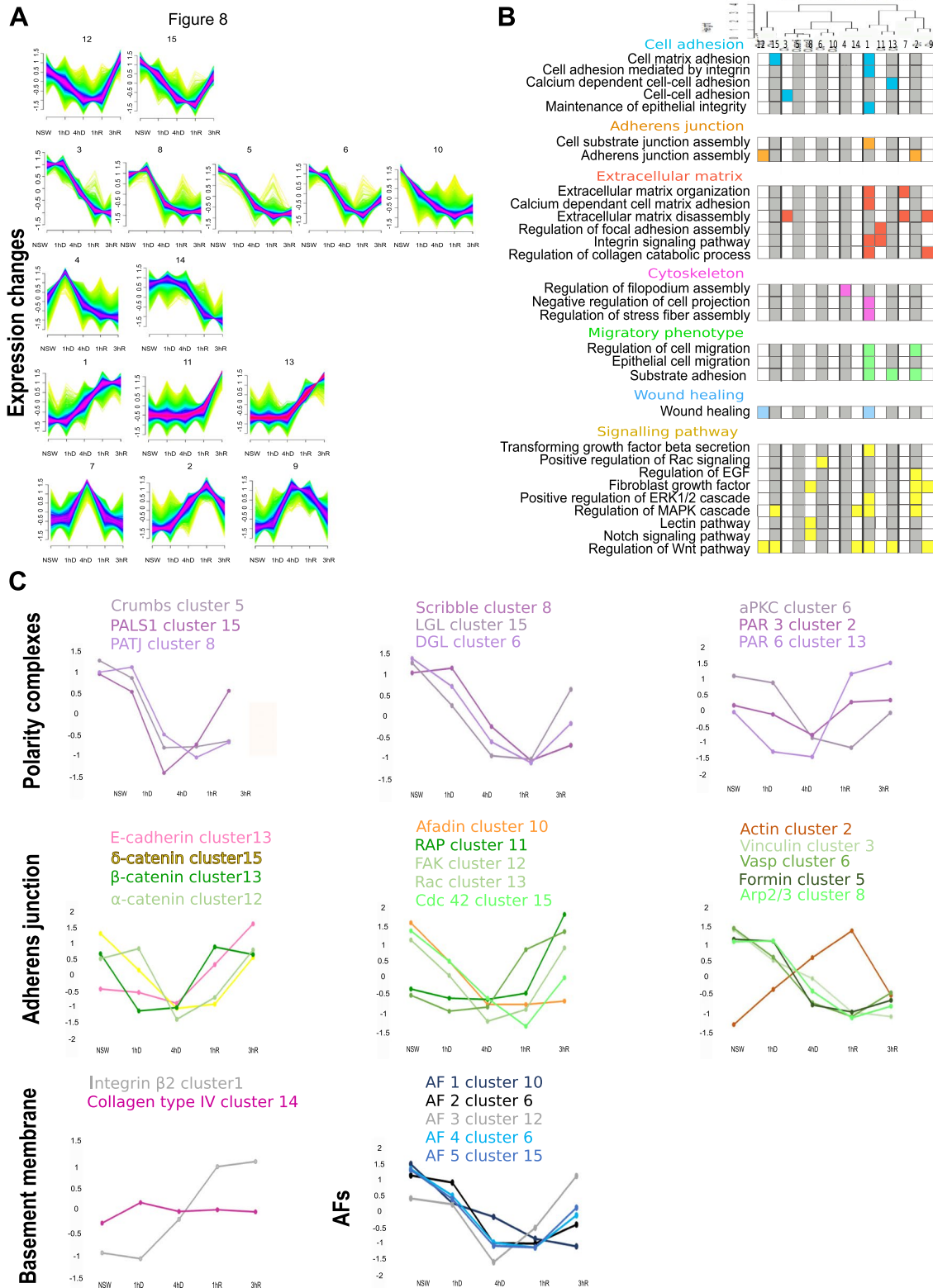
#### Gene expression dynamics during dissociation and reagggregation of *Oscarella lobularis* buds

To identify genes with a similar temporal expression profile, we performed Mfuzz clustering [51] on standardized transcript per million read counts (TPM) of all expressed genes. In this analysis, the expression profile over time, rather than the degree of differential expression is decisive, allowing to monitor the overall expression dynamics of genes during dissociation-reagggregation and to identify co-regulated gene groups. Mfuzz cluster profiles are shown in Fig. 8A (Additional file 13: Table S10 for the list of the core cluster genes). We could discriminate different types of expression profiles in our dataset: those with a decrease in expression during dissociation and upregulation of genes during reagggregation (clusters 12, 15); those with decreasing expression during the

time-course (clusters 3, 8, 5, 6, 10); those with a strong increase of expression during reagggregation (clusters 1, 11, 13); finally, those with a peak in expression either during late dissociation or early reagggregation (clusters 7, 2, 9, 4, 14). Cluster cores, defined by a membership value  $\alpha > 0.7$ , contained between 43 and 654 genes.

We performed GO enrichment of core Mfuzz clusters using TopGO [4] (Fig. 8B; complete list of GO terms are available in Additional file 11: Table S8C). During both, the dissociation and reagggregation process, a large set of terms related to locomotion, cell migration, wound healing, adherens junction regulation, EGF-receptor signaling as well as MAPK signaling were enriched in different clusters: during dissociation we observed a decrease of transcripts related to lectins and the Notch pathway (Cluster 8), FGF signaling (Cluster 8) and an increase of transcripts related to filopodium assembly (Cluster 4), regulation of focal adhesion assembly (Cluster 11), TGF $\beta$  signaling (Cluster 1), as well as EGF and FGF signaling (Cluster 2,9). During reagggregation, genes belonging to enriched GO terms related to cell adhesion (involving Calcium-dependent mechanisms) (Cluster 13) were increased, and genes related to the regulation of extracellular matrix disassembly (Cluster 7, 9) and filopodium assembly (Cluster 4) were decreased.

In addition, we searched in clusters with common expression profiles for transcripts with conserved domains related to cadherin, integrin and immunoglobulins as major domains involved in the Cell Adhesion Molecules (CAM) [62] (Additional file 14: Figure S4). According to the phylostratigraphy analysis, more than 30% of these transcripts were Sponge-specific (Additional file 10: Table S7D). Furthermore, these specific



**Fig. 8** (See legend on previous page.)

transcripts showed interesting profiles with decreased expression at 4hD and an increase during reaggregation.

We looked more specifically at the expression profiles of genes involved in epithelial organization in bilaterians, including genes encoding proteins involved in cell polarity, basement membrane, as well as adherens junctions (Fig. 8C; standardized TPM values are provided in Additional file 15: Table S11A). The nine genes that are part of the three polarity complexes Crumbs, Scribble and Par had similar expression patterns: they decreased during dissociation and increased again during reaggregation (Fig. 8C). Par complex genes showed slightly different dynamics and recovered earlier in reaggregation to initial levels. Expression profiles of genes involved in forming adherens junctions followed a similar trend, with downregulation at early stages followed by increasing expression levels during the reaggregation process, though the variability of observed profiles was higher for this gene group. Notably, Actin showed a reverse trend to all adherens junction genes, with increasing expression levels during dissociation and a drop in expression levels between 1 and 3 h of reaggregation (Fig. 8C). The two genes coding for the basement membrane components Integrin  $\beta$ 2 and type IV collagen showed differential expression dynamics: type IV collagen was mostly stable during the time-course with only a slight increase in expression at 1hD, suggesting that loss of protein immunostaining we observed (Fig. 3) was likely controlled on protein level. In contrast, the expression of Integrin  $\beta$ 2 was induced during late dissociation and became stable at 1hR.

Finally, we looked at Aggregation Factors (AFs) that are thought to be key players of cell adhesion processes in sponges [63]. The five putative Aggregation Factors (belonging to AF-Group3 according to the analyses of [59]), were recovered in clusters 10, 6, 12, 15 (Fig. 8C, TPM standardized values are shown in Additional file 15: Table S11B; AF domain composition is shown in Additional file 16: Figure S5 and Additional file 17: Table S12). Nearly all AFs showed similar expression dynamics, with a decrease during dissociation and an increase in expression during reaggregation, with one AF being the exception that did not recover its transcript levels by 3hR (AF 1 in cluster 10).

In conclusion, we found an enrichment of terms related to cellular adhesion, locomotion, adherens junction, as well as cytoskeleton organization in several profile clusters. Moreover, several proteins known to be required for epithelial structures or epithelial organization in bilaterians, as well as the Sponge-specific aggregation factors were differentially regulated during the dissociation and reaggregation process.

## Discussion

### A new transcriptomic resource exploring the complement of the ancestral epithelium.

In this study, we present an improved version of the transcriptome of *Oscarella lobularis* [124]. We used a mix of stages, which included an adult in reproduction which contained pre-larvae and embryos, as well as buds under different environmental conditions (calcium-magnesium deprivation or not) and combined different methods of sequencing (454+Illumina). This new transcriptomic resource is therefore a valuable addition to the few (2 sequenced species out of 126 described homoscleromorph species) transcriptomic and genomic resources currently available for the homoscleromorph sponge class [42, 110, 119], reviewed in [117]. The de novo assembled transcriptome of *O. lobularis* revealed a high degree of completeness (82% BUSCO score) and substantial similarity to other sponges and metazoans (56.6% of orthologs with the closest sponge species). It also contained a high percentage (72%) of conserved domain arrangements as determined by DOGMA. This is well within the range for other non-standard model organisms and other sponges. *A. queenslandica* has a conserved domain arrangement of 81.65%, *A. vastus* of 60.81% and this score is 75.65% for the closely related *O. pearsei*.

Nonetheless, it is probable that the *O. lobularis* transcriptome presented here is incomplete, best reflected in the total number of contigs and the percentage of conserved genes with other sponges and metazoan species. For example, it is currently estimated based on near-complete sponge genomes that sponges have a large gene repertoire and gene numbers range between ~30,000 and ~40,000 [42, 72]. The closest neighbor of *O. lobularis*, *O. pearsei* has about 29,000 predicted genes [42]. Based on the analysis of our transcriptome, we only find 56% shared genes between these two closely related species. The *O. lobularis* transcriptome shows a relatively high number of species-specific genes (>40%) compared to other Porifera [119]. While we tried to reduce redundancy in our assembly by clustering similar contigs using CD-HIT, remaining redundant contigs, or only partially assembled transcripts could obscure the annotation process and fail to assign orthologs between species. It is therefore desirable to aim for a chromosome-level assembly of this sponge. Due to the relatively high number of repetitive elements this genome seems to harbor, this could be achieved by combining long-read sequencing with partial genome sequence data already available, partial genome sequence data [11]. This resource presented here should then help in predicting genes with high accuracy. Furthermore, for sponge-specific transcripts, it is desirable to study those with interesting

expression patterns derived from our experimental setup functionally.

Beside the de novo transcriptome provided here, this study is the first transcriptomic analysis designed to identify genes involved in cell dissociation–reaggregation in a sponge, despite the long-standing interest for this process in sponges [137]. This particular experimental context is especially suited to explore molecular mechanisms involved in epithelial morphogenetic processes in sponge, to determine, whether these processes are conserved between sponges and bilaterians and thus, to study the evolutionary origin of epithelium formation.

#### Calcium as key regulator for epithelium morphogenesis

Cell dissociation in the homoscleromorph sponge *Oscarella lobularis* was induced by incubating biological samples in Calcium and Magnesium Free Sea Water. While this type of experiment was already performed in numerous calcareous sponges and demosponges [24, 25, 52, 66, 68, 78, 79, 81, 91, 132, 137], this study is the first one (i) performed on a homoscleromorph, (ii) with such a level of precision regarding the cellular description, and (iii) using transcriptomic approaches to decipher molecular mechanisms involved during these processes. This taxonomic choice is key since homoscleromorph sponges are the only sponges in which the basement membrane has been established. This work is a preliminary study on the effect of depletion of calcium and magnesium ions on *Oscarella lobularis*. This depletion can affect more general processes, like metabolism, and induce a change in the expression of genes involved in these processes. Here, we chose to focus on potential changes of expression due to epithelial reorganization. Consistent with the role of calcium deprivation described in other metazoans, cell–cell and cell–matrix interactions and epithelial integrity are affected in *O. lobularis* in the same way as in other animals resulting in the loss of epithelial characteristics.

Under calcium-magnesium deprived conditions, the choanocytes, forming the choanoderm epithelial layer (1) change their cell shape (from conical to round shape) and lose their basal–apical polarity; (2) reorganize their actin cytoskeleton (loss of apical ring of actin microvilli and increase of actin-based protrusions at the basal pole); (3) lose their cell–cell and cell–matrix contacts; and (4) start to migrate (towards the choanocyte chamber internal cavity or in the mesohyl).

Both the loss of tight cell–cell and cell–matrix contacts and the reorganization of the actin cytoskeleton are highly reminiscent of the phenotype resulting from calcium and magnesium deprivation in mammalian and other bilaterian epithelial cells. It is well known that, in bilaterians, cell contacts are calcium-dependent and removing bivalent cations from the medium induces

junction disassembly and detachment from the basal lamina [20, 21, 32, 55, 75, 108, 114].

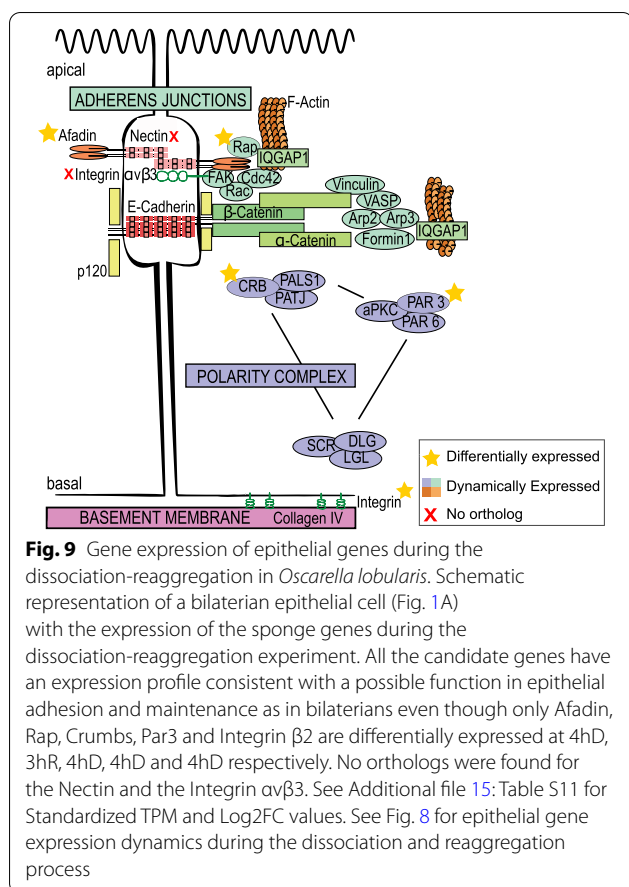
In addition, this process is reversible, as is also reported for bilaterians: incubation in natural sea water induces cell-reaggregation and re-epithelialization with a complete restoration of cell polarity, cell–cell and cell–matrix contacts. Such a plasticity of epithelial tissues is essential since epithelial cell movements are required during major morphogenetic processes such as, for example, wound healing and regeneration, tissue renewal, or gastrulation [31, 73, 113, 115]. In bilaterians, calcium signaling is commonly and naturally involved in the regulation of epithelial tissue properties (motility, differentiation, cilia beating) during developmental processes [20]. Interestingly even if calcium-signaling is an ancient feature of Eukaryotes, the emergence of a complex epithelial level of organization coincides with the multiplication and diversification of calcium-signaling proteins involved in cell–cell contacts in animals [92]. The calcium depletion implemented in this study provides an invaluable opportunity to study the evolution of calcium-dependent molecular mechanisms involved in epithelial morphogenesis during animal evolution.

#### An ancestral metazoan molecular toolkit involved in epithelium organization and morphogenesis

In bilaterians, the regulation of epithelial features involves an intimate interplay between adhesion proteins, polarity protein complexes and regulators of the actin cytoskeleton. The current knowledge on sponge epithelia suggests that they present similarities with those of bilaterians: chemical and mechanical properties [36, 86], histological likeness [38, 83–85], supposed involvement of Wnt signaling in their regulation [1–3, 77, 124, 125, 139, 140] and shared epithelial genes and proteins [11, 44, 58, 97, 98, 110, 119, 128]. Nevertheless, the homology and functional conservation of proteins involved in the composition or in the regulation of sponge epithelia remains an open question. The molecular data we present that derives from transcriptomic analysis captured during an epithelial morphogenetic process brings important new insights on this key evolutionary question.

#### Crumbs as an ancestral regulator of epithelial features?

The protein Crumbs is a metazoan innovation [7, 10, 11, 80, 116]. In bilaterians, Crumbs is not only involved in the establishment and maintenance of cellular apico-basal polarity and the positioning and stability of adherens junctions, but is an essential player in the dynamics of actin remodeling by interacting with many actin binding partners [10]. These partners can interact with the FERM or the PDZ-binding domains of Crumbs directly or indirectly. For example, Crumbs proteins can regulate



actin dynamics by recruiting ARP2/3 or by regulating Cdc42 [10]. In *O. lobularis*, orthologs shared with the bilaterian proteins involved in the Crumbs complex (namely Crumbs, Patj, Pals1) have been reported, and key functional motifs required for their functional interactions have been clearly identified [11]. Here we show that the genes encoding Crumbs, Pals1 and Patj are downregulated during dissociation and upregulated during the reaggregation (only Crumbs is differentially expressed at 4hD) (Fig. 9, Additional file 15: Table S11A for Log2 Fold Change values). Both, our previous in silico analysis [11] and their very similar expression dynamics during dissociation and reaggregation, are consistent with the possible conservation of their joint functional involvement in epithelium integrity, as has been shown in other animals [7]. Because these genes, as well as an epithelium in general, are specific to metazoans, it suggests that Crumbs and its partners have an ancestral role in epithelial organization in Metazoa.

Similarly, the genes encoding the proteins involved in the two other bilaterian polarity complexes, the Scribble and Par complexes, are also downregulated while epithelium characteristics are lost; and then upregulated

during the re-acquisition of epithelial features. The facts that (i) all these genes (except Dlg) are metazoan innovations [116], (ii) their protein domain structures are highly conserved [11] and (iii) their expression dynamics during the epithelial morphogenetic process studied here, are in favor of their ancestral joint involvement in metazoan epithelia. Furthermore, the presence of Dlg in Choanoflagellata, Filasterea and Ichthyosporea [11, 44, 53, 80, 95, 107, 118], suggested a Holozoan origin of this gene. This therefore suggests a co-optation of this gene for epithelial functions in the Last Metazoan Common Ancestor (LMCA). Further expression, localization and functional studies will help decipher the exact role of these genes in sponges, and thereby help resolve their respective ancestral roles.

### Cadhesome, integrin-based “adhesomes” and cytoskeleton

In bilaterians, the crucial cross talk between cell–cell and cell–matrix adhesion sites for epithelium stability and plasticity is well described. Cadherin- and integrin-dependent adhesion and signaling functions intersect and interact. Changes in adhesion/force transmission at one site may affect membrane trafficking, cytoskeletal association, avidity or binding affinity [9, 28, 136]. Integrins and cadherins are both transmembrane adhesion receptors. They have many signaling effector molecules in common and are linked to common scaffolding and cytoskeletal elements. Both adhesion molecules are essential for the formation of cell–cell and cell–matrix contacts through adherens junctions (AJs) and focal adhesions (FAs), respectively. Interestingly, the expression of most of the genes involved in both cadherin and integrin adhesion complexes are downregulated during dissociation and upregulated during reaggregation suggesting an involvement in epithelial organization in sponges as they do in bilaterian species. In addition, our results show that many genes encoding extracellular matrix proteins displayed differential expression, such as Glycoproteins, Laminins or Disintegrin and Metalloproteases (ADAMs), which are integrin ligands [19, 56]. These results suggest that, like in bilaterians, the Extracellular Matrix (ECM) of *O. lobularis* could play an important role as substrate for migration, assembly, and regulation of the epithelial cells [122, 129].

Together with previous genomic surveys [11, 109] and recent biochemical studies [97, 98, 128], our results support the hypothesis that cadherin and integrin-based mechanisms are ancient features of metazoans, as seems their ancestral roles in cell and tissue cohesion in the LMCA.

Interestingly, the E-cadherin gene is barely changing expression levels during cell dissociation while it shows an increase during reaggregation and expression of the



afadin gene shows opposite regulation. This unexpected finding according to previous results [8, 71] questions the relative contribution of Cadherin and Afadin in the establishment and the maintenance of adherent junctions in *O. lobularis*. It cannot be excluded that the expression of these genes is regulated at protein level. Immunolocalization of their encoded proteins will shed light on this issue.

Beyond what is known about other metazoans, different genes may be involved in adhesion in sponges. During dissociation-reaggregation, some genes encoding other proteins involved in adhesion show a differential expression: this is the case for genes encoding Fat 4 proto-cadherin, non-classical Cadherins 23 and C-type lectins. For example, the increase of C-type lectin domain containing genes during reaggregation is consistent with its reported functions in bilaterians: (i) as suppressor of Epithelial-Mesenchymal Transition (EMT) [121], (ii) as being involved in adhesion in budding tunicates [94]; (iii) as necessary for cell-aggregation in choanoflagellates (formation of rosettes) [82]; and (iv) as being involved in cell-reaggregation in the sponge *Aphrocallistes vastus* (Hexactinellida) [60, 106]. In vertebrates, C-type lectin domain proteins are known to facilitate the Ca<sup>2+</sup>-dependent cell–matrix and cell–cell interactions [121]. In the same way, Fat 4 and Cadherin 23 proteins were shown to regulate cell–cell adhesion and to be involved in dynamic changes leading to epithelial fluid-like attributes in bilaterians [6, 76].

#### **Sponge-specific genes involved in epithelium organization open new perspectives on epithelium evolution.**

Among sponge-specific genes involved in epithelial processes, we chose to focus on Aggregation Factors (AFs), because of the historical significance of these extracellular proteoglycans concerning sponge dissociation-reaggregation experiments [64, 66, 67, 81, 100–105]. AFs were shown to be involved not only in allorecognition but also in physical cell–cell bridges-like connections. They play key species-specific cell adhesion roles. A recent study questioned the presence of AFs in the homoscleromorph class because of the too low sequence similarity with demosponge AF sequences and because of the lack of evidence of reaggregation properties in this sponge class [59]. In this study we identified 5 genes in *O. lobularis* considered as possible AFs candidates (AFs of Group 3) by Grice and co-workers [59]. Interestingly, the transcripts of the 5 AF candidates decreased during the dissociation process and 4 of them were upregulated during reaggregation. This expression is consistent to observations described for demosponge AFs [59, 78]. However, biochemical studies are needed to confirm that

these genes encode *bona fide* aggregation factors in this homoscleromorph sponge.

Interestingly, these aggregation factors have been proposed to be able to interact with the integrin RGD motif and to activate integrin signaling [45, 63, 98, 138]. Therefore, models that will be proposed for the establishment of cell–cell contacts and adhesion in sponges will have to consider and integrate the involvement of AFs and their possible interactions with other adhesion receptors.

In addition, the search for genes harboring cadherin, integrin and immunoglobulin conserved domains (usually found in the Cell Adhesion Molecules (CAM) [62]) in the different clusters of expression has highlighted the presence of these domains in numerous Sponge-specific transcripts (more than 30%). While some of those might be resulting from partial redundancy in our de novo assembled transcriptome, or biases induced by the design of the experiment, some could be potential new candidates to play a role in the epithelial cell adhesion because of their increase in expression during reaggregation. However, to test the validity of these candidates, additional in-depth analyses of their domain composition and functional analyses will be needed.

#### **Conclusions**

The transcriptomic data collected during dynamic epithelial processes provide new clues for understanding the molecular elements of epithelization in *Oscarella lobularis*, and, more generally, for investigating the evolution of the mechanisms involved in epithelia formation in bilaterians. This study supports the hypothesis that epithelium organization in *O. lobularis* involves both genes and processes homologous to bilaterians, as well as sponge-specific ones. Biochemical and functional experiments will be needed to assess to what extent the complex signaling cascades and protein interactions described in bilaterian epithelia organization predate the cnidarian-bilaterian emergence.

#### **Methods**

##### **Animal husbandry**

Adult specimens of *Oscarella lobularis* (Schmidt 1862) were collected by SCUBA diving in the bay of Marseille (France) and kept in a thermostatic chamber at 17 °C. Budding was induced by cutting adults and each fragment was placed in a well containing 8 ml Natural Sea Water (NSW) as described in [14, 120]. Free buds are transferred into Petri dishes containing NSW (renewed once a week) and maintained at 17 °C (Fig. 2).

##### **Cell staining**

To monitor choanocytes, fluorescein labeled lectin PhaE (*Phaseolus vulgaris* Erythroagglutinin) (Vector lab) was

used to specifically stain this cell type (1: 500 [120]). Buds were incubated overnight (ON) at 17 °C then rinsed three times in NSW.

#### Cell dissociation and reaggregation experiments

To induce cell-dissociation, stage 3 buds (8–29 days, [120]) from *O. lobularis* (n = 3) were placed in 24-wells culture plates containing 2 mL of Calcium Magnesium Free Sea Water (CMFSW) at 17 °C. Choanocyte chambers integrity was monitored (see next section) at 20 min, 1 h, 2, 3, 4 h. According to our observations (see Result section) the pivotal time-points chosen for further analyses were 1 h of dissociation (time-point hereafter referred to as 1hD) and 4 h (4hD).

After 4 h of dissociation in CMFSW, cell-reaggregation was triggered by putting back buds into 2 mL of natural sea water (NSW). The reaggregation process was monitored after 30 min, 1 h, 2 h, 3 h, 5 h, 6 h and 24 h. According to our observations the pivotal time-points chosen for further analyses were 1 h (1hR), 3 h (3hR) and 24 h of reaggregation (24hR). For each time-point (except 24hR), one part of the treated and NSW control samples was collected and frozen (−80 °C) for RNA-sequencing and one part preserved in fixative solutions for imaging (see next section and Fig. 2).

#### Microscopy observations

At each stage of the dissociation (1hD, 4hD) and reaggregation (1hR, 3hR, 24hR) processes treated and control buds were observed by fluorescent light and electronic microscopy (except for reaggregation process), focusing on the choanocyte chamber.

For Scanning Electron microscopy (SEM), samples were prepared using the NCMIR protocol for SBF-SEM [27]. Imaging was carried out on a FEI Teneo VS running in low vacuum (30 Pa), at 2 kV and using a backscattered electron detector. Acquisition pixel size was 20 × 20 × 60 nm.

For confocal observations, samples were fixed overnight in 3% paraformaldehyde (PFA) in PBS at 4 °C, rinsed twice and staining with DAPI (1:500) (Thermo Fisher Scientific) and 1:1000 (Alexa fluor 647 coupled-phalloidin (Santa Cruz Biotechnology) and mounted in Prolong Diamond antifading mounting medium (Thermo Fisher Scientific) as described in [120] and [14]. Confocal acquisition was done with a Zeiss LSM 880 with a 63x Oil objective and images were processed using FIJI [127]. All experiments were repeated at least 3 times.

#### Type IV collagen antibody production and validation

Antibodies against the type IV collagen of *O. lobularis* were obtained by immunizing rabbits using the speedy protocol from Eurogentec (Seraing, Belgium) with a

synthetic peptide (QTISDPGEEDPPVSKC) coupled to the KLH (keyhole limpet hemacyanin) carrier. This peptide is present in the C-terminal NCI domain of *O. lobularis* type IV collagen and was used to purify antibodies from immunized rabbit sera.

To validate this antibody, we performed competition assays. During these assays, the antibodies against type IV collagen were incubated ON at 4 °C with either the immunizing peptide, or an unrelated peptide (CSTVS-VAQTGLKGGGI, from the Crumbs protein, negative control) or only PBS buffer (positive control). The molar ratio between antibodies and peptides was 1 mol for 25 mol, respectively, and antibodies were used at 1:200 (2.6 µg/mL). After centrifugation (21000 g, 4 h, 4 °C) to get rid of potential antibody-peptide complexes, supernatants were used to perform the immunostaining protocol (see next section). Incubation of antibodies with the immunizing peptide yielded to the loss of immunostaining unlike incubation with irrelevant peptide antibodies α- type IV Collagen (Additional file 18: Figure S6).

#### Acetylated tubulin and type IV collagen immunostaining

For immunofluorescence, buds are fixed in 3% PFA in PBS ON at 4 °C. To monitor the epithelial features of the choanoderm, the type IV collagen was immunolocalized to observe cell–matrix adhesion, acetylated tubulin (Sigma T6793) was immunolocalized to observe alteration of cell polarity and Alexa fluor-647 coupled-phalloidin (Santa Cruz Biotechnology) was used to follow actin-rich adhesive cell–cell junctions and cytoskeleton remodeling. Immunofluorescence was performed as described in Rocher et al. [120] and additional technical details are available in [14].

#### Statistical analyses of confocal acquisition

Perturbations of the choanoderm layer in CMFSW-treated buds compared to the controls were estimated according to 2 criteria (1) after 1 h of dissociation (1hD) by the presence or absence of microvilli on choanocytes (2) after 4 h of dissociation (4hD) and 24 h of reaggregation by the integrity or not of the choanocyte chambers (cell–cell and cell–matrix adhesion, presence/absence of central cavity). The experiments were at least repeated 3 times for each time-points with at least 3 buds. The integrity of the choanoderm, according to the previously defined criteria, was observed by counting the choanocyte chambers on two distinct part of the buds (1hD: 3 × 4 buds, 178 choanocyte chambers counted; 4hD: 6 × 2/3 buds; 216 choanocyte chambers counted; 24hR: 3 × 3 buds; 94 choanocyte chambers counted) (Additional

file 3: Table S1). A Wilcoxon test was performed using R [123].

### RNA-sequencing

Biological triplicates of NSW control buds, 1hD, 4hD, 1hR and 3hR were collected and RNA was directly extracted with the RNAeasy Mini Kit (Quiagen). For each replicate, 40 clonal stage 3 buds per condition were pooled to perform RNA extraction. Sequencing was realized by the Transcriptomic Genomic platform of Marseille Luminy (TGML), using the Illumina Truseq stranded mRNA kit for library preparation and 75 nt single-end sequencing with the Illumina NextSeq 500 instrument, and 400 million reads per run. Raw sequencing data were submitted to the Submission Raw Archives (SRA) (SRR12538867, SRR12538868, SRR12538869, SRR12538870, SRR12538871, SRR12538872, SRR12538873, SRR12538874, SRR12538875, SRR12538876, SRR12538877, SRR12538878, SRR12538879, SRR12538880, SRR12538881, SRR12538882).

### De novo assembly of the *O. lobularis* transcriptome

The de novo transcriptome assembly strategy is shown in Fig. 5. In brief, 454 sequencing data from the adult in sexual reproduction with embryos and larvae [124] and illumine sequencing data from *O. lobularis* buds were first assembled individually and then combined to obtain a transcriptome containing transcripts from all life stages. The 454 dataset of adult *O. lobularis* was assembled with the Staden package and GAP4 [13] (v1.7.0) by Eurofins. Illumina reads were first quality filtered using FASTQC (v 0.11.8, available at <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and trimmed using TrimGalore (v0.6.4, available at <https://github.com/FelixKrueger/TrimGalore>). Trimmed sequences were de novo assembled using Trinity (v2.9.1, [57]). The two datasets were concatenated and clustered using CD-HIT EST (v4.8.1, [87]), whereby all sequences with 80% sequence identity were clustered together and the longest transcript was kept. We also used GMAP (v2019.12.01, [141]) to remove all shorter sequences that were 100% identical to a longest sequence. Vector contamination was removed with Vecscreen ([61], available at <https://www.ncbi.nlm.nih.gov/tools/vecscreen/>) using default parameters: this tool performs a BLAST+(BlastN [23]) search against the UniVec database. To remove sequences that might have contaminated our biological specimen, we performed BlastN searches against a custom database composed of Archaea, Bacteria, Viridiplantae, Fungi and Dinoflagellata (e-value  $10^{-6}$  and a percentage of identity of 80%). Protein coding sequences were predicted using Transdecoder (v5.5.0, available at <https://github.com/TransDecoder>).

By default, Transdecoder will identify ORFs that are at least 100 amino acids long. We cross-validated Transdecoder results with results from HMMScan ([33], available at <http://hmmer.org/>) against the PFAM database ([35], available at <http://pfam.xfam.org/>) for predicting the most likely coding regions. The completeness of the transcriptome was evaluated by BUSCO (v4.1.2, [130]). BUSCO scores were calculated based on the BUSCO Metazoan gene set. In addition we tested the completeness of conserved domain arrangements with DOGMA ([29], <https://domainworld-services.uni-muenster.de/dogma/>) (Additional file 7: Table S4 and Additional file 8: Table S5).

### Gene functional annotation of the de novo assembled *O. lobularis* transcriptome

We annotated gene functions by information transfer from orthologous sequences. Orthologous groups were identified with eggNOG mapper (v5.0, [65]). The percentage of orthologous groups that *O. lobularis* has in common with other metazoans species was inferred by Orthofinder searches (v2.2.7, [37]) against a set of selected species: *Homo sapiens* (GRCh38.pep from ENSEMBL), *Drosophila melanogaster* (BDGP6.28.pep from ENSEMBL), *Nemastostella vectensis* (ASM20922v1.pep from ENSEMBL), *Trichoplax adhaerens* (ASM15027v1.pep from ENSEMBL), *Sycon ciliatum* (SCIL\_T-PEP\_130802 from COMPAGEN), *Oscarella pearsei* (OCAR\_T-PEP\_130911 from COMPAGEN) *Amphimedon queenslandica* (Aqu1.pep from ENSEMBL) and *Aphrocallistes vastus* (AV4\_1\_Trinity from ERA archive of Alberta university) and *Salpingoeca rosetta* (UP000007799 from Uniprot). Functional annotation with GO terms was done using eggNOG and InterProScan (v5.40, [69]). Identification of conserved domains was done using CD-search (CDD-database version 3.18, [93]).

Identification of Aggregation Factors from the *Oscarella lobularis* transcriptome *Oscarella lobularis* aggregation factors were identified using BlastP (e-value = 10) with *Oscarella carmela* aggregation factor candidates from group 3 identified in the study of Grice [59] as queries. Protein domain predictions were done using Interproscan online with default parameters for each sequence (Additional file 16: Figure S5 and Additional file 17: Table S12).

### Differential and time-series analysis of RNA-seq data

Illumina sequences were mapped using Kallisto (v0.46.2, [18]) and Tximport [134] to resolve ambiguities in read mapping. Detection of differentially expressed genes was performed with DESeq2 [90]. Lowly expressed genes, defined by having less than 10 read counts across all samples, were excluded. Genes with  $\text{padj} < 0.05$  and  $\log_2$

fold change ( $\log_2|FC|$ ) of at least 1.5 were defined as differentially expressed. Pairwise comparisons were done as follows: dissociation time-points 1hD and 4hD were compared to NSW control and reaggregation time-points 1hR and 3hR were compared to the last time-point of dissociation, 4hD.

Soft clustering of time-course data was performed in R with Mfuzz [51], using the transcripts per million (TPM) count table (Additional file 19: Table S13) after a standardization. Hierarchical clustering, using the hclust function in R (Additional file 20: Figure S7), was performed to evaluate the number of clusters parameter needed as input for the Mfuzz clustering. By inspecting the dendrogram, 15 clusters showed to well discriminated the data. The fuzzifier value  $m$  prevents clustering of random data, is the second parameter needed for the algorithm, was calculated with the function mestimate. This value has been estimated at 1.99. Mfuzz clustering was repeated 10 times to test the clusters stability by calculating the Jaccard index to evaluate the similarity of the core cluster between each run (Additional file 21: Table S14). Mfuzz core clusters were restricted to genes with a membership value ( $\alpha$  greater or equal to 0.7; to visualize the relationship between clusters (Fig. 8), average expression profiles of member genes were calculated as the average standard-normal expression and then were clustered in R using Euclidean distance and complete linkage.

### Enrichment analysis

Enrichment analysis was performed with the TopGO package in R (available at <https://bioconductor.org/packages/release/bioc/html/topGO.html>) using the weight 01 algorithm, a classic Fisher's exact test on the Biological Process GO category with the annotated transcriptome as background. Enrichment of Conserved Domains was done with the clusterProfileR package on custom made Conserved Domain IDs [142] using a Fisher's exact test. Full results and gene lists are available in the (Additional file 11: Table S8).

### Phylostratigraphy analysis of the *O. lobularis* transcriptome and differentially expressed genes during dissociation and reaggregation

The evolutionary emergence of genes of the transcriptome and those regulated during the dissociation and reaggregation of *O. lobularis*, was estimated using the method proposed by [133] with a custom database containing 27 transcriptomes and proteomes (Additional file 10: Table S7). Identification of orthologous groups with eggNOG [65] and BlastP [23] searches was performed for every coding sequence of *O. lobularis* against this custom database, and its evolutionary age was

inferred based on the oldest Blast hit relative to the pre-determined phylogenetic classification of all 27 species (Additional file 12: Table S9).

### Data availability

This Transcriptome Shotgun Assembly project has been deposited at DDBJ/EMBL/GenBank under the Bioproject PRJNA659410. The version described in this paper is the first version publicly available. Additional file 5: Table S2 contains all the information available for the transcriptome. Translated version of the transcriptome is available in Additional file 22: File S1.

### Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12862-021-01866-x>.

**Additional file 1: Figure S1.** Specificity of the PhaE staining.

**Additional file 2: Figure S2.** S Comparison of the aspect of choanocyte chambers under different conditions (confocal microscopy views).

**Additional file 3: Table S1.** Estimated percentage of perturbations of the choanoderm after (A) 1 h of dissociation (1hD) regarding the aspect of the microvilli collar, (B) 4 h of dissociation (4hD) and (C) 24 h of reaggregation (24hR) concerning the shape of the choanocyte chambers. Statistical tests were performed with Wilcoxon rank sum test with continuity correction. Table related to Figs. 3, 4 and Additional file 2: Figure S2.

**Additional file 4: Figure S3.** Acetylated tubulin immunostaining of the choanocyte flagella during dissociation and reaggregation.

**Additional file 5: Table S2.** Functional transcriptome annotation using groups of orthologs (eggNOG), gene ontology terms (GO), conserved domains (CDs) and analyses of gene expression with DeSeq2. Differentially Expressed Genes (DEG) were identified with a threshold of  $pvalue \leq 0.05$  and with a  $\log_2$  Fold Change ( $\log_2|FC|$ ) of at least 1.5 at different times of the dissociation/reaggregation process. (A) List of transcripts with orthologs. (B) List of transcripts with GO terms. (C) List of transcripts with CDs. (D) List of transcripts without any annotations. Related to Figs. 6, 7 and Additional file 14: Figure S4.

**Additional file 6: Table S3.** Number of orthologs shared between species estimated with Orthofinder and Percentage of orthologs between each species and *O. lobularis* Related to Fig. 6.

**Additional file 7: Table S4.** General information about available sponge transcriptomes (according to Riesgo et al, 2014; [47, 134], This study).

**Additional file 8: Table S5.** Completeness of conserved domain arrangements (CDAs) estimated by the proportion of the CDAs found among the expected conserved CDAs in different metazoan species using DOGMA. Related to Fig. 6E.

**Additional file 9: Table S6.** Differential Expressed Genes (DEG) during (A) Dissociation process (B) dissociation and reaggregation processes and (C) reaggregation process. DEG extracted from DeSeq2 analysis with a threshold of  $pvalue \leq 0.05$  and  $\log_2|FC|$  of at least 1.5 at different times of dissociation (1hD, 4hD) and reaggregation (1hR, 3hR). Related to Fig. 7.

**Additional file 10: Table S7.** Phylostratigraphy analysis. (A) Transcriptomes and Proteomes used in Phylostratigraphy analyses. Each phylostratum represents a node. (B) Accession numbers and links of the different transcriptomes and proteomes used. (C) Phylostratigraphy analysis of the *Oscarella lobularis* transcriptome and of the differentially expressed genes during dissociation and reaggregation. (D) Number of genes with conserved CAMs domains (Cadherin, Integrin, Immunoglobulin) in each phylogenetic category with common expression profiles (see Fig. 8 for cluster numbers). Related to Fig. 7 and Additional file 1: Fig. 4.

**Additional file 11: Table S8.** GO term enrichment analysis. **(A)** Enriched categories in Differential Expressed Genes (DEG) during dissociation only, reaggregation only and both (combined). **(B)** Enriched categories in the first 100 DEGs. **(C)** Enriched categories in each core clusters (see Fig. 8 for cluster numbers and Additional file 11: Table S10 for complete list of genes included in the core cluster). Related to Figs. 7 and 8. Selected GO biological processes of interest are highlighted in yellow.

**Additional file 12: Table S9.** Identification of orthologous groups with eggNOG and BlastP searches **(A)** for significantly deregulated transcripts associated with GO biological processes of interest (complete list of GO terms in Additional file 9: Table S8) during dissociation, reaggregation and both (combined) processes shown in Table S6, **(B)** for the top 100 differentially expressed genes. Related to Fig. 7.

**Additional file 13: Table S10.** List of genes included in the core clusters (1 to 15) with orthologs (eggNOG) and gene ontology terms (GO). Related to Fig. 8.

**Additional file 14: Figure S4.** Distribution of the transcripts with CAMs conserved domain during dissociation and reaggregation.

**Additional file 15: Table S11:** Transcripts Per Million (TPM) standardized and Log<sub>2</sub> Fold Change values for **(A)** epithelial genes (see Fig. 1A) and **(B)** Aggregation factors (AFs). Related to Fig. 8 and 9.

**Additional file 16: Figure S5.** Domain composition of putative Aggregation Factors in *Oscarella lobularis*.

**Additional file 17: Table S12.** Domain predictions for putative group 3 Aggregation Factors (AFs) using InterproScan Website (<https://www.ebi.ac.uk/interpro/>). Related to Additional file 16: Figure S5.

**Additional file 18: Figure S6.** Validation of the antibody against *O. lobularis* type IV Collagen by peptide competition assay.

**Additional file 19: Table S13.** Raw file including mRNA-Seq Transcript Per Million count table using Kallisto.

**Additional file 20: Figure S7.** Hierarchical clustering on TPM values.

**Additional file 21: Table S14.** Mfuzz clusters stability. For each cluster the Jaccard index was calculated to evaluate the similarity of each cluster between each run (10 runs, Run 1 was taken as reference). Related to Fig. 8.

**Additional file 22: File S1.** Translated version of the Transcriptome obtain with Transdecoder (available at <https://github.com/TransDecoder>).

## Acknowledgements

We would like to thank the OEB Team of the IMBE lab for the financial support concerning the RNA sequencing; the IMBE molecular biology service, the IMBE morphology service, the diving service of OSU Institut Pytheas and divers from the IMBE lab (Christian Marschal, Sandrine Chenesseau) for sampling and their time and advice. We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees (Bioinfo Genotoul) and High-Performance Computing (HPC) cluster of the Institut Pytheas (OSU) for providing computing resources. A special thanks to all members of the Habermann Team (IBDM) for all their precious advice and technical support.

## Authors' contributions

AV performed the experiments, participated in transcriptome assembly and analysis, and writing of the manuscript. MMP participated in transcriptome assembly and analysis, and in writing of the manuscript. FM participated in transcriptome and time-series analysis helped edit the manuscript. HD was involved in validation of the antibody. NB participated in electron microscopy (SEM 3D) imaging and helped edit the manuscript. DMH participated in validation of the antibody and manuscript editing. CR provided technical support and participated in manuscript editing. ER was involved in project design, data analysis and manuscript writing. ALB was involved in project design, data analysis and manuscript writing. BHH was involved in bioinformatic project design, data analysis and manuscript writing. CB was involved in project design, data analysis and manuscript writing. All authors read and approved the final manuscript.

## Funding

This work was supported by the French National Center for Scientific Research (CNRS) and the Aix-Marseille University (AMU), the Amidex Foundation Project (AMX-18-INT-021) awarded to CB, the ANR-grant ANR-18-CE45-0016-01 MITO-DYNAMICS awarded to BHH, the CNRS PICS Grant (STRAS project) awarded to ER and the LabEx INFORM (ANR-11-LABX-0054) to DMH and ALB. AV's salary was supported by the French Ministry for Higher Education and Research (MESRI) and HD's salary was supported by the LabEx INFORM (ANR-11-LABX-0054). DMH salary was supported by the National Institute for Biomedical Research (INSERM). The project leading to this publication has received funding from the « Investissements d'Avenir » French Government program managed by the French National Research Agency (ANR-16-CONV-0001) and from Excellence Initiative of Aix-Marseille University - A\*MIDEX. Light and electron microscopy experiments were performed by the optical imaging and electron microscopy (PICS-FBI core facility) platforms of the Institute for Developmental Biology of Marseille (IBDM), a France Bio-Imaging infrastructure, supported by the French National Research Agency (ANR-10-INBS-04-01, «Investments for the future»).

## Availability of data and materials

All data generated or analysed during this study are included in this published article and its Additional files. Raw files of the sequencing and transcriptome assembly has been deposited at DDBJ/EMBL/GenBank under the Bioproject PRJNA659410.

## Declarations

### Ethics approval and consent to participate

'Not applicable' for that section.

### Consent for publication

'Not applicable' for that section.

### Competing interests

The authors declare that they have no competing interests.

## Author details

<sup>1</sup>Aix Marseille Univ, CNRS, IRD, IMBE UMR 7263, Avignon Université, Institut Méditerranéen de Biodiversité et d'Ecologie Marine et Continentale, Station Marine d'Endoume, Marseille, France. <sup>2</sup>Aix Marseille Univ, CNRS, UMR 7288, Developmental Biology Institute of Marseille Luminy (IBDM), Marseille, France. <sup>3</sup>Aix Marseille Univ, CNRS, UMR 7288, Developmental Biology Institute of Marseille Luminy (IBDM), Turing Center for Living Systems (CENTURI), Marseille, France.

Received: 16 March 2021 Accepted: 18 June 2021

Published online: 21 August 2021

## References

- Adamska M. Developmental signalling and emergence of animal multicellularity. In evolutionary transitions to multicellular life: principles and mechanisms, I. Ruiz-Trillo, and A.M. Nedelcu, eds. (Dordrecht: Springer Netherlands), 2015; pp. 425–450.
- Adamska M, Degnan SM, Green KM, Adamski M, Craigie A, Larroux C, Degnan BM. Wnt and TGF-beta expression in the sponge *Amphimedon queenslandica* and the origin of metazoan embryonic patterning. *PLoS ONE*. 2007;2:e1031.
- Adamska M, Larroux C, Adamski M, Green K, Lovas E, Koop D, Richards GS, Zwafink C, Degnan BM. Structure and expression of conserved Wnt pathway components in the demosponge *Amphimedon queenslandica*. *Evol Dev*. 2010;12:494–518.
- Alexa A. Rahnenfuhrer, J. topGO: enrichment analysis for gene ontology (R package version 2.40.0). Bioconductor. 2020.
- Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 1997;25:3389–402.

6. Apostolopoulou M, Ligon L. Cadherin-23 mediates heterotypic cell-cell adhesion between breast cancer epithelial cells and fibroblasts. *PLOS ONE*. 2012;7:e33289.
7. Assémat E, Bazellères E, Pallesi-Pocachard E, Le Bivic A, Massey-Harroche D. Polarity complex proteins. *Biochimica et Biophysica Acta (BBA) Biomembranes*. 2008;1778:614–30.
8. Baj J, Korona-Główniak I, Forma A, Maani A, Sitarz E, Rahnama-Hezavah M, Radzikowska E, Portincasa P. Mechanisms of the epithelial-mesenchymal transition and tumor microenvironment in *Helicobacter pylori*-induced gastric cancer. *Cells*. 2020;9:1055.
9. Bays JL, DeMali KA. Vinculin in cell–cell and cell–matrix adhesions. *Cell Mol Life Sci*. 2017;74:2999–3009.
10. Bazellères E, Aksenova V, Barthélémy-Requin M, Massey-Harroche D, Le Bivic A. Role of the Crumbs proteins in ciliogenesis, cell migration and actin organization. *Semin Cell Dev Biol*. 2018;81:13–20.
11. Belahbib H, Renard E, Santini S, Jourda C, Claverie J-M, Borchiellini C, Le Bivic A. New genomic data and analyses challenge the traditional vision of animal epithelium evolution. *BMC Genomics*. 2018;19:393.
12. Bich L, Pradeu T, Moreau J-F. Understanding multicellularity: the functional organization of the intercellular space. *Front Physiol*. 2019. <https://doi.org/10.3389/fphys.2019.01170>.
13. Bonfield JK, Smith KF, Staden R. A new DNA sequence assembly program. *Nucl Acids Res*. 1995;23:4992–9.
14. Borchiellini C, Degnan SM, Le Goff E, Rocher C, Vernale A, Baghdiguian S, Séjourné N, Marschal F, Le Bivic A, Godefroy N, et al. Staining and Tracking Methods for Studying Sponge Cell Dynamics. In: *Developmental Biology of the Sea Urchin and Other Marine Invertebrates: Methods and Protocols*, D.J. Carroll, and S.A. Stricker, eds. (New York, NY: Springer US), 2021; pp. 81–97.
15. Boury-Esnault N, Rützler K. *Thesaurus of sponge morphology*. Virginia: Smithsonian Institution Press; 1997.
16. Boury-Esnault N, Ereskovsky A, Bézac C, Tokina D. Larval development in the Homoscleromorpha (Porifera, Demospongiae). *Invertebr Biol*. 2003;122:187–202.
17. Boute N, Exposito J-Y, Boury-Esnault N, Vacelet J, Noro N, Miyazaki K, Yoshizato K, Garrone R. Type IV collagen in sponges, the missing link in basement membrane ubiquity. *Biol Cell*. 1996;88:37–44.
18. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic RNA-seq quantification. *Nat Biotechnol*. 2016;34:525–7.
19. Bridges L, Bowditch R. ADAM-integrin interactions: potential integrin regulated ectodomain shedding activity. *Curr Pharm Des*. 2005;11:837–47.
20. Brodskiy PA, Zartman JJ. Calcium as a signal integrator in developing epithelial tissues. *Phys Biol*. 2018;15:051001.
21. Brown RC, Davis TP. Calcium modulation of adherens and tight junction function: a potential mechanism for blood-brain barrier disruption after stroke. *Stroke*. 2002;33:1706–11.
22. Brunet T, King N. The origin of animal multicellularity and cell differentiation. *Dev Cell*. 2017;43:124–40.
23. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST+: architecture and applications. *BMC Bioinformatics*. 2009;10:421.
24. Curtis ASG. Pattern and mechanism in the reaggregation of sponges. *Nature*. 1962;196:245–8.
25. Curtis ASG. Problems and some solutions in the study of cellular aggregation. In: *Symp Zool Soc Lond*, 1970; pp. 335–352.
26. Daley AC, Antcliffe JB. Evolution: the battle of the first animals. *Curr Biol*. 2019;29:R257–9.
27. Deerinck TJ, Bushong E, Thor A, Ellisman MH. NCMIR methods for 3D EM: a new protocol for preparation of biological specimens for serial block-face SEM. *Microscopy*. 2010; pp. 6–8. <http://ncmir.ucsd.edu/sbfsem-protocol.pdf>.
28. DeMali KA, Sun X, Bui GA. Force transmission at cell-cell and cell-matrix adhesions. *Biochemistry*. 2014;53:7706–17.
29. Dohmen E, Kremer LPM, Bornberg-Bauer E, Kemena C. DOGMA: domain-based transcriptome and proteome quality assessment. *Bioinformatics*. 2016;32:2577–81.
30. Domazet-Lošo T, Tautz D. A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature*. 2010;468:815–8.
31. Donati G, Watt FM. Stem cell heterogeneity and plasticity in epithelia. *Cell Stem Cell*. 2015;16:465–76.
32. D'Souza RS, Lim JY, Turgut A, Servage K, Zhang J, Orth K, Sosale NG, Lazara MJ, Allegood JC, Casanova JE. Calcium-stimulated disassembly of focal adhesions mediated by an ORP3/IQSec1 complex. *BioRxiv*. 2019; 866392.
33. Eddy SR. Profile hidden Markov models. *Bioinformatics*. 1998;14:755–63.
34. Eerkes-Medrano D, Feehan CJ, Leys SP. Sponge cell aggregation: checkpoints in development indicate a high level of organismal complexity. *Invertebr Biol*. 2015;134(1):1–18.
35. El-Gebali S, Mistry J, Bateman A, Eddy SR, Luciani A, Potter SC, Qureshi M, Richardson LJ, Salazar GA, Smart A, et al. The Pfam protein families database in 2019. *Nucleic Acids Res*. 2019;47:D427–32.
36. Elliott GRD, Leys SP. Coordinated contractions effectively expel water from the aquiferous system of a freshwater sponge. *J Exp Biol*. 2007;210:3736–48.
37. Emms DM, Kelly S. OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biol*. 2019;20:238.
38. Ereskovsky AV. *The comparative embryology of sponges*. Berlin: Springer Science & Business Media; 2010.
39. Ereskovsky AV, Boury-Esnault N. Cleavage pattern in *Oscarella* species (Porifera, Demospongiae, Homoscleromorpha): transmission of maternal cells and symbiotic bacteria. *J Nat Hist*. 2002;36:1761–75.
40. Ereskovsky AV, Tokina DB. Asexual reproduction in homoscleromorph sponges (Porifera; Homoscleromorpha). *Mar Biol*. 2007;151:425–34.
41. Ereskovsky AV, Borchiellini C, Gazave E, Ivanisevic J, Lapébie P, Perez T, Renard E, Vacelet J. The Homoscleromorph sponge *Oscarella lobularis*, a promising sponge model in evolutionary and developmental biology. *BioEssays*. 2009;31:89–97.
42. Ereskovsky AV, Richter DJ, Lavrov DV, Schippers KJ, Nichols SA. Transcriptome sequencing and delimitation of sympatric *Oscarella* species (*O. carmela* and *O. pearsei* sp. nov) from California, USA. *PLOS ONE*. 2017;12:e183002.
43. Erives A, Fritzsch B. A screen for gene paralogs delineating evolutionary branching order of early metazoa. *G3 (Bethesda)*. 2019;10:811–26.
44. Fahey B, Degnan BM. Origin of animal epithelia: insights from the sponge genome: evolution of epithelia. *Evol Dev*. 2010;12:601–17.
45. Fernández-Busquets X, Gerosa D, Hess D, Burger MM. Accumulation in marine sponge grafts of the mRNA encoding the main proteins of the cell adhesion system. *J Biol Chem*. 1998;273:29545–53.
46. Feuda R, Dohrmann M, Pett W, Philippe H, Rota-Stabelli O, Lartillot N, Wörheide G, Pisani D. Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Curr Biol*. 2017;27(24):3864–3870.e4.
47. Fidler AL, Darris CE, Chetyrkin SV, Pedchenko VK, Boudko SP, Brown KL, Gray Jerome W, Hudson JK, Rokas A, Hudson BG. Collagen IV and basement membrane at the evolutionary dawn of metazoan tissues. *Elife*. 2017;18(6):e24176.
48. Fierro-Constain L, Schenkelaars Q, Gazave E, Haguenaer A, Rocher C, Ereskovsky A, Borchiellini C, Renard E. The conservation of the germline multipotency program, from sponges to vertebrates: a stepping stone to understanding the somatic and germline origins. *Genome Biol Evol*. 2017;9(3):474–88.
49. Francis WR, Canfield DE. Very few sites can reshape the inferred phylogenetic tree. *PeerJ*. 2020;8:e8865.
50. Francis WR, Eitel M, Vargs S, Adamski M, Haddock SH, Krebs S, Blum H, Erpenbeck D, Wörheide G. The genome of the contractile Demosponge *Tethya wilhelma* and the evolution of metazoan neural signalling pathways. *BioRxiv*. 2017;120998.
51. Futschik ME, Carlisle B. Noise-Robust soft clustering of gene expression time-course data. *J Bioinform Comput Biol*. 2005;03:965–88.
52. Galtsoff PS. Regeneration after dissociation (an experimental study on sponges). II. Histogenesis of microcilia proliferata, verr. *J Exp Zool*. 1925;42:223–55.
53. Ganot P, Zoccola D, Tambutté E, Voolstra CR, Aranda M, Allemand D, Tambutté S. Structural molecular components of septate junctions in cnidarians point to the origin of epithelial junctions in eukaryotes. *Mol Biol Evol* 2015;32:44–62.
54. Gazave E, Lapébie P, Ereskovsky AV, Vacelet J, Renard E, Cárdenas P, Borchiellini C. No longer Demospongiae: Homoscleromorpha formal nomination as a fourth class of Porifera. *Hydrobiologia*. 2012;687:3–10.
55. Giannone G, Rondé P, Gaire M, Beaudouin J, Haiech J, Ellenberg J, Takeda K. Calcium rises locally trigger focal adhesion disassembly and

- enhance residency of focal adhesion kinase at focal adhesions. *J Biol Chem*. 2004;279:28715–23.
56. Giebeler N, Zigrino P. A disintegrin and metalloprotease (ADAM): historical overview of their functions. *Toxins (Basel)*. 2016;8:122.
  57. Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol*. 2011;29:644–52.
  58. Green KJ, Roth-Carter Q, Niessen CM, Nichols SA. Tracing the evolutionary origin of desmosomes. *Curr Biol*. 2020;30:R535–43.
  59. Grice LF, Gauthier MEA, Roper KE, Fernández-Busquets X, Degnan SM, Degnan BM. Origin and evolution of the sponge aggregation factor gene family. *Mol Biol Evol*. 2017;34:1083–99.
  60. Gundacker D, Leys SP, Schröder HC, Müller IM, Müller WEG. Isolation and cloning of a C-type lectin from the hexactinellid sponge *Aphrocallistes vastus*: a putative aggregation factor. *Glycobiology*. 2001;11:21–9.
  61. Hancock JM, Bishop MJ. VecScreen. In *Dictionary of bioinformatics and computational biology*. Georgia: American Cancer Society; 2014.
  62. Harjunpää H, Lloret Asens M, Guenther C, Fagerholm SC. Cell adhesion molecules and their roles and regulation in the immune and tumor microenvironment. *Front Immunol*. 2019;10:1078.
  63. Harwood A, Coates JC. A prehistory of cell adhesion. *Curr Opin Cell Biol*. 2004;16:470–6.
  64. Henkart P, Humphreys S, Humphreys T. Characterization of sponge aggregation factor. A unique proteoglycan complex. *Biochemistry*. 1973;12:3045–50.
  65. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res*. 2019;47:D309–14.
  66. Humphreys T. Chemical dissolution and in vitro reconstruction of sponge cell adhesions: I. Isolation and functional demonstration of the components involved. *Dev Biol*. 1963;8:27–47.
  67. Humphreys T. Cell surface components participating in aggregation: evidence for a new cell particulate. *Exp Cell Res*. 1965;40:539–43.
  68. Huxley JS. Differences in viability in different types of regeneration from dissociated sponges, with a note on the entry of somatic cells by spermatozoa. *Biol Bull*. 1921;40:127–9.
  69. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. InterProScan 5: genome-scale protein function classification. *Bioinformatics*. 2014;30:1236–40.
  70. Kapli P, Telford MJ. Topology-dependent asymmetry in systematic errors affects phylogenetic placement of Ctenophora and Xenacoelomorpha. *Science Advances*. 2020;6:eabc5162.
  71. Kaszak I, Witkowska-Pilaszewicz O, Niewiadomska Z, Dworecka-Kaszak B, Ngosa Toka F, Jurka P. Role of cadherins in cancer—a review. *Int J Mol Sci*. 2020;21:7624.
  72. Kenny NJ, Francis WR, Rivera-Vicéns RE, Juravel K, Mendoza A, de Diez-Vives C, Lister R, Bezares-Calderon L, Grombacher L, Roller M, et al. Tracing animal genomic evolution with the chromosomal-level assembly of the freshwater sponge *Ephydatia muelleri*. *BioRxiv*. 2020; 2020.02.18.954784.
  73. Kim DH, Xing T, Yang Z, Dudek R, Lu Q, Chen Y-H. Epithelial mesenchymal transition in embryonic development, tissue repair and cancer: a comprehensive overview. *J Clin Med*. 2017;7:1.
  74. King N, Rokas A. Embracing uncertainty in reconstructing early animal evolution. *Curr Biol*. 2017;27:R1081–8.
  75. Ko KS, Arora PD, Bhide V, Chen A, McCulloch CA. Cell-cell adhesion in human fibroblasts requires calcium signaling. *J Cell Sci*. 2001;114:1155–67.
  76. Kumar A, Rizvi MS, Athilingam T, Parihar SS, Sinha P. Heterophilic cell-cell adhesion of atypical cadherins Fat and Dachous regulate epithelial cell size dynamics during *Drosophila* thorax morphogenesis. *MBoC*. 2019;31:546–60.
  77. Lapébie P, Gazave E, Ereskovsky A, Derelle R, Bézac C, Renard E, Houliston E, Borchiellini C. WNT/beta-catenin signalling and epithelial patterning in the homoscleromorph sponge *Oscarella*. *PLoS ONE*. 2009;4:e5823.
  78. Lavrov AI, Kosevich IA. Sponge cell reaggregation: mechanisms and dynamics of the process. *Russ J Dev Biol*. 2014;45:205–23.
  79. Lavrov AI, Kosevich IA. Sponge cell reaggregation: cellular structure and morphogenetic potencies of multicellular aggregates. *J Exp Zool*. 2016;325:158–77.
  80. Le Bivic A. Evolution and cell physiology. 4. Why invent yet another protein complex to build junctions in epithelial cells? *Am J Phys Cell Physiol*. 2013;305:C1193–201.
  81. Leith A. Role of aggregation factor and cell type in sponge cell adhesion. *Biol Bull*. 1979;156:212–23.
  82. Levin TC, Greaney AJ, Wetzel L, King N. The Rosetteless gene controls development in the choanoflagellate *S. rosetta*. *Elife*. 2014. <https://doi.org/10.7554/eLife.04070>.
  83. Leys SP, Hill A. The physiology and molecular biology of sponge tissues. In: *Advances in marine biology*, 2012; Elsevier. pp. 1–56.
  84. Leys SP, Riesgo A. Epithelia, an evolutionary novelty of metazoans. *J Exp Zool (Mol Dev Evol)*. 2012;318:438–47.
  85. Leys SP, Nichols SA, Adams EDM. Epithelia and integration in sponges. *Integr Comp Biol*. 2009;49:167–77.
  86. Leys SP, Mah JL, McGill PR, Hamonic L, De Leo FC, Kahn AS. Sponge behavior and the chemical basis of responses: a post-genomic view. *Integr Comp Biol*. 2019;59:751–64.
  87. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*. 2006;22:1658–9.
  88. Li Y, Shen XX, Evans B, Dunn CW, Roskas A, Tamura K. Rooting the animal tree of life. *Mol Biol Evol*. <https://doi.org/10.1093/molbev/msab170>.
  89. Littlewood DTJ. Animal evolution: last word on sponges-first? *Curr Biol*. 2017;27:R259–61.
  90. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15:550.
  91. MacLennan AP, Dodd RY. Promoting activity of extracellular materials on sponge cell reaggregation. 1967; 10.
  92. Marchadier E, Oates ME, Fang H, Donoghue PCJ, Hetherington AM, Gough J. Evolution of the calcium-based intracellular signaling system. *Genome Biol Evol*. 2016;8:2118–32.
  93. Marchler-Bauer A, Lu S, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, et al. CDD: a Conserved Domain Database for the functional annotation of proteins. *Nucleic Acids Res*. 2011;39:D225–9.
  94. Matsumoto J, Nakamoto C, Fujiwara S, Yubisui T, Kawamura K. A novel C-type lectin regulating cell growth, cell adhesion and cell differentiation of the multipotent epithelium in budding tunicates. *Development*. 2001;128:3339–47.
  95. de Mendoza A, Suga H, Ruiz-Trillo I. Evolution of the MAGUK protein gene family in premetazoan lineages. *BMC Evol Biol*. 2010;10:93.
  96. Miller PW, Clarke DN, Weis WI, Lowe CJ, Nelson WJ. The evolutionary origin of epithelial cell-cell adhesion mechanisms. *Curr Top Membr*. 2013;72:267–311.
  97. Miller PW, Pokutta S, Mitchell JM, Chodaparambil JV, Clarke DN, Nelson WJ, Weis WI, Nichols SA. Analysis of a vinculin homolog in a sponge (phylum Porifera) reveals that vertebrate-like cell adhesions emerged early in animal evolution. *J Biol Chem*. 2018;293:11674–86.
  98. Mitchell JM, Nichols SA. Diverse cell junctions with unique molecular composition in tissues of a sponge (Porifera). *EvoDevo*. 2019;10:26.
  99. Miyoshi J, Takai Y. Adherens junctions. In: Schwab M, editor. *Encyclopedia of cancer*. Berlin, Heidelberg: Springer; 2011. p. 57–61.
  100. Moscona AA. Studies on cell aggregation: demonstration of materials with selective cell-binding activity. *Proc Natl Acad Sci U S A*. 1963;49:742–7.
  101. Moscona AA. Cell aggregation: properties of specific cell-ligands and their role in the formation of multicellular systems. *Dev Biol*. 1968;18:250–77.
  102. Müller WEG, Zahn RK. Purification and characterization of a species-specific aggregation factor in sponges. *Exp Cell Res*. 1973;80:95–104.
  103. Müller WE, Müller I, Zahn RK, Kurelec B. Species-specific aggregation factor in sponges. VI. Aggregation receptor from the cell surface. *J Cell Sci*. 1976;21:227–41.
  104. Müller WE, Beyer R, Pondeljak V, Müller I, Zahn RK. Species-specific aggregation factor in sponges. XIII. Entire and core structure of the

- large circular proteid particle from *Geodia cydonium*. *Tissue Cell*. 1978;10:191–9.
105. Müller WE, Conrad J, Zahn RK, Gramzow M, Kurelec B, Uhlenbruck G. Identification and isolation of the primary aggregation factor from the cell membrane of the sponge *Geodia cydonium*. *Mol Cell Biochem*. 1985;67:55–64.
  106. Müller WEC, Conrad J, Zahn RK, Steffen R, Uhlenbruck G, Miller I. Cell adhesion molecule in the hexactinellid *Aphrocallistes vastus*: species-unspecific aggregation factor. *Differentiation*. 1984;26:30–5.
  107. Murray PS, Zaidel-Bar R. Pre-metazoan origins and evolution of the cadherin adhesome. *Biol Open*. 2014;3(12):1183–1195.
  108. Naik MU, Naik UP. Calcium-and integrin-binding protein regulates focal adhesion kinase activity during platelet spreading on immobilized fibrinogen. *Blood*. 2003;102:3629–36.
  109. Nichols SA, Dirks W, Pearse JS, King N. Early evolution of animal cell signaling and adhesion genes. *PNAS*. 2006;103:12451–6.
  110. Nichols SA, Roberts BW, Richter DJ, Fairclough SR, King N. Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/ $\beta$ -catenin complex. *PNAS*. 2012;109:13046–51.
  111. Nielsen C. Six major steps in animal evolution: are we derived sponge larvae? *Evol Dev*. 2008;10(2):241–257.
  112. Nielsen C. *Animal evolution: interrelationships of the living phyla*. Oxford: OUP Oxford; 2012.
  113. Nieto MA. Epithelial plasticity: a common theme in embryonic and cancer cells. *Science*. 2013;342:1234850.
  114. O'Keefe EJ, Briggaman RA, Herman B. Calcium-induced assembly of adherens junctions in keratinocytes. *J Cell Biol*. 1987;105:807–17.
  115. Plygawko AT, Kan S, Campbell K. Epithelial–mesenchymal plasticity: emerging parallels between tissue morphogenesis and cancer metastasis. *Philos Trans R Society B Biol Sci*. 2020;375:20200087.
  116. Renard E, Bivic AL, Borchiellini C. Origin and evolution of epithelial cell types. In: Leys S, Hejnal A, editors. *Origin and evolution of metazoan cell types*. Boca Raton: CRC Press; 2021. p. 75–100.
  117. Renard E, Leys SP, Wörheide G, Borchiellini C. Understanding animal evolution: the added value of sponge transcriptomics and genomics. *BioEssays*. 2018;40:1700237.
  118. Richter DJ, Fozouni P, Eisen MB, King N. Gene family innovation, conservation and loss on the animal stem lineage. *Elife*. 2018;7:e34226.
  119. Riesgo A, Farrar N, Windsor PJ, Giribet G, Leys SP. The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. *Mol Biol Evol*. 2014;31:1102–20.
  120. Rocher C, Vernale A, Fierro-Constain L, Sejourne N, Chenesseau S, Marschal C, Golf EL, Dutilleul M, Matthews C, Marschal F, et al. The buds of *Oscarella lobularis* (Porifera): a new convenient model for sponge cell and developmental biology. *BioRxiv*. 2020; 2020.06.23.167296.
  121. Rodriguez-Teja M, Gronau JH, Minamidate A, Darby S, Gaughan L, Robson C, Mauri F, Waxman J, Sturge J. Survival outcome and EMT suppression mediated by a lectin domain interaction of Endo180 and CD147. *Mol Cancer Res*. 2015;13:538–47.
  122. Rozario T, DeSimone DW. The extracellular matrix in development and morphogenesis: a dynamic view. *Dev Biol*. 2010;341:126–40.
  123. RStudio Team. *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA. 2020. <http://www.rstudio.com/>.
  124. Schenkelaars Q, Fierro-Constain L, Renard E, Hill AL, Borchiellini C. Insights into Frizzled evolution and new perspectives. *Evol Dev*. 2015;17:160–9.
  125. Schenkelaars Q, Fierro-Constain L, Renard E, Borchiellini C. Retracing the path of planar cell polarity. *BMC Evol Biol*. 2016;16:69.
  126. Schenkelaars Q, Vernale A, Fierro-Constain L, Borchiellini C, Renard E. A Look Back Over 20 Years of Evo-Devo Studies on Sponges: A Challenged View of Urmetazoa. In: Pontarotti P, editor. *Evolution, origin of life, concepts and methods*. Cham: Springer International Publishing; 2019. p. 135–60.
  127. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods*. 2012;9:676–82.
  128. Schippers KJ, Nichols SA, Wittkopp P. Evidence of signaling and adhesion roles for  $\beta$ -Catenin in the sponge *Ephydatia muelleri*. *Mol Biol Evol*. 2018;35:1407–21.
  129. Scott LE, Weinberg SH, Lemmon CA. Mechanochemical signaling of the extracellular matrix in epithelial-mesenchymal transition. *Front Cell Dev Biol*. 2019;7:135.
  130. Seppey M, Manni M, Zdobnov EM. BUSCO: Assessing Genome Assembly and Annotation Completeness. In: Kollmar M, editor. *Gene prediction: methods and protocols*. New York: Springer; 2019. p. 227–45.
  131. Simion P, Philippe H, Baurain D, Jager M, Richter DJ, Di Franco A, Roure B, Satoh N, Quéinnec É, Ereskovsky A, et al. A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Curr Biol CB*. 2017;27:958–67.
  132. Simpson TL. *The cell biology of sponges*. New York: Springer-Verlag; 1984.
  133. Sogabe S, Hatleberg WL, Kocot KM, Say TE, Stoupin D, Roper KE, Fernandez-Valverde SL, Degnan SM, Degnan BM. Pluripotency and the origin of animal multicellularity. *Nature*. 2019;570:519–22.
  134. Soneson C, Love MI, Robinson MD. Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res*. 2015;4:1521.
  135. Tyler S. Epithelium—the primary building block for metazoan complexity 1. *Integr Comp Biol*. 2003;43:55–63.
  136. Weber GF, Bjerke MA, DeSimone DW. Integrins and cadherins join forces to form adhesive networks. *J Cell Sci*. 2011;124:1183–93.
  137. Wilson HV. On some phenomena of coalescence and regeneration in sponges. *J Exp Zool*. 1907;5:245–58.
  138. Wimmer W, Perovic S, Kruse M, Schröder HC, Krasko A, Batel R, Müller WE. Origin of the integrin-mediated signal transduction. Functional studies with cell cultures from the sponge *Suberites domuncula*. *Eur J Biochem*. 1999;260:156–165.
  139. Windsor PJ, Leys SP. Wnt signaling and induction in the sponge aquiferous system: evidence for an ancient origin of the organizer. *Evol Dev*. 2010;12:484–93.
  140. Windsor Reid PJ, Matveev E, McClymont A, Posfai D, Hill AL, Leys SP. Wnt signaling and polarity in freshwater sponges. *BMC Evol Biol*. 2018;18:12.
  141. Wu TD, Watanabe CK. GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics*. 2005;21:1859–75.
  142. Yu G, Wang L, Han Y, He Q. ClusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–87. <https://doi.org/10.1089/omi.2011.0118>.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.