

# Functional specialization of cellulose synthase genes of prokaryotic origin in chordate larvaceans

Yoshimasa Sagane<sup>1</sup>, Karin Zech<sup>1</sup>, Jean-Marie Bouquet<sup>1</sup>, Martina Schmid<sup>1</sup>, Ugur Bal<sup>1,\*</sup> and Eric M. Thompson<sup>1,2,†</sup>

## SUMMARY

Extracellular matrices play important, but poorly investigated, roles in morphogenesis. Extracellular cellulose is central to regulation of pattern formation in plants, but among metazoans only tunicates are capable of cellulose biosynthesis. Cellulose synthase (*CesA*) gene products are present in filter-feeding structures of all tunicates and also regulate metamorphosis in the ascidian *Ciona*. *Ciona* *CesA* is proposed to have been acquired by lateral gene transfer from a prokaryote. We identified two *CesA* genes in the sister-class larvacean *Oikopleura dioica*. Each has a mosaic structure of a glycosyltransferase 2 domain upstream of a glycosyl hydrolase family 6 cellulase-like domain, a signature thus far unique to tunicates. Spatial-temporal expression analysis revealed that *Od-CesA1* produces long cellulose fibrils along the larval tail, whereas *Od-CesA2* is responsible for the cellulose scaffold of the post-metamorphic filter-feeding house. Knockdown of *Od-CesA1* inhibited cellulose production in the extracellular matrix of the larval tail. Notochord cells either failed to align or were misaligned, the tail did not elongate properly and tailbud embryos also exhibited a failure to hatch. Knockdown of *Od-CesA2* did not elicit any of these phenotypes and instead caused a mild delay in pre-house formation. Phylogenetic analyses including *Od-CesAs* indicate that a single lateral gene transfer event from a prokaryote at the base of the lineage conferred biosynthetic capacity in all tunicates. Ascidians possess one *CesA* gene, whereas duplicated larvacean genes have evolved distinct temporal and functional specializations. Extracellular cellulose microfibrils produced by the pre-metamorphic *Od-CesA1* duplicate have a role in notochord and tail morphogenesis.

**KEY WORDS:** Extracellular matrix, Notochord, Tunicate, Lateral gene transfer, Appendicularian, Metamorphosis, *Oikopleura*

## INTRODUCTION

Cellulose is the most abundant natural product in the biosphere with a variety of functional roles. Despite this abundance, the capacity to synthesize cellulose is restricted to relatively few phyla. Among prokaryotes, soil bacteria of the family Rhizobiaceae (*Agrobacterium tumefaciens* and *Rhizobium* spp) use cellulose in anchoring to host plant tissues during infection (Matthysse 1983; Smith et al., 1992). In *Acetobacter xilinum*, cellulose fibrils maintain bacterial cells in an aerobic environment in liquid and protect the cells from UV radiation (Williams and Cannon, 1989). Within the plant kingdom, cellulose plays a key role in structural support and the oriented deposition of cellulose microfibrils is crucial to patterning through anisotropic growth during development (Smith and Oppenheimer, 2005). The social amoeba, *Dictyostelium*, requires cellulose for stalk and spore formation (Blanton et al., 2000), and cellulose synthesis is also present in some fungi, although its function remains unclear (Stone, 2005). Among metazoans, cellulose biosynthesis is found only in the tunicate subphylum (Brown, 1999).

Cellulose is produced by multimeric cellulose synthase terminal complexes (TCs) inserted in the plasma membrane (Brown, 1996). In plants, the TCs are in the form of a rosette that moves through the

cell membrane as cellulose fibrils are extruded (Paredes et al., 2006). In bacteria (Brown et al., 1976), various algae (Brown and Montezinos, 1976; Itoh, 1990) and tunicate ascidians (Kimura and Itoh, 1996), the TCs are disposed in a stationary, linear organization. The tunicates comprise larvaceans, ascidians and thaliaceans. Post-metamorphic stages of the latter two groups incorporate cellulose into a tough integument, the tunic, which surrounds the animal and forms in part the filter-feeding apparatus (Hirose et al., 1999; Kimura and Itoh, 2004). Pre-metamorphic, non-feeding, larval ascidians are also surrounded by a tunic composed in part of cellulose, and in addition to its protective function, cellulose has a role in the control of *Ciona* metamorphosis. Insertional mutagenesis in the promoter of the *C. intestinalis* cellulose synthase (*CesA*) gene caused a drastic reduction of cellulose in the larval tunic resulting in a *swimming juvenile* (*sj*) mutant, where the order of metamorphic events was disrupted (Sasakura et al., 2005). *Sj* larvae initiated metamorphosis in the trunk without prior tail resorption. Further analysis of metamorphic pathways in *C. intestinalis* (Nakayama-Ishimura et al., 2009) revealed that cellulose represses initiation of papillae retraction and body axis rotation until larval settlement has occurred.

Larvaceans do not live inside a rigid tunic, but instead repetitively secrete and discard a complex, gelatinous filter-feeding house. The house comprises cellulose (Kimura et al., 2001) and on the order of at least 30 proteins (Spada et al., 2001; Thompson et al., 2001), and is secreted by a polyploid oikoplastic epithelium (Ganot and Thompson, 2002). Of 11 characterized house proteins, none show significant similarity with proteins in the sequenced ascidian genomes of *C. intestinalis* or *C. savignyi*, or in a broader sense, with any proteins in public databases, suggesting that these innovations are specific to the larvacean

<sup>1</sup>Sars International Centre for Marine Molecular Biology, Thormøhlensgate 55, N-5008 Bergen, Norway. <sup>2</sup>Department of Biology, University of Bergen, PO Box 7800, N-5020 Bergen, Norway

\*Present address: Department of Agricultural Biotechnology, Faculty of Agriculture, Namik Kemal University, Tekirdag, 59 030, Turkey

†Author for correspondence (eric.thompson@sars.uib.no)

lineage. Another important distinguishing feature of larvaceans is that they are the only tunicates to retain the chordate tail after metamorphosis. Traditionally, this has been considered neotenic, with both the common chordate and common tunicate ancestors viewed as having a free-swimming larval stage and sessile adult stage (Garstang, 1928; Nielsen, 1999; Lacalli, 2005; Stach, 2008a). Sequence analysis of rRNA genes (Wada and Satoh, 1994; Wada, 1998) and more extensive molecular phylogenomic datasets (Delsuc et al., 2006; Delsuc et al., 2008) support an opposing view that the primitive life cycle in chordates was entirely free-living, as in extant larvaceans, and place the larvaceans basal to ascidians and thaliaceans in the tunicate lineage.

To date, tunicate *CesA*s have only been identified in two ascidians, *Ciona savignyi* and *C. intestinalis*, each with one *CesA* gene encoding a protein with mosaic structure comprising an N-terminal cellulose synthase core domain and C-terminal cellulase-like domain. This mosaic domain organization has not been found in plant or bacterial *CesA*s. Phylogenetic analyses of ascidian *CesA*s suggested that the genes might have been acquired from a prokaryote by horizontal gene transfer prior to the split of *C. savignyi* and *C. intestinalis* (Matthysse et al., 2004; Nakashima et al., 2004). Matthysse et al. concluded that information on larvacean cellulose synthases would be essential to resolving whether a single horizontal gene transfer event was responsible for acquisition of cellulose synthetic capability in the entire tunicate lineage. Here, we have characterized two *CesA* genes in the larvacean *Oikopleura dioica*, and phylogenetic analyses of the two major domains support the hypothesis of a single lateral gene transfer event from a prokaryote at the base of the tunicate lineage. Spatial-temporal expression and knockdown experiments demonstrate that the two *O. dioica CesA* genes have distinct functional roles, one acting in the pre-metamorphic, and the second in the post-metamorphic, phase of the life cycle.

## MATERIALS AND METHODS

### Animal collection and culture

*Oikopleura dioica* were maintained in culture at 15°C (Bouquet et al., 2009). For in vitro fertilizations, females were collected in watch glasses, washed with artificial seawater (Red Sea, final salinity 30.4–30.5 g/l) and left to spawn. Sperm from 3–5 males was checked for viability and used for fertilization. Embryos were left to develop at room temperature.

### Cellulose synthase cloning

Putative *O. dioica* cellulose synthase genes were identified using the amino acid sequence of the *C. intestinalis* cellulose synthase (BAD10864) as a query in Tblastn against the *O. dioica* genomic shotgun dataset. Total RNA was isolated from 4 hours post-fertilization (hpf) and day 4 animals using Trizol (Invitrogen) according to manufacturer's instructions and cDNA was synthesized using GeneRacer (Invitrogen). Full-length cDNAs of the cellulose synthase genes were isolated by PCR using gene-specific primers.

### In silico analyses

To generate phylogenetic trees, amino acid sequences of enzymes with GT-2 or GH-6 domains were gathered from the NCBI protein database and aligned (ClustalW). Gaps, comprising less than 20% of the dataset, were deleted as missing data and remaining sequences were realigned (ClustalW). Phylogenetic analyses used Bayesian inference (Ronquist and Huelsenbeck, 2003). Analyses were done using the Jones amino acid model (Jones et al., 1992) with 1,000,000 generations sampled every 1000 generations. Putative transmembrane domains and topology of *CesA* proteins were predicted using the TMHMM algorithm (Krogh et al., 2001). Initiation ATG codons were predicted by NetStart 1.0 (Pedersen and Nielsen, 1997) and ATGpr (Salamov et al., 1998).

### Quantitative reverse transcriptase-polymerase chain reaction (RT-PCR)

Total RNAs were isolated from each stage using RNeasy (Qiagen). For RT-PCR, 1 µg of total RNA was subjected to RT using M-MLV RT (Invitrogen). Real-time PCRs (DNA Engine Opticon 2; MJ Research Waltham, MA, USA) were run using cDNA templates synthesized from an equivalent of 10 ng total RNA, 10 µl of Quantitect qPCR 2× Master Mix (Qiagen), 0.2 µM primers (see Table S1 in the supplementary material) in a total volume of 20 µl. After initial denaturation for 15 minutes at 95°C, 40 cycles of 95°C for 15 seconds, 58°C for 30 seconds and 72°C for 30 seconds were conducted, with a final extension for 5 minutes at 72°C. RT negative controls were run to 40 cycles of amplification. In all qRT-PCRs, 18S rRNA was used as a normalization control.

### Whole mount in situ hybridization

Fragments of 470, 466 and 861 bp for the *Od-CesA1*, *Od-CesA2* and *Od-Brachyury* (AF204208) genes, respectively, were PCR-amplified using specific primers (see Table S1 in the supplementary material) and cDNA libraries generated from 4 hpf (*Od-CesA1* and *Od-Brachyury*) and day 4 (*Od-CesA2*) animals. PCR products were cloned into pCRII-TOPO vector (Invitrogen). Sense and antisense RNA probes were synthesized by in vitro transcription of linearized plasmids with either T7 (Promega) or SP6 (Takara Bio) RNA polymerase in the presence of digoxigenin-labeled UTP (digoxigenin RNA Labeling Mix; Roche Molecular Biochemicals). Embryos at 5 hpf and day 4 animals were fixed in 4% paraformaldehyde, 0.1 M MOPS (pH 7.5) and 0.5 M NaCl at 4°C overnight, rinsed with 0.1 M MOPS (pH 7.5) and 0.5 M NaCl and then transferred to fresh 70% ethanol for storage at –20°C. Prior to hybridization, embryos were rehydrated in 50 mM Tris-HCl (pH 8.0) containing 0.1% Triton X-100. Hybridization and detection of probes were performed as described by Seo et al. (Seo et al., 2004).

### Confocal analysis of cellulose microfibrils

Embryos and animals were fixed in 4% paraformaldehyde, 0.1% saponin, 0.1 M MOPS (pH 7.5) and 0.5 M NaCl at 4°C overnight. Fixed animals were rinsed with PBS/0.1% saponin/0.1% Tween 20 (S/PBS-T) and then blocked with 3% BSA+S/PBS-T at 4°C overnight. Cellulose content was probed by incubation in 1% BSA+S/PBS-T containing rCBD-Protein L (10 µg/ml; Fluka) and mouse IgG (10 µg/ml; Sigma) at 4°C overnight, followed by incubation in Rhodamine-Red-X-conjugated goat anti-mouse IgG (1:200 in 1% BSA+S/PBS-T) at 4°C overnight. The rCBD-Protein L reagent can recognize other polysaccharides, notably, chitin. We therefore also performed specific staining for chitin using a chitin-binding probe (New England BioLabs) that recognized other structures distinct from those that we determined as cellulose with the rCBD-Protein L reagent. Finally, digestion with cellulase specifically eliminated the cellulose staining detected by rCBD-Protein L. To visualize cell shapes, cortical F-actin was stained with Alexa Fluor 488 Phalloidin (10 units/ml; Molecular Probes). Nuclei were counterstained with 1 µM To-Pro-3 iodide (Molecular Probes). Specimens were mounted in Vectashield (Vector Laboratories) and analyzed at 20°C with a Leica TCS laser scanning confocal microscope (Plan Apo 40× oil immersion 1.25 NA objective) using Leica v2.5 and Zeiss LSM 5 software.

### Morpholino knockdown experiments

Nucleotide sequences of the morpholino oligonucleotides (MOs) are given in Table S1 in the supplementary material. For *Od-CesA2* knockdown, a mixture of two MOs was used. The concentration of MO in the microinjection solution was 0.75 mM. MOs were injected into fertilized eggs before the first cleavage. In vitro fertilization of eggs and method of injection were as previously described (Clarke et al., 2007), except that sperm was obtained from pools of 10 males in 50 mm diameter petri dishes maintained on ice, and siliconized quartz capillaries (Sutter, QF100-70-10) pulled on a Sutter P2000 laser puller were used to prepare injection needles in place of aluminosilicate capillaries. The volume of injected solution was ~4 pl. To detect the splice modification of *Od-CesA1* and *Od-CesA2* genes, total RNA was extracted from the 1- and 4-hpf embryos for *Od-CesA1* and 10-hpf embryos for *Od-CesA2* by using Lysis II Buffer in the Cells-to-cDNA Kit (Ambion) according to

manufacturer's instructions. To generate cDNA, 10 µl of cell lysate was subjected to RT using M-MLV RT (Invitrogen). Nested PCR was performed using Dynazyme (Finnzymes) and specific primers (see Table S1 in the supplementary material).

### Rescue and phenocopy experiments

To rescue the effect of the splicing-blocking MO on *Od-CesA1* gene expression, a full-length cDNA containing three point mutations in the target region of the MO was synthesized by PCR (for primer sets, see Table S1 in the supplementary material). The PCR fragments were digested with restriction endonucleases, ligated and cloned into pCRII-TOPO TA (Invitrogen). To attempt to mimic the effect of the *Od-CesA1* splicing-blocking MO, a cDNA with a premature stop codon to generate an mRNA truncation similar to that generated by the splice-block MO was synthesized by PCR using the primer set cCesA1-01/cCesA1-d01 (see Table S1 in the supplementary material) and then cloned into pCRII-TOPO TA cloning vector. Capped mRNA (cmRNA) was synthesized using mMessage mMachine Sp6 (Ambion), tailed by Poly (A) Tailing (Ambion), precipitated with 5 M lithium chloride, washed four times with 70% ethanol and resuspended in nuclease-free water.

## RESULTS

### Duplicated *CesA* genes in *O. dioica*

Two loci homologous to *C. intestinalis CesA* (*Ci-CesA*) were identified in the *O. dioica* genomic database. To clone both genes (*Od-CesA1* and *Od-CesA2*), primers designed from these regions were used in a series of PCRs with cDNA from pools of 4-hpf embryos for *Od-CesA1* and day-4 animals for *Od-CesA2*. As a result, a 9-exon *Od-CesA1* gene and a 10-exon *Od-CesA2* gene, coding 1143 and 1252 amino acid residues, respectively, were identified (Fig. 1). Both encoded proteins had a mosaic structure with a cytoplasmic cellulose synthase core region featuring a glycosyltransferase 2 (GT-2) domain and a C-terminal extracellular glycosyl hydrolase family 6 (GH-6) cellulase-like domain (Fig. 1C). This organization was similar to *Ci-CesA* and *C. savignyi CesA* (*Cs-CesA*), and the *O. dioica* sequences had 51–57% amino acid similarity to the *Ciona* enzymes. This mosaic structure is not found in any other *CesA*s, and at present is unique to tunicate *CesA*s. In bacterial and fungal GH-6 cellulases, two aspartic acid residues are implicated in catalytic function (Rouvien et al., 1990), with the most C-terminal residue demonstrated to be crucial (Koivula et al., 2002). Alignment of urochordate GH-6 domains with bacterial and fungal domains (see Fig. S1 in the supplementary material) reveals both of these residues to be modified in *Ciona* *CesA*s and *Od-CesA1*, whereas *Od-CesA2* lacks the most C-terminal one. This raises questions as to the functional activity of the urochordate GH-6 domains. Cellulase activity is essential to cellulose biosynthesis in both prokaryotes and eukaryotes, although its precise role is unclear (Delmer, 1999). BLAST searches of the *O. dioica* genomic database revealed several putative GHF-9 cellulases, a family known in plants, bacteria, fungi and animals, including *Ciona* (Davidson and Blaxter, 2005). Among these, Korrigan is essential for cell-wall biosynthesis in *Arabidopsis* (Nicol et al., 1998). It is probable that some urochordate GHF-9 cellulases are active in cellulose biosynthesis as opposed to only being involved in digestion of dietary cellulose.

Bayesian phylogenetic analysis revealed higher phylogenetic affinity of tunicate GT-2 domains with corresponding bacterial domains than those in plants (Fig. 2A). GH-6 family proteins are found only in bacteria and fungi and are absent in plants and all animals except tunicates. Tunicate *CesA* GH-6 domains showed an affinity intermediate to bacterial and fungal cellulases (Fig. 2B). It has been proposed that ascidians acquired the *CesA* gene by horizontal transfer from bacteria (Matthysse et al., 2004; Nakashima et al., 2004; Sasakura et al., 2005). Based on phylogenetic analyses,

placing larvaceans nearer the base of the urochordate lineage than ascidians and thaliaceans (Wada and Satoh, 1994; Delsuc et al., 2006), the findings here indicate that horizontal transfer of the *CesA* gene occurred in the urochordate ancestor prior to divergence of the sister classes.

We identified two *CesA* paralogs in *O. dioica*, whereas *C. intestinalis* and *C. savignyi* each possess only one *CesA* gene. There are no conserved splice sites in the *O. dioica CesA* paralogs. This contrasts with the *Ciona CesA* homologs that share 14 conserved splice sites, including one that is conserved with *Od-CesA2*. In Table 1, amino acid similarities among the tunicate *CesA* proteins are shown for the whole sequence and the GT-2 and GH-6 domains. Within the urochordate lineage, the GH-6 cellulase domains are evolving more rapidly than the GT-2 glycosyl transferase domains. Overall, the *Od-CesA1* and *Od-CesA2* proteins exhibit slightly higher similarity to each other than either does to the individual *Ciona CesA* proteins. However, Bayesian trees using the GT-2 domains (Fig. 2A) or the GH-6 domains (Fig. 2B), yield different topologies. The GT-2 domain analysis suggests that *Od-CesA2* has greater affinity to the ascidian *CesA*s than does *Od-CesA1*. The GH-6 domain analysis suggests duplication of an ancestral *CesA* gene in the larvacean lineage.

### *Od-CesA1* and *Od-CesA2* form different extracellular structures

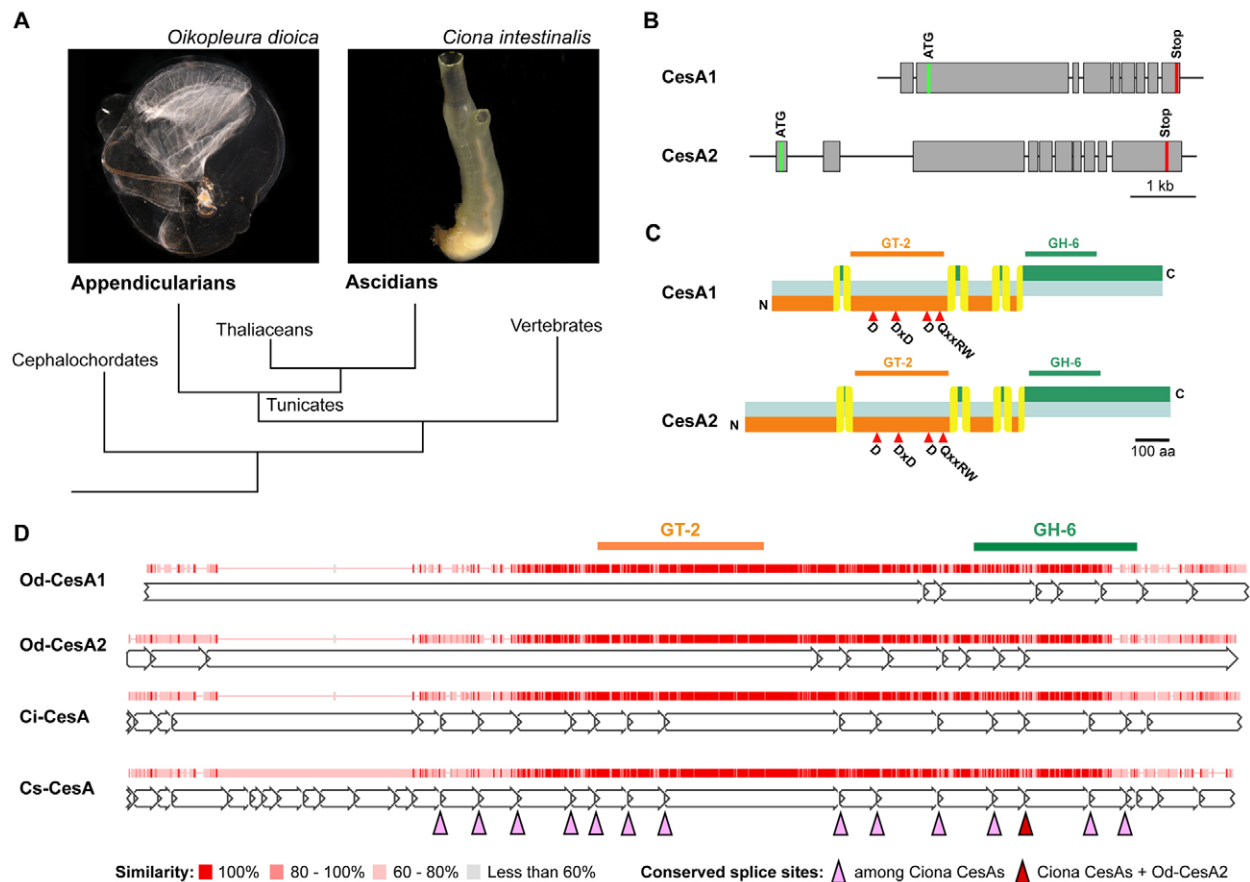
To analyze the temporal expression patterns of *Od-CesA1* and *Od-CesA2*, qRT-PCR was performed using cDNAs at twelve different development stages from oocyte to day 6. Expression of *Od-CesA1* was restricted to embryonic stages from 1 hpf to the hatching stage, whereas *Od-CesA2* was expressed at later stages from hatching to day 6 (Fig. 3A). Respective spatial expression patterns of these genes were identified by in situ hybridization (Fig. 3B). *Od-CesA1* was expressed at the lateral sides of the tail in tailbud embryos, whereas *Od-CesA2* was expressed in the oikoplastic epithelium, responsible for secretion of the filter-feeding house.

Cellulose microfibrils were first observed in pre-hatching tailbud embryos at 3 hpf (Fig. 3C). After hatching (4 hpf), the fibrils were seen to emerge laterally from the tail epidermis and aligned in an anterior-to-posterior orientation towards the tail tip. At 8 hpf, disintegration of the cellulose fibrils commenced in the anterior-most region of the tail and proceeded towards posterior regions of the tail over the next 2 hours. In parallel with the disappearance of cellulose fibrils, fin-like structures delimited by actin staining appeared along the tail margins. Cellulose staining in the oikoplastic epithelium initiated in local patches at 8 hpf and had spread over the entire trunk surface by 10 hpf. The appearance and disappearance of the cellulose fibrils corresponded very well to the spatial-temporal

**Table 1. Pairwise identities among tunicate *CesA*s**

A	Od CesA1	Od CesA2	Ci CesA	Cs CesA
Od CesA1	–	62.8	55.0	50.6
Od CesA2		–	56.7	56.7
Ci CesA			–	63.8
Cs CesA				–
B	Od CesA1	Od CesA2	Ci CesA	Cs CesA
Od CesA1	–	68.9	60.5	59.9
Od CesA2	83.2	–	59.3	58.1
Ci CesA	77.2	82.1	–	82.0
Cs CesA	77.2	83.2	97.8	–

The pairwise identities (%) were calculated using ClustalW and whole *CesA* (A), GT-2 domain (below the table diagonal in B) and GH-6 domain (above the table diagonal in B) amino acid sequences. Od, *O. dioica*; Ci, *C. intestinalis*; Cs, *C. savignyi*.



**Fig. 1. *Oikopleura dioica* cellulose synthase genes.** (A) Among metazoans, cellulose synthesis is restricted to the urochordates. Cellulose is present in the repetitively synthesized larvacean house and in the ascidian tunic. (B) Scaled schematic of the *O. dioica* cellulose synthase genes CesA1 (FN432362) and CesA2 (AM157749). (C) Domain organizations of the encoded cellulose synthase proteins indicating the intracellular glycosyltransferase 2 (GT-2, orange), transmembrane (yellow) and extracellular glycosylhydrolase family 6 (GH-6, green) domains. The GT-2 domain is traversed by seven predicted (<http://www.cbs.dtu.dk/services/TMHMM/>) transmembrane domains. Red arrowheads indicate conserved catalytic residues for GT-2 activity. (D) Conserved splice sites among tunicate CesAs. Entire amino acid sequences of *Od-CesA1* and *Od-CesA2* were aligned (ClustalW) with those of CesAs from *C. intestinalis* (*Ci-CesA*) and *C. savignyi* (*Cs-CesA*). Extent of similarity is indicated by degree of red shading of vertical bars. Gaps are indicated by light red horizontal lines. Exons are represented by white block arrows. Conserved splice sites are shown with arrowheads.

expression pattern of the *Od-CesA1* gene. Similarly, appearance of cellulose on the epithelium coincided with spatial-temporal expression of the *Od-CesA2* gene.

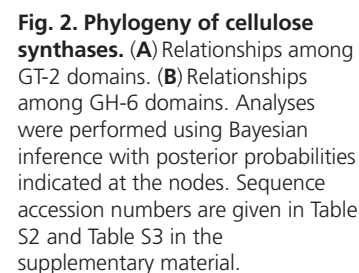
We further compared the cellulose structures in *O. dioica* with those in *C. intestinalis* (Fig. 3D). In *C. intestinalis* tadpoles, the entire animal was surrounded by cellulose. Fibers aligned in an anterior-to-posterior orientation as in *O. dioica*, were not observed on the lateral side of the tail and were only present at the tail tip. Similar to *O. dioica* tadpoles, actin staining delimited the fin-like structure along the tail margins.

### ***Od-CesA1* is required for embryo hatching, notochord alignment and tail elongation**

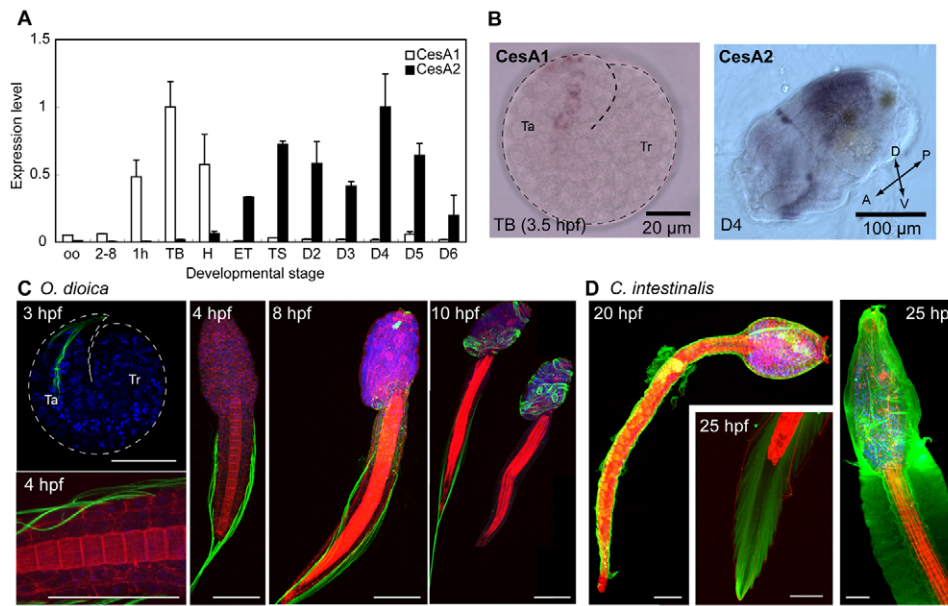
We designed MOs to block either *Od-CesA1* protein translation (*cesa1start*) or mRNA splicing (*cesa1e2i2*) (Fig. 4A). To assess whether MOs targeting the splice junctions could interfere with endogenous *Od-CesA1* transcripts in vivo, we injected *cesa1e2i2* MO or a 5-mismatched control MO into one-cell stage embryos, allowed them to develop until 1 hpf or 4 hpf, and then performed RT-PCR using primers located in exon2 and exon4 of the *Od-CesA1* gene. Control MO-injected and uninjected embryos yielded

expected wild-type products of 413 bp, whereas embryos injected with *cesa1e2i2* MO yielded a product of 318 bp (Fig. 4B). The nucleotide sequence of the shorter product extracted from *cesa1e2i2* MO-injected embryos revealed an excision of 95 bp from the 3'-end of exon2 due to activation of a cryptic splice donor site (Fig. 4C). This modification resulted in a frame shift downstream, creating a premature stop codon in exon3. This causes deletion of the last of the seven transmembrane helices and the entire GH-6 domain in the translation product of the incorrectly spliced mRNA (Fig. 4D). The ratio of the modified 318-bp product to the native 413-bp product was highest in 1-hpf embryos and decreased in 4-hpf embryos (Fig. 4B), suggesting that the efficiency of splice blocking decreased over this time interval.

Both translation blocking and splice-blocking MOs targeting *Od-CesA1* caused embryonic phenotypes (Fig. 5). In *cesa1start* MO-injected embryos, the predominant phenotype was a failure to elongate the tail and, additionally, an increase in failure of embryo hatching was observed. In *cesa1e2i2* MO-injected embryos there was an extensive failure of embryo hatching. To further assess the specificity of the MO effects we generated capped mRNA (cmRNA) from a rescue cDNA construct in which we had mutated three



Given the multimeric structure of cellulose synthase complexes, the different degree of severity of phenotypes is perhaps not surprising. The splice-blocking MO created a prematurely truncated form of Od-CesA1 in which almost the entire GT-2 domain was still present but the last transmembrane domain and the cellulase domain were deleted. This might have created a dominant-negative form that could have efficiently poisoned multimeric complexes that also

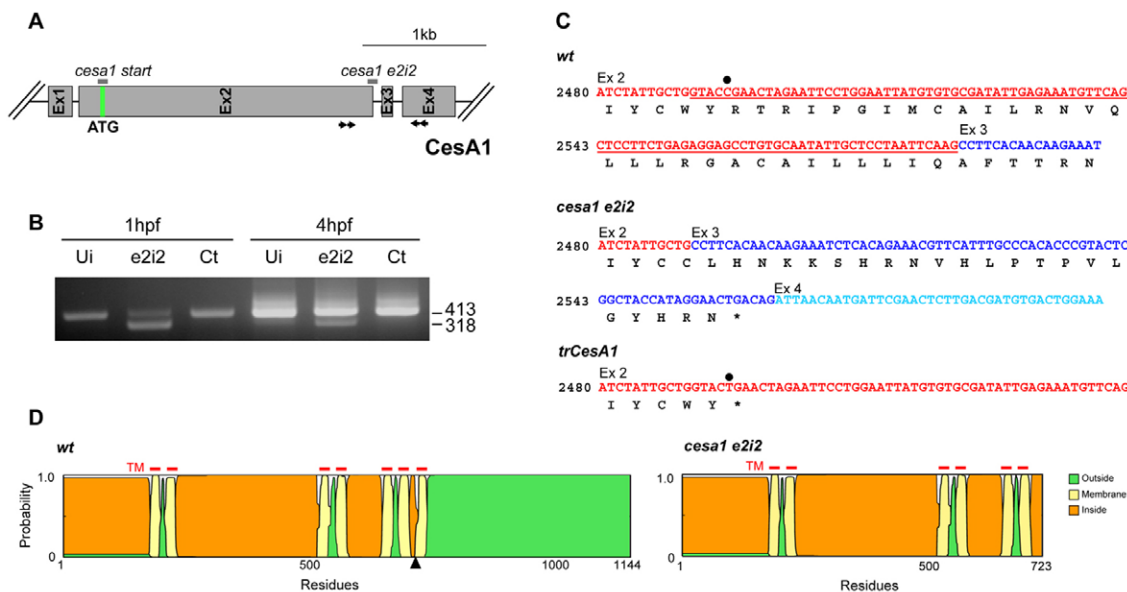


**Fig. 3. Cellulose synthase expression and patterning in *Oikopleura dioica* and *Ciona intestinalis*.** (A) Developmental expression profiles of CesA1 (white bars) and CesA2 (black bars) determined by qRT-PCR. oo, oocytes; 2-8, 2- to 8-cell embryos; 1h, 1 hour post-fertilization (hpf) embryos; TB, tail bud; H, hatching tadpole; ET, early tadpole; TS, tail shift; D2-D6, day 2 to day 6 animals. (B) Wholemount in situ hybridization patterns for CesA1 and CesA2 in 3.5-hpf embryos and day 4 animals. Ta, tail; Tr, trunk; A-P, anteroposterior axis; D-V, dorsoventral axis. (C) Confocal image stacks of cellulose-staining (green) in *O. dioica* embryos showing actin (red) and DNA (blue). (D) Confocal image stacks of cellulose staining in *C. intestinalis* embryos. Actin and DNA staining as in C. Scale bars: 50  $\mu$ m in C,D.

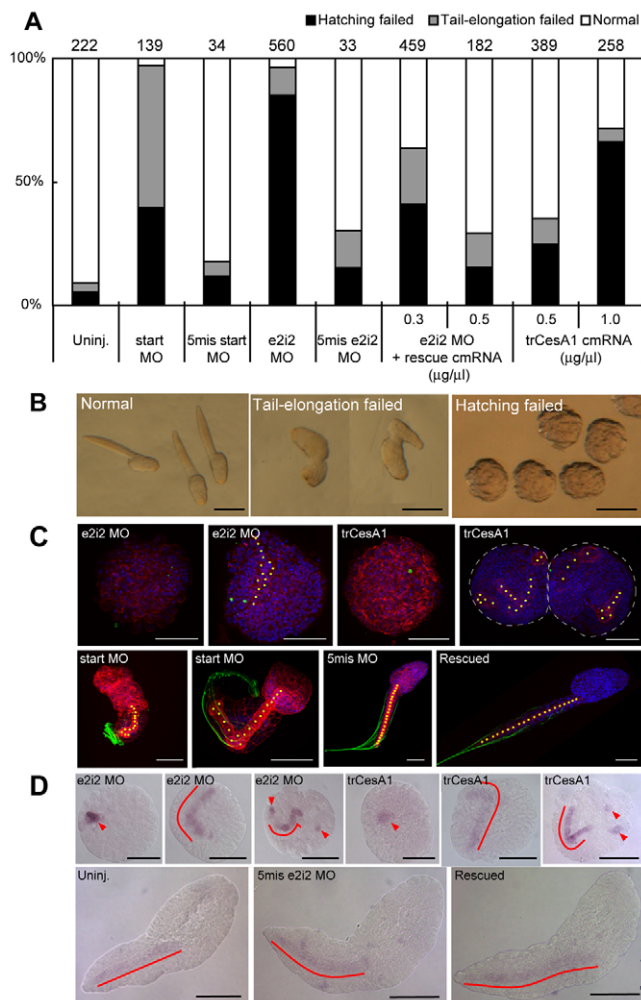
contained unmutated Od-CesA1 subunits. Conversely, the translation-blocking MO would reduce the quantity but not the quality of Od-CesA1 subunits produced. To test this idea, we created a truncated cDNA to mimic the RNA species produced by the *cesa1e2i2* MO (Fig. 4). Injection of this *trCesA1* cDNA did result in an increased ratio of hatching

failure to improper tail elongation and did so in a dose-dependent manner (Fig. 5A), consistent with poisoning of multimeric complexes by dominant-negative subunits.

Production of cellulose fibrils under the different experimental conditions was assessed with cellulose staining. In *cesa1start* MO-injected embryos, cellulose production was restricted to the



**Fig. 4. Knockdown of the *Oikopleura dioica* CesA1 gene.** (A) Schema showing the target locations for the translation-blocking morpholino (MO; *cesa1start*) and splice-blocking MO (*cesa1e2i2*). Nested primers used for RT-PCR are indicated by arrowheads. (B) RT-PCR of *cesa1e2i2* MO-injected embryos. The mRNA population isolated from MO-injected embryos yielded a smaller 318-bp band in addition to the wild-type 413-bp band at both 1 and 4 hpf. Ui, uninjected embryos; e2i2, *cesa1e2i2* MO-injected embryos; Ct, 5-mismatch *cesa1e2i2* MO-injected embryos. (C) Nucleotide sequences around the exon2 to exon3 junction in cDNAs generated from wild-type (wt) 4-hpf embryos and *cesa1e2i2* MO-injected 4-hpf embryos. A 95-bp sequence (underlined) was deleted from exon2 in MO-injected embryos through use of a cryptic splice donor site upstream of the MO-targeted splice donor. A truncated cDNA (*trCesA1*) was created by introducing a C-to-T point mutation (black dot) in order to produce truncated capped mRNAs to test whether this construct mimicked the effect of the *cesa1e2i2* MO. (D) Predicted (http://www.cbs.dtu.dk/services/TMHMM) transmembrane domains (TM, red bars) in wild-type and MO-disrupted CesA1. The vertical axes indicate average values of the posterior probabilities of inside, outside and transmembrane helix. MO injection results in a truncated protein lacking the seventh transmembrane helix and the entire GH-6 domain of CesA1. The black arrowhead in the wt representation indicates the position of the introduced premature stop codon in the *trCesA1* construct.



**Fig. 5. Phenotypes of embryos following manipulation of *CesaA1* expression by MO and capped mRNA (cmRNA) injections, or combinations of both.** (A,B) Injected embryos were scored morphologically 5 hpf and categorized as normal development, failure of tail elongation or failure to hatch. The number of embryos analyzed for each treatment is indicated above each column. Normal embryos were predominant in control uninjected and 5-mismatch MO-injected embryos. Failure to elongate the tail was predominant in *cesa1start* MO-injected embryos and an increase in the failure of embryo hatching was also observed. In *cesa1e2i2* MO-injected embryos there was extensive failure of embryo hatching. When a mutated *CesaA1* cmRNA refractory to base-pairing with the *cesa1e2i2* MO was co-injected with the MO, hatching was rescued. Degree of rescue was dose-dependent on the quantity of cmRNA co-injected. Injection of truncated cmRNA (*trCesA1*) mimicking mRNAs resulting from *cesa1e2i2* MO injection resulted in increased failure of hatching in a dose-dependent manner. (C) Confocal image stacks of the cellulose (green)-based lateral fin-like structure in *O. dioica* embryos. Red, actin; blue, DNA. 5-mismatch MO-injected embryos (both constructs) displayed no effects on cellulose structure compared with uninjected embryos shown in Fig. 3. By contrast, *cesa1start* MO-injected embryos demonstrated cellulose production restricted to the tail tip, whereas *cesa1e2i2* MO and *trCesA1* cmRNA-injected embryos exhibited an absence of cellulose production and failed to hatch. Yellow dots indicate notochord cells. In the upper right panel, two juxtaposed embryos are outlined with dashed lines. (D) Wholemount in situ hybridization patterns for the notochord marker, brachyury, in 5-hpf embryos. Red lines and arrowheads indicate expression domains. Perturbations in both alignment and continuity of the notochord cells were observed in embryos injected with *cesa1e2i2* MO or *trCesA1* cmRNA. Correct continuity and linear alignment were observed in uninjected and 5-mismatch MO-injected embryos, as well as those injected with a combination of the *cesa1e2i2* MO and rescue cmRNA. Scale bars: 100 μm in B; 50 μm in C,D.

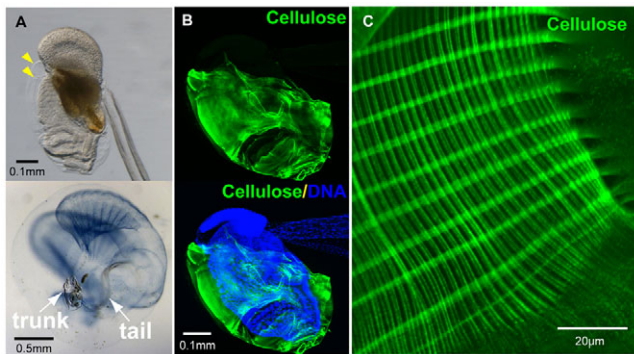
posterior portion of the tail, whereas in *cesa1e2i2* MO- or *trCesA1* mRNA-injected embryos no cellulose production was observed (Fig. 5C). In 5-mismatched MO-injected embryos, actin staining revealed a single linear row of notochord cells in the tail. In *cesa1start* MO-injected embryos, the alignment of notochord cells was perturbed, with some cells forming a ball-like agglomeration, and the shape of the cells was non-uniform. The point of notochord cell misalignment corresponded with the position of the anterior-most emergence of cellulose fibrils from the tail epidermis. Conversely, in *cesa1e2i2* MO- or *trCesA1* mRNA-injected embryos, where no cellulose fibril production was detected, Phalloidin staining revealed no typical linear arrangement of notochord cells. Co-injection of the rescue cmRNA with the *cesa1e2i2* MO recovered tail cellulose expression domains and the correct linear alignment of notochord cells. None of the constructs used in this study caused a failure of embryos to produce cells expressing the notochord differentiation-specific marker brachyury (Fig. 5D), but the ability to correctly align these notochord cells was clearly impaired by reduced or failed extracellular cellulose production.

### The cellulose-based filter-feeding house in post-metamorphic *O. dioica*

After tail elongation, metamorphosis occurs, with the tail switching from a posterior orientation to a final arrangement where the tail is orthogonal to the trunk and retains the notochord as its axial

structure. Then the first filter-feeding house is inflated. The filter-feeding house is initially secreted as a compact rudiment by a specialized oikoplasmic epithelium and several rudiment layers are often observed stacked above the trunk (Fig. 6A, upper panel). Upon escape of the animal from an inflated house, the outermost rudiment swells and is subsequently expanded by specific movements of the trunk and tail until the entire animal is contained within the mature structure (Fig. 6A, lower panel). Cellulose staining revealed the skeletal structure of the house rudiments and the food-concentrating filter and inlet filter are readily identified in pre-house rudiments (Fig. 6B). The inlet filter exhibited a meshwork composed of a single-warp and double-weft thread (Fig. 6C). The termini of each cellulose bundle branched into smaller fibrils.

We also designed MOs to block *Od-CesA2* mRNA splicing (*cesa2i2e3* and *cesa2e3i3*) and injected a mixture of these MOs into one-cell embryos. RT-PCR using primers located in exon2 and exon4 of the *Od-CesA2* gene (see Fig. S2A in the supplementary material) on cDNAs isolated from 10-hpf embryos revealed successful targeting of the *Od-CesA2* mRNA, with deletion of the entire GT-2 domain (see Fig. S2B in the supplementary material). Cellulose production in the *cesa2i2e3/cesa2e3i3* MO-injected embryos was analyzed by cellulose staining and compared with that in 5-mismatched control MO-injected embryos. Injection of these MOs had no effect on hatching, notochord formation, tail elongation or the production of cellulose fibrils along the tail in early embryos as observed when *Od-CesA1* was targeted. Instead, a minor phenotype was noted, where delayed cellulose



**Fig. 6. Cellulose structures in post-metamorphic *Oikopleura dioica*.** (A) Upper: day-5 animal with gonad at the top and mouth at the bottom has two uninflated pre-house rudiments (arrowheads) secreted around the trunk. Lower: day-3 animal inside an inflated house stained with India ink. The ribbed food-concentrating filter is visible at the top. (B) Confocal image stack of cellulose (green) in the rudiment (upper) superimposed on stained nuclei (blue) of the oikoplastic epithelium responsible for secretion of house components (lower). (C) Confocal image stacks of mesh formed by cellulose microfibrils in the maturing inlet filter.

production on the epithelium retarded pre-house formation (see Fig. S2C in the supplementary material). This suggests that MO injection into one-cell zygotes exhibited relatively limited penetrance on the *Od-CesA2* gene, which is expressed at high levels at later developmental stages than *Od-CesA1*.

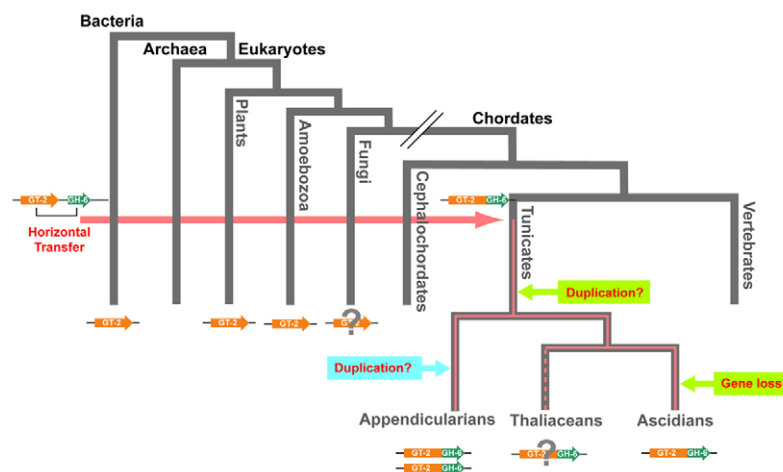
## DISCUSSION

We identified two larvacean *CesA* genes that show very distinct functional specializations. *Od-CesA1* has a pre-metamorphic function to produce long cellulose fibrils along the larval tail,

whereas *Od-CesA2* is responsible for primarily post-metamorphic production of the cellulose scaffold that forms, in part, the complex filter-feeding house. Knockdown of *Od-CesA1* using a splice-blocking MO resulted in a failure to produce cellulose fibrils along the tail and yielded a penetrant phenotype in which most embryos failed to hatch. Targeting of the same mRNA with a translation-blocking MO resulted in reduced production of cellulose fibrils and an elevated proportion of embryos that failed to hatch, but the major effect was failure to properly elongate the tail post-hatching.

Disruption of cellulose production in *Ciona sj* mutants did not impair embryo hatching. In this regard, it is notable that there is a large space between the *Ciona* chorion and embryo, whereas the *Oikopleura* embryo is tightly juxtaposed to the chorion. Thus, mechanical forces generated by the embryo might play a more significant role in hatching in *Oikopleura*. Cellulose fibrils might be implicated in these forces through facilitating sliding of trunk and tailbud cells against one another or through involvement in correct formation and ensuring sufficient rigidity of the tail. Relevant to this idea is that both MOs targeting *Od-CesA1* exhibited clear effects on disrupting correct notochord formation. In splice-blocking MO-injected embryos, we failed to observe a typical linear arrangement of any notochord cells, whereas in translation-blocking MO-injected embryos, alignment of notochord cells was disrupted. The shape of the notochord cells was non-uniform and some cells formed a ball-like agglomeration corresponding positionally with emergence of the anterior-most cellulose fibrils from the tail epidermis.

Phylogenetic analyses of the *Od-CesAs* show affinity with bacterial GT-2 domains and an intermediate affinity with bacterial and fungal GH-6 domains. These data support the hypothesis (Matthysse et al., 2004; Nakashima et al., 2004) of lateral gene transfer from a prokaryote to the chordate ancestor of the tunicate lineage (Fig. 7). However, the two domains give alternative topologies with respect to the evolution of *CesA* genes within tunicates. Analysis using the GT-2 domain suggests that gene duplication occurred in the common tunicate ancestor. Subsequently,



**Fig. 7. Origin and evolution of cellulose synthase (*CesA*) genes in the tunicate lineage.** Cellulose synthesizing genes containing glycosyltransferase 2 (GT-2) domains are found in bacteria, plants, amoebae and tunicates. Cellulose microfibrils are found in some fungal species but the gene(s) responsible for cellulose production have not yet been isolated. In many bacteria, cellulose synthase and endoglucanase genes are contained within single operons (Römling, 2002). The majority of these endoglucanases belong to family 8 of the glycosylhydrolases, although in *Streptomyces coelicolor*, a GH-6 glycosylhydrolase is present downstream of the glycosyltransferase gene, albeit in the opposite orientation (Xu et al., 2008). Our data support horizontal transfer of a prokaryotic *CesA*-like gene to the common ancestor of the tunicates. At this point, two scenarios are possible. The horizontally transferred gene underwent gene duplication at the base of the tunicate lineage and was retained in larvaceans (Appendicularians), while being lost in ascidians. Alternatively, gene duplication occurred specifically in the larvacean lineage, with ascidians retaining the ancestral single-copy state. The gene(s) responsible for cellulose production in thaliaceans have not yet been isolated. Further details are discussed in the text.

the ascidians would have lost the homolog of *Od-CesA1*, whereas larvaceans retained it. Trees based on the more rapidly evolving GH-6 domain suggest that the gene duplication event occurred in the larvacean lineage after their split from ascidians. Molecular phylogeny of the GT-2 domain indicates that *Od-CesA2* has more affinity to the *Ci-CesAs* than *Od-CesA1* and this is corroborated by a respective degree of retention of intron positions and function, with *Od-CesA2* being required for adult house formation in larvaceans and *Ci-CesA* necessary for adult tunic formation in ascidians. Characterization of the *CesA* complement in thaliaceans should help to resolve these alternative gain/loss scenarios.

Interestingly, in *O. dioica*, we found that cellulose is progressively degraded and lost along the larval tail (Fig. 3) and this precedes metamorphosis. This is at least superficially reminiscent of the loss of tail cellulose in ascidians (Nakayama-Ishimura et al., 2009) required for correct ordering of metamorphic events. In larvaceans, metamorphosis involves much less extensive morphological change than in ascidians. The longitudinal axis of the larval tail is aligned with the anteroposterior axis of the trunk in both groups. Whereas in ascidians the larval tail is resorbed and lost during metamorphosis, in larvaceans it merely undergoes migration to the ventral side of the trunk such that its longitudinal axis becomes orthogonal to the trunk. It remains a point of debate as to whether tail loss or retention is more representative of the ancestral tunicate. The morphological data suggests larvaceans are neotenic (Stach, 2008a), whereas molecular phylogenetic data (Delsuc et al., 2006; Delsuc et al., 2008) and a filter-feeding hypothesis on urochordate evolution (Satoh, 2009) places them basal to ascidians. Whereas repression of metamorphic initiation by *Ci-CesA* and/or cellulose in ascidians is alleviated through tail loss, in larvaceans the tail must be retained in juveniles and adults as an integral part of the feeding mechanism. Instead, tail cellulose is lost through developmental regulation of the *Od-CesA1* paralog. It is possible that cellulose fibrils emerging from the larval tail of larvaceans would simply impair the supple sinusoidal movement of the juvenile and adult tail required to regulate the flow of water through the filter-feeding house and to inflate new houses, rendering the timing of cellulose loss merely coincidental with the initiation of metamorphosis in this lineage. Experiments to prolong the expression of *Od-CesA1* could be informative as to whether this would delay the metamorphic tailshift, suggesting a conserved role for cellulose in regulating timing of tunicate metamorphosis, or only impair post-metamorphic tail function.

The horizontal transfer of a prokaryotic gene giving rise to the extant tunicate *CesAs* is more than a mere curiosity. It has been speculated that the ability to secrete a protective covering could have significantly impacted life history strategies by prohibiting larval feeding and increasing evolutionary pressure on speed of development (Stach, 2008b). Thus, relative to other chordates, the notable acceleration of tunicate development, greatly accentuated in the fully planktonic larvaceans, might have been triggered by the ability to secrete a tunic after the lateral gene transfer event. In a larger sense, tunicates, which are uniformly filter-feeders, have combined the ability to synthesize cellulose with cellular mechanisms enabling the elaboration of complex extracellular structures, some of which are invariably associated with the filter-feeding mechanism. The sister vertebrates, lacking cellulose synthetic capability, exhibit a variety of more active feeding mechanisms, including filter-feeding, and have undergone considerable elaboration of skeletal, sensory and nervous systems compared with tunicates and the common chordate ancestor. Arguably therefore, the lateral gene transfer event has had a

profound influence on the tunicate lineage, which has undergone secondary morphological simplification and is evolving at faster evolutionary rates than their vertebrate cousins (Delsuc et al., 2006). It will be of considerable interest to investigate how tunicate *CesAs* have been integrated into metazoan cell machinery in order to scaffold complex extracellular structures and to further explore the roles of *CesA* and cellulose in tunicate notochord formation and metamorphosis.

#### Acknowledgements

We thank the staff from Appendic Park for supplying animals, and Daniel Chourrout, Sars and Genoscope, France for the development of *Oikopleura* genomic resources. This work was supported by grant 17541/S10 NFR-FUGE from the Norwegian Research Council (E.M.T.).

#### Competing interests statement

The authors declare no competing financial interests.

#### Supplementary material

Supplementary material for this article is available at <http://dev.biologists.org/lookup/suppl/doi:10.1242/dev.044503/-/DC1>

#### References

- Blanton, R. L., Fuller, D., Iranfar, N., Grimson, M. J. and Loomis, W. F. (2000). The cellulose synthase gene of *Dictyostelium*. *Proc. Natl. Acad. Sci. USA* **97**, 2391-2396.
- Bouquet, J.-M., Spriet, E., Troedsson, C., Otterå, H., Chourrout, D. and Thompson, E. M. (2009). Culture optimization for the emergent zooplanktonic model organism *Oikopleura dioica*. *J. Plankton Res.* **31**, 359-370.
- Brown, R. M., Jr (1996). The biosynthesis of cellulose. *Pure Appl. Chem.* **10**, 1345-1373.
- Brown, R. M., Jr (1999). Cellulose structure and biosynthesis. *Pure Appl. Chem.* **71**, 767-776.
- Brown, R. M., Jr and Montezinos, D. (1976). Cellulose microfibrils: visualization of biosynthetic and orienting complexes in association with the plasma membrane. *Proc. Natl. Acad. Sci. USA* **73**, 143-147.
- Brown, R. M., Jr, Willison, J. H. and Richardson, C. L. (1976). Cellulose biosynthesis in *Acetobacter xylinum*: visualization of the site of synthesis and direct measurement of the in vivo process. *Proc. Natl. Acad. Sci. USA* **73**, 4565-4569.
- Clarke, T., Bouquet, J.-M., Fu, X., Kalløe, T. and Thompson, E. M. (2007). Rapidly evolving lamins in a chordate, *Oikopleura dioica*, with unusual nuclear architecture. *Gene* **396**, 159-169.
- Davidson, A. and Blaxter, M. (2005). Ancient origin of glycosyl hydrolase family 9 cellulase genes. *Mol. Biol. Evol.* **22**, 1273-1284.
- Delmer, D. P. (1999). Cellulose biosynthesis: exciting times for a difficult field of study. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **50**, 245-276.
- Delsuc, F., Brinkmann, H., Chourrout, D. and Philippe, H. (2006). Tunicates and not cephalochordates are the closest living relatives of vertebrates. *Nature* **439**, 965-968.
- Delsuc, F., Tsagkogeorga, G., Lartillot, N. and Philippe, H. (2008). Additional molecular support for the new chordate phylogeny. *Genesis* **46**, 592-604.
- Ganot, P. and Thompson, E. M. (2002). Patterning through differential endoreduplication in epithelial organogenesis of the chordate, *Oikopleura dioica*. *Dev. Biol.* **252**, 59-71.
- Garstang, W. (1928). The morphology of the Tunicate, and its bearings on the phylogeny of the Chordata. *Q. J. Microsc. Sci.* **72**, 51-187.
- Hirose, E., Kimura, S., Itoh, T. and Nishikawa, J. (1999). Tunic morphology and cellulose components of pyrosomas, doliods, and salps (thaliacea, urochordate). *Biol. Bull.* **196**, 113-120.
- Itoh, T. (1990). Cellulose-synthesizing complexes in some giant marine algae. *J. Cell Sci.* **95**, 309-319.
- Jones, D. T., Taylor, W. R. and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Comput. Appl. Biosci.* **8**, 275-282.
- Kimura, S. and Ito, T. (1996). New cellulose-synthesizing complexes (=terminal complexes) involved in animal cellulose biosynthesis in the tunicate, *Metandrocarpa uedai*. *Protoplasma* **194**, 151-163.
- Kimura, S. and Itoh, T. (2004). Cellulose synthesizing terminal complexes on the ascidians. *Cellulose* **11**, 377-383.
- Kimura, S., Ohshima, C., Hirose, E., Nishikawa, J. and Itoh, T. (2001). Cellulose in the house of the appendicularian *Oikopleura rufescens*. *Protoplasma* **216**, 71-74.
- Koivula, A., Ruohonen, L., Wohlfahrt, G., Reinikainen, T., Teeri, T. T., Piens, K., Claeysens, M., Weber, M., Vasella, A., Becker, D. et al. (2002). The active site of cellobiohydrolase Cel6A from *Trichoderma reesei*: the roles of aspartic acids D221 and D175. *J. Am. Chem. Soc.* **124**, 10015-10024.

- Krogh, A., Larsson, B., von Heijne, G. and Sonnhammer, E. L. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567-580.
- Lacalli, T. C. (2005). Protochordate body plan and the evolutionary role of larvae: Old controversies resolved? *Can. J. Zool.* **83**, 216-224.
- Matthyse, A. G. (1983). Role of bacterial cellulose fibrils in *Agrobacterium tumefaciens* infection. *J. Bacteriol.* **154**, 906-915.
- Matthyse, A. G., Deschet, K., Williams, M., Marry, M., White, A. R. and Smith, W. C. (2004). A functional cellulose synthase from ascidian epidermis. *Proc. Natl. Acad. Sci. USA* **101**, 986-991.
- Nakashima, K., Yamada, L., Satou, Y., Azuma, J. and Satoh, N. (2004). The evolutionary origin of animal cellulose synthase. *Dev. Genes Evol.* **214**, 81-88.
- Nakayama-Ishimura, A., Chambon, J. P., Horie, T., Satoh, N. and Sasakura, Y. (2009). Delineating metamorphic pathways in the ascidian *Ciona intestinalis*. *Dev. Biol.* **326**, 357-367.
- Nicol, F., His, I., Jauneau, A., Vernhettes, S., Canut, H. and Höfte, H. (1998). A plasmamembrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in *Arabidopsis*. *EMBO J.* **17**, 5563-5576.
- Nielsen, C. (1999). Origin of the chordate central nervous system – and the origin of the chordates. *Dev. Genes. Evol.* **209**, 198-205.
- Parédez, A. R., Somerville, C. R. and Ehrhardt, D. W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. *Science* **312**, 1491-1495.
- Pedersen, A. G. and Nielsen, H. (1997). Neural network prediction of translation initiation sites in eukaryotes: perspectives for EST and genome analysis. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* **5**, 225-233.
- Römling, U. (2002). Molecular biology of cellulose production in bacteria. *Res. Microbiol.* **153**, 205-212.
- Ronquist, F. and Huelsenbeck, J. P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.
- Rouvainen, J., Bergfors, T., Teeri, T., Knowles, J. K. C. and Jones, T. A. (1990). Three-dimensional structure of cellobiohydrolase II from *Trichoderma reesei*. *Science* **249**, 380-386.
- Sasakura, Y., Nakashima, K., Awazu, S., Matsuoka, T., Nakayama, A., Azuma, J.-I. and Satoh, N. (2005). Transposon-mediated insertional mutagenesis revealed the functions of animal cellulose synthase in the ascidian *Ciona intestinalis*. *Proc. Natl. Acad. Sci. USA* **102**, 15134-15139.
- Satoh, N. (2009). An advanced filter-feeder hypothesis for urochordate evolution. *Zool. Sci.* **26**, 97-111.
- Seo, H.-C., Edvardsen, R. B., Maeland, A. D., Bjordal, M., Jensen, M. F., Hansen, A., Flaatt, M., Weissenbach, J., Lehrach, H., Wincker, P. et al. (2004). Hox cluster disintegration with persistent anteroposterior order of expression in *Oikopleura dioica*. *Nature* **431**, 67-71.
- Smith, G., Swart, S., Lugtenberg, B. J. J. and Kijne, J. W. (1992). Molecular mechanisms of attachment of Rhizobium bacteria to plant roots. *Mol. Microbiol.* **6**, 2897-2903.
- Smith, L. G. and Oppenheimer, D. G. (2005). Spatial control of cell expansion by the plant cytoskeleton. *Annu. Rev. Cell. Dev. Biol.* **21**, 271-295.
- Spada, F., Steen, H., Troedsson, C., Kallesøe, T., Spriet, E., Mann, M. and Thompson, E. M. (2001). Molecular patterning of the oikoplasmic epithelium of the larvacean tunicate *Oikopleura dioica*. *J. Biol. Chem.* **276**, 20624-20632.
- Stach, T. (2008a). Chordate phylogeny and evolution: a not so simple three taxon problem. *J. Zool.* **276**, 117-141.
- Stach, T. (2008b). Anatomy of the trunk mesoderm in tunicates: homology considerations and phylogenetic interpretation. *Zoomorphology* **126**, 203-214.
- Stone, B. (2005). Cellulose: structure and distribution. In *Encyclopedia of Life Sciences*. Chichester: John Wiley & Sons. <http://www.els.net/> (doi: 10.1038/npg.els.0003892).
- Thompson, E. M., Kallesøe, T. and Spada, F. (2001). Diverse genes expressed in distinct regions of the trunk epithelium define a monolayer cellular template for construction of the oikopleurid house. *Dev. Biol.* **238**, 260-273.
- Wada, H. (1998). Evolutionary history of free-swimming and sessile lifestyles in urochordates as deduced from 18S rDNA molecular phylogeny. *Mol. Biol. Evol.* **15**, 1189-1194.
- Wada, H. and Satoh, N. (1994). Details of the evolutionary history from invertebrates to vertebrates, as deduced from the sequences of 18S rDNA. *Proc. Natl. Acad. Sci. USA* **91**, 1801-1804.
- Williams, W. S. and Cannon, R. E. (1989). Alternative environmental roles for cellulose produced by *Acetobacter xylinum*. *Appl. Environ. Microbiol.* **55**, 2448-2452.
- Xu, H., Chater, K. F., Deng, Z. and Tao, M. (2008). A cellulose synthase-like protein involved in hyphal tip growth and morphological differentiation in *Streptomyces*. *J. Bacteriol.* **190**, 4971-4978.

Table S1. Nucleotide sequences of primers and MOs used in this study

Primer	Sequence*	Direction
qRT-PCR for developmental stages		
<u>CesA1</u>		
qCesA2-01	3'-atttcatgccgcggcgaaactatc-5'	Sense
qCesA1-02	3'-tgatcccaaacatcttcgtccgca-5'	Anti-sense
<u>CesA2</u>		
YS001	3'-acgtctacttgactacgttgcca-5'	Sense
YS002	3'-tcagcatcgaagataacgacggct-5'	Anti-sense
PCR for ISH templates		
<u>CesA1</u>		
CesA1-p03	3'-gatttcatgccgcggcgaaactat-5'	Sense
CesA1-p04	3'-atccgctgtgacattgtgagacct-5'	Anti-sense
<u>CesA2</u>		
CesA-F	3'-tttcgcgaggaatcgaaacg-5'	Sense
CesAL-R	3'-ccgtagttgggatccttg-5'	Anti-sense
<u>Brachyury</u>		
OdiT-03	3'-tggccgacgaatgtttccggttat-5'	Sense
OdiT-04	3'-agctctggcgtcttccaatccggta-5'	Anti-sense
RT-PCR for MO-injected embryos		
<u>CesA1</u>		
CesA1e2-F05	3'-agatgatgacgatgactctggcgt-5'	Sense
CesA1e4-R02	3'-acgatcaggcgcatgtaatgaga-5'	Anti-sense
CesA1e2-F06	3'-tcgcgcatcagtcagttt-5'	Sense
CesA1e4-R03	3'-aacagggacagacgaatcttcgca-5'	Anti-sense
<u>CesA2</u>		
cesa2e2-01	3'-tgaaatcgctcccttagaagcca-5'	Sense
cesa2e4-01	3'-acgtcggacaccgttgaagtatgt-5'	Anti-sense
cesa2e2-02	3'-atttcgcgaggaatcgaaacgctg-5'	Sense
cesa2e4-02	3'-attggttgacgccaactctcgtt-5'	Anti-sense
CesA2e5-F01	3'-tcaactgccgtcgaagtcaagttct-5'	Sense
CesA2e6-R01	3'-gtggcgatcgaccaatttcagcaa-5'	Anti-sense
Cloning of full-length CesA1 cDNA with introduced point mutations		
cCesA1-01	3'-tcgcagttagcgaagatccgact-5'	Sense
cCesA1-04	3'-tgccctgaataatcaagcttgggc-5'	Anti-sense
cCesA1-05	3'-tcgatgtcctggactgaatcgaa-5'	Sense
cCesA1-06	3'-agaagaactgaggcgtctgcacta-5'	Anti-sense
cCesA1-07	3'-atctctggcgagtgttctctgtt-5'	Sense
cCesA1-08	3'-caccaacgattctctgacgccat-5'	Anti-sense
cCesA1-me01	3'-ttcgctttcacgttcttagttactacactctcgaaacctgctgcttctcta-5'	Sense
cCesA1-me02	3'-aacgttctgtgagatttctgtgtgaaggcCtgGatGaggagcaatattgc-5'	Anti-sense
cCesA1-09	3'-gaaatgttcagctccttctgagag-5'	Sense
cCesA1-3R03	3'-aatgtaaaaatactttattcatggcaga-5'	Anti-sense
Introducing a premature stop codon into CesA1 cDNA		
cCesA1-d01	3'-tcAgtaccagcaatagataacgacgaca-5'	Anti-sense
MO	Sequence	Type**
<u>CesA1</u>		
cesa1 start	5'-tgggctttttgattcctccatttcg-3'	TB
5-mis cesa1 start	5'-tgcgcattttgattgctcgattacg-3'	Control for TB
cesa1 e2i2	5'-taaaattgagtttaccttgaattag-3'	SB
5-mis cesa1 e2i2	5'-tataattcagtttagcttcaataag-3'	Control for SB
<u>CesA2</u>		
cesa2 i2e3	3'-gatccttaatttttggtagaatta-5'	SB
cesa2 e3i3	3'-tacgtggtcagtgatttaccttgac-5'	SB
5-mis cesa2 i2e3	3'-tataattcagtttagcttcaataag-5'	Control for SB
5-mis cesa2 e3i3	3'-taggtgctcagtcatttagcttcac-5'	Control for SB
*Nucleotides in upper case are introduced point mutations. **TB, translation-blocking; SB, splice-blocking.		

**Table S2. Protein sequences used for phylogenetic analyses in Fig. 2A**

Abbreviation	Species and protein name	Accession numbers
Hsa HS	<i>Homo sapiens</i> hyaluronan synthase	AAC50706
Mmu HS	<i>Mus musculus</i> hyaluronan synthase	BAA11654.1
Afu CS	<i>Aspergillus fumigatus</i> chitin synthase	CAA63928.1
Ani	<i>Aspergillus niger</i> hypothetical protein An02g05730	XP_001399709.1
Cac	<i>Clostridium acetobutylicum</i> ATCC 824 cell wall biosynthesis glycosyltransferase	NP_348113.1
Cdi	<i>Clostridium difficile</i> QCD-37x79 cellulose synthase	ZP_02726444.1
Pcv	<i>Paramecium bursaria</i> <i>Chlorella virus 1</i> similar to <i>Acetobacter</i> cellulose synthase	AAC96840.1
Sgr	<i>Streptomyces griseus</i> subsp. <i>griseus</i> NBRC 13350 putative glycosyl transferase	YP_001826277.1
Cin	<i>Ciona intestinalis</i> cellulose synthase	BAD10864
Csa	<i>Ciona savignyi</i> cellulose synthase	AAR89623.1
Agr	<i>Agrobacterium</i> sp. ATCC 31749 putative beta 1, 3 glucan synthase	AAD20440.2
Mae	<i>Microcystis aeruginosa</i> PCC 7806 unnamed protein	CAO87270.1
Sme	<i>Sinorhizobium meliloti</i> 1021 putative cellulose synthase	NP_436917.1
Mpo	<i>Methylobacterium populi</i> BJ001 cellulose synthase	YP_001923985.1
Msp	<i>Methylobacterium</i> sp. 4-46 cellulose synthase	YP_001770326.1
Ret	<i>Rhizobium etli</i> CIAT 652 putative cellulose synthase	YP_001985895.1
Rle	<i>Rhizobium leguminosarum</i> bv. <i>trifolii</i> WSM2304 cellulose synthase	ZP_02858951.1
Atu	<i>Agrobacterium tumefaciens</i> cellulose synthase	AAC41436.1
Axy1	<i>Gluconacetobacter xylinus</i> cellulose synthase	AAA85264.1
Axy2	<i>Gluconacetobacter xylinus</i> cellulose synthase	O82859.1
Axy3	<i>Gluconacetobacter xylinus</i> cellulose synthase 1	Q9WX61.1
Ppu1	<i>Pseudomonas putida</i> F1 cellulose synthase	YP_001267459.1
Ppu2	<i>Pseudomonas putida</i> F1 cellulose synthase	NP_744779.1
Sen1	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Typhi str. CT18 cellulose synthase	NP_458301.1
Sen2	<i>Salmonella enterica</i> subsp. <i>enterica</i> serovar Weltevreden str. HI_N05-537 cellulose synthase	ZP_02830558.1
Eco1	<i>Escherichia coli</i> E110019 cellulose synthase	ZP_03049480.1
Eco2	<i>Escherichia coli</i> SMS-3-5 cellulose synthase	YP_001745808.1
Eco3	<i>Escherichia coli</i> O157:H7 str. Sakai putative cellulose synthase	BAB37836.1
Ddi	<i>Dictyostelium discoideum</i> cellulose synthase	AAF00200.1
Npu	<i>Nostoc punctiforme</i> PCC 73102 cellulose synthase	YP_001865112.1
Sco	<i>Streptomyces coelicolor</i> A3(2) glycosyl transferase	NP_627065.1
Ssv	<i>Streptomyces svceus</i> ATCC 29083 glycosyl transferase	YP_002208804.1
Ath CesA1-10	<i>Arabidopsis thaliana</i> cellulose synthase 1-10	NP_194967.1 NP_195645.1 NP_196136.1 NP_199216.2 NP_196549.1 NP_201279.1 Q9SWW6.1 NP_567564.1 NP_179768.1 NP_180124.1
Ath CsID1-5	<i>Arabidopsis thaliana</i> cellulose synthase-like protein D1-5	O49323.1 Q9LFL0.1 Q9M9M4.1 Q9SZL9.1 Q9SRW9.1
Osa CesA1-9	<i>Oryza sativa</i> Japonica Group cellulose synthase 1-9	Q6AT26.1 Q84M43.1 Q69V23.1 Q5JN63.1 Q851L8.1 Q6YVM4.1 Q9AV71.1 Q84ZN6.1 Q69P51.1
Osa CsID1-5	<i>Oryza sativa</i> Japonica Group cellulose synthase-like protein D1-5	Q8W3F9.1 Q9LHZ7.1 Q7EZW6.2 Q2QN56.1 Q5Z6E5.1
Zma CesA1-12	<i>Zea mays</i> cellulose synthase 1-12	AAF89961.1 AAF89962.1 AAF89963.1 AAF89964.1 AAF89965.1 AAF89966.1 AAF89967.1 AAF89968.1 AAF89969.1 AAR23310.1 AAR23311.1 AAR23312.1
Nal CesA	<i>Nicotiana glauca</i> cellulose synthase	AAK49454.1
Nal CsID	<i>Nicotiana glauca</i> cellulose synthase D-like protein	AAK49455.1

**Table S3. Protein sequences used for phylogenetic analyses in Fig. 2B**

Abbreviation	Species and protein name	Accession number
Hko	<i>Hypocrea koningii</i> cbh2	ABG48766.1
Tsp	<i>Trichoderma</i> sp. XST1 cellobiohydrolase II	ACH96126.1
Hje	<i>Hypocrea jecorina</i> cellobiohydrolase II	AAG39980.1
Tvi	<i>Trichoderma viride</i> cellobiohydrolase II	AAQ76094.1
Tpa	<i>Trichoderma parceramosum</i> cellobiohydrolase II precursor	AAU05379.2
Aor	<i>Aspergillus oryzae</i> RIB40 hypothetical protein	XP_001825360.1
Nfi	<i>Neosartorya fischeri</i> NRRL 181 cellobiohydrolase, putative	XP_001264772.1
Ate	<i>Aspergillus terreus</i> NIH2624 exoglucanase 2 precursor	XP_001210279.1
Ani1	<i>Aspergillus niger</i> hypothetical protein An08g01760	XP_001392295.1
Ani2	<i>Aspergillus nidulans</i> FGSC A4 hypothetical protein AN1273.2	XP_658877.1
Eni	<i>Emericella nidulans</i> beta-1,4-glucan-cellobiohydrolyase	ABF50873.1
Fox	<i>Fusarium oxysporum</i> putative endoglucanase type B	P46236.1
Gze	<i>Gibberella zeae</i> glycoside hydrolase 6-like protein	AAQ72468.1
Cth	<i>Chaetomium thermophilum</i> cellobiohydrolase family 6	AAV88915.1
Hin	<i>Humicola Insolens</i> chain A, cellobiohydrolase li	1BVW_A
Cgl	<i>Chaetomium globosum</i> CBS 148.51 hypothetical protein CHGG_10762	XP_001226029.1
Pan	<i>Podospora anserina</i> unnamed protein	XP_001903170.1
Mgr	<i>Magnaporthe grisea</i> 70-15 hypothetical protein MGG_05520	XP_360146.1
Acl	<i>Aspergillus clavatus</i> NRRL 1 cellobiohydrolase, putative	XP_001273717.1
Ace	<i>Acremonium cellulolyticus</i> Y-94 cellobiohydrolase II	BAA74458.1
Pfu	<i>Penicillium funiculosum</i> cellulase	ACH91035.1
Bfu	<i>Botryotinia fuckeliana</i> B05.10 hypothetical protein	XP_001552807.1
Tem	<i>Talaromyces emersonii</i> cellobiohydrolase II	JC7931
Afu	<i>Aspergillus fumigatus</i> Af293 cellobiohydrolase	XP_748511.1
Pno	<i>Phaeosphaeria nodorum</i> SN15 hypothetical protein	XP_001796781.1
Ila	<i>Irpep lacteus</i> cellobiohydrolase II	BAG48183.1
Pch	<i>Phanerochaete chrysosporium</i> exocellobiohydrolase CBHII	AAB32942.1
Par	<i>Polyporus arcularius</i> cellobiohydrolase II	BAF80327.1
Led	<i>Lentinula edodes</i> cellobiohydrolase	AAK28357.1
Abi	<i>Agaricus bisporus</i> cellobiohydrolase	AAA50608.1
Vvo	<i>Volvariella volvacea</i> cellobiohydrolase II-I	AAT64008.1
Lsa	<i>Lentinus sajor-caju</i> cellobiohydrolase II	AAL15038.1
Cok	<i>Coprinopsis cinerea</i> okayama7#130 hypothetical protein CC1G_01107	XP_001833045.1
Ptr	<i>Pyrenophora tritici-repentis</i> Pt-1C-BFP exoglucanase 3 precursor	XP_001934153.1
Psp	<i>Piromyces</i> sp. E2 cellobiohydrolase Cel6B	AAP30749.1
Sce	<i>Sorangium cellulosum</i> cellulose 1,4-beta-cellobiosidase	YP_001618727.1
Osp	<i>Orpinomyces</i> sp. PC-2 cellulase A	AAB92678.1
Prh	<i>Piromyces rhizinflatus</i> 1,4-beta-D-glucan-cellobiohydrolase	ABY52799.1
Cin	<i>Ciona intestinalis</i> cellulose synthase	BAD10864
Csa	<i>Ciona savignyi</i> cellulose synthase	AAR89623.1
Nsp	<i>Nocardioides</i> sp. JS614 cellulase	YP_925799.1
Sav	<i>Streptomyces avermitilis</i> MA-4680 endo-1,4-beta-glucanase	NP_828072.1
Ssp	<i>Streptomyces</i> sp cellulase precursor	AAA26776.1
Mce	<i>Micromonospora cellulolyticum</i> endo-beta-1,4-glucanase McenA	AAC60491.1
Tbi	<i>Thermobispora bispore</i> endoglucanase A	P26414.1
Pst	<i>Providencia stuartii</i> ATCC 25827 hypothetical protein PROSTU_03046	ZP_02961057.1
Mbo	<i>Mycobacterium bovis</i> AF2122/97 endo-1,4-beta-glucanase	NP_853732.1
Mtu	<i>Mycobacterium tuberculosis</i> endoglucanase Cel6	1UP2_A
Mma	<i>Mycobacterium marinum</i> M cellobiohydrolase a	YP_001848433.1
Mav	<i>Mycobacterium avium</i> 104 endoglucanase A	YP_879613.1
Msm	<i>Mycobacterium smegmatis</i> str. MC2 155 endoglucanase A	YP_890960.1
Mab	<i>Mycobacterium abscessus</i> cellulase CelA	YP_001705499.1
Msp	<i>Mycobacterium</i> sp. MCS cellulase	YP_642458.1
Mgi	<i>Mycobacterium gilvum</i> PYR-GCK cellulase	YP_001132274.1
Mva	<i>Mycobacterium vanbaalenii</i> PYR-1 cellulase	YP_956608.1
Ppa	<i>Plesiocystis pacifica</i> SIR-1 cellulase	ZP_01907667.1
Sco	<i>Streptomyces coelicolor</i> A3(2) secreted endoglucanase	NP_627067.1
Ssv	<i>Streptomyces svceus</i> ATCC 29083 secreted endoglucanase	YP_002208806.1
Spr	<i>Streptomyces pristinaespiralis</i> ATCC 25486 secreted endoglucanase	YP_002196808.1
Sna	<i>Streptomyces nanchangensis</i> NanG8	AAP42880.1
Fal	<i>Frankia alni</i> ACN14a endoglucanase 1 precursor	YP_715139.1
Fsp	<i>Frankia</i> sp. EAN1pec cellulase	YP_001511422.1
Cfi	<i>Cellulomonas fimi</i> endoglucanase A	P07984.1
Kra	<i>Kineococcus radiotolerans</i> SR530216 cellulase	YP_001468205.1
Sgr	<i>Streptomyces griseus</i> subsp. griseus NBRC 13350 putative cellulase	YP_001823957.1
Sha	<i>Streptomyces halstedii</i> endoglucanase 1	P33682.2
Mxa	<i>Myxococcus xanthus</i> beta-1,4-glycanase	CAA54086.1
Ser	<i>Saccharopolyspora erythraea</i> NRRL 2338 endoglucanase	YP_001108158.1