BMC Plant Biology

RESEARCH

Open Access

Transcriptional dynamics of gametogenesis in the green seaweed *Ulva mutabilis* identifies an RWP-RK transcription factor linked to reproduction

Xiaojie Liu[1], Jonas Blomme[1,2,3], Kenny A. Bogaert[1], Sofie D'hondt[1], Thomas Wichard[4], Dieter Deforce[5], Filip Van Nieuwerburgh[5] and Olivier De Clerck[1*]

Abstract

Background: The molecular mechanism underlying sexual reproduction in land plants is well understood in model plants and is a target for crop improvement. However, unlike land plants, the genetic basis involved in triggering reproduction and gamete formation remains elusive in most seaweeds, which are increasingly viewed as an alternative source of functional food and feedstock for energy applications.

Results: Gametogenesis of *Ulva mutabilis*, a model organism for green seaweeds, was studied. We analyzed transcriptome dynamics at different time points during gametogenesis following induction of reproduction by fragmentation and removal of sporulation inhibitors. Analyses demonstrated that 45% of the genes in the genome were differentially expressed during gametogenesis. We identified several transcription factors that potentially play a key role in the early gametogenesis of *Ulva* given the function of their homologs in higher plants and microalgae. In particular, the detailed expression pattern of an evolutionarily conserved transcription factor containing an RWP-RK domain suggested a key role during *Ulva* gametogenesis.

Conclusions: Transcriptomic analyses of gametogenesis in the green seaweed *Ulva* highlight the importance of a conserved RWP-RK transcription factor in the induction of sexual reproduction. The identification of putative master regulators of gametogenesis provides a starting point for further functional characterization.

Keywords: *Ulva*, Green seaweeds, Reproduction, Gametogenesis, Transcriptome, RWP-RK transcription factor

Background

Sexual reproduction is one of the most important and conserved processes in eukaryotes. In essence, sexual reproduction encompasses the fusion of two haploid gametes of the opposite sex to form a diploid zygote. Zygote formation is either followed by meiosis to restore the haploid state (zygotic meiosis) or by the development of a diploid life stage through mitotic divisions (gametic meiosis) [1]. Molecular genetic studies of reproduction in land plants resulted in an excellent understanding of the fundamental biological principles and genes involved. However, unlike land plants, the molecular mechanisms involved in the onset of reproduction, the formation of gametes (gametogenesis), fusion and meiosis remain elusive for most algae and seaweeds. In the green algal lineage, the molecular mechanisms underlying gametogenesis and fusion have been extensively studied in *Chlamydomonas reinhardtii*, a unicellular freshwater

*Correspondence: Olivier.declerck@ugent.be
[1] Phycology Research Group and Center for Molecular Phylogenetics and Evolution, Ghent University, Ghent, Belgium
Full list of author information is available at the end of the article

green alga [2–5], but to which extent key regulatory genes are conserved across green algae or even land plants remains unclear.

Here, we study the expression of genes during the various stages of gamete formation in *Ulva*, an emerging model for green seaweeds [6–8]. *Ulva* species are abundant in coastal benthic communities around the world and form a potential source of biomass that can be used for food, feed, biofuel and nutraceuticals [8–13]. Given its commercial value, reproducibly controlling the life cycle of *Ulva* is important in the development of this crop. Under natural conditions, large parts of the *Ulva* thallus are converted into gametangia or sporangia in mature individuals and the entire cell content is converted into reproductive cells. Numerous studies have investigated the effect of different endogenous (e.g., production of sporulation and swarming inhibitors) and environmental factors (e.g. light, temperature or nutrient conditions) on gametogenesis in different *Ulva* species [14–18]. Although we are still far removed from a complete understanding of the interplay of the various factors that appear to play a role, thallus fragmentation has been recognized as the most efficient way to induce gametogenesis [15, 19–22].

The ability to efficiently induce gametogenesis in laboratory cultures of *Ulva mutabilis* provides an elegant time course to study both phenotypic and molecular factors involved in this process [16]. Two sporulation inhibitors, SI-1 and SI-2 control the onset of gametogenesis of *Ulva mutabilis*. SI-1 is a species-specific glycoprotein secreted in the culture medium, which depletes with maturation of the thallus. SI-2 is a nonprotein molecule existing within the inner space between the two blade cell layers. Thallus fragmentation, followed by a washing step, whereby spent culture medium is replaced, results in the total removal of sporulation inhibitors and induces gametogenesis [15]. Following fragmentation and washing, SI-1 and SI-2 are removed and vegetative cells enter a determination phase (~ 0-26 h, Fig. 1) in which a 'swarming inhibitor' (SWI) is produced [16]. SWI functions as a mechanism to synchronize gamete release. The start of the differentiation phase coincides with the completion of the S-phase and the G2-phase (~ 36 h) in the first cell cycle (Fig. 1). At the end of the determination phase, the cells enter the next S-phase and become irreversibly committed to gametangium differentiation. The differentiation phase (~ 26-70 h) is characterized by a reorientation of the chloroplast, followed by four consecutive cell divisions forming sixteen progametes per cell and their maturation [15, 16]. Mature gametes are eventually released during the swarming phase if the

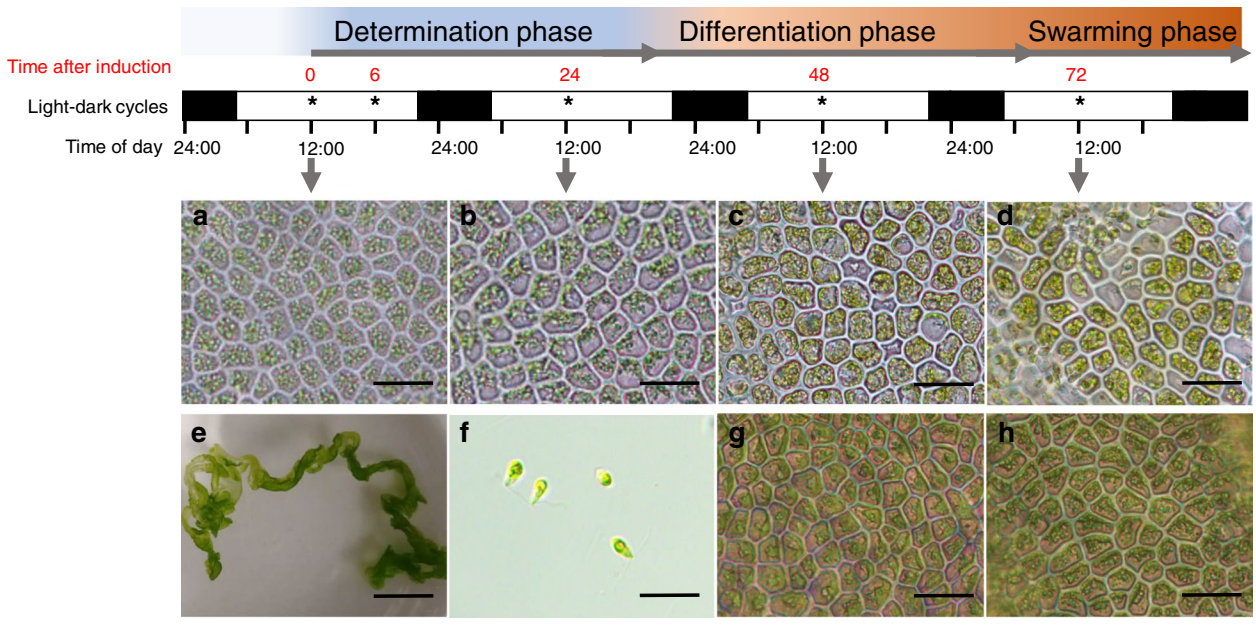

Fig. 1 Time course of gametogenesis in *Ulva mutabilis*. "determination phase" and "differentiation phase" described previously by Wichard and Oertel [16]. The "swarming phase" is the time when the change of medium can release the gametes. Asterisks indicate the sampled time points for RNA-seq. (**a**) vegetative cells before induction, (**b**) gametogenesis induction after 24 h with chloroplast reorientation, (**c**) gametangia after 48 h, (**d**) reproductive cells with fully formed gametes and gametes can be released upon removal of the SWI, (**e**) *Ulva* thallus before induction, (**f**) released gametes, (**g**, **h**) induced *Ulva* tissue without refreshment of growth medium after 48 h (**g**) and 72 h (**h**), demonstrating that no gametes were formed as long as sporulation inhibitors are not removed (control experiment). Scale bars: 2 cm (**e**) or 20 µm (others)

SWI declines in concentration on the third day after induction, following a further light stimulus [16, 23].

Given that *Ulva* gametogenesis can be subdivided into discrete phases, the gene regulatory networks involved in the different stages of gametogenesis should be distinct. As a first step toward understanding the molecular mechanisms of gametogenesis, we established the transcriptional structure for *Ulva mutabilis* gametogenesis in this research. We make use of differential removal of SI's to separate the effect of fragmentation on transcription from gametogenesis. Transcription patterns are compared with those of other green algae and land plants.

Results

To dissect the molecular mechanisms underlying *Ulva* gametogenesis, the global gene expression levels of *U. mutabilis* were measured by RNA-seq as a function of time. We sampled five-time points: 0 h, 6 h, 24 h, 48 h and 72 h after induction of gametogenesis (Fig. 1) under standard induction conditions (CW) and two-time points (6 h and 24 h after induction of gametogenesis) that omitted the washing step (CNW), with three biological replicates for each time point and condition. RNA-sequencing resulted in 215×10^6 reads, > 70% mapped to the *U. mutabilis* genome. The total number of mapped reads for each sample and the raw reads counts for each gene were listed in Table S2 and Table S3, respectively. Transcript profiles were highly reproducible among the three biological replicates at each time point based on the Pearson R^2 test and PCA analysis (Table S4, Fig. S1).

Gene expression patterns during gametogenesis

In total, 8296 distinct genes (62.2% of annotated genes) were expressed during gametogenesis, with relatively low variation in the total number of expressed genes between the time points: ranging between 7146 (0 h) and 7949 expressed genes (72 h). We identified 6056 differentially expressed genes (DEGs) between any two of the five time points (0 h, 6 h, 24 h, 48 h, 72 h) during the gametogenesis process, representing 45% of genes of the annotated genome. Each time point could be characterized by a set of genes describing the various steps of gametogenesis (Fig. 1). Hierarchical clustering grouped genes in 5 clusters for further analyses based on the stabilization of the SSE and the average silhouette width values (Fig. S2).

The identified clusters were characterized by nearly unique sets of enriched GO terms and reflected the identified phases of the gametogenesis process well (Fig. 2). Cluster 1 grouped genes expressed in the vegetative phase with decreasing expression levels during gametogenesis. Cluster 2 was the only cluster with high expression during the determination phase but low expression during the differentiation and swarming phase. Cluster 3 and 4 grouped genes with high expression during the differentiation phase. The main difference between both clusters consisted of genes of cluster 3 being down-regulated during the swarming phase. In contrast, genes in clusters 4 were characterized by a high expression level during the swarming phase also. Cluster 5 contains the least number of genes showing high expression levels only in the swarming phase.

DEGs related to photosynthesis and carboxylic acid biosynthesis were upregulated during the determination phase. Next, genes involved in DNA replication and microtubule-based movements have a high expression during the early differentiation process of gametogenesis and gamete formation. However, the different enriched GO terms between the two clusters apply primarily to DNA replication which indicates that DNA replication is completed in the early phase of the differentiation process before gametes formation starts. Moreover, the final swarming phase was mainly relevant to the pigment synthesis and cellular amino acid biosynthesis which was the final step for the gamete's formation.

Initiation of gametogenesis versus response to fragmentation

By adding an additional control group whereby thalli were fragmented but the sporulation inhibitors were not removed (CNW, chopping but no washing), we aimed to disentangle the effect of fragmentation on gene expression from the induction of gametogenesis. By analyzing DEGs between CW and CNW groups at 6 h and 24 h, we identified 901 and 1137 significantly DEGs (FDR < 0.05, log2FC > 1 or < −1) at 6 h and 24 h, respectively. DEGs included 493 upregulated and 408 downregulated genes at 6 h and 652 upregulated and 485 downregulated genes at 24 h compared to the CNW group. Nearly half of the DEGs at 6 h have the same expression profile at 24 h compared to the CNW group (Fig. 3). Of the 493 upregulated and 408 downregulated genes at 6 h, 206 and 219 genes showed the same expression profiles at 24 h, indicating that DEGs with a function in the determination phase tended to have continuous expression profiles during early gametogenesis. The enriched GO terms for the upregulated DEGs at 6 h are mainly related to the photosynthesis, ATP biosynthetic process and oxidation-reduction processes (Fig. 4), in accordance with cluster 2 enrichment. In contrast, GO terms significantly enriched at 24 h are mainly related to the microtubule-based movement, DNA replication and cilium organization, suggesting that the DNA replication initiation for the gametogenesis starts from 24 h, consistent with the expression pattern of cluster 3 (Fig. 2). The over-represented GO terms for the downregulated genes were related to the response to oxidative stress and

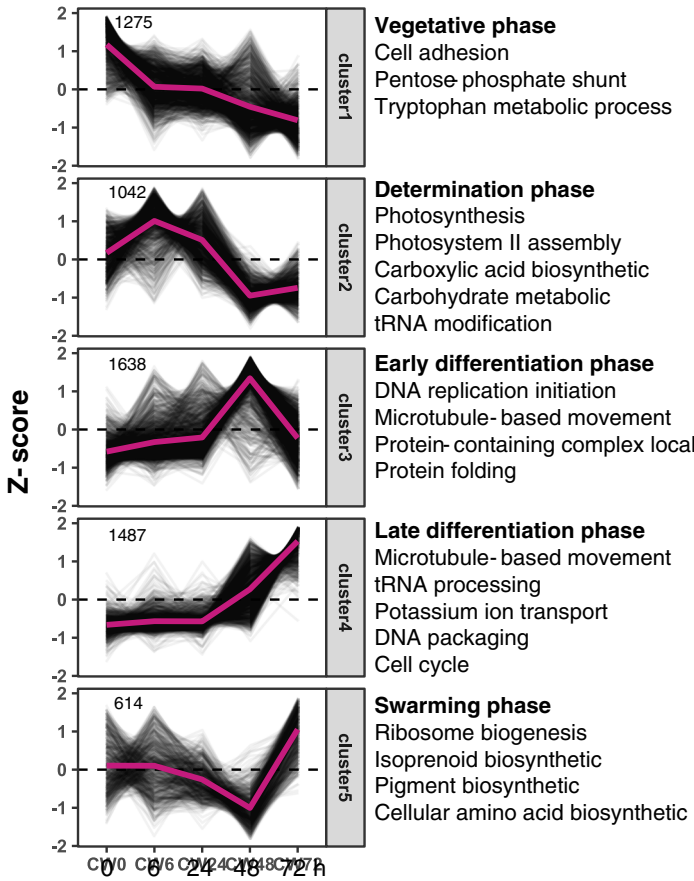

Fig. 2 Clusters of differentially expressed genes (DEG) identified during the different phases of gametogenesis at 0 h, 6 h, 24 h, 48 h and 72 h. Clusters are presented according to their maximal expression in the different phases of gametogenesis. A selection of the most relevant overrepresented GO terms ($p < 0.01$ and summarized by REVIGO) is presented for each cluster. We refer to Table S5 for the full list of enriched GO terms

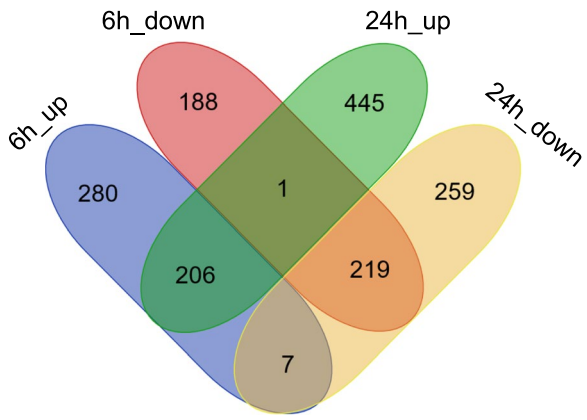

Fig. 3 Differentially expressed genes (DEGs) number of fully induced thalli (CW) compared to thalli in the presence of the sporulation inhibitors (CNW). The Venn diagram depicts the number of up- and down-regulated DEGs at 6 h and 24 h

oxidation-reduction process both at 6 h and 24 h after induction of gametogenesis.

To identify the possible key initiator for *Ulva* gametogenesis, we analyzed the differentially expressed TFs at 6 h, which signifies the early responder to the washing treatment exclusively by comparing the chopping and washing (CW) group with the chopping without washing (CNW) group. We identified 12 upregulated and 6 downregulated transcription factors (TFs) at 6 h (Table 1). Several homologs of these TFs, as represented by the top BLAST hits in the *Ulva*-PLAZA database, are involved in gamete formation in *Arabidopsis* or *Chlamydomonas*.

We acquired the homologs of the *Arabidopsis* and *Chlamydomonas* of the upregulated TFs at 6 h by BLAST. With the reported function of the homologs, we tried to find the conserved TFs function in plants gametogenesis. Notably, an extensive studied TF containing RWP-RK domain aroused our interests for further characterization. The RWP-RK gene in *Chlamydomonas*,

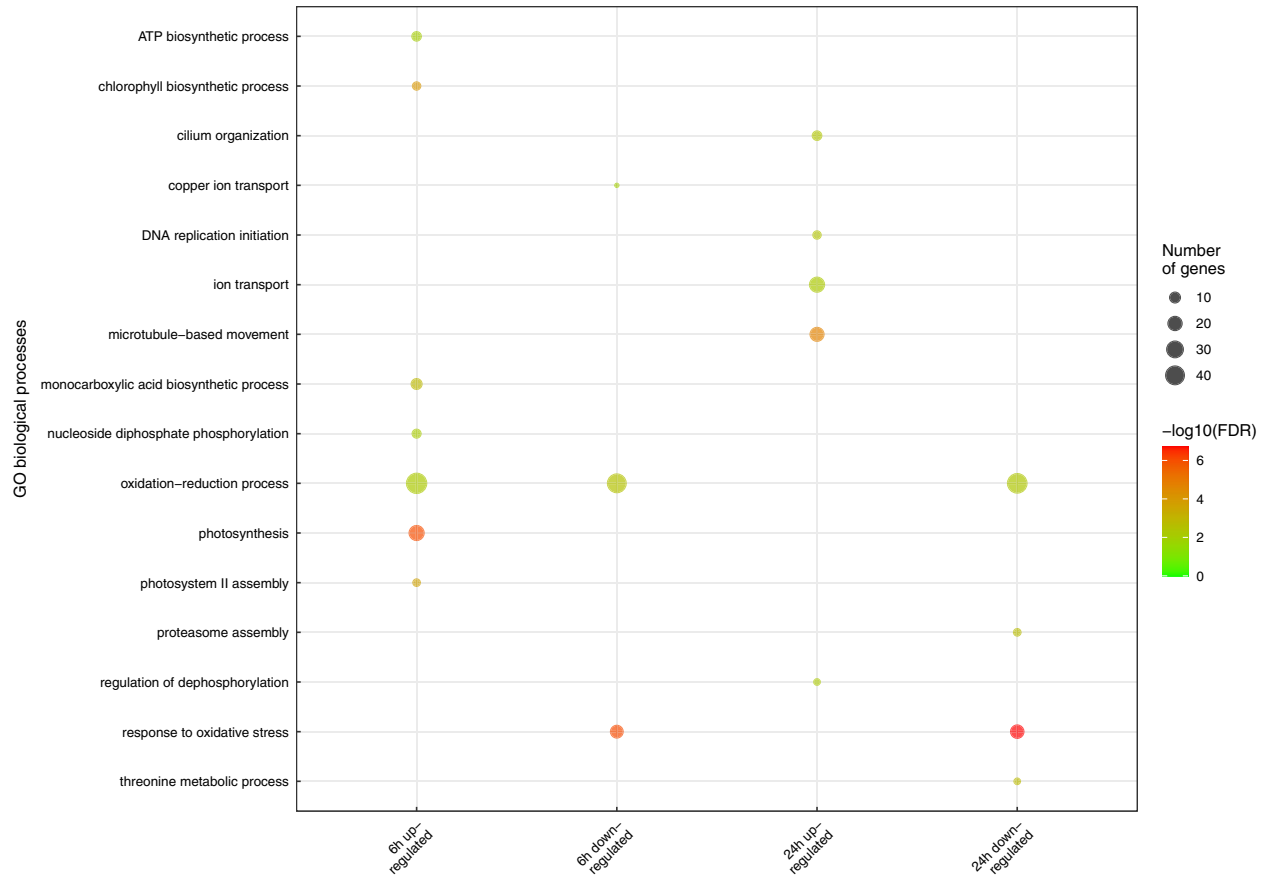

Fig. 4 Enriched GO terms associated with DEGs of fully induced thalli (CW) compared to thalli in the presence of the sporulation inhibitors (CNW) at 6 h and 24 h, respectively

which is minus dominance (*MID*), is responsible for switching on the minus-programme and switching off the plus-programme in gamete differentiation [2]. Studies in *Arabidopsis* indicate that RWP-RK TFs control cell differentiation during female gametophyte development [24]. We investigated the expression pattern of the RWP-RK TF (*UMSL057_0048*) in more detail using qRT-PCR analysis, sampling tissue fragments undergoing gametogenesis every 2 h. At each timepoint the expression was compared against a control treatment not undergoing gametogenesis. *UMSL057_0048* is upregulated in the early determination phase and throughout the determination and differentiation phase, with maximum expression levels at 6 h (Fig. 5a). More specifically, *UMSL057_0048* shows a quick response to the chopping and washing treatment with slight upregulation at 2 h after induction and significantly upregulation at 4 h.

In an additional experiment, we contrasted the expression at 6 h of *UMSL057_0048* to fragmented thalli from which the SI were not removed (CNW) and unfragmented thalli that were washed to remove the SI in the medium and hence were partially induced (NCW). qRT-PCR results indicate *UMSL057_0048* is not upregulated at 6 h in treatments where gametogenesis is not induced (CNW and NCNW). In contrast, in both treatments where gametogenesis is induced (CW and NCW), *UMSL057_0048* is significantly upregulated (Fig. 5b). The *Ulva* slender genome contains one additional gene with an RWP-RK domain, *UMSL048_0014*. The latter, however, is not expressed (CPM value ~0) during gametogenesis.

Discussion

The transition from a vegetative to a reproductive state is a key step in the life cycle of *Ulva*. Analysis of the molecular mechanisms underpinning this process is pivotal for efficient seeding technologies and putative genetic improvement for *Ulva* as a commercial crop species. In this study, RNA-seq was applied to analyze the dynamic changes of gene expression of *U. mutabilis* and to identify key TFs involved in the initiation of *Ulva* gametogenesis. Transcriptional analysis showed that nearly half

Table 1 Differential expressed TFs of induced (CW) versus non-induced (CNW) thalli at 6 h. Log2FC > 1: upregulated, Log2FC<-1: down-regulated. The homologs of the TFs of *A. thaliana* and *C. reinhardtii* are listed in the "AtGID" and "CrID" columns respectively. Gene function annotations are summarized from TAIR or Phytozome databases, respectively

GeneID	AtGID	CrID	Family/Domain	Log2FC	Gene function annotation
UMSL012_0085	*AT4G18770*	*CR03G00530*	MYB	2.55	MYB98: expressed in the synergid cells, affect the female gametophyte[a]
UMSL129_0009	*N/A*	*N/A*	MYB	1.85	N/A
UMSL039_0026	*AT2G43770*	*CR12G01870*	DEAD	2.21	FAP 52: flagellar associated protein[b]
UMSL054_0061	*N/A*	*N/A*	DEAD	2.81	N/A
UMSL003_0532	*AT5G58080*	*CR02G02460*	GARP_G2-like	1.08	ARR18: member of response regulator[a]
UMSL103_0003	*AT5G43990*	*N/A*	SET	1.53	SUVR2: regulation of eukaryotic gene expression and chromatin structure[a]
UMSL008_0028	*AT5G57390*	*CR01G02250*	AP2/EREBP	1.49	AIL5: involved in germination and seedling growth[a]
UMSL032_0028	*AT2G01830*	*CR06G00550*	DEAD	1.08	AHK4: cytokinin-binding receptor that transduces cytokinin signals[a]
UMSL057_0048	*AT5G53040*	*CR03G09530*	RWP-RK	1.62	GR: promotes zygote elongation and basal cell fates[a]
UMSL024_0043	*AT3G28730*	*CR12G13130*	HMG	4.77	ATHMG: binds to the promoter of a repressor of flowering: FLC[a]
UMSL033_0026	*AT2G25170*	*CR03G03890*	SWI/SNF_SNF2	1.51	CHD3: involved in post-germination repression of embryonic development[a]
UMSL003_0659	*AT3G27730*	*CR10G06300*	DEAD	1.52	MER3: DNA helicase required for interference-sensitive meiotic crossover[a]
UMSL151_0016	*AT3G16770*	*CR16G05640*	AP2	−1.32	ATEBP: ethylene response factor[a]
UMSL085_0029	*AT2G17520*	*CR08G03590*	zf-CCCH	−1.27	ATIRE1–2: involved in the regulation of the ER stress responsive genes[a]
UMSL051_0112	*AT5G41370*	*CR06G12570*	DEAD	−1.14	ATXPB1: encodes a DNA repair protein[a]
UMSL003_0378	*AT1G19270*	*CR10G01560*	DEAD	−1.11	DA1: controls the initiation of axillary meristems[a]
UMSL053_0033	*AT3G58680*	*CR02G09000*	MBF1	−1.03	FAP280: flagellar associated protein[b]
UMSL011_0269	*AT1G01040*	*N/A*	DEAD	−1.03	ABNORMAL SUSPENSOR 1: encodes a Dicer homolog[a]

[a] Access from TAIR database (https://www.arabidopsis.org/)

[b] Access from Phytozome database (https://phytozome.jgi.doe.gov/pz/portal.html)

(45.4%) of the protein-coding genes predicted in the draft genome are differentially expressed between at least two timepoints during gametogenesis. This large-scale reprogramming of the transcriptome likely reflects the full transformation of a vegetative cell into 16 gametes [16].

Dynamic changes in gene regulation during gametogenesis

GO enrichment analysis of clusters of DEGs sheds light on the most enriched up-or downregulated biological process during each of the three phases that characterize gametogenesis in *Ulva*. Cluster 1 contains DEGs that are down-regulated during the determination phase and genes in this cluster are enriched in the cell adhesion and pentose-phosphate shunt activity. For cluster 2 which is upregulated during the determination phase, there is a high expression of genes that are associated with photosynthesis, carboxylic acid biosynthesis and ribose phosphate metabolism. This correlates with significant increases in maltose during gametogenesis of *U. prolifera* at 12 h and 24 h is, which are approximately 12-fold change upregulated compared to the control group [25]. Similarly, the high expression level of photosynthesis-related genes at 24 h and 48 h is linked to the increase in chlorophyll content subsequent to the divisions of the chloroplast, as demonstrated for *Ulva pertusa* during sporulation [26]. All clusters represent a marked changes of gene expression in the transition from the determination phase (24 h) to the differentiation phase (48 h). 36 h after induction is regarded an important checkpoint during gametogenesis, after which blade cells are irreversibly committed to gametangium differentiation [15, 16, 23]. Genes related to DNA replication and microtubule-based movement were significantly upregulated after this irreversible timepoint.

Electron microscope observations indeed demonstrate drastic rearrangements in microtubule morphology between 36 h and 48 h after induction [27]. The microtubule cytoskeleton of somatic cells consists of parallel microtubule bundles arranged mainly in the cortical cytoplasm parallel to the plasmalemma. However, 36 h after induction, the microtubule bundles traverse the cortical cytoplasm converging on a particular area. Then, the microtubules start from a pointed site at the basal part of gametangium and form a basket-like configuration with a circular opening at the top after 48 h [27]. Stratmann et al. [15] analyzed DNA synthesis during gametogenesis and concluded that DNA replication starts from 25 h after induction, which is consistent with our RNA-seq data. The enriched GO terms of microtubule-based movement in cluster 3 and

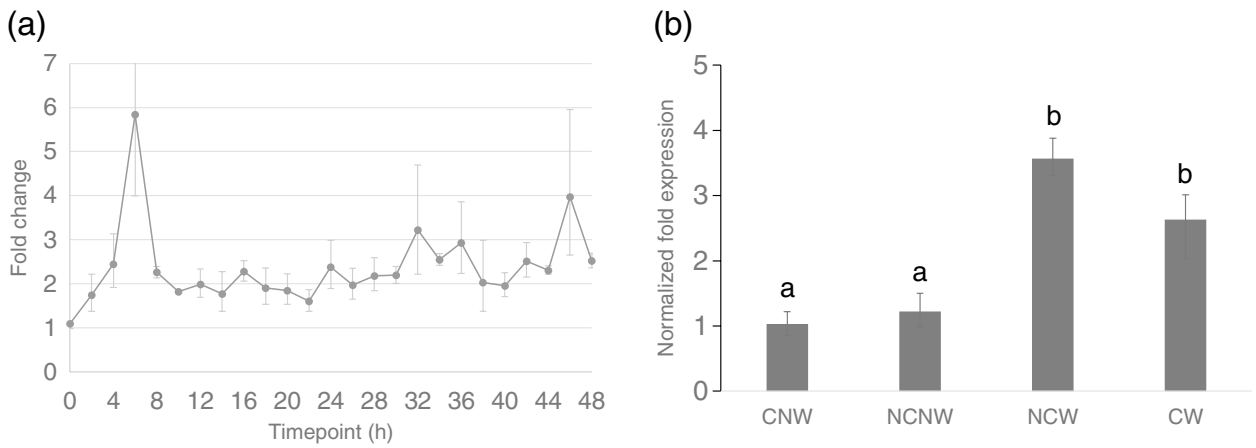

Fig. 5 Expression of RWP-RK transcription factor *UMSL057_0048* during gametogenesis. **a** RT-qPCR analysis of *UMSL057_0048* expression levels compared to the control group (NCNW) in the first 48 h of gametogenesis. Three biological replicates were used for analysis and the gene expression level was calculated using the standard ddCt method. **b** Expression level of *UMSL057_0048* at 6 h in induced (CW and NCW) versus not induced (CNW and NCNW) thalli. Three biological replicates were used for the experiment and the standard ddCt method was applied for analysis. Different lower case letters indicate significant differences among treatments (one-way ANOVA, $P < 0.05$). Error bars represent the standard error ($n = 3$)

4 are the molecular basis of the cytological observation of the cytoskeleton organization. Genes in cluster 5 are upregulated only during the swarming phase and the enriched GO terms mainly involve the synthesis of amino acid and isoprenoid pigment biosynthesis. The isoprenoid pigments biosynthetic pathway catalyzes the synthesis of essential pigments of the photosystem structure in plant cells [28]. We speculate that upregulation of the isoprenoid biosynthetic pathway is related to the reconstruction of the photosynthesis system for the gamete formation.

Transcription factors associated with gametogenesis

Transcription factors play a pivotal role in gene expression networks. TFs were surveyed in the developmental program of gametogenesis that involves extensive cellular morphogenesis and subsequent cell division and proliferation. *Ulva* gametogenesis was initiated by fragmentation and subsequent removal of SI-1 and SI-2. As shown by Stratmann et al. [15], fragmentation of growing gametophytes itself does not initiate gametogenesis and the cells remain in a "vegetative" state after 72 h because of the rapidly excreted SI-1 in the culture medium suppresses gametogenesis during the determination phase. To remove the effect of fragmentation and to identify a putative regulator initiating gametogenesis in *Ulva*, we designed an experiment where thalli are fragmented but the sporulation inhibitors are not removed. With this control group, we identified 12 upregulated and 6 downregulated TFs at 6 h after induction of gametogenesis which constitutes about 4–5% of the TF repertoire of *U. mutabilis*.

We selected the RWP-RK TF (*UMSL057_0048*) for further analysis because studies indicate a conserved role of RWP-RK domain containing TFs in the gametogenesis process [24, 29–31]. Nevertheless, the other TFs identified in this study might be potentially interesting as well. For example, Myb domain protein 98 (MYB98, *UMSL012_0085*) is involved in the regulation of synergid differentiation in angiosperms [32, 33]. Similarly, homologs of high mobility group transcription factor (*UMSL024_0043*) are relevant for the transition of a vegetative to a reproductive phase [34–36]. The homolog of *UMSL024_0043*, which is the most upregulated gene compared to the control group at 6 h, encodes SSRP1, which is a component of the facilitates chromatin transcription (FACT) complex in *Arabidopsis*. FACT is a conserved heterodimeric histone chaperone among eukaryotes and facilitates the expression of *FLC* (FLOWERING LOCUS C) by binding to the promoter of *FLC*, which adjusts the switch from vegetative to reproductive development in *Arabidopsis* [37, 38]. Reduced amounts of SSRP1 result in decreased expression of *FLC,* thus accelerating flowering [38]. Even though, *UMSL024_0043* is upregulated during gametogenesis in *Ulva*, and we did not identify a homolog of *FLC* in the genome. Thus, the specific function and the target genes of *UMSL024_0043* will need to be confirmed experimentally.

One particular TF (encoded by *UMSL057_0048*) containing a RWP-RK protein motif that is important in DNA binding, is upregulated at 6 h. Remarkably, minus

dominance (MID), homolog of RWP-RK TFs in *C. reinhardtii*, has been described to be required for gametogenesis [39]. In addition, similar functions of RWP-RK TFs during gametogenesis were reported in many other plants [24, 40]. RWP-RK transcription factors are found throughout the Viridiplantae. Phylogenetic analysis divided the RWP-RK homologs into 4 subfamilies, RKD(A), RKD(B), RKD(C) and NLP subfamily [29]. Previous studies on algae and higher plants showed that members of NLP subfamily are early regulators of cellular response to N supply [41, 42] and RKD homologs plays evolutionarily conserved roles in germ cell differentiation [2, 29, 31, 41–43]. Unlike other organisms which have multiple RWP-RK transcription factors, only two RWP-RK transcription factors are identified in the *U. mutabilis* slender genome. A recent study on the *Ulva partita* mating type locus structure reported three RWP-RK TFs in the *U. partita* genome, including one RWP-RK TF located in the mating type minus locus with weak homology with the *Chlamydomonas MID* gene [44]. The genome applied in our analysis is mating type plus, so the gene located in the mating type minus is not characterized. UMSL048_0014 contains a conserved GAF domain that is shared in the NLP subfamily. Two members of this subfamily were identified as nitrogen response TFs in *C. reinhardtii* (CreNIT2) and *Volvox carteri* (VcaNIT2) [41, 42]. Interestingly, gametogenesis of *C. reinhardtii* is activated by the N-removal induction, while nitrogen concentration seems to have a positive effect on reproduction in *Ulva* [45]. This may suggest that *UMSL048_0014* may have a function in nitrogen metabolism pathway only.

UMSL057_0048 identified in our study shares a conserved protein domain with RKD(A) family. In our study, *UMSL057_0048* is exclusively upregulated in the gametogenesis group (CW and NCW) and shows upregulation at 2 h after fragmentation and washing treatments. These findings indicate that RKD family member *UMSL057_0048* plays an essential role in *Ulva* gametogenesis.

Compared to *UMSL057_0048*, *UMSL048_0014* contains a truncated GAF domain, which is conserved in the NIN-like proteins (NLPs) subfamily and this subfamily is stated to be the early regulators of cellular response to N supply [29, 41, 42]. In contrast, *UMSL057_0048* contains a conserved protein domain, which Chardin et al. [29] referred to as motif 12, which is conserved in RKD subfamily (Fig. 6) and this subfamily is involved in gametogenesis [2, 29, 41–43].

Conclusions

Our present work provides an important resource to study gametogenesis regulatory framework in a model organism from which insights can be translated into commercially important *Ulva* species. Transcriptomic profiling helps us understand the gene regulatory networks during *Ulva* gametogenesis. The CNW group provides useful information to understand *Ulva*'s response to fragmentation and combined fragmentation and washing factors respectively. The genetic transformation system is available in *U. mutabilis* and further genome editing technology is under development [46], which will allow functional characterization of gametogenesis-related TFs in future experiments. The understanding of the mechanisms of inducing *Ulva* gametogenesis would be helpful to the commercial cultivation of *Ulva*, extending to other seaweeds. In addition, *Ulva* has developed multicellularity independently from land plants, and our study suggests that some TFs play a conserved role in reproduction throughout the green lineage. Therefore, further study of gametogenesis in *Ulva* should provide us with evolutionary insights into control of the sexual reproduction process for the Viridiplantae in general.

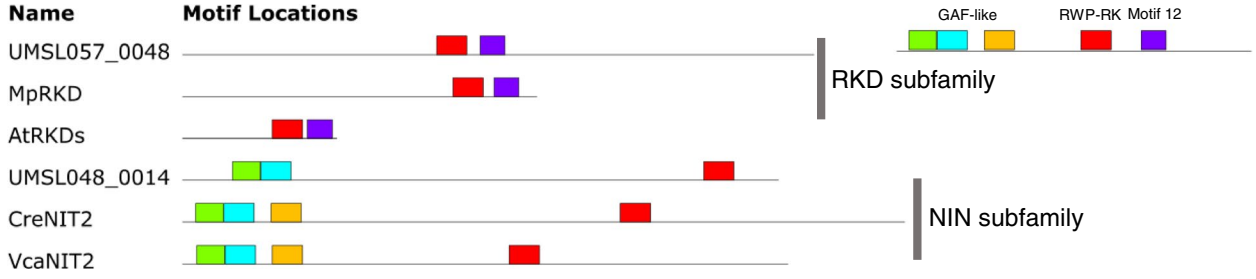

Fig. 6 Comparison of domain organization between RWP-RK proteins. "Motif 12" refers to the same motif identified by Chardin et al. [29]. Model organisms in which the RWP-RK protein were functionally validated were chosen as reference. UMSL, *Ulva mutabilis* slender (+); Mp, *Marchantia polymorpha*; At, *Arabidopsis thaliana*; Cre, *Chlamydomonas reinhardtii*; Vca, *Volvox carteri*; Amino acid sequences and accession numbers are listed in Table S6

Methods

Strains and culture conditions

Haploid gametophytes of *Ulva mutabilis* [strain 'wildtype' (wt-)] were collected initially in southern Portugal by B. Føyn [47]. During the last half century, *U. mutabilis* has been cultivated and permanently propagated under laboratory conditions and has been used as a model system for studying algal development by classical methods of plant physiology and genetics [6, 7]. The strain used in this study was acquired from the lab of Thomas Wichard, Friedrich Schiller University Jena in 2015 [15, 16]. Gametophytes of *U. mutabilis* were raised parthenogenetically from unmated gametes and cultured in transparent plastic boxes containing 2 L Provasoli enriched seawater (PES) medium [48] with the following specific parameters: light intensity, 70 µmol photons $m^{-2} \cdot s^{-1}$; temperature, 18 ± 1 °C; and 17:7 h light:dark cycle in our lab. The medium was completely changed every 2 weeks until fertility (~ 10 weeks). Afterwards, the medium was partially (20%) changed to avoid induction of gametogenesis.

Induction of gametogenesis

Intact mature thalli of *U. mutabilis* were cut into 1–3 mm^2 fragments using a herb chopper. Fragmentation was carried out in the morning. Fragments derived from a single individual were separated into 2 groups. Fragments were washed 3 times for 15 min with autoclaved seawater and transferred to new medium (CW, chopping and washing treatment). The washing step removes sporulation inhibitors, thereby inducing differentiation of the gametangia [15, 16]. Triplicates of 0.2 g fragmented thallus were collected at 0 h, 6 h, 24 h, 48 h and 72 h following gametogenesis induction. Samples at 72 h were collected before the removal of SWI and the gametes still stayed in the gametangia [16]. Control treatments consisted of thalli that were fragmented but not washed and transferred immediately to the original culture medium (CNW, chopping but no washing). The control was sampled at 6 h and 24 h, respectively. The presence of SI-I inhibits gametogenesis and control thalli therefore remain 'vegetative' after 72 h. A third and a fourth treatment, used for expression analyses of RWP-RK transcription factors, consisted of thalli with normal growth that were neither chopped nor washed (NCNW) or not chopped but washed (NCW), respectively. Samples were flash-frozen in liquid nitrogen and stored at − 80 °C for RNA extraction.

RNA extraction and RNAseq

Total RNA was extracted using a CTAB method [49]. The quality and quantity of total RNA was evaluated using a NanoDrop™ 2000c spectrophotometer (Thermo Fisher Scientific) and Bioanalyzer RNA6000 (Agilent Technologies). cDNA libraries were constructed with a Quantseq™ 3′ mRNA-Seq library prep kit (Lexogen) following the manufacturer's instructions. Illumina sequencing was performed using the Illumina Nextseq 500 platform to produce 50 bp single-end reads. RNA reads are available with accession number PRJNA773495. Reads were mapped to the *U. mutabilis* genome [7]. We used the version of the slender strain which has UTR-regions annotated, as opposed to the wildtype genome. Reads were mapped using TopHat ver. 2.1.1 [50] with default parameters. The number of mapped reads was calculated using HTseq [51]. Differentially expressed genes (DEG) were identified between each treatment and control groups using the R package edgeR [52, 53]. The sample variation was estimated by tag-wise dispersion. Raw counts were normalized by CPM (counts per million reads mapped). Genes whose CPM value was > 1 were considered as 'expressed'. A false discovery rate (FDR) of 0.05 and absolute value of log 2-fold change of > 1 were adopted as thresholds for differential expressed genes (DEGs) detection.

Hierarchical clustering analysis

To study the transcriptional dynamics during the gametogenesis process, DEGs between any two of five time points (0 h, 6 h, 24 h, 48 h, 72 h) of the treatment group were identified. The identified DEGs were subjected to a hierarchical clustering analysis by Pearson correlation [54] based on their CPM values in the R (ver. 3.5.0) programming environment using the hclust function. Prior to the analysis, the optimal number of clusters was identified and investigated by performing a SSE merit analysis and an R-based average Silhouette Width analysis [55]. SSE is defined as the sum of the squared distance between each cluster member and its centroid cluster. As the number of clusters increases, the distance between each point and its centroid will be smaller. The optimal number of clusters is suggested when the addition of a new cluster does not significantly decrease the SSE. The silhouette value describes how similar a gene is to its own cluster (cohesion) compared to other clusters (separation).

Functional annotation and GO enrichment

Transcription factor (TF) identification and GO terms for gene models were retrieved from the *Ulva mutabilis* genome annotation file (downloaded from https://bioinformatics.psb.ugent.be/orcae/). The top hit of homologs acquired by the integrative orthology method of the differential expressed TFs of *A. thaliana* and *C. reinhardtii* were obtained using a custom-built PLAZA-*Ulva* version [56]. The functional annotation of these homologs

was acquired on TAIR (https://www.arabidopsis.org/) or Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html). GO enrichment analysis for DEGs was based on a Fisher's exact test implemented in the *TopGO* package in R [57]. The enriched GO terms ($p < 0.01$) in the category "biological function" were summarized using the REVIGO web server, which performed a clustering algorithm that relies on the semantic similarity method [58].

TF identification analysis

To verify RNA-seq results and to obtain more detailed expression pattern for specific TFs during gametogenesis, we performed RT-qPCR analysis of the expression of upregulated TFs. We sampled the treatment group (CW, described above) and thalli with normal growth that were neither chopped nor washed (NCNW) every 2 h in the first 48 h during gametogenesis to obtain detailed expression profiles for TFs identified in the initial RNA-seq experiment. We designed an experiment consisting of four groups: normal vegetative thalli without chopping nor washing (NCNW); chopped thalli that were not washed (CNW); chopped and washed thalli (CW, normal induction); intact thalli that were washed to remove of SI in the medium (NCW). Gametogenesis was partly induced in the latter treatment. Thalli were sampled at 6 h for the four treatments to identify the expression profile of the TFs. RNA was extracted as described above and cDNA was synthesized using an iScript™ cDNA Synthesis Kit (Bio-Rad). PCR was performed using Bio-Rad CFX96 Real-Time PCR systems. Reactions were performed in a final volume of 10 µL containing 5 µL of SYBR Green Master Mix, 0.5 µM each primer, and 10 ng of cDNA under following program: 5 min at 95 °C followed by 35 cycles of 95 °C for 10 s and 55 °C for 30 s. Gene expression levels were calculated using the ddCt method [59]. *UBIQUITIN* (*UMSL012_0173*) and *ELFa* (*UMSL 016_0119*) were stable under a wide range of time points and were taken as reference. Primer sequences are listed in Table S1.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-021-03361-3.

Additional file 1.

Additional file 2.

Acknowledgements
Not applicable.

Authors' contributions
X.L., and O.D.C. designed the study. X.L., T.W. and S.D. performed experiments. D.D. and F.V.N generated sequences. X.L., J.B. and K.B. performed the transcriptional analysis. X.L. wrote the manuscript with contributions from all authors. The author(s) read and approved the final manuscript.

Funding
This work makes use of resources and facilities provided by UGent as part of the Belgian contribution to EMBRC-ERIC (FWO GOH3817N and I001621N) and the Betty Moore Foundation for a SYMBIOSIS grant (nr. 4546891618). X.L. is indebted to the China Scholarship Council (201504910698) and Ghent University Special Research Fund (BOF-16/CHN/023) for a PhD grant and to COST Action FA1406 (Phycomorph) for a Short-Term Scientific Mission to Friedrich Schiller University Jena. J.B. thanks the Research Foundation – Flanders (FWO) for a postdoctoral fellowship (12T3418N) and Ghent University Special Research Grant (BOF20/PDO/016). T.W. was funded by the Deutsche Forschungsgemeinschaft through Grant No. SFB 1127/2 ChemBioSys–239748522.

Availability of data and materials
The raw reads generated in this study are available with accession number PRJNA773495.

Declaration

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Phycology Research Group and Center for Molecular Phylogenetics and Evolution, Ghent University, Ghent, Belgium. [2]Department of Plant Biotechnology and Bioinformatics, Ghent University, 9052 Ghent, Belgium. [3]VIB Center for Plant Systems Biology, 9052 Ghent, Belgium. [4]Institute for Inorganic and Analytical Chemistry, Jena School for Microbial Communication, Friedrich Schiller University Jena, Jena, Germany. [5]Laboratory of Pharmaceutical Biotechnology, Ghent University, 9000 Ghent, Belgium.

Received: 19 August 2021 Accepted: 17 November 2021
Published online: 06 January 2022

References

1. Coelho SM, Peters AF, Charrier B, Roze D, Destombe C, Valero M, et al. Complex life cycles of multicellular eukaryotes: new approaches based on the use of model organisms. Gene. 2007;406(1–2):152–70.
2. Lin H, Goodenough UW. Gametogenesis in the *Chlamydomonas reinhardtii* minus mating type is controlled by two genes, *MID* and *MTD1*. Genetics. 2007;176(2):913–25.
3. Goodenough U, Lin H, Lee JH. Sex determination in *Chlamydomonas*. Semin Cell Dev Biol. 2007;18(3):350–61.
4. Abe J, Kubo T, Takagi Y, Saito T, Miura K, Fukuzawa H, et al. The transcriptional program of synchronous gametogenesis in *Chlamydomonas reinhardtii*. Curr Genet. 2004;46(5):304–15.
5. Abe J, Kubo T, Saito T, Matsuda Y. The regulatory networks of gene expression during the sexual differentiation of *Chlamydomonas reinhardtii*, as analyzed by mutants for gametogenesis. Plant Cell Physiol. 2005;46(2):312–6.
6. Wichard T, Charrier B, Mineur F, Bothwell JH, Clerck OD, Coates JC. The green seaweed *Ulva*: a model system to study morphogenesis. Front Plant Sci. 2015;6:72.
7. De Clerck O, Kao S-M, Bogaert KA, Blomme J, Foflonker F, Kwantes M, et al. Insights into the evolution of multicellularity from the sea lettuce genome. Curr Biol. 2018;28(18):2921–33 e2925.
8. Balar NB, Mantri VA. Insights into life cycle patterns, spore formation, induction of reproduction, biochemical and molecular aspects of sporulation in green algal genus *Ulva*: implications for commercial cultivation. J Appl Phycol. 2020;32(1):473–84.
9. Nisizawa K, Noda H, Kikuchi R, Watanabe T. The main seaweed foods in Japan. Hydrobiologia. 1987;151(1):5–29.
10. Neori A, Chopin T, Troell M, Buschmann AH, Kraemer GP, Halling C, et al. Integrated aquaculture: rationale, evolution and state of the art

emphasizing seaweed biofiltration in modern mariculture. Aquaculture. 2004;231(1–4):361–91.
11. Alsufyani T, Engelen AH, Diekmann OE, Kuegler S, Wichard T. Prevalence and mechanism of polyunsaturated aldehydes production in the green tide forming macroalgal genus *Ulva* (Ulvales, Chlorophyta). Chem Phys Lipids. 2014;183:100–9.
12. Bolton JJ, Cyrus MD, Brand MJ, Joubert M, Macey BM. Why grow *Ulva*? Its potential role in the future of aquaculture. Perspect Phycol. 2016;3(3):113–20.
13. Wang Y, Liu F, Liu X, Shi S, Bi Y, Moejes FW. Comparative transcriptome analysis of four co-occurring *Ulva* species for understanding the dominance of *Ulva prolifera* in the Yellow Sea green tides. J Appl Phycol. 2019;31(5):3303–16.
14. Lüning K, Kadel P, Pang S. Control of reproduction rhythmicity by environmental and endogenous signals in *Ulva pseudocurvata* (Chlorophyta). J Phycol. 2008;44(4):866–73.
15. Stratmann J, Paputsoglu G, Oertel W. Differentiation of *Ulva mutabilis* (Chlorophyta) gametangia and gamete release are controlled by extracellular inhibitors. J Phycol. 1996;32:1009–21.
16. Wichard T, Oertel W. Gametogenesis and gamete release of *Ulva mutabilis* and *Ulva lactuca* (Chlorophyta): regulatory effects and chemical characterization of the "swarming inhibitor". J Phycol. 2010;46(2):248–59.
17. Sousa AI, Martins I, Lillebø AI, Flindt MR, Pardal MA. Influence of salinity, nutrients and light on the germination and growth of *Enteromorpha sp.* spores. J Exp Mar Biol Ecol. 2007;341(1):142–50.
18. Mantri VA, Singh RP, Bijo A, Kumari P, Reddy C, Jha B. Differential response of varying salinity and temperature on zoospore induction, regeneration and daily growth rate in *Ulva fasciata* (Chlorophyta, Ulvales). J Appl Phycol. 2011;23(2):243–50.
19. Hiraoka M, Enomoto S. The induction of reproductive cell formation of *Ulva pertusa* Kjellman (Ulvales, Ulvophyceae). Phycol Res. 1998;46(3):199–203.
20. Zhang X, Xu D, Mao Y, Li Y, Xue S, Zou J, et al. Settlement of vegetative fragments of *Ulva prolifera* confirmed as an important seed source for succession of a large-scale green tide bloom. Limnol Oceanogr. 2011;56(1):233–42.
21. Vesty EF, Kessler RW, Wichard T, Coates JC. Regulation of gametogenesis and zoosporogenesis in *Ulva linza* (Chlorophyta): comparison with *Ulva mutabilis* and potential for laboratory culture. Front Plant Sci. 2015;6:15.
22. Gao S, Chen X, Yi Q, Wang G, Pan G, Lin A, et al. A strategy for the proliferation of *Ulva prolifera*, main causative species of green tides, with formation of sporangia by fragmentation. PLoS One. 2010;5(1):e8571.
23. Kessler RW, Crecelius AC, Schubert US, Wichard T. In situ monitoring of molecular changes during cell differentiation processes in marine macroalgae through mass spectrometric imaging. Anal Bioanal Chem. 2017;409(20):4893–903.
24. Tedeschi F, Rizzo P, Rutten T, Altschmied L, Baumlein H. RWP-RK domain-containing transcription factors control cell differentiation during female gametophyte development in *Arabidopsis*. New Phytol. 2017;213(4):1909–24.
25. He Y, Wang Y, Hu C, Sun X, Li Y, Xu N. Dynamic metabolic profiles of the marine macroalga *Ulva* prolifera during fragmentation-induced proliferation. PLoS One. 2019;14(5):e0214491.
26. Park J. Photosynthetic and biochemical traits change in the green-tide-forming macroalga *Ulva pertusa* during sporulation. J Phycol. 2020;56(2):549–57.
27. Katsaros C, Weiss A, Llangos I, Theodorou I, Wichard T. Cell structure and microtubule organisation during gametogenesis of *Ulva mutabilis* Føyn (Chlorophyta). Bot Mar. 2017;60(2):123–35.
28. Couso I, Vila M, Vigara J, Cordero BF, Vargas MÁ, Rodríguez H, et al. Synthesis of carotenoids and regulation of the carotenoid biosynthesis pathway in response to high light stress in the unicellular microalga *Chlamydomonas reinhardtii*. Eur J Phycol. 2012;47(3):223–32.
29. Chardin C, Girin T, Roudier F, Meyer C, Krapp A. The plant RWP-RK transcription factors: key regulators of nitrogen responses and of gametophyte development. J Exp Bot. 2014;65(19):5577–87.
30. Rovekamp M, Bowman JL, Grossniklaus U. *Marchantia* MPRKD regulates the gametophyte-sporophyte transition by keeping egg cells quiescent in the absence of fertilization. Curr Biol. 2016;26(13):1782–9.
31. Yin M, Zhang Z, Xuan M, Feng H, Ye W, Zheng X, et al. Conserved subgroups of the plant-specific RWP-RK transcription factor family are present in oomycete pathogens. Front Microbiol. 2020;11:1724.
32. Kasahara RD, Portereiko MF, Sandaklie-Nikolova L, Rabiger DS, Drews GN. MYB98 is required for pollen tube guidance and synergid cell differentiation in *Arabidopsis*. Plant Cell. 2005;17(11):2981–92.
33. Punwani JA, Rabiger DS, Drews GN. MYB98 positively regulates a battery of synergid-expressed genes encoding filiform apparatus–localized proteins. Plant Cell. 2007;19(8):2557–68.
34. Klepikova AV, Logacheva MD, Dmitriev SE, Penin AA. RNA-seq analysis of an apical meristem time series reveals a critical point in *Arabidopsis thaliana* flower initiation. BMC Genomics. 2015;16(1):466.
35. Searle I, He Y, Turck F, Vincent C, Fornara F, Kröber S, et al. The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in *Arabidopsis*. Genes Dev. 2006;20(7):898–912.
36. Pfab A, Grønlund JT, Holzinger P, Längst G, Grasser KD. The *Arabidopsis* histone chaperone FACT: role of the HMG-box domain of SSRP1. J Mol Biol. 2018;430(17):2747–59.
37. Grasser KD. The FACT histone chaperone: tuning gene transcription in the chromatin context to modulate plant growth and development. Font Plant Sci. 2020;11:85.
38. Lolas IB, Himanen K, Grønlund JT, Lynggaard C, Houben A, Melzer M, et al. The transcript elongation factor FACT affects *Arabidopsis* vegetative and reproductive development and genetically interacts with HUB1/2. Plant J. 2010;61(4):686–97.
39. Ferris PJ, Pavlovic C, Fabry S, Goodenough UW. Rapid evolution of sex-related genes in *Chlamydomonas*. Proc Natl Acad Sci. 1997;94(16):8634–9.
40. Waki T, Hiki T, Watanabe R, Hashimoto T, Nakajima K. The *Arabidopsis* RWP-RK protein RKD4 triggers gene expression and pattern formation in early embryogenesis. Curr Biol. 2011;21(15):1277–81.
41. Camargo A, Llamas A, Schnell RA, Higuera JJ, Gonzalez-Ballester D, Lefebvre PA, et al. Nitrate signaling by the regulatory gene NIT2 in *Chlamydomonas*. Plant Cell. 2007;19(11):3491–503.
42. Castaings L, Camargo A, Pocholle D, Gaudon V, Texier Y, Boutet-Mercey S, et al. The nodule inception-like protein 7 modulates nitrate sensing and metabolism in *Arabidopsis*. Plant J. 2009;57(3):426–35.
43. Koi S, Hisanaga T, Sato K, Shimamura M, Yamato KT, Ishizaki K, et al. An evolutionarily conserved plant RKD factor controls germ cell differentiation. Curr Biol. 2016;26(13):1775–81.
44. Yamazaki T, Ichihara K, Suzuki R, Oshima K, Miyamura S, Kuwano K, et al. Genomic structure and evolution of the mating type locus in the green seaweed *Ulva partita*. Sci Rep. 2017;7(1):11679.
45. Gao G, Clare AS, Rose C, Caldwell GS. Eutrophication and warming-driven green tides (*Ulva rigida*) are predicted to increase under future climate change scenarios. Mar Pollut Bull. 2017;114(1):439–47.
46. Oertel W, Wichard T, Weissgerber A. Transformation of *Ulva mutabilis* (Chlorophyta) by vector plasmids integrating into the genome. J Phycol. 2015;51(5):963–79.
47. Føyn B. Über die Sexualität und den Generationswechsel von *Ulva mutabilis*. Arch Protistenkd. 1958;102:473–80.
48. Provasoli L: Media and prospects for the cultivation of marine algae. In: Cultures and Collections of Algae Proceedings of US-Japan Conference, Hakone, September 1966: 1968. Japan Society of Plant Physiology.
49. Le Bail A, Dittami SM, de Franco P-O, Rousvoal S, Cock MJ, Tonon T, et al. Normalisation genes for expression analyses in the brown alga model *Ectocarpus siliculosus*. BMC Mol Biol. 2008;9(1):75.
50. Trapnell C, Salzberg SL. How to map billions of short reads onto genomes. Nat Biotechnol. 2009;27(5):455.
51. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31(2):166–9.
52. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 2010;26(1):139–40.
53. Team RC. R: A language and environment for statistical computing. 2013.
54. Corpet F. Multiple sequence alignment with hierarchical clustering. Nucleic Acids Res. 1988;16(22):10881–90.
55. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the dynamic tree cut package for R. Bioinformatics. 2008;24(5):719–20.

56. Vandepoele K, Van Bel M, Richard G, Van Landeghem S, Verhelst B, Moreau H, et al. Pico-PLAZA, a genome database of microbial photosynthetic eukaryotes. Environ Microbiol. 2013;15(8):2147–53.
57. Alexa A, Rahnenführer J. Gene set enrichment analysis with topGO. 2009. https://bioconductor.org/packages/release/bioc/vignettes/topGO/inst/doc/topGOpdf.
58. Supek F, Bošnjak M, Škunca N, Šmuc T. REVIGO summarizes and visualizes long lists of gene ontology terms. PLoS One. 2011;6(7):e21800.
59. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001;25(4):402–8.

Publisher's Note