

# **Transcriptome analysis reveals the spatial and temporal differentiation of gene expression in the sporophyte of *Undaria pinnatifida***

Louis Graf<sup>1</sup>, Younhee Shin<sup>1,2</sup>, Ji Hyun Yang<sup>1</sup>, Il Ki Hwang<sup>3</sup>, Hwan Su Yoon<sup>1,\*</sup>

<sup>1</sup>Department of Biological Sciences, Sungkyunkwan University, Suwon 16419, Korea

<sup>2</sup>Research and Development Center, Insilicogen Inc., Yongin-si 16954, Korea

<sup>3</sup>Aquaculture Management Division, National Institute of Fisheries Science, Busan 46083, Korea

\*corresponding author: hsyoon2011@skku.edu

## **Abstract**

Multicellular organisms are differentiated into specialized organs having specific functions. Kelps, morphologically diverse brown algae of the order Laminariales, exhibit specialized morphological features with a holdfast, a stipe and a blade. However, little is known about the expression profiles and their biological functions from the different tissues. Among the kelp species, the edible Pacific kelp, *Undaria pinnatifida* is an ecologically and economically important species native to the Northeastern Asia. It presents the typical organization of kelps but is characterized by the formation of its sporophyll, the sexually reproductive tissue, on the stipe. To study their respective gene expression profiles, we generated transcriptome sequences from four tissues; the sporophyll, meristem, healthy blade, and decaying apex blade were sampled at the three time points in the development of the sporophyte. We demonstrate an up-regulation of genes related to photosynthetic genes in the healthy blade, of genes related to cell proliferation in the meristem and of genes related to flagella development in the sporophyll. Overall, the study

revealed clear spatial and temporal variation in the expression patterns within and between the different tissues that were well correlated with the biological functions in the different tissues.

**Keywords:** *Undaria pinnatifida*, kelp, multicellularity, differential gene expression, spatio temporal gene expression, WGCNA

## Introduction

Multicellularity describes living organisms that are composed of more than one cell. Simple multicellular life forms, such as colonial clusters, balls, sheets, or filaments of cells, are common and evolved in at least 25-30 lineages (Knoll, 2011). These organisms gain selective advantages from their multicellular assemblage but they have limited cell-to-cell communication, cell differentiation and cell-to-cell transfer of resources. Furthermore, in simple multicellular organisms, all the cells are exposed to the external environment. On the other hand, complex multicellular organisms have well developed cell-to-cell communication, cell differentiation, a three-dimensional organization leaving only some cells in contact with the external environment and cell-to-cell transfer of resources. Evolution of such complexity only evolved six times: in the animals, ascomycete and basidomycete fungi, green algae and land plants, red algae and brown algae (Knoll, 2011). Importantly, in complex multicellular organisms, tissue differentiation is much more marked, resulting in specialized cells and organs having specific functions. This differentiation is the result of a genetic development program that leads to the specific expression of particular networks of genes in each tissue. In recent years, the function of these differentiated tissues was investigated through genome-wide analyses of their respective transcriptomes. These analyses led to the development of gene expression databases for many model organisms such as PEATmoss (Fernandez-Pozo *et al.*, 2020), fungi.guru (Jia *et al.*, 2020), or bgee (Bastian *et al.*, 2021). However, non-model organisms, such as brown algae, remain understudied and very little is known about their differentiation and tissue specificity.

The brown algae, class Phaeophyceae, is a distinct lineage within the heterokont algae (Stramenopiles or Heterokontophyta) (Andersen, 2004, Yang *et al.*, 2012). Brown algae have the defining characters for complex multicellularity, cell-to-cell communication (Cock *et al.*, 2010, Charrier *et al.*, 2012, Deniaud-Bouët *et al.*, 2017), internal transport (Fritsch, 1935, Schmitz & Srivastava, 1976), growth in three dimensions (Fritsch, 1935, Starko & Martone, 2016), and tissue differentiation (Fritsch, 1935, Kloareg & Quatrano, 1988). Brown algal complex multicellularity evolved relatively recently (around 200 Mya; (Kawai *et al.*, 2015, Silberfeld *et al.*, 2010) when compared to Basidiomycete fungi around 450 My (Berbee & Taylor, 2010), animals around 600 Mya (Peterson & Butterfield, 2005), and plants around 650 Mya (Munakata *et al.*, 2016). Notably, the order Laminariales, commonly named kelps, contains the largest and most complex morphologies among the algae. This lineage diverged even more recently, around 30 Mya (Silberfeld *et al.*, 2010, Kawai *et al.*, 2015) and little is known about the development and emergence of its morphologies.

The biology and ecology of the Pacific kelp *Undaria pinnatifida* has been extensively studied (Epstein & Smale, 2017). It has annual growth and it is native to the littoral regions of Northeast Asia (Epstein & Smale, 2017). Its life cycle is divided into two phases, a microscopic haploid phase named the gametophyte and a macroscopic diploid phase named the sporophyte, i.e. it has a typical kelp haplodiplontic life cycle. *Undaria* growth is principally constrained by sea water temperature. The gametophytes grow from July to September when temperatures exceed 25°C, and they reach maturity in October-November when temperatures decrease under 25°C (Morita *et al.*, 2003). The mature gametophyte produces egg and sperm, and after fertilization, the zygote develops into sporophyte. Sporophytes grow from October-November to July, and at maturity the sporophyll tissue undergoes meiosis to produce haploid meiospores. Sporophyte growth can be broadly divided in four stages: 1) the juvenile stage, 2) the vegetative stage, 3) the mature stage and 4) the decaying stage (Akiyama & Kurogi, 1982, Ohno & Matsuoka, 1993).

The juvenile stage represents the newly formed zygotes commonly observed from October to December. The juvenile stage, which develops from the zygote, resembles a small leaf with a very short stipe and holdfast. The juveniles occur from October to December, and they do not exceed 10 cm in length. The vegetative stage has more pronounced differentiation. The holdfast anchors the alga to a rocky substrate, an elongate stipe extends upward from the substrate and a pinnated blade grows on each side of the stipe. This stage with its deeply divided pinnae gives the blade its characteristic shape reflected in the name of the species. The intercalary meristem, situated at the base of the blade, provides the new growth (Charrier et al., 2012). Cell proliferation in the meristem compensates for biomass loss at the blade apex, which shreds and decays due to wave action. Therefore, the youngest cells are found at the base of the blade and oldest cells are at the blade apex. The highly differentiated regions of thallus play different ecological and physiological roles, such as growth, reproduction, photosynthesis, primary metabolism and chemical composition (Nimura & Mizuta, 2001, Schmid & Stengel, 2015, Küppers & Kremer, 1978, Lüning *et al.*, 1973, Wang *et al.*, 2013). Variations in biochemical concentrations have been observed for fatty acids (Barbosa *et al.*, 2020), pigments (Wang *et al.*, 2013), iodine (Hiroyuki & Hajime, 2010, Nitschke *et al.*, 2011), radical oxygen species (Hiroyuki & Hajime, 2010), and auxins (Kai *et al.*, 2006). Sporophytes in vegetative stage are generally observed from December to April.

As in all kelps, the meiosis takes place in the reproductive sporophyll, a specialized tissue that differentiates during the vegetative stage of the sporophyte. The sporophyll develops in a very distinctive way, forming a ruffled and folded tissue on the stipe between the base of the blade and the holdfast. The sporophyll forms on the stipe near the holdfast and newly formed sporophyll is continuously formed by meristematic activity (Schaffelke *et al.*, 2019). The sporophyll appears when daylength reaches approximately 9h (Pang & Lüning, 2004), i.e. late February early March. At the time of sporophyll formation, the vegetative sporophytes are 30-40 cm in length (Ohno & Matsuoka, 1993, Alvarez & Boraso, 2020). The first appearance of

sporophyll tissue is not synonymous with sexual maturity, i.e. meiospores are not produced until later.

The mature sporophyte stage begins when the sporophyll starts releasing meiospores. The mature sporophyll contains a large number of sporangia that form meiospores that are released from April to July (Akiyama & Kurogi, 1982, Ohno & Matsuoka, 1993). The sporophyll development is not homogeneous, and maturity occurs gradually from the holdfast to the blade (Schaffelke et al., 2019). Recent investigation of different regions of the sporophyll revealed that gene expression varied as a function of the level of development (Shan et al., 2020a). However, generally the gene expression of the sporophyll was notably linked to meiosis and meiospore development (Shan et al., 2020a). The meiospores are motile, biflagellate cells that are responsible for dispersal in the species (Forrest et al., 2000). The meiospores eventually attach to a substrate and develop into gametophytes. Interestingly, the blade growth significantly decreases during the reproductive stage (Dean & Hurd, 2007). In the final decaying stage, the blade tissue dies off and mechanically dissolves, and during this stage the meristem is inactive and does not compensate for blade erosion (Akiyama & Kurogi, 1982).

Recently, the nuclear genome of *Undaria pinnatifida* was completed independently by two teams (Graf et al., 2021, Shan et al., 2020b). Because the availability of an annotated gene model greatly facilitated transcriptome analysis, we designed a transcriptomic study of the sporophyte during its entire lifespan. Previously, transcriptomic analysis of the kelps (and brown algae in general) focused on environmental stress responses to heat (Liu et al., 2013), light (Li et al., 2020), salinity (Rugiu et al., 2020), grazing (Ritter et al., 2017), depth (Paul et al., 2020, Salavarria et al., 2018) and multiple stresses (Machado Monteiro et al., 2019, Zhang et al., 2021b). Only in the past few years, transcriptomic studies in the kelp *Saccharina japonica* included different tissues and/or different development stages (Lu et al., 2020, Shao et al., 2019, Zhang et al., 2020, Zhang et al., 2021a). However, despite using genome-wide transcriptome sequencing approaches, these studies largely focused on specific biological processes such as the

biosynthesis of alginate and mannitol (Shao et al., 2019), mannuronan C-5 epimerase (Zhang et al., 2021a) or sulfotransferase (Lu et al., 2020). Only the study by Shao et al. (2019) attempted to integrate alginate biosynthesis in a larger biological context through weighted gene co-expression network. This integration revealed complex and unsuspected interactions between biosynthesis pathways; however, this analysis did not include the important sporophyll and meristematic tissues (Shao et al., 2019). Our analysis was designed to address this omission.

We produced transcriptomes from the meristem, sporophyll, healthy blade and decaying blade tissues for juvenile, vegetative and reproductive stages (Figure 1). We hypothesized that important gene expression regulation occurs in the different regions of the blade, and further that these regulations might decrease overtime toward the end of the lifespan of the sporophyte. We also aimed at investigating maturation of the sporophyll before and after it reaches maturity (i.e. releasing meiospores). To that aim, we produced a RNAseq collection for four tissues at three time points in the life cycle of *Undaria pinnatifida*. We analyzed the expression of 19,678 genes both in space and time and using a weighted gene co-expression network. The gene expression patterns of the different tissues and their link to physiological functions are discussed. Notably, we report the up-regulation of (a) photosynthetic genes in the blade, (b) cell proliferation genes in the meristem, and (c) genes related to flagellar development. Overall, our results represent the first expression atlas for a brown alga, and because our data is from a complex multicellular organism, the results will likely be of interest to the broader scientific community.

## Material and methods

### Algal specimen and sample preparation

Specimens of *Undaria pinnatifida* were harvested on a single outdoor culture line in Wando, Korea (approx. 34°19'30.2"N 126°39'05.5"E). Sporophyte development was sampled by collecting entire specimens on November 19, 2016, February 26, 2017, and April 25, 2017. All

the samples were harvested on the line, rapidly washed in autoclaved seawater, dried with paper towels and frozen in liquid nitrogen. To dislodge the tightly compacted sporophytes on the line, we had to cut above the holdfast, and therefore we could not investigate the transcriptome in the holdfast tissue. The November specimens represented the juvenile stage and were ca. 10 cm with a short stipe with no morphological differentiation of the blade; no sporophyll was present (Figure S1). The November individuals were termed “undifferentiated sporophytes”. Conversely, the February and April specimens were fully differentiated with a long stipe containing sporophyll tissue and a pinnated blade with marked decomposition and shredded apex (Figure S2). The February samples represent the vegetative sporophyte stage while the April samples represent the reproductive sporophyte stage. We dissected ca. 5 cm<sup>2</sup> from each of four tissues: (1) the sporophyll along the stipe, (2) the meristem from each side of the stipe (ca. 2 cm on each side) at the base of the blade, (3) the blade tissue ca. 10 cm above the meristem (termed "healthy blade"), and (4) the old decaying blade apex tissue (termed "apex blade"). The healthy blade was dissected a few cm away from the stipe and not in the pinnae as they also present eroding regions at their apex. The apex blade was dissected from damaged tissue at the blade tip and represented the oldest cells of the sporophyte. The February and April specimens differed in their level of maturation of the sporophyte: February sporophylls were restricted to a small portion of the stipe whereas April sporophylls totally covered the stipe (Figure S3). Therefore, the February and April specimens were named thereafter according to their level of maturation: “vegetative sporophyte” and “reproductive sporophyte” for the February and April specimens, respectively. Frozen samples were maintained at -80°C until RNA extraction.

## RNA extraction

Frozen samples were ground using an Automill TK-AM5 frozen crusher (<http://www.tokken.jp>). Total RNA from each tissue was extracted independently using the GeneMatrix Universal RNA

Purification kit (EurX, Gdansk, Poland) following manufacturer's instructions. Total RNA was loaded and resolved through a 1.0% agarose gel in order to check its integrity. After a 30 min electrophoresis at 100V, the gel was stained in a solution of ethidium bromide (0.5 µg/ml) for 30 min and unstained in distilled water. Samples presenting DNA contamination were treated with DNase. RNA quantification and qualification was performed on the 2100 Expert Bioanalyzer platform (<https://www.genomics.agilent.com>) using the RNA 6000 Nano Kit (Agilent, CA, USA).

### **Illumina paired-end cDNA library and sequencing**

A total of 27 RNA libraries were prepared from the different tissues sampled. NGS sequencing libraries were generated from one microgram of total RNA using the TruSeq RNA Sample Prep Kit (Illumina, CA, USA) according to the manufacturer's instructions. The resulting libraries were then paired-end sequenced (2x101 bp) for all of samples with HiSeq2500 system.

### **Preprocessing and DEG analysis**

To remove low-quality sequences and adapter contaminations, sequencing reads were assessed using Trimmomatic-0.32 (Bolger *et al.*, 2014). Preprocessed clean reads were mapped to the *Undaria pinnatifida* Kr2015 gene model (Graf *et al.*, 2021) using the option “--sjdbGTFfile” in STAR (Dobin *et al.*, 2013). The expression levels of all samples were estimated by calculating the transcripts per kilobase million (TPM) values for each library using RSEM v1.2.31 (Li & Dewey, 2011). Level of expression threshold was set to 0.3 TPM and genes covered by less than 5 reads were also discarded. The distances between the samples was assessed by principal component analysis and by Euclidean distances using the TPM of each sample as calculated with the statistical language R v3.6.3 (R Core Team, 2020). The TPM values were subjected to edgeR

(Robinson *et al.*, 2010) to obtain the differential expressed genes (DEGs) from all samples. DEGs were defined at absolute log<sub>2</sub> fold change  $\geq 2$  and *p*-value  $\leq 0.001$ .

### **Weight gene-gene co-expression network analysis (WGCNA)**

Pattern discovery was performed to identify genes involved in tissue development by growth period. Log<sub>2</sub> transformed TPM values for all expressed genes were used to estimate the similarity score between genes and tree height was cut at 0.8. Total of six signed clusters were defined using WGCNA (Langfelder & Horvath, 2008) with min cluster size=100. Eigenvectors for each sample by cluster were calculated using ‘moduleEigengene’ function of WGCNA package in R.

### **Accession numbers**

All the sequencing reads were deposited in the National Center for Biotechnology Information Sequence Read Archive under the Bioproject PRJNA81331. The genes expression levels in each transcriptomes library in transcript per million (TPM) value are reported in Supplementary File 1. Differential gene expression and assigned cluster from the WGCNA analysis are reported in Supplementary File 2.

## **Results and Discussion**

### **Sequencing and mapping**

We collected biological triplicates of each tissue under study for a total of 27 cDNA libraries and generated a total of 142 Gb of paired-end sequencing reads (average 63 million reads per library; Table S1) with mapping rates ranging from 69.28% to 82.57% (average 74.87%; Table S1). These mapping rates were comparable to that of most brown algal transcriptomic analysis; e.g.

77.29% in *Sargassum vulgare* (Kumar *et al.*, 2017); 85.5-87.6% in *Saccharina latissima* (Li *et al.*, 2020, Machado Monteiro *et al.*, 2019), or 80.4% in *Saccharina japonica* (Shao *et al.*, 2019).

The TPM density distributions in the 27 samples were consistent (Figure S4), as did the normalized TPM counts (Figure S5). All the sequencing reads were deposited in the National Center for Biotechnology Information Sequence Read Archive under the Bioproject PRJNA81331.

### Gene expression patterns

To identify transcriptional profiles of the different tissues, we performed clustering analysis of the tissues and genes, as it is common for transcriptomic analysis (Liu & Si, 2014). Principal components analysis (PCA) and hierarchical clustering analysis were used to explore the expression across the 27 samples. Generally, the PCA revealed a clear segregation of samples according to the tissue type; there was a loose segregation of samples based upon sampling date (Figure 2a). The PC1 explained 28.21% of the variance and separated the samples in two large clusters: 1) the apex blade, the healthy blade, and the undifferentiated sporophyte samples (Figure 2a, left) and (2) the meristem and sporophyll samples (Figure 2a, right). The second principle component (PC2, Figure 2a) and the hierarchical cluster analysis (Figure 2b) helped to further delimitate four clusters of highly correlated samples: 1) the sporophyll in the reproductive sporophyte (April) (Figure 2a, lower right; Figure 2b, upper left), 2) the meristem and sporophyll in the vegetative sporophyte (February) (Figure 2a, mostly upper right; Figure 2b, upper left), 3) the apex blade in both the vegetative and reproductive sporophytes (Figure 2a, lower left; Figure 2b, midregion), and 4) the healthy blade and apex blades in undifferentiated, vegetative and reproductive sporophytes (Figure 2a, left; Figure 2b, lower right).

The expression across the triplicates of the nine conditions was generally uniform, but we detected a few incongruences. For example, in PC2 the sporophyll 1 of the vegetative sporophyte grouped with the reproductive tissues in the lower component (Figure 2a) while the healthy blade

1 of the reproductive sporophyte 1 and healthy blade samples from two reproductive sporophytes also grouped with the reproductive tissues in the lower component (Figure 2a). These incongruencies could be explained by our choice to work with individuals grown in a natural environment, i.e. on a cultivation rope at sea. First, weather conditions on the days of sampling in February and April varied greatly and could have affected the expression of genes, notably related to response to environmental cues. Second, the maturation of the sporophyte is not exactly synchronized between the individuals in a population, building noise into our data set. This appears to be particularly clear for the sporophyll tissue of vegetative sporophyte 1, which had an expression profile more similar to that of sporophylls of the reproductive sporophytes. Third, unknowingly to us at the initiation of this project, the genotype of the individuals considered in our study could have affected the level of gene expression. To that end, we had chosen to work from cultivated individuals assuming that their genotypes would be less diverse than in wild populations. However, the vegetative and reproductive sporophytes that were genotyped as part of a population genomics study of *Undaria pinnatifida* had unexpectedly high genetic diversity (Graf et al., 2021). Despite these exceptions, these results overall support the consistency of our biological replicate for downstream analysis.

The gene model of the *U. pinnatifida* Kr2015 predicted that there were 20,716 genes (Graf et al., 2021), and we found that 19,678 genes (94.99%) were expressed in at least one of the samples in this study. The 1,038 genes that were not expressed might represent genes expressed uniquely in tissues not included in this study (e.g. sporophyte holdfast, gametophyte tissues, gametes). To explore the distribution of the 19,678 expressed genes among the different tissues, we performed an UpSet plot (Figure 3a). There was no substantial difference between the tissues in the number of gene expressed (Table 1). The sporophylls (either from the vegetative or reproductive sporophytes) expressed the most genes (18,152; 92.25% and 18,595; 94.49%, respectively), followed by all other tissues expressing roughly the same number of genes (meristem of reproductive sporophytes: 16,935; 86.06%, undifferentiated sporophytes: 16,819;

85.47%, healthy blade of vegetative sporophytes: 16,779; 85.53%, healthy blade of reproductive sporophytes: 16,660; 84.66%, meristem of vegetative sporophytes: 16,569; 84.20%, apex blade of reproductive sporophytes: 16,551; 84.11%, apex blade of vegetative sporophytes: 16,311; 82.89%). A large majority of the genes were expressed in all tissue samples (14,360; 72.97%; Figure 3a). Genes expressed in only one tissue ranged from 27 genes (healthy and apex blades of the reproductive sporophytes) to 366 genes (sporophylls of the reproductive sporophytes). Furthermore, there were 355 genes that were expressed only in the sporophylls of the vegetative and reproductive sporophytes. Thus, in terms of gene expression, the sporophyll tissues were the most distinct (Table 1). Therefore, this multicellular kelp was similar to multicellular animals and plants, where sexual reproductive tissues had very distinct and specific gene expression profiles (e.g. (Bentz *et al.*, 2019, Kudapa *et al.*, 2018, Machado *et al.*, 2020, Xanthopoulou *et al.*, 2021)).

## Housekeeping genes

Taking advantage of our collection of transcriptomes from different tissues, we attempted to identify housekeeping genes (HKGs), i.e. genes expressed consistently across a wide range of conditions and/or in a variety of tissues and organs (Hounkpe *et al.*, 2021). We followed the criteria presented in Machado et al. (2020) and used the product of the coefficient of variation (CoV) and the ratio of maximum to minimum expression to identify potential HKGs. The 718 genes within the fifth percentile of this product were classified as HKGs, and they showed constant expression across the tissues (Figure S6a). We further used the tissue specificity index *Tau* to compare the specificity of the HKGs and the remaining genes. The specificity index ranged from 0 to 1, where low value indicated a broad expression and high value a tissue specific expression. The HKG had a *Tau* score ranging from 0 to 0.25 (mean = 0.048), while the remaining genes generally had higher *Tau* scores (mean = 0.376; Figure S6b). This result corroborated our classification of the HKGs based on the CoV and the maximum-minimum expression ratio.

Enrichment analysis of the putative 718 HKGs revealed that an important number of these genes were involved in biological processes related to translation, ribosomal biogenesis, or intracellular traffic (Table S2). We identified 50 ribosomal proteins and seven elongation factor genes. To date, despite their importance to gene expression analysis, there has been no identification of *Undaria pinnatifida* HKGs. In a recent study, Xing *et al.* (2020) attempted to identify and evaluate such genes in the sugar kelp *Saccharina latissima*. They identified three genes (i.e. *ATPase* [F-type H-ATPase beta subunit]; *EIF5B* [eukaryotic translation initiation factor 5B]; *ND11* [NADH dehydrogenase (ubiquinone) sub-unit 11]) as the best choices for RT-qPCR normalization (Xing *et al.*, 2020). Interestingly, the two homologs of *ATPase* and the single homolog of *ND11* found in the genome of *U. pinnatifida* were present in our HKGs. Furthermore, the *U. pinnatifida* homolog of *EIF5B* was excluded only by an extremely slight margin ( $1.942 > 1.934$ ) and therefore it too could be considered an HKG (Figure S6c). This overlap of our HKG and previously proposed genes validates both datasets, even though the two copies of the *ATPase* found in *U. pinnatifida* make this gene improper in this species. More importantly, this suggests that the remaining HKGs uncovered in our analysis will be good candidates in future studies (e.g. to normalize gene expression under specific conditions, to investigate developmental stages, or to study specific tissues).

### **Spatio-temporal differential gene expression in the sporophyte of *Undaria pinnatifida***

An expression level heatmap across genes and tissues revealed clear differences among the tissues, most notably the sporophyll of the reproductive sporophytes (Figure 3b). Pairwise analysis of the different tissues within the vegetative sporophyte and the reproductive sporophyte samples identified differentially expressed genes (DEGs, i.e. up- and down-regulated collectively) between the different tissue types while comparisons of the undifferentiated,

vegetative, and reproductive sporophyte samples identified seasonal DEGs (Figure 3c). The number of DEGs varied greatly between the comparisons. In the vegetative sporophyte, the smallest difference was between the healthy blade tissue and the apex blade tissue (577 DEGs) while the largest difference was between the meristem tissue and the apex blade tissue (2,938 DEGs). In the reproductive sporophyte, similarly, the smallest difference was between the healthy blade tissue and the apex blade tissue (271 DEGs), but the largest difference was between the sporophyll tissue and the apex blade tissue (4,406 DEGs).

We investigated the evolution of gene expression in different tissues over the November to April time period by further analyzing the DEGs (Figure 3c). We used the undifferentiated sporophyte as a comparison for the tissues of the vegetative sporophytes and we compared tissues between the vegetative and reproductive sporophytes. Generally, there were more DEGs in the comparison between undifferentiated and vegetative sporophytes (2,504 DEGs) than between vegetative and reproductive sporophytes (1,507 DEGs). In the meristem and the apex blade, important variations were detected between the undifferentiated and vegetative sporophytes (3,198 and 1,731 DEGs, respectively), but not between the vegetative and reproductive sporophytes (405 and 99 DEGs, respectively). Finally, in the healthy blade, low variations were detected between the undifferentiated and vegetative sporophytes, as well as between the vegetative and reproductive sporophytes (571 and 180 DEGs, respectively).

Overall, our DEGs analysis demonstrated that there were specialized expression patterns in the different tissues and at different stages of the life cycle. However, we also revealed that the numbers of DEGs varied greatly between the comparisons. The sporophylls appeared to have the most specialized expression pattern with the highest number of DEGs when compared with other tissues of the sporophyte or at different time points (Figure 3c). On the other hand, there were notably less differences in the expression between the healthy blade and the apex blade. Furthermore, the expression in these tissues was the closest to that of the undifferentiated

sporophyte and it remained relatively unchanged in the vegetative and reproductive sporophytes (Figure 3c).

### Gene co-expression network analysis

Due to the complex multi-dimensional nature of our dataset (i.e. including temporal and spatial DEGs comparisons), it was difficult to analyze tissue functions using only DEGs analysis. Therefore, we performed a co-expression network analysis (WGCNA) to determine the different functional groups of genes expressed in different tissues in which they were up- and down-regulated (Figure 4). The analysis, including the triplicate samples from the nine tissues, resulted in a network presenting seven principal clusters regrouping a total of 9,027 genes. The different clusters contained varying numbers of genes, 267 to 2,734, and they confirmed the differential expression across the different tissues (Figure 4). To explore the functions encoded by the genes in these clusters, we analyzed their annotation and performed GO term enrichment analysis.

#### *Cluster-A: up-regulated genes in the undifferentiated sporophyte*

The 267 genes of the cluster-A were up-regulated in the undifferentiated November sporophytes and down-regulated in all other samples (Figure 4, Table S4). Of the 267 genes, 190 were annotated, but only 47 had a functional annotation; therefore, the enrichment analysis did not recover a large number of functions (Table S3). However, the genes might show similarities to embryonic plants. For example, the cluster contained seven endo-1,3-beta-glucanase genes out of 25 copies, a large gene family having various physiological functions in plants, and notably in the embryonic development of plants (Dong & Dunstan, 1997). Thus, the undifferentiated sporophyte of *U. pinnatifida* is similar to the undifferentiated plant embryo (Epstein & Smale, 2017). Furthermore, the cluster-A contained three imm (immediately upright) upregulated proteins that have been proposed to play roles in the development of the brown algae (Peters *et al.*, 2008).

Therefore, the up-regulation of some genes corroborated their potential role during early differentiation and development of the sporophyte. The search for unique, homologous or analogous genes during embryonic development constitute interesting targets for future studies with controlled growing conditions.

### *Cluster B and C: photosynthesis related genes*

Cluster-B and cluster-C contained 484 and 2,083 genes, respectively (Figure 4, Table S3). In both clusters, the most significantly enriched GO terms were linked to functions localized in the thylakoid of the chloroplast and related to photosynthesis, pigment biosynthesis, energy, and carbon metabolism (Table S3). These genes notably encoded light harvesting complex proteins, electron transport proteins, magnesium chelatase, and chlorophyll synthase, phytoene synthase/desaturase, or geranylgeranyl pyrophosphate synthetase. Other enriched GO terms were related to the energy metabolism and carbon metabolism associated with photosynthesis, notably with genes encoding enolase, ATP synthase gamma chain, glyceraldehyde-3-phosphate dehydrogenase, fructose-1,6-bisphosphate aldolase, fructose-bisphosphatase, phosphoribulokinase, mannitol 1-phosphate dehydrogenase or phosphoglycerate kinase, and NADH dehydrogenase activity. These genes were generally up-regulated in the undifferentiated sporophytes as well as in the healthy and apex blade tissues of the vegetative and reproductive sporophytes (even though the cluster B was not significantly up-regulated in the apex of the reproductive sporophyte).

Generally, the cluster-B and -C indicated that photosynthetic genes were up-regulated in the vegetative tissues and down-regulated in the meristem and sporophyll tissues. In the tangle kelp *Laminaria hyperborea*, the chlorophyll fluorescence (Fv/Fm) and relative electron transport rate (rETR) were estimated for young sporophytes (i.e. equivalent to our undifferentiated sporophytes), vegetative old sporophytes (i.e. equivalent to our blades), and old fertile

sporophytes (i.e. equivalent to our sporophylls) in controlled culture conditions (Olischläger *et al.*, 2012). This study showed that the highest photosynthesis activity occurred in the young sporophyte and that the lowest activity occurred in the fertile sporophyte tissues. Similarly, in the kelp *Saccharina japonica*, net photosynthesis was lower in fertile sporophytes than in the vegetative sporophytes (Nimura & Mizuta, 2001). These results support a correlation between the observed gene expressions in the sporophyte of *Undaria pinnatifida* and the primary metabolism. Interestingly, the apex blade and healthy blade showed a combination of slightly down-regulated (i.e. cluster-B) and up-regulated genes (i.e. cluster-C) in link with photosynthesis, suggesting that in the oldest part of the sporophyte (i.e. at its apex), the photosynthesis activity was diminished when compared to the youngest part of the sporophytes (i.e. near the meristem). However, further validations by measures of Fv/Fm and rETR are required to test this hypothesis.

#### *Cluster-D: up-regulated genes in the meristems and sporophylls*

The 2,098 genes grouped in the cluster-D were up-regulated in the meristem of the vegetative and reproductive sporophytes as well as in the sporophyll of the vegetative sporophytes, but cluster D genes were only slightly up-regulated in sporophyll tissue of the reproductive sporophytes, and they were down-regulated in the remaining tissues, including November samples (Figure 4, Table S3). The GO term enrichment analysis of the cluster-D revealed numerous terms linked to DNA metabolic processes such as DNA replication and DNA repair and GO localization (CC) to the nucleus and in link with the chromosomes (Table S3). The genes associated to these functions notably encoded DNA polymerase, DNA primase small and large subunits, helicases of the MCM complex or the genes forming the chromosome passenger complex. Cell proliferation of *Undaria pinnatifida* occurs in an intercalary meristem placed at the base of the blade and is typical of brown algae of the order Laminariales (Charrier *et al.*, 2012). Therefore, it is not surprising to find a general up-regulation of genes involved in DNA biosynthesis, in tissues where cells are actively dividing. Furthermore, the sporophyll of *U. pinnatifida* is a differentiated region located

on the stipe above the holdfast and under the meristem. The development of the sporophyll begins relatively early in the lifespan of *U. pinnatifida* and continues until the senescence of the alga (Alvarez & Boraso, 2020). This development also occurs through cell proliferation, therefore explaining the up-regulation of DNA biosynthesis genes like that observed in the meristem. The cell division in photosynthetic organisms is known to be linked to light and notably by gene of the cyclin family (Moulager *et al.*, 2007) and in marine environment, blue light appears to be a key factor in growth and development of kelps (Shi *et al.*, 2005, Wang *et al.*, 2010, Zhang *et al.*, 2005). Up-regulation of cyclin genes in sporophytes of *Saccharina japonica* exposed to blue light was reported (Deng *et al.*, 2012), and a blue light specific photoreceptor, aureochrome, has been shown in diatoms to regulate cell divisions (Huysman *et al.*, 2013). In our transcriptome dataset, the aureochrome and cyclin genes were up-regulated principally in the meristem and sporophyll (Figure S7), suggesting a link in the regulation of the aureochrome and cyclin genes in kelps, and therefore of cell divisions. The mechanisms of maintenance of the cell pluripotency in the meristem of the brown algae remains elusive, nevertheless, our collection of transcriptomes could potentially serve to identify or validate potential candidate genes for this function.

Another corollary of cell division is the biosynthesis of daughter cell walls. Unsurprisingly, many genes related to cell wall biosynthesis were found in the cluster-D, notably cellulose synthase, GDP-fucose fucokinase, GDP-mannose 4,6-dehydratase, fucosyltransferase and sulfotransferase (Table S3). These genes are involved in the biosynthesis of alginates, cellulose and fucose-containing sulphated polysaccharides, all essential components of the brown algal cell wall (Michel *et al.*, 2010, Charrier *et al.*, 2019). Furthermore, we found that the cluster-D contained 64 genes encoding putative pectin lyase along with 25 mannuronan C5-epimerase (MC5E), the two most abundant gene families in this cluster. These two gene families were extensively expanded in the common ancestor of the kelps (Ye *et al.*, 2015, Graf *et al.*, 2021). The MC5E are responsible for the conversion of alginate into guluronate, which alters the cell wall's flexibility and its proprieties (Rabillé *et al.*, 2019). Interestingly, an alginate lyase protein

recently discovered in *Saccharina japonica* (Inoue & Ojima, 2019), showed high similarity with the pectin-lyase identified in the cluster-D. This suggests that intense growth (i.e. cell division) in the meristem and sporophyll of *Undaria pinnatifida* is accompanied by a modification of the cell wall, notably through the modification and lysis of the alginates. However, great care should be taken with such interpretations. The role of the cell wall components in the growth of brown algae remains largely unexplored and thus far no clear correlation between alginates alternation and growth have been established (Charrier et al., 2019). Furthermore, alginates are not the only polysaccharides found in the cell wall of brown algae; therefore, the complex interaction between different polysaccharides needs to be considered (Charrier et al., 2019).

#### *Cluster E and F: up-regulated genes in the sporophylls*

The clusters-E and -F contained 277 and 1,084 genes, respectively. The cluster genes were up-regulated in the sporophyll tissue of both the vegetative and reproductive sporophytes, but they were down-regulated in the other tissues (Figure 4, Table S3). The cluster-E contained few functionally annotated genes and recovery of enriched GO terms was limited. Cluster-F genes were enriched in GO terms that were related to cell cycle, DNA recombination, and meiotic processes; furthermore, cluster F genes played an important cytoskeletal role and other microtubule-based processes (Table S3).

*U. pinnatifida* propagate and disseminate by releasing flagellate meiospores from their sporophyll tissues. Meiospores are produced by meiosis and are sexually distinct (male or female), but they are not gametes. Rather, a flagellate meiospore (= zoospores) gives rise to microscopic male or female gametophyte that in turn produces either male and female gametes. Therefore, upregulation of meiotic genes is predicted for the sporophyll tissue of reproductive sporophytes but interestingly it was also detected in the vegetative sporophytes (Figure S8). These genes notably included homologues of the meiotic recombination protein DMC1/LIM15

and of the meiosis-specific protein HOP1 and the meiotic nuclear division protein 1. This last gene was also identified to be up-regulated during the development of the sporophyll of *U. pinnatifida* (Shan et al., 2020a). Together with our results, the up-regulation of meiotic genes in the vegetative and reproductive sporophyll could suggest that meiotic genes are not only up-regulated during the meiosis (i.e. in the mature sporophyll) but also before the meiosis. However, as the sporophyll 1 of the vegetative sporophyte appeared to have an expression more similar to that of sporophyll of the reproductive sporophytes than the vegetative sporophytes (Figure 2) this upregulation could well be an artifact.

Furthermore, in kelps, differentiation of the sporophyll is associated with morphological, chemical, and physiological changes. For example, as stated above (cluster-B and -C), photosynthesis rates decline in the sporophyll tissue of kelps (Nimura & Mizuta, 2001). In *U. pinnatifida* specifically, the sporophyll is characterized by low alginate concentration (Apoya et al., 2005, Olischläger et al., 2012) and high concentration of fucoidans (Skriptsova et al., 2010) when compared to the rest of the thallus. Fucoidans biosynthesis in the kelps involved mannose-6-phosphate isomerase (MPI), phosphomannomutase (PMM), GDP-mannose 4, 6-dehydrogenase (GM46D), fucose kinase (FK), GDP-fucose pyrophosphorylase (GFPP), fucosyltransferase (FUT), and sulfotransferase (ST) as proposed by genomic comparative analysis (Skriptsova, 2015, Chi et al., 2018, Lu et al., 2020, Michel et al., 2010). Despite a higher concentration of fucoidans in the sporophylls, the cluster-E and -F were not significantly enriched in genes encoding these proteins (Table S3), and in general, these genes were not up-regulated in the sporophyll (Figure S9). This was somewhat contrary to previous work that was observed in transcriptomes analysis of the developing sporophyll of *U. pinnatifida* (Shan et al., 2020a). In this study, the FK, GFPP, FUT and ST genes appeared up-regulated in the sporophyll, notably in the early developing sporophyll. Interestingly, these genes were actually down-regulated in the fully developed sporophyll when compared to the developing sporophyll (Shan et al., 2020a). Together with our results, this could suggest that the accumulation of fucoidans in the sporophyll of *U.*

*pinnatifida* occurs during its differentiation and that up-regulation of the fucoidans biosynthesis gene is only transient. Similarly, we did not observe a significant down-regulation of the genes involved in the alginate biosynthesis (i.e. GDP-mannose 6- dehydrogenase [GMD], Mannuronan synthase [MS], Mannuronate C5-epimerase [MC5E]) in the sporophyll (Figure S9).

Interestingly, in *Undaria pinnatifida* beside their higher concentration, the fucoidans of the sporophyll were also distinguished from the fucoidans of the rest of the thallus by the higher proportion of galactose (Skriptsova et al., 2010). The process through which binding of galactose to the fucose chain occurs in the brown algae remains not fully understood but the role of a galactosyltransferase has been proposed (Skriptsova, 2015). Interestingly, a single copy beta- 1,3 galactosyltransferase was significantly up-regulated in the sporophyll (and meristem) of *U. pinnatifida*, making it an interesting target to further elucidate the fucoidans biology in the brown algae.

#### *Cluster-G: up-regulated genes in the sporophyll of the reproductive sporophyte*

The cluster-G was the largest cluster, containing 2,734 genes that were up-regulated specifically in the sporophyll of the reproductive sporophyte and down-regulated in all the other tissues (Figure 4, Table S3). The principal enriched GO terms were similar to the one found in the cluster-F, notably related to cytoskeleton and microtubule processes. However, GO terms specific to cluster-G were related to the assembly of the flagellum and cell motility (Table S3). Interestingly, in *Saccharina japonica*, the DNA content of the fertile sporophyte was twice high as in the sessile sporophyte upon sporophyll maturation (Nimura & Mizuta, 2001).

Meiospores of *U. pinnatifida* are released from mature sporangia in the sporophyll and disseminate the species. These specialized cells are mobile and possess the typical heteromorphic flagella of the heterokont algae: a short smooth flagellum and a long flagellum bearing two rows of tripartite tubular hairs (Henry & Cole, 1982, Fu et al., 2014). Therefore, the up-regulation of genes related to flagellum in the sporophyll of the reproductive sporophyte is well explained by

the function of this tissue (Figure S10). The absence of up-regulation of these genes in the sporophyll of the vegetative sporophyte seems to indicate that the development of the flagellum occurs later in the life cycle of *U. pinnatifida*. The up-regulation of genes related to the flagella was also reported by Shan et al. (2020a) in their transcriptome analysis of *U. pinnatifida* sporophyll. Somewhat similarly to our observation, the up-regulation of the genes related to the flagella was found to be more important at a later stage of sporophyll development than at earlier stages (Shan et al., 2020a).

Overall, our WGCNA analysis helped to characterize the biological functions of the different tissues of the sporophyte of *Undaria pinnatifida* at a molecular level. This simply correspond to what is known on the biology of the kelps, with the blade in its entirety (healthy blade and apex blade in our analysis) as the photosynthetic organ, the meristem, the center of cell divisions and the sporophyll as the tissue into which gametes differentiate (Figure 5). Furthermore, this unprecedented collection of transcriptomes constitutes the first step toward the establishment of a gene expression database for the ecologically and economically important Pacific kelp *Undaria pinnatifida*. The addition of transcriptomes from the female and male gametophytes as well as some missing tissues such as the stipe or the holdfast will likely be the next steps towards this goal. A gene expression database could be an invaluable tool to further our understanding of the biology of the kelps and of their complex multicellularity.

## Conclusion

In this study, we established a collection of 27 transcriptomes from *Undaria pinnatifida* sporophytes. Our dataset covers the entire development of the sporophyte and includes the four major tissues that constitute it, namely the sporophyll, the meristematic region, the healthy blade, and the apex blade. The combination of tissue specificity, DEGs, and co-expression analysis revealed important spatial and temporal variation of the gene expression in link with the

biological function of these tissues. Comparisons of the undifferentiated embryonic sporophyte with the vegetative sporophyte tissues revealed a clear specialization of the gene expression in the differentiated vegetative sporophyte. Within the differentiated vegetative sporophyte, the healthy blade and apex blade tissues expressed one set of genes while the meristem and sporophyll were undergoing rapid cell division and expressed a different set of genes. Comparison between the vegetative and reproductive sporophytes revealed that, apart from the reproductive sporophyll, tissue specific gene expression was maintained during the lifespan. In the reproductive sporophyll, we observed a switch in gene expression from the typical cell proliferation and vegetative growth genes to a meiospore-producing set of genes.

## Acknowledgments

The authors would like to express their gratitude to Kyungwoon Kim who provided samples from his farming facility in Wando, Korea. The authors would like to warmly thank Cátia Machado Monteiro, Kyounghee Lee and Robert A. Andersen for their valuable and meaningful reviews of the manuscript that greatly helped for its scientific content and readability.

## Authors' contribution

LG and HSY designed the study. LG, JHY and IKH collected the samples and extracted RNA. LG and YS generated and analyzed the transcriptome data. All the authors read and approved the final manuscript.

## Funding information

This project was funded by the Collaborative Genome Program of the Korea Institute of Marine Science and Technology Promotion (KIMST) funded by the Ministry of Oceans and Fisheries (MOF) (20180430), and the National Research Foundation of Korea (2017R1A2B3001923, 2022R1A2B5B03002312) awarded to HSY and the NRF (2020R1C1C1008173) awarded to LG.

## Reference

- Akiyama, K. & Kurogi, M. 1982. Cultivation of *Undaria pinnatifida* (Harvey) Suringar, the decrease in crops from natural plants following crop increase from cultivation. *Bulletin of the Tohoku Regional Fisheries Research Laboratory* **44**:91-100.
- Alvarez, M. V. & Boraso, A. 2020. Development morphology of *Undaria pinnatifida* sporophytes (Phaeophyceae, Alariaceae) in Caleta Cordova (Chubut, Argentina). *Marine and Fishery Sciences* **33**:77-94.
- Andersen, R. A. 2004. Biology and systematics of heterokont and haptophyte algae. *American Journal of Botany* **91**:1508-22.
- Apoya, M., Ogawa, H. & Nanba, N. 2005. Alginate content of farmed *Undaria pinnatifida* (Harvey) Suringar from the Three Bays of Iwate, Japan during harvest period. *Botanica Marina* **45**:445-52.
- Barbosa, M., Fernandes , F., Pereira , D. M., Azevedo, I. C., Sousa-Pinto, I., Andrade , P. B. & Valentão , P. 2020. Fatty acid patterns of the kelps *Saccharina latissima*, *Saccorhiza polyschides* and *Laminaria ochroleuca*: Influence of changing environmental conditions. *Arabian Journal of Chemistry* **13**:45-58.
- Bastian, F. B., Roux, J., Niknejad, A., Comte, A., Fonseca Costa, S. S., de Farias, T. M., Moretti, S., Parmentier, G., de Laval, V. R., Rosikiewicz, M., Wollbrett, J., Echchiki, A., Escoriza, A., Gharib, W. H., Gonzales-Porta, M., Jarosz, Y., Laurenczy, B., Moret, P., Person, E., Roelli, P., Sanjeev, K., Seppey, M. & Robinson-Rechavi, M. 2021. The Bgee suite: integrated curated expression atlas and comparative transcriptomics in animals. *Nucleic Acids Research* **49**:D831-D47.
- Bentz, A. B., Thomas, G. W. C., Rusch, D. B. & Rosvall, K. A. 2019. Tissue-specific expression profiles and positive selection analysis in the tree swallow (*Tachycineta bicolor*) using a de novo transcriptome assembly. *Scientific Reports* **9**:15849.
- Berbee, M. L. & Taylor, J. W. 2010. Dating the molecular clock in fungi - how close are we? *Fungal Biology Reviews* **24**:1-16.
- Bolger, A. M., Lohse, M. & Usadel, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**:2114-20.
- Charrier, B., Le Bail, A. & de Reviers, B. 2012. Plant Proteus: brown algal morphological plasticity and underlying developmental mechanisms. *Trends in Plant Science* **17**:468-77.
- Charrier, B., Rabillé, H. & Billoud, B. 2019. Gazing at cell wall expansion under a golden light. *Trends in Plant Science* **24**:130-41.
- Chi, S., Liu, T., Wang, X., Wang, R., Wang, S., Wang, G., Shan, G. & Liu, C. 2018. Functional genomics analysis reveals the biosynthesis pathways of important cellular components (alginate and fucoidan) of *Saccharina*. *Current Genetics* **64**:259-73.
- Cock, J. M., Sterck, L., Rouzé, P., Scornet, D., Allen, A. E., Amoutzias, G., Anthouard, V., Artiguenave, F., Aury, J.-M., Badger, J. H., Beszteri, B., Billiau, K., Bonnet, E., Bothwell, J. H., Bowler, C., Boyen, C., Brownlee, C., Carrano, C. J., Charrier, B., Cho,

- G. Y., Coelho, S. M., Collén, J., Corre, E., Da Silva, C., Delage, L., Delaroque, N., Dittami, S. M., Doulbeau, S., Elias, M., Farnham, G., Gachon, C. M. M., Gschloessl, B., Heesch, S., Jabbari, K., Jubin, C., Kawai, H., Kimura, K., Kloareg, B., Küpper, F. C., Lang, D., Le Bail, A., Leblanc, C., Lerouge, P., Lohr, M., Lopez, P. J., Martens, C., Maumus, F., Michel, G., Miranda-Saavedra, D., Morales, J., Moreau, H., Motomura, T., Nagasato, C., Napoli, C. A., Nelson, D. R., Nyvall-Collén, P., Peters, A. F., Pommier, C., Potin, P., Poulain, J., Quesneville, H., Read, B., Rensing, S. A., Ritter, A., Rousvoal, S., Samanta, M., Samson, G., Schroeder, D. C., Ségurens, B., Strittmatter, M., Tonon, T., Tregear, J. W., Valentin, K., von Dassow, P., Yamagishi, T., Van de Peer, Y. & Wincker, P. 2010. The Ectocarpus genome and the independent evolution of multicellularity in brown algae. *Nature* **465**:617-21.
- Dean, P. R. & Hurd, C. L. 2007. Seasonal growth, erosion rates, and nitrogen and photosynthetic ecophysiology of *Undaria pinnatifida* (Heterokontophyta) in southern New Zealand. *Journal of Phycology* **43**:1138-48.
- Deng, Y., Yao, J., Wang, X., Guo, H. & Delin, D. 2012. Transcriptome sequencing and comparative analysis of *Saccharina japonica* (Laminariales, Phaeophyceae) under blue light induction. *PloS one* **7**:e39704.
- Deniaud-Bouët, E., Hardouin, K., Potin, P., Kloareg, B. & Hervé, C. 2017. A review about brown algal cell walls and fucose-containing sulfated polysaccharides: Cell wall context, biomedical properties and key research challenges. *Carbohydrate polymers* **175**:395-408.
- Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T. R. 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**:15-21.
- Dong, J.-Z. & Dunstan, D. I. 1997. Endochitinase and β-1,3-glucanase genes are developmentally regulated during somatic embryogenesis in *Picea glauca*. *Planta* **201**:189-94.
- Epstein, G. & Smale, D. A. 2017. *Undaria pinnatifida*: A case study to highlight challenges in marine invasion ecology and management. *Ecology and Evolution* **7**:8624-42.
- Fernandez-Pozo, N., Haas, F. B., Meyberg, R., Ullrich, K. K., Hiss, M., Perroud, P.-F., Hanke, S., Kratz, V., Powell, A. F., Vestey, E. F., Daum, C. G., Zane, M., Lipzen, A., Sreedasyam, A., Grimwood, J., Coates, J. C., Barry, K., Schmutz, J., Mueller, L. A. & Rensing, S. A. 2020. PEATmoss (*Physcomitrella* Expression Atlas Tool): a unified gene expression atlas for the model plant *Physcomitrella patens*. *the plant journal* **102**:165-77.
- Forrest, B. M., Brown, S. N., Taylor, M. D., Hurd, C. L. & Hay, C. H. 2000. The role of natural dispersal mechanisms in the spread of *Undaria pinnatifida* (Laminariales, Phaeophyceae). *Phycologia* **39**:547-53.
- Fritsch, F. E. 1935. *The structure and reproduction of the algae*. University Press, Cambridge,
- Fu, G., Nagasato, C., Oka, S., Cock, J. M. & Motomura, T. 2014. Proteomics analysis of heterogeneous flagella in brown algae (Stramenopiles). *Protist* **165**:662-75.
- Graf, L., Shin, Y., Yang, J. H., Choi, J. W., Hwang, I. K., Nelson, W., Bhattacharya, D., Viard, F. & Yoon, H. S. 2021. A genome-wide investigation of the effect of farming and human-mediated introduction on the ubiquitous seaweed *Undaria pinnatifida*. *Nature Ecology & Evolution* **5**:360-68.
- Henry, E. C. & Cole, K. M. 1982. Ultrastructure of swarmers in the Laminariales (Phaeophyceae). i. zoospores. *Journal of Phycology* **18**:550-69.
- Hiroyuki, M. & Hajime, Y. 2010. Significance of radical oxygen production in sorus development and zoospore germination in *Saccharina japonica* (Phaeophyceae). **53**:409-16.
- Hounkpe, B. W., Chenou, F., de Lima, F. & De Paula, Erich V. 2021. HRT Atlas v1.0 database: redefining human and mouse housekeeping genes and candidate reference transcripts by mining massive RNA-seq datasets. *Nucleic Acids Research* **49**:D947-D55.

- Huysman, M. J. J., Fortunato, A., E., Matthijs, M., Costa, B. S., Vanderhaeghen, R., Van den Daele, H., Sachse, M., Inzé, D., Bowler, C., Kroth, P. G., Wilhelm, C., Falciatore, A., Vyverman, W. & De Veylder, L. 2013. AUREOCHROME1a-mediated induction of the diatom-specific cyclin dsCYC2 controls the onset of cell division in diatoms (*Phaeodactylum tricornutum*). *The Plant cell* **25**:215-28.
- Inoue, A. & Ojima, T. 2019. Functional identification of alginate lyase from the brown alga *Saccharina japonica*. *Scientific Reports* **9**:4937.
- Jia, L. J. J., Jace, K., Rong, M. J., Fajardo, V. E. M., Anindya, P. D., Shan, L. Y., Song, S. W., Sriya, M., Zhen, N. J. W., Uyen, N. N. L., Rinta, R., Herman, F., Xuan, Z. M., Ling, C. T., Edric, R. E., Singh, K. R., Min, H. K., Cassandra, C. N., Don, L. A., Xin, C. Z. H., Chun, L. S., Vanessa, L., Choy, F. T., Xin'Er, C. C., Sween, K. K. T., Irene, J., Riccardo, D.-P., Xiong, N. J. W. & Marek, M. 2020. Fungi.guru: Comparative genomic and transcriptomic resource for the fungi kingdom. *Computational and structural biotechnology journal* **18**:3788-95.
- Kai, T., Nimura, K., Yasui, H. & Mizuta, H. 2006. Regulation of sorus formation by auxin in Laminariales sporophyte. *Journal of Applied Phycology* **18**:95.
- Kawai, H., Hanyuda, T., Draisma, S. G. A., Wilce, R. T. & Andersen, R. A. 2015. Molecular phylogeny of two unusual brown algae, *Phaeostrophion irregulare* and *Platysiphon glacialis*, proposal of the Stschapoviales ord. nov. and Platysiphonaceae fam. nov., and a re-examination of divergence times for brown algal orders. *Journal of phycology* **51**:918-28.
- Kloareg, B. & Quatrano, R. S. 1988. Structure of the cell walls of marine algae and ecophysiological functions of the matrix polysaccharides. *Oceanography and Marine Biology* **26**:259-315.
- Knoll, A. H. 2011. The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences* **39**:217-39.
- Kudapa, H., Garg, V., Chitikineni, A. & Varshney, R. K. 2018. The RNA-Seq-based high resolution gene expression atlas of chickpea (*Cicer arietinum* L.) reveals dynamic spatio-temporal changes associated with growth and development. *Plant, Cell & Environment* **41**:2209-25.
- Kumar, A., Castellano, I., Patti, F. P., Delledonne, M., Abdelgawad, H., Beemster, G. T. S., Asard, H., Palumbo, A. & Buia, M. C. 2017. Molecular response of *Sargassum vulgare* to acidification at volcanic CO<sub>2</sub> vents: insights from de novo transcriptomic analysis. *Molecular Ecology* **26**:2276-90.
- Küppers, U. & Kremer, B. P. 1978. Longitudinal profiles of carbon dioxide fixation capacities in marine macroalgae. *Plant Physiology* **62**:49-53.
- Langfelder, P. & Horvath, S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**:559.
- Li, H., Scheschonk, L., Heinrich, S., Valentin, K., Harms, L., Glöckner, G., Corre, E. & Bischof, K. 2020. Transcriptomic responses to darkness and the survival strategy of the kelp *Saccharina latissima* in the early polar night. *Frontiers in Marine Science* **7**:592033.
- Liu, F., Wang, W., Sun, X., Liang, Z. & Wang, F. 2013. RNA-Seq revealed complex response to heat stress on transcriptomic level in *Saccharina japonica* (Laminariales, Phaeophyta). *Journal of Applied Phycology* **26**:1585-96.
- Liu, P. & Si, Y. 2014. Cluster Analysis of RNA-Sequencing Data. In: Datta, S. & Nettleton, D. [Eds.] *Statistical analysis of next generation sequencing data*. Springer, Cham, pp. 191-217.
- Lu, C., Shao, Z., Zhang, P. & Duan, D. 2020. Genome-wide analysis of the *Saccharina japonica* sulfotransferase genes and their transcriptional profiles during whole developmental periods and under abiotic stresses. *BMC Plant Biology* **20**:271.

- Lüning, K., Schmitz, K. & Willenbrink, J. 1973. CO<sub>2</sub> fixation and translocation in benthic marine algae. III. Rates and ecological significance of translocation in *Laminaria hyperborea* and *L. saccharina*. *Marine Biology* **23**:275-81.
- Machado, F. B., Moharana, K. C., Almeida-Silva, F., Gazara, R. K., Pedrosa-Silva, F., Coelho, F. S., Grativil, C. & Venancio, T. M. 2020. Systematic analysis of 1298 RNA-Seq samples and construction of a comprehensive soybean (*Glycine max*) expression atlas. *The Plant Journal* **103**:1894-909.
- Machado Monteiro, C. M., Li, H., Bischof, K., Bartsch, I., Valentim, K. U., Corre, E., Collén, J., Harms, L., Glöckner, G. & Heinrich, S. 2019. Is geographical variation driving the transcriptomic responses to multiple stressors in the kelp *Saccharina latissima*? *BMC Plant Biology* **19**:513.
- Michel, G., Tonon, T., Scornet, D., Cock, M. J. & Bernard, K. 2010. The cell wall polysaccharide metabolism of the brown alga *Ectocarpus siliculosus*. Insights into the evolution of extracellular matrix polysaccharides in Eukaryotes. *New Phytologist* **188**:82-97.
- Morita, T., Kurashima, A. & Maegawa, M. 2003. Temperature requirements for the growth and maturation of the gametophytes of *Undaria pinnatifida* and *U. undarioides* (Laminariales, Phaeophyceae). *Phycological Research* **51**:154-60.
- Moulager, M., Monnier, A., Jesson, B., Bouvet, R., Mosser, J., Schwartz, C., Garnier, L., Corellou, F. & Bouget, F.-Y. 2007. Light-dependent regulation of cell division in *Ostreococcus*: evidence for a major transcriptional input. *Plant physiology* **144**:1360-69.
- Munakata, H., Nakada, T., Nakahigashi, K., Nozaki, H. & Tomita, M. 2016. Phylogenetic position and molecular chronology of a colonial green flagellate, *Stephanosphaera pluvialis* (Volvocales, Chlorophyceae), among unicellular algae. *The Journal of eukaryotic microbiology* **63**:340-48.
- Nimura, K. & Mizuta, H. 2001. Differences in photosynthesis and nucleic acid content between sterile and fertile parts of the sporophyte of *Laminaria japonica* (Phaeophyceae). *Algae* **16**:151-55.
- Nitschke, U., Ruth, A. A., Dixneuf, S. & Stengel, D. B. 2011. Molecular iodine emission rates and photosynthetic performance of different thallus parts of *Laminaria digitata* (Phaeophyceae) during emersion. *Planta* **233**:737-48.
- Ohno, M. & Matsuoka, M. 1993. *Undaria* cultivation "Wakame". In: Ohno, M. & Critchley, A. T. [Eds.] *Seaweed cultivation and marine ranching*. Japanese International Cooperation Agency (JICA), Yokosuka, pp. 41-50.
- Olischläger, M., Bartsch, I., Gutow, L. & Wiencke, C. 2012. Effects of ocean acidification on different life-cycle stages of the kelp *Laminaria hyperborea* (Phaeophyceae). *Botanica Marina* **55**:511-25.
- Pang, S. & Lüning, K. 2004. Photoperiodic long-day control of sporophyll and hair formation in the brown alga *Undaria pinnatifida*. *Journal of Applied Phycology* **16**:83-92.
- Paul, S., Salavarria, E., Gil-Kodaka, P. & Villena, G. K. 2020. A de novo transcriptomic approach to study the influence of marine water depth in *Macrocystis pyrifera* alginate production. *Aquatic Botany* **163**:103211.
- Peters, A. F., Scornet, D., Ratin, M., Charrier, B., Monnier, A., Merrien, Y., Corre, E., Coelho, S. M. & Cock, J. M. 2008. Life-cycle-generation-specific developmental processes are modified in the immediate upright mutant of the brown alga *Ectocarpus siliculosus*. *Development* **135**:1503-12.
- Peterson, K. J. & Butterfield, N. J. 2005. Origin of the Eumetazoa: Testing ecological predictions of molecular clocks against the Proterozoic fossil record. *Nature* **436**:9547-52.
- R Core Team 2020. R: A language and environment for statistical computing. *R foundation for Statistical Computing* <https://www.R-project.org/>.

- Rabillé, H., Torode, T. A., Tesson, B., Le Bail, A., Billoud, B., Rolland, E., Le Panse, S., Jam, M. & Charrier, B. 2019. Alginates along the filament of the brown alga *Ectocarpus* help cells cope with stress. *Scientific Reports* **9**:12956.
- Ritter, A., Cabioch, L., Brillet-Guéguen, L., Corre, E., Cosse, A., Darteville, L., Duruflé, H., Fasshauer, C., Goulitquer, S., Thomas, F., Correa, J. A., Potin, P., Faugeron, S. & Leblanc, C. 2017. Herbivore-induced chemical and molecular responses of the kelps *Laminaria digitata* and *Lessonia spicata*. *PLOS One* **12**:e0173315.
- Robinson, M. D., McCarthy, D. J. & Smyth, G. K. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics (Oxford, England)* **26**:139-40.
- Rugiu, L., Panova, M., Pereyra, R. T. & Jormalainen, V. 2020. Gene regulatory response to hypersalinity in the brown seaweed *Fucus vesiculosus*. *BMC Genomics* **21**:42.
- Salavarria, E., Paul, S., Gil-Kodaka, P. & Villena, G. K. 2018. First global transcriptome analysis of brown algae *Macrocystis integrifolia* (Phaeophyceae) under marine intertidal conditions. *3 Biotech* **8**:185.
- Schaffelke, B., Campbell, M. L. & Hewitt, C. L. 2019. Reproductive phenology of the introduced kelp *Undaria pinnatifida* (Phaeophyceae, Laminariales) in Tasmania, Australia. *Phycologia* **44**:84-94.
- Schmid, M. & Stengel, D. B. 2015. Intra-thallus differentiation of fatty acid and pigment profiles in some temperate Fucales and Laminariales. *Journal of Phycology* **51**:25-36.
- Schmitz, K. & Srivastava, L. M. 1976. The fine structure of sieve elements of *Nereocystis lütkeana*. *American Journal of Botany* **63**:679-93.
- Shan, T., Li, Q., Wang, X. & Pang, S. 2020a. Full-length transcriptome sequencing and comparative transcriptomic analysis of different developmental stages of the sporophyll in *Undaria pinnatifida* (Laminariales: Alariaceae). *Journal of Applied Phycology* **32**:2081-92.
- Shan, T., Yuan, J., Su, L., Li, J., Leng, X., Zhang, Y., Gao, H. & Pang, S. 2020b. First genome of the brown alga *Undaria pinnatifida*: Chromosome-level assembly using PacBio and Hi-C technologies. *Frontiers in Genetics* **11**:140.
- Shao, Z., Zhang, P., Lu, C., Li, S., Chen, Z., Wang, X. & Duan, D. 2019. Transcriptome sequencing of *Saccharina japonica* sporophytes during whole developmental periods reveals regulatory networks underlying alginate and mannitol biosynthesis. *BMC Genomics* **20**:975.
- Shi, C., Kataoka, H. & Duan, D. 2005. Effects of blue light on gametophyte development of *Laminaria japonica* (Laminariales, Phaeophyta). *Chinese Journal of Oceanology and Limnology* **23**:323-29.
- Silberfeld, T., Leigh, J. W., Verbruggen, H., Cruaud, C., de Reviers, B. & Rousseau, F. 2010. A multi-locus time-calibrated phylogeny of the brown algae (Heterokonta, Ochrophyta, Phaeophyceae): Investigating the evolutionary nature of the "brown algal crown radiation". *Molecular phylogenetics and evolution* **56**:659-74.
- Skripsova, A. V. 2015. Fucoidans of brown algae: Biosynthesis, localization, and physiological role in thallus. *Russian Journal of Marine Biology* **41**:145-56.
- Skripsova, A. V., Shevchenko, N. M., Zvyagintseva, T. N. & Imbs, T. I. 2010. Monthly changes in the content and monosaccharide composition of fucoidan from *Undaria pinnatifida* (Laminariales, Phaeophyta). *Journal of Applied Phycology* **22**:79-86.
- Starko, S. & Martone, P. T. 2016. An empirical test of 'universal' biomass scaling relationships in kelps: evidence of convergence with seed plants. *New Phytologist* **212**:719-29.
- Wang, W.-J., Sun, X.-T. & Wang, F.-J. 2010. Effect of blue light on early sporophyte development of *Saccharina japonica* (Phaeophyta). *Marine Biology* **157**:1811-17.
- Wang, Y., Xu, D., Fan, X., Zhang, X., Ye, N., Wang, W., Mao, Y., Mou, S. & Cao, S. 2013. Variation of photosynthetic performance, nutrient uptake, and elemental composition of

- different generations and different thallus parts of *Saccharina japonica*. *Journal of Applied Phycology* **25**:631-37.
- Xanthopoulou, A., Montero-Pau, J., Picó, B., Boumpas, P., Tsaliki, E., Paris, H. S., Tsafaris, A., Kalivas, A., Mellidou, I. & Ganopoulos, I. 2021. A comprehensive RNA-Seq-based gene expression atlas of the summer squash (*Cucurbita pepo*) provides insights into fruit morphology and ripening mechanisms. *BMC Genomics* **22**:341.
- Xing, Q., Rousvoal, S. & Leblanc, C. 2020. Transcriptome-wide identification and evaluation of optimal reference genes for RT-qPCR expression analysis of *Saccharina latissima* responses to biotic and abiotic stress. *Journal of Applied Phycology* **33**:617-27.
- Yang, E. C., Boo, G. H., Kim, H. J., Cho, S. M., Boo, S. M., Andersen, R. A. & Yoon, H. S. 2012. Supermatrix data highlight the phylogenetic relationships of photosynthetic Stramenopiles. *Protist* **163**:217-31.
- Ye, N., Zhang, X., Miao, M., Fan, X., Zheng, Y., Xu, D., Wang, J., Zhou, L., Wang, D., Gao, Y., Wang, Y., Shi, W., Ji, P., Li, D., Guan, Z., Shao, C., Zhuang, Z., Gao, Z., Qi, J. & Zhao, F. 2015. Saccharina genomes provide novel insight into kelp biology. *Nature Communications* **6**:6986.
- Zhang, L., Cao, Z., Liang, G., Li, X., Wu, H. & Yang, G. 2020. Comparative transcriptome analysis reveals candidate genes related to structural and storage carbohydrate biosynthesis in kelp *Saccharina japonica* (Laminariales, Phaeophyceae). *Journal of Phycology* **56**:1168-83.
- Zhang, P., Lu, C., Shao, Z., Liu, F., Yao, J. & Duan, D. 2021a. Genome-wide transcriptome profiling and characterization of mannuronan C5-epimerases in *Saccharina japonica*. *Algal Research* **60**:102491.
- Zhang, X., Fan, X., Wang, Y., Xu, D., Zhang, J. & Ye, N. 2021b. Exploring core response mechanisms to multiple environmental stressors via a genome-wide study in the brown alga *Saccharina japonica* (Laminariales, Phaeophyceae). *Journal of Phycology* **57**:345-54.
- Zhang, X., Li, D., Hu, H. & Tan, T. 2005. Growth promotion of vegetative gametophytes of *Undaria pinnatifida* by blue light. *Biotechnology letters* **27**:1467-75.

## Table

**Table 1.** Gene expression specificity within the sporophyte of *Undaria pinnatifida* and during its development.

Sampling date	Tissue	No. of expressed genes	No. of uniquely expressed genes	No. of uniquely tissue specific genes ( <i>Tau</i> )	No. of tissue specific genes ( <i>Tau</i> )	TPM < 10	TPM > 100
November	Undifferentiated	16,819	80	263	272	10,218	1,005
(vegetative sporophyte)	Sporophyll	18,152	103	353	384	9,768	1,093
	Meristem	16,569	32	84	92	9,684	962
	Healthy Blade	16,779	42	107	113	9,934	1,042
	Apex blade	16,311	32	113	120	9,500	1,151

April (reproductive sporophyte)	Sporophyll	18,595	366	1,493	1,524	8,321	1,302
	Meristem	16,935	34	86	94	9,037	1,133
	Healthy Blade	16,660	27	73	81	9,389	1,183
	Apex blade	16,551	27	84	90	9,043	1,350

## Figure Legends

**Figure 1.** Summary of the sampling strategy to explore the gene expression in different tissues within the sporophyte and its developmental stages of *Undaria pinnatifida*. The pictures of the different tissues are representative and are voluntarily larger than the one used for the RNA sequencing.

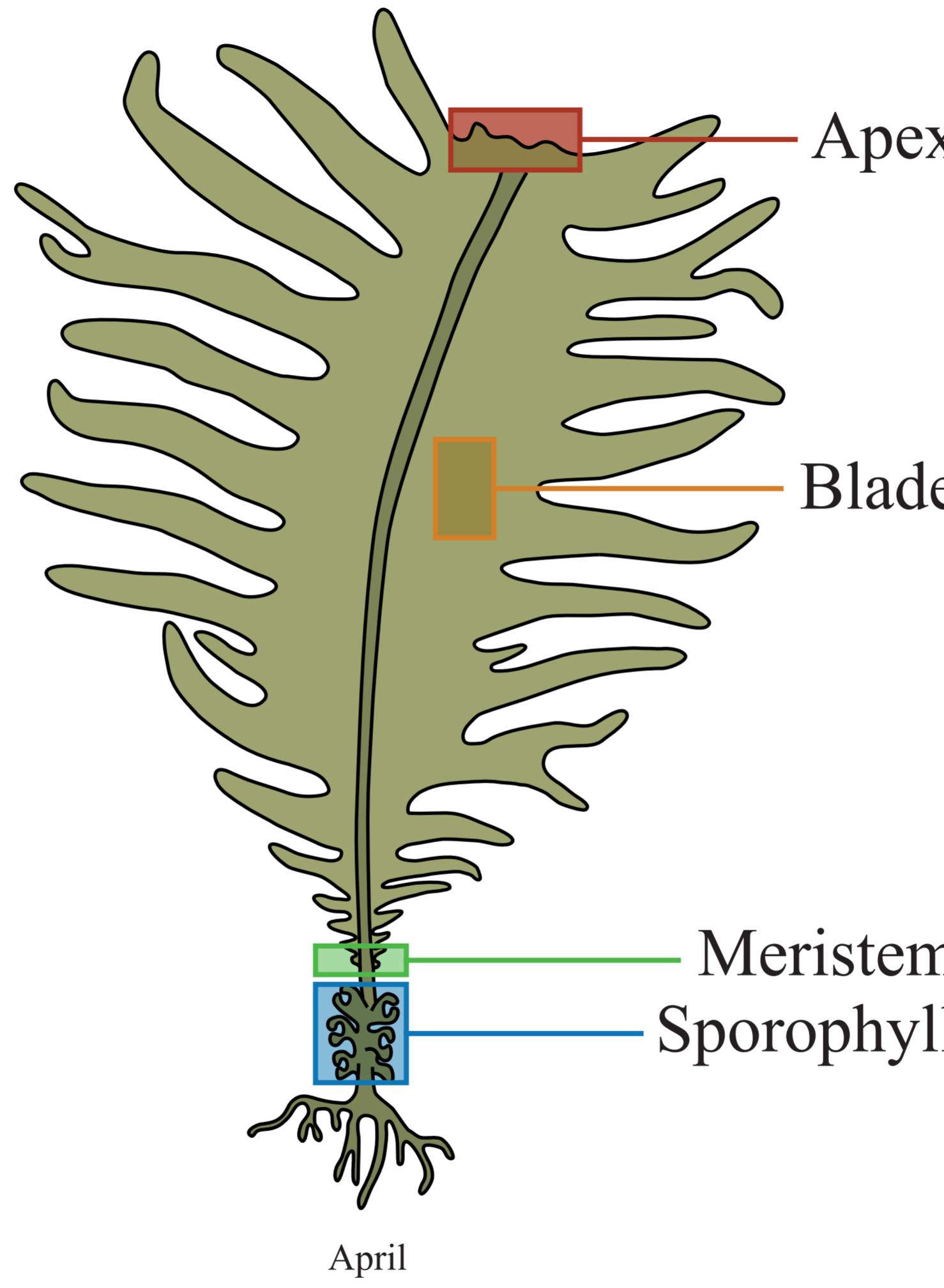
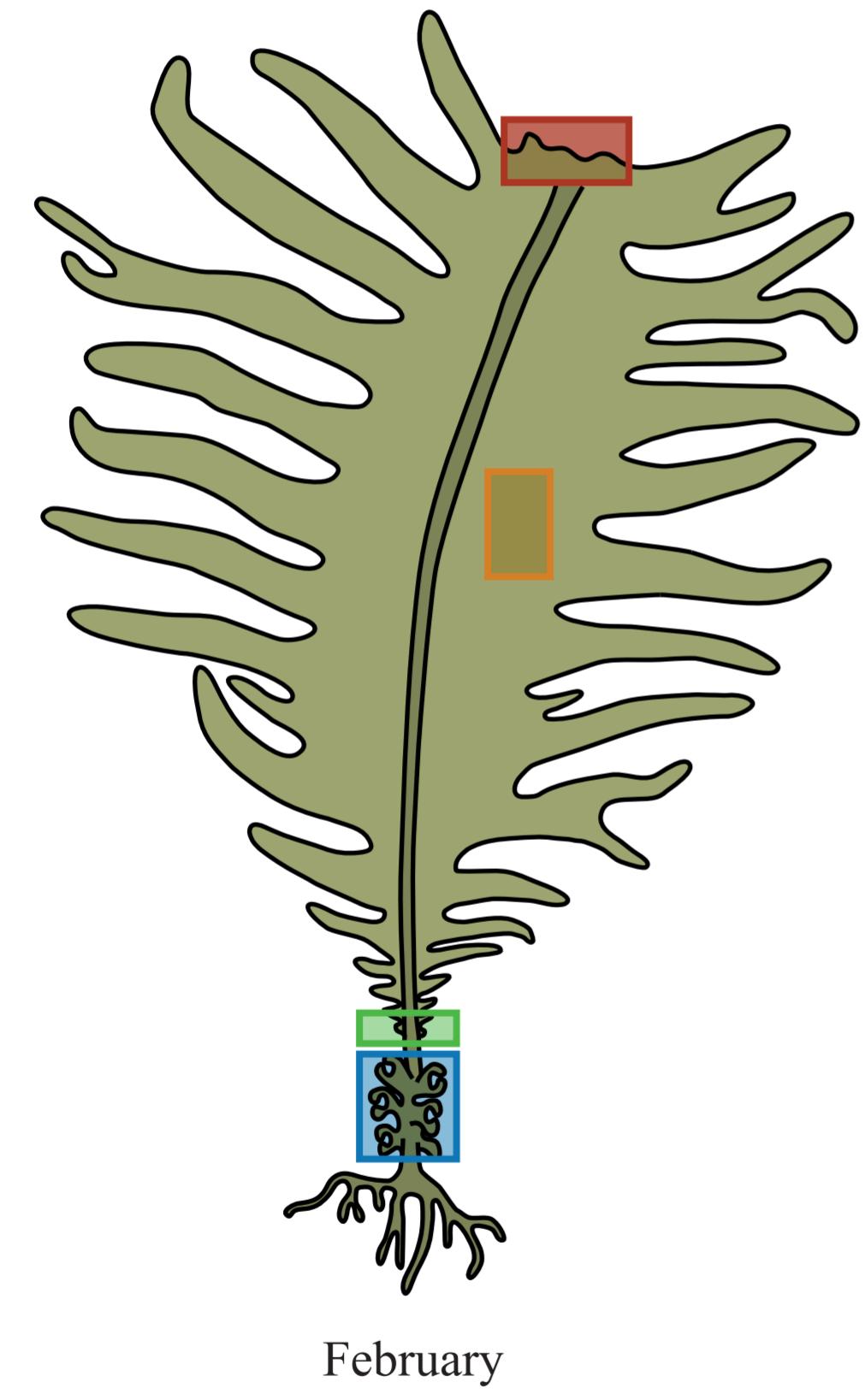
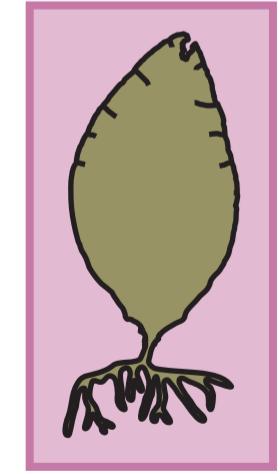
**Figure 2.** Exploratory analysis of the expression in the RNA sequencing of the sporophyte of *Undaria pinnatifida*. **a** Principal components analysis based on the transcript count per kilobase (TPM) from the 27 RNA libraries. **b** Correlation between the count of the transcript count per kilobase (TPM) of the 27 RNA libraries. **c** Comparison of the expression profile in the different tissues of the sporophyte of *Undaria pinnatifida*. Repro. s. = reproductive sporophyte (Apr.); Vege. s. = vegetative sporophyte (Feb.); sporo. = sporophyte; V.s. S. = Vegetative sporophyte sporophyll (Feb.); V.s. M. = Vegetative sporophyte meristem (Feb.); V.s. B. = Vegetative sporophyte healthy blade (Feb.); V.s. A. = Vegetative sporophyte apex blade (Feb.); R.s. S. = Reproductive sporophyte sporophyll (Apr.); R.s. M. = Reproductive sporophyte meristem (Apr.); R.s. B. = Reproductive sporophyte healthy blade (Apr.); R.s. A. = Reproductive sporophyte apex blade (Apr.).

**Figure 3.** Gene expression profiles and differentially expressed genes within the sporophyte and its developmental stages of *Undaria pinnatifida*. **a** UpSet plot representing the number of genes expressed in the a single tissue or in a combination of tissues. On top are the total number of genes expressed in each tissue and the intersection size is in number of genes expressed. **b**

Heatmap of the expression profiles of the nine tissues. Expression of genes is colored according to their level from high (red) to low (blue). Tissues are presented on the x-axis and genes on the y-axis. Genes were clustered on the y-axis using hierarchical clustering. c Number of differentially expressed genes identified in pairwise comparison within the sporophyte and during the developmental stage of the sporophyte of *Undaria pinnatifida*.

**Figure 4.** Summary of the weighted gene co-expression network analysis on the sporophyte transcriptome of *Undaria pinnatifida*. The seven clusters detected in the WGCNA are represented by a colored box linked to the position of the tissue(s) in which genes were over-expressed. In each box, a graphical summary of the genes identified in the cluster is presented. At the bottom, Z-transformed TPM distribution of the genes is presented for the three stages of the sporophyte. On top, heatmap map of the differentially expressed genes found in the cluster. Repro. s. = reproductive sporophyte (Apr.); Vege. s. = vegetative sporophyte (Feb.).

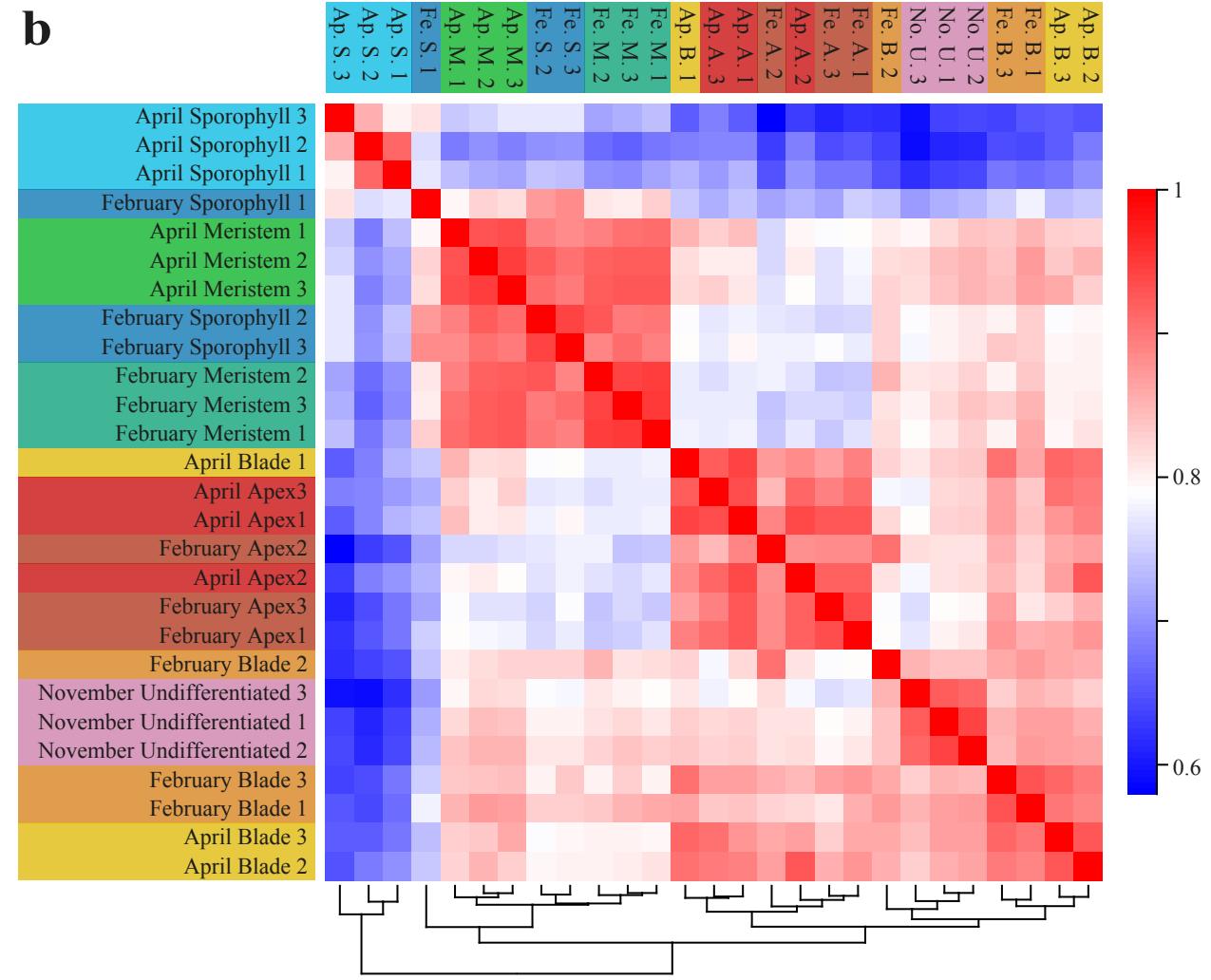
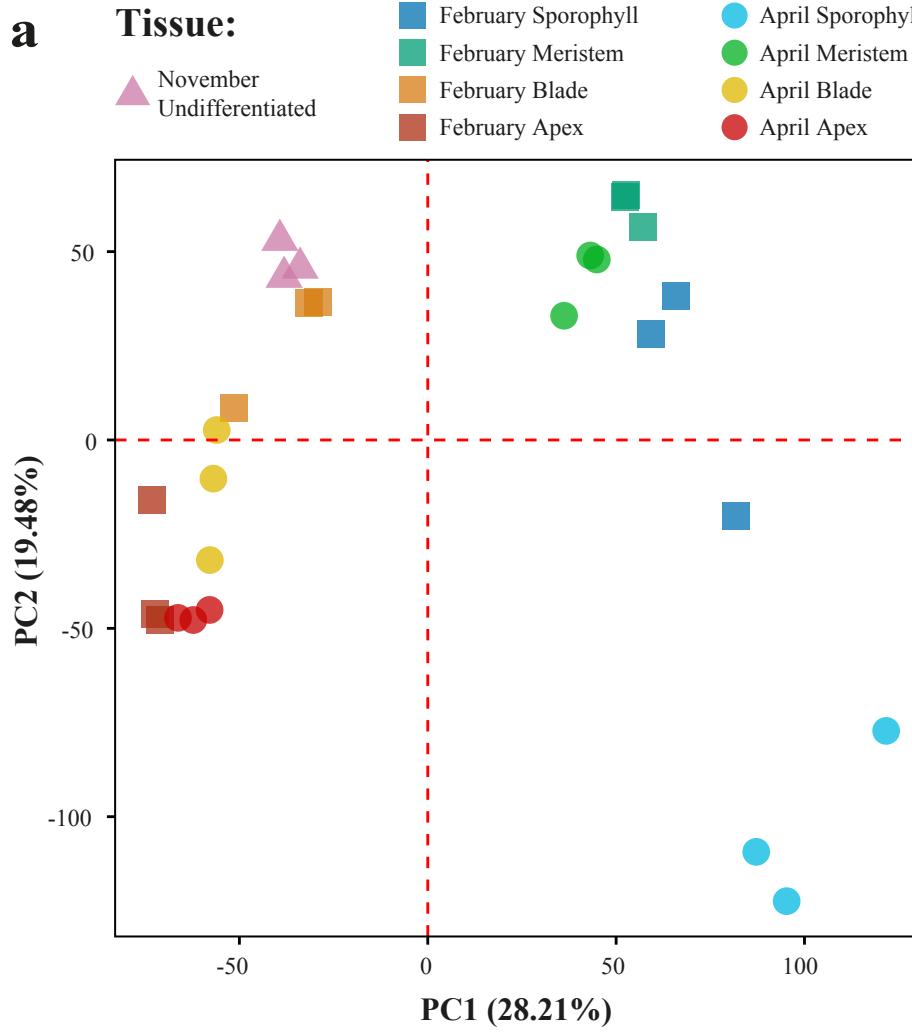
**Figure 5.** Summary of the biological functions expressed across the sporophyte of *Undaria pinnatifida*. The different tissues are on the x-axis and the genes are classified according to their GO biological process term annotation in the y-axis. Genes were clustered using hierarchical clustering within each GO biological process class, respectively.



November

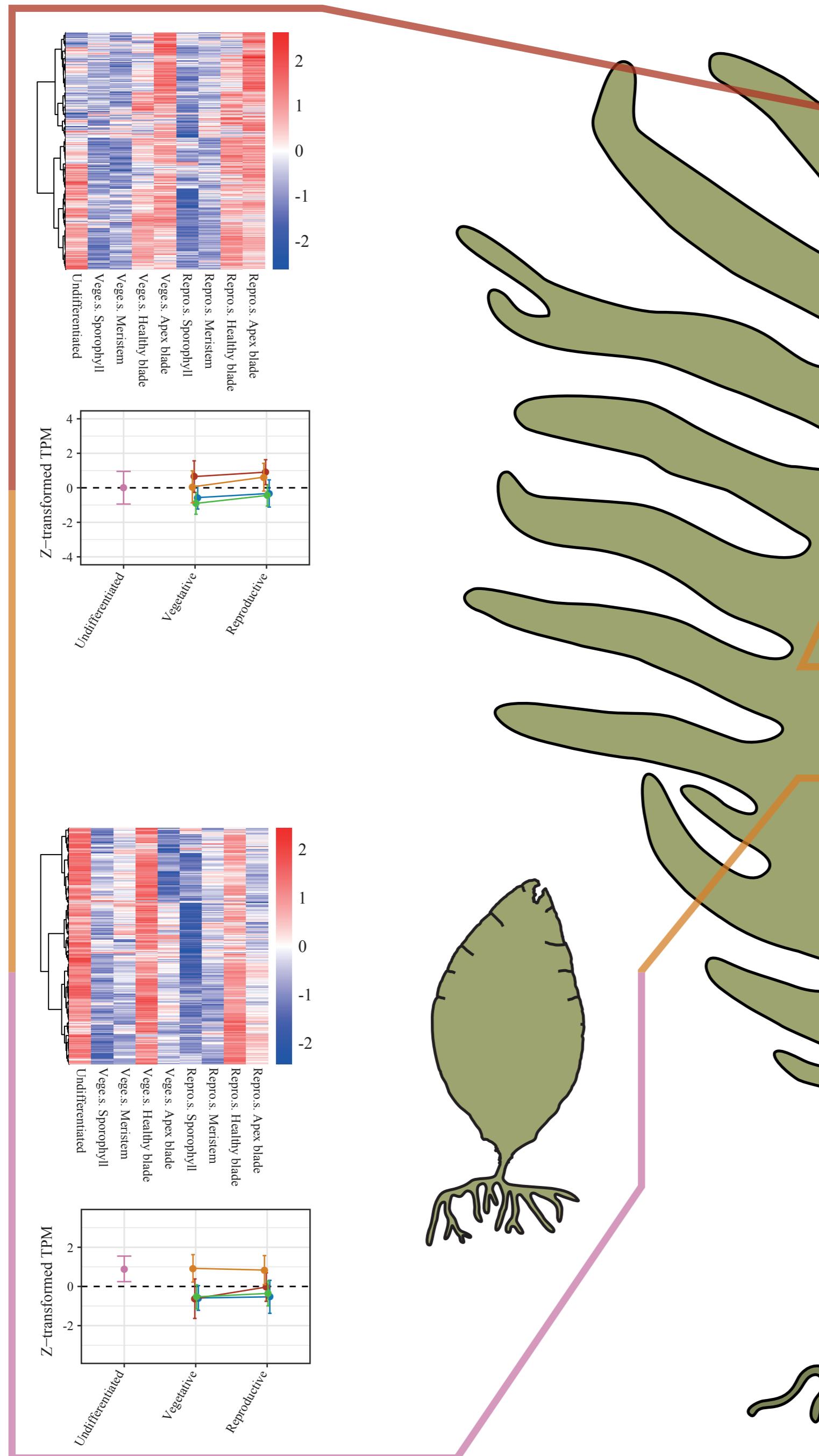
February

April

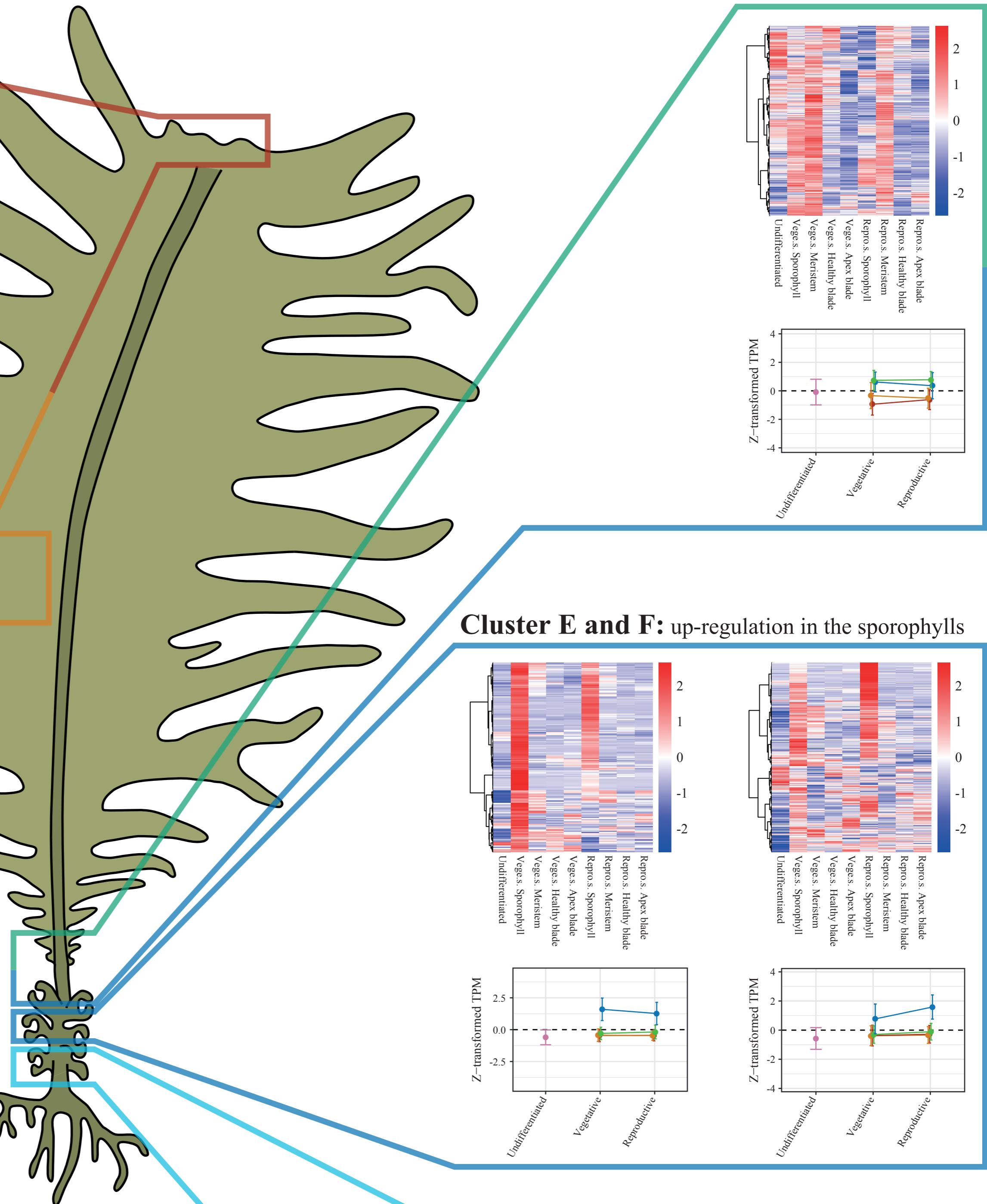


**a**

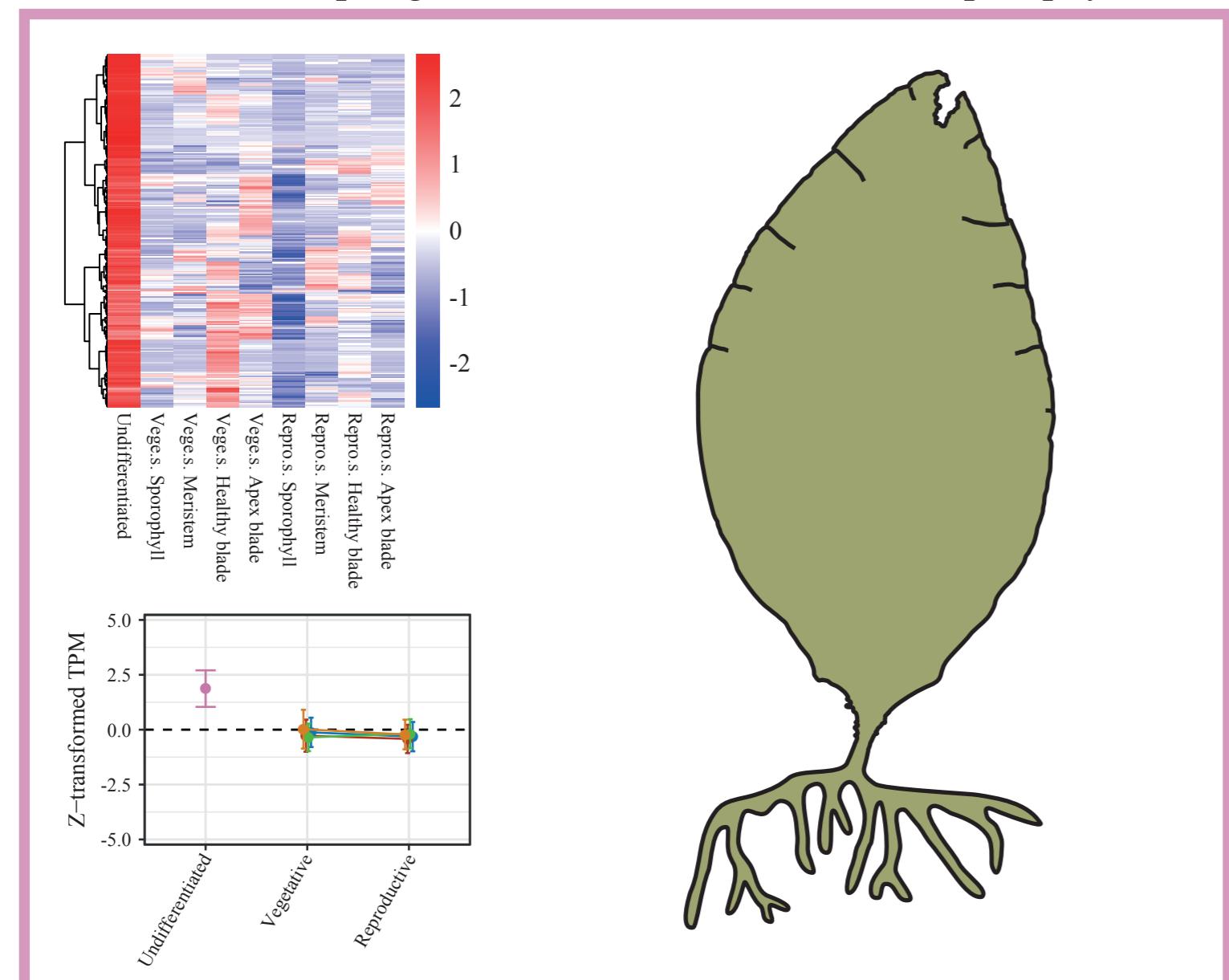
### **Cluster B and C:** up-regulation in the undifferentiated sporophyte and blade



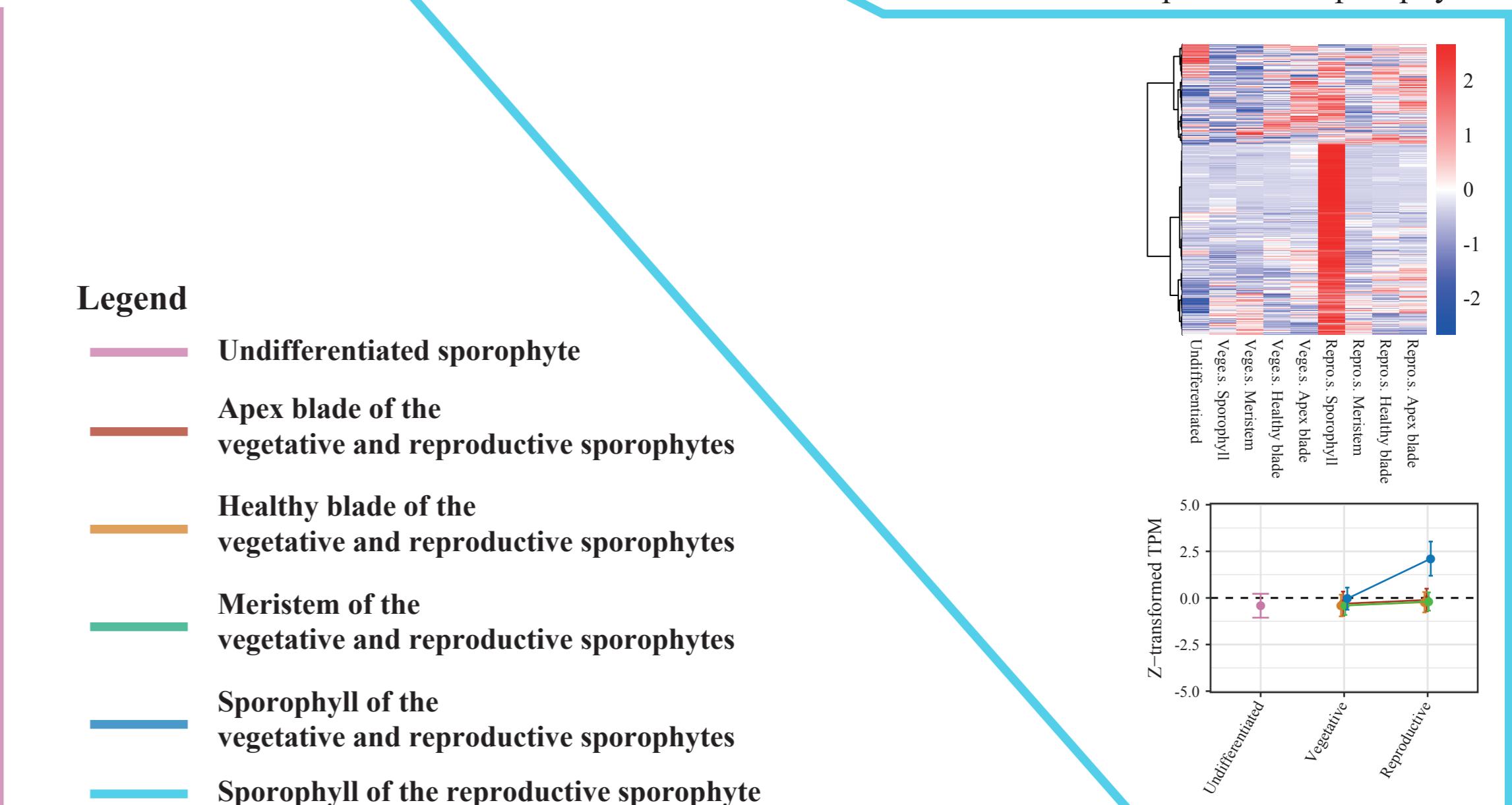
## **Cluster D:** up-regulation in the meristems and sporophylls

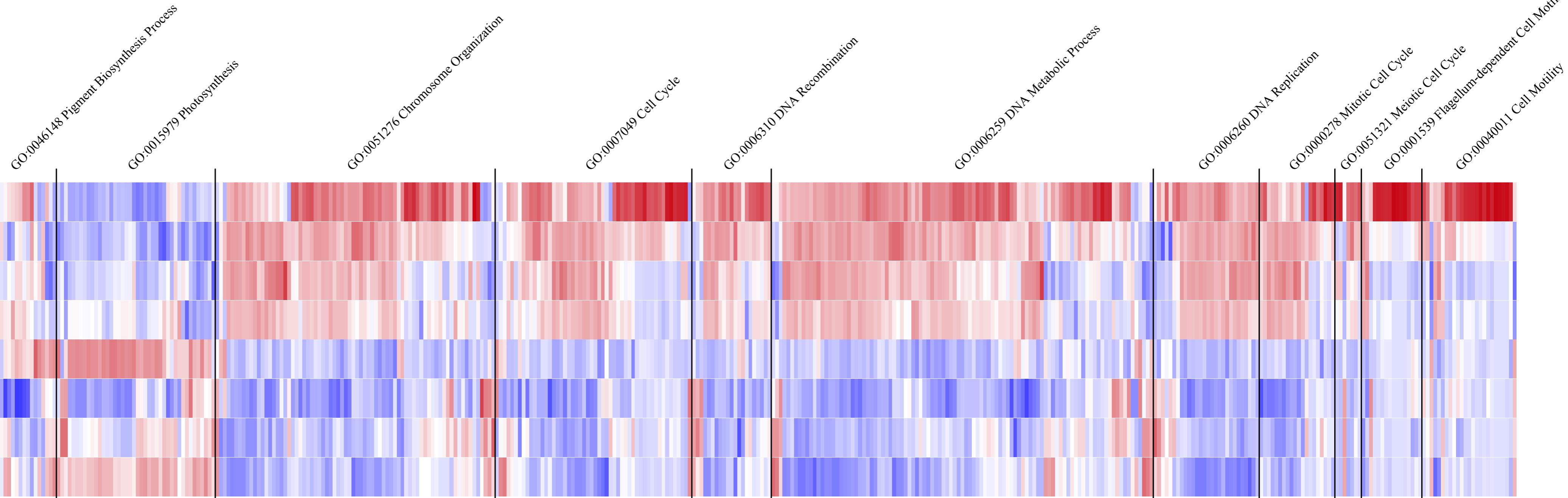
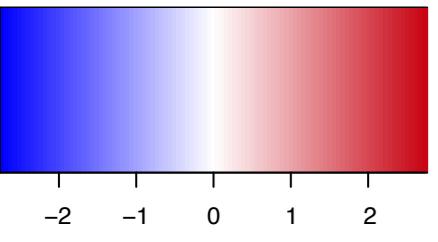


## **Cluster A:** up-regulation in the undifferentiated sporophyte



## **Cluster G:** up-regulation in the sporophylls of the reproductive sporophyte







**Figure S8:** Voucher specimen of an undifferentiated sporophyte of *Undaria pinnatifida* sampled on November 19, 2016; representative of the undifferentiated sporophytes used for RNA sequencing.

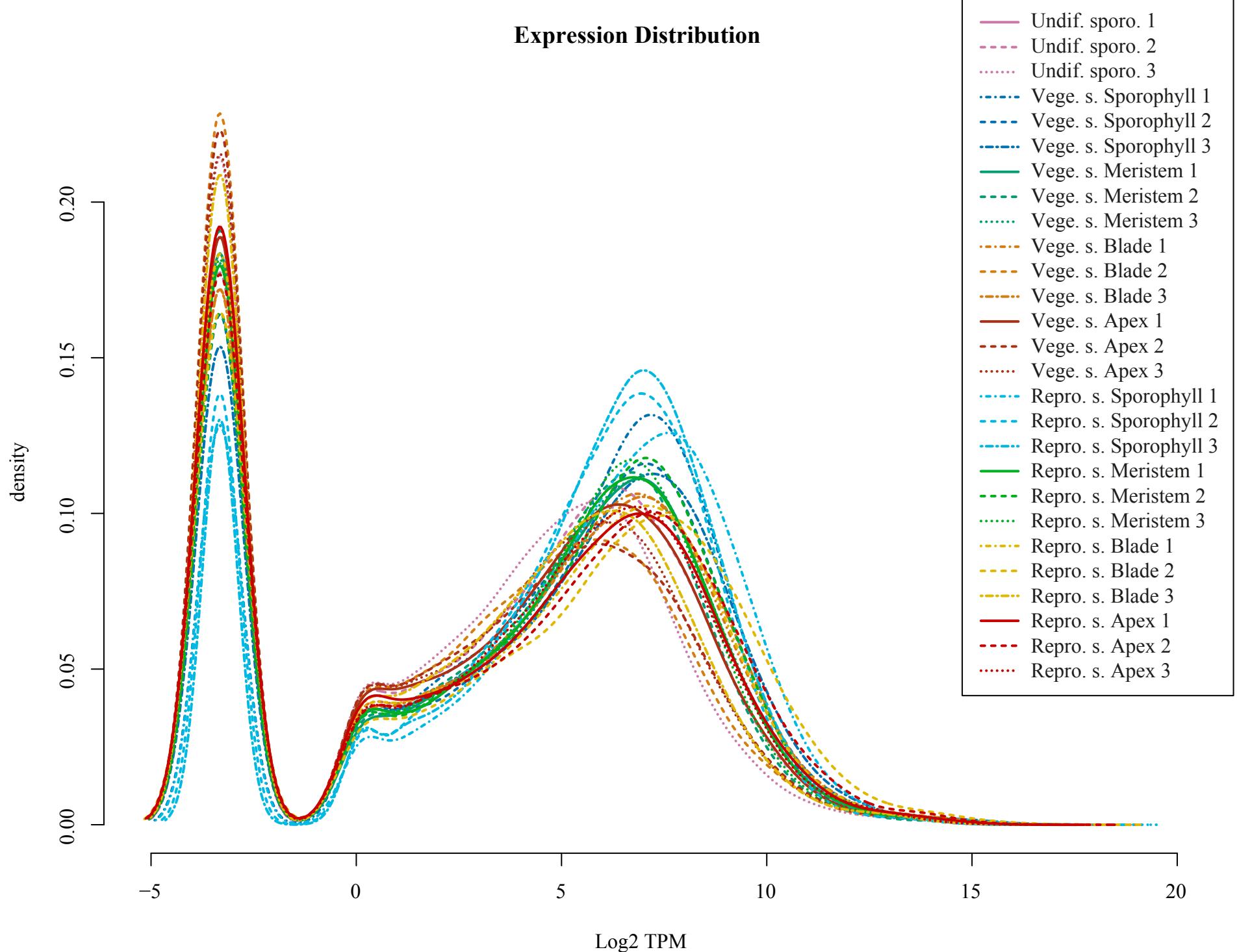


**Figure S9:** Voucher specimen of a vegetative sporophyte of *Undaria pinnatifida* sampled on February 26, 2017; representative of the vegetative sporophytes used for RNA sequencing. The sporophyll can be seen limited to a small portion of the stipe.

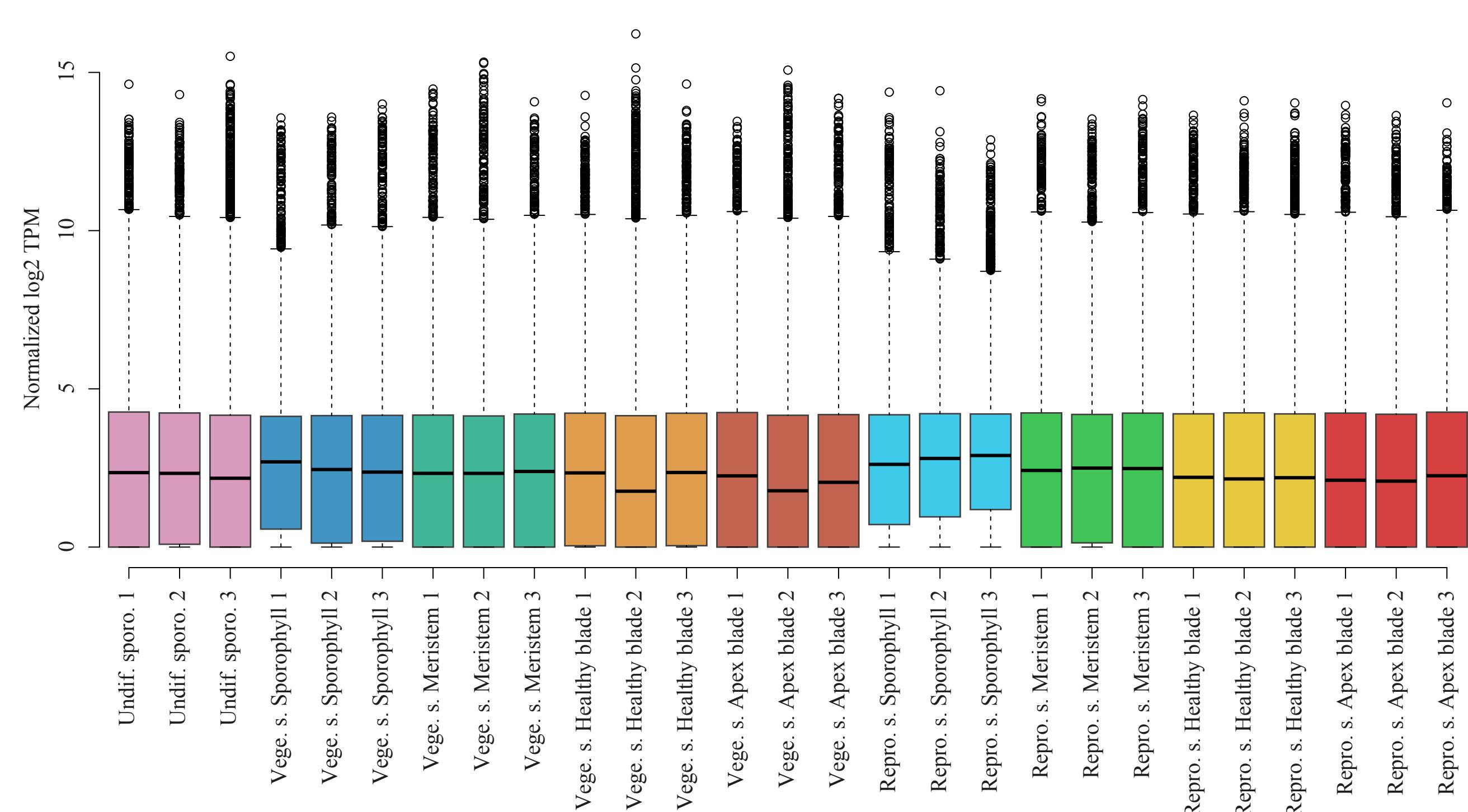


**Figure S10:** Detail of the sporophyll of a reproductive sporophyte of *Undaria pinnatifida* sampled on April 25, 2017; representative of the sporophylls used for RNA sequencing.

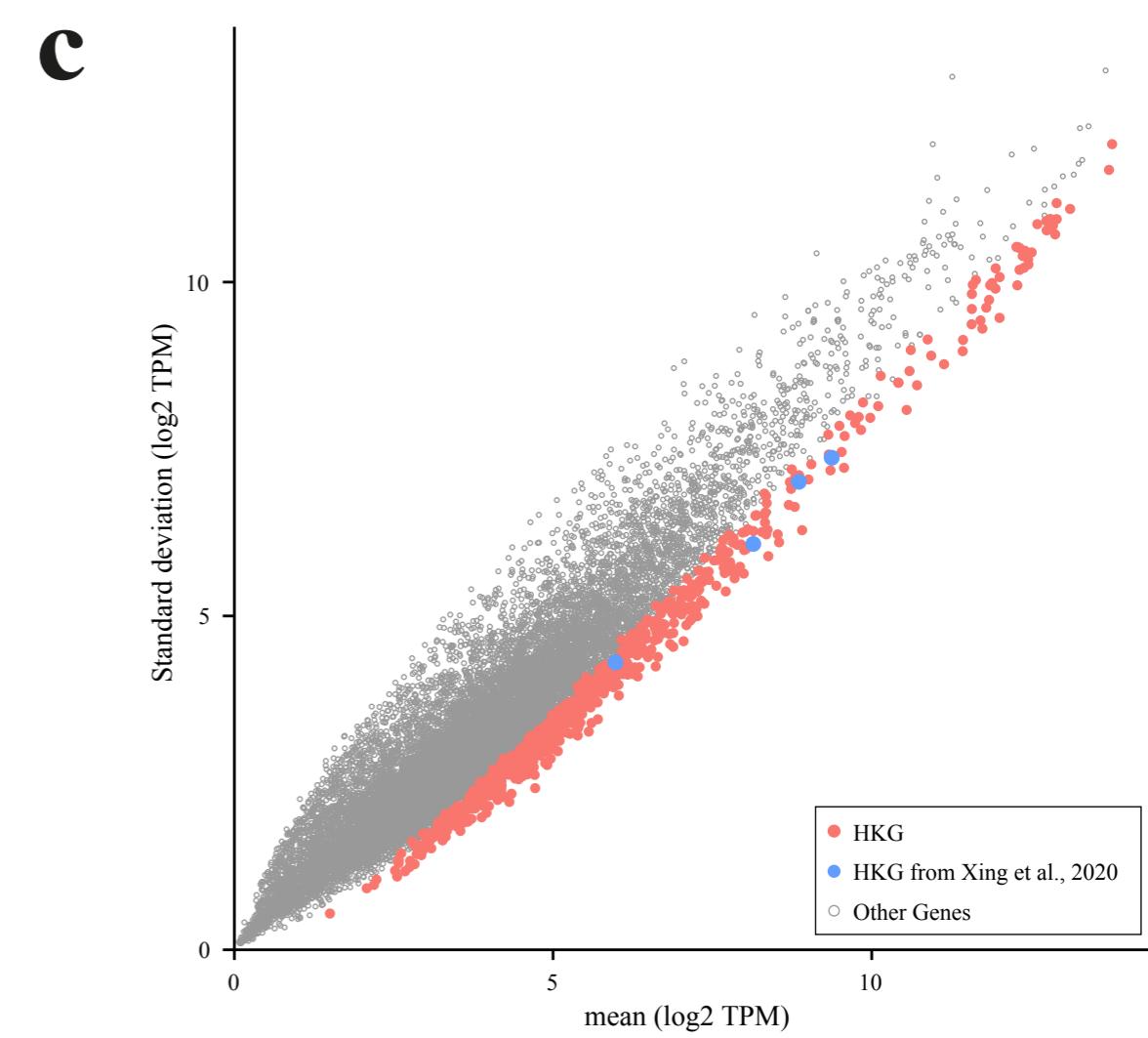
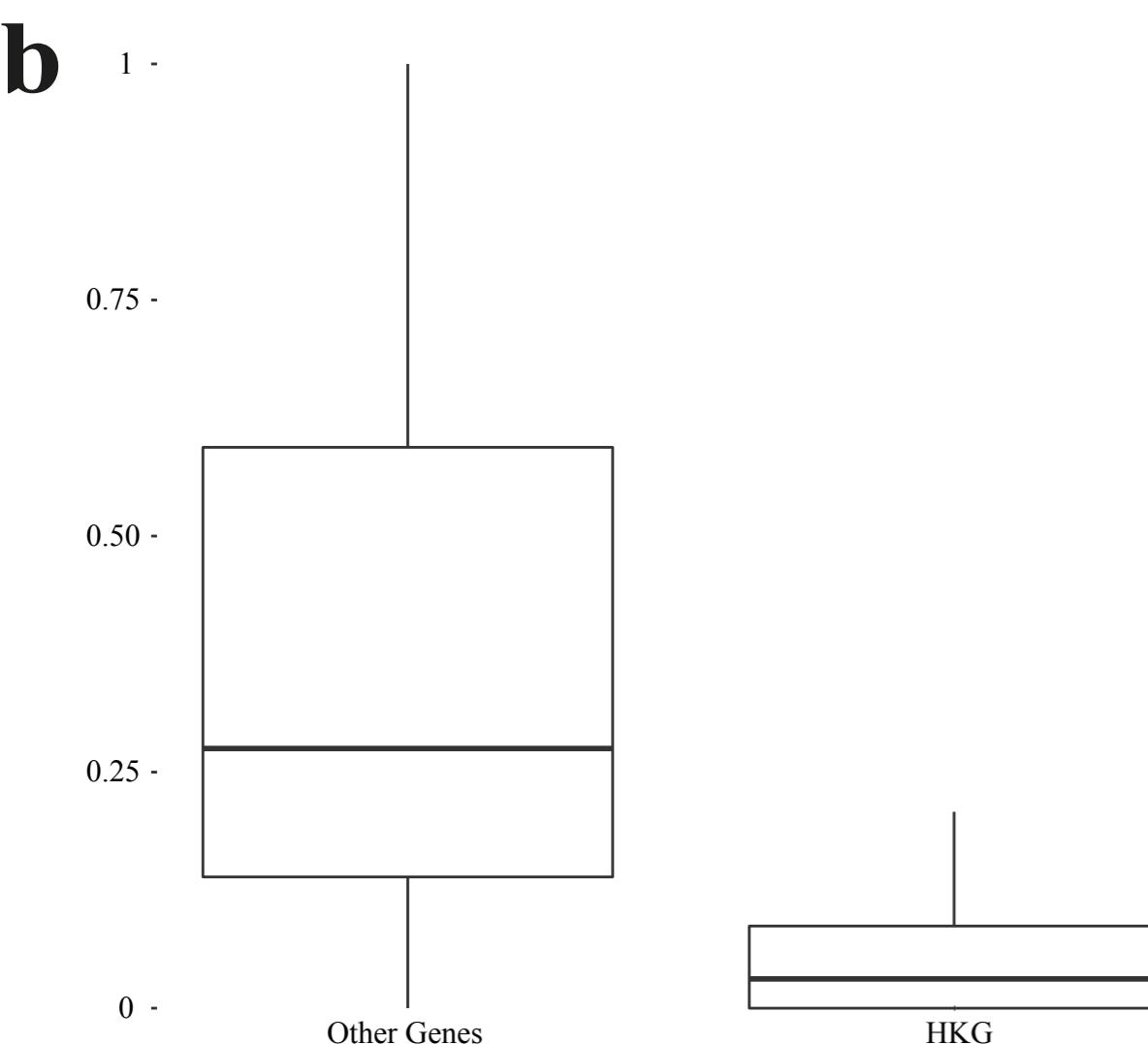
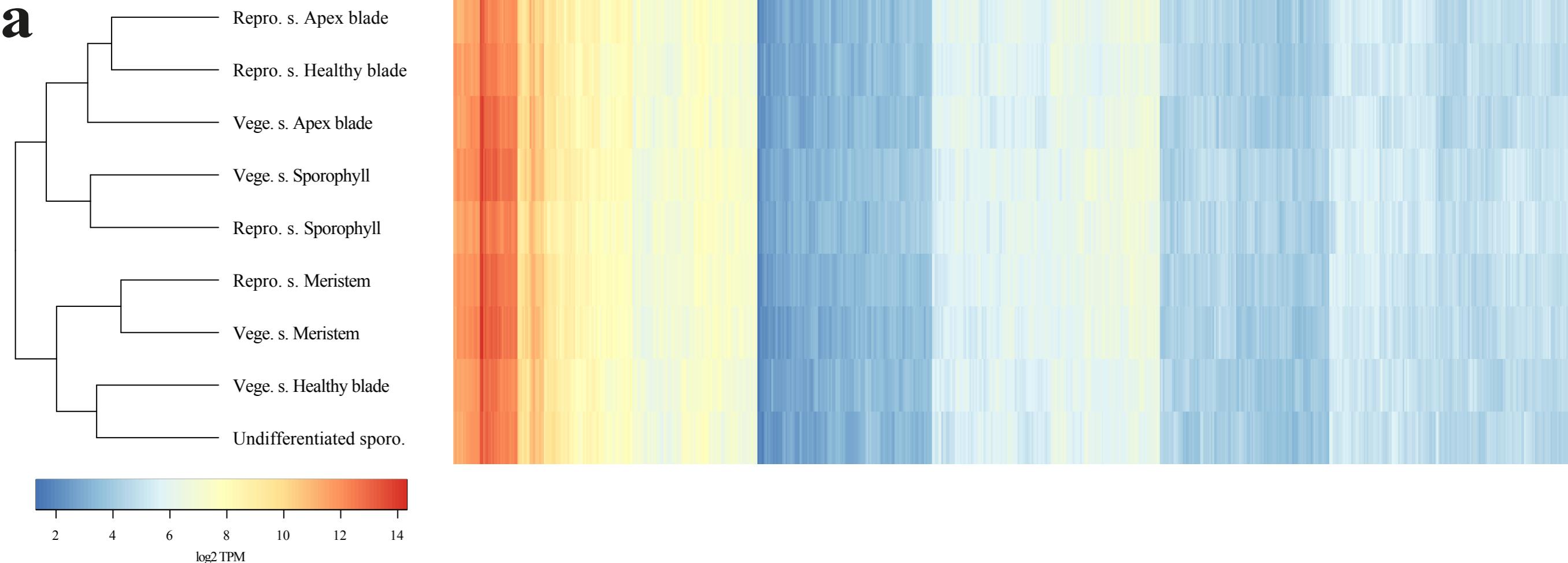
## Expression Distribution



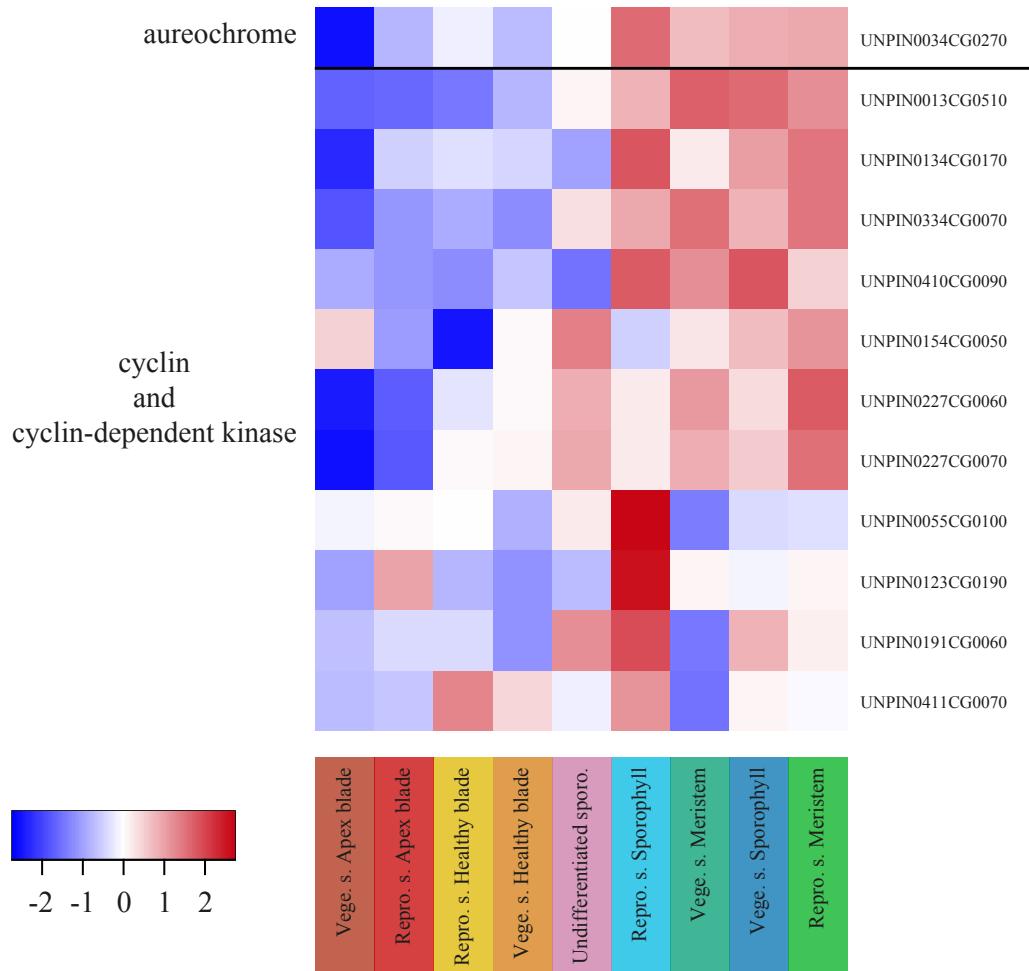
**Figure S1:** TPM density distributions of each of the 27 samples.



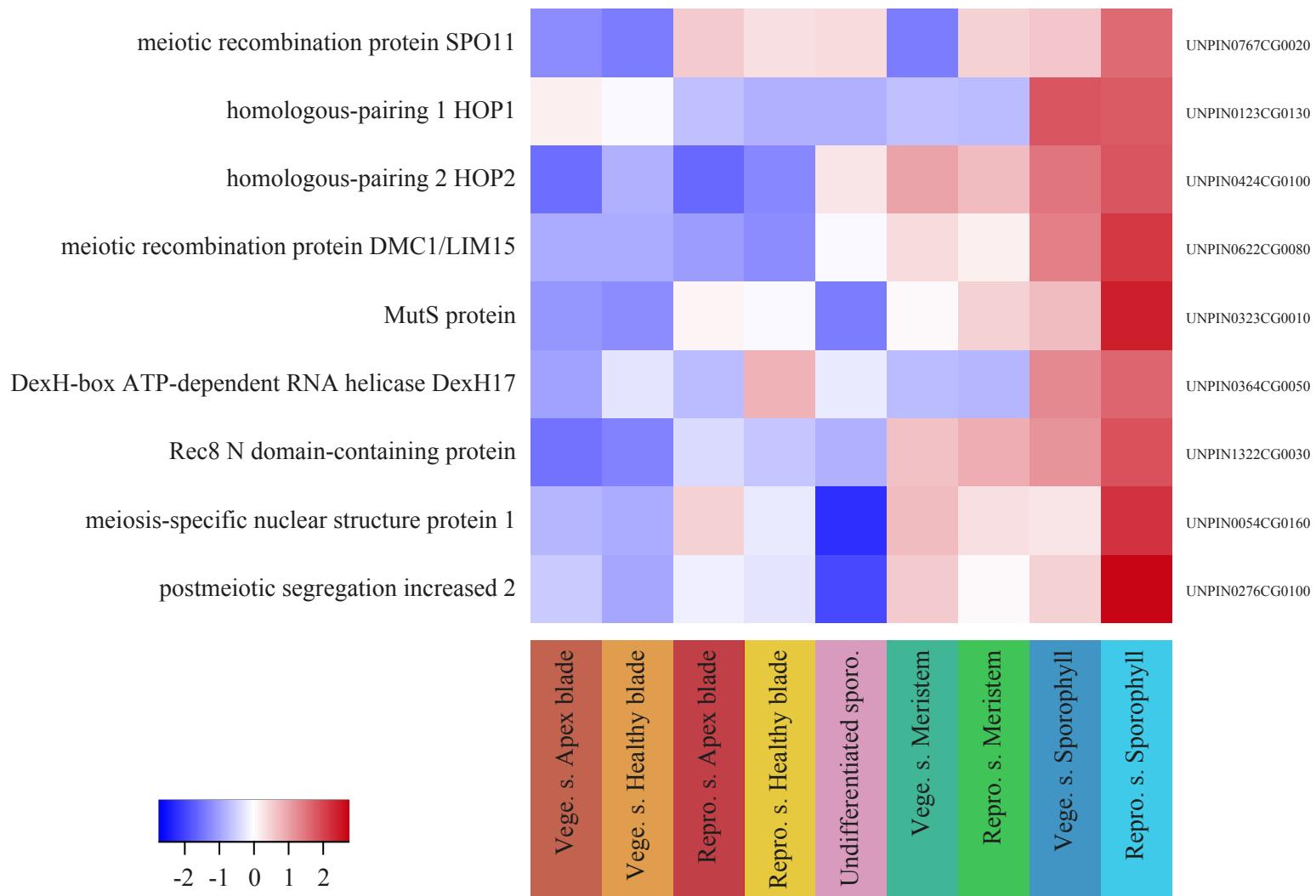
**Figure S2:** Boxplot of the normalized TPM counts of the 27 samples.



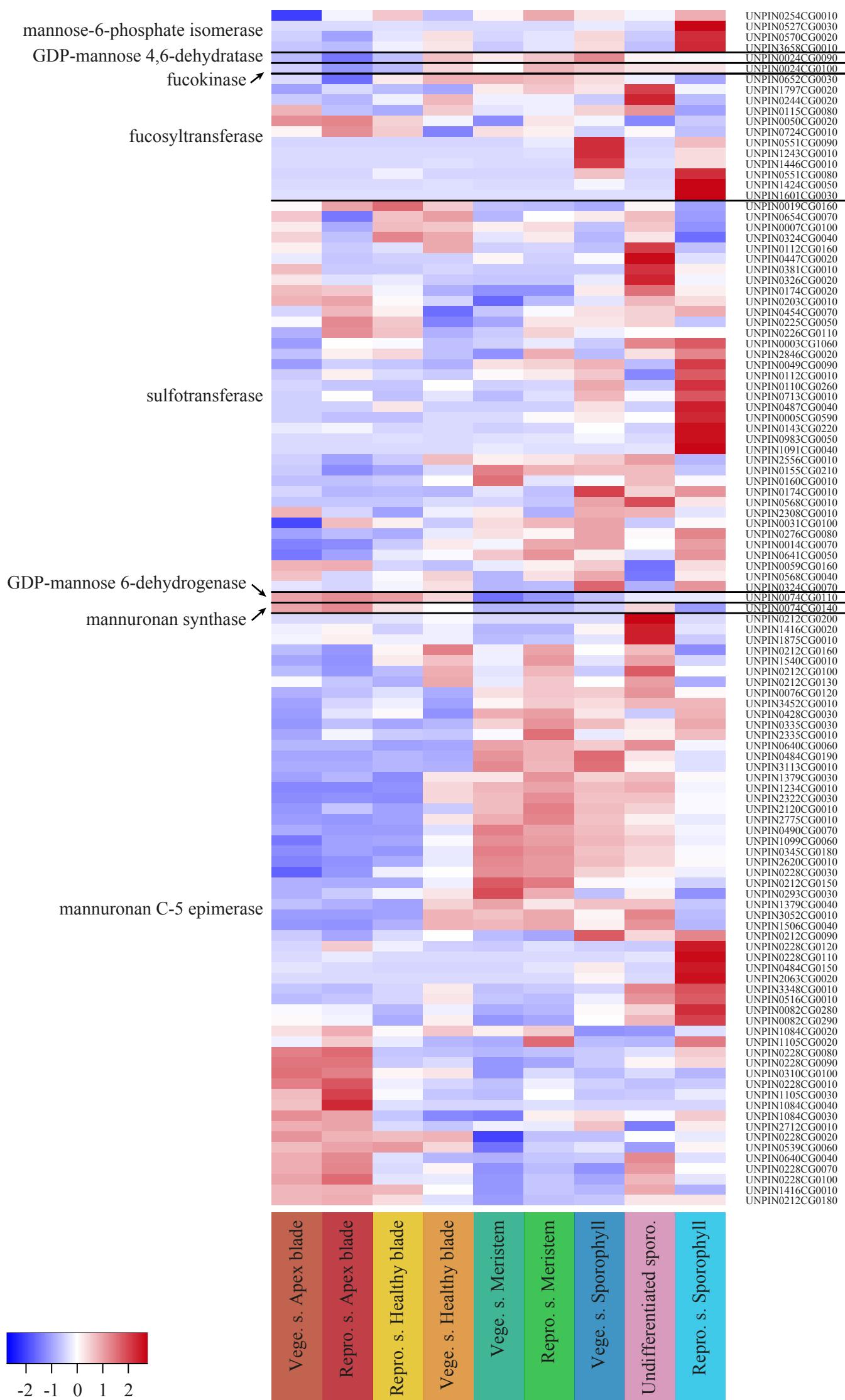
**Figure S3:** Expression and specificity of the Housekeeping Genes (HKGs). (a) Heatmap of the  $\log_2 \text{TPM}$  of the 718 genes identified as HKGs. (b) Boxplot of the  $Tau$  specificity index in the HKGs and all other genes. (c) Plot of the  $\log_2 \text{TPM}$  standard deviation as function of the mean  $\log_2 \text{TPM}$  of the gene expressed in *Undaria pinnatifida*.



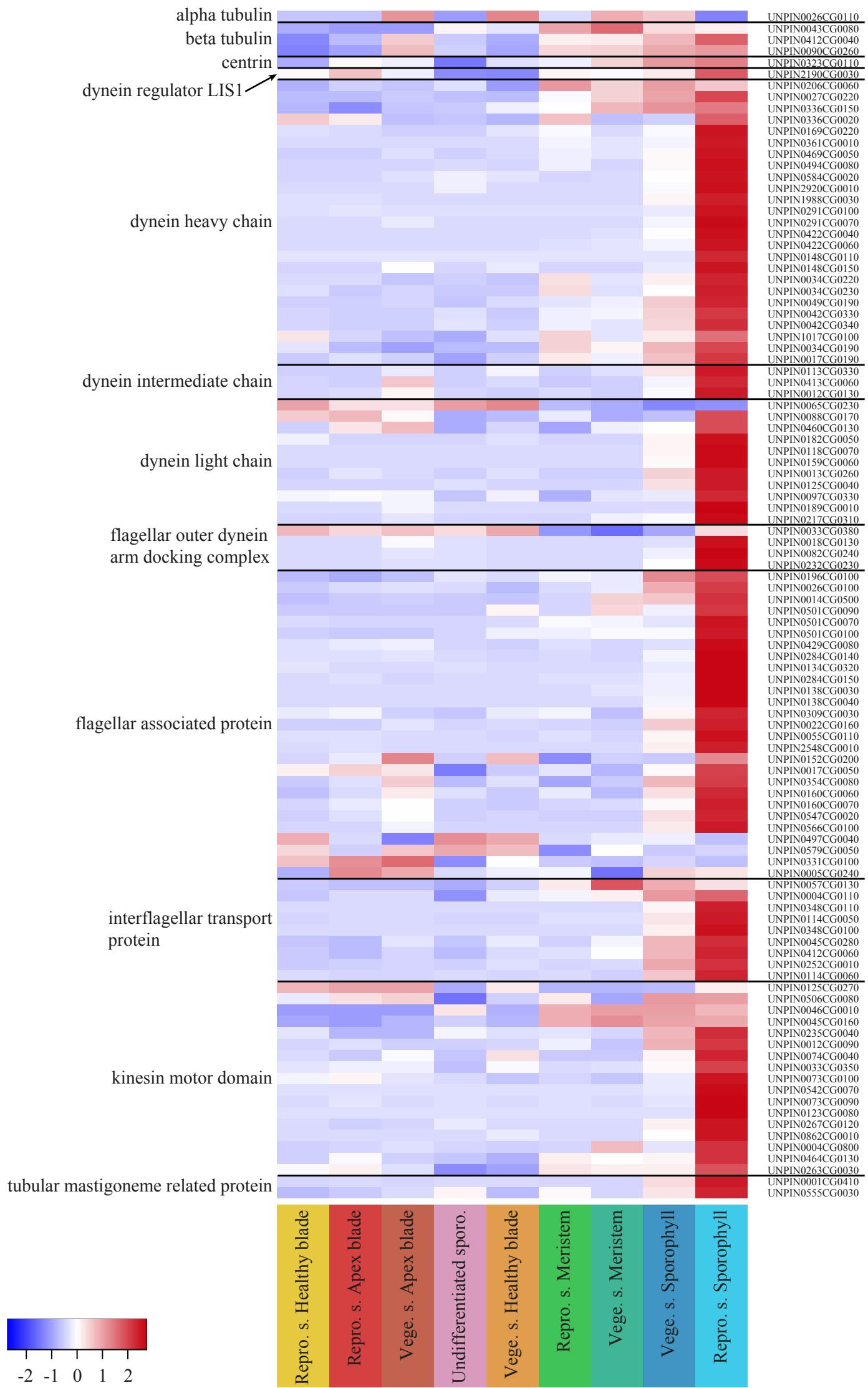
**Figure S4:** Heatmap of the differentially expressed genes of the cyclin and cyclin-dependent kinase gene families. Expression of genes is colored according to their level from high (red) to low (blue).



**Figure S5:** Heatmap of the differentially expressed genes related to meiosis. Expression of genes is colored according to their level from high (red) to low (blue).



**Figure S6:** Heatmap of the differentially expressed genes involved in the fucoidans biosynthesis in *Undaria pinnatifida*. Expression of genes is colored according to their level from high (red) to low (blue).



**Figure S7:** Heatmap of the differentially expressed genes related to flagellum. Expression of genes is colored according to their level from high (red) to low (blue).