

## ORIGINAL RESEARCH



# Insight into the nodal cells transcriptome of the streptophyte green alga *Chara braunii* S276

Daniel Heß<sup>1</sup> | Anja Holzhausen<sup>2</sup> | Wolfgang R. Hess<sup>1</sup>

<sup>1</sup>Genetics and Experimental Bioinformatics Group, Faculty of Biology, University of Freiburg, Freiburg, Germany

<sup>2</sup>Plant Cell Biology, Department of Biology, Philipps University Marburg, Marburg, Germany

## Correspondence

Wolfgang R. Hess, Genetics and Experimental Bioinformatics Group, Faculty of Biology, University of Freiburg, Schänzlestr. 1, 79104 Freiburg, Germany.  
 Email: [wolfgang.hess@biologie.uni-freiburg.de](mailto:wolfgang.hess@biologie.uni-freiburg.de)

## Funding information

Deutsche Forschungsgemeinschaft, Grant/Award Number: HE2544/18-1; Collaborative Research Centre 992 Medical Epigenetics, Grant/Award Number: SFB 992/1 2012; German Federal Ministry of Education and Research, Grant/Award Numbers: 031L0106 de.STAIR (de.NBI), 031L0101B/031L0101C de.NBI-epi, 031A538A/A538C RBC

Edited by S. de Vries

## Abstract

Charophyceae are the most complex streptophyte algae, possessing tissue-like structures, rhizoids and a cellulose-pectin-based cell wall akin to embryophytes. Together with the Zygnematophyceae and the Coleochaetophyceae, the Charophyceae form a grade in which the Zygnematophyceae share a last common ancestor with land plants. The availability of genomic data, its short life cycle, and the ease of non-sterile cultivation in the laboratory have made the species *Chara braunii* an emerging model system for streptophyte terrestrialization and early land plant evolution. In this study, tissue containing nodal cells was prepared under the stereomicroscope, and an RNA-seq dataset was generated and compared to transcriptome data from whole plantlets. In both samples, transcript coverage was high for genes encoding ribosomal proteins and a homolog of the putative PAX3- and PAX7-binding protein 1. Gene ontology was used to classify the putative functions of the differently expressed genes. In the nodal cell sample, main upregulated molecular functions were related to protein, nucleic acid, ATP- and DNA binding. Looking at specific genes, several signaling-related genes and genes encoding sugar-metabolizing enzymes were found to be expressed at a higher level in the nodal cell sample, while photosynthesis- and chloroplast-related genes were expressed at a comparatively lower level. We detected the transcription of 21 different genes encoding DUF4360-containing cysteine-rich proteins. The data contribute to the growing understanding of Charophyceae developmental biology by providing a first insight into the transcriptome composition of *Chara* nodal cells.

## 1 | INTRODUCTION

Originally considered as a model organism for plant electrophysiology (reviewed in Beilby, 2016) and studies of plant polarized growth and gravity sensing (Braun & Limbach, 2006), the Charophyceae alga *Chara braunii* Gmel. 1826 (Gmelin, 1826) has recently gained considerable interest as a model organism to study early land plant evolution. Charophyceae, along with the Zygnematophyceae and the

Coleochaetophyceae, form the ZCC grade, from which the common ancestor of land plants evolved (de Vries & Archibald, 2018; Hess et al., 2022; Wickett et al., 2014; Wodniok et al., 2011). While Zygnematophyceae are the sister group to, and share a last common ancestor with, land plants (Cheng et al., 2019; Ruhfel et al., 2014; Wickett et al., 2014), the earlier branching Charophyceae form the only class that includes streptophyte algae with tissue-like structures and functional rhizoids (Bonnot et al., 2019), traits that were probably already

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 The Authors. *Physiologia Plantarum* published by John Wiley & Sons Ltd on behalf of Scandinavian Plant Physiology Society.

possessed by the common ancestor of land plants, at least in a rudimentary fashion (Fürst-Jansen et al., 2020). Moreover, Charophyceae of the ZCC grade possess a huge cellulose-pectin-based cell wall with many polymers similar to land plants as well as plasmodesmata (Mikkelsen et al., 2014; Sørensen et al., 2011; Umen, 2014), which are features relevant for the successful colonization of the terrestrial habitat. Charophyceae have indeed been considered as the sister lineage to land plants early on (Pringsheim, 1863).

The model organism *Chara braunii* produces both a high number of oospores within a few months and has a short annual life cycle compared to other Charophyceae species. Reproduction can occur generatively by means of oospores, but can also proceed vegetatively by fragmentation of thalli (Casanova, 2015; Holzhausen et al., 2022). The sequenced draft genome (Nishiyama et al., 2018) revealed evolutionarily significant, plant-like features as well as a striking secondary complexity. Furthermore, standardized protocols for cultivation and generative germination have been established, enabling a more detailed insight through transcriptomic studies (Holzhausen et al., 2022).

Roughly 500 million years ago, an algal concestor of Zygnemato-phyceae and embryophytes colonized land and gave rise to land plants (Martin & Allen, 2018). Generally, a comprehensive understanding of early molecular innovations towards the dynamic terrestrial environment is still limited (Zhang et al., 2022). The extant Charales, such as *Chara braunii*, bear striking similarities to the morphology of land plants, yet also display important differences in cell biology (reviewed in Domozych & Bagdan, 2022). Therefore, analysis of the *Chara braunii* developmental program can shed light on the evolutionary adaptations associated with these morphological features.

The monoecious macroalgae grows from a terminal apical cell and consists of a central stem with branchlets radiating from axial nodes at regular intervals. Nodal and elongated, multinucleate internodal cells are derived from the apical cell in alternating fashion, separating each whorl by an internodal cell. The nodes contain at their center a pair of small cells, called central cells, and an additional 6 to 20 cells surrounding this pair. These nodal cells derive from a nodal initial cell (Kuczewski, 1906; Schubert et al., 2016) and can develop into lateral branches (Nishiyama et al., 2018), suggesting a stem cell-like character for the initial cells. Along its branchlet nodes, the resulting *Chara braunii* bears both male (antheridia) and female (oogonia) gametangia (Moody, 2020; Nishiyama et al., 2018).

Characeae present certain evolutionary novelties of streptophyte algae, leading to three-dimensional growth: apical cell growth, tip growth and division plane rotation (involving the phragmoplast and cell plates). The nodal cells, found at morphologically key positions between the long internodal cells and branchlets along central cells, can still undergo cell divisions. In contrast to internodal cells, this enables them to generate distinct nodal discs and the founder cells for branches (reviewed by Buschmann, 2020). Compared to oogonia and antheridia, this de novo formation of apical cells also enables asexual propagation (Nishiyama et al., 2018).

Due to their prevailing ability to divide and differentiate, we hypothesized that the transcriptional profiles of nodal cells might

differ from vegetative cells in *Chara braunii* and, therefore, might provide insight into cell-type specific gene expression.

Here, we performed a transcriptome analysis of a *Chara braunii* sample enriched for nodal cells compared to total plantlet material, providing the first insight into the gene expression in these cells. Whole transcriptome sequencing of *Chara braunii* central and nodal cells revealed the relative enrichment of several signaling-related mRNAs and of mRNAs encoding sugar-metabolizing enzymes, while photosynthesis- and chloroplast-related genes were expressed at a comparatively lower level.

## 2 | MATERIALS AND METHODS

### 2.1 | Culture of *Chara braunii*, growth conditions and dissection of central and nodal cells

Cultures of the non-axenic freshwater strain *Chara braunii* S276 (KU-MACC) were vegetatively propagated. *Chara braunii* S276 was originally isolated from Lake Kasumigaura (Ibaraki, Japan), then maintained at Kobe University (Kawai et al., 2022; Sato et al., 2014) and consecutively propagated at the Universities of Marburg and Freiburg. Plantlets were grown using the protocol of Holzhausen et al. (2022) in double autoclaved culture vessels containing sieved compost (Gardol Pure Nature, Bauhaus), lime (Gardol Garten- & Rasenkalk, Bauhaus), quartz sand (0.4–0.8 mm in diameter, Carl Roth GmbH) and distilled water, sealed with surgical tape (Micropore, 3 M). Algae were kept under long-day conditions at 22°C in a 16 h light: 8 h dark cycle with white light lateral illumination (30–70  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ , Lumilux L36W/840, OSRAM).

Algae were cut along their central thallus below and above each node, removing internodal and lateral cells (including branchlets), at the University of Marburg using razor blades and a stereomicroscope (Leica DM6000 CS). Harvested nodes from multiple algae were pooled, collected in 1.5 ml Eppendorf tubes, frozen in liquid nitrogen and stored at  $-80^{\circ}\text{C}$  until further analysis.

### 2.2 | Preparation of total RNA and northern hybridizations

Total RNA was extracted from complete thalli utilizing a modified acid guanidinium thiocyanate-phenol-chloroform protocol (Chomczynski & Sacchi, 1987, 2006) using PGTX (Pinto et al., 2009), but omitting Triton X-100.

At least 100 mg FW total *Chara braunii* samples were ground into a homogenate using mortar and pestle under liquid nitrogen. Samples were suspended in 1 ml Z6 buffer (8 M guanidiniumhydrochloride, 20 mM MES, 20 mM EDTA, 50 mM 2-mercaptoethanol, pH 7), mortar rinsed with 500  $\mu\text{l}$  Z6 buffer, and transferred to a screw-cap tube. 3 ml PGTX was added and samples were incubated for 30 min at room temperature while vortexing every 5 min. Two chloroform extractions were performed consecutively: each adding 3 ml

chloroform/isoamylalcohol (24:1), incubating samples for 10 min at room temperature under occasional vortexing, centrifugation for 3 min (3.273 g, at room temperature) and transferring the aqueous layer to a fresh tube. After chloroform extractions, RNA was precipitated using one volume 2-propanol overnight at  $-20^{\circ}\text{C}$ . RNA was collected by centrifugation for 30 min (13,000 g,  $4^{\circ}\text{C}$ ), and washed two times using 70% ethanol and air-dried for 10 min. RNA was resuspended in 100  $\mu\text{l}$  RNase-free  $\text{H}_2\text{O}$  and stored at  $-80^{\circ}\text{C}$  until further analysis.

Nodal RNA was extracted similarly. Pooled nodal cells were suspended in 250  $\mu\text{l}$  Z6 buffer in ice-cooled 2 ml reaction tubes. Cell disruption was performed using a mixer mill (MM 400, Retsch; 5 cycles of 90 s with varying frequencies: 30 s 8/s, 30 s 16/s, 30 s 25/s, and 2 min cooled on ice) using one glass bead (2.85–3.45 mm diameter, Carl Roth GmbH). The supernatant was transferred into a new reaction tube, together with 125  $\mu\text{l}$  of Z6 buffer used for rinsing the glass beads. 750  $\mu\text{l}$  PGTX was added and samples incubated for 30 min at room temperature while vortexing every 5 min. Addition of 750  $\mu\text{l}$  chloroform/isoamylalcohol (24:1) was followed by incubation for 10 min at room temperature under occasional vortexing. Afterwards, sample tubes were centrifuged for 3 min (room temperature, 3273 g). Following another chloroform extraction, the upper aqueous phase was precipitated with one volume 2-propanol overnight at  $-20^{\circ}\text{C}$ . RNA was collected by centrifugation for 30 min (13,000 g,  $4^{\circ}\text{C}$ ) and air-dried for 10 min. RNA was resuspended in 20  $\mu\text{l}$  RNase-free  $\text{H}_2\text{O}$  and stored at  $-80^{\circ}\text{C}$  until further analysis.

RNA concentration and purity were measured using a NanoDrop ND-1000 spectrophotometer according to manufacturer instructions (PEQLAB Biotechnologie GmbH). Residual DNA was removed using the Turbo DNase-free™ Kit (Thermo Fisher Scientific) following manufacturer instructions. RNA was recovered using RNA Clean & Concentrator kits (Zymo Research). RNA integrity was controlled on a 5200 Fragment Analyzer System, according to manufacturer instructions (Agilent).

Eight  $\mu\text{g}$  total RNA was separated by 1.4% denaturing agarose gel electrophoresis and transferred by capillary transfer on Hybond N+ nylon membranes (Cytiva Europe GmbH) overnight. Northern hybridization was performed with radioactively labeled probe for chlorophyll

*a/b* binding protein (*cab*) mRNA (primers for template; see Table 1 for sequences), generated using [ $\alpha$ - $^{32}\text{P}$ ]-UTP and the Maxiscript T7 in vitro transcription kit (Thermo Fisher Scientific). Blotted RNA was crosslinked to the membrane via 240 mJ using a UV-Stratalinker (Stratagene). Hybridizations were performed in Northern buffer (50% deionized formamide, 7% SDS, 250 mM NaCl and 120 mM  $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$  pH 7.2) overnight at  $62^{\circ}\text{C}$ . The membranes were washed at  $57^{\circ}\text{C}$  in buffer 1 ( $2\times$  SSC (3 M NaCl, 0.3 M sodium citrate, pH 7.0), 1% SDS), buffer 2 ( $1\times$  SSC, 0.5% SDS) and buffer 3 ( $0.1\times$  SSC, 0.1% SDS) for 10, 5, and 2 min. Hybridization signals were detected via Typhoon FLA 9500 imaging system (GE Healthcare) using phosphorimaging and quantified using Quantity One software (Bio-Rad Laboratories Inc.).

Templates were amplified from *Chara* genomic DNA using OneTaq Quick-Load  $2\times$  Master Mix with Standard Buffer. To facilitate the use of PCR products as templates for in vitro transcription, the T7 RNA polymerase promoter (TAATACGACTCACTATAGGG) was included in the 5' sections of the corresponding reverse primers (Table 1).

## 2.3 | Preparation of total DNA from *Chara braunii*

Genomic DNA was extracted from at least 100 mg FW total *Chara braunii* samples ground into a homogenate using mortar and pestle under liquid nitrogen. Samples were suspended in 1 ml SET buffer (1 mM EDTA, 25% sucrose, 50 mM Tris) and lysed overnight at  $50^{\circ}\text{C}$  in a screw-cap tube after adding 1/4 volume 0.5 M EDTA, 1/10 volume 20% SDS and 100  $\mu\text{g}/\text{ml}$  proteinase K (Scholz et al., 2019).

After the addition of 2 volumes of ROTI Aqua-Phenol (DNA; Carl Roth GmbH) and 2 volumes of chloroform/isoamylalcohol (24:1), samples were incubated for 5 min at room temperature while vortexing intermediately. Samples were centrifuged for 5 min (3273 g, room temperature), the aqueous layer was transferred to a fresh tube and the chloroform extraction step was repeated. Afterwards, DNA was precipitated using one volume 2-propanol overnight at  $-20^{\circ}\text{C}$ , collected by centrifugation for 30 min (13,000 g,  $4^{\circ}\text{C}$ ), and washed two times using 70% ethanol and air dried for 10 min. DNA was resuspended in 100  $\mu\text{l}$  RNase-free  $\text{H}_2\text{O}$  and stored at  $-20^{\circ}\text{C}$  until further analysis.

**TABLE 1** Oligonucleotides used in this study

ID	Sequence	Description
CBr-CABconsensus-F	<b>TAATACGACTCACTATAGGG</b> TAGGCCAGGCGTTGTT	Northern Blot
T7-CBr-CABconsensus-R	GACCGCCCAAGTACC	Northern Blot
TruSeq_Sense_nodal	AATGATACGGCGACCACCGAGATCTACAC-CAAGTGGG- ACACTCTTTCCCTACACGACGCTCTTCCGATCT	TruSeq_Sense_primer i5 Barcode
TruSeq_Sense_control	AATGATACGGCGACCACCGAGATCTACAC-AGATTGG- ACACTCTTTCCCTACACGACGCTCTTCCGATCT	TruSeq_Sense_primer i5 Barcode
TruSeq_Antisense_nodal	CAAGCAGAAGACGGCATACGAGAT-TCCGCGAA- GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	TruSeq_Antisense_primer i7 Index
TruSeq_Antisense_control	CAAGCAGAAGACGGCATACGAGAT-TCCGCGAA- GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	TruSeq_Antisense_primer i7 Index

Note: An added T7 promoter sequence is highlighted in boldface letters.

## 2.4 | cDNA sequencing

cDNA libraries were constructed and sequenced as a service provided by Vertis Biotechnologie AG (Germany). Poly(A)<sup>+</sup> RNAs were fragmented using ultrasound (one pulse of 30 s at 4°C), after which an oligonucleotide adapter was ligated to the 3' termini. First-strand cDNA synthesis was performed using M-MLV reverse transcriptase and the 3' adapter as primer. First-strand cDNA was purified and the 5' Illumina TruSeq sequencing adapter ligated to the 3' end of the antisense cDNA. The resulting cDNA was PCR-amplified using a high fidelity DNA polymerase. The cDNA was purified using an Agencourt AMPure XP kit (Beckman Coulter Genomics).

For Illumina NextSeq sequencing, samples were pooled in approximately equimolar amounts, purified using the Agencourt AMPure XP kit (Beckman Coulter Genomics) and analyzed by capillary electrophoresis. Primers used for PCR amplification were designed for TruSeq sequencing according to manufacturer instructions. The cDNA pool was single-read sequenced on an Illumina NextSeq 500 system using 75 bp read length. Single replicates were sequenced for the pooled nodal cells and the control cell samples, respectively.

## 2.5 | Bioinformatic analyses

Data analysis was performed on the European instance galaxy server (The Galaxy Community, 2022) following the available guidelines for reference-based RNA-seq analysis (Batut et al., 2018, 2022). The quality of raw reads was assessed using FastQC (Andrews, 2010) and MultiQC (Ewels et al., 2016). Trimming of adapter and barcode contamination was performed using cutadapt (Martin, 2011). Mapping and counting of reads against the published *Chara* genome (Nishiyama et al., 2018) was performed via RNA STAR (Dobin et al., 2013). Differential gene expression analysis was performed via log<sub>2</sub>FC calculation following the given guidelines. Differentially expressed genes were determined with a cutoff as log<sub>2</sub>FC ≥ |1|. Gene ontology (GO) analysis of RNA-seq data was performed using goseq (Galaxy version 1.44.0; Young et al., 2010).

## 2.6 | Phylogenetic analysis

A set of amino acid sequences (Table S1) was selected based on position within the green lineage and sequence similarity to *Chara braunii* putative amylase homologous genes of interest, g31563 and g41182, and putative MurE-like PAP11, g32224. The sequences were aligned with M-coffee (Di Tommaso et al., 2011; Notredame et al., 2000) and further analyzed using the BEAST 2 software package version 2.7.3 (Bouckaert et al., 2019). Calculations were performed using standard BEAUTi settings utilizing the tree prior yule model, substitution model Blosum62 and MCMC chain length of 1e6, with logged parameters every 1e4 steps. Tracer v1.7.2 software (Rambaut et al., 2018) was used to validate the effective sample sizes (ESS > 200). TreeAnnotator was used to build maximum clade credibility trees using burnIn of

50% and a posterior probability limit of 0.5 for median node heights. FigTree v1.4.4 software (Rambaut, 2018) was used to visualize the generated dendrograms.

## 3 | RESULTS

### 3.1 | The most abundantly transcribed genes in *Chara* nodal cells

The nodal cells of *Chara braunii* are embedded in the zone between two adjacent large internodal cells from where the branchlets are emanating (Figure 1). By cutting the algae below and above each node, only the nodal cells from the central thallus were collected. Separated nodes were kept in liquid nitrogen until RNA extraction. RNA was extracted from 28 pooled nodes, yielding 0.68 µg of total RNA after DNase treatment. In parallel, RNA was extracted from fresh total algal material. We decided to use complete above-ground thalli, including nodal cells, as control. The inclusion of nodal cells in the control allowed, after cDNA sequencing, the direct calculation of log<sub>2</sub> fold change (log<sub>2</sub>FC) factors for the differentially expressed genes. Accordingly, positive log<sub>2</sub>FC factors indicated mRNAs enriched in nodal cells. Judged by the visible sharp rRNA bands and the test hybridization for *cab* mRNA, the prepared RNA samples were of high quality (Figure S1). After polyA-mRNA enrichment, cDNA libraries were generated for the two different *Chara braunii* RNA sample types and sequenced, yielding a total of 54,042,636 raw single read counts for enriched nodal cells and 53,423,073 for total cells. From these, 10,991,291 (20.6%) reads for total and 13,763,747 (25.5%) for the enriched sample remained unmapped. After trimming and mapping steps, 4,994,360 nodal (9.4%) and 2,628,740 total (4.9%) reads were uniquely mapped to the *Chara* genome (Figure 2). Mapped reads matched 11,319 putative genes for the nodal sample and 15,627 genes for total sample. The 15 most highly expressed genes for each tissue sample are given in Tables 2 and 3, while the full list of genes is given in Table S2.

In the control sample, the 15 most expressed genes were mostly associated with photosynthesis and protein synthesis, such as chlorophyll *a-b* binding proteins (5/15 instances), the photosystem (PS) II 5 kDa protein PsbT, or the ribosomal proteins Rpl6-like, Rpl35a-3-like, Rpl10, Rps19-3, and Rps8 (Table 2). Moreover, mRNA for g46838, a gene encoding a GC-rich sequence DNA-binding factor-like protein with similarity to the PAX3- and PAX7-binding proteins, was found enriched in both samples (Tables 2 and 3). Last, three of these highly expressed mRNAs encode unknown proteins.

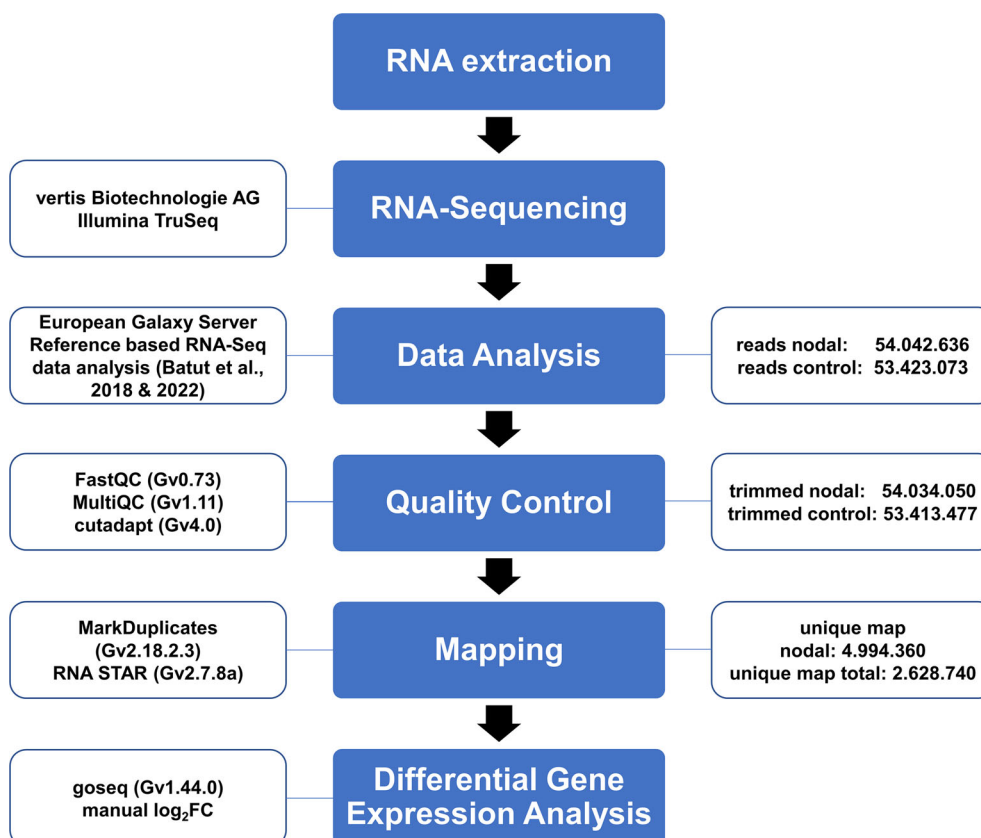
In the sample enriched for nodal cells, the 15 most strongly expressed genes differed from the control by lacking all genes related to photosynthesis. However, even 9/15 mRNAs were encoding ribosomal proteins, indicating that the drive for protein synthesis was high in this tissue (Table 3). g26356, encoding a thioredoxin-encoding mRNA, and g46838, a protein with similarity to the PAX3- and PAX7-binding proteins, were also found enriched. Another four enriched mRNAs encode unknown proteins.



**FIGURE 1** Specimen of *Chara braunii*. Detailed view of *Chara braunii* S276. Shown is the area with the interior node cells confined by two internodal cells. Nodal cells pinched off peripheric cells from which branchlets developed with a structure analogous to the main stem consisting of branchlet nodal and internodal cells. Branchlet nodal cells are the origin for the development of antheridia (shown) and the typical charophytic oogonia after proliferation.



**FIGURE 2** RNA-seq data analysis workflow. Overview of steps performed for RNA-seq analysis of *Chara braunii* transcriptome. Blue boxes (middle) illustrate steps within the workflow from RNA extraction to gene expression analysis, transparent boxes (left) illustrate the methods and bioinformatic tools used, transparent boxes (right) illustrate the number of reads at various points of the bioinformatic analysis of transcriptome data. Data analysis was performed on the European galaxy server (The Galaxy Community, 2022) based on the published guidelines for reference-based RNA-seq data analysis (Batut et al., 2018, 2022).



When looking at the full list of genes expressed, 117 different mRNAs for proteins with domains of unknown function (DUF) were identified in both samples. Among them, we found 21 instances of

genes encoding DUF4360-containing cysteine-rich proteins. While DUF4360-containing proteins exist in bacteria and eukaryotes (Paysan-Lafosse et al., 2023), it is striking that these *Chara braunii*

Gene ID	Description	Control counts	Nodal counts
g49772	n/a	820,153	1875
g50349	Chlorophyll a-b binding protein	76,820	7262
g49773	n/a	74,655	178
g48962	n/a	56,585	128
g20157	Chlorophyll a-b binding protein 151, chloroplast	53,471	1408
g18875	PSII 5 kDa PsbT protein, chloroplast	50,667	5196
g20098	60S ribosomal L6-like protein	39,190	38,394
g20154	Major chlorophyll binding protein	37,336	2194
g20151	Chlorophyll a-b binding protein 151, chloroplast	36,381	740
g50044	60S ribosomal protein L35a-3-like	33,211	20,144
g29937	60S ribosomal protein L10	29,890	21,826
g21828	40S ribosomal protein S19-3	29,726	19,115
g32357	40S ribosomal protein S8	29,627	18,184
g19118	Chlorophyll a-b binding protein 151, chloroplast	26,204	1230
g46838	PAX3- and PAX7-binding protein 1	25,206	30,172

Note: The 15 genes for which the highest read counts in the control sample were detected in descending order. The gene ID in the first column is followed by the annotation and the normalized read counts in the control and in the nodal samples.

**TABLE 2** Most abundant mRNAs in control.

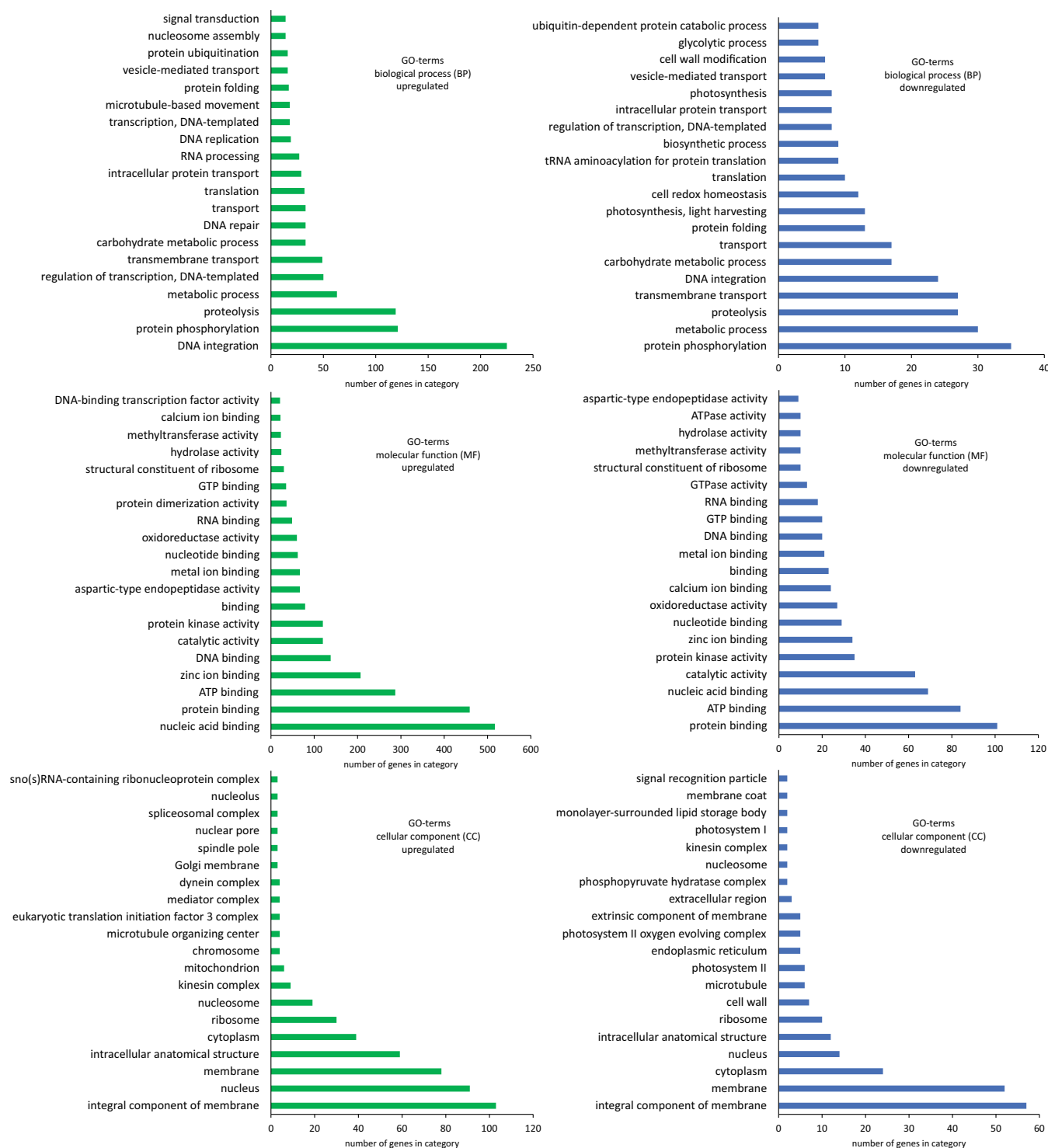
Gene ID	Description	Nodal counts	Control counts
g20098	60S ribosomal L6-like protein	38,394	39,190
g46838	PAX3- and PAX7-binding protein 1	30,172	25,206
g29937	60S ribosomal protein L10	21,826	29,890
g10845	n/a	20,269	19,117
g50044	60S ribosomal protein L35a-3-like	20,144	33,211
g21828	40S ribosomal protein S19-3	19,115	29,726
g17012	ribosomal protein L3, conserved site	18,638	20,362
g32357	40S ribosomal protein S8	18,184	29,627
g41408	60S ribosomal protein L5	13,381	23,292
g26356	thioredoxin H-type	13,127	6048
g12280	n/a	13,114	9764
g38728	n/a	13,063	22,161
g23412	60S ribosomal protein L17-2-like	12,797	25,171
g8281	40S ribosomal protein S21-2	11,504	10,917
g36718	n/a	10,474	5103

Note: The 15 genes for which the highest read counts in the nodal sample were detected, in descending order. The gene ID in the first column is followed by the annotation and the normalized read counts in the nodal and in the control samples.

**TABLE 3** Most abundant mRNAs in central and nodal cells.

DUF4360-containing proteins are more similar to homologs from fungi (mainly Ascomycota) than to any other. Although potential instances of horizontal gene transfer events between algae and fungi or soil bacteria were reported (Donoghue & Paps, 2020), we cannot unambiguously exclude the possibility of fungal contamination given the non-axenic nature of the *Chara* cultures. However, the genes encoding DUF4360-containing proteins are part of the *Chara braunii*

genome annotation. Therefore, if these reads resulted from a contaminating fungus, the same contamination had to be present already in the cultures that were taken for the initial genome analysis. Most of the mRNAs for DUF4360-containing proteins were more abundant in the *Chara* total control sample (e.g., genes g9217, g68682, g2717, g2719, g58375, g11068, g36724, g68700, g9208, g68689, g11072, and g76151; for details see Table S1).



**FIGURE 3** GO term analysis of transcripts up- or downregulated in *Chara* nodal cells. Chart representations of gene ontology terms enriched among genes up- or downregulated in nodal cells compared to total *Chara braunii* within the  $\log_2FC \geq |1|$  data set. The 20 most abundant GO terms for biological processes (BP), cellular components (CC) and molecular function (MF) are shown, separated by up- (green bars) and down-regulated genes (blue bars). The complete list is provided in Table S3.

### 3.2 | Functional assignments to differentially expressed genes

To obtain an overview of genes differentially expressed between the two samples, we set cutoffs of  $\log_2FC \geq |1|$  and reads per kb (RPK)

$\geq 10$ , leaving a set of 3145 differentially regulated genes (2762 up, 383 down in the sample enriched for nodal cells compared to control). The full list of genes can be found in Table S2.

Gene ontology (GO) assignments were used to classify the putative functions of the differentially expressed *Chara braunii* genes

**TABLE 4** Selected up- and down-regulated genes in nodal cells compared to the control.

Gene ID	Description	log <sub>2</sub> FC
g31563	Putative alpha-amylase, 1,4-alpha-glucan-branching protein	5.19
g29739	Caffeoylshikimate esterase	4.93
g48048	Mediator of RNA polymerase II transcription subunit 19a-like	4.61
g19216	Spastin	4.59
g41182	Beta-amylase 3, chloroplast	4.38
g23946	Protein TRIGALACTOSYLDIACYLGLYCEROL 3 (TGD3), chloroplast	4.07
g19355	Serine/threonine-protein kinase EDR1 isoform X1	2.60
g55470	Serine/threonine-protein kinase EDR1-like isoform X1	2.59
g17647	ETHYLENE INSENSITIVE 3-like 1 protein	2.38
g197	Probable ADP-ribosylation factor GTPase-activating protein AGD11	2.34
g84830	START-like domain homeobox protein Wuschel-like	2.33
g30098	ETHYLENE INSENSITIVE2	2.11
g32224	PAP11, MurE-like_UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-2, 6-diaminopimelate ligase	1.17
g19139	Chlorophyll a-b binding protein 50, chloroplast	-1.44
g51450	PAP10, thioredoxin-like protein CITRX, chloroplast	-1.45
g40841	Chlorophyll a-b binding protein CP26, chloroplast	-1.91
g20154	Major chlorophyll binding protein	-2.10
g72716	Chlorophyll a-b binding protein P4, chloroplast	-2.27
g19118	Chlorophyll a-b binding protein 151, chloroplast	-2.43
g17032	PAP9, Superoxide Dismutase, FeSOD	-2.57
g19140	Chlorophyll a-b binding protein 50, chloroplast	-2.58
g31512	Photosystem I reaction center subunit VI, chloroplast	-2.89
g19133	Major chlorophyll binding protein	-3.20
g20157	Chlorophyll a-b binding protein 151, chloroplast	-3.26
g50764	Photosystem II repair protein PSB27-H1, chloroplast	-3.30
g61495	Peroxisome protein 2B	-3.40
g48045	PAP5	-3.63
g20151	Chlorophyll a-b binding protein 151, chloroplast	-3.64
g602	Protein LHCP TRANSLOCATION DEFECT	-3.93
g19141	Chlorophyll a-b binding protein 3C, chloroplast	-5.16
g20155	Major chlorophyll binding protein	-5.37
g19126	Chlorophyll a-b binding protein 50, chloroplast	-6.79

Note: The gene ID in the first column is followed by the annotation and the calculated differential expression (log<sub>2</sub>FC). A positive value for the log<sub>2</sub>FC indicates higher read count in the nodal sample (shaded in light brown), while a negative value indicates a higher read count in the control (shaded in light green).

(Figure 3). We sorted the GO assignments according to biological process, cellular components and molecular function, and into the up- or down-regulated categories. Most prevalent among the biological

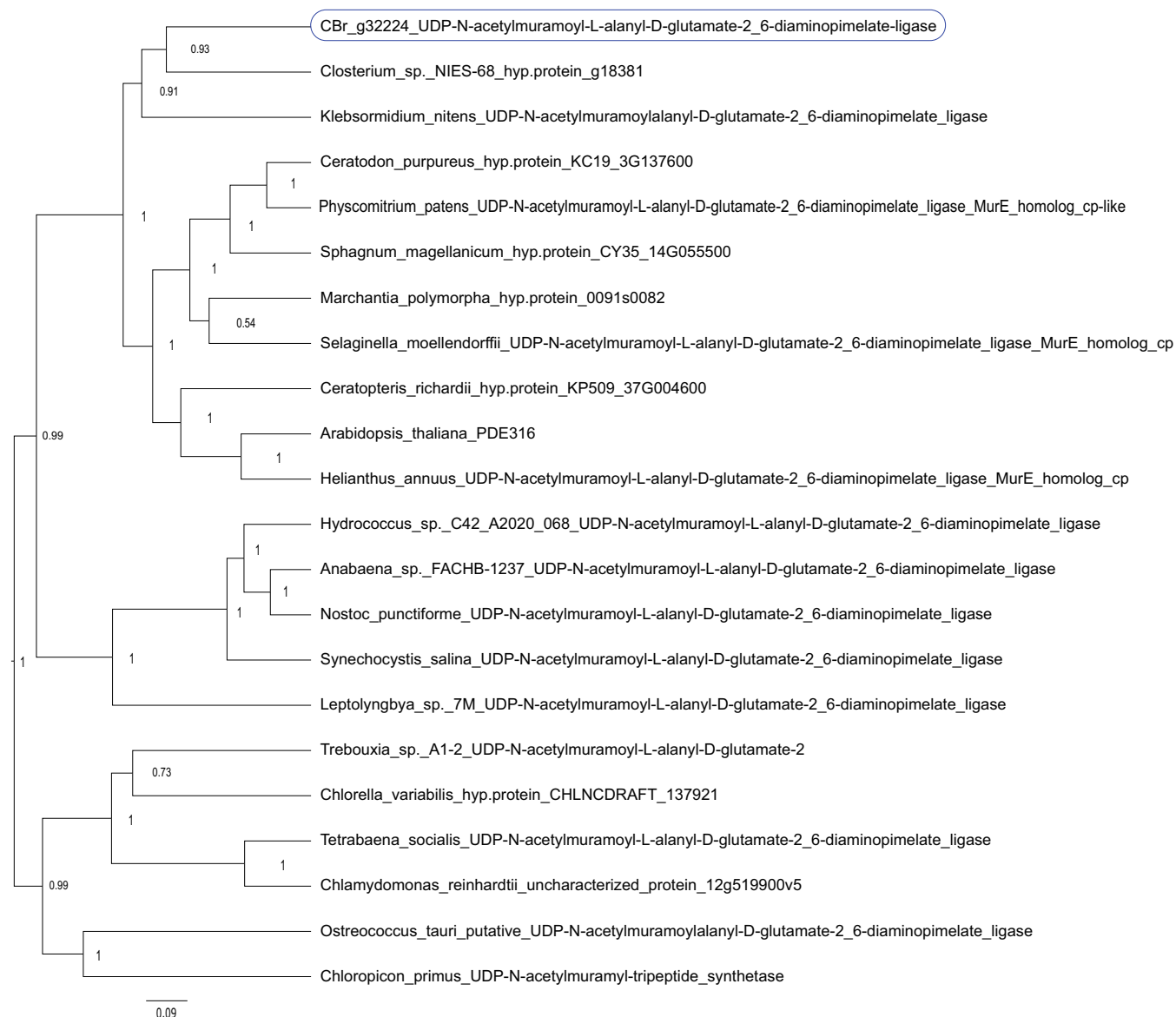
processes upregulated in the nodal sample were the GO terms “DNA integration” (225 genes), “protein phosphorylation” (121 genes), “proteolysis” (119 genes), “metabolic process” (63 genes) and “regulation of transcription, DNA-templated” (50 genes). Prominent cellular component GO terms were “integral component of membrane” (103 genes), “nucleus” (91 genes), “membrane” (78 genes), “intracellular anatomical structure” (59 genes) and “cytoplasm” (39 genes). Among the upregulated GO terms for molecular function, “protein binding” (459 genes), “nucleic acid binding” (517 genes), “ATP binding” (287 genes), “zinc ion binding” (138 genes), and “DNA binding” (138 genes) were highly enriched.

In contrast, the biological process GO terms “protein phosphorylation” (35 genes), “metabolic process” (30 genes), “proteolysis” (27 genes), “transmembrane transport” (27 genes) and “DNA integration” (24 genes) were downregulated. Among downregulated cellular component GO terms, “integral component of membrane” (57 genes), “membrane” (52 genes), “cytoplasm” (24 genes), “nucleus” (14 genes), and “intracellular anatomical structure” (12 genes) were most prevalent. Prominent GO terms for downregulated molecular functions were “protein binding” (101 genes), “ATP binding” (84 genes), “nucleic acid binding” (69 genes), “catalytic activity” (63 genes), and “protein kinase activity” (35 genes). Interestingly, 30 genes associated with the GO term “ribosome” were upregulated and 10 genes downregulated. Within this subset, two paralogous genes encoding 40S ribosomal S21-2 proteins were inversely regulated (g8281, log<sub>2</sub>FC 2.06 and g28482, log<sub>2</sub>FC -1.59), similarly two genes for the 60S ribosomal L22-2 protein were inversely regulated (g50176, log<sub>2</sub>FC 1.40 and g41227, log<sub>2</sub>FC -1.20). Besides their crucial function in translation, ribosomal proteins can have extraribosomal functions. One intriguing example of this is provided by RPL18aB, which is involved in sexual reproduction and male gametophyte development in *Arabidopsis* (Xiong et al., 2021; Yan et al., 2016). A putative 60S ribosomal protein L18-2 was found downregulated within the ribosomal subset (g50118, log<sub>2</sub>FC -3.06). A full list of GO terms can be found in Table S3.

Because there were many genes with unclear annotation or similarity to retrotransposons, we focused on genes with an available functional annotation (Table 4). Among these genes, g31563 (log<sub>2</sub>FC + 5.19) and g41182 (log<sub>2</sub>FC + 4.38), which encode proteins with similarity to 1,4-alpha-glucan-branching enzymes/alpha/maltogenic amylase and beta amylase, were enriched in nodal cells. While it is known that streptophyte algae use starch as a storage product (García, 1994), the upregulation of both genes in the nodal cell sample indicates active carbohydrate storage molecule metabolism in the respective cells of *Chara braunii*.

With their publication of the *Chara* draft genome, Nishiyama et al. (2018) predicted the presence of certain factors involved in phytohormone biosynthesis and signaling pathways. Within our data set, we detected several upregulated genes belonging to these pathways. Related to ethylene signaling, we detected three upregulated genes in the nodal samples. The genes g30098 (log<sub>2</sub>FC +2.1) and g17647 (log<sub>2</sub>FC +2.4) encode putative homologs of the ETHYLENE INSENSITIVE 2 and 3 proteins (EIN2 and EIN3) (Nishiyama et al., 2018),



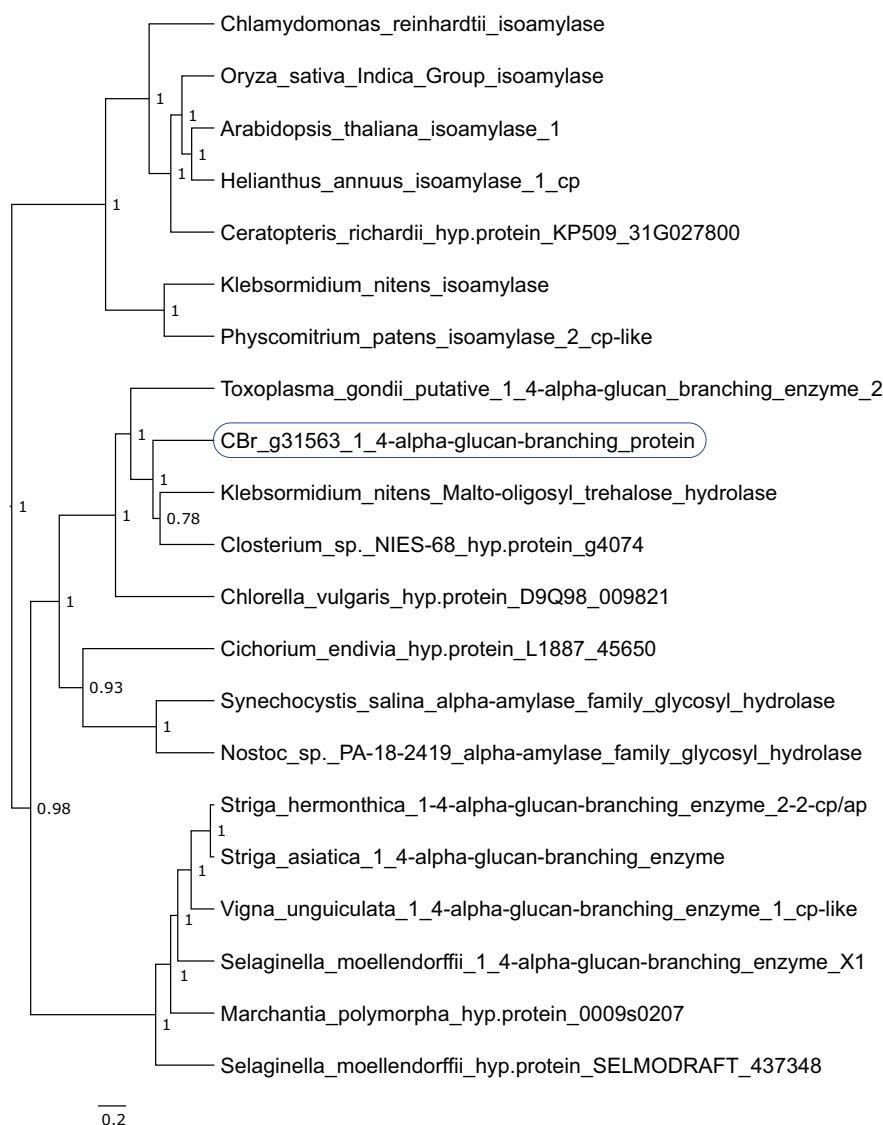


**FIGURE 4** Phylogenetic analysis of a putative *Chara braunii* UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-2, 6-diaminopimelate ligase (PAP11) homolog. Bayesian inference of *Chara braunii* g32224 UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-2, 6-diaminopimelate ligase. Selected homologous proteins (Table S1) were aligned via M-Coffee and analyzed using BEAST 2. The depicted maximum clade credibility tree, generated via tree prior yule model and 1e4 logged MCMC chain length of 1e6, was built using TreeAnnotator with 50% burnin and a posterior probability limit of 0.5 for median node heights. The generated dendrogram was visualized using FigTree (Rambaut, 2018) with posterior probabilities labeled on their respective nodes. The position of the putative *Chara braunii* g32224 UDP-N-acetylmuramoyl-L-alanyl-D-glutamate-2, 6-diaminopimelate ligase homolog is boxed.

proteins that are central in the regulation of ethylene-mediated responses (Chao et al., 1997; Solano et al., 1998). Furthermore, we detected putative serine/threonine-protein kinases EDR1, g19355 ( $\log_2FC +2.6$ ) and g55470 ( $\log_2FC +2.59$ ), homologs of two proteins involved in crosstalk between ethylene-, ABA- and salicylic acid signaling in *A. thaliana* (Tang et al., 2005; Wawrzynska et al., 2008). Additionally, we detected with a  $\log_2FC$  of +4.1 enrichment of g23946, encoding a homolog of the trigalactosyldiacylglycerol 3 (TGD3) protein, involved in chloroplast lipid import (Lu et al., 2007), jasmonic acid signaling and plant defense responses via phosphatidic acid signaling (Tagami et al., 2021).

Regarding further upregulated, signaling-related genes, g197 ( $\log_2FC +2.3$ ), an mRNA encoding a putative ADP-ribosylation factor GTPase-activating protein (ARF), was detected. Another upregulated gene, g48048 ( $\log_2FC +4.6$ ), encoding a putative mediator of RNA polymerase II transcription subunit 19a-like, points towards an increase in transcriptional activity for the actively dividing nodal cells. Another potential regulatory factor enriched in the nodal sample ( $\log_2FC +4.6$ ) is a putative spastin-like protein encoded by g19216.

Within our data set, we detected several genes encoding proteins involved in photosynthetic processes expressed at lower levels in nodal cells. Among them, the genes g19126, g20155, g19141, g20151,



**FIGURE 5** Phylogenetic analysis of a *Chara braunii* protein with similarity to 1,4-alpha-glucan-branching enzymes and alpha amylases. Bayesian inference of *Chara braunii* g31563 protein. Identified homologous proteins were aligned via M-Coffee and analyzed using BEAST 2. The depicted maximum clade credibility tree, generated via tree prior yule model and 1e4 logged MCMC chain length of 1e6, was built using TreeAnnotator with 50% burnin and a posterior probability limit of 0.5 for median node heights. The generated dendrogram was visualized using FigTree (Rambaut, 2018) with posterior probabilities labeled on their respective nodes. The position of the putative *Chara braunii* g31563 alpha amylase homolog is boxed. Two cyanobacterial sequences (*Nostoc* sp. and *Synechocystis salina*) were added as outgroups. The compared protein sequences can be found in Table S1.

g20157 and g19133 ( $\log_2FC$  -6.8, -5.4, -5.2, -3.6, -3.3, -3.2) encode chlorophyll *a/b* binding proteins, apoproteins of the light-harvesting complex of photosystem II (Liu et al., 2013). Additionally, g31512 ( $\log_2FC$  -2.9; encoding the PSI reaction center subunit VI), g50764 ( $\log_2FC$  -3.3; encoding the putative PSII repair protein PSB27) and g602 ( $\log_2FC$  -3.9; encoding a putative LHCP translocation defect protein) point at less active photosynthetic processes in central and nodal cells. Interestingly, g61495 ( $\log_2FC$  -3.4), a putative peroxiredoxin involved in cellular protection from reactive oxygen species (ROS), was downregulated as well. Besides their crucial scavenging capabilities, however, peroxiredoxins also function as modulators of stress response pathways with their interaction with transcription factors (Bréhélin et al., 2003; Hopkins & Neumann, 2019).

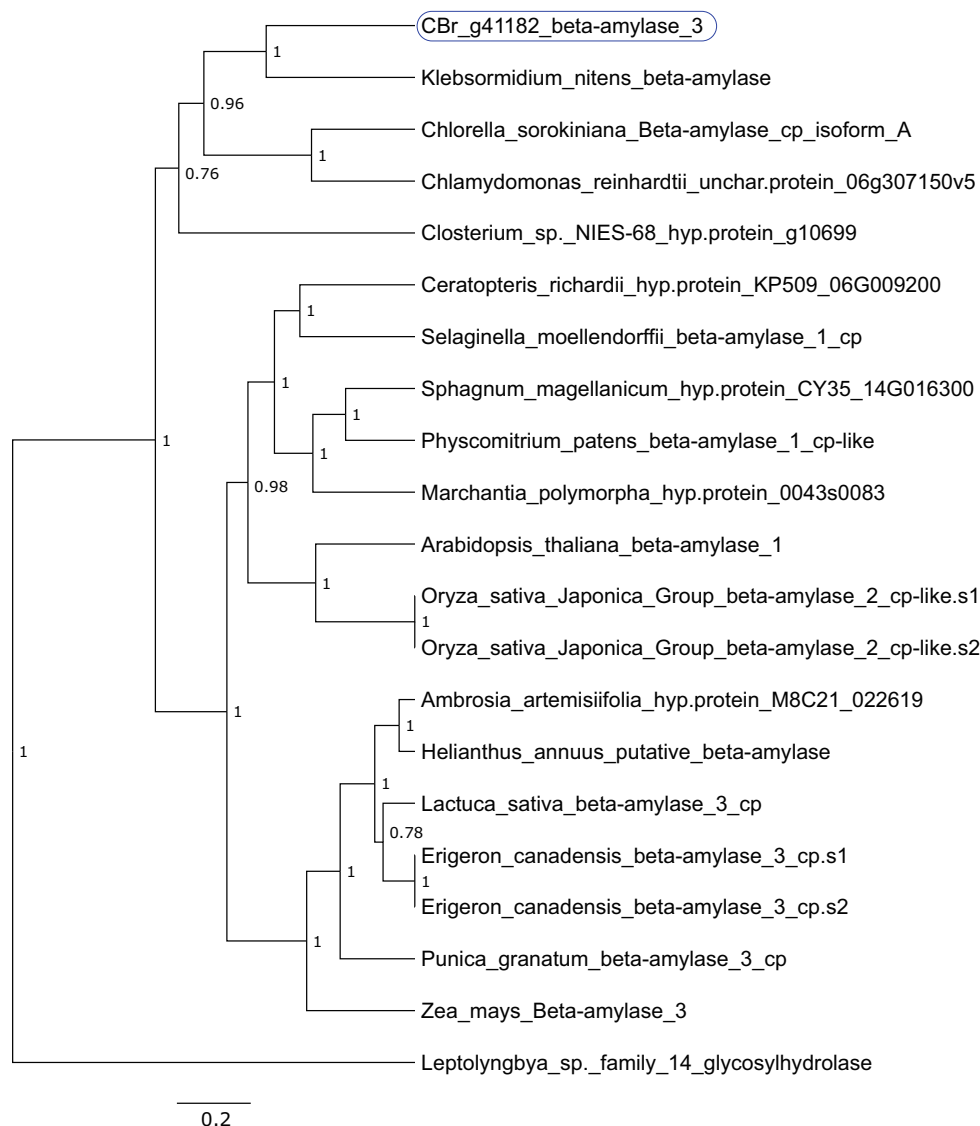
The primary cell walls of streptophyte algae contain pectin, hemicellulose and cellulose, and resemble a basal version of those in land plants (Domozych & Bagdan, 2022). Within our data set, we detected mRNAs of several cell wall-associated genes. We found putative cellulose synthase catalytic subunits (g8592 and g8591) upregulated in the nodal sample, as well as an expansin-like protein

(g49779) and an mRNA encoding a putative polygalacturonate 4-alpha-galacturonosyltransferase-like protein (g11196). Additionally, a putative cellulose synthase catalytic subunit (g31308) was downregulated, as well as an expansin like protein (g3302), and a putative polygalacturonate 4-alpha-galacturonosyltransferase-like protein (g3195). Furthermore, we detected six putative pectine esterases; Four of which were downregulated (g9040, g11990, g12130, and g17556) and one upregulated (g20314), while one putative pectinesterase/pectinesterase inhibitor 24-like protein was upregulated (g32217). Finally, we detected g29739 ( $\log_2FC$  +4.9), a putative caffeoylshikimate esterase (CSE) involved in the biosynthesis of lignin as enriched in the nodal material (Table 4).

### 3.3 | Divergent phylogenies for paralogous genes highlight evolutionary innovations

Another set of proteins with central functions in the chloroplast are the PEP-associated proteins (PAPs), which control the activity of the

**FIGURE 6** Phylogenetic analysis of a putative *Chara braunii* beta amylase homolog. Bayesian inference of *Chara braunii* g41182 beta amylase homolog phylogeny. Selected homologous proteins (Table S1) were aligned via M-Coffee and analyzed using BEAST 2. The depicted maximum clade credibility tree, generated via tree prior yule model and 1e4 logged MCMC chain length of 1e6, was built using TreeAnnotator with 50% burnin and a posterior probability limit of 0.5 for median node heights. The generated dendrogram was visualized using FigTree (Rambaut, 2018) with posterior probabilities labeled on their respective nodes. The position of the putative *Chara braunii* g41182 beta amylase homolog is boxed. A protein from the cyanobacterium *Leptolyngbya* served as outgroup.



plastid-encoded RNA polymerase (PEP) (Pfalz & Pfannschmidt, 2013). Here, we detected a lower expression for several of these PAPs in the nodal sample compared to the control, with  $\log_2FC$  of  $-3.63$ ,  $-2.57$ , and  $-1.45$  for PAP5, PAP9 and PAP10 (genes g48045, g17032, and g51450), which hence is consistent with their role in photosynthetically active cells. However, we detected for one of these proteins, the MurE-like PAP11 (gene g32224), enrichment in the nodal material ( $\log_2FC +1.17$ ). Phylogenetic analysis of the putative MurE-like PAP11 gene (g32224) indicated that it is most closely related to homologs from *Closterium* sp. NIES-68 (67/83% identical/similar amino acids in a 531-residue overlap between g32224 and protein GJP33877.1) and *K. nitens* (60/75% identical/similar amino acids in a 586-residue overlap with protein GAQ90408.1), which form a shallow sister group to mosses, ferns and land plant PAPs (Figure 4) and, therefore, do not contradict a divergent functionality of the PAP11 homolog in *Chara braunii*. Surprisingly, we detected a deep divergence among the homologs in some green algae. These homologs, found in *Chloropicon primus*, *Trebouxia*, *Chlorella variabilis*, *Tetrahena socialis*,

*Chlamydomonas reinhardtii*, and *Ostreococcus tauri* form a distinct group against homologs from all other organisms (Figure 4, 43/56% identical/similar amino acids in a 589-residue overlap between g32224 and protein QDZ19900.1 of *Chloropicon primus*).

The putative 1,4-alpha-glucan-branching enzyme/alpha amylase g31563 and the beta amylase g41182 highly activated in nodal cells also showed an interesting phylogenetic divergence (Figures 5 and 6). g31563 encodes a protein with high sequence conservation to homologs from other algae (51% identical and 68% similar amino acids in a 572-residues overlap with the homolog in the Klebsormidiophyceae alga *K. nitens* (accession no. GAQ89829) and 47% identical and 65% similar amino acids in a 585 residues-long overlap with the homolog in the Zygnematophyceae algae *Closterium* sp. NIES-68 (accession no. GJP44716)). The next-closely related proteins are mainly from Apicomplexa, such as *Toxoplasma gondii*, and other Alveolata, the Chlamydomonadales and other green algae.

The most closely related protein to the protein encoded by g41182 is, again, a homolog from *K. nitens* (52/69% identical/similar

amino acids in a 485-residues overlap with protein GAQ79925) and homologs from *Chlorella vulgaris*, *Chlamydomonas reinhardtii*, and *Closterium* sp. NIES-68 (Figure 6). However, these form a shallow sister group to proteins from land plants, indicating a closer relationship to homologs in land plants (e.g., 49/61% identical/similar amino acids in a 505 residues-long overlap between the *Chara braunii* g41182 gene product and protein KAJ0630873 from sunflower (*Helianthus annuus*)).

## 4 | DISCUSSION

Transcriptome profiling has provided comprehensive and detailed insight into the stress responses of streptophyte alga, including *Chara globularis* (de Vries et al., 2018; Liang et al., 2020). Using this powerful approach, we aimed to obtain a first insight into possible cell-type dependent differences in gene expression in *Chara braunii* by studying the transcriptome of *Chara* nodal cells under optimized, basal growth conditions. We found that the two transcriptome samples differed enormously with 3145 differentially regulated genes. The fact that most of the differentially regulated genes were upregulated in the nodal cell sample compared to the control indicates active transcription of genes across a broader functional spectrum in these cells.

In the control sample, the enriched mRNAs were mostly associated with photosynthesis and protein synthesis, consistent with the photosynthetic lifestyle and fast growth of the *Chara braunii* algae. Accordingly, chloroplasts are of fundamental physiological relevance. Regarding regulation in the chloroplast, it is noteworthy that we detected the enrichment of mRNAs encoding several PEP-associated proteins (PAPs) in the total thallus material, consistent with their control function over the activity of the plastid-encoded RNA polymerase (PEP) (Pfalz & Pfannschmidt, 2013). In land plants such as *Arabidopsis thaliana*, 12 PAPs were described (Pfalz & Pfannschmidt, 2013). These PAPs are associated with the development of chloroplasts in photosynthetically active cells and therefore constitute highly selective cell type-specific markers. Ten PAP orthologs were previously detected in *C. braunii*, in contrast to only 5 or 8 in *C. reinhardtii* and *Klebsormidium nitens*, respectively (Nishiyama et al., 2018). Here, we detected the transcription of PAP5, PAP9, PAP10, and PAP11. Transcripts for the first three were more abundant in the whole thallus sample, consistent with their role in PEP regulation in the chloroplast. In contrast, PAP11 was upregulated in the nodal cells, pointing at its divergent function in these cells. This is consistent with its phylogenetic separation, along with homologs from other algae, from land plants homologs (Figure 4). The essential role of PAP11 in chloroplast development in *Arabidopsis* has been established (Garcia et al., 2008). However, PAP11 possesses a MurE ligase domain (Garcia et al., 2008; Pfalz & Pfannschmidt, 2013; Pfannschmidt et al., 2015), suggesting involvement in the biosynthesis of peptidoglycans. Its divergent regulation suggests that PAP11 in *Chara braunii* may be less relevant in the regulation of PEP in the chloroplasts, and may rather represent a non-PAP version (Nishiyama et al., 2018).

We detected high transcript levels for several ribosomal proteins or proteins associated to ribosomes. Some of the ribosomal proteins encoded by the detected mRNAs are universally found throughout the tree of life (Lan et al., 2022), and some also facilitate extraribosomal functions (reviewed in Xiong et al., 2021). Rpl6 has been shown to promote growth, productivity and tolerance towards salt stress in rice (Moin et al., 2021). Rpl35a in humans has been shown to inhibit cell death (Lopez et al., 2002). Finally, Rpl10 is not only highly conserved due to its critical function in joining the 40S and 60S ribosomal subunits into a functional 80S ribosome, but is also involved in cell growth and resistance to non-host disease and reactive oxygen species in plants (Ferreyra et al., 2010; Ramu et al., 2020). As reviewed in (Li & Wang, 2020), ribosomes can be highly heterogeneous and able to preferentially translate specific subsets of mRNAs. Paralogues of ribosomal protein are seldom found in mammals but exist widely in plants, yeast and bacteria (Li & Wang, 2020). Variations in the building blocks and assembly of ribosomal subunits could therefore enable ribosomes to function as key players in the regulation of translational control and an organism's broader development. Our data are an indication that these mechanisms are also active in *Chara braunii*.

The high transcript level of g46838, encoding a DNA-binding factor-like protein with similarity to the PAX3- and PAX7-binding proteins, in both samples indicates a likely more general role of this factor. Homologs of this protein in animals are involved in epigenetic regulation (Diao et al., 2012), but there is no information on putative plant homologs.

The enrichment of g19216, encoding a putative spastin-like protein, in the nodal sample has probably broader implications for its stated function in the coordination of microtubule and endoplasmic reticulum dynamics. While spastins modulate the lipid droplet network and their dispersion throughout the cell in animals, this has been speculated to be an ancestral process in eukaryotes, as antecedent ER elongation along microtubules is a conserved mechanism also described in plants (Arribat et al., 2020; Hamada et al., 2014). Hence, this view is supported here.

A substantial part of a plant genome is devoted to genes encoding enzymes for the biosynthesis of cell wall components. The primary cell walls of streptophyte algae and their mode of non-enzymatic expansion, as well as the precise wall composition, have been extensively studied (Anderson & King, 1961a, b; Boyer, 2009, 2016; Domozych & Bagdan, 2022; Höfte et al., 2012; Pfeifer et al., 2023). They consist of cellulose embedded within a pectin and hemicellulose matrix and therefore resemble those of land plants in many aspects (Sørensen et al., 2011). We detected mRNAs for several enzymes associated with cell wall synthesis in our data set. While most of these genes were down-regulated in the nodal sample, some were upregulated, pointing at possibly divergent roles of the respective proteins in nodal compared to other cells. Among these upregulated genes was g29739, a putative caffeoylshikimate esterase (CSE) involved in the biosynthesis of lignin, which was already confirmed to be present in *Chara* through biochemical analysis (Rekha & Sujathamma, 2020).

The precise functional characterization of the mentioned upregulated gene functions in nodal cells is of great interest for future analyses.

Overall, the analysis of the nodal cell sample indicated the enrichment of transcripts associated with signaling, protein biosynthesis, starch metabolism and ethylene signaling pathways, while the whole thallus sample showed enrichment of transcripts associated with photosynthesis, chloroplasts and primary metabolism.

## AUTHOR CONTRIBUTIONS

Wolfgang R. Hess conceptualized the study. Anja Holzhausen established the dissecting method, Daniel Heß and Anja Holzhausen dissected nodal cells, Daniel Heß generated RNA samples and performed all bioinformatic analyses. Daniel Heß and Wolfgang R. Hess drafted and Anja Holzhausen edited the manuscript and all authors read and approved the final manuscript.

## ACKNOWLEDGMENTS

We thank the Freiburg Galaxy team for running this great infrastructure, Madeleine Tarika Maier for maintaining vegetative cultures of *Chara braunii*, Manuel Brenes-Álvarez for assistance with bioinformatic analyses and the research group of Annegret Wilde (all University of Freiburg) for use of their mixer mill. This work was supported by the Deutsche Forschungsgemeinschaft (DFG) priority program 2237 “MAdLand”, <http://madland.science>, grant HE2544/18-1 to Wolfgang R. Hess. The Galaxy server that was used for some calculations is in part funded by the Collaborative Research Centre 992 Medical Epigenetics (DFG grant SFB 992/1 2012) and the German Federal Ministry of Education and Research (BMBF grants 031 A538A/A538C RBC, 031L0101B/031L0101C de.NBI-epi, and 031L0106 de. STAIR (de.NBI)). Open Access funding enabled and organized by Projekt DEAL.

## DATA AVAILABILITY STATEMENT

Raw Illumina RNA-seq data produced in this study have been deposited in the Sequence Read Archive (SRA) with the accession numbers SRX19048645 and SRX19048646 under BioProject PRJNA924555 at <https://www.ncbi.nlm.nih.gov/sra/?term=PRJNA924555>.

## ORCID

Wolfgang R. Hess  <https://orcid.org/0000-0002-5340-3423>

## REFERENCES

- Anderson, D.M.W. & King, N.J. (1961a) Polysaccharides of the characeae II. The carbohydrate content of *Nitella translucens*. *Biochimica et Biophysica Acta*, 52, 441–449.
- Anderson, D.M.W. & King, N.J. (1961b) Polysaccharides of the characeae III. The carbohydrate content of *Chara australis*. *Biochimica et Biophysica Acta*, 52, 449–454.
- Andrews, S. (2010) FastQC: a quality control tool for high throughput sequence data. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> Accessed 13 January 2023.
- Arribat, Y., Grepper, D., Lagarrigue, S., Qi, T., Cohen, S. & Amati, F. (2020) Spastin mutations impair coordination between lipid droplet dispersion and reticulum. *PLoS Genetics*, 16, e1008665.
- Batut, B., Freeberg, M., Heydarian, M., Erxleben, A., Videm, P., Blank, C. et al. (2022) Reference-based RNA-seq data analysis. <https://training.galaxyproject.org/training-material/topics/transcriptomics/tutorials/reference-based/tutorial.html> Accessed 13 January 2023.
- Batut, B., Hiltmann, S., Bagnacani, A., Baker, D., Bhardwaj, V., Blank, C. et al. (2018) Community-driven data analysis training for biology. *Cell Systems*, 6, 752–758.e1.
- Beilby, M.J. (2016) Multi-scale characean experimental system: from electrophysiology of membrane transporters to cell-to-cell connectivity, cytoplasmic streaming and auxin metabolism. *Frontiers in Plant Science*, 7, 1052.
- Bonnot, C., Hetherington, A.J., Champion, C., Breuninger, H., Kelly, S. & Dolan, L. (2019) Neofunctionalisation of basic helix-loop-helix proteins occurred when embryophytes colonised the land. *The New Phytologist*, 223, 993–1008.
- Bouckaert, R., Vaughan, T.G., Barido-Sottani, J., Duchêne, S., Fourment, M., Gavryushkina, A. et al. (2019) BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. *PLoS Computational Biology*, 15, e1006650.
- Boyer, J.S. (2009) Evans review: cell wall biosynthesis and the molecular mechanism of plant enlargement. *Functional Plant Biology*, 36, 383–394.
- Boyer, J.S. (2016) Enzyme-less growth in *Chara* and terrestrial plants. *Frontiers in Plant Science*, 7, 866.
- Braun, M. & Limbach, C. (2006) Rhizoids and protonemata of characean algae: model cells for research on polarized growth and plant gravity sensing. *Protoplasma*, 229, 133–142.
- Bréhélin, C., Meyer, E.H., de Souris, J.-P., Bonnard, G. & Meyer, Y. (2003) Resemblance and dissemblance of *Arabidopsis* type II peroxiredoxins: similar sequences for divergent gene expression, protein localization, and activity. *Plant Physiology*, 132, 2045–2057.
- Buschmann, H. (2020) Into another dimension: how streptophyte algae gained morphological complexity. *Journal of Experimental Botany*, 71, 3279–3286.
- Casanova, M.T. (2015) A revision of *Chara* sect. *Charopsis* (Characeae: Charophyceae) in Australia, including specimens collected for Bush Blitz. *Australian Systematic Botany*, 27, 403–414.
- Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W. & Ecker, J.R. (1997) Activation of the ethylene gas response pathway in *Arabidopsis* by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. *Cell*, 89, 1133–1144.
- Cheng, S., Xian, W., Fu, Y., Marin, B., Keller, J., Wu, T. et al. (2019) Genomes of subaerial Zygnematophyceae provide insights into land plant evolution. *Cell*, 179, 1057–1067.e14.
- Chomczynski, P. & Sacchi, N. (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Analytical Biochemistry*, 162, 156–159.
- Chomczynski, P. & Sacchi, N. (2006) The single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction: twenty-something years on. *Nature Protocols*, 1, 581–585.
- de Vries, J. & Archibald, J.M. (2018) Plant evolution: landmarks on the path to terrestrial life. *New Phytologist*, 217, 1428–1434.
- de Vries, J., Curtis, B.A., Gould, S.B. & Archibald, J.M. (2018) Embryophyte stress signaling evolved in the algal progenitors of land plants. *Proceedings of the National Academy of Sciences USA*, 115, E3471–E3480.
- Di Tommaso, P., Moretti, S., Xenarios, I., Orobítz, M., Montanyola, A., Chang, J.-M. et al. (2011) T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Research*, 39, W13–W17.
- Diao, Y., Guo, X., Li, Y., Sun, K., Lu, L., Jiang, L. et al. (2012) Pax3/7BP is a Pax7- and Pax3-binding protein that regulates the proliferation of muscle precursor cells by an epigenetic mechanism. *Cell Stem Cell*, 11, 231–241.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S. et al. (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, 29, 15–21.



- Domozych, D.S. & Bagdan, K. (2022) The cell biology of charophytes: exploring the past and models for the future. *Plant Physiology*, 190, 1588–1608.
- Donoghue, P. & Paps, J. (2020) Plant evolution: assembling land plants. *Current Biology*, 30, R81–R83.
- Ewels, P., Magnusson, M., Lundin, S. & Käller, M. (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*, 32, 3047–3048.
- Ferreira, M.L.F., Pezza, A., Biarc, J., Burlingame, A.L. & Casati, P. (2010) Plant L10 ribosomal proteins have different roles during development and translation under ultraviolet-B stress. *Plant Physiology*, 153, 1878–1894.
- Fürst-Jansen, J.M.R., de Vries, S. & de Vries, J. (2020) Evo-physio: on stress responses and the earliest land plants. *Journal of Experimental Botany*, 71, 3254–3269.
- García, A. (1994) Charophyta: their use in paleolimnology. *Journal of Paleolimnology*, 10, 43–52.
- García, M., Myouga, F., Takechi, K., Sato, H., Nabeshima, K., Nagata, N. et al. (2008) An Arabidopsis homolog of the bacterial peptidoglycan synthesis enzyme MurE has an essential role in chloroplast development. *The Plant Journal*, 53, 924–934.
- Gmelin, K.C. (1826) Flora Badensis alsatica: et confinium regionum Cis et Transrhenana plantas a Lacu Bodamico usque ad confluentem Mosellae et Rheni sponte nascentes exhibens secundum systema sexuale cum iconibus ad naturam delineatis. Mülleriana.
- Hamada, T., Ueda, H., Kawase, T. & Hara-Nishimura, I. (2014) Microtubules contribute to tubule elongation and anchoring of endoplasmic reticulum, resulting in high network complexity in *Arabidopsis*. *Plant Physiology*, 166, 1869–1876.
- Hess, S., Williams, S.K., Busch, A., Irisarri, I., Delwiche, C.F., de Vries, S. et al. (2022) A phylogenomically informed five-order system for the closest relatives of land plants. *Current Biology*, 32, 4473–4482.e7.
- Höfte, H., Peaucelle, A. & Braybrook, S. (2012) Cell wall mechanics and growth control in plants: the role of pectins revisited. *Frontiers in Plant Science*, 3, 121.
- Holzhausen, A., Stingl, N., Rieth, S., Kühn, C., Schubert, H. & Rensing, S.A. (2022) Establishment and optimization of a new model organism to study early land plant evolution: germination, cultivation and oospore variation of *Chara braunii* Gmelin, 1826. *Frontiers in Plant Science*, 13, 987741.
- Hopkins, B.L. & Neumann, C.A. (2019) Redoxins as gatekeepers of the transcriptional oxidative stress response. *Redox Biology*, 21, 101104.
- Kawai, H., Hanyuda, T., Akita, S. & Uwai, S. (2022) The macroalgal culture collection in Kobe University (KU-MACC) and a comprehensive molecular phylogeny of macroalgae based on the culture strains. *Applied Phycology*, 3, 159–166.
- Kuczewski, O. (1906) Morphologische und biologische Untersuchungen an *Chara delicatula* f. *bulbillifera* A. Braun. Arbeiten aus dem Laboratorium für allgemeine Botanik und Pflanzenphysiologie der Universität Zürich 1–51.
- Lan, T., Xiong, W., Chen, X., Mo, B. & Tang, G. (2022) Plant cytoplasmic ribosomal proteins: an update on classification, nomenclature, evolution and resources. *The Plant Journal*, 110, 292–318.
- Li, D. & Wang, J. (2020) Ribosome heterogeneity in stem cells and development. *Journal of Cell Biology*, 219, e202001108.
- Liang, Z., Geng, Y., Ji, C., Du, H., Wong, C.E., Zhang, Q. et al. (2020) *Mesostigma viride* genome and transcriptome provide insights into the origin and evolution of Streptophyta. *Advancement of Science*, 7, 1901850.
- Liu, R., Xu, Y.-H., Jiang, S.-C., Lu, K., Lu, Y.-F., Feng, X.-J. et al. (2013) Light-harvesting chlorophyll a/b-binding proteins, positively involved in abscisic acid signalling, require a transcription repressor, WRKY40, to balance their function. *Journal of Experimental Botany*, 64, 5443–5456.
- Lopez, C.D., Martinovsky, G. & Naumovski, L. (2002) Inhibition of cell death by ribosomal protein L35a. *Cancer Letters*, 180, 195–202.
- Lu, B., Xu, C., Awai, K., Jones, A.D. & Benning, C. (2007) A small ATPase protein of Arabidopsis, TGD3, involved in chloroplast lipid import. *Journal of Biological Chemistry*, 282, 35945–35953.
- Martin, M. (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, 17, 10–12.
- Martin, W.F. & Allen, J.F. (2018) An algal greening of land. *Cell*, 174, 256–258.
- Mikkelsen, M.D., Harholt, J., Ulvskov, P., Johansen, I.E., Fangel, J.U., Doblin, M.S. et al. (2014) Evidence for land plant cell wall biosynthetic mechanisms in charophyte green algae. *Annals of Botany*, 114, 1217–1236.
- Moin, M., Saha, A., Bakshi, A., Madhav, M.S. & Kirti, P. (2021) Constitutive expression of ribosomal protein L6 modulates salt tolerance in rice transgenic plants. *Gene*, 789, 145670.
- Moody, L.A. (2020) Three-dimensional growth: a developmental innovation that facilitated plant terrestrialization. *Journal of Plant Research*, 133, 283–290.
- Nishiyama, T., Sakayama, H., de Vries, J., Buschmann, H., Saint-Marcoux, D., Ullrich, K.K. et al. (2018) The *Chara* genome: secondary complexity and implications for plant terrestrialization. *Cell*, 174, 448–464.e24.
- Notredame, C., Higgins, D.G. & Heringa, J. (2000) T-coffee: a novel method for fast and accurate multiple sequence alignment. *Journal of Molecular Biology*, 302, 205–217.
- Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B.L., Salazar, G.A. et al. (2023) InterPro in 2022. *Nucleic Acids Research*, 51, D418–D427.
- Pfalz, J. & Pfannschmidt, T. (2013) Essential nucleoid proteins in early chloroplast development. *Trends in Plant Science*, 18, 186–194.
- Pfannschmidt, T., Blanvillain, R., Merendino, L., Courtois, F., Chevalier, F., Liebers, M. et al. (2015) Plastid RNA polymerases: orchestration of enzymes with different evolutionary origins controls chloroplast biogenesis during the plant life cycle. *Journal of Experimental Botany*, 66, 6957–6973.
- Pfeifer, L., Mueller, K.-K., Utermöhlen, J., Erdt, F., Zehge, J.B.J., Schubert, H. et al. (2023) The cell walls of different *Chara* species (Charophyceae) are characterized by branched galactans rich in 3-O-methylgalactose and absence of arabinogalactan-proteins. 2023.05.22.541140.
- Pinto, F.L., Thapper, A., Sontheim, W. & Lindblad, P. (2009) Analysis of current and alternative phenol based RNA extraction methodologies for cyanobacteria. *BMC Molecular Biology*, 10, 79.
- Pringsheim, N. (1863) Ueber die Vorkeime und die nacktfüssigen Zweige der Charen. *Jahrbücher für wissenschaftliche Botanik*, 3, 294–324.
- Rambaut, A. (2018) FigTree software, version 1.4.4. <http://tree.bio.ed.ac.uk/software/figtree/> Accessed 10 February 2023.
- Rambaut, A., Drummond, A.J., Xie, D., Baele, G. & Suchard, M.A. (2018) Posterior summarization in bayesian phylogenetics using tracer 1.7. *Systematic Biology*, 67, 901–904.
- Ramu, V.S., Dawane, A., Lee, S., Oh, S., Lee, H.-K., Sun, L. et al. (2020) Ribosomal protein QM/RPL10 positively regulates defence and protein translation mechanisms during nonhost disease resistance. *Molecular Plant Pathology*, 21, 1481–1494.
- Rekha, A. & Sujathamma, P. (2020) Biochemical, phytochemical and antibacterial analysis of *Chara braunii* c.c. gmelin. *International Journal of Botany Studies*, 5, 358–360.
- Ruhfel, B.R., Gitzendanner, M.A., Soltis, P.S., Soltis, D.E. & Burleigh, J.G. (2014) From algae to angiosperms-inferring the phylogeny of green plants (Viridiplantae) from 360 plastid genomes. *BMC Evolutionary Biology*, 14, 23.
- Sato, M., Sakayama, H., Sato, M., Ito, M. & Sekimoto, H. (2014) Characterization of sexual reproductive processes in *Chara braunii* (Charales, Charophyceae). *Phycological Research*, 62, 214–221.
- Scholz, I., Lott, S.C., Behler, J., Gärtner, K., Hagemann, M. & Hess, W.R. (2019) Divergent methylation of CRISPR repeats and cas genes in a subtype I-D CRISPR-Cas-system. *BMC Microbiology*, 19, 147.1–147.11.
- Schubert, H., Holzhausen, A. & Nowak, P. (2016) *Individualentwicklung der Characeen*. Armleuchteralgen: Die Characeen Deutschlands. Springer, Berlin, Heidelberg, pp. 57–78.

- Solano, R., Stepanova, A., Chao, Q. & Ecker, J.R. (1998) Nuclear events in ethylene signaling: a transcriptional cascade mediated by ETHYLENE-INSENSITIVE3 and ETHYLENE-RESPONSE-FACTOR1. *Genes & Development*, 12, 3703–3714.
- Sørensen, I., Pettolino, F.A., Bacic, A., Ralph, J., Lu, F., O'Neill, M.A. et al. (2011) The charophycean green algae provide insights into the early origins of plant cell walls. *The Plant Journal*, 68, 201–211.
- Tagami, S., Ohnishi, K., Hikichi, Y. & Kiba, A. (2021) Trigalactosyldiacylglycerol 3 protein orthologs are required for basal disease resistance in *Nicotiana benthamiana*. *Plant Biotechnol (Tokyo)*, 38, 373–378.
- Tang, D., Christiansen, K.M. & Innes, R.W. (2005) Regulation of plant disease resistance, stress responses, cell death, and ethylene signaling in *Arabidopsis* by the EDR1 protein kinase. *Plant Physiology*, 138, 1018–1026.
- The Galaxy Community. (2022) The galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research*, 50, W345–W351.
- Umen, J.G. (2014) Green algae and the origins of multicellularity in the plant kingdom. *Cold Spring Harbor Perspectives in Biology*, 6, a016170.
- Wawrzynska, A., Christiansen, K.M., Lan, Y., Rodibaugh, N.L. & Innes, R.W. (2008) Powdery mildew resistance conferred by loss of the ENHANCED DISEASE RESISTANCE1 protein kinase is suppressed by a missense mutation in KEEP ON GOING, a regulator of abscisic acid signaling. *Plant Physiology*, 148, 1510–1522.
- Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N. et al. (2014) Phylotranscriptomic analysis of the origin and early diversification of land plants. *Proceedings of the National Academy of Sciences USA*, 111, E4859–E4868.
- Wodniok, S., Brinkmann, H., Glöckner, G., Heide, A.J., Philippe, H., Melkonian, M. et al. (2011) Origin of land plants: do conjugating green algae hold the key? *BMC Evolutionary Biology*, 11, 104.
- Xiong, W., Lan, T. & Mo, B. (2021) Extraribosomal functions of cytosolic ribosomal proteins in plants. *Frontiers in Plant Science*, 12, 607157.
- Yan, H., Chen, D., Wang, Y., Sun, Y., Zhao, J., Sun, M. et al. (2016) Ribosomal protein L18aB is required for both male gametophyte function and embryo development in *Arabidopsis*. *Scientific Reports*, 6, 31195.
- Young, M.D., Wakefield, M.J., Smyth, G.K. & Oshlack, A. (2010) Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biology*, 11, R14.
- Zhang, Z., Ma, X., Liu, Y., Yang, L., Shi, X., Wang, H. et al. (2022) Origin and evolution of green plants in the light of key evolutionary events. *Journal of Integrative Plant Biology*, 64, 516–535.

## SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Heß, D., Holzhausen, A. & Hess, W.R. (2023) Insight into the nodal cells transcriptome of the streptophyte green alga *Chara braunii* S276. *Physiologia Plantarum*, 175(5), e14025. Available from: <https://doi.org/10.1111/ppl.14025>