



HAL
open science

Comparative RNA-Seq of Ten *Phaeodactylum tricornutum* Accessions: Unravelling Criteria for Robust Strain Selection from a Bioproduction Point of View

Charlotte Toustou, Isabelle Boulogne, Anne-Alicia Gonzalez, Muriel Bardor

► To cite this version:

Charlotte Toustou, Isabelle Boulogne, Anne-Alicia Gonzalez, Muriel Bardor. Comparative RNA-Seq of Ten *Phaeodactylum tricornutum* Accessions: Unravelling Criteria for Robust Strain Selection from a Bioproduction Point of View. *Marine drugs*, 2024, 22 (8), pp.353. 10.3390/md22080353. hal-04671512

HAL Id: hal-04671512

<https://normandie-univ.hal.science/hal-04671512v1>

Submitted on 15 Aug 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Article

Comparative RNA-Seq of Ten *Phaeodactylum tricornutum* Accessions: Unravelling Criteria for Robust Strain Selection from a Bioproduction Point of View

Charlotte Toustou¹, Isabelle Boulogne^{1,*} , Anne-Alicia Gonzalez²  and Muriel Bardor^{1,3,*} 

¹ Laboratoire GlycoMEV UR 4358, Université de Rouen Normandie, SFR Normandie Végétal FED 4277, Innovation Chimie Carnot, 76000 Rouen, France; charlotte.toustou1@univ-rouen.fr

² MGX-Montpellier GenomiX, Univ. Montpellier, CNRS, INSERM, 34094 Montpellier, France; anne-alicia.gonzalez@inserm.fr

³ ALGA BIOLOGICS, CURIB, 25 rue Tesnières, 76821 Mont-Saint-Aignan, France

* Correspondence: isabelle.boulogne@univ-rouen.fr (I.B.); muriel.bardor@univ-rouen.fr (M.B.)

Abstract: The production of biologics in mammalian cells is hindered by some limitations including high production costs, prompting the exploration of other alternative expression systems that are cheaper and sustainable like microalgae. Successful productions of biologics such as monoclonal antibodies have already been demonstrated in the diatom *Phaeodactylum tricornutum*; however, limited production yields still remain compared to mammalian cells. Therefore, efforts are needed to make this microalga more competitive as a cell biofactory. Among the seventeen reported accessions of *P. tricornutum*, ten have been mainly studied so far. Among them, some have already been used to produce high-value-added molecules such as biologics. The use of “omics” is increasingly being described as useful for the improvement of both upstream and downstream steps in bioprocesses using mammalian cells. Therefore, in this context, we performed an RNA-Seq analysis of the ten most used *P. tricornutum* accessions (Pt1 to Pt10) and deciphered the differential gene expression in pathways that could affect bioproduction of biologics in *P. tricornutum*. Our results highlighted the benefits of certain accessions such as Pt9 or Pt4 for the production of biologics. Indeed, these accessions seem to be more advantageous. Moreover, these results contribute to a better understanding of the molecular and cellular biology of *P. tricornutum*.

Keywords: *P. tricornutum*; diatom; microalgae; RNA-Seq; bioproduction; biologics; pharmaceutical; recombinant proteins



Citation: Toustou, C.; Boulogne, I.; Gonzalez, A.-A.; Bardor, M.

Comparative RNA-Seq of Ten

Phaeodactylum tricornutum Accessions:

Unravelling Criteria for Robust Strain

Selection from a Bioproduction Point

of View. *Mar. Drugs* **2024**, *22*, 353.

[https://doi.org/10.3390/](https://doi.org/10.3390/md22080353)

[md22080353](https://doi.org/10.3390/md22080353)

Academic Editor: Min Zhao

Received: 30 April 2024

Revised: 26 July 2024

Accepted: 29 July 2024

Published: 30 July 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Biologics include a wide variety of drugs used to treat life-threatening diseases, ranging from gene or cell therapies to recombinant proteins. Today, recombinant protein therapies represent more than 60% of biopharmaceutical products and their peculiarity is that they must be produced by living systems, either mammalian cells or microorganisms, due to the complexity of the molecules that includes post-translational modifications [1–3]. Currently, recombinant proteins are mainly produced in mammalian systems, especially in CHO (Chinese Hamster Ovary) cells. Although considered as a gold standard, this type of production presents several drawbacks. The heterogeneity of the proteins and the possible contamination of the culture by viruses potentially transmissible to humans or by host cell proteins (HCPs) and DNA imply costly downstream processing steps in order to make the product compliant and safe according to pharmaceutical industry standards [4–6]. In this context, alternatives such as plants and microalgae are being explored [7–10]. Among these systems, microalgae seem to be advantageous due to their cheap culture conditions, their rapid growth rate in large-scale photobioreactors, and their ability to perform protein post-translation modifications [3,11–14]. In addition, using microalgae as cell biofactories

to produce biologics contributes to the decarbonization of the pharmaceutical industry, especially biologics production. Moreover, some microalgae species, such as the diatom *Phaeodactylum tricornerutum*, are considered safe by the Food and Drug Administration authority due to the lack of common pathogens between them and humans. Some of them have also been described as promising for the production of different biologics [3,7,9,12,15–17]. However, the use of microalgae for industrial-scale production of biologics is still hindered by their limited production yields, which currently do not reach the level of CHO cells. Therefore, much work needs to be performed to better understand the biology of these species in order to increase their biologics production yields.

The diatom *Phaeodactylum tricornerutum* (Bohlin, 1897) has been described as a promising alternative for the production of biologics such as monoclonal antibodies (mAbs) directed against the antigens of deadly viruses [16–18] or recombinant antigens that have been used as sensitive diagnostic tools for the detection of the SARS-CoV-2 [19]. It is a well-established model diatom whose first genome sequencing and assembly in 2008 [20] allowed the development of powerful genetic tools like genome editing with TALENs [21] or CRISPR technologies [22–24]. In 2020, the sequencing of the ten most studied accessions of *P. tricornerutum* (Pt1 to Pt10) revealed genetic diversity among these accessions, including CNVs (copy number variations) and INDELS (insertions and deletions) [25]. The analysis of molecular marker genes of these ten accessions allows their classification into four genetic clusters. In addition to genetic diversity, these accessions also differ phenotypically. Indeed, *P. tricornerutum* has the unique ability to be pleiomorphic, and a single accession can switch between three main morphotypes: fusiform, oval, and triradiate [26]. While the fusiform morphotype is predominantly observed in natural conditions, some ecotypes show a natural tendency towards alternative morphotypes. For example, in some culture conditions, the Pt8 accession has been described to be found in triradiate morphotypes [26]; however, this morphotype seems to be unstable [27]. The Pt3 strain, derived from Pt2, was selected for its ability to grow in freshwater media, with cells that predominantly adopt an oval shape [26]. The Pt9 accession, isolated in Micronesia, exhibits fusiform morphotypes at temperatures ranging from 25 to 28 °C but is able to switch to an oval shape when cultured at temperatures between 15 and 19 °C. Similar temperature-dependent transitions are observed for Pt3 and Pt8 when the temperature decreases from 19 to 15 °C. Thus, the appearance of the oval morphotype shows the ability of *P. tricornerutum* to adapt to stress conditions (salinity, temperature, or light), that differ from one strain to another [26,28]. In contrast, some accessions originally isolated from a low-salinity environment, such as Pt7, display a preference for the fusiform morphotype [26]. Recently, seven newly isolated accessions have been phenotypically characterized [29].

In addition to genetic tools, some “omics” data have also been generated: either numerous genomic [30–33], transcriptomic [27,29,34–38], proteomic [39–44], or metabolomic data [45,46]. “Omics” technologies are currently widely used to optimize the production of biologics in CHO cells, either in terms of titer, quality, or cell productivity [47,48]. To date, most of the “omics” data generated on *P. tricornerutum* focus on how environmental conditions—such as nutrient depletion or starvation or exposure to stresses such as light/dark, toxic compounds, or predators—can affect the cell biology of a single accession compared to a control condition. Given the genetic differences and the possible variation in gene expression in response to different culture conditions, the question arises of the differences in gene expression between these accessions grown under identical conditions and growth phases. In a previous meta-analysis, we attempted to identify the best *P. tricornerutum* accession for biologics production. The results of this work suggest that the Pt4, Pt3 oval, and Pt8 strains could be interesting chassis for optimizing the production of recombinant proteins and glycoproteins in *P. tricornerutum* (Boulogne et al. 2024, submitted, Université de Rouen Normandie, Rouen, France, 2024 [49]); however, the limitation of this meta-analysis is the use of datasets from different RNA-Seq studies with similar but not exactly the same culture conditions. Moreover, due to the need to have the closest culture conditions, only three accessions (Pt3, Pt4, and Pt8) and cultures enriched in the

three morphotypes of Pt3 (oval, triradiate, and fusiform) have been compared to Pt1 in this previous meta-analysis.

In this context, the aim of the present study is to analyze the transcriptome of the ten most commonly used accessions of *P. tricornutum* (Pt1 to Pt10), cultivated under the exact same conditions and compared at the same exponential growth phase. This includes analyzing the expression of genes involved in biological pathways that are known to have a significant impact on the production of biologics like protein biosynthesis, protein export and secretion, *N*-glycan biosynthesis, quality control, and proteasomes or those-encoding proteases. All these pathways were compared between the *P. tricornutum* accessions, in order to propose the most suitable one(s) (Figure 1).

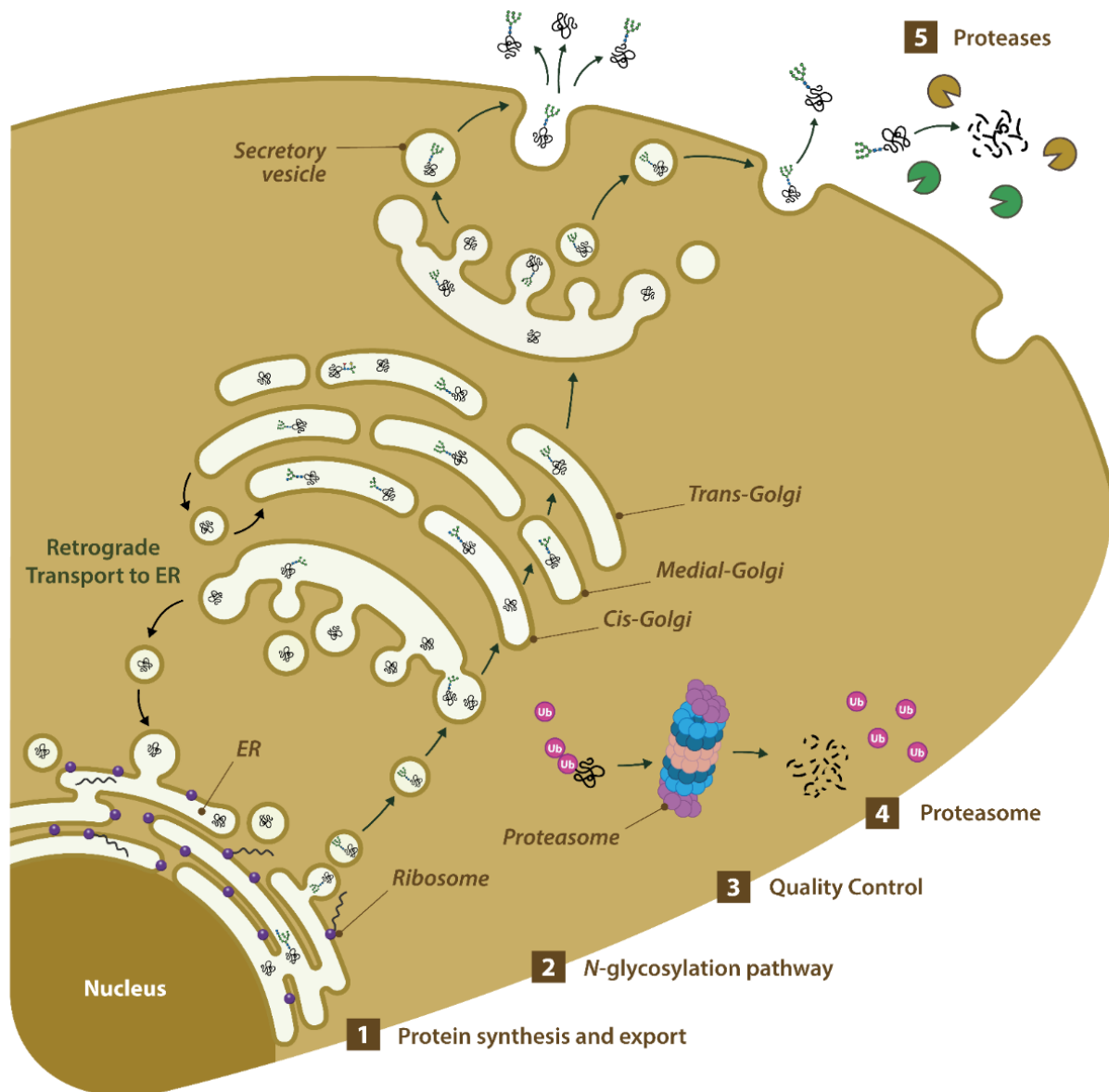


Figure 1. Representation of biological processes that have been shown in other organisms to have a notable influence on the production of biologics and have been studied in *P. tricornutum*.

All these biological pathways were analyzed by comparing the ten accessions with the Pt1 accession, which was used as a reference. Indeed, Pt1 is the strain that the genome was sequenced for in 2008 [20] and represents the lab-scale model. Finally, a focus was made on the expression of specific genes whose promoters have already been used for the production of recombinant proteins in *P. tricornutum*. The results obtained provide key elements to consider when selecting a *P. tricornutum* accession strain for the production of molecules of interest.

2. Results and Discussion

2.1. Growth Kinetics

In order to determine when the cells from the 10 accessions were in the same physiological growth state using our culture conditions, growth kinetics were performed on the 10 accessions of *P. tricornutum* (Figure 2A), allowing us to calculate and compare the growth rate (μ) (Figure 2B) and generation time (G) (Figure S1). In order to compare the gene expression of the 10 accessions in a fair manner, we decided to consider these accessions in the exponential growth phase. Based on this data set, we decided to extract mRNA from cultures at day 4, which represents the time when the 10 accessions were in the same exponential growth phase and where the growth rates are at their highest, as it is under these conditions that protein synthesis is at its highest.

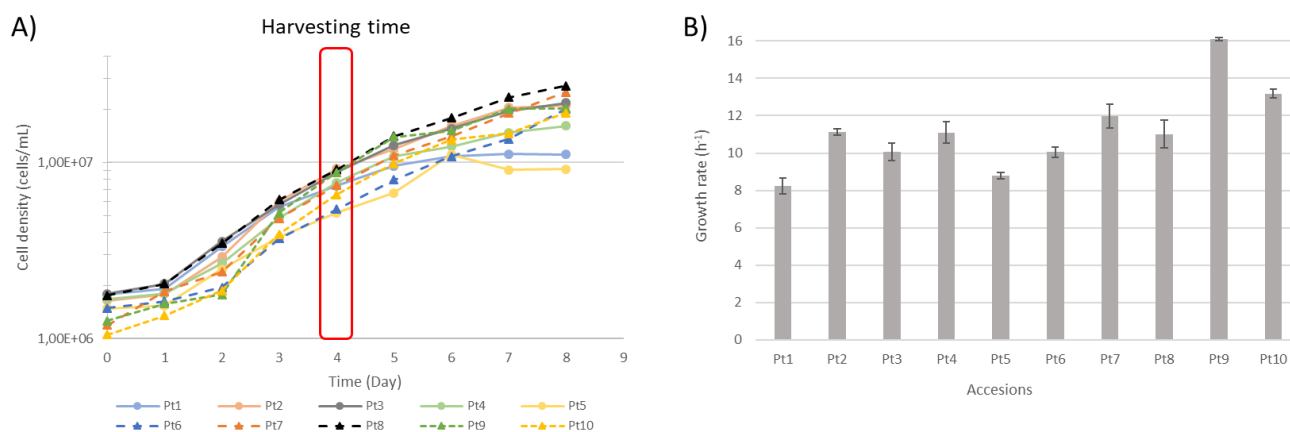


Figure 2. Growth parameters of the 10 most used accessions (Pt1 to Pt10). (A) Growth kinetics of the 10 accessions measured from day 0 (inoculation) to day 8. Cells were harvested at the same growth phase on day 4, which represents the exponential growth phase. (B) Growth rate of the 10 accessions, error bars indicate standard deviations of triplicate measurements.

Under our conditions, the highest growth rates were observed for the Pt9 and Pt10 accessions, with generation times of 1.03 and 1.26 days, respectively. In contrast, the Pt1 and Pt5 accessions grew the slowest, with generation times of 1.89 and 2.02 days, respectively. Surprisingly, another recent publication comparing the phenotypic traits of 17 accessions of *P. tricornutum* showed different results under their culture conditions [29]. Indeed, in their work, although Pt10 was one of the accessions with the fastest growth rate, Pt9 was the accession with the slowest growth rate; however, although the behavior of the accessions is different between their and our culture conditions, the growth rates are quite close for the strain Pt9, with 15.6 h⁻¹ for Chaumier et al. and 16.1 h⁻¹ in our experiment. When comparing the culture conditions in the two studies, the temperature and light in their and our culture conditions are quite close. Therefore, the observed differences can be attributed to either the composition of the culture media or to the inoculum size that was different. Indeed, in the study of Chaumier et al., the accessions were grown in an enriched artificial seawater (EASW) medium with an initial concentration of 1 × 10⁵ cells/mL [29], while in our study *P. tricornutum* accessions were grown in artificial seawater supplemented with Conway's solution with an initial inoculum of 1 × 10⁶ cells/mL. In addition, it has been

shown on other microalgae species such as *Chlorella* and *Scenedesmus* that the concentration of the initial inoculum dose can influence the growth behavior [50–52]. Moreover, a specific study conducted on *C. sorokiniana* showed that higher levels of photosynthesis-related proteins and a high turnover rate were obtained with an inoculum size of 10^6 cells/mL. The authors of this study suggested that the culture with this inoculum size could exploit the potential for more protein synthesis in this microalgae [52]. Thus, it is tempting to hypothesize that a higher expression of endogenous protein may lead to a higher expression of recombinant proteins used as biologics. Considering its faster growth rate, we can speculate that Pt9 could present a greater synthesis of proteins in comparison to other strains, and so could be more of interest.

2.2. Differentially Expressed Genes Analysis

The differential expression of genes (DEGs) analysis in our culture conditions was performed with the ten ecotypes of *P. tricornutum*. Therefore, we compared the accessions from Pt2 to Pt10 with Pt1 as a reference (Figure S2). The accession Pt1 was chosen as a reference since it is the first strain that has its genome sequenced [20] and it is the most used accession so far. The volcano plots showed that there were fewer differences in Pt2 and Pt3. About 2000 genes were differentially expressed between the accessions Pt4, Pt5, Pt6, Pt8, and Pt10 compared to Pt1. Finally, the largest differences in the number of DEGs are for accessions Pt7 and Pt9, with 4219 and 5216 genes that are differentially expressed between these accessions and Pt1, respectively. The genetic diversity of the 10 accessions has been previously studied, which has allowed them to be classified into four clades, regardless of the method used for examination (internal transcribed spacer 2, amplified fragment length polymorphism, or Fixation index) [25,26,29]. Therefore, these studies show that Pt1, Pt2, Pt3, and Pt9 are clustered in genotype A; Pt4 is the only representative of genotype B; Pt5 and Pt10 belong to the genotype C; and, finally, Pt6, Pt7, and Pt8 are grouped in genotype D [25,26,29]. Although the number of DEGs in Pt2 and Pt3 is very low when compared to Pt1, Pt9—which is genetically clustered with Pt1, like Pt2 and Pt3—presents the highest number of DEGs. The observed difference in the number of DEGs may be related to epigenetic [33,53,54] and/or environmental factors such as temperature and probably not to polymorphisms, considering the few differences in terms of SNPs or INDELS within the same clade. For example, under low-temperature conditions, compared with its original environment, a phenotypic change was observed in the Pt9 accession from fusiform to oval morphotypes [28]. Therefore, the large number of DEGs observed in this work could reflect an adaptation of the diatom to the culture conditions by regulating gene expression. Overall, most of the DEGs are under-expressed when comparing the nine other accessions with Pt1 as a reference. The exceptions are Pt3 and Pt7, where the number of over- and under-expressed genes is approximately equal. In addition, the Pt9 strain shows a higher number of over-expressed genes compared to under-expressed genes. The Log₂ FoldChange (Log₂FC) values for under-expressed genes mostly vary between -1 (threshold) and around -11 . For over-expressed genes, the Log₂FC values vary between 1 (threshold) and about 6. Accessions Pt4, Pt6, and Pt10 had a greater number of genes with significantly higher differences in expression compared to the other accessions.

A GO enrichment analysis was then performed on DEGs that were specifically over- and under-expressed using all available gene set databases on ShinyGO (Table S1). Regarding down-regulated genes, the pathways found to be enriched were related to DNA (integration, metabolic processing, or nucleic acid binding) for all the accessions, except for Pt9, which was enriched for pathways related to the plasma membrane cell periphery and protein binding. Regarding up-regulated genes, no significant enrichment was found for Pt2, Pt3, and Pt10. For some accessions, enriched pathways are mainly related to transport or assimilation of elements essential for algal growth. For example, in Pt4, pathways related to nitrate were enriched, while in Pt6, four pathways related to phosphate were enriched. Such observations for Pt4 are in agreement with previous reports. Indeed, the

whole genome sequencing of the 10 accessions of *P. tricornutum* revealed the presence of a higher copy number in the genome of Pt4 of the gene Phatr3_EGO2286, which is involved in nitrate assimilation [25]. In addition, a recent study highlighted a higher nitrogen uptake in Pt4 compared to Pt1 [55]. For Pt5, Pt7, and Pt8, the enriched pathways are mainly related to chlorophyll and photosynthesis. Finally, concerning Pt9, which has the more enriched pathway for up-regulated genes, a lot of enriched pathways were related to translation or peptide biosynthesis.

This first level of analysis may suggest that Pt9 could be of interest for the production of biologics, particularly for increasing production yields, considering the over-representation of genes over-expressed in pathways related to translation and peptide synthesis. To confirm this first level of analysis, we then turned our attention to biological processes that have been shown in other organisms to have a notable influence on the production of biologics (Figure 1): (1) protein synthesis and export (translation, protein export, and secretion), (2) N-glycosylation, (3) quality control, (4) proteasomes, and (5) occurrence of proteases that may accumulate in the culture media and may be responsible for proteolytic degradation of mAbs.

2.3. Targeted Analysis of Interesting Pathway from a Bioproduction Point of View

2.3.1. Protein Synthesis: Translation and Ribosome Biogenesis

In eukaryotes, cytosolic ribosomes are composed of ribosomal RNA and proteins arranged in two subunits, which collaborate to guarantee accurate decoding of the mRNA sequence and the proper assembly of amino acids into a polypeptide chain [56,57]. There, we can suppose that an up-regulation of genes involved in metabolic pathways such as translation or ribosome biogenesis may result in enhanced protein synthesis and could be beneficial for improving the production yield of recombinant proteins in *P. tricornutum*. Using KEGG, 75 homologous genes encoding putative cytosolic 40S and 60S proteins were identified in the genome of *P. tricornutum* (Table S2). No significant DEGs were observed for Pt2 and Pt3. These two accessions are clustered together and are the more distant from the cluster formed by Pt7 and Pt9, which are the accessions with more DEGs (Figure 3A) in comparison to Pt1. Pt7 is the accession with the most under-expressed genes, while Pt9 is the accession with the most over-expressed genes. These observations are consistent with the previous ShinyGO analysis, which indicates that Pt9 has some enriched pathways with up-related genes related to ribosomes and translation (Table S1).

Organellar ribosomes, such as mitochondrial/chloroplastic ribosomes, are also found throughout the Eukarya and Archaeplastida kingdoms. Thirteen homologous genes encoding mitochondrial ribosomal proteins were identified in the *P. tricornutum* genome (Table S2). The clustering profile of the different accessions is the same between the expression of genes encoding cytosolic ribosomes and those encoding mitochondrial/chloroplastic ribosomes. The cluster formed by Pt2 and Pt3 is still the most distant from that formed by Pt7 and Pt9. Once again, Pt9 is the strain with the most over-expressed genes, with 8 out of 13 over-expressed genes (Figure 3B). Based on these initial observations, we can suggest that Pt9 may be a promising candidate for improving the synthesis and production of recombinant proteins. Nevertheless, further studies, such as proteomic analysis of this accession under these conditions, would be necessary in the future to ascertain whether the observed up-regulation of over 70% of the genes involved in ribosome biogenesis will result in an enhanced production yield of proteins.

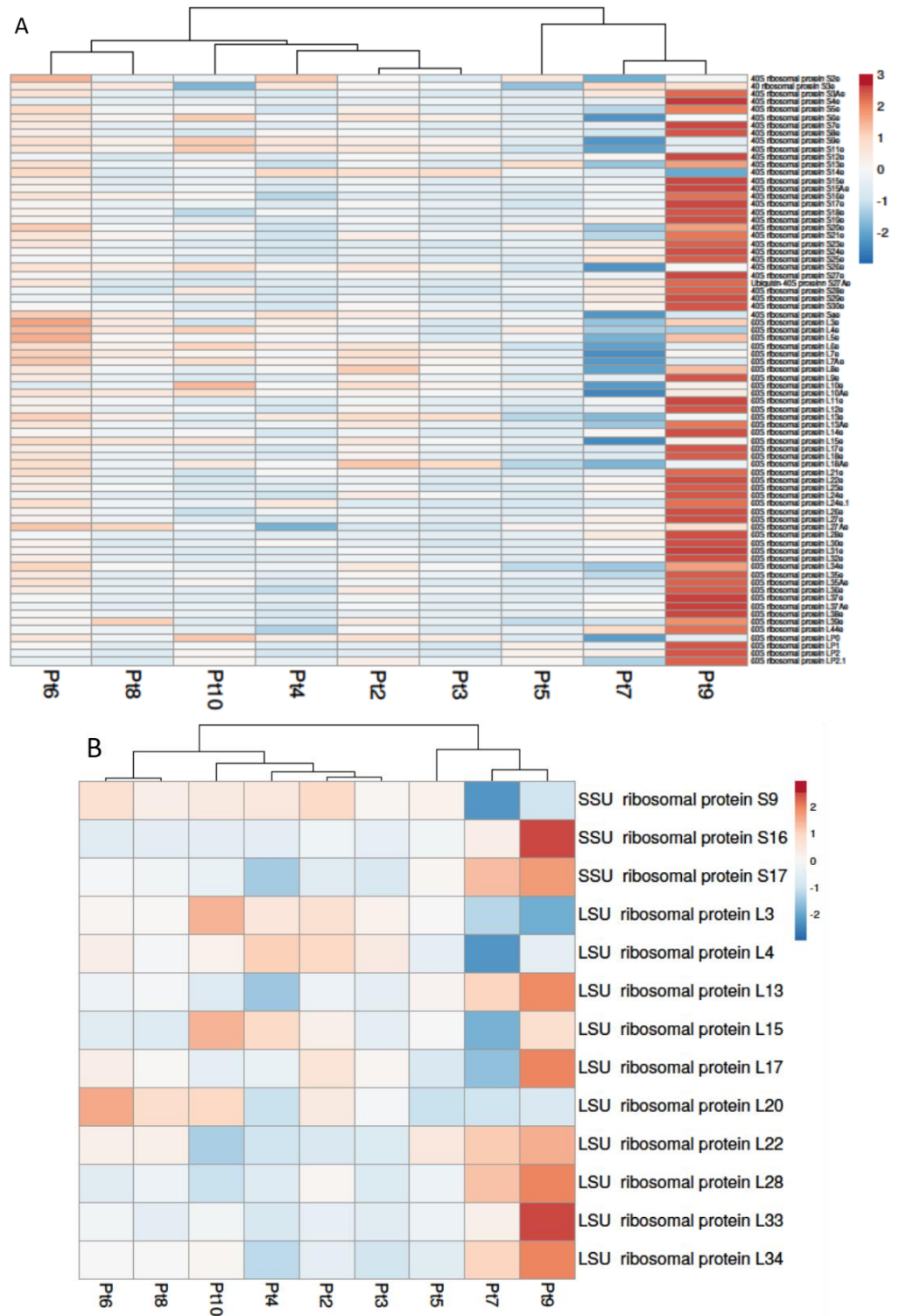


Figure 3. Differential expression of genes involved in ribosome pathway of *P. tricornutum*. (A) Expression of homologous cytosolic ribosomal genes. (B) Expression of homologous mitochondrial/chloroplastic ribosome genes. Under-expressed genes are shown in blue and over-expressed genes in red. Columns are clustered using correlation distance and average linkage. Gene expressions in Pt2 to Pt10 are compared to Pt1.

2.3.2. Protein Export

To achieve higher production yields of biologics, efficient protein expression is undoubtedly a crucial factor but not the only critical point. Indeed, in some systems such as

plants, processing steps like cell harvesting, lysis, and purification are time-consuming and contribute to higher production costs [58]. Thus, the optimal expression system should not only facilitate high protein expression but also involve efficient protein secretion, reducing costly downstream processing steps that represent approximately 84% of the total cost in expression systems like plants [58]. With regards to mammalian cells, microalgae are capable of secreting proteins into the culture medium [43,59]. The secretion of biologics has also been reported in two species of microalgae [17,18,60,61], including *P. tricornutum*, for which the production of two secreted anti-viral mAbs has been described [17,18].

Protein export, the active transport of the proteins outside the cell, occurs through two main mechanisms: the co- and post-translational delivery mechanisms. During the co-translational mechanism, the nascent polypeptides are transported to the export membrane along with the ribosome during protein synthesis. In contrast, in the post-translational mechanism, proteins are delivered to the outside of the cell after the synthesis is completed. Orthologues for molecular actors involved in these two mechanisms have been identified in the genome of *P. tricornutum* (Table S3). Their expression profiles were studied and compared in the course of the present work for the 10 accessions. It was observed that genes involved in the protein export are differentially expressed mostly in Pt4, Pt7, and Pt9. It is somewhat surprising that the Pt4 accession was not hierarchically closer to the cluster formed by the Pt7 and Pt9 accessions. The accession with the greatest number of over-expressed genes is Pt9. In contrast, Pt4 has the greatest number of down-regulated genes (Figure 4).

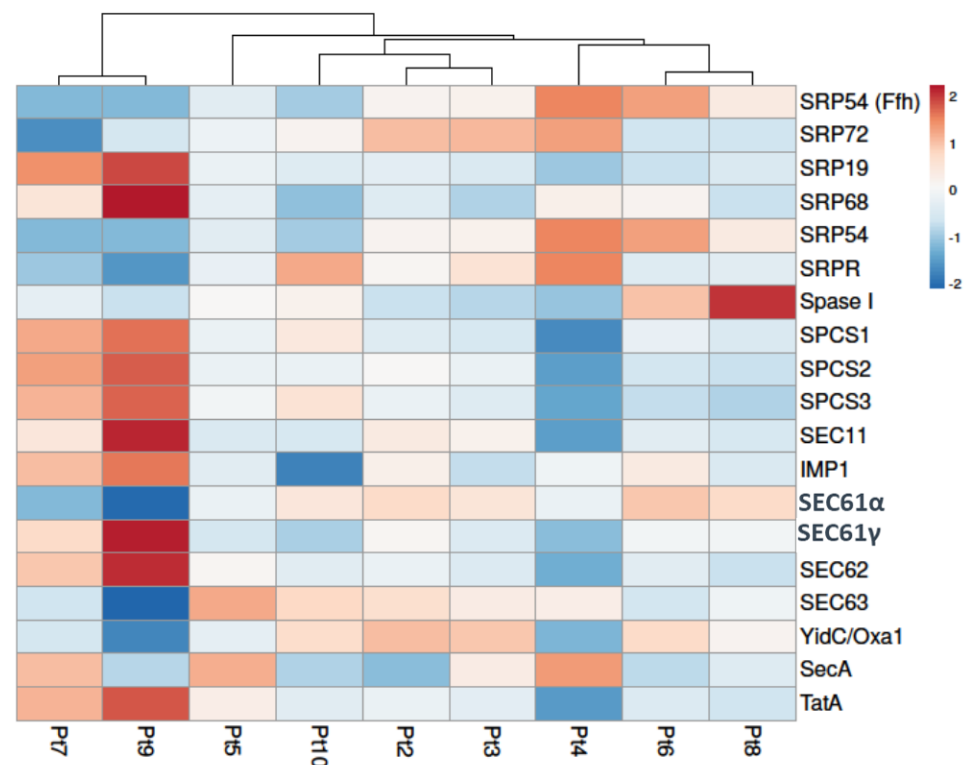


Figure 4. Differential expression of genes involved in protein export and secretion pathways of *P. tricornutum*. Under-expressed genes are shown in blue and over-expressed genes in red. Columns are clustered using correlation distance and average linkage. Genes in Pt2 to Pt10 accessions are compared to Pt1 used as a reference in this study.

In the co-translational mechanism, peptide chains are first synthesized by the ribosome. Signal recognition particles (SRPs) simultaneously recognize and bind to the signal peptide and the SRP receptor (SRPR), forming a complex with the ribosome on the endoplasmic reticulum (ER) membrane. Two SRPs (SRP19 and SRP68) are over-expressed in Pt9, while in Pt7, SRP72 is under-expressed in these two accessions. This complex allows the peptide

to enter the ER lumen. At the same time, signal peptides are removed by signal peptidases. According to the KEGG database, orthologues of both eukaryotic (SPCS1, SPCS2, and SPCS3) and prokaryotic (Spase I) signal peptidases are described in *P. tricornutum* (Table S3). These genes are under-expressed in Pt4, while all of them are over-expressed in Pt9.

Eukaryotic, prokaryotic, and mitochondrial orthologues of important actors of the post-translational mechanism are also described in *P. tricornutum* (Table S3). Three components of the eukaryotic translocation channel are differentially expressed in Pt9, Sec61 α is under-expressed, while Sec61 γ and Sec62 are over-expressed. A homolog of the bacterial chaperone YidC/Oxa1 is under-expressed in Pt4, Pt7, and Pt9. Another protein transport system, the twin-arginine translocation (Tat) pathway, is responsible for the transport of folded proteins in bacteria, archaea, and chloroplasts. TatA, one of the membrane proteins involved in this pathway, is under-expressed in Pt4 and over-expressed in Pt9.

Interestingly, the results for Pt4 diverge significantly from those of our previous meta-analysis study of the same genes, in which Pt4 exhibited 12 genes over-expressed (Boulogne et al., submitted [49]). This discrepancy may be attributed to the difference in culture conditions used to generate the datasets selected for the meta-analysis and the conditions selected for this present work. Indeed, in the meta-analysis, cells were harvested at 6 days of culture following 2 days in the dark [36]. In contrast, the present work examined mRNA expression extracted after 4 days of culture under a light/dark cycle of 16/8 h. The overexpression of genes involved in the protein export pathway in Pt9 may also be linked to a possible shift in morphotype from fusiform to oval of this accession due to the temperature of 19 °C used in this study [26]. Studies have demonstrated a higher and faster expression and secretion of proteins in oval cell cultures in comparison to triradiate and fusiform cell cultures of the same accession [43,62]. The overexpression of genes involved in the co-translational mechanism, in which ribosomes are strongly involved, appears to be consistent with the GO enrichment analysis that indicated an overexpression of genes involved in ribosome biogenesis in Pt9.

This new level of analysis focused on protein export also indicated the potential of the Pt9 strain as a means to enhance protein production.

2.3.3. N-Glycosylation

Most biologics are glycoproteins—meaning proteins bearing glycans, either O- or N-linked—that can significantly impact their biological activity and immunogenicity [1,63]. Therefore, special attention must be paid to the glycosylation processing that is required when choosing an organism for the production of glycosylated biologics. Currently, there is no known information regarding the O-glycosylation processing in *P. tricornutum*. In contrast, the addition of N-glycans onto proteins occurs like the other eukaryotes, in the secretory system as the proteins progress through the ER and the Golgi apparatus to their final compartment [64] (Figure 1). The recombinant mAbs produced in Pt4 have been shown to be predominantly glycosylated with mammalian-like oligomannosides. No complex-type N-glycans or immunogenic glycoepitopes were identified on the produced mAbs [65], although *P. tricornutum* possesses the putative glycoenzymes that are involved in the synthesis of such complex N-glycans [66].

With regards to the overall genes involved in the N-glycosylation pathway of *P. tricornutum* (Table S4), it is observed that there are more differences in accessions Pt7 and Pt9 compared to the other accessions (Figure 5). In fact, Pt7 is the accession with the greatest number of under-expressed genes compared to Pt1 and Pt9 is the one with the greatest number of over-expressed genes. No DEGs related to the N-glycosylation pathway were identified in Pt2 or Pt3. Once again, the hierarchical cluster formed by those accessions is more distant from the one constituted by Pt7 and Pt9.

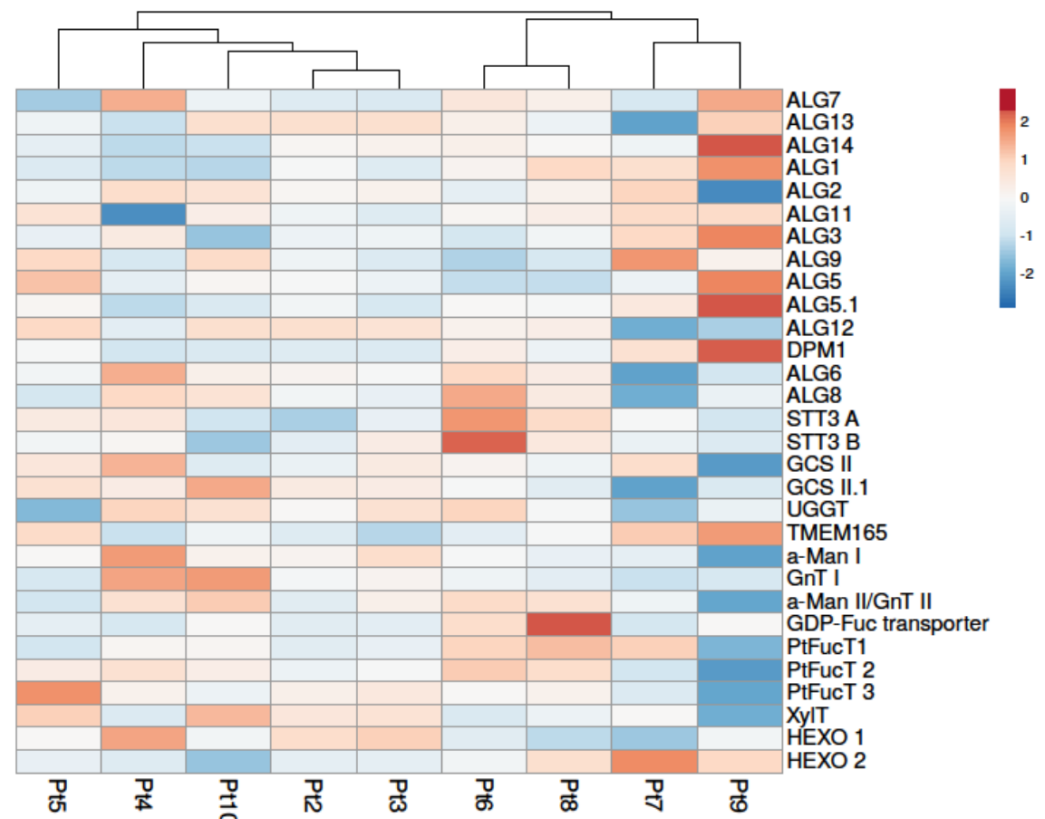


Figure 5. Differential expression of genes involved in the protein *N*-glycosylation biosynthesis pathway of *P. tricornutum*. Under-expressed genes are shown in blue and over-expressed genes in red. Columns are clustered using correlation distance and average linkage. Genes in Pt2 to Pt10 are compared to Pt1.

The synthesis of *N*-glycans starts in the ER with the synthesis of a precursor called the lipid-linked oligosaccharide (LLO). This precursor is synthesized by the action of ALGs (asparagine-linked glycosylation) enzymes that use nucleotide sugars and their specific acceptors to elongate the LLO [67,68]. Among the 13 genes involved in the synthesis of the LLO, some are differentially expressed in different accessions. The most notable differences are observed in Pt7 and in Pt9 which present, respectively, two significant down-regulated genes and six significant up-regulated genes. These results suggest an increased synthesis of the LLO in Pt9.

The synthesized LLO precursor is subsequently transferred *en bloc* to the asparagine at the consensus *N*-glycosylation site (Asn-X-Thr/Ser/Cys) by an enzyme complex known as oligosaccharyltransferase (OST) [69,70]. Two genes encoding putative catalytic subunits (SST3A and SST3B) of the OST have been identified in the *P. tricornutum* genome. In our conditions, the only differential expression observed concerned the gene Phatr3_J55198 encoding for the catalytic subunit SST3B, which was found to be differentially expressed in Pt6 and Pt10. The maturation of the *N*-glycan attached to the glycoprotein is then carried out by two ER-resident enzymes: glucosidase II (GCS II) and UDP-glucose—glycoprotein glycosyltransferase (UGGT). Once correctly folded, the glycoproteins leave the ER to reach the Golgi apparatus, where maturation steps involve a variety of glycoenzymes that result in species-specific complex *N*-glycan structures [71].

Regarding the *N*-glycosylation steps that occur within the Golgi apparatus, our attention was focused on the expression of specific genes encoding for *N*-acetylglucosaminyltransferase I (GnT I), GDP-Fucose (GDP-Fuc) transporters, fucosyltransferases (PtFucT), xylosyltransferases (XylIT), and *N*-acetylhexosaminidases (HEXO) (Figure 5) as they are key glycoenzymes for the synthesis of complex-type *N*-glycans, which are critical quality attributes for the ac-

tivity and the safety of biopharmaceuticals. GnT I is a key enzyme for the initiation of the synthesis of complex-type *N*-glycans [72]. It is over-expressed in Pt4 and Pt10 but under-expressed in Pt5, Pt7, and Pt9. The over-expression of this enzyme could favor the selection of Pt4 and Pt10 accessions when complex-type *N*-glycans are required on the recombinant glycoproteins. The over-expression of GnT I has already been observed in the Pt4 accession in our meta-analysis (Boulogne et al., submitted [49]), although the culture conditions used were different from our current conditions, particularly in terms of pre-harvest culture time and exposure to darkness. However, the characterization of the anti-VHB mAbs produced in Pt4 showed that only oligomannosidic *N*-glycans but not complex-type *N*-glycans were detected [65]. The absence of complex-type *N*-glycans on mAbs could be explained by the nature of this protein, which has its *N*-glycosylation sites embedded, limiting the access of glycosyltransferases to the *N*-glycans, or by the culture conditions used for the production of such mAbs, which do not allow the over-expression of GnT I and the other glycoenzymes involved in the synthesis of complex-type *N*-glycans in *P. tricornutum*.

Core α -(1,3)-fucose and core β (1,2)-xylose glycoepitopes present on plant *N*-glycans have been reported to be immunogenic in humans [73,74]. Moreover, it is well-known that the core fucosylation can affect the biological function of mAbs. A recent study highlighted an increase in α -(1,3)-fucosylation on endogenous glycoproteins of *P. tricornutum* when PtFucT1 is over-expressed [75]. The GDP-Fuc transporter plays a critical role in the fucosylation by allowing the import of GDP-Fuc into the Golgi apparatus where it is the substrate of FucT. Therefore, specific attention was paid to the 5 genes related to core fucosylation and xylosylation in *P. tricornutum*. Among them, only the genes encoding the GDP-Fuc transporter and PtFucT1 were significantly over-expressed and only in the Pt8 accession. Finally, putative *N*-acetylhexosaminidases that can lead to truncated *N*-glycan structures called paucimannoses that lack terminal GlcNAc residue have been identified in *P. tricornutum* [65,66]. Two genes encoding *N*-acetylhexosaminidases are predicted in *P. tricornutum*. The gene Phatr3_J45073 (HEXO1) is under-expressed in Pt7 and Pt8 but the gene Phatr3_J49563 (HEXO2) is over-expressed in Pt7. However, the activity of these enzymes has not yet been demonstrated in *P. tricornutum*.

Based on all this information, the Pt4, Pt9, and Pt10 accessions appear to be slightly more interesting in terms of *N*-glycosylation to produce biologics. Indeed, if the cell is to produce more recombinant proteins, the *N*-glycosylation processing must be able to keep up to ensure that the synthesized proteins are correctly glycosylated. In this case, Pt9 is of interest because it has six genes that are over-expressed in the ER. Pt4 and Pt10 are interesting because of the over-expression of the GnT I that could lead to the presence of more complex type *N*-glycans. However, further glycomic studies need to be performed to compare the protein *N*-glycosylation of the 10 accessions and to decipher whether or not the over-expression of these genes can impact *N*-glycan structures.

2.3.4. Quality Control and Proteasomes

Proteins undergo a quality control process within the ER before being accumulated into various intracellular compartments or released into the extracellular media. This process involves interactions with multiple quality control factors to ensure the release of properly folded and functional proteins. In particular, nascent proteins interact with chaperones such as calreticulin, calnexin, and heat shock proteins (HSPs), including BiP (binding immunoglobulin protein) and GRPs (glucose-related proteins) [76,77]. The analysis of the DEGs involved in the quality control (QC) in *P. tricornutum* revealed that most of the differentially expressed genes tended to be under-expressed (Figure 6). More differentially expressed genes are observed for Pt4, Pt7, and Pt9, while no differential expression of genes is observed in Pt2 and Pt3. Regarding luminal chaperones (Table S5), the gene Phatr3_J16786 that encodes the protein GRP94—a chaperone specialized for protein folding in the ER—is down-regulated in Pt4 and Pt9; however, GRP-94 seems to be essential primarily for metazoans but not for unicellular organisms such as mammalian cell culture [78]. The gene Phatr3_EG02643 encoding BiP is also down-regulated in Pt4, Pt7, and Pt9. Recently,

it has been highlighted that when the basal expression level of BiP in host cell lines is low, it facilitates the generation of recombinant CHO cell lines (rCHO) that give a higher true positive rate during random integration-based pool selection [79]. However, some studies have shown in cell lines producing mAb that the expression level of BiP and mAb production were correlated [79–81]. For *N*-glycosylated proteins, recognition of the oligosaccharide on their surface occurs by the ER luminal enzymes glucosidases II (GCSII), which sequentially cleave terminal glucose residues. Prior to the removal of the third glucose by the GCSII, chaperone proteins assess the correct folding of glycoproteins, and direct properly folded proteins to the secretory pathway. Misfolded proteins, with all three glucose residues removed, are identified by UDP-glucose/glycoprotein glucosyltransferase (UGGT), leading to the addition of glucose residues and subsequent re-engagement with chaperones [77–79]. Abiotic or biotic stresses (Difficult To Express (DTE) protein production or environmental stresses) can provoke a deregulation of the homeostasis in the ER leading to an accumulation of misfolded/unfolded proteins in the ER and to the activation of the unfolding protein response (UPR) or ER-associated degradation (ERAD) pathway [79,80]. In mammals, it has been shown that the production of DTE proteins such as mAbs can lead to the activation of the UPR pathway to restore ER homeostasis. Among all the mechanisms involved in the ERAD pathway, ubiquitination plays a crucial role by affecting the stability of the protein. Ubiquitin-related enzymes have been shown to be important for stress regulation in mammals [82–84], plants [85–87], and also in *P. tricornutum* [88]. Regarding the genes involved in the ERAD pathway, most of the genes are also down-regulated. Recently, 18 UBC genes (PtUBC1 to PtUBC18) were identified in *P. tricornutum* [88]. Among these genes, half of them were differentially expressed in Pt9, with a majority of up-regulated genes. The up-regulation of genes related to ubiquitination could suggest that the ER is under stress and that the cell responds accordingly. As mentioned above, the Pt9 accession seems to adopt an oval morphotype (stress-related) at temperatures between 15 and 19 °C, the latter being the temperature used for the cultures analyzed in this study. This could explain why more ERAD-related genes are up-regulated in this accession.

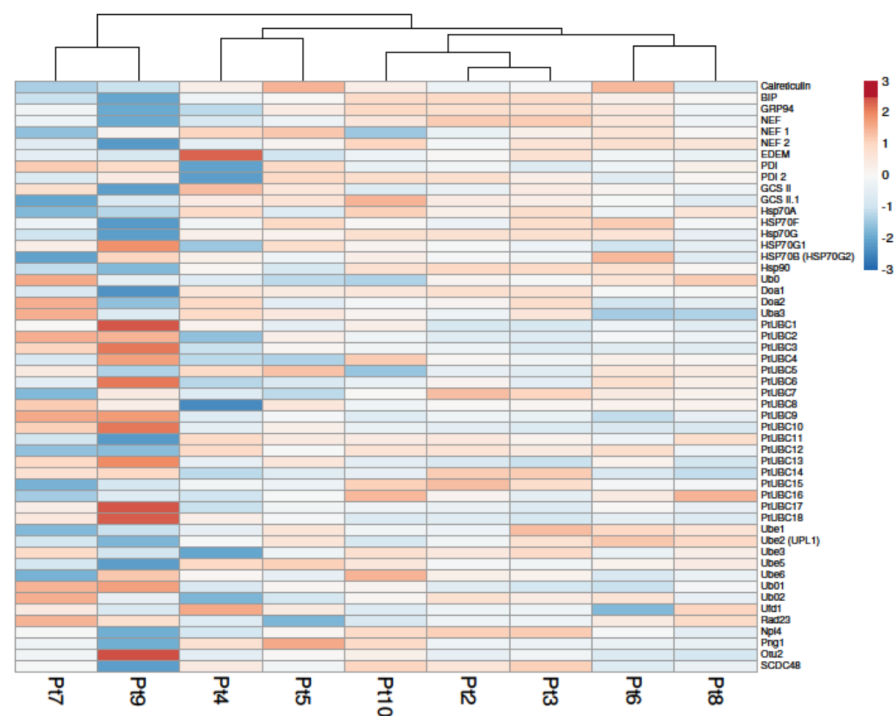


Figure 6. Differential expression of genes involved in protein quality control pathway of *P. tricornutum*. Under-expressed genes are shown in blue and over-expressed genes in red. Columns are clustered using correlation distance and average linkage. Expression of genes in Pt2 to Pt10 was compared to that of Pt1, used as a reference accession.

Taking into consideration the elements provided by the study of the regulation of genes related to QC and proteasomes (Figures 6 and 7), we can suggest that the Pt4, Pt7, and Pt9 accessions could be of interest for the production of biologics. Indeed, the down-regulation of genes related to these pathways could indicate that these accessions would have a better capacity to produce correctly folded proteins and would have less need to recruit the various players involved in the regulation of ER homeostasis. Thus, under more stressful conditions, such as the production of DTE proteins, these accessions would have greater leeway to regulate ER homeostasis. It would be interesting in future studies to compare these different accessions—genetically transformed for the production of biologics—to monitor the expression of QC-related genes and to see if mAb production could be correlated with the expression of chaperone proteins, as it has been previously shown in mammalian cells [79–81].

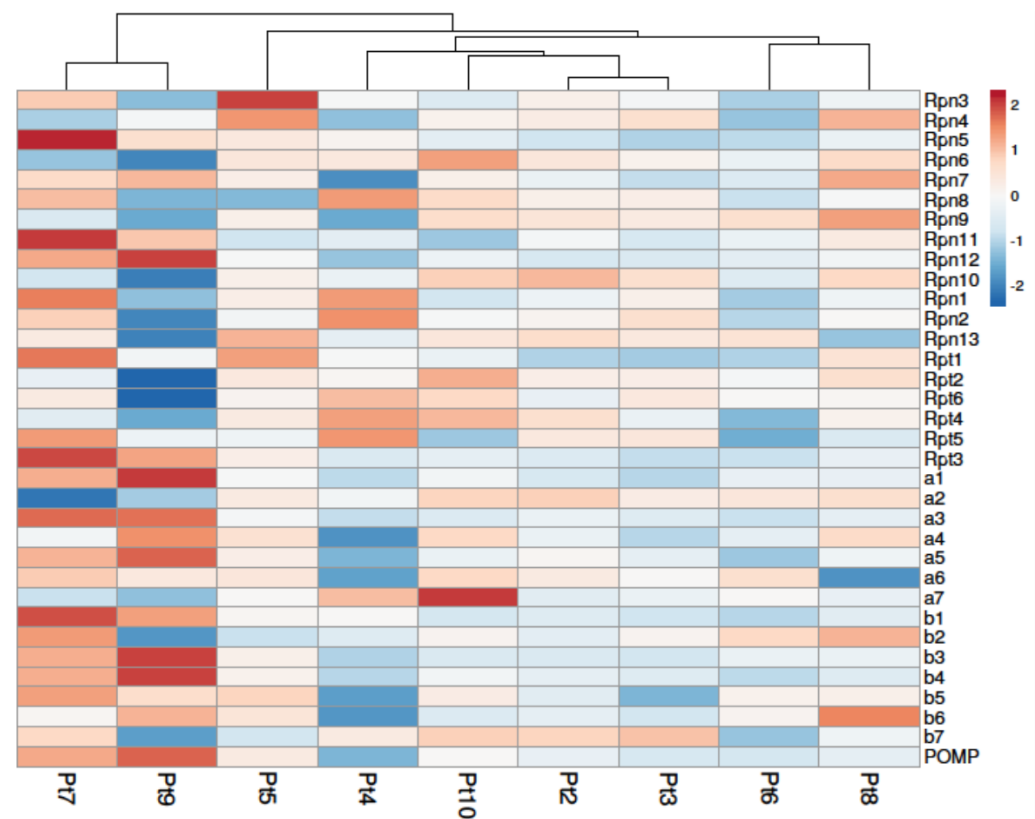


Figure 7. Differential expression of genes involved in the proteasome pathway of *P. tricornutum*. Under-expressed genes are shown in blue and over-expressed genes in red. Columns are clustered using correlation distance and average linkage. Expression of genes in Pt2 to Pt10 was compared to that of Pt1, used as a reference accession.

2.3.5. Proteases

The use of microalgae as a cost-effective production platform on an industrial scale is currently limited by their low production yields, which are lower than those obtained with optimized CHO cells [18,89]. These lower yields, reflecting a low accumulation of secreted proteins in the culture medium, could be partly attributed to the potential proteolytic degradation of mAbs in the culture medium. Indeed, while host cell proteases play an essential role in catalytic and metabolic pathways [90], their presence in the culture media can lead to the degradation of the product of interest [91]. The presence of proteases in the culture media can result from either host cell secretion and/or cell lysis. The negative impact of proteolytic degradation on the yield of recombinant proteins accumulated in culture media has already been demonstrated in CHO cells [91,92] and in plant systems [93]. In addition, the rigorous purification steps of biologics do not always allow for complete removal of

residual protease activity from CHO cell lines [94]. To date, nothing is known regarding the proteolytic degradation of recombinant proteins accumulated in *P. tricornutum*'s culture medium. Two recent studies on the secretome of *P. tricornutum* indicate the presence of several putative proteases in the culture medium of this microalgae [43,95].

One hundred and seventy-three putative proteases were identified in *P. tricornutum*. For all accessions, few genes encoding putative proteases are significantly differentially expressed compared with Pt1 (Figure 8a). The largest differences were observed for Pt4, Pt5, Pt7, and Pt9. Due to their larger number of DEGs, we decided to focus only on these four accessions for the remaining analysis regarding proteases (Figure 9). Firstly, we were interested in the expression of proteases that were specifically over-expressed (Figure 8b) or under-expressed (Figure 8c) in each accession.

a)		Up-regulated	Down-regulated	Non-significant FC
	Pt2	2	2	169
	Pt3	1	1	171
	Pt4	12	16	145
	Pt5	14	17	142
	Pt6	6	17	150
	Pt7	17	21	135
	Pt8	7	11	155
	Pt9	21	43	109
	Pt10	4	11	158

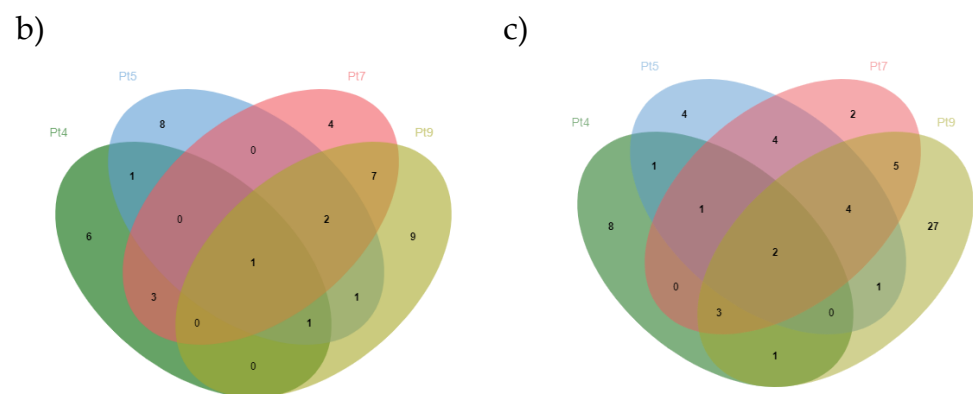


Figure 8. Differential expression of genes encoding proteases identified in *P. tricornutum*. (a) Number of differentially expressed proteases or with a non-significant fold change in Pt2 to Pt10 accessions compared to Pt1. Genes involved in protease activities were obtained from UniprotKB (keyword: KW-064 Protease) and from DiatomicBase. Up-regulated genes have a significant fold change > 2 (i.e., $\log_2FC > 1$), $p < 0.05$; down-regulated genes have a significant fold change < 0.5 (i.e., $-1 < \log_2FC$), p -value < 0.05 . (b) Venn diagram of up-regulated proteases in Pt4 (dark green), Pt5 (blue), Pt7 (red), and Pt9 (light green). (c) Venn diagram of down-regulated proteases in Pt4 (dark green), Pt5 (blue), Pt7 (red), and Pt9 (light green).

Among these specifically over- or under-expressed proteases, we used recent data generated on the *P. tricornutum* secretome [43,95] to investigate whether any of them could be secreted into the culture medium. Looking at the down-regulated genes, some are specifically expressed in one accession, while others like the Phatr3_EG02535 gene are under-expressed in all four accessions. Among up-regulated genes, similar observations can be noticed. The type of secreted proteases whose genes are differentially expressed are predominantly serine-type endopeptidases, followed by metalloproteases or metallopeptidases. Serine-type and metalloproteases have already been shown to be problematic in the context of recombinant protein production in CHO cells [96–99]. Released into the culture medium as HCPs, these proteases can lead to protein degradation [96–99]. Furthermore, their elimination during purification steps is not always guaranteed. In fact, some HCPs such as the serine protease HTRA1 have been shown to co-elute with mAbs using protein A chromatography [100–102].

	Down-regulated genes encoding for secreted proteases		Up-regulated genes encoding for secreted proteases	
	Gene ID	Predicted protease	Gene ID	Predicted protease
Pt4	Phatr3_J50346	Serine-type endopeptidase	Phatr3_J40462	Serine-type endopeptidase
	Phatr3_EG02535	Serine-type endopeptidase	Phatr3_J50210	Serine-type endopeptidase
			Phatr3_J48262	Serine-type endopeptidase
			Phatr3_J45995	Metalloprotease
Pt5	Phatr3_J54319	Serine-type endopeptidase	Phatr3_J48492	Serine-type endopeptidase
	Phatr3_EG02535	Serine-type endopeptidase	Phatr3_J42588	Metalloendopeptidase
			Phatr3_J46548	Serine-type endopeptidase
		Phatr3_J50210	Serine-type endopeptidase	
		Phatr3_J45995	Metalloprotease	
Pt7	Phatr3_J49604	Metalloprotease	Phatr3_J29309	Metallopeptidase
	Phatr3_J50346	Serine-type endopeptidase	Phatr3_J48262	Serine-type endopeptidase
	Phatr3_EG02535	Serine-type endopeptidase		
Pt9	Phatr3_J48493	Serine-type endopeptidase	Phatr3_J45995	Metalloprotease
	Phatr3_J46087	Serine-type endopeptidase		
	Phatr3_J40462	Serine-type endopeptidase		
	Phatr3_J38966	Serine-type endopeptidase		
	Phatr3_J48492	Serine-type endopeptidase		
	Phatr3_J46872	Serine-type endopeptidase		
	Phatr3_J48496	Serine-type endopeptidase		
	Phatr3_EG02320	Metallopeptidase		
	Phatr3_J49604	Metalloprotease		
Phatr3_J50346	Serine-type endopeptidase			
Phatr3_EG02535	Serine-type endopeptidase			

Figure 9. Genes of secreted proteases (according to [43,95]) up- and down-regulated found in Pt4, Pt5, Pt7, and Pt9 (compared to Pt1). Genes encoding proteases specifically found in the corresponding accession are indicated for further information in bold.

In addition, biochemical assays of protease activities were carried out on the 10 *P. tricornutum* accessions. The results show that protease activities are lower inside the cells (Figure 10A) than in the extracellular medium (Figure 10B). Activity profiles between strains are fairly similar when comparing the intracellular protease activities to the ones present in the culture medium. In contrast to the analysis of DEGs, the protease activities of all accessions are stronger than the one observed for Pt1. The highest activities are observed in Pt3 and Pt10 accessions, even though they were among those with fewer over-expressed protease-encoding genes in comparison to Pt1; lower activities are observed in the Pt1 and Pt7 accessions. The discrepancy in results between differential gene expression and biochemical characterization of protease activities can be attributed to the fact that the assay kit is targeting different types of proteases that are serine-type, cysteine-type, and acid-type. Moreover, during our analysis of DEGs, we only analyzed genes identified as encoding for proteases. However, it is important to remember that approximately 25–30% of *P. tricornutum* genes are not well-characterized and are still annotated to encode for “predicted proteins”. Therefore, one can hypothesize that additional protease activities are still to be discovered in *P. tricornutum*.

If we look specifically at the genes encoding secreted proteases, the Pt7 and especially the Pt9 accession could be of interest for the production of biologics because they have more related genes down-regulated than up-regulated. In line with this, a better amount of recombinant proteins can be expected to accumulate in the culture media of such accessions. The biochemical characterization of protease activities tends to confirm the DEGs analysis. Indeed, Pt1, Pt6, Pt7, and Pt9 present lesser protease activities in the culture media in comparison to the other strains.

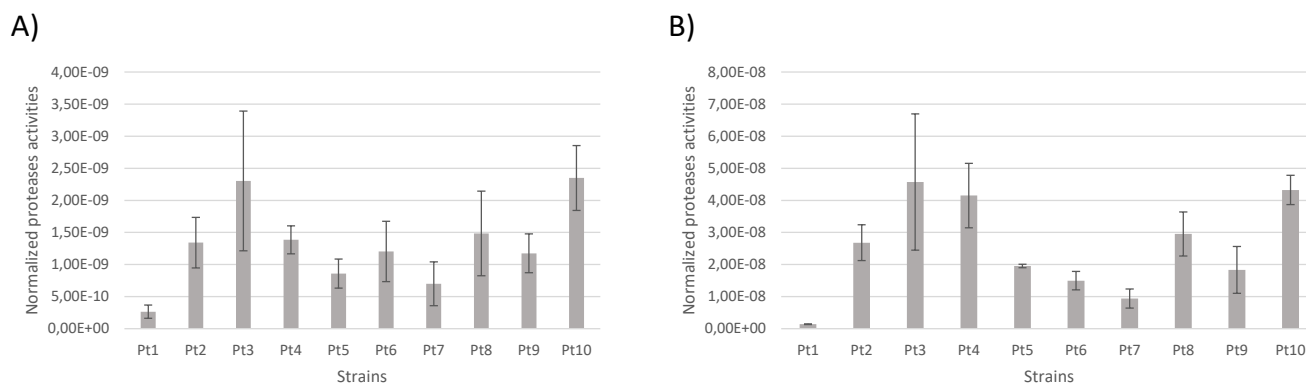


Figure 10. Protease activities normalized per cell of the 10 accessions of *P. tricornutum*. (A) Activities of intracellular serine-type, cysteine-type, and acidic proteases of the 10 accessions of *P. tricornutum*. (B) Activities of secreted serine-type, cysteine-type, and acidic proteases of the 10 accessions of *P. tricornutum*. Assays were performed on three biological replicates.

2.3.6. Genetic Tools

The controlled expression of foreign genetic material is a fundamental requirement to produce recombinant proteins such as biologics. In this context, the identification of genetic elements that ensure a strong and consistent expression of genes encoding the protein of interest is crucial. A lot of work has already been published regarding the design and enhancement of the CHO expression system, including the characterization of elements that improve the yield of recombinant protein production [103–107]. For now, the available information on *P. tricornutum* is limited to elements such as promoters and 5′ untranslated regions (UTR) sequences (Figure 11) or terminators and 3′UTR elements [24,108,109]. With regard to this latter aspect, we postulated that if a gene exhibits a high level of expression, this could potentially indicate that the corresponding promoter is a strong promoter, and therefore, a useful tool for recombinant protein production. This hypothesis has already been tested by Garza and coworkers, who examined the promoter upstream of the gene Phatr3_J49202, which was “consistently observed to be one of the most highly expressed genes in transcriptomic datasets from *P. tricornutum*” [109]. The results indicate that this promoter can drive a higher level of activity in comparison to the promoter currently utilized in the Pt1 accession, such as the *fcpb*, H4, or NR promoters [109]. However, in our conditions, the expression of this gene was found to be under-expressed in Pt6, Pt7, and Pt9. Thus, it seems that the strength of the promoter must be tested in accordance with the selected accession in order to identify the optimal tandem promoter/accession combination. It would be of interest in further studies to ascertain whether the expression of a given protein differs when this promoter is tested in these accessions in comparison to Pt1.

Firstly, a survey was conducted on the promoters used for the production of biologics in *P. tricornutum*. To date, only a limited number of biologics have been produced in this microalga. Two antibodies directed against lethal viral epitopes were produced with each chain under the control of the *nitrate reductase* (NR) promoter [16–18], and the RBD of the SARS-CoV-2 was produced under the control of the HASP1 promoter [19,110]. With regard to these genes, the only differential expression observed concerns the gene Phatr3_J54983, which encodes for nitrate reductase that is over-expressed in Pt4 (Figure 11). This result seems to be consistent with the enrichment analysis, which showed an over-representation of over-expressed genes in nitrogen-related pathways in Pt4.

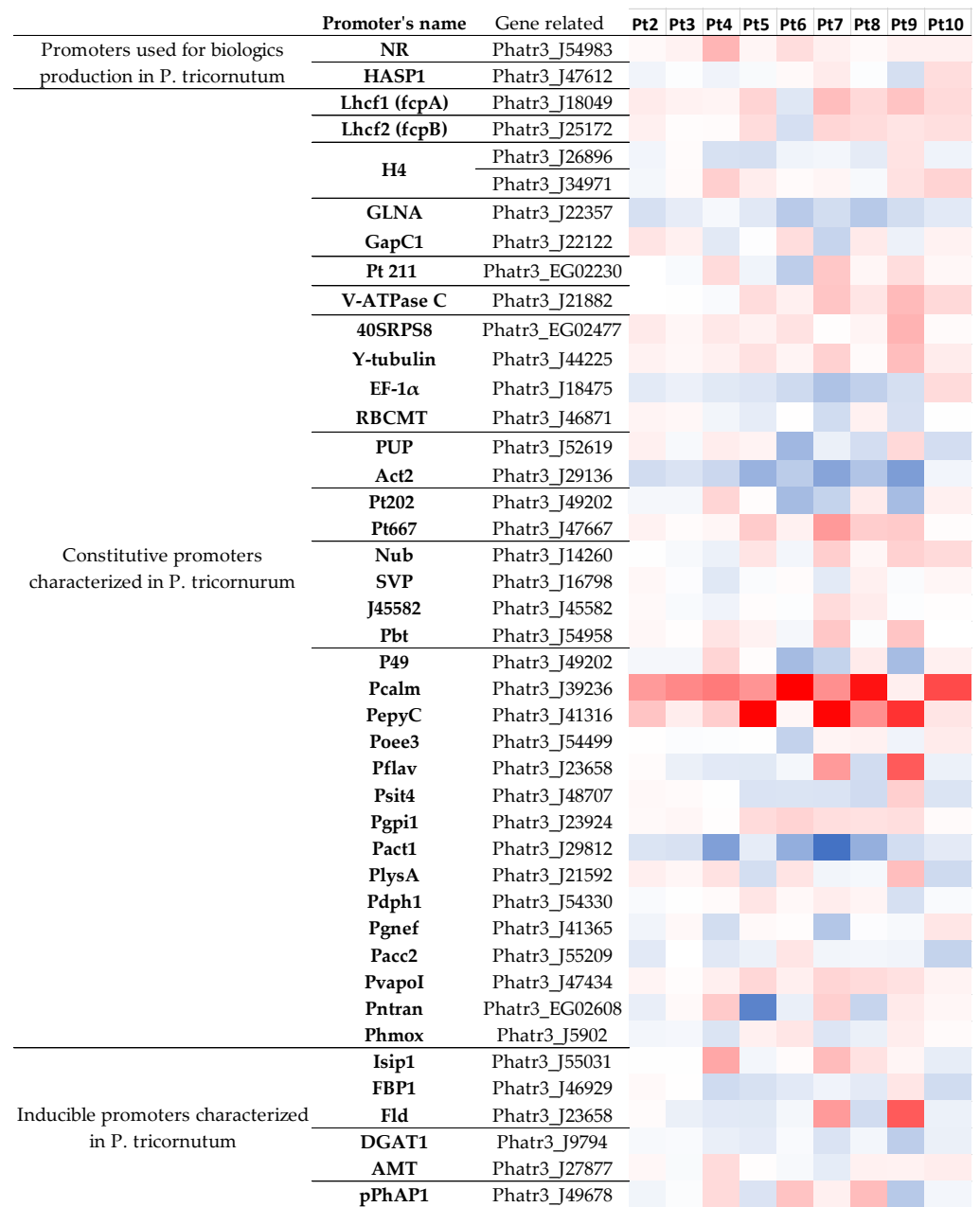


Figure 11. Differential gene expression of genes whose promoters have been characterized in *P. tricornutum*. Promoters that have already been used to produce biologics in *P. tricornutum*, as well as constitutive and inducible promoters that have already been characterized, are presented. Under-expressed genes are shown in blue and over-expressed genes in red. The more intense the color, the stronger the differential expression is. Genes in Pt2 to Pt10 are compared to Pt1.

Then, the gene expression driven by other promoters already characterized in *P. tricornutum*—which were either constitutive [24,108,109,111–118] or inducible [117,119–121]—were analyzed as these promoters represent useful tools for the production of recombinant proteins. Most differences were observed for Pt7 and Pt9 (Figure 11). These two accessions are also the ones with the most up-regulated genes, with seven and nine up-regulated genes for Pt7 and Pt9, respectively, while Pt6 has the most down-regulated genes. Some genes appear to be more universally differentially expressed between different accessions. This is exemplified by Phatr3_J39236 which is over-expressed in all accessions, except Pt9, but also for Phatr3_J29136 which is down-regulated in 6 accessions. Consequently, it can be proposed that the promoter of the gene Phatr3_J39236 (Pcalm) could be employed as

a versatile promoter as its overexpression, is observed in most of the accessions, thus, it represents an interesting promoter to enhance the production of the recombinant protein of interest in *P. tricornutum*. The relevance of inducible promoters for the production of biologics may be questioned, particularly when the biologics are not lethal for the host cell (in such cases, it is necessary to use such promoter to produce the product of interest without interfering with the growth of the cell). This is especially true when the promoters rely on the starvation of a key element, such as phosphate for the alkaline phosphatase promoter (pPhAP1) [121] or the HASP1 promoter [19]. Indeed, certain elements such as phosphorus or nitrogen have been demonstrated to be essential for cell growth and fitness. Furthermore, when deprived of these elements, it has been observed that *P. tricornutum* tends to redirect these metabolic pathways towards lipid production as a stress regulation mechanism, thus limiting protein synthesis [35,41,44,45,122,123]. Nevertheless, despite the demonstrated efficacy of these methods for the production of recombinant proteins, the question remains as to their robustness when applied at an industrial scale.

A comprehensive analysis of the differentially expressed genes (DEGs) revealed the overexpression of several genes in multiple accessions. Among these, the gene Phatr3_J339237 is one of the top over-expressed genes in the accessions Pt2, Pt3, Pt5, Pt6, Pt8, and Pt10. With regards to genetic elements, it appears that no strain is of particular interest in comparison with another and that no promoter among those tested is truly universal. In this case, it would be advisable to adapt the choice of promoter to the strain used for the production of proteins of interest. Future studies may benefit from evaluating the efficacy of this promoter for recombinant protein production in these accessions. Nevertheless, it is also important to consider the combination of promoter and terminator, as some studies have demonstrated that protein expression can vary depending on the terminator used in combination with a given promoter [24,108,109].

3. Materials and Methods

3.1. Cell Culture and Growth Conditions

The ten *P. tricornutum* accessions Pt1 (CCMP 2561), Pt2 (CCMP 2557), Pt3 (CCMP 2558), Pt4 (CCMP 2559), Pt5 (CCMP 630), Pt6 (CCMP 631), Pt7 (CCMP 1327), Pt8 (CCMP 2560), Pt9 (CCMP 633), and Pt10 (CCMP 2928) used in this study were purchased from Bigelow Laboratory (National Center for Marine Algae and Microbiota). Cells were cultured at 19 °C with a 16 h/8 h light–dark cycle in 100% artificial seawater (33.3 g·L⁻¹ of sea salt (Instant Ocean[®], Blacksburg, VA, USA) filtered through 0.45 µm filters and autoclaved. The sterilized culture medium was supplemented by 1 mL·L⁻¹ of adjusted Conway medium (Na₂EDTA·2H₂O: 45 g·L⁻¹; NaNO₃: 1000 g·L⁻¹; H₃BO₃: 33.6 g·L⁻¹; NaH₂PO₄: 20 g·L⁻¹; FeCl₃: 0.768 g·L⁻¹; ZnCl₂: 21 mg·L⁻¹; CoCl₂, 6H₂O: 20 mg·L⁻¹; (NH₄)₆Mo₇O₂₄, 4H₂O: 9 mg·L⁻¹; CuSO₄, 5H₂O: 20 mg·L⁻¹; MnCl₂, 4H₂O: 360 mg·L⁻¹; B1 vitamin: 200 mg·L⁻¹; B12 vitamin: 10 mg·L⁻¹).

3.2. Growth Curves

The ten accessions were adapted over several months using the same culture conditions. Then, the cultures were inoculated at an initial concentration of around 1 × 10⁶ cells per mL to perform the experiments. To generate the growth curves, samples were taken every 24 h for 8 days and cell density was measured using the Countess 3 Cell Counter according to the instructions of the manufacturer (ThermoFisher Scientific[®], Carlsbad, CA, USA). Growth rates were determined with Excel using the slope of the exponential growth phase, corresponding to the following formula:

$$\text{Growth rate } \mu = \frac{\log 2N_2 - \log 2N_1}{t_2 - t_1}$$

3.3. Preparation of RNA Samples

For each accession, 150 mL of cell culture was inoculated at 1 × 10⁶ cells·mL⁻¹ in triplicate. Cells were harvested the fourth day after inoculation by centrifugation at

5000× *g* for 5 min at room temperature. The total cell pellet was resuspended in 1 mL of Nucleozol[®] (Macherey-Nagel, Hoerd, France). Total RNA was extracted using lysing beads (E-matrix lysing tubes, MP Biomedicals[®], Fisher Scientific, Illkirch, France) and grounded for 4 cycles of 30 s at 6.5 m·s⁻¹ in a FastPrep-24[™] homogenizer (MP Biomedicals[®]). Between each run, lysing tubes were placed on ice. Lysis tubes were incubated at room temperature (RT) for 5 min, then centrifuged at 12,000× *g* for 5 min, and the supernatant was collected in RNase-free collection tubes. A total of 400 µL of RNase-free water was added and tubes were mixed by inversion. Tubes were incubated at RT for 15 min, then centrifuged at 16,000× *g* for 15 min at 4 °C. The aqueous phases were collected and pooled in a new RNase-free tube, and 900 µL of ultrapure ethanol was added. Samples were homogenized by pipetting and 650 µL of the samples was loaded onto a NucleoSpin RNA Plus Column (Macherey-Nagel[®]). Tubes were centrifuged at 11,000× *g* for 30 s. These steps were repeated until the entire sample had passed through the column. From this stage onwards, the washing and elution steps were carried out using the Nucleospin RNA Plus kit (Macherey-Nagel[®]) according to the manufacturer's instructions. The DNase treatment was performed following the two-step incubation procedure from "rigorous DNase treatment" of the Turbo DNA-free[™] kit (Invitrogen, Carlsbad, CA, USA). Total RNA concentration was measured by spectrophotometry using the NanoDrop[™] One (ThermoFisher Scientific). The 30 RNA samples were diluted to a maximum of 3 ng·µL⁻¹ and heated at 70 °C for 2 min to minimize RNA secondary structures. A total of 1 µL of each sample was used to assess the quality of the RNAs, using the Agilent 2100 Bioanalyzer system (Agilent, Santa Clara, CA, USA) and RNA 6000 Pico Kit (Agilent).

3.4. Generation of cDNA Libraries and Quality Control of Reads

After quantity and quality checks, RNA sequencing was performed at the MGX facility (Montpellier, France). Libraries were prepared using the Stranded mRNA Prep Ligation kit (Illumina, San Diego, CA, USA) according to the manufacturer's instructions. Briefly, polyadenylated RNAs were selected using oligo-dT magnetic beads from 1000 ng of total RNA. The polyA+ RNAs were fragmented and reverse-transcribed using random hexamers, Super Script II (Thermo Fisher Scientific, Carlsbad, CA, USA), and actinomycin D. Deoxy-TTP was replaced by dUTP during the second-strand synthesis to prevent its amplification by PCR. Double-stranded cDNAs were adenylated at their 3' ends and ligated to Illumina's adapters containing unique dual indexes (UDI). Ligated cDNAs were PCR amplified (11 cycles) and the PCR products were purified using AMPure XP Beads (Beckman Coulter Genomics, Brea, CA, USA). The size distribution of the resulting libraries was monitored using a Fragment Analyzer (Agilent Technologies, Santa Clara, CA, USA) and the libraries were quantified using the KAPA Library quantification kit (Roche, Basel, Switzerland). Library preparation was realized on three biological replicates for each *P. tricornutum*'s accession. The libraries were denatured with NaOH, neutralized with Tris-HCl, and diluted to 150 pM. Clustering and sequencing were performed on a NovaSeq 6000 (Illumina, San Diego, CA, USA) using the paired-end 2 × 150 nt protocol on one lane of an S4 flow cell. Image analysis and base calling were realized in real-time using the NovaSeq Control Software and the Real-Time Analysis 3 software, respectively (Illumina, San Diego, CA, USA). Demultiplexing and FASTQ file generation were carried out using Illumina's bcl2fastq software (v2.20.0.422). The quality control of raw data and demultiplexed reads were assessed, respectively, using Illumina's Sequencing Analysis Viewer (SAV) software and the FastQC (v.0.11.9, Illumina Inc., San Diego, CA, USA) software from the Babraham Institute. Contaminant screening was performed using FastQ Screen (v0.14.0, Babraham Institute). The quality control highlighted that the large majority of the sequence was of very good quality with a Phred score of 33, and no contaminants were detected. The percentage of sequences passing the Illumina PF (Purity Filter) (Illumina Inc.) compared to the raw clusters was greater than 76% for the whole lane. A total of 2.79 × 10⁹ sequences were obtained for the 30 samples.

3.5. Data Analysis

3.5.1. Trimming, Alignment, and Gene Count

FastQ files were uploaded on the Galaxy Europe platform [124] and reads were trimmed with Trimmomatic v0.38.1 [125]. Reads were aligned on the *Phaeodactylum tricorutum* Pt1.8.6 genome (*Phaeodactylum_tricorutum*.ASM15095v2) with HISAT2 v2.2.1 [126]. Gene count was performed with featureCounts v2.0.1 [127], and 72.1 to 77.1% of the reads were assigned on the reference genome.

3.5.2. Differential Gene Expression Analysis

Differential expression of genes (DEG) analysis (Pt2, Pt3, Pt4, Pt5, Pt6, Pt7, Pt8, Pt9, and Pt10 against Pt1) was conducted with DESeq2 v2.11.40.7 [128]. Visualization of DEGs was made by Volcano Plot with VolcanoPlot v0.0.3 (<https://bioconductor.org/packages/devel/bioc/vignettes/EnhancedVolcano/inst/doc/EnhancedVolcano.html>, accessed on February 2024) and heatmaps with Clustvis (<http://biit.cs.ut.ee/clustvis>, accessed on February 2024) [129]. The chosen fold change (FC) was $0.5 < FC < 2$ (i.e., $-1 < \log_2FC < 1$) with a $p < 0.05$ (i.e., $\log_{10}p > 1.3$).

3.5.3. Gene Ontology

Gene ontology was explored using the overrepresentation test performed through ShinyGO 0.77 [130] (pathway database: all available gene set, FDR cutoff 0.05). KEGG annotation was obtained using pathway maps of the KEGG PATHWAY database (<https://www.genome.jp/kegg/pathway.html>, accessed on February 2024). The genes in all other processes were obtained from UniprotKB and from DiatomicBase (<https://www.diatomicsbase.bio.ens.psl.eu/resources>, accessed on February 2024) by BLAST after identification in the pathway maps of the KEGG PATHWAY database. Venn diagrams were created using JVenn [131].

3.6. Protease Activities Assay

For each accession, 5 mL of cell culture was inoculated at 1×10^6 cells.mL⁻¹ in triplicate. On day four after inoculation, cell density was measured as described above and 1 mL of each culture was harvested by centrifugation at $2000 \times g$ for 15 min at room temperature. Culture supernatants were harvested and cell pellets were resuspended in 500 μ L of Tris Buffer 0.1 M pH 7.5. Intracellular proteins were extracted using lysing beads (E-matrix lysing tubes, MP Biomedicals[®], Fisher Scientific, Illkirch, France) and grounded for 4 cycles of 30 s at $6.5 \text{ m}\cdot\text{s}^{-1}$ in a FastPrep-24TM homogenizer (MP Biomedicals[®] Santa Ana, CA, USA). Tubes were centrifugated for 5 min at $10,000 \times g$. The supernatant was collected and the previous step was repeated. Protease activities were measured using the "Protease Assay Kit" (Calbiochem[®] San Diego, CA, USA) according to the manufacturer's instructions, using 200 μ L of culture supernatant and 200 μ L of cell extracts.

4. Conclusions

Altogether, these results demonstrate that omics approaches are powerful tools to help choose the right chassis and strain for the production of biologics in microalgae. Indeed, in this work, applying RNA-Seq analyses on 10 accessions of *P. tricorutum* allowed us to identify the drawbacks and advantages of each of them for the production of secreted recombinant proteins and glycoproteins, and such approach can contribute to decision-making. Our analysis reveals that the Pt9 accession stands out as the most interesting. In fact, the results show its potential in protein synthesis and export, N-glycosylation, as well as in secreted proteases. Other accessions, such as Pt4, were shown to be of interest in terms of N-glycosylation, quality control, and proteasomes. Finally, some accessions, such as Pt2 and Pt3, showed no significant differences in any of the pathways analyzed. Concerning hierarchical clustering, in all the pathways explored, Pt2 and Pt3 were always clustered together and this cluster was always more distant from the cluster formed by Pt7 and Pt9. With regard to the genetic elements that enable the expression of recombinant proteins,

such as promoters, we showed that in this case, no strain seemed to be universal but that the choice of promoter had to be matched more closely to the accession used, or vice versa. It would be interesting to perform further multi-omics studies, adding proteomics or glycomics, to have a more complete picture and thus a better understanding of the physiology of the *P. tricornutum* accessions from a bioproduction point of view.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/md22080353/s1>, Figure S1: Generation time of the ten accessions; Figure S2: DEGs analysis in the different accessions of *P. tricornutum*; Table S1: Fold enrichment pathways analysis with DEGs found in Pt2 to Pt10 compared to Pt1. Table S2: Correspondence between proteins names and genes involved in ribosome pathway of *P. tricornutum*. Table S3: Correspondence between proteins names and genes involved in protein export and secretion pathways of *P. tricornutum*. Table S4: Correspondence between enzymes names and genes involved in the protein *N*-glycosylation biosynthesis pathway of *P. tricornutum*. Table S5: Correspondence between proteins names and genes involved in protein quality control pathway of *P. tricornutum*. Table S6: Correspondence between proteins names and genes involved in ribosome pathway of *P. tricornutum*.

Author Contributions: Conceptualization, I.B., M.B. and C.T.; methodology, I.B., M.B., A.-A.G. and C.T.; investigation, I.B., M.B., A.-A.G. and C.T.; data curation, I.B., A.-A.G. and C.T.; writing—original draft preparation, I.B., M.B. and C.T.; writing—review and editing, I.B., M.B., A.-A.G. and C.T.; supervision, M.B. and I.B.; project administration and funding acquisition, M.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was financially supported by the French government through the ANR agency under the program «Grand défi Biomédicament: améliorer les rendements et maîtriser les coûts de production: Nouveaux Systèmes d’Expression–2020» (PHAEOMABS project—ANR-21-F2II-0005). MGX acknowledges financial support from France Génomique National infrastructure, funded as part of “Investissement d’Avenir” program managed by the Agence Nationale pour la Recherche (contract ANR-10-INBS-09). The Galaxy server that was used for some calculations is partly funded by 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, 031L0106 de.STAIR (de.NBI)).

Data Availability Statement: The data for this study have been deposited in the European Nucleotide Archive (ENA) at EMBL-EBI under accession number PRJEB75330 (<https://www.ebi.ac.uk/ena/browser/view/PRJEB75330> accessed on 23 July 2024).

Acknowledgments: The authors would like to thank Mounir Ichalalen for his technical assistance in generating the growth curves for the 10 accessions, Olivier Perruchon and Carole Burel for their technical support concerning protease activities. The authors also thank Hugues Aroux, University of Rouen Normandie, for drawing the model (drafted by C.T.) presented in Figure 1. The authors also thank Anaïs Louis from MGX for proofreading the materials and methods concerning the generation of RNA libraries.

Conflicts of Interest: M.B. is employed at the University of Rouen Normandie and co-funder and shareholder of ALGA BIOLOGICS. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Lingg, N.; Zhang, P.; Song, Z.; Bardor, M. The Sweet Tooth of Biopharmaceuticals: Importance of Recombinant Protein Glycosylation Analysis. *Biotechnol. J.* **2012**, *7*, 1462–1472. [[CrossRef](#)] [[PubMed](#)]
2. Butler, M. Animal Cell Cultures: Recent Achievements and Perspectives in the Production of Biopharmaceuticals. *Appl. Microbiol. Biotechnol.* **2005**, *68*, 283–291. [[CrossRef](#)] [[PubMed](#)]
3. Banerjee, A.; Ward, V. Production of Recombinant and Therapeutic Proteins in Microalgae. *Curr. Opin. Biotechnol.* **2022**, *78*, 102784. [[CrossRef](#)] [[PubMed](#)]
4. Kornecki, M.; Mestmäcker, F.; Zobel-Roos, S.; Heikaus de Figueiredo, L.; Schlüter, H.; Strube, J. Host Cell Proteins in Biologics Manufacturing: The Good, the Bad, and the Ugly. *Antibodies* **2017**, *6*, 13. [[CrossRef](#)] [[PubMed](#)]

5. Djemal, L.; Fournier, C.; von Hagen, J.; Kolmar, H.; Deparis, V. Review: High Temperature Short Time Treatment of Cell Culture Media and Feed Solutions to Mitigate Adventitious Viral Contamination in the Biopharmaceutical Industry. *Biotechnol. Prog.* **2021**, *37*, e3117. [[CrossRef](#)] [[PubMed](#)]
6. Prade, E.; Zeck, A.; Stiefel, F.; Unsoeld, A.; Mentrup, D.; Arango Gutierrez, E.; Gorr, I.H. Cysteine in Cell Culture Media Induces Acidic IgG1 Species by Disrupting the Disulfide Bond Network. *Biotechnol. Bioeng.* **2021**, *118*, 1091–1104. [[CrossRef](#)] [[PubMed](#)]
7. Arias, C.A.D.; de Oliveira, C.F.M.; Molino, J.V.D.; Ferreira-Camargo, L.S.; Matsudo, M.C.; de Carvalho, J.C.M. Production of Recombinant Biopharmaceuticals in *Chlamydomonas Reinhardtii*. *Int. J. Plant Biol.* **2023**, *14*, 39–52. [[CrossRef](#)]
8. Daniell, H.; Streatfield, S.J.; Wycoff, K. Medical Molecular Farming: Production of Antibodies, Biopharmaceuticals and Edible Vaccines in Plants. *Trends Plant Sci.* **2001**, *6*, 219–226. [[CrossRef](#)]
9. Taunt, H.N.; Stoffels, L.; Purton, S. Green Biologics: The Algal Chloroplast as a Platform for Making Biopharmaceuticals. *Bioengineered* **2018**, *9*, 48–54. [[CrossRef](#)]
10. Moon, K.-B.; Park, J.-S.; Park, Y.-I.; Song, I.-J.; Lee, H.-J.; Cho, H.S.; Jeon, J.-H.; Kim, H.-S. Development of Systems for the Production of Plant-Derived Biopharmaceuticals. *Plants* **2020**, *9*, 30. [[CrossRef](#)]
11. Barolo, L.; Abbriano, R.M.; Commault, A.S.; George, J.; Kahlke, T.; Fabris, M.; Padula, M.P.; Lopez, A.; Ralph, P.J.; Pernice, M. Perspectives for Glyco-Engineering of Recombinant Biopharmaceuticals from Microalgae. *Cells* **2020**, *9*, 633. [[CrossRef](#)]
12. Gimpel, J.A.; Hyun, J.S.; Schoepp, N.G.; Mayfield, S.P. Production of Recombinant Proteins in Microalgae at Pilot Greenhouse Scale. *Biotechnol. Bioeng.* **2015**, *112*, 339–345. [[CrossRef](#)]
13. Mathieu-Rivet, E.; Kiefer-Meyer, M.-C.; Vanier, G.; Ovide, C.; Burel, C.; Lerouge, P.; Bardor, M. Protein N-Glycosylation in Eukaryotic Microalgae and Its Impact on the Production of Nuclear Expressed Biopharmaceuticals. *Front. Plant Sci.* **2014**, *5*, 359. [[CrossRef](#)]
14. Mathieu-Rivet, E.; Lerouge, P.; Bardor, M. *Chlamydomonas Reinhardtii*: Protein Glycosylation and Production of Biopharmaceuticals. In *Chlamydomonas: Biotechnology and Biomedicine*; Hippler, M., Ed.; Microbiology Monographs; Springer International Publishing: Cham, Switzerland, 2017; pp. 45–72; ISBN 978-3-319-66360-9.
15. Dehghani, J.; Movafeghi, A.; Mathieu-Rivet, E.; Mati-Baouche, N.; Calbo, S.; Lerouge, P.; Bardor, M. Microalgae as an Efficient Vehicle for the Production and Targeted Delivery of Therapeutic Glycoproteins against SARS-CoV-2 Variants. *Mar. Drugs* **2022**, *20*, 657. [[CrossRef](#)]
16. Hempel, F.; Lau, J.; Klingl, A.; Maier, U.G. Algae as Protein Factories: Expression of a Human Antibody and the Respective Antigen in the Diatom *Phaeodactylum tricornutum*. *PLoS ONE* **2011**, *6*, e28424. [[CrossRef](#)]
17. Hempel, F.; Maurer, M.; Brockmann, B.; Mayer, C.; Biedenkopf, N.; Kelterbaum, A.; Becker, S.; Maier, U.G. From Hybridomas to a Robust Microalgal-Based Production Platform: Molecular Design of a Diatom Secreting Monoclonal Antibodies Directed against the Marburg Virus Nucleoprotein. *Microb. Cell Factories* **2017**, *16*, 131. [[CrossRef](#)]
18. Hempel, F.; Maier, U.G. An Engineered Diatom Acting like a Plasma Cell Secreting Human IgG Antibodies with High Efficiency. *Microb. Cell Factories* **2012**, *11*, 126. [[CrossRef](#)]
19. Slattery, S.S.; Giguere, D.J.; Stuckless, E.E.; Shrestha, A.; Briere, L.-A.K.; Galbraith, A.; Reaume, S.; Boyko, X.; Say, H.H.; Browne, T.S.; et al. Phosphate-Regulated Expression of the SARS-CoV-2 Receptor-Binding Domain in the Diatom *Phaeodactylum tricornutum* for Pandemic Diagnostics. *Sci. Rep.* **2022**, *12*, 7010. [[CrossRef](#)]
20. Bowler, C.; Allen, A.E.; Badger, J.H.; Grimwood, J.; Jabbari, K.; Kuo, A.; Maheswari, U.; Martens, C.; Maumus, F.; Otilar, R.P.; et al. The *Phaeodactylum* Genome Reveals the Evolutionary History of Diatom Genomes. *Nature* **2008**, *456*, 239–244. [[CrossRef](#)]
21. Serif, M.; Lepetit, B.; Weißert, K.; Kroth, P.G.; Rio Bartulos, C. A Fast and Reliable Strategy to Generate TALEN-Mediated Gene Knockouts in the Diatom *Phaeodactylum tricornutum*. *Algal Res.* **2017**, *23*, 186–195. [[CrossRef](#)]
22. Nymark, M.; Sharma, A.K.; Sparstad, T.; Bones, A.M.; Winge, P. A CRISPR/Cas9 System Adapted for Gene Editing in Marine algae. *Sci. Rep.* **2016**, *6*, 24951. [[CrossRef](#)] [[PubMed](#)]
23. Stukenberg, D.; Zauner, S.; Dell’Aquila, G.; Maier, U.G. Optimizing CRISPR/Cas9 for the Diatom *Phaeodactylum tricornutum*. *Front. Plant Sci.* **2018**, *9*, 740. [[CrossRef](#)]
24. Slattery, S.S.; Diamond, A.; Wang, H.; Therrien, J.A.; Lant, J.T.; Jazey, T.; Lee, K.; Klassen, Z.; Desgagné-Penix, I.; Karas, B.J.; et al. An Expanded Plasmid-Based Genetic Toolbox Enables Cas9 Genome Editing and Stable Maintenance of Synthetic Pathways in *Phaeodactylum tricornutum*. *ACS Synth. Biol.* **2018**, *7*, 328–338. [[CrossRef](#)]
25. Rastogi, A.; Vieira, F.R.J.; Deton-Cabanillas, A.-F.; Veluchamy, A.; Cantrel, C.; Wang, G.; Vanormelingen, P.; Bowler, C.; Piganeau, G.; Hu, H.; et al. A Genomics Approach Reveals the Global Genetic Polymorphism, Structure, and Functional Diversity of Ten Accessions of the Marine Model Diatom *Phaeodactylum tricornutum*. *ISME J.* **2020**, *14*, 347–363. [[CrossRef](#)]
26. De Martino, A.; Meichenin, A.; Shi, J.; Pan, K.; Bowler, C. Genetic and Phenotypic Characterization of *Phaeodactylum tricornutum* (Bacillariophyceae) Accessions1. *J. Phycol.* **2007**, *43*, 992–1009. [[CrossRef](#)]
27. Zhao, X.; Rastogi, A.; Deton Cabanillas, A.F.; Ait Mohamed, O.; Cantrel, C.; Lombard, B.; Murik, O.; Genovesio, A.; Bowler, C.; Bouyer, D.; et al. Genome Wide Natural Variation of H3K27me3 Selectively Marks Genes Predicted to Be Important for Cell Differentiation in *Phaeodactylum tricornutum*. *New Phytol.* **2021**, *229*, 3208–3220. [[CrossRef](#)]
28. De Martino, A.; Bartual, A.; Willis, A.; Meichenin, A.; Villazán, B.; Maheswari, U.; Bowler, C. Physiological and Molecular Evidence That Environmental Changes Elicit Morphological Interconversion in the Model Diatom *Phaeodactylum tricornutum*. *Protist* **2011**, *162*, 462–481. [[CrossRef](#)]

29. Chaumier, T.; Yang, F.; Manirakiza, E.; Ait-Mohamed, O.; Wu, Y.; Chandola, U.; Jesus, B.; Piganeau, G.; Groisillier, A.; Tirichine, L. Genome-Wide Assessment of Genetic Diversity and Transcript Variations in 17 Accessions of the Model Diatom *Phaeodactylum tricorutum*. *ISME Commun.* **2024**, *4*, ycad008. [[CrossRef](#)]
30. Montsant, A.; Jabbari, K.; Maheswari, U.; Bowler, C. Comparative Genomics of the Pennate Diatom *Phaeodactylum tricorutum*. *Plant Physiol.* **2005**, *137*, 500–513. [[CrossRef](#)]
31. Rastogi, A.; Maheswari, U.; Dorrell, R.G.; Vieira, F.R.J.; Maumus, F.; Kustka, A.; McCarthy, J.; Allen, A.E.; Kersey, P.; Bowler, C.; et al. Integrative Analysis of Large Scale Transcriptome Data Draws a Comprehensive Landscape of *Phaeodactylum tricorutum* Genome and Evolutionary Origin of Diatoms. *Sci. Rep.* **2018**, *8*, 4834. [[CrossRef](#)]
32. Rogato, A.; Richard, H.; Sarazin, A.; Voss, B.; Cheminant Navarro, S.; Champeimont, R.; Navarro, L.; Carbone, A.; Hess, W.R.; Falciatore, A. The Diversity of Small Non-Coding RNAs in the Diatom *Phaeodactylum tricorutum*. *BMC Genom.* **2014**, *15*, 698. [[CrossRef](#)] [[PubMed](#)]
33. Huguin, A.; Rastogi, A.; Bowler, C.; Tirichine, L. Genome-Wide Analysis of Allele-Specific Expression of Genes in the Model Diatom *Phaeodactylum tricorutum*. *Sci. Rep.* **2021**, *11*, 2954. [[CrossRef](#)] [[PubMed](#)]
34. Sapriel, G.; Quinet, M.; Heijde, M.; Jourden, L.; Tanty, V.; Luo, G.; Crom, S.L.; Lopez, P.J. Genome-Wide Transcriptome Analyses of Silicon Metabolism in *Phaeodactylum tricorutum* Reveal the Multilevel Regulation of Silicic Acid Transporters. *PLoS ONE* **2009**, *4*, e7458. [[CrossRef](#)] [[PubMed](#)]
35. Cruz de Carvalho, M.H.; Sun, H.-X.; Bowler, C.; Chua, N.-H. Noncoding and Coding Transcriptome Responses of a Marine Diatom to Phosphate Fluctuations. *New Phytol.* **2016**, *210*, 497–510. [[CrossRef](#)] [[PubMed](#)]
36. König, S.; Eisenhut, M.; Bräutigam, A.; Kurz, S.; Weber, A.P.M.; Büchel, C. The Influence of a Cryptochrome on the Gene Expression Profile in the Diatom *Phaeodactylum tricorutum* under Blue Light and in Darkness. *Plant Cell Physiol.* **2017**, *58*, 1914–1923. [[CrossRef](#)]
37. Li, S.; Ismar, S.M.H. Transcriptome, Biochemical and Growth Responses of the Marine Phytoplankter *Phaeodactylum tricorutum* Bohlin (Bacillariophyta) to Copepod Grazer Presence. *CPB* **2018**, *46*, 1091–1111. [[CrossRef](#)] [[PubMed](#)]
38. Ovide, C.; Kiefer-Meyer, M.-C.; Bérard, C.; Vergne, N.; Lecroq, T.; Plasson, C.; Burel, C.; Bernard, S.; Driouich, A.; Lerouge, P.; et al. Comparative in Depth RNA Sequencing of *P. Tricorutum*'s Morphotypes Reveals Specific Features of the Oval Morphotype. *Sci. Rep.* **2018**, *8*, 14340. [[CrossRef](#)]
39. Feng, T.-Y.; Yang, Z.-K.; Zheng, J.-W.; Xie, Y.; Li, D.-W.; Murugan, S.B.; Yang, W.-D.; Liu, J.-S.; Li, H.-Y. Examination of Metabolic Responses to Phosphorus Limitation via Proteomic Analyses in the Marine Diatom *Phaeodactylum tricorutum*. *Sci. Rep.* **2015**, *5*, 10373. [[CrossRef](#)]
40. Bai, X.; Song, H.; Lavoie, M.; Zhu, K.; Su, Y.; Ye, H.; Chen, S.; Fu, Z.; Qian, H. Proteomic Analyses Bring New Insights into the Effect of a Dark Stress on Lipid Biosynthesis in *Phaeodactylum tricorutum*. *Sci. Rep.* **2016**, *6*, 25494. [[CrossRef](#)]
41. Longworth, J.; Wu, D.; Huete-Ortega, M.; Wright, P.C.; Vaidyanathan, S. Proteome Response of *Phaeodactylum tricorutum*, during Lipid Accumulation Induced by Nitrogen Depletion. *Algal Res.* **2016**, *18*, 213–224. [[CrossRef](#)]
42. Poirier, I.; Pallud, M.; Kuhn, L.; Hammann, P.; Demortière, A.; Jamali, A.; Chicher, J.; Caplat, C.; Gallon, R.K.; Bertrand, M. Toxicological Effects of CdSe Nanocrystals on the Marine Diatom *Phaeodactylum tricorutum*: The First Mass Spectrometry-Based Proteomic Approach. *Ecotoxicol. Environ. Saf.* **2018**, *152*, 78–90. [[CrossRef](#)]
43. Chuberre, C.; Chan, P.; Walet-Balieu, M.-L.; Thiébert, F.; Burel, C.; Hardouin, J.; Gügi, B.; Bardor, M. Comparative Proteomic Analysis of Q12 the Diatom *Phaeodactylum tricorutum* Reveals New Insights Into Intra- and Extra-Cellular Protein Contents of Its Oval, Fusiform, and Triradiate Morphotypes. *Front. Plant Sci.* **2022**, *13*, 673113. [[CrossRef](#)]
44. Yang, Z.-K.; Ma, Y.-H.; Zheng, J.-W.; Yang, W.-D.; Liu, J.-S.; Li, H.-Y. Proteomics to Reveal Metabolic Network Shifts towards Lipid Accumulation Following Nitrogen Deprivation in the Diatom *Phaeodactylum tricorutum*. *J. Appl. Phycol.* **2014**, *26*, 73–82. [[CrossRef](#)]
45. Huang, A.; Li, Y.; Duan, J.; Guo, S.; Cai, X.; Zhang, X.; Long, H.; Ren, W.; Xie, Z. Metabolomic, Proteomic and Lactylated Proteomic Analyses Indicate Lactate Plays Important Roles in Maintaining Energy and C:N Homeostasis in *Phaeodactylum tricorutum*. *Biotechnol. Biofuels Bioprod.* **2022**, *15*, 61. [[CrossRef](#)]
46. Duarte, B.; Feijão, E.; Cruz de Carvalho, R.; Duarte, I.A.; Marques, A.P.; Maia, M.; Hertzog, J.; Matos, A.R.; Cabrita, M.T.; Caçador, I.; et al. Untargeted Metabolomics Reveals Antidepressant Effects in a Marine Photosynthetic Organism: The Diatom *Phaeodactylum tricorutum* as a Case Study. *Biology* **2022**, *11*, 1770. [[CrossRef](#)]
47. Stolfa, G.; Smonskey, M.T.; Boniface, R.; Hachmann, A.-B.; Gulde, P.; Joshi, A.D.; Pierce, A.P.; Jacobia, S.J.; Campbell, A. CHO-Omics Review: The Impact of Current and Emerging Technologies on Chinese Hamster Ovary Based Bioproduction. *Biotechnol. J.* **2018**, *13*, 1700227. [[CrossRef](#)]
48. Lewis, A.M.; Abu-Absi, N.R.; Borys, M.C.; Li, Z.J. The Use of 'Omics Technology to Rationally Improve Industrial Mammalian Cell Line Performance. *Biotechnol. Bioeng.* **2016**, *113*, 26–38. [[CrossRef](#)]
49. Boulogne, I.; Toustou, C.; Bardor, M. Meta-Analysis of RNA-Seq Datasets Helps to Better Understand *P. tricorutum* Cellular Biology, a Requirement to Improve the Production of Biologics. Université de Rouen Normandie, Rouen, France, 2024. *submitted*.
50. Bohutskyi, P.; Kligerman, D.C.; Byers, N.; Nasr, L.K.; Cua, C.; Chow, S.; Su, C.; Tang, Y.; Betenbaugh, M.J.; Bouwer, E.J. Effects of Inoculum Size, Light Intensity, and Dose of Anaerobic Digestion Centrate on Growth and Productivity of *Chlorella* and *Scenedesmus* Microalgae and Their Poly-Culture in Primary and Secondary Wastewater. *Algal Res.* **2016**, *19*, 278–290. [[CrossRef](#)]

51. Lu, S.; Wang, J.; Ma, Q.; Yang, J.; Li, X.; Yuan, Y.-J. Phospholipid Metabolism in an Industry Microalga *Chlorella Sorokiniana*: The Impact of Inoculum Sizes. *PLoS ONE* **2013**, *8*, e70827. [[CrossRef](#)]
52. Ma, Q.; Wang, J.; Lu, S.; Lv, Y.; Yuan, Y. Quantitative Proteomic Profiling Reveals Photosynthesis Responsible for Inoculum Size Dependent Variation in *Chlorella Sorokiniana*. *Biotechnol. Bioeng.* **2013**, *110*, 773–784. [[CrossRef](#)]
53. Gibney, E.R.; Nolan, C.M. Epigenetics and Gene Expression. *Heredity* **2010**, *105*, 4–13. [[CrossRef](#)]
54. Veluchamy, A.; Lin, X.; Maumus, F.; Rivarola, M.; Bhavsar, J.; Creasy, T.; O'Brien, K.; Sengamalay, N.A.; Tallon, L.J.; Smith, A.D.; et al. Insights into the Role of DNA Methylation in Diatoms by Genome-Wide Profiling in *Phaeodactylum tricornutum*. *Nat. Commun.* **2013**, *4*, 2091. [[CrossRef](#)]
55. Scarsini, M.; Thurotte, A.; Veidl, B.; Amiard, F.; Niepceron, F.; Badawi, M.; Lagarde, F.; Schoefs, B.; Marchand, J. Metabolite Quantification by Fourier Transform Infrared Spectroscopy in Diatoms: Proof of Concept on *Phaeodactylum tricornutum*. *Front. Plant Sci.* **2021**, *12*, 756421. [[CrossRef](#)]
56. Henras, A.K.; Plisson-Chastang, C.; O'Donohue, M.-F.; Chakraborty, A.; Gleizes, P.-E. An Overview of Pre-Ribosomal RNA Processing in Eukaryotes. *WIREs RNA* **2015**, *6*, 225–242. [[CrossRef](#)]
57. Jüttner, M.; Ferreira-Cerca, S. A Comparative Perspective on Ribosome Biogenesis: Unity and Diversity Across the Tree of Life. In *Ribosome Biogenesis: Methods and Protocols*; Entian, K.-D., Ed.; Springer: New York, NY, USA, 2022; pp. 3–22; ISBN 978-1-07-162501-9.
58. Schillberg, S.; Raven, N.; Spiegel, H.; Rasche, S.; Buntru, M. Critical Analysis of the Commercial Potential of Plants for the Production of Recombinant Proteins. *Front. Plant Sci.* **2019**, *10*, 720. [[CrossRef](#)]
59. Ves-Urai, P.; Krobthong, S.; Thongsuk, K.; Roytrakul, S.; Yokthongwattana, C. Comparative Secretome Analysis between Salinity-Tolerant and Control *Chlamydomonas Reinhardtii* Strains. *Planta* **2021**, *253*, 68. [[CrossRef](#)]
60. Eichler-Stahlberg, A.; Weisheit, W.; Ruecker, O.; Heitzer, M. Strategies to Facilitate Transgene Expression in *Chlamydomonas Reinhardtii*. *Planta* **2009**, *229*, 873–883. [[CrossRef](#)]
61. Kiefer, A.M.; Niemeyer, J.; Probst, A.; Erkel, G.; Schroda, M. Production and Secretion of Functional SARS-CoV-2 Spike Protein in *Chlamydomonas Reinhardtii*. *Front. Plant Sci.* **2022**, *13*, 988870. [[CrossRef](#)]
62. Galas, L.; Burel, C.; Schapman, D.; Ropitiaux, M.; Bernard, S.; Bénard, M.; Bardor, M. Comparative Structural and Functional Analyses of the Fusiform, Oval, and Triradiate Morphotypes of *Phaeodactylum tricornutum* Pt3 Strain. *Front. Plant Sci.* **2021**, *12*, 638181. [[CrossRef](#)]
63. van Beers, M.M.C.; Bardor, M. Minimizing Immunogenicity of Biopharmaceuticals by Controlling Critical Quality Attributes of Proteins. *Biotechnol. J.* **2012**, *7*, 1473–1484. [[CrossRef](#)]
64. Colley, K.J.; Varki, A.; Kinoshita, T. Cellular Organization of Glycosylation. In *Essentials of Glycobiology*; Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Darvill, A.G., Kinoshita, T., Packer, N.H., Prestegard, J.H., et al., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2015.
65. Vanier, G.; Hempel, F.; Chan, P.; Rodamer, M.; Vaudry, D.; Maier, U.G.; Lerouge, P.; Bardor, M. Biochemical Characterization of Human Anti-Hepatitis B Monoclonal Antibody Produced in the Microalgae *Phaeodactylum tricornutum*. *PLoS ONE* **2015**, *10*, e0139282. [[CrossRef](#)] [[PubMed](#)]
66. Baïet, B.; Burel, C.; Saint-Jean, B.; Louvet, R.; Menu-Bouaouiche, L.; Kiefer-Meyer, M.-C.; Mathieu-Rivet, E.; Lefebvre, T.; Castel, H.; Carlier, A.; et al. N-Glycans of *Phaeodactylum tricornutum* Diatom and Functional Characterization of Its N-Acetylglucosaminyltransferase I Enzyme. *J. Biol. Chem.* **2011**, *286*, 6152–6164. [[CrossRef](#)] [[PubMed](#)]
67. Aebi, M. N-Linked Protein Glycosylation in the ER. *Biochim. Biophys. Acta* **2013**, *1833*, 2430–2437. [[CrossRef](#)]
68. Lucas, P.-L.; Dumontier, R.; Loutelier-Bourhis, C.; Mareck, A.; Afonso, C.; Lerouge, P.; Mati-Baouche, N.; Bardor, M. User-Friendly Extraction and Multistage Tandem Mass Spectrometry Based Analysis of Lipid-Linked Oligosaccharides in Microalgae. *Plant Methods* **2018**, *14*, 107. [[CrossRef](#)] [[PubMed](#)]
69. Silberstein, S.; Gilmore, R. Biochemistry, Molecular Biology, and Genetics of the Oligosaccharyltransferase. *FASEB J.* **1996**, *10*, 849–858. [[CrossRef](#)]
70. Breitling, J.; Aebi, M. N-Linked Protein Glycosylation in the Endoplasmic Reticulum. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, a013359. [[CrossRef](#)]
71. Toustou, C.; Walet-Balieu, M.-L.; Kiefer-Meyer, M.-C.; Houdou, M.; Lerouge, P.; Foulquier, F.; Bardor, M. Towards Understanding the Extensive Diversity of Protein N-Glycan Structures in Eukaryotes. *Biol. Rev.* **2022**, *97*, 732–748. [[CrossRef](#)]
72. Stanley, P.; Taniguchi, N.; Aebi, M. N-Glycans. In *Essentials of Glycobiology*; Varki, A., Cummings, R.D., Esko, J.D., Stanley, P., Hart, G.W., Aebi, M., Darvill, A.G., Kinoshita, T., Packer, N.H., Prestegard, J.H., et al., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2015.
73. Bardor, M.; Faveeuw, C.; Fitchette, A.-C.; Gilbert, D.; Galas, L.; Trottein, F.; Faye, L.; Lerouge, P. Immunoreactivity in Mammals of Two Typical Plant Glyco-Epitopes, Core Alpha(1,3)-Fucose and Core Xylose. *Glycobiology* **2003**, *13*, 427–434. [[CrossRef](#)] [[PubMed](#)]
74. Jin, C.; Altmann, F.; Strasser, R.; Mach, L.; Schähs, M.; Kunert, R.; Rademacher, T.; Glössl, J.; Steinkellner, H. A Plant-Derived Human Monoclonal Antibody Induces an Anti-Carbohydrate Immune Response in Rabbits. *Glycobiology* **2008**, *18*, 235–241. [[CrossRef](#)]
75. Zhang, P.; Burel, C.; Plasson, C.; Kiefer-Meyer, M.-C.; Ovide, C.; Gügi, B.; Wan, C.; Teo, G.; Mak, A.; Song, Z.; et al. Characterization of a GDP-Fucose Transporter and a Fucosyltransferase Involved in the Fucosylation of Glycoproteins in the Diatom *Phaeodactylum tricornutum*. *Front. Plant Sci.* **2019**, *10*, 610. [[CrossRef](#)]

76. Hebert, D.N.; Molinari, M. In and out of the ER: Protein Folding, Quality Control, Degradation, and Related Human Diseases. *Physiol. Rev.* **2007**, *87*, 1377–1408. [[CrossRef](#)] [[PubMed](#)]
77. Lamriben, L.; Graham, J.B.; Adams, B.M.; Hebert, D.N. N-Glycan-Based ER Molecular Chaperone and Protein Quality Control System: The Calnexin Binding Cycle. *Traffic* **2016**, *17*, 308–326. [[CrossRef](#)] [[PubMed](#)]
78. Marzec, M.; Eletto, D.; Argon, Y. GRP94: An HSP90-like Protein Specialized for Protein Folding and Quality Control in the Endoplasmic Reticulum. *Biochim. Biophys. Acta* **2012**, *1823*, 774–787. [[CrossRef](#)] [[PubMed](#)]
79. Kyeong, M.; Lee, J.S. Endogenous BiP Reporter System for Simultaneous Identification of ER Stress and Antibody Production in Chinese Hamster Ovary Cells. *Metab. Eng.* **2022**, *72*, 35–45. [[CrossRef](#)] [[PubMed](#)]
80. Du, Z.; Treiber, D.; McCoy, R.E.; Miller, A.K.; Han, M.; He, F.; Domnitz, S.; Heath, C.; Reddy, P. Non-Invasive UPR Monitoring System and Its Applications in CHO Production Cultures. *Biotechnol. Bioeng.* **2013**, *110*, 2184–2194. [[CrossRef](#)] [[PubMed](#)]
81. Kober, L.; Zehe, C.; Bode, J. Development of a Novel ER Stress Based Selection System for the Isolation of Highly Productive Clones. *Biotechnol. Bioeng.* **2012**, *109*, 2599–2611. [[CrossRef](#)]
82. Braakman, I.; Hebert, D.N. Protein Folding in the Endoplasmic Reticulum. *Cold Spring Harb. Perspect. Biol.* **2013**, *5*, a013201. [[CrossRef](#)]
83. Fornace, A.J., Jr.; Alamo, I., Jr.; Hollander, M.C.; Lamoreaux, E. Ubiquitin mRNA Is a Major Stress-Induced Transcript in Mammalian Cells. *Nucleic Acids Res.* **1989**, *17*, 1215–1230. [[CrossRef](#)]
84. Sheng, X.; Xia, Z.; Yang, H.; Hu, R. The Ubiquitin Codes in Cellular Stress Responses. *Protein Cell* **2024**, *15*, 157. [[CrossRef](#)]
85. Li, G.-Q.; Zang, X.-N.; Zhang, X.-C.; Lu, N.; Ding, Y.; Gong, L.; Chen, W.-C. Cloning of Ubiquitin-Activating Enzyme and Ubiquitin-Conjugating Enzyme Genes from *Gracilaria lemaneiformis* and Their Activity under Heat Shock. *Gene* **2014**, *538*, 155–163. [[CrossRef](#)]
86. Li, W.; Schmidt, W. A Lysine-63-Linked Ubiquitin Chain-Forming Conjugase, UBC13, Promotes the Developmental Responses to Iron Deficiency in Arabidopsis Roots. *Plant J.* **2010**, *62*, 330–343. [[CrossRef](#)] [[PubMed](#)]
87. Jue, D.; Sang, X.; Lu, S.; Dong, C.; Zhao, Q.; Chen, H.; Jia, L. Genome-Wide Identification, Phylogenetic and Expression Analyses of the Ubiquitin-Conjugating Enzyme Gene Family in Maize. *PLoS ONE* **2015**, *10*, e0143488. [[CrossRef](#)] [[PubMed](#)]
88. Ma, W.; Du, H.; Kazmi, S.S.U.H.; Chen, J.; Chen, W.; Fan, Y.; Liu, Z.; Luo, H.; Fang, H.; Wang, Z.; et al. UBC Gene Family and Their Potential Functions on the Cellular Homeostasis under the Elevated pCO_2 Stress in the Diatom *Phaeodactylum tricornutum*. *Ecol. Indic.* **2023**, *148*, 110106. [[CrossRef](#)]
89. Kunert, R.; Reinhart, D. Advances in Recombinant Antibody Manufacturing. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 3451–3461. [[CrossRef](#)]
90. Bond, J.S. Proteases: History, Discovery, and Roles in Health and Disease. *J. Biol. Chem.* **2019**, *294*, 1643–1651. [[CrossRef](#)]
91. Gilgunn, S.; Bones, J. Challenges to Industrial MAb Bioprocessing—Removal of Host Cell Proteins in CHO Cell Bioprocesses. *Curr. Opin. Chem. Eng.* **2018**, *22*, 98–106. [[CrossRef](#)]
92. Yang, B.; Li, W.; Zhao, H.; Wang, A.; Lei, Y.; Xie, Q.; Xiong, S. Discovery and Characterization of CHO Host Cell Protease-Induced Fragmentation of a Recombinant Monoclonal Antibody during Production Process Development. *J. Chromatogr. B* **2019**, *1112*, 1–10. [[CrossRef](#)]
93. Jutras, P.V.; Marusic, C.; Lonoce, C.; Deflers, C.; Goulet, M.-C.; Benvenuto, E.; Michaud, D.; Donini, M. An Accessory Protease Inhibitor to Increase the Yield and Quality of a Tumour-Targeting MAb in Nicotiana Benthamiana Leaves. *PLoS ONE* **2016**, *11*, e0167086. [[CrossRef](#)]
94. Gao, S.X.; Zhang, Y.; Stansberry-Perkins, K.; Buko, A.; Bai, S.; Nguyen, V.; Brader, M.L. Fragmentation of a Highly Purified Monoclonal Antibody Attributed to Residual CHO Cell Protease Activity. *Biotechnol. Bioeng.* **2011**, *108*, 977–982. [[CrossRef](#)]
95. Dorrell, R.G.; Villain, A.; Perez-Lamarque, B.; Audren de Kerdel, G.; McCallum, G.; Watson, A.K.; Ait-Mohamed, O.; Alberti, A.; Corre, E.; Frischkorn, K.R.; et al. Phylogenomic Fingerprinting of Tempo and Functions of Horizontal Gene Transfer within Ochrophytes. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2009974118. [[CrossRef](#)]
96. Clincke, M.-F.; Guedon, E.; Yen, F.T.; Ogier, V.; Goergen, J.-L. Characterization of Metalloprotease and Serine Protease Activities in Batch CHO Cell Cultures: Control of Human Recombinant IFN- γ Proteolysis by Addition of Iron Citrate. *BMC Proc.* **2011**, *5*, P115. [[CrossRef](#)] [[PubMed](#)]
97. Li, S.W.; Yu, B.; Byrne, G.; Wright, M.; O'Rourke, S.; Mesa, K.; Berman, P.W. Identification and CRISPR/Cas9 Inactivation of the C1s Protease Responsible for Proteolysis of Recombinant Proteins Produced in CHO Cells. *Biotechnol. Bioeng.* **2019**, *116*, 2130–2145. [[CrossRef](#)] [[PubMed](#)]
98. Sandberg, H.; Lütkemeyer, D.; Kuprin, S.; Wrangel, M.; Almstedt, A.; Persson, P.; Ek, V.; Mikaelsson, M. Mapping and Partial Characterization of Proteases Expressed by a CHO Production Cell Line. *Biotechnol. Bioeng.* **2006**, *95*, 961–971. [[CrossRef](#)] [[PubMed](#)]
99. Laux, H.; Romand, S.; Nuciforo, S.; Farady, C.J.; Tapparel, J.; Buechmann-Moeller, S.; Sommer, B.; Oakeley, E.J.; Bodendorf, U. Degradation of Recombinant Proteins by Chinese Hamster Ovary Host Cell Proteases Is Prevented by Matriptase-1 Knockout. *Biotechnol. Bioeng.* **2018**, *115*, 2530–2540. [[CrossRef](#)] [[PubMed](#)]
100. Aboulaich, N.; Chung, W.K.; Thompson, J.H.; Larkin, C.; Robbins, D.; Zhu, M. A Novel Approach to Monitor Clearance of Host Cell Proteins Associated with Monoclonal Antibodies. *Biotechnol. Prog.* **2014**, *30*, 1114–1124. [[CrossRef](#)] [[PubMed](#)]
101. Zhang, Q.; Goetze, A.M.; Cui, H.; Wylie, J.; Trimble, S.; Hewig, A.; Flynn, G.C. Comprehensive Tracking of Host Cell Proteins during Monoclonal Antibody Purifications Using Mass Spectrometry. *mAbs* **2014**, *6*, 659–670. [[CrossRef](#)] [[PubMed](#)]

102. Zhang, Q.; Goetze, A.M.; Cui, H.; Wylie, J.; Tillotson, B.; Hewig, A.; Hall, M.P.; Flynn, G.C. Characterization of the Co-Elution of Host Cell Proteins with Monoclonal Antibodies during Protein A Purification. *Biotechnol. Prog.* **2016**, *32*, 708–717. [[CrossRef](#)]
103. Ho, S.C.L.; Bardor, M.; Feng, H.; Mariati, Tong, Y.W.; Song, Z.; Yap, M.G.S.; Yang, Y. IRES-Mediated Tricistronic Vectors for Enhancing Generation of High Monoclonal Antibody Expressing CHO Cell Lines. *J. Biotechnol.* **2012**, *157*, 130–139. [[CrossRef](#)] [[PubMed](#)]
104. Ho, S.C.L.; Bardor, M.; Li, B.; Lee, J.J.; Song, Z.; Tong, Y.W.; Goh, L.-T.; Yang, Y. Comparison of Internal Ribosome Entry Site (IRES) and Furin-2A (F2A) for Monoclonal Antibody Expression Level and Quality in CHO Cells. *PLoS ONE* **2013**, *8*, e63247. [[CrossRef](#)]
105. Bayat, H.; Hoseinzadeh, S.; Pourmaleki, E.; Ahani, R.; Rahimpour, A. Evaluation of Different Vector Design Strategies for the Expression of Recombinant Monoclonal Antibody in CHO Cells. *Prep. Biochem. Biotechnol.* **2018**, *48*, 160–164. [[CrossRef](#)]
106. Chng, J.; Wang, T.; Nian, R.; Lau, A.; Hoi, K.M.; Ho, S.C.L.; Gagnon, P.; Bi, X.; Yang, Y. Cleavage Efficient 2A Peptides for High Level Monoclonal Antibody Expression in CHO Cells. *mAbs* **2015**, *7*, 403–412. [[CrossRef](#)] [[PubMed](#)]
107. Cheng, J.K.; Morse, N.J.; Wagner, J.M.; Tucker, S.K.; Alper, H.S. Design and Evaluation of Synthetic Terminators for Regulating Mammalian Cell Transgene Expression. *ACS Synth. Biol.* **2019**, *8*, 1263–1275. [[CrossRef](#)] [[PubMed](#)]
108. Windhagauer, M.; Abbriano, R.M.; Ashworth, J.; Barolo, L.; Jaramillo-Madrid, A.C.; Pernice, M.; Doblin, M.A. Characterisation of Novel Regulatory Sequences Compatible with Modular Assembly in the Diatom *Phaeodactylum tricornutum*. *Algal Res.* **2021**, *53*, 102159. [[CrossRef](#)]
109. Garza, E.A.; Bielinski, V.A.; Espinoza, J.L.; Orlandi, K.; Alfaro, J.R.; Bolt, T.M.; Beeri, K.; Weyman, P.D.; Dupont, C.L. Validating a Promoter Library for Application in Plasmid-Based Diatom Genetic Engineering. *ACS Synth. Biol.* **2023**, *12*, 3215–3228. [[CrossRef](#)]
110. Erdene-Ochir, E.; Shin, B.-K.; Kwon, B.; Jung, C.; Pan, C.-H. Identification and Characterisation of the Novel Endogenous Promoter HASP1 and Its Signal Peptide from *Phaeodactylum tricornutum*. *Sci. Rep.* **2019**, *9*, 9941. [[CrossRef](#)]
111. Apt, K.E.; Kroth-Pancic, P.G.; Grossman, A.R. Stable Nuclear Transformation of the Diatom *Phaeodactylum tricornutum*. *Mol. Gen. Genet.* **1996**, *252*, 572–579.
112. Falciatore, A.; Casotti, R.; Leblanc, C.; Abrescia, C.; Bowler, C. Transformation of Nonselectable Reporter Genes in Marine Diatoms. *Mar. Biotechnol.* **1999**, *1*, 239–251. [[CrossRef](#)]
113. Siaux, M.; Heijde, M.; Mangogna, M.; Montsant, A.; Coesel, S.; Allen, A.; Manfredonia, A.; Falciatore, A.; Bowler, C. Molecular Toolbox for Studying Diatom Biology in *Phaeodactylum tricornutum*. *Gene* **2007**, *406*, 23–35. [[CrossRef](#)]
114. Erdene-Ochir, E.; Shin, B.-K.; Huda, M.N.; Kim, D.H.; Lee, E.H.; Song, D.-G.; Kim, Y.-M.; Kim, S.M.; Pan, C.-H. Cloning of a Novel Endogenous Promoter for Foreign Gene Expression in *Phaeodactylum tricornutum*. *Appl. Biol. Chem.* **2016**, *59*, 861–867. [[CrossRef](#)]
115. Zou, L.-G.; Chen, J.-W.; Zheng, D.-L.; Balamurugan, S.; Li, D.-W.; Yang, W.-D.; Liu, J.-S.; Li, H.-Y. High-Efficiency Promoter-Driven Coordinated Regulation of Multiple Metabolic Nodes Elevates Lipid Accumulation in the Model Microalga *Phaeodactylum tricornutum*. *Microb. Cell Factories* **2018**, *17*, 54. [[CrossRef](#)]
116. Watanabe, Y.; Kadono, T.; Kira, N.; Suzuki, K.; Iwata, O.; Ohnishi, K.; Yamaguchi, H.; Adachi, M. Development of Endogenous Promoters That Drive High-Level Expression of Introduced Genes in the Model Diatom *Phaeodactylum tricornutum*. *Mar. Genom.* **2018**, *42*, 41–48. [[CrossRef](#)] [[PubMed](#)]
117. Adler-Agnon (Shemesh), Z.; Leu, S.; Zarka, A.; Boussiba, S.; Khozin-Goldberg, I. Novel Promoters for Constitutive and Inducible Expression of Transgenes in the Diatom *Phaeodactylum tricornutum* under Varied Nitrate Availability. *J. Appl. Phycol.* **2018**, *30*, 2763–2772. [[CrossRef](#)]
118. Zou, L.-G.; Balamurugan, S.; Zhou, T.-B.; Chen, J.-W.; Li, D.-W.; Yang, W.-D.; Liu, J.-S.; Li, H.-Y. Potentiation of Concurrent Expression of Lipogenic Genes by Novel Strong Promoters in the Oleaginous Microalga *Phaeodactylum tricornutum*. *Biotechnol. Bioeng.* **2019**, *116*, 3006–3015. [[CrossRef](#)]
119. Yoshinaga, R.; Niwa-Kubota, M.; Matsui, H.; Matsuda, Y. Characterization of Iron-Responsive Promoters in the Marine Diatom *Phaeodactylum tricornutum*. *Mar. Genom.* **2014**, *16*, 55–62. [[CrossRef](#)] [[PubMed](#)]
120. Shemesh, Z.; Leu, S.; Khozin-Goldberg, I.; Didi-Cohen, S.; Zarka, A.; Boussiba, S. Inducible Expression of Haematococcus Oil Globule Protein in the Diatom *Phaeodactylum tricornutum*: Association with Lipid Droplets and Enhancement of TAG Accumulation under Nitrogen Starvation. *Algal Res.* **2016**, *18*, 321–331. [[CrossRef](#)]
121. Lin, H.-Y.; Yen, S.-C.; Kuo, P.-C.; Chung, C.-Y.; Yeh, K.-L.; Huang, C.-H.; Chang, J.; Lin, H.-J. Alkaline Phosphatase Promoter as an Efficient Driving Element for Exogenic Recombinant in the Marine Diatom *Phaeodactylum tricornutum*. *Algal Res.* **2017**, *23*, 58–65. [[CrossRef](#)]
122. Lupette, J.; Tardif, M.; Brugière, S.; Couté, Y.; Salvaing, J.; Maréchal, E. Quantitative Proteomic Analyses Reveal the Impact of Nitrogen Starvation on the Proteome of the Model Diatom *Phaeodactylum tricornutum*. *Proteomics* **2022**, *22*, 2200155. [[CrossRef](#)]
123. Abida, H.; Dolch, L.-J.; Mei, C.; Villanova, V.; Conte, M.; Block, M.A.; Finazzi, G.; Bastien, O.; Tirichine, L.; Bowler, C.; et al. Membrane Glycerolipid Remodeling Triggered by Nitrogen and Phosphorus Starvation in *Phaeodactylum tricornutum*. *Plant Physiol.* **2015**, *167*, 118–136. [[CrossRef](#)]
124. The Galaxy Community The Galaxy Platform for Accessible, Reproducible and Collaborative Biomedical Analyses: 2022 Update. *Nucleic Acids Res.* **2022**, *50*, W345–W351. [[CrossRef](#)]
125. Bolger, A.M.; Lohse, M.; Usadel, B. Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014**, *30*, 2114–2120. [[CrossRef](#)]
126. Kim, D.; Langmead, B.; Salzberg, S.L. HISAT: A Fast Spliced Aligner with Low Memory Requirements. *Nat. Methods* **2015**, *12*, 357–360. [[CrossRef](#)] [[PubMed](#)]

127. Liao, Y.; Smyth, G.K.; Shi, W. FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* **2014**, *30*, 923–930. [[CrossRef](#)]
128. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [[CrossRef](#)] [[PubMed](#)]
129. Metsalu, T.; Vilo, J. ClustVis: A Web Tool for Visualizing Clustering of Multivariate Data Using Principal Component Analysis and Heatmap. *Nucleic Acids Res.* **2015**, *43*, W566–W570. [[CrossRef](#)] [[PubMed](#)]
130. Ge, S.X.; Jung, D.; Yao, R. ShinyGO: A Graphical Gene-Set Enrichment Tool for Animals and Plants. *Bioinformatics* **2020**, *36*, 2628–2629. [[CrossRef](#)]
131. Bardou, P.; Mariette, J.; Escudié, F.; Djemiel, C.; Klopp, C. Jvenn: An Interactive Venn Diagram Viewer. *BMC Bioinform.* **2014**, *15*, 293. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.