

Université de Montréal  
Département des Sciences Biologiques, Faculté des Arts et Sciences

*Cette thèse intitulée*

***Variations diurnes dans l'abondance et la vitesse de  
synthèse de protéines chez le dinoflagellé *Lingulodinium****

Présentée par

**Carl Bowazolo**

A été évaluée par un jury composé des personnes suivantes

**Annie Angers**

Président-rapporteur

**David Morse**

Directeur de recherche

**Anne-Lise Routier**

Membre du jury

**Connie Lovejoy**

Examineur externe

## Résumé

Les urgences écologiques actuelles, résultant des dérives de l'industrialisation et de la mondialisation, mènent la recherche fondamentale à se préoccuper hâtivement de la biologie de certains organismes comme les dinoflagellés. Ces organismes sont à la base de la chaîne alimentaire, et certaines espèces sont impliquées dans la formation des récifs coralliens. De plus, la prolifération incontrôlée de certaines espèces de dinoflagellés engendrée par l'eutrophisation industrielle est responsable des marées rouges qui menacent d'une part la flore et la faune aquatiques. Certaines des toxines produites par ces efflorescences algales ont un effet sur la santé publique, car elles peuvent rendre dangereuse la consommation alimentaire de poissons ou de fruits de mer contaminés.

Nous avons examiné le comportement de l'espèce *Lingulodinium polyedra* à travers des cycles diurnes. Dans un premier projet, le protéome dans des extraits d'algues récoltés sur un cycle de 24 heures en conditions de 12 heures de jour suivies de 12 heures de nuit (un cycle LD) a été examiné. Nous avons identifié treize protéines clés qui ont montré des variations quantitatives synchronisées avec les temps de réalisation des rythmes biologiques les impliquant. Deux protéines déjà connues de varier faisait partie de ce groupe, tandis que les autres protéines sont impliquées dans des processus rythmiques nouveaux. Nous avançons l'hypothèse qu'une augmentation quantitative des protéines clés permettrait à l'accomplissement des différents processus cellulaires à différents moments de la journée.

Dans un second projet, nous avons examiné la vitesse de synthèse des protéines qui pourrait expliquer ces variations quantitatives protéiques. En appliquant la technique du profilage ribosomal (« *ribosome profiling* ») sur des échantillons algaux collectés d'abord à trois temps, puis ensuite sur tout le cycle LD, nous avons confirmé des variations de vitesse de

synthèse des protéines identifiées dans notre étude protéomique et des protéines dont la vitesse de synthèse a déjà été mesurée par l'incorporation de méthionine radioactive *in vivo*. De plus, nous avons identifié plusieurs milliers d'autres séquences dont la vitesse de synthèse varie. Une classification des séquences dans les catégories GO nous a permis d'identifier un rythme diurne dans la synthèse des acides aminés qui pourrait aider à comprendre le métabolisme du nitrate. Nous proposons que des variations de vitesse de synthèse entraînent des augmentations quantitatives des protéines aux niveaux inférieurs à ce que l'on a pu détecter par la spectroscopie de masse. La variation dans les niveaux de protéines clés pourrait aider l'accomplissement des rythmes diurnes chez le dinoflagellé *Lingulodinium*.

Ces avancées dans la compréhension de la régulation des rythmes biologiques du dinoflagellé *Lingulodinium* d'une part permettront de mieux penser la recherche concernant la lutte écologique contre les marées rouges et d'autre part ouvriront de nouvelles perspectives dans l'entendement de la régulation des rythmes biologiques des autres organismes y compris ceux de l'Homme dont les troubles impliquent de nombreuses maladies.

**Mots-clés :** Dinoflagellé, *Lingulodinium polyedra*, profilage des ribosomes, protéome, rythmes biologiques, traduction.

## Abstract

Industrialization and globalization has led to industrial eutrophication in many regions of the oceans resulting in different types of ecological emergencies. One of these is the uncontrolled proliferation of dinoflagellates which are responsible for the harmful algal blooms called “red tides”. Certain species release toxins which can harm aquatic flora and fauna and, through consumption of contaminated fish or seafood, can affect public health. Dinoflagellates are also at the base of the food chain in the marine environment, and certain species form symbioses with anthozoans thus allowing growth of coral reefs. There is thus a strong need for fundamental research on dinoflagellate biology.

Using the dinoflagellate *Lingulodinium polyedra* as a model, we have examined how cells adapt to the daily changes in light and dark. In particular we have examined changes in protein amounts as well as changes in protein synthesis rates to gain insight into how cells differ during the day and night.

In a first project, the daily proteome was examined by mass spectrometry using algae extracts harvested over a 24 hour cycle in conditions of 12 hours of day followed by 12 hours of night. Key proteins involved in biological rhythms have shown quantitative variations synchronized with the times of the biological rhythms involving them, ie a quantitative increase in key proteins is necessary for the accomplishment of the various cellular processes. Thirteen proteins were identified, of which two were previously documented to change in amount over the daily cycle.

In a second project we addressed the origin of these quantitative protein variations. By applying the technique of ribosome profiling on algal samples, first collected in triplicate at three times and then at two hour intervals throughout the 24 h cycle, we have identified

variations in the synthesis rate of thousands of proteins. These include the proteins found in the proteomic analysis as well as four proteins whose synthesis rate variations had already been observed using *in vivo* metabolic labeling. Interestingly, classification of the regulated proteins into GO categories also revealed a late night increase in the synthesis of many amino acid biosynthetic enzymes, potentially linking amino acid synthesis associated with the daily metabolism of nitrate.

These advances in our understanding of the regulation of the daily rhythms of the dinoflagellate *Lingulodinium* will provide new tools in the ecological fight against red tides. They may also open new perspectives in understanding the daily rhythms of other organisms.

**Keywords:** Dinoflagellate, *Lingulodinium polyedra*, ribosome profiling, proteome, biological rhythms, translation.

## Table des matières

RÉSUMÉ.....	I
ABSTRACT.....	III
TABLE DES MATIÈRES .....	V
LISTE DES TABLEAUX.....	VII
LISTE DES FIGURES.....	VIII
LISTE DES ABRÉVIATIONS .....	X
REMERCIEMENTS.....	XV
CHAPITRE 1 : INTRODUCTION ET REVUE DE LITTÉRATURE .....	1
1.1 ROLES ET PHENOMENES CIRCADIENS DU DINOFLAGELLE <i>LINGULODINIUM</i> .....	2
1.1.1 PRÉSENTATION ET RÔLES DU DINOFLAGELLÉ <i>LINGULODINIUM</i> .....	2
1.1.2 LES PHÉNOMÈNES CIRCADIENS CHEZ LE DINOFLAGELLÉ <i>LINGULODINIUM</i> .....	2
1.2 STRATEGIES DE CONTROLE BIOCHIMIQUE CHEZ <i>LINGULODINIUM</i> .....	4
1.2.1 CONTRÔLE BIOCHIMIQUE DE LA BIOLUMINESCENCE.....	4
1.2.2 CONTRÔLE BIOCHIMIQUE DE LA PHOTOSYNTHÈSE .....	7
1.3 STRATEGIES DE CONTROLE DE L'EXPRESSION GENIQUE CHEZ <i>LINGULODINIUM</i> .....	10
1.3.1 CONTRÔLE DE LA TRANSCRIPTION ET DE LA QUANTITÉ DES ARNm .....	10
1.3.2 CONTRÔLE DU TEMPS DE SYNTHÈSE PROTÉIQUE ET DES CHANGEMENTS DU NIVEAU PROTÉIQUE .....	16
1.3.3 CONTRÔLE DE L'ACTIVITÉ ENZYMATIQUE.....	18
1.4 UN MÉCANISME POSSIBLE D'UNE HORLOGE CIRCADIEN CHEZ <i>LINGULODINIUM</i> .....	21
1.5 OBJECTIFS .....	27
CHAPITRE 2: LABEL-FREE MS/MS ANALYSES OF THE DINOFLAGELLATE <i>LINGULODINIUM</i> IDENTIFIES RHYTHMIC PROTEINS FACILITATING ADAPTATION TO A DIURNAL LD CYCLE .....	33
LABEL-FREE MS/MS ANALYSES OF THE DINOFLAGELLATE <i>LINGULODINIUM</i> IDENTIFIES RHYTHMIC PROTEINS FACILITATING ADAPTATION TO A DIURNAL LD CYCLE.....	34
2.1. ABSTRACT .....	35
2.2. INTRODUCTION .....	36

<b>2.3. RESULTS</b> .....	<b>38</b>
<b>2.4. DISCUSSION</b> .....	<b>40</b>
<b>2.5 METHODS</b> .....	<b>47</b>
<b><i>DATA ANALYSIS</i></b> .....	<b>49</b>
<b>CHAPITRE 3: ORCHESTRATED TRANSLATION SPECIALIZES DINOFLAGELLATE METABOLISM THREE TIMES A DAY.</b> .....	<b>60</b>
<b>3.1 ABSTRACT</b> .....	<b>62</b>
<b>3.2. INTRODUCTION</b> .....	<b>62</b>
<b>3.3 RESULTS AND DISCUSSION</b> .....	<b>64</b>
<b>3.4. METHODS</b> .....	<b>69</b>
<b>CHAPITRE 4 : DISCUSSION GÉNÉRALE</b> .....	<b>118</b>
<b>CHAPITRE 5. PROPOSITION D'ÉVENTUELS PROCHAINS PROJETS À PRIORISER</b> .....	<b>127</b>
<b>CHAPITRE 6. CONCLUSION</b> .....	<b>129</b>
<b>RÉFÉRENCES</b> .....	<b>130</b>

## Liste des tableaux

Tableau 1. Variations quotidiennes des concentrations de luciférase (en µg/mL) dans les extraits protéiques de <i>Lingulodinium</i> , déterminées par ELISA. Inspiré de (Johnson, Roeber et al. 1984). .....	30
Tableau 2. 1. Proteins with significant changes in levels over the LD cycle .....	51
Tableau 2. 2. Proportion of rhythmic proteins in a range of systems. ....	52
Tableau 3. 1. Number of reads before and after trimming (test run) with an LD14 sample. .	110
Tableau 3. 2. Number of reads before and after trimming (3 times). ....	111
Tableau 3. 3. Number of reads before and after trimming (12 times). ....	112
Tableau 3. 4. Average RPF counts for diurnally regulated proteins at three times. ....	113
Tableau 3. 5. Consensus motifs detected in transcripts whose synthesis is highest near LD0, LD12, or LD16.....	114
Tableau 3. 6. Amino acid levels (pmol/mg FW, n=5) at four times in an LD cycle. ....	115
Tableau 3. 7. Number of sequences and enzymes in various KEGG annotated pathways.....	116
Tableau 4. 1. Évolution quantitative de la vitesse de synthèse de la chlorophylle synthase et de la phytoène désaturase à ZT6, 14 et 20 chez le dinoflagellé <i>Lingulodinium</i> par ribosome profiling.....	123



## Liste des figures

Figure 1. 1. Modèle simplifié de boucle de rétrocontrôle.....	31
Figure 1. 2. Cycle biologique de <i>Lingulodinium</i> sous un cycle LD12: 12 avec les heures des pics des deux principaux phénomènes diurne et nocturne. ....	32
Figure 2. 1. Range of protein abundance. ....	53
Figure 2. 2. Proteins with diurnal changes fall into two groups. ....	54
Figure 2. 3. The diurnally varying proteins plotted by abundance. ....	55
Figure 2. 4. Schematic overview of the cell localisation for diurnally varying proteins.....	56
Figure 2. 5. (Supplementary) Comparison of LBP levels measured by Western blotting and by MS/MS.....	57
Figure 2. 6. (Supplementary) Comparison of TPI levels measured by coupled enzyme assays and by MS/MS. ....	58
Figure 3. 1. A complex temporal landscape of protein synthesis rates with three peak times. ....	73
Figure 3. 2. Enzymes involved in carbon metabolism are regulated around LD0.....	75
Figure 3. 3. Known night phase rhythms are concurrently regulated around LD12. ....	77
Figure 3. 4. Enzymes involved in regulation of gene expression and amino acid biosynthesis are translated around LD16.....	79
Figure 3. 5. (Supplementary) Electrophoretic analysis of RPF. ....	81
Figure 3. 6. (Supplementary) Size distribution of Ribosome Protected Fragments (RPF).....	83
Figure 3. 7. (Supplementary) Similarity of triplicate samples to one another.....	85
Figure 3. 8. (Supplementary) Read distribution for LBP.....	86
Figure 3. 9. (Supplementary) Comparison of RPF read counts and <sup>35</sup> S-methionine <i>in vivo</i> labeling.....	87
Figure 3. 10. (Supplementary) Comparison of RPF read counts with diurnal variations in protein levels. ....	89
Figure 3. 11. (Supplementary) Principal GO categories for enriched RPF at LD6. ....	91
Figure 3. 12. (Supplementary) Heat maps for proteins related to the light reaction of photosynthesis.....	93
Figure 3. 13. (Supplementary) Pentose phosphate shunt.....	94

Figure 3. 14. (Supplementary) Fatty acid metabolism.....	96
Figure 3. 15. (Supplementary) Starch metabolism. ....	97
Figure 3. 16. (Supplementary) Biosynthetic pathways for valine, leucine and isoleucine. ....	98
Figure 3. 17. (Supplementary) Cellulose metabolism. ....	99
Figure 3. 18. (Supplementary) Fructose/Mannose Metabolism.....	100
Figure 3. 19. (Supplementary) Principal GO categories for enriched RPF at LD14. ....	101
Figure 3. 20. (Supplementary) Pyruvate metabolism. ....	102
Figure 3. 21. (Supplementary) Purine Metabolism.....	103
Figure 3. 22. (Supplementary) Pyrimidine metabolism.....	104
Figure 3. 23. (Supplementary). Principal GO categories for enriched RPF at LD20. ....	105
Figure 3. 24. (Supplementary) Amino acid biosynthetic enzymes regulated in the 3 times experiment.....	107
Figure 3. 25. (Supplementary) Amino acid profiles over a daily cycle shows increases during the dark phase. ....	108
Figure 3. 26. (Supplementary) Correlation between RPF read counts with transcript abundance. ....	109
Figure 4. 1. Rapport journalier du nombre de copies en cours de synthèse entre la phytoène désaturase et la chlorophylle synthase chez <i>Lingulodinium</i> . ....	124
Figure 4. 2. A. Schéma proposé de la régulation des rythmes biologiques chez le dinoflagellé <i>Lingulodinium</i> . B. Modèle simplifié de boucle de rétrocontrôle dans une cellule de l'horloge de la drosophile.....	125

## Liste des abréviations

ACON: Aconitase

ALA: Alanine

ARG: Arginine

ASN: Asparagine

ASP: Aspartate

BMAL1: Brain and Muscle ARNT-Like 1

CCA1: Circadian Clock Associated 1

cCL: conventionnelle CrossLinking

CK: Caséine Kinase

CoAT: Oxalate CoA transferase

CS: Citrate Synthase

Cyb2: Cytochrome b2

CYS: Cystéine

DEPC: Diéthyl pyrocarbonate

DHAP: Dihydroxyacetone phosphate

DIA: data-independent acquisition

DMSO: Diméthylsulfoxyde

DTT: Dithiothréitol

EC: Enzyme Commission numbers

eEF2: Eukaryotic elongation factor 2

eEF3: Eukaryotic elongation factor 3

eIF4E: Eukaryotic translation initiation factor 4E

ELISA: Enzyme-Linked Immunosorbent Assay

ER: Endoplasmic Reticulum

ESI: Electrospray ionization

EVS: 2-epi-5-epi-valiolone synthase

F1,6BP: Fructose 1,6 bis phosphate

F6P: Fructose 6 Phosphate

FA: Formic Acid

FDR: False Discovery Rate

FRQ: Frequency

GAP ou G3P: Glyceraldehyde-3-phosphate

GAPDH: glycéraldéhyde-3-phosphate déshydrogénase

Glc: glucose

GLN:Glutamine

GLU:Glutamate

GLY: Glycine

GO: Gene ontogeny

HIS: Histidine

HPLC: High-Performance Liquid Chromatography

IAA: Iodoacetic acid

ICL: IsoCitrate Lyase

ILE: Isoleucine

KEGG: Kyoto Encyclopedia of Genes and Genomes

LBP: Luciferin Binding Protein

LCF ou LUC: Luciferase

LC-MS/MS: Liquid Chromatography-Mass Spectrometry/ Mass Spectrometry

LD: Light/Dark

LEU: Leucine

LHC: Light-Harvesting Complex

LHY: Late Elongated Hypocotyls

LL: Light/Light

LYS:Lysine

MAA: Mycosporine-like amino acids

MAS: Malate Synthase

MES (buffer): 2-(N-morpholino)ethanesulfonic acid

MET: Méthionine

NADP-ICDH: IsoCitrate Déshydrogénase

NAF: Sodium fluoride

nt: nucleotide

OEE: Oxygen-evolving enhancer protein 1

PAR-CL: Photo-Activable-Ribonucléoside CrossLinking

PCNA: Proliferating cell nuclear antigen

PCP: Peridinin-Chlorophyll a binding protein

PDX1: Pyridoxal phosphate synthase protein

PEP: Phosphoenol pyruvate

PEPC: Phosphoénolpyruvate carboxylase

PER: Period circadian protein

PFK: Phosphofructokinase  
PFP: Pyrophosphate fructose-1,6-phosphate phosphotransferase  
PHE: Phénylalanine  
PMSF: Phenylmethylsulfonyl fluoride  
PNK: Polynucleotide kinase  
PRO: Proline  
PSII D1: Photosystem II D protein  
PTM: Post-translation modification  
RBP: RNA Binding Protein  
RPF: Ribosome protected fragments  
RPKM: Reads per Kilobase per Million Reads  
RUB: Rubisco  
SAR: Stramenopiles-Alveolates-Rhizaria  
Sar1: Secretion-associated RAS-related protein 1  
SCN: Suprachiasmatic nucleus  
SER: Serine  
TBE: Tris-Borate-EDTA  
TCA: Tricarboxylic acid  
TFA: Acide trifluoroacétique  
THR: Threonine  
TPI: Triosephosphate isomerase  
TRP: Tryptophane  
TYR: Tyrosine

UTR: Untranslated region

VAL: Valine

WCC: White Collar Complex

ZT : Zeitgeber Time

## **Remerciements**

Tout d'abord une immense gratitude et reconnaissance envers mon directeur David Morse pour son inestimable confiance et son excellent encadrement m'ayant permis d'être à la hauteur des exigences de cette opportunité doctorale, grâce à lui me voilà spécialiste mondial du riboprofilage des dinoflagellés.

Ensuite, un grand merci à Mathieu Beauchemin pour l'accueil enthousiaste et l'initiation doctorale remarquable, à Steve Dagenais-Bellefeuille et Sébastien Dumont pour leurs conseils et leur sollicitude, à Julie Augustin pour le travail harmonieux effectué ensemble lors des démos, à Bahar pour sa douce et bienveillante compagnie... et à Nicolas Boivin pour sa générosité et son altruisme...

Enfin, merci à Jean Rivoal et Sonia Dorion pour leur attention et leurs fructueuses collaborations.



Malade et seul en exil,  
Tu étais là depuis Brazzaville...  
Confiné et isolé en exil  
Tu étais encore là depuis Brazzaville...  
Ton soutien comme ton amour  
Ici m'aura été d'un vital secours  
Chère tendre maman  
Encore merci pour tous ces moments...

Quand les épreuves se faisant chorale  
Ont fait chanter mon moral  
Tu me fus bon conseiller  
Tonton Guénolé merci de m'avoir éveillé...

Profonds et infinis remerciements  
À ma chère tante Helda qui m'avait chaleureusement  
Si bien accueilli  
Et hébergé durant ma première semaine dans ce lointain pays...  
Tu étais fière de me savoir doctorant,  
Aujourd'hui tu l'aurais été encore plus si tu avais été au courant...  
Ta joie a tropicalisé mon Québec en plein hiver  
Alors je te devais bien ces quelques vers...

*À Maître Emmanuel qui me disait: "Tous ces méandres mènent au doctorat,  
Accroche-toi et tu l'auras..."*

## **Chapitre 1 : Introduction et revue de littérature**

## **1.1 Rôles et phénomènes circadiens du dinoflagellé *Lingulodinium***

### **1.1.1 Présentation et rôles du dinoflagellé *Lingulodinium***

*Lingulodinium polyedra* (autrefois appelé *Gonyaulax polyedra*) (Dodge 1989) est un protiste unicellulaire photosynthétique, de couleur rouge orangé mesurant en moyenne 35 x 45µm (Schmitter 1971). Ce dinoflagellé marin est recouvert d'une thèque cellulosique (Adamich and Sweeney 1976) et est doté de deux flagelles lui permettant d'être dynamique dans les milieux aquatiques (Taylor 1975). Ce phytoplancton est retrouvé aussi bien dans les milieux marins tempérés comme les côtes nord-américaines que dans les milieux marins équatoriaux le long des côtes africaines (Zonneveld 2013). Dans les océans, les dinoflagellés participent à la vie marine en étant à la base de la chaîne alimentaire et en contribuant à la formation des récifs coralliens (Stat, Morris et al. 2008) (Field, Behrenfeld et al. 1998). En conditions optimales de prolifération rapide (abondance de nutriments et de soleil) ces protistes sont également responsables de phénomènes époustouffants: de nuit le phénomène de bioluminescence (production d'une lumière bleue à la surface des eaux) (Harvey 1952) et de jour le phénomène de marée rouge (coloration rouge des eaux) (Morales-Ramirez, Viquez et al. 2001). Malheureusement, leurs concentrations très élevées (Bennouna 2002) peuvent parfois conduire à des problèmes majeurs de santé publique, car certaines espèces produisent des toxines (Honner, Kudela et al. 2012).

### **1.1.2 Les phénomènes circadiens chez le dinoflagellé *Lingulodinium***

La plupart des dinoflagellés peuvent aussi être cultivés dans des milieux aquatiques artificiels (Guillard and Ryther 1962). Dans ce cas, les salles de culture sont normalement sous des cycles LD12:12, à savoir 12h de jour suivies de 12h d'obscurité (par convention, le début du jour s'appelle LD0 et le début de la nuit LD12). Cette culture algale en laboratoire a permis l'étude

de leurs rythmes biologiques. Ainsi, *Lingulodinium* est devenu un modèle pour étudier les cycles circadiens, qui sont les processus biologiques ayant une périodicité d'environ 24h même en conditions constantes d'obscurité ou de lumière (Hastings, Astrachan et al. 1961). Ces rythmes, dits rythmes circadiens, sont d'origine endogène et sont contrôlés par un chronomètre biochimique appelé une horloge circadienne. Des chronomètres endogènes furent découverts par Jean-Jacques d'Ortous de Mairan en 1729, en observant la persistance d'un rythme d'ouverture des feuilles de mimosa lorsque les plantes étaient placées en obscurité constante. Chez *Lingulodinium*, on distingue plusieurs phénomènes circadiens, dont l'acrophase (ou le point maximal) peut être pendant le jour ou la nuit.

Parmi les phénomènes circadiens nocturnes, on observe la bioluminescence (Hastings and Sweeney 1958) qui semble leur permettre de se défendre face aux prédateurs (Lindstrom, Grebner et al. 2017). Certaines nuits, lorsque la mer est agitée ou quand un bateau passe, l'eau s'illumine magistralement (Harvey 1952), et ce spectacle bioluminescent peut aussi être observé au laboratoire en agitant les flasques d'algues pendant leur période l'obscurité. On voit également la réplication de l'ADN, appelé la phase S du cycle cellulaire, qui a lieu aux alentours de LD18 (Dagenais-Bellefeuille, Bertomeu et al. 2008).

Le principal phénomène circadien diurne est la photosynthèse qui leur permet de synthétiser et d'emmagasiner les sucres dont ils ont besoin pour leur énergie (Calvin and Benson 1948). Ce rythme se manifeste autant par un rythme dans la quantité de carbone fixé que par la quantité d'oxygène libéré.

À ces phénomènes circadiens on peut rajouter la migration verticale diurne, car les algues se répartissent en profondeur juste avant la période de nuit, et remontent à la surface juste avant le jour (Eppley, Holm-Hansen et al. 1968), même dans les conditions constantes. Cette

différence quotidienne de leur position dans la colonne d'eau s'explique en partie parce que les algues se positionnent de façon à bénéficier au maximum de la lumière du jour nécessaire à la réalisation de leur photosynthèse. La raison de leur descente nocturne est moins claire et a originalement été proposée comme une façon d'obtenir plus d'azote (Cullen and Horrigan 1981) (Cullen 1985). Une autre possibilité serait que cette migration leur permet de fuir certains prédateurs qui eux migrent massivement en surface la nuit pour se nourrir et pour échapper à leurs propres prédateurs (Hays, Kennedy et al. 2001). Cette migration verticale (océanique, marine, fluviale ou lacustre) est aussi observée en laboratoire, car ce comportement a été décrit dans plusieurs études notamment celle de Moorthi (Moorthi, Countway et al. 2006).

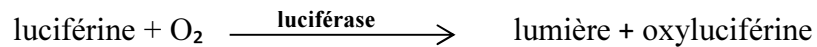
## **1.2 Stratégies de contrôle biochimique chez *Lingulodinium***

### **1.2.1 Contrôle biochimique de la bioluminescence**

Pendant la phase de nuit, *Lingulodinium polyedra* peut émettre un éclair, ou « flash » lumineux, lorsqu'une cellule détecte le mouvement dans le milieu externe. L'intensité du flash est environ  $10^9$  quanta/sec/cellule, et dure environ 0.1 seconde. La bioluminescence d'une population de cellules dans une flasque peut durer plusieurs secondes quand le contenant est agité continuellement à l'obscurité et la bioluminescence est facilement visible à l'œil nu. Curieusement, une émission lumineuse de faible intensité (lueur, ou « glow ») d'environ  $10^4$  quanta/sec/cellule peut également être détectée en fin de nuit. L'origine de ce mode d'émission de la lumière n'est pas complètement claire, mais correspond dans le temps au moment où les protéines de la bioluminescence sont en train d'être dégradées par la cellule (voir plus bas). Les pics de productivité de la lumière sont observés durant la phase de nuit, soit dans le nombre des éclairs ou dans l'intensité de la lueur (Hastings and Sweeney 1958). Ainsi certaines nuits, cette

capacité bioluminescente permet aux étendues aquatiques de briller splendidement où ces phytoplanctons abondent (Harvey 1952).

Cette lumière bleue est produite par l'action de la luciférase lors de l'oxydation de son substrat la luciférine en présence de dioxygène (Dunlap and Hastings 1981, Dunlap, Taylor et al. 1981) selon la réaction biochimique suivante:



La luciférine est oxydée en un intermédiaire excité capable d'émettre un photon qui se dégrade par la suite en oxyluciférine (Hastings 1989) (Nakamura, Kishi et al. 1989). La rythmicité de la bioluminescence chez *Lingulodinium* s'explique en partie parce que les niveaux cellulaires de luciférase varient en fonction d'un rythme circadien, c'est-à-dire en plein jour ces quantités sont faibles tandis que la nuit elles sont relativement élevées (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989). En phase d'obscurité, les quantités de luciférase sont 7 à 17 fois plus importantes qu'en phase de lumière (Tableau 1), et les oscillations protéiques liées aux rythmes de bioluminescence sont similaires en LD et LL. Ces variations cycliques qui continuent même en conditions lumineuses constantes (LL) démontrent leur contrôle par l'horloge circadienne. Les variations dans la quantité de la luciférase sont le résultat d'une augmentation dans la vitesse de synthèse au début de la nuit et d'une augmentation de la vitesse de dégradation vers la fin de la nuit. Les deux seraient donc contrôlées par l'horloge endogène. L'activité d'une réaction enzymatique pouvant être proportionnelle à la quantité de l'enzyme, les variations quantitatives enzymatiques (luciférase) contribuent à la quantité de lumière émise.

LBP, la protéine de liaison de luciférine, est l'autre protéine clé de la bioluminescence chez *Lingulodinium* (Morse, Pappenheimer et al. 1989). LBP est localisée, comme la luciférase, dans les scintillons, le nom donné à l'organite qui produit de la bioluminescence (Nicolas, Morse

et al. 1991). Ces structures cellulaires principalement nocturnes se trouvent tout près de la membrane vacuolaire vers la périphérie de la cellule. Au cours de leur synthèse, ces deux protéines bioluminescentes se réunissent au centre de la cellule sous une forme « préscintillon » avant de migrer vers la périphérie pour former les scintillons matures (Nicolas, Morse et al. 1991) (Johnson, Inoué et al. 1985) (Desjardins and Morse 1993) (Fritz, Morse et al. 1990). LBP peut lier la luciférine de façon réversible, et cette liaison est médiée par le pH. Une baisse du pH fait en sorte que la luciférine est relâchée par la LBP et liée par la luciférase (Morse, Pappenheimer et al. 1989). Le rôle de LBP semble être de protéger la luciférine contre l'oxydation non bioluminescente, car la luciférine seule en présence d'oxygène est rapidement dégradée dans une réaction qui ne produit pas la lumière (Nakamura, Kishi et al. 1989). La protéine LBP varie elle aussi à travers le cycle de 24 heures. Le pic dans la protéine a la même phase que la luciférase, et l'amplitude du rythme est similaire aussi (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989). LBP détermine les niveaux de substrat disponible, et parce que la quantité de substrat contribue aussi à l'activité d'une réaction, cette protéine aussi contribue à l'amplitude du rythme de la bioluminescence. La quantité de la protéine LBP peut atteindre 1% de protéines totales durant la phase de nuit, ce qui correspond à environ  $10^8$  molécules. Ceci concorde assez bien au nombre de quanta de la quantité de lumière produite par une cellule dans un flash lumineux (Morse, Pappenheimer et al. 1989).

La conclusion de l'ensemble de ces études est que le rythme de la production de la lumière peut être le résultat de changements dans la quantité de protéines clés dans une réaction. Ceci représente une stratégie très importante qui pourrait soutenir d'autres rythmes circadiens d'un organisme.

### 1.2.2 Contrôle biochimique de la photosynthèse

La photosynthèse oxygénique est un processus bioénergétique essentiel des organismes photoautotrophes comme *Lingulodinium*. Elle est constituée de deux phases, la phase photochimique (les réactions lumineuses) et la phase d'incorporation (les réactions sombres). La phase photochimique permet la production de coenzymes réduits (NADPH) et de l'énergie (ATP), toutes les deux utilisées pour la fixation du dioxyde de carbone sous forme de sucres pendant le cycle de Calvin. La phase d'incorporation fait intervenir une enzyme clé appelée la ribulose-1,5-bisphosphate carboxylase/oxygénase, ou Rubisco (EC 4.1.1.39) (Raven 1995). Rubisco catalyse la condensation d'une molécule de dioxyde de carbone avec un sucre à cinq carbonés (Calvin and Benson 1948) formant du glucose. D'autre part, cette même enzyme peut aussi catalyser l'oxygénation du ribulose-1,5-bisphosphate initiant la photorespiration (Lorimer, Andrews et al. 1973). La photorespiration ajoute un O<sub>2</sub> au sucre au lieu d'un CO<sub>2</sub>, et est donc un processus qui réduit le rendement de la photosynthèse. La cause est la capacité de la Rubisco de lier soit le CO<sub>2</sub> soit l'O<sub>2</sub>. La capacité de distinguer les deux gaz s'appelle le ratio de sélectivité, S<sub>C/O</sub>.

Structurellement et fonctionnellement quatre types de Rubisco sont distingués. Le type I se retrouve chez la plupart des phototrophes oxygéniques et est le plus abondant. Elle est formée de huit grandes et huit petites sous unités, et démontre une grande affinité pour le dioxyde de carbone. Le type II est formé de grandes sous unités, typiquement 2, et son affinité pour le dioxyde de carbone est inférieure à ce que l'on voit avec le type I. Le type II se retrouve chez les bactéries anaérobies et chez les dinoflagellés tels que *Lingulodinium* dont les chloroplastes contiennent de la péridinine comme pigment photosynthétique (Tabita 1995)



(Raven 1997) (Morse, Salois et al. 1995). Le ratio  $S_{C/O}$  est environ dix fois moins pour la forme II que pour la forme I, et le problème de photorespiration dû à Rubisco risque donc d'être plus important chez les dinoflagellés. La forme III se trouve uniquement dans quelques archéobactéries, et la forme IV ne catalyse pas la carboxylation ou l'oxygénation du ribulose.

Les niveaux de la protéine Rubisco type II sont constants au cours de la journée chez *Lingulodinium*, et donc différent de la luciférase. À la place, c'est la répartition dans les chloroplastes qui varie au cours du temps, et cette variation est corrélée avec le rythme circadien de fixation du dioxyde de carbone (Nassoury, Fritz et al. 2001). Quand la capacité de fixer du dioxyde de carbone est élevée, la Rubisco type II est localisée au centre de la cellule dans une partie des chloroplastes appelés des pyrénoides (Nassoury, Fritz et al. 2001). Par contre, quand la vitesse de fixation du carbone est moindre, la Rubisco se trouve distribuée partout à la surface du chloroplaste. La différence dans la distribution est importante, car lorsque Rubisco se trouve dans les pyrénoides, la majeure partie de la protéine PCP (Péridinine-Chlorophylle-Protéine) se trouve dans les régions corticales du plastide. PCP est la principale protéine récolteuse de lumière, et sa présence indique le site de production d'oxygène par les réactions photochimiques. La Rubisco de forme II, plus susceptible d'utiliser l'oxygène dans la photorespiration que la forme I, se trouve éloignée de la source d'oxygène pendant le jour. La Rubisco plus loin de la source d'oxygène pourrait bénéficier d'un ratio plus favorable de  $CO_2/O_2$  autour de cette enzyme.

Les variations circadiennes de la capacité maximale de fixation du dioxyde de carbone et celles des niveaux de production maximale de dioxygène ont été suivies pendant plusieurs jours (Nassoury, Fritz et al. 2001). En conditions de lumière continue les deux rythmes se

poursuivent avec une amplitude d'environ trois fois chacun et des périodes similaires. Les phases des deux rythmes sont différentes par contre: les taux maximaux de fixation du carbone précèdent les taux maximaux d'évolution de l'oxygène d'environ 6 h. De plus, les formes des courbes des deux rythmes sont différentes, avec une courbe de fixation maximale du carbone en dents de scie et une courbe sinusoïdale pour l'évolution maximale des niveaux d'oxygène. La différence entre les deux rythmes s'explique par le système de dosage pour la fixation de carbone, qui a utilisé une concentration du carbone et une intensité lumineuse saturante. Cette courbe suit la formation des pyrénoides, les structures qui facilitent la fixation de carbone. Par contre, quand les mesures de fixation de carbone sont faites via une mesure de pH, ce qui permet de visualiser le rythme de fixation du carbone dans les cultures normales (avec les concentrations de carbone et de l'intensité lumineuse normale), on voit que les rythmes de fixation du carbone et l'évolution d'oxygène ont la même phase et la forme des rythmes est la même (Mackenzie and Morse 2011).

L'étude du rythme de la photosynthèse nous a permis de déceler une deuxième stratégie pour produire un rythme. Dans ce cas, il s'agit d'un contrôle sur la disponibilité du substrat au lieu de la quantité de l'enzyme. L'ensemble des processus biochimiques qui constitue la base moléculaire du rythme circadien d'un organisme peut donc être contrôlé par différentes stratégies. Il est aussi possible que la cellule ait à ajuster l'activité spécifique d'une enzyme. On n'a pas encore trouvé un exemple de ceci chez *Lingulodinium*, mais l'activité de l'enzyme PEPC (Phosphoénolpyruvate carboxylase) chez la plante supérieure *Kalanchoe* est modulée par phosphorylation (Theng, Agarie et al. 2007) et contribue au rythme de photosynthèse dans cette espèce.

### 1.3 Stratégies de contrôle de l'expression génique chez *Lingulodinium*

#### 1.3.1 Contrôle de la transcription et de la quantité des ARNm

La transcription est une étape clé dans le processus de l'expression génique, car elle permet le passage de l'information génétique présente dans les gènes sous forme acide désoxyribonucléique (ADN) à la forme acide ribonucléique messenger (ARNm). Ce processus est typiquement contrôlé au niveau de la vitesse de transcription par des facteurs de transcription spécifiques. L'information génétique sous forme d'ARNm sera ensuite traduite en protéines dont dépendra le fonctionnement de pratiquement tous les processus cellulaires. La synthèse des protéines dépend en partie de la quantité d'ARNm, déterminée par la balance entre la vitesse de transcription et la vitesse de dégradation. La vitesse de traduction dépend également de facteurs spécifiques, tels que les protéines ou des petits ARN, qui permettent de régler chaque transcrit indépendamment des autres. En principe, toutes ces étapes peuvent être réglées par les horloges biologiques.

Les horloges circadiennes dans la plupart des cellules eucaryotes exploitent elles-mêmes des contrôles sur l'expression génique. En règle générale, on parle d'un (ou plusieurs) gène de l'horloge dont les protéines synthétisées sont capables d'inhiber leur propre transcription par un rétrocontrôle négatif (Figure 1.1). Avec ce contrôle sur la transcription des gènes de l'horloge, il n'est pas surprenant de voir que les horloges peuvent aussi contrôler la transcription d'autres gènes, génériquement appelés des gènes contrôlés par l'horloge. En effet, la plupart des organismes contrôlent la quantité de plusieurs ARNm, affectant ainsi la synthèse des protéines nécessaires aux différents phénomènes circadiens. Par exemple chez *Neurospora* une variation de 5 à 20% des ARNm a été notée (Nowrousian, Duffield et al. 2003) (Correa, Lewis et al. 2003) (Hurley, Dasgupta et al. 2014). D'autre part, cette variation est de 10%, de 5 à 10%, et de 30 à

65%, respectivement, chez *Arabidopsis thaliana* (Harmer, Hogenesch et al. 2000) (Schaffer, Landgraf et al. 2001), la souris (Akhtar, Reddy et al. 2002) (Panda, Antoch et al. 2002) et chez les cyanobactéries *Synechococcus* (Woelfle and Johnson 2006) (Ito, Mutsuda et al. 2009). Une variation des quantités d'ARNm a été notée chez deux espèces de dinoflagellés par puce d'ADN (*microarray*). Chez *Pyrocystis lunula*, 3% des gènes ont montré une variation quantitative (Okamoto and Hastings 2003), et chez *Karenia brevis*, 0,7% des ARNm changent d'abondance à travers la journée en condition de lumière ou d'obscurité constante (Okamoto and Hastings 2003).

En revanche, une situation différente a été vue suite à l'étude par ARN Seq des ARNm chez *Lingulodinium* et *Symbiodinium*. Chez *Symbiodinium*, seulement sept transcrits sur environ 70,000 avaient une abondance différente à LD0 et LD12 par ARN Seq (Zaheri et al 2019). Chez *Lingulodinium*, une étude a d'abord comparé deux temps, midi (LD6) et minuit (LD18), et ce dans les conditions LD et LL. L'ARN Seq n'a révélé aucune variation statistiquement significative des niveaux de transcrits dans le transcriptome (plus que 74,000 transcrits). Dans une deuxième expérience avec la même espèce, la quantité d'ARNm a été mesurée à midi (LD6), à minuit (LD18), et aux temps intermédiaires LD2 et LD14. Il y avait 131 transcrits avec des niveaux variables parmi les quatre temps, mais lorsqu'une douzaine de ces transcrits sélectionnés aléatoirement ont été testés par des analyses Northern, aucune différence n'a été détectée.

Dans l'ensemble, les niveaux de transcrits chez *Lingulodinium* semblent être constants au cours du cycle journalier en condition LD tout comme en condition LL. Ceci suggère que la génération de rythmie, contrairement à la situation chez d'autres eucaryotes, n'implique pas l'oscillation des quantités d'ARNm. Ceci est appuyé par l'observation que l'inhibition à 95%

de la transcription chez *Lingulodinium* ne perturbe pas les rythmes de la bioluminescence ou de la photosynthèse. Les auteurs (Roy, Beauchemin et al. 2014) ont suggéré la possible existence d'une boucle de rétroaction traduction/post-traduction avec une rythmicité de 24h permettant de réguler le cycle circadien (Figure 1.1). Cette hypothèse est cohérente avec le fait que les inhibiteurs des kinases et des phosphatases ont une influence sur les propriétés rythmiques de l'horloge (Comolli and Hastings 1999) (Comolli, Taylor et al. 1996) (Comolli, Taylor et al. 1994) (Comolli, Fagan et al. 2003). Il est à noter que la phosphorylation de plusieurs protéines qui peuvent être utilisées pour contrôler la traduction chez *Lingulodinium* et des protéines de liaison à l'ARN ont des sites de caséine kinase 2 (CK2) (Roy and Morse 2014). Ceci est important, car CK2 est impliquée dans la régulation de cycles circadiens chez les animaux, les végétaux et les champignons (Nakajima, Imai et al. 2005). Il a également été observé que certains inhibiteurs de la traduction (l'anisomycine, la puromycine et le cycloheximide) perturbent effectivement les rythmes de bioluminescence (Karakashian and Hastings 1963) (Taylor, Krasnow et al. 1982). L'effet de ces inhibiteurs semble être sur l'horloge plutôt que sur les rythmes seulement, car des expositions aux inhibiteurs peuvent changer le temps du rythme, mais pas l'amplitude du rythme.

Actuellement, une absence d'implication de la transcription dans la régulation des cycles circadiens chez *Lingulodinium* pourrait s'expliquer par la compaction du génome. Les génomes des dinoflagellés sont souvent gigantesques avec des centaines de milliards de paires de bases (Veldhuis, Cucci et al. 1997). Ce n'est qu'en 2015 que le premier génome de ces organismes unicellulaires a été enfin séquencé. Avec 1,18 gigabases le dinoflagellé symbiotique *Symbiodinium kawagutii* avait le génome le plus petit découvert à cette date, et donc était le plus accessible au séquençage (Lin, Cheng et al. 2015). Par comparaison, le génome de

*Lingulodinium* est plus proche de 250 gigabases (Hou and Lin 2009). Hormis sa taille particulièrement grande, le génome des dinoflagellés est aussi remarquable par sa forme de cristal liquide maintenu par des ARNs et des cations métalliques à la place d'une compaction classique médiée souvent majoritairement par les protéines histones (Chow, Yan et al. 2010) (Herzog and Soyer 1983). L'absence d'histones entraîne aussi l'absence des nucléosomes. Curieusement, des ARNm pour des histones ont été trouvés dans les transcriptomes des dinoflagellés, et ce fut aussi le cas pour les transcrits encodant les enzymes modificateurs des histones (Roy and Morse 2012). Les histones elles-mêmes ont été détectées chez les dinoflagellés, mais dans de très faibles quantités (Beauchemin and Morse 2018). À la place des histones se trouvent les Histone-like proteins (HLPs) (Chan and Wong 2007) (Wong, New et al. 2003) et des Dinoflagellate/Viral Nucleoproteins (DVNPs) (Gornik, Ford et al. 2012). Ces protéines basiques contribuent à la structuration et possiblement à la régulation de la structure de l'ADN (Sun, Liu et al. 2013) (Chudnovsky, Li et al. 2002) (Sala-Rovira, Geraud et al. 1991) (Sigeo 1983). Le rôle joué par le peu d'histones présentes n'est pas connu, mais elles sont probablement importantes étant donné la conservation de leurs séquences.

La structure très compacte de la chromatine pourrait créer un environnement peu propice à la transcription d'où une transcription constante et sans variation. Un faible taux de transcription pourrait être compensé par une longue durée de vie des ARNm tel qu'observé chez le dinoflagellé *Karenia* (Morey and van Dolah 2013). On ignore jusqu'à présent si ceci est aussi le cas chez *Lingulodinium*, car aucune étude n'a mesuré la durée de vie des différents ARNm.

La régulation de la transcription par l'horloge circadienne a été observée dans les cinq systèmes modèles malgré une différence dans le nombre de transcrits réglés (Li and Zhang

2015). Chez les cyanobactéries il y a une régulation de la transcription de pratiquement tous les gènes. Ceci est contrôlé par une horloge constituée principalement de trois gènes, kaiA, kaiB et kaiC. Les trois gènes sont essentiels, car l'inactivation de l'un d'eux inhibe l'horloge circadienne (Ishiura, Kutsuna et al. 1998). On a originellement suggéré que l'horloge chez *Synechococcus* exploite une boucle de transcription/traduction de l'expression du gène kaiC, dans laquelle la protéine KaiC inhibe sa propre transcription et entraîne des rythmes d'expression génomique. Les expériences subséquentes ont démontré l'importance de la phosphorylation de KaiC, et il est important de noter que non seulement l'état de phosphorylation de la protéine KaiC oscille de façon circadienne (Tomita, Nakajima et al. 2005), mais cette oscillation de l'état de phosphorylation de KaiC, en présence de seulement KaiA, KaiB et l'ATP, peut même se faire *in vitro* avec une période de 24 heures. En fait, KaiC a une activité ATPase, et les mutants affectant l'activité ATPase de KaiC ont des périodes rythmiques *in vitro* qui sont corrélées avec cette activité (Terauchi, Kitayama et al. 2007). Les trois protéines Kai interagissent entre elles et régulent les rythmes circadiens en contrôlant la condensation des chromosomes (Kageyama, Kondo et al. 2003) (Xu, Mori et al. 2003). Une étude de Hosokawa (Hosokawa, Hatakeyama et al. 2011) a démontré que l'oscillateur circadien post-traductionnel, basé sur les protéines Kai, régule l'activité transcriptionnelle dans l'obscurité, même si la transcription des gènes Kai est inhibée. Donc, en obscurité continue, l'expression de la plupart des gènes y compris la transcription/traduction de KaiBC est fortement supprimée, tandis que l'oscillateur à la base de la phosphorylation de KaiC est maintenu et renforce l'expression des gènes induits par l'obscurité et entraîne des oscillations amorties dans un sous-ensemble de gènes contrôlés par l'horloge sans nouvelle expression des gènes de l'horloge eux-mêmes. Ces résultats ont établi que le promoteur KaiBC n'est pas l'unique cible transcriptionnelle primaire de l'horloge

(Hosokawa, Hatakeyama et al. 2011). Dans les cyanobactéries *Synechococcus elongatus* PCC7942, entre 30% à 64% du transcriptome est exprimé circadiennement d'après les résultats des études par microarray (Ito, Mutsuda et al. 2009) (Vijayan, Zuzow et al. 2009). Il n'y a pas d'homologues des protéines Kai chez les dinoflagellés.

Chez *Neurospora*, la transcription circadienne de plusieurs gènes est contrôlée directement par des facteurs de transcription formant l'horloge. Dans le mécanisme de l'horloge, le complexe WCC agit comme facteur de transcription de façon positive et le complexe FRQ comme inhibiteur de sa propre transcription ce qui forme une boucle de rétroaction négative (Cheng, Yang et al. 2002) (Froehlich, Loros et al. 2003). WCC, comme facteur de transcription, agit également sur d'autres gènes, ceux dits contrôlés par l'horloge. Dans *Neurospora*, les micropuces à haute densité ont démontré qu'environ 20% à 25% du transcriptome peut être exprimé circadiennement (Correa, Lewis et al. 2003) (Dong, Tang et al. 2008). Les études d'ARN Seq confirment les résultats des microarrays (Hurley, Dasgupta et al. 2014).

Certains organismes ont plusieurs boucles de rétrocontrôle transcriptionnel. L'horloge de la plante *Arabidopsis*, par exemple, a trois boucles de rétroaction transcriptionnelle, appelées la boucle du matin, la boucle du soir et la boucle centrale (Alabadi, Oyama et al. 2001) (Matsushika, Makino et al. 2000) (Pokhilko, Mas et al. 2013) (Schaffer, Ramsay et al. 1998) (Wang and Tobin 1998). Chez *Arabidopsis*, entre 6% et 15% du transcriptome est régulé par l'horloge circadienne (Covington and Harmer 2007) (Edwards, Anderson et al. 2006) (Harmer, Hogenesch et al. 2000). Les ARNm cycliques peuvent avoir des pics dans toutes les phases (Harmer, Hogenesch et al. 2000) suggérant un contrôle complexe. Ceci semble être le cas aussi chez *Neurospora*. Un autre exemple avec plusieurs boucles de rétroaction est celui de la



drosophile, chez qui il y a deux boucles de rétroaction interconnectées qui permettent de réguler son cycle circadien (Allada, White et al. 1998) (Rutila, Suri et al. 1998) et plus d'un tiers de ces transcrits présentent une oscillation avec des phases diverses comme chez *Arabidopsis* (Ueda, Matsumoto et al. 2002).

### **1.3.2 Contrôle du temps de synthèse protéique et des changements du niveau protéique**

On a déjà discuté du rôle de la luciférase et de la LBP dans la bioluminescence chez *Lingulodinium* (partie 1.2.1). Ces protéines semblent être réglées de la même façon, car les temps de synthèse protéique et de la dégradation sont identiques. Par contre, il n'y a pas d'homologie de séquence dans les transcrits qui pourrait expliquer les changements coordonnés dans la vitesse de synthèse. Il n'y a pas non plus d'homologies de séquence protéique qui pourraient expliquer la dégradation coordonnée. Malgré cela, il est clair que l'abondance des deux protéines est contrôlée par l'horloge (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989).

La vitesse de synthèse élevée de LBP ne dépend pas des niveaux d'ARNm, car ceux-ci sont relativement constants au cours du cycle circadien et n'ont aucune variation corrélée avec l'augmentation de la synthèse de la protéine LBP (Morse, Milos et al. 1989). Ce contrôle est donc au niveau traductionnel. L'un des mécanismes suggérés serait que la synthèse de LBP était régulée par la liaison de facteurs d'inhibition dans la région 3'UTR des ARNm (Mittag, Lee et al. 1994). Malheureusement, d'autres études n'ont pu confirmer ce mécanisme dans la régulation de LBP, car aucun facteur d'inhibition se liant spécifiquement au 3'UTR de l'ARNm de LBP n'a été détecté (Lapointe and Morse 2008). Il n'y a pas non plus d'évidence permettant de croire que la traduction de l'ARN LBP soit affectée par des micro ARNs (Dagenais-

Bellefeuille, Beauchemin et al. 2017). Des mécanismes alternatifs peuvent impliquer l'activation ou l'inactivation de facteurs de traduction spécifiques à la synthèse des protéines et pourrait inclure le pic circadien de phosphorylation du facteur de traduction eEF-2 qui coïncide avec la traduction rythmique de certains ARNm spécifiques dont celui de l'enzyme glutathion S transférase chez *Neurospora* (Caster, Castillo et al. 2016).

À première vue il est curieux que les protéines impliquées dans la bioluminescence chez *Lingulodinium polyedra* subissent une synthèse et une dégradation quotidienne (Dunlap and Hastings 1981) (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989). Ceci est encore plus curieux étant donné que le recyclage de protéines de bioluminescence n'est pas commun à tous les dinoflagellés. Chez *Pyrocystis lunula*, le rythme dans la bioluminescence n'est pas corrélé avec des variations dans les niveaux de la luciférase. Possiblement en fonction de son environnement, soit riche ou pauvre en énergie ou en azote, le dinoflagellé adopte ou non la méthode du recyclage afin d'allouer des ressources à différentes protéines (Hastings 2013).

Dans un cycle de 24 heures, trois phases de synthèse protéique ont été observées suite à des marquages *in vivo* par l'incorporation d'une méthionine radioactive (Markovic, Roenneberg et al. 1996). La première en début de la phase de nuit correspondait à la transition entre le jour et la nuit, et comprend la luciférase et LBP et le début de la bioluminescence. La deuxième a eu lieu au milieu de la phase de nuit et comprend la Rubisco et coïncide avec la formation des pyrenoïdes. La troisième qui se passe à la fin de la phase de nuit correspondait à la transition entre la nuit et le jour et comprend la PCP impliquée dans les réactions lumineuses de la photosynthèse. La restriction de la synthèse de certaines protéines à des temps spécifiques de jour est peut-être une façon d'économiser l'énergie dans le cas des protéines dont la synthèse

n'a pas besoin d'être continue. Il est aussi possible que ce phénomène reflète le contrôle de l'horloge sur le comportement de la cellule, dans le cas où la quantité de protéine permettrait d'initier un processus biochimique donné plus rapidement.

### **1.3.3 Contrôle de l'activité enzymatique**

L'une des étapes préalables à la réalisation d'un phénomène cellulaire dans un moment circadien particulier est le contrôle de l'activité des enzymes, car ce sont normalement les enzymes qui catalysent l'ensemble des réactions biochimiques de la cellule. Le contrôle de l'activité enzymatique est donc d'une importance capitale pour l'organisme, car la régulation d'un rythme circadien peut passer par le contrôle de l'activité enzymatique.

Plusieurs autres études chez *Lingulodinium*, en plus de celles sur la bioluminescence, ont démontré que l'horloge peut contrôler le niveau de la quantité des enzymes. Par exemple, l'isocitrate déshydrogénase (NADP-ICDH), une enzyme du cycle de Krebs, montre des variations liées à l'horloge circadienne avec un pic en période nocturne (Akimoto, Kinumi et al. 2005). Ceci suggère que l'activité de l'enzyme NADP-ICDH va être régulée par le cycle circadien, contrairement à celles de plusieurs autres enzymes dans le cycle de Krebs. Cette étude a également confirmé que les quantités d'ARNm NADP-ICDH des cellules de *Lingulodinium* placées en condition LD 12:12 et LD 8:16 ou en lumière constante ne montrent aucune variation quantitative, tout comme les enzymes Citrate Synthase, IsoCitrate Lyase, et Malate Synthase. Ces résultats permettent de dire que le rythme cellulaire de l'enzyme NADP-ICDH est régulé par un contrôle circadien de la synthèse protéique au niveau traductionnel. Les auteurs ont proposé que le cycle de Krebs de *Lingulodinium* serait contrôlé par le cycle circadien via le

contrôle circadien de l'enzyme NADP-ICDH, ce qui peut constituer une base moléculaire pour le rythme circadien de respiration.

Une autre étude sur la glycéraldéhyde-3-phosphate déshydrogénase (GAPDH) des chloroplastes (Fagan, Morse et al. 1999) a démontré que la quantité protéique et l'activité enzymatique varient de façon concertée au cours de la journée. Encore une fois, les niveaux d'ARNm ne sont pas corrélés avec la synthèse, car la synthèse a lieu entre LD0 et LD4 tandis que les quantités d'ARNm sont constantes. Le lien entre la synthèse et l'activité enzymatique mène à dire que la synthèse protéique est responsable de l'augmentation de la quantité de la protéine (Fagan, Morse et al. 1999) et l'activité de cette enzyme dépend de sa quantité. La GAPDH analysée dans cette étude est une protéine qui est dirigée vers le chloroplaste, laissant penser que la quantité de la protéine peut influencer le rythme de fixation du CO<sub>2</sub>.

Tel que mentionné plus haut, Rubisco ne démontre aucune variation quantitative de la protéine durant le cycle circadien malgré les changements dans la vitesse de fixation du dioxyde de carbone (Nassoury, Fritz et al. 2001). À la place, les différences dans la vitesse de fixation du dioxyde de carbone sont dues aux différences dans la répartition de l'enzyme à l'intérieur des chloroplastes. Le mécanisme n'est pas connu à cette date, mais la séquestration de la Rubisco dans les pyrénoides est corrélée dans le temps avec la synthèse protéique de la Rubisco. De plus, lorsque le transport de Rubisco de l'appareil de Golgi vers les chloroplastes est bloqué, les pyrénoides ne se forment pas et la quantité de dioxyde de carbone fixée n'augmente pas (Nassoury, Wang et al. 2005). Toutes les protéines dirigées vers les chloroplastes des dinoflagellés avec la péridinine comme photopigment doivent passer par le Golgi car ces organelles sont entourées par trois membranes. La concordance entre la synthèse de la Rubisco et la formation des pyrénoides supporte l'hypothèse selon laquelle la redistribution de Rubisco

est à l'origine de l'accroissement de l'activité carboxylase responsable de la forte quantité de dioxyde de carbone fixée. Il n'est pas connu si c'est la Rubisco fraîchement synthétisée ou une autre protéine synthétisée au même moment que la Rubisco qui agit comme signal pour la redistribution de la protéine.

Les niveaux protéiques constants de Rubisco paraissent paradoxaux au vu du rythme de la vitesse de synthèse. La vitesse de synthèse de Rubisco est croissante en phase de nuit avec un sommet en fin de nuit tandis qu'en phase de jour cette vitesse tend à diminuer avec un minimum durant le moment transitoire entre la fin du jour et le début de nuit. Ce paradoxe peut être résolu de deux façons: d'une part il peut être supposé que les niveaux protéiques de Rubisco sont si élevés que les oscillations quantitatives dues à la variation de synthèse ne sont pas détectées, d'autre part on peut supposer que la dégradation contrebalance la nouvelle synthèse (Markovic, Roenneberg et al. 1996). Par contre, la vitesse de dégradation de la Rubisco, mesurée lorsque la synthèse a été inhibée par le cycloheximide, semble être très lente (Nassoury, Wang et al. 2005).

Dans une étude concernant les effets de l'irradiance circadienne sur l'expression des transcrits de PCP du dinoflagellé *Prorocentrum donghaiense*, les quantités d'ARNm de PCP avaient tendance à diminuer en phase sombre tandis que celles-ci augmentaient en phase lumineuse avec un début de croissance des quantités d'ARNm autour de LD2 (Shi, Li et al. 2018). Cette variation quotidienne d'expression des ARNm PCP serait en accord avec l'étude indiquant une synthèse de PCP en fin de phase nocturne chez *Lingulodinium* (Markovic, Roenneberg et al. 1996). Par contre, chez *Lingulodinium* il n'y avait aucune différence dans les quantités de l'ARNm PCP, et les quantités protéiques ne variaient pas au cours du cycle LD. Pour cette protéine, la formation des pyrénoides provoque un changement de la localisation dans

le chloroplaste, car PCP se trouve dans les thylacoïdes qui deviennent très espacés par la présence de la Rubisco dans le stroma (Nassoury, Fritz et al. 2001).

Les études d'ARN Seq n'ont pas réussi à démontrer les différences diurnes dans la quantité d'ARN pour les différents transcrits (Roy, Beauchemin et al. 2014). En conséquence, pour des protéines dont les quantités démontrent des variations, on doit avoir des changements dans la vitesse de synthèse ou la vitesse de dégradation. Certaines protéines, comme la LBP, démontrent les deux. Il est vrai que certaines protéines ne démontrent pas de différences dans leur quantité cellulaire, malgré qu'elles soient impliquées dans les rythmes par exemple la Rubisco. Par contre, cette protéine a une vitesse de synthèse cyclique et la conservation d'un mécanisme de régulation suggère qu'il soit important pour la cellule. Il semble donc important pour l'étude du comportement circadien ou diurne de regarder non seulement les protéines dont les niveaux cellulaires changent, mais aussi des protéines dont la vitesse de synthèse change.

#### **1.4 Un mécanisme possible d'une horloge circadien chez *Lingulodinium***

Le système circadien a deux parties différentes, les rythmes que l'on détecte dans la biochimie, la physiologie ou le comportement, et l'engrenage moléculaire qui génère les signaux de synchronisation. La discussion plus haut a été concentrée sur l'expression génique impliquée dans différents rythmes circadiens chez *Lingulodinium*, et en particulier le fait que cette espèce n'utilise pas de contrôles transcriptionnels pour contrôler l'expression génique. Effectivement, la quantité de transcrits de chaque gène présent dans le transcriptome (presque 75 000 séquences) reste stable pendant un cycle LD ou LL (Roy, Beauchemin et al. 2014) malgré le fait que certains transcrits soient traduits à des temps circadiens différents (Milos, Morse et al.

1990). Ces observations suggèrent que non seulement les ARNm codants des protéines utilisées dans des rythmes biologiques mais aussi ceux des protéines impliquées dans l'horloge générant les rythmes circadiens ont des niveaux constants. Dans ce cas, le mécanisme de l'horloge ne peut pas utiliser un cycle d'accumulation et de dégradation des transcrits comme on l'observe chez les autres eucaryotes. Cette affirmation est appuyée par le fait qu'une inhibition de la transcription provoquée chimiquement ne permet d'arrêter ni les rythmes de photosynthèse ni ceux de bioluminescence. En fait le pic de ces deux rythmes se voit aux temps normaux, et seule une variation dans l'amplitude de ces deux phénomènes est observée (Roy, Beauchemin et al. 2014).

Si la transcription ne fait pas partie du mécanisme fondamental de régulation de l'horloge circadienne chez *Lingulodinium*, on pourrait donc se demander quelles autres étapes de l'expression des gènes pourraient être impliquées. Une des hypothèses que l'on peut avancer concernant le mécanisme de l'horloge chez *Lingulodinium* est qu'une régulation traductionnelle et post-traductionnelle des gènes de l'horloge remplacerait une régulation transcriptionnelle et post-transcriptionnelle. On a dans les deux cas deux processus différents qui se suivent dans le temps pour un gène donné et qui pourraient être contrôlés indépendamment l'un de l'autre. Les études antérieures ont démontré que la répression de la synthèse protéique par des inhibiteurs chimiques perturbe le déroulement du cycle circadien chez *Lingulodinium*, soit en reculant soit en avançant la phase du rythme de bioluminescence en fonction du moment de l'inhibition (Taylor, Dunlap et al. 1982) (Dunlap, Taylor et al. 1980). D'autres études ont également démontré que les inhibiteurs de protéines kinases et de protéines phosphatases peuvent influencer la période rythmique de la bioluminescence (Comolli, Taylor et al. 1996) (Comolli, Taylor et al. 1994) (Comolli, Fagan et al. 2003) il y a donc des évidences expérimentales

attestant l'implication d'une régulation traductionnelle et post-traductionnelle dans le mécanisme de l'horloge moléculaire chez *Lingulodinium*.

L'hypothèse selon laquelle la régulation de l'horloge circadienne implique un contrôle traductionnel suggère qu'en plus des ARNm dont la traduction circadienne contribue aux rythmes, il y aura aussi des ARNm dont la traduction est régulée et qui sont des composants de l'horloge circadienne elle-même. Pour générer un catalogue de tous les ARNm dont la traduction est régulée, la technique récemment développée du profilage ribosomique pourrait être utile. Cette technique permet de séquencer à haut débit les ARNm en cours de traduction par les ribosomes (Ingolia, Brar et al. 2012), et d'estimer les variations dans la vitesse de synthèse de différents transcrits par un comptage du nombre de séquences associées aux ribosomes aux temps différents. Nous espérons ainsi identifier l'ensemble des gènes dont la traduction est restreinte à différents moments du cycle circadien. L'utilisation de cette technique dans les études concernant les animaux et les plantes indique qu'elle permet l'identification des gènes traduits pendant le cycle circadien (Janich, Arpat et al. 2015) (Jang, Lahens et al. 2015) (Juntawong, Girke et al. 2014). En parallèle, une analyse du protéome sera effectuée à différents moments du cycle, afin de voir si les variations dans la vitesse de traduction prédites par le profilage des ribosomes ont un impact sur la quantité des protéines. Nous anticipons que les changements du protéome dépendront des vitesses de synthèse et de dégradation, et donc certaines des protéines dont la synthèse prédit de varier par profilage ribosomique devront changer en abondance. Les études de protéome menées chez la souris ont montré qu'une grande partie du protéome hépatique changeait au cours du cycle circadien sans qu'il y ait variation d'abondance des transcrits pour certains gènes. Cette observation confirme chez une espèce



modèle que les variations quantitatives et qualitatives du protéome peuvent se faire sans variation d'abondance des transcrits (Robles, Cox et al. 2014) (Mauvoisin, Wang et al. 2014). Afin de tester la procédure, il va être nécessaire de vérifier les trois temps dans le cycle circadien durant lesquels la Rubisco, la PCP et la LBP sont spécifiquement synthétisées (Markovic, Roenneberg et al. 1996, Lapointe and Morse 2008) (Milos, Morse et al. 1990). Ces temps sont LD6, où la synthèse de la protéine PCP atteint le sommet, LD14 quand la protéine LBP est synthétisée au maximum, et LD20 lorsque la Rubisco est synthétisée de façon maximale. En plus de ces trois protéines, qui serviront de contrôle, cette expérience devrait aussi identifier d'autres protéines dont la synthèse est réglée par l'horloge et qui sont peut-être impliquées dans les mêmes rythmes que les trois protéines mentionnées ci-haut. On peut donc estimer le nombre de gènes dont l'expression est réglée au niveau traductionnelle. À partir de ces résultats, nous évaluerons si nous élargissons l'étude à d'autres temps voire à l'ensemble du cycle.

La deuxième partie de l'hypothèse, selon laquelle le mécanisme du cycle circadien chez *Lingulodinium* impliquerait des modifications post-traductionnelles, pourrait être abordée par une analyse du phosphoprotéome. La phosphorylation est la principale modification post-traductionnelle dans les cellules et les phosphoprotéines peuvent être sélectivement enrichies à partir des extraits protéiques. Ces expériences vont utiliser les techniques à haut débit pour connaître les différentes protéines phosphorylées à différents moments du cycle circadien. Pour ce faire, il faudra purifier les phosphoprotéines à différents moments du cycle circadien puis les identifier par spectrométrie de masse. Une étude a déjà été menée chez *Lingulodinium* à « petite échelle » (Roy and Morse 2014). Ici, il faudra amplifier et élargir l'étude en choisissant des temps d'étude recouvrant l'ensemble du cycle circadien chez *Lingulodinium* pour connaître au complet les variations quantitatives et qualitatives concernant les protéines phosphorylées qui

pourraient être impliquées dans les rythmes circadiens ou dans l'horloge. Chez la souris, une étude sur les phosphoprotéines a conclu que près d'un quart des sites de phosphorylation montraient des rythmes circadiens de phosphorylation et que ces variations cycliques de phosphorylation influençaient fortement les réseaux de signalisation et le métabolisme hépatique (Robles, Humphrey et al. 2017). De façon complémentaire il serait intéressant de voir l'impact des inhibiteurs de kinases et de phosphatases sur le phosphoprotéome, car ces inhibiteurs ont la capacité de perturber le cycle circadien (Comolli and Hastings 1999) (Comolli, Taylor et al. 1996) (Comolli, Taylor et al. 1994) (Comolli, Fagan et al. 2003). Ceci permettra de connaître les phosphoprotéines clés du rythme circadien en dehors des inhibiteurs de CK1 et CK2 qui sont impliquées dans les rythmes circadiens (Nakajima, Imai et al. 2005).

La modification post-traductionnelle par phosphorylation peut être impliquée dans la génération des rythmes circadiens, car tel que mentionné plus haut, les inhibiteurs des protéines kinases ou des protéines phosphatases perturbent le chronométrage du rythme de bioluminescence chez *Lingulodinium* (Comolli and Hastings 1999) (Comolli, Taylor et al. 1996) (Comolli, Taylor et al. 1994). En fait, des études préliminaires du phosphoprotéome de cette espèce à deux moments du cycle circadien (midi et minuit) ont montré des différences suggérant que la phosphorylation des protéines est effectivement différente pendant le cycle circadien du dinoflagellé *Lingulodinium* (Liu, Lo et al. 2012). Dans cette étude, le phosphoprotéome a été analysé par deux méthodes, l'une utilisant des gels 2D couplés avec une coloration de ProQ Diamond (Agrawal and Thelen 2006) et l'autre utilisant une purification par affinité des phosphopeptides par chromatographie au dioxyde de titane (Larsen, Thingholm et al. 2005). Par gel 2D, deux protéines avec un degré de phosphorylation différent ont été identifiées: une protéine du complexe de récolte de la lumière (LHC) et Rad24. Rad24, une protéine impliquée

dans la réplication de l'ADN, est phosphorylée plus fortement la nuit, ce qui est conforme à l'observation selon laquelle la phase S commence à minuit (Dagenais-Bellefeuille, Bertomeu et al. 2008). La première protéine identifiée, LHC, est aussi connue comme étant une phosphoprotéine dont la phosphorylation dépend de la lumière dans les plantes et les algues vertes (Bonardi, Pesaresi et al. 2005). La phosphorylation entraîne une modification de l'affinité de cette protéine pour les photosystèmes (PS) I et II. La forme phosphorylée de LHCII s'associe préférentiellement avec PSI, tandis que la forme déphosphorylée s'associe à PSII (Allen and Forsberg 2001). La phosphorylation du LHC permet d'équilibrer la quantité de lumière entrant dans chacun des photosystèmes (Rochaix 2007). Par chromatographie au dioxyde de titane, six protéines phosphorylées différemment ont été identifiées, une polycétide synthase, un transporteur non caractérisé, un domaine LIM (liaison d'actine) et trois protéines de domaine de liaison d'ARN (Liu, Lo et al. 2012). Il n'y avait aucune protéine commune aux deux méthodes.

Dans cette étude, il y avait de nombreux peptides phosphorylés chez *Lingulodinium* avec la même intensité à LD6 et à LD18. Par contre, il est toujours possible que certains parmi eux varient dans leur degré de phosphorylation à des moments autres que ceux étudiés. En élargissant les points d'étude, au cours du cycle circadien en condition LD, on pourrait éclairer les zones d'ombre entre ces temps et commencer à dresser les mécanismes de régulation des rythmes circadiens chez *Lingulodinium* liés à la phosphorylation. Récemment, par une étude du protéome et du phosphoprotéome chez la souris il a été découvert qu'il y a plusieurs rythmes de phosphorylation dans l'hippocampe. Lors des analyses des 3052 protéines totales et des 2688 phosphosites, 1368 protéines ont été quantifiées avec précision, et 1,7% des protéines et 5,2% des phosphorylations présentaient des profils dépendant du temps. La majorité des phosphopeptides circadiens présentaient des fluctuations du milieu à la fin de la journée sans

rythme d'abondance protéique (Chiang, Xu et al. 2017). Dans une autre étude, Cao et ses collègues ont découvert que la phosphorylation du facteur d'initiation de la traduction eucaryote eIF4E augmente pendant la phase du jour. La forme phosphorylée de eIF4E favorise la traduction des ARNm encodant les protéines de l'horloge PER 1 et PER 2, ce qui facilite la réinitialisation de l'horloge circadienne (Cao, Gkogkas et al. 2015).

Finalement on doit mentionner que le système circadien chez *Lingulodinium* est compliqué par la possibilité que la cellule puisse avoir plus qu'une horloge circadienne. Cette possibilité a été envisagée pour expliquer l'observation selon laquelle deux rythmes différents peuvent avoir dans certaines conditions expérimentales des périodes rythmiques différentes (Roenneberg and Morse 1993).

## **1.5 Objectifs**

Les objectifs généraux élaborés au début de mon travail étaient de regarder, à travers le cycle d'un jour de *Lingulodinium* sous condition LD12: 12 (Figure 1.2), les protéines dont les niveaux cellulaires varient, les protéines dont la vitesse de synthèse varie, et les protéines dont le niveau de phosphorylation change. Le but ultime était d'identifier des composants protéiques qui pourraient nous aider à comprendre la base biochimique des rythmes diurnes et aussi d'identifier des candidats pour les composants de l'horloge elle-même. La technique de séquençage par LC-MS/MS a été sélectionnée pour les études sur les protéines et les phosphoprotéines, et la technique de profilage des ribosomes a été sélectionnée pour regarder la vitesse de synthèse.

Nous espérons non seulement mieux comprendre les rythmes diurnes chez cette espèce mais aussi commencer à comprendre l'horloge elle-même. L'ensemble de ces résultats permettra aussi de comprendre les rythmes biologiques des autres organismes.

### **1.5.1 Le séquençage haut débit de protéines par spectroscopie de masse**

Cette technique vise l'identification et la quantification de plusieurs milliers de protéines et les détails de la méthodologie se trouvent dans l'article concerné. Ces études permettent de déterminer quantitativement et qualitativement les variations des protéines présentes aux différents temps du cycle circadien. Pour l'analyse protéomique, les échantillons ont été récoltés à travers un cycle diurne et pourraient inclure des protéines réglées par la lumière ou par l'horloge circadienne.

Après une digestion des protéines par l'enzyme trypsine, les peptides tryptiques sont analysés par LC-MS/MS. Cette technique permet non seulement de déterminer la quantité de chaque peptide, mais aussi la séquence du peptide et donc sa protéine d'origine. L'examen, aux différents temps circadiens, de plusieurs milliers de protéines peut permettre de découvrir de nouvelles protéines impliquées dans les rythmes ou dans l'horloge circadienne chez *Lingulodinium*. Une étude similaire a été menée chez la souris, et les analyses du protéome du foie de souris ont montré que parmi les 3000 protéines quantifiées avec précision pendant deux cycles consécutifs, 6% avaient des oscillations circadiennes avec une phase d'expression définie. Curieusement, près de 25% des protéines qui variaient de façon cyclique n'avaient aucune variation quantitative de transcrits (Robles, Cox et al. 2014) dans cette étude. Il n'y avait pas de transcrits variables pour les 13 protéines connues pour avoir des niveaux variants chez *Lingulodinium*.

### **1.5.2 Mesures de la synthèse protéique à grande échelle**

Le profilage ribosomique, appelé « *ribosome profiling* » en anglais, est une technique qui permet de voir les ARNm qui sont en train d'être synthétisés. Pour ce faire, les ARNases sont ajoutées à des extraits cellulaires afin de digérer tout l'ARNm sauf la partie protégée par la présence d'un ribosome qui est en train de le traduire (Ingolia, Brar et al. 2012). Ceci produit des fragments d'ARN d'environ trente nucléotides qui sont purifiés et rétro-transcrits en ADN complémentaire pour séquençage à haut débit (Ingolia, Brar et al. 2012). Cette technique permet donc d'estimer la vitesse de synthèse des différentes protéines par le nombre de fois qu'une séquence d'un ARNm donné se trouve dans la fraction protégée par les ribosomes. Il ne peut pas être utilisé pour comparer la vitesse de synthèse de deux séquences différents à un moment donné, mais peut permettre de connaître les changements de vitesse de synthèse de chaque ARNm aux différents moments du cycle circadien. Dans plusieurs versions de la technique, la traduction a été inhibée par le cycloheximide qui permet de fixer les ribosomes sur les ARNm en cours de traduction.

### **1.5.3 Le séquençage haut débit de phosphoprotéines par spectroscopie de masse**

La dernière étude planifiée consistait à chercher l'influence de la phosphorylation pendant le cycle circadien en analysant le phosphoprotéome. Ces expériences nous permettront de comprendre éventuellement quelles protéines subissent des variations dans leur état de phosphorylation.

Il est important de noter ici que ce travail a été fait en collaboration avec un groupe de protéomique à Hong Kong. Les échantillons ont été préparés et envoyés, mais suite aux

nombreuses difficultés, reliées à la situation politique et ensuite sanitaire, seulement un des trois replicats a pu être séquencé. Cette partie de l'étude ne fait donc pas partie de cette thèse.

#### Tableau et figures

Tableau 1. Variations quotidiennes des concentrations de luciférase (en µg/mL) dans les extraits protéiques de *Lingulodinium*, déterminées par ELISA. Inspiré de (Johnson, Roeber et al. 1984).

Expériences	Concentrations en luciférase jour	Concentrations en Luciférase nuit	Ratios nuit-jour en Luciférase
LD-1	0.4	2,7	6,8
LD-2	0.06	1	16,7
LL-1	2.0	15,3	7,7
LL-2	0.8	7,8	9,8
LL-3	1.3	9,6	7,4

Figure 1. 1. Modèle simplifié de boucle de rétrocontrôle.

(A) régulation par les protéines synthétisées, Inspiré de (Lakin-Thomas 2006). (B) régulation par modifications posttraductionnelles, inspiré par (Roy, Beauchemin et al. 2014).

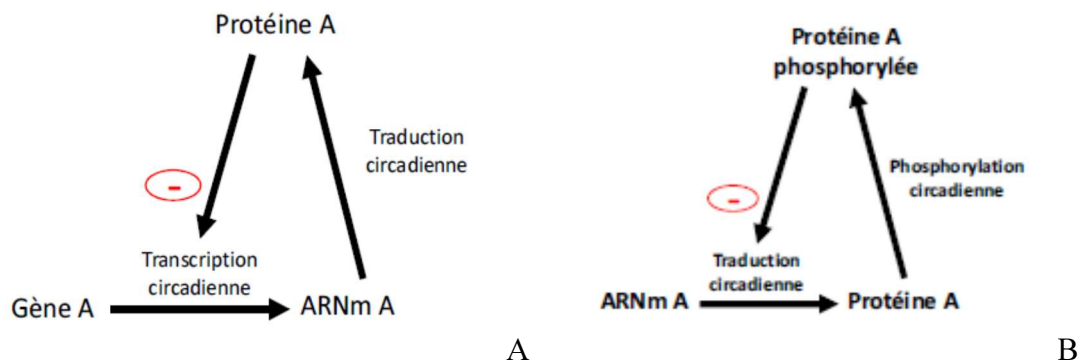
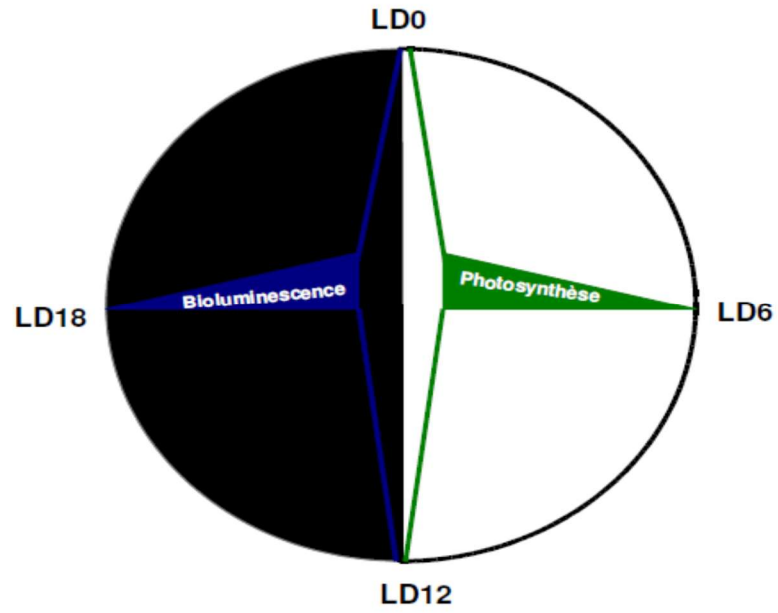




Figure 1. 2. Cycle biologique de *Lingulodinium* sous un cycle LD12: 12 avec les heures des pics des deux principaux phénomènes diurne et nocturne.



**Chapitre 2: Label-free MS/MS analyses of the dinoflagellate *Lingulodinium* identifies rhythmic proteins facilitating adaptation to a diurnal LD cycle**

**Label-free MS/MS analyses of the dinoflagellate *Lingulodinium* identifies rhythmic proteins facilitating adaptation to a diurnal LD cycle**

Carl Bowazolo, Sirius P.K Tse, Mathieu Beauchemin, Samuel C-L. Lo, Jean Rivoal and David Morse

Publié dans Science of the Total Environment 704 (2020) 135430.

Contribution des auteurs:

Avec M. Beauchemin, j'ai préparé les échantillons. Avec D. Morse j'ai participé à l'analyse des résultats. S. Tse et S. Lo ont fait des analyses LC-MS/MS. J'ai fait certaines figures et rédigé la première version de l'article. J'ai aussi effectué les expériences ayant servi de contrôle avec l'aide de Mathieu et d'un stagiaire. David Morse a fait la figure 4 et contribué à la rédaction de l'article. Jean Rivoal a contribué au matériel et méthodes nécessaire pour les mesures du TPI.

## 2.1. Abstract

Protein levels were assessed in the dinoflagellate *Lingulodinium polyedra* over the course of a diurnal cycle using a label-free LC-MS/MS approach. Roughly 1700 proteins were quantitated in a triplicate dataset over a daily period, and 13 were found to show significant rhythmic changes. Included among the proteins found to be most abundant at night were the two bioluminescence proteins, luciferase and luciferin binding protein, as well as a proliferating cell nuclear protein involved in the nightly DNA replication. Aconitase and a pyrophosphate fructose-6-phosphate-1-phosphotransferase were also found to be more abundant at night, suggestive of an increased ability to generate ATP by glucose catabolism when photosynthesis does not occur. Among the proteins more abundant during the day were found a 2-epi-5-epi-valiolone synthase, potentially involved in synthesis of mycosporin-like amino acids that can act as a “microbial sunscreen”, and an enzyme synthesizing vitamin B6 which is known to protect against oxidative stress. A lactate oxidoreductase was also found to be more abundant during the day, perhaps to counteract the pH changes due to carbon fixation by facilitating conversion of pyruvate to lactate. This unbiased proteomic approach reveals novel insights into the daily metabolic changes of this dinoflagellate. Furthermore, the observation that only a limited number of proteins vary support a model where metabolic flux through pathways can be controlled by variations in a select few, possibly rate limiting, steps. Data are available via ProteomeXchange with identifier PXD006994.

## 2.2. Introduction

The unicellular dinoflagellate *Lingulodinium polyedra* has been a model for the study of circadian rhythms since the late 1950s when a bioluminescence rhythm persisting in constant conditions was discovered (Hastings and Sweeney 1958). This seminal discovery was followed by the observations that cell division (Sweeney and Hastings 1958) and photosynthesis (Sweeney 1960) were also daily rhythms that persisted without environmental cues. These biological rhythms, now called circadian rhythms, are found in all walks of life, and have been extensively studied in cyanobacteria, fungi, plants and animals. Since the initial rhythms were discovered in *Lingulodinium*, much further work has attempted to understand the biochemical changes that allow the cells to specialise for different tasks at different times during the day. These studies have uncovered several strategies used by cells to produce daily changes in their behavior or physiology. For example, changes in protein levels can allow the cell to develop new abilities, as evidenced by the correlation between the nightly bioluminescence rhythm and levels of the two proteins, luciferase and luciferin binding protein, required for the bioluminescence reaction (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989). In addition, the subcellular location of a protein can also influence its catalytic ability, as evidenced by the rhythmic changes in Rubisco distribution within individual chloroplasts (Nassoury, Fritz et al. 2001).

These results have led to a number of different experimental approaches aimed at identifying proteins likely to show rhythmic changes. One sensitive approach has used *in vivo* labeling with radiolabelled amino acids to identify proteins whose synthesis rates change over the daily cycle, and it was found that synthesis of many proteins is indeed under circadian control (Markovic, Roenneberg et al. 1996, Fagan, Morse et al. 1999, Le, Jovine et al. 2001, Hollnagel, Pinto et al. 2002). However, it was also found that protein levels did not change

appreciably despite the changes in synthesis rates. This was suggested to mean that increased synthesis rates provided only a small change in total protein levels, and that marked changes in protein levels may require high turnover rates for the proteins involved.

Recently, advances in protein identification by mass spectrometry (MS) have made high throughput comparative proteomics more accessible for research (Morse, Tse et al. 2018), and circadian (or diurnal) MS-based proteomic studies have been carried out on many of the systems whose molecular clocks have been characterized. For example, in the cyanobacteria *Synechococcus*, 5% of the 1537 proteins quantified were found to cycle (Guerreiro, Benevento et al. 2014). A similar number of cycling proteins (2.5% out of a total of 653) were found to cycle in the green alga *Ostreococcus*, an organism with a green plant-like clock (Le Bihan, Martin et al. 2011). In mouse liver, between 3 to 6% of the proteins detected varied over the circadian cycle (Mauvoisin, Wang et al. 2014, Robles, Cox et al. 2014), and *Drosophila* heads had 1.4% diurnally varying proteins out of a total of 3154 (Du, Zhang et al. 2018). Fungi, curiously, had a far greater proportion, with 1273 clock regulated proteins found among the 4742 *Neurospora* proteins detected (27%) (Hurley, Jankowski et al. 2018). Several of these studies have also compared changes in transcript abundance with the circadian proteome, and this has led to the surprising conclusion that not all cycling transcripts are correlated with cycling proteins. For example, between one fifth and one half the cycling proteins had constant transcript levels (Mauvoisin, Wang et al. 2014, Robles, Cox et al. 2014). Clearly, effects on cell behavior are more likely to be predicted when proteomes rather than transcriptomes are examined.

In *Lingulodinium*, previous work used two-dimensional (2D) gel electrophoresis-based methods to examine protein changes in *Lingulodinium* over a daily cycle. 2D gels have a relatively cheap set-up platform and a facile sample preparation, and in *Lingulodinium* allowed 900 spots to be visualised (Akimoto, Kinumi et al. 2005) (Akimoto, Wu et al. 2004). However, many of the spots identified had lower molecular weights than they should and may thus represent partially degraded proteins. Here we have used LC-MS/MS to assess protein expression in crude protein extracts from *Lingulodinium* over a diurnal cycle. This technique is more powerful than the two-D gel approach, as we have quantitated roughly 1700 proteins, almost all of which are identified by 3 or more peptides. We found that 13 (0.8%) display significant changes in abundance over the daily cycle. We propose that these proteins may provide insight into more general

biochemical changes within dinoflagellates that underlie previously unsuspected daily rhythms adapting these organisms to daily changes in light and dark.

### 2.3. Results

In order to identify *Lingulodinium* proteins whose levels vary reproducibly over the course of the daily light:dark cycle, three independent time series were used to prepare total protein extracts every three hours during the light period, and every two hours during the dark period. The shorter interval during the night phase was intended to accommodate the larger number of different processes that vary during this period, including both S and M phases of the cell cycle (Dagenais-Bellefeuille, Bertomeu et al. 2008), the synthesis and destruction of scintillons (Fritz, Morse et al. 1990), and the formation of pyrenoids in the plastid (Nassoury, Fritz et al. 2001). Ten samples were taken per series (thirty samples in all), and all were digested with trypsin and the peptides analyzed by LC-MS/MS. Progenesis peak picking identified 18759 peptides, and these were identified by comparison to a transcriptome database (Beauchemin, Roy et al. 2012) using Mascot server. After entering these identifications into Progenesis, the peptides were found to be derived from 1701 quantifiable proteins, with 99.5% of the proteins identified by three or more peptides. Of these, thirteen proteins were found to have significantly different levels over the diurnal cycle ( $\geq 1.5$  fold changes;  $p \leq 0.05$ ). Average intensities for the 13 proteins spanned a range of just under 3 orders of magnitude (Fig. 2.1).

The identity of the different proteins encoded by the transcriptome sequences was determined by BLASTx searches against the non-redundant protein database (Table 2.1). The confidence in the identifications, as estimated from the e-values for the best blast hits, is generally good, with the poorest hit having an e-value of  $6 \times 10^{-41}$ . As expected, the two bioluminescence proteins whose levels are known to vary over a circadian cycle in *Lingulodinium*, luciferase (Johnson, Roeber et al. 1984) and luciferin binding protein (LBP) (Morse, Milos et al. 1989), were among the proteins varying over the diurnal cycle. Furthermore, we also recovered proliferating cell nuclear antigen (PCNA), whose levels have been shown to vary as a function of the cell cycle in another dinoflagellate (Brunelle and van Dolah 2011). PCNA is implicated in DNA replication which begins at the middle of the night phase (Dagenais-Bellefeuille, Bertomeu et

al. 2008). Also as expected, the proteins glyceraldehyde-3-phosphate dehydrogenase (GAPDH) (Fagan, Morse et al. 1999), Rubisco (Nassoury, Fritz et al. 2001) and peridinin chlorophyll *a* protein (PCP) (Hollnagel, Pinto et al. 2002) did not vary. Somewhat surprisingly, our variable proteins also did not include the glucose metabolic enzymes pyruvate kinase, isocitrate dehydrogenase and succinate dehydrogenase, three proteins previously reported to be variable over the daily cycle during a 2D gel evaluation of proteins in *Lingulodinium* (Akimoto, Wu et al. 2004). Instead, we recovered five different proteins potentially implicated in carbon metabolism (pyrophosphate fructose-6-phosphate-1-phosphotransferase, triosephosphate isomerase, aconitase and lactate oxidoreductase).

A number of unexpected proteins showing diurnal variations were found. These include p43, a cell surface protein of unknown function (Bertomeu, Hastings et al. 2003); translation elongation factor 3; Sar1, a monomeric GTPase regulating vesicle formation at the ER membrane (Nakano and Muramatsu 1989) and PDX1, a protein involved in pyridoxal phosphate (vitamin B6) biosynthesis (Gengenbacher, Fitzpatrick et al. 2006). An acetyl CoA oxalyl CoenzymeA transferase and a bacterial 2-epi-5-epi-valiolone synthase thought to be involved in mycosporine-like amino acid (“microbial sunscreen”, (Balskus and Walsh 2010)) were also found.

To assess the relative timing of the different proteins, protein abundance values at each time point were first used to construct a hierarchically clustered heat map (Fig. 2.2). This analysis places the different proteins into two main clusters, one whose levels are highest during the day and the other whose levels are highest during the night. The variations in the day phase and night phase groups were then plotted separately to evaluate the relative error in the measurements. Data were also plotted without normalization to emphasize the wide range in protein levels (Fig. 2.3). To allow variations to be seen for all proteins, proteins with different relative abundance were plotted separately.

To provide a more global picture of the different diurnally varying proteins, the proteins were overlaid on top of a cell schematic illustrating both their subcellular location and the phase of their peak occurrence (Fig. 2.4), with red corresponding to those proteins more abundant during the day phase and blue to night phase proteins. For the most part, the different proteins are disparate, and appear to function in different pathways. The three exceptions to this are for bioluminescence, photoprotection and carbon metabolism. In the case of bioluminescence, LBP



and luciferase levels change in concert, and are highest at night when both bioluminescence capacity and the number of scintillons is greatest (Fritz, Morse et al. 1990). In the case of photoprotection, the enzymes PDX1 and EVS are involved in synthesizing vitamin B6 and mycosporine-like amino acids, respectively, both of which have demonstrated ability to protect cells against damage to high light conditions. Lastly, with respect to carbon metabolism, five different proteins show diurnal variations. Proteins more abundant during the night include cytoplasmic pyrophosphate fructose-1,6-phosphate phosphotransferase (PFK), which catalyzes the formation of fructose-1,6-bisphosphate in the glycolytic pathway, and mitochondrial aconitase which catalyzes the biosynthesis of isocitrate in the TCA cycle. The observation that these proteins have higher levels at night, when there is no photosynthetic ATP production, would be consistent with increased production of ATP by glycolysis and the TCA cycle at this time. The three other enzymes involved in carbon metabolism are more abundant during the day suggesting they are not involved in ATP generation. One, lactate oxidoreductase, may act to increase lactate synthesis to counteract the increased cellular pH caused by carbon fixation (Mackenzie and Morse 2011). Another, triosephosphate isomerase (TPI), which controls the balance between glyceraldehyde-3-phosphate (GAP) and dihydroxyacetone phosphate (DHAP) produced during glycolysis, may have higher levels during the day to ensure more DHAP is available for glycerol-3-phosphate and triglyceride synthesis. Lastly, increased day time levels of acetyl-CoA : oxalate CoA transferase (CoAT) may be involved in catabolism of oxalate, a metabolite potentially produced by photorespiration.

## **2.4. Discussion**

### *Night phase proteins involved in bioluminescence and DNA replication*

LBP and luciferase are the two most abundant proteins whose levels change over the daily cycle. LBP levels can be as high as 1% the total soluble protein (Morse, Pappenheimer et al. 1989), and on SDS gels, LBP appears roughly twice as abundant as luciferase (Desjardins

and Morse 1993). This agrees remarkably well with the protein levels measured here by MS/MS. In fact, LBP and luciferase were expected to be among the diurnally variable proteins, as circadian variation in the levels of both have been previously documented (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989). LBP was also found as a protein more abundant during night phase during 2D-gel based proteomic analysis of the diurnal cycle (Akimoto, Wu et al. 2004), although luciferase was not found in this study.

PCNA is also among the proteins whose levels were expected to change over the diurnal cycle. PCNA has been described as a sliding clamp implicated in DNA replication, and it is thought that PCNA wraps around the end of the double-stranded DNA molecule and slides along to the polymerase  $\delta$  where it stimulates the processivity necessary for DNA synthesis (Prosperi 1997). High levels have been reported during S-phase in mammalian cells (Bravo and Celis 1980) and the phytoplankton *Dunaliella tertiolecta* and *Isochrysis galbana* (Lin and Carpenter 1998), consistent with this proposed role. Previous work in the dinoflagellate *Prorocentrum donghaiense* has shown that this protein is roughly 2-3 times more abundant during the nightly S-phase (Wang, Zhang et al. 2013). Surprisingly, though, in the dinoflagellate *Karenia brevis*, PCNA levels are actually highest during the day phase, and the nightly DNA replication is correlated with the presence of a higher molecular weight form of PCNA (~ 8 kDa larger) (Brunelle and van Dolah 2011). It is not known if the PCNA identified in our samples has been post-translationally modified, as additional purification procedure and mass spectrometry settings were required for post-translation modification (PTM) studies, so they are not included in protein identity searches in Mascot server. In any event, it seems reasonable to assume that either native PCNA or a post-translationally modified form would be more abundant during the S-phase, which occurs during the night phase in *Lingulodinium* (Dagenais-Bellefeuille, Bertomeu et al. 2008).

#### *Day phase proteins with a photoprotective role*

The enzyme 2-epi-5-epi-valiolone synthase (EVS) catalyses the cyclization of sedoheptulose-7-phosphate to 2-epi-5-epi-valiolone, and in *Streptomyces hygroscopicus* is involved in synthesis of the antifungal validamycin (Yu, Bai et al. 2005). In cyanobacteria, EVS is part of a pathway that makes mycosporine-like amino acids (MAA), which act as small molecule sunscreens protecting cells from the harmful effects of UV radiation (Balskus and

Walsh 2010). The gene is found in several species of dinoflagellates, coral and haptophytes, with phylogenetic analyses suggesting lateral gene transfer to the coral hosts of symbiotic dinoflagellates (Lin, Cheng et al. 2015). Studies on the dinoflagellate *Scrippsiella sweeneyae* have shown that MAA do protect photosystem II against damage caused by UV radiation (Taira and Taguchi 2017). The increased EVS protein abundance in day phase cells would thus be consistent with an increased requirement for UV protection during the day.

The enzyme PDX1 was also found to be more abundant during the day. This enzyme catalyzes the biosynthesis of pyridoxal 5-phosphate (vitamin B6), an important cofactor for many enzymes, from ribose 5-phosphate, glyceraldehyde-3-phosphate and glutamine in plants and bacteria (Ehrenshaft, Bilski et al. 1999, Raschke, Amrhein et al. 2005). In plants, vitamin B6 has an important antioxidant role, as a *pdx* mutant plant has lower levels of the PSII D1 protein when grown in high light unless the plants are also supplied with exogenous vitamin B6 (Titiz, Tambasco-Studart et al. 2006). Furthermore, plants engineered to overexpress PDX1 are more resistant to oxidative stress and photoinhibition (Raschke, Boycheva et al. 2011).

#### *Proteins involved in carbon metabolism*

The enzyme pyrophosphate fructose-6-phosphate-1-phosphotransferase (PFK) catalyzes an alternative biosynthetic route to the synthesis of fructose-1,6-bisphosphate that is also catalyzed by the glycolytic enzyme phosphofructokinase (PFK). Interestingly, PFK was previously thought to be the rate limiting step in glycolysis and gluconeogenesis, although new evidence suggests that the rate limiting step may actually be catalyzed by GAPDH when fructose-1,6-bisphosphate is abundant (Shestov, Liu et al. 2014). However, when levels of fructose-1,6-bisphosphate are low, it is possible that PFK (or PFP) may influence glycolytic flux. Phosphate is an inhibitor of the reaction in the forward direction (fructose 1,6-bisphosphate production) (Kombrink, Kruger et al. 1984). PFP can also allow PPi-dependent formation of fructose-1,6-bisphosphate, perhaps acting to spare ATP at night.

Aconitase is a TCA cycle enzyme that catalyzes the conversion of citrate to isocitrate, and thus also implicated in glucose metabolism as well as in fatty acid degradation. However, it is unclear what impact changing aconitase levels might have, as the important regulatory steps in the TCA cycle are thought to be isocitrate dehydrogenase and  $\alpha$ -ketoglutarate dehydrogenase since both are allosterically regulated by ADP (a positive regulator) and ATP (a negative

regulator). It is possible that the regulatory steps in the TCA cycle of the dinoflagellates differ from what is normally found, as dinoflagellates have found novel means to solve other common biochemical problems. For example, DNA is organized in the nucleus without the use of histones (Wargo and Rizzo 2001), and the normally found form I Rubisco in the plastid has been replaced with an oxygen sensitive form II enzyme (Morse, Salois et al. 1995).

The day phase protein lactate cytochrome c oxidoreductase (Cyb2) catalyzes the interconversion of pyruvate and lactate. In animal cells, lactate can be an important driver of the TCA cycle (Hui, Ghergurovich et al. 2017). However, in photosynthetic cells, the role of lactate may differ. Photosynthetic carbon fixation uses CO<sub>2</sub> and this has been observed to cause an increase in the pH in the culture medium (Eisensamer and Roenneberg 2004, Mackenzie and Morse 2011). This pH change results from the equilibrium between CO<sub>2</sub> and H<sub>2</sub>CO<sub>3</sub> in seawater, and is essentially the opposite of the ocean acidification resulting from rising atmospheric CO<sub>2</sub> levels. In day phase cells, there is ample ATP produced during photosynthesis, and pyruvate may not be required to fuel the TCA cycle to the same extent as in night phase cells. Thus, an increase in lactate oxidoreductase activity during the day may allow conversion of pyruvate to lactic acid which could act to counterbalance the increased pH due to CO<sub>2</sub> fixation.

The acetyl-CoA : oxalate CoA transferase (CoAT) was also observed to be more abundant during the day. Although we do not know the direction of the reaction catalyzed by this enzyme, two lines of evidence suggest it is likely to function in synthesis of oxalyl-CoA. First, there would be less need for acetyl-CoA to fuel ATP production in the mitochondria during the day, when photosynthesis is active, so acetyl-CoA would be available for the forward reaction. Second, photosynthesis itself may produce oxalate as a byproduct, as the glycolate which results from photorespiration can be oxidized to oxalate by glycolate oxidase (Franceschi and Nakata 2005). This suggests that the ultimate goal is degradation of oxalate, as recent work has shown that oxalyl-CoA is an intermediate in oxalate catabolism (Foster, Luo et al. 2016).

Triosephosphate isomerase (TPI) catalyzes the reversible conversion of glyceraldehyde-3-phosphate (G3P) and dihydroxyacetone phosphate (DHAP) and in plants is used in glycolysis, gluconeogenesis, glycerol metabolism and the Calvin cycle. During glycolysis, fructose-1,6-bisphosphate is catabolized to form G3P and DHAP, and since only G3P can continue in the glycolytic pathway, TPI serves the important role of converting DHAP to G3P. However, DHAP can also be reduced to glycerol-3-phosphate which serves as the starting point for

glycerolipid biosynthesis. TPI levels are highest during the day, when demand for ATP produced by glycolysis is low, so it is possible that an increased use of DHAP in glycerolipid synthesis might require increased DHAP production from G3P. Lipid synthesis is likely to occur during the day, along with starch synthesis, as both are fueled by carbon fixation (Dagenais Bellefeuille, Dorion et al. 2014).

#### *Proteins involved in unknown rhythms*

Sar1 is a monomeric GTPase that catalyzes the formation of vesicles at the ER membrane (Nakano and Muramatsu 1989). Clearly, given the many proteins that transit through the ER, and the many possible destinations for these proteins, it is impossible to distinguish with any certainty which proteins might be affected by increased Sar1 levels. One intriguing hypothesis is that increasing Sar1 levels may facilitate Rubisco transport to the chloroplast. In dinoflagellates, Rubisco transits through the ER and the Golgi on its way to the triple membrane bound plastid (Nassoury, Cappadocia et al. 2003), and blocking vesicular transport between the ER and the Golgi by Brefeldin A has a major impact on Rubisco distribution in the chloroplast (Nassoury, Wang et al. 2005). When carbon fixation capacity is high, Rubisco is localized in specialized regions of the plastid called pyrenoids. Pyrenoid formation and Rubisco synthesis both occur during the middle of the night phase, and increased levels of Sar1 may help facilitate the transfer of Rubisco to the plastid (Markovic, Roenneberg et al. 1996, Nassoury, Fritz et al. 2001).

For decades, eukaryotic elongation factor 3 (eEF3), a ribosome-dependent ATPase, was considered “fungal-specific”. However, recent bioinformatics analysis indicates it may be more widely distributed and also be present in other unicellular eukaryotes (Mateyak, Pupek et al. 2018). While this represents the first time it has been found in an alveolate, it has been found in other members of the so-called SAR (stramenopiles, alveolates, rhizaria) group. The role of the protein is not yet known, but is unlikely to be involved in global protein synthesis rates as these have been reported to be higher in night phase cells (Rensing, Cornelius et al. 1985).

The protein p43 was first identified as a membrane protein exposed at the external face of the cell by labelling with radioactive iodine that does not pass the cell membrane (Bertomeu, Hastings et al. 2003). This protein is found in a number of dinoflagellates, but has no homology to other known proteins by BLAST searches. The web server CCTOP (<http://cctop.enzim.ttk.mta.hu>) (Dobson, Remenyi et al. 2015) predicts a single N terminal

transmembrane domain with the bulk of the protein on the exterior face of the cell, so if p43 acts as a receptor its cytoplasmic signaling domain is small. It is nonetheless intriguing that levels of p43 vary over the course of the diurnal cycle. If the variations in levels are due to a daily synthesis, then the constitutive rate of protein degradation must be quite high as p43 has the greatest day-night difference in amount of all the proteins studied. However, it is possible that synthesis is constitutive and that the variations are due to an increased degradation rate at night. *Lingulodinium*, like many dinoflagellates, undergoes a diurnal vertical migration in the water column (Eppley, Holm-Hansen et al. 1968) and in culture flasks the cells are more concentrated at night than they are during the day phase (Roenneberg, Colfax et al. 1989). Perhaps cell concentration or location in the water column are related to p43 levels.

#### *General features of rhythmic proteins in dinoflagellates and perspectives*

It is at first sight remarkable that so few *Lingulodinium* proteins were observed to have significantly different levels over the daily cycle, as less than 0.8% of the proteins detected change in abundance. However, to give this figure some perspective, between 3 and 6% of the proteome in mammalian liver showed rhythmic changes (Reddy, Karp et al. 2006, Mauvoisin, Wang et al. 2014, Robles, Cox et al. 2014), and there are other systems that also have between 2% to 5% of the detected proteome showing rhythms (Table 2.2). Our values compared to these are reasonable, especially given that we restricted our analysis to those proteins with 3 or more peptides identified, while some other studies report on proteins identified by only a single peptide. It should be noted though, that other systems can display a far greater proportion of rhythmic proteins, with proportions between 20% and 77% being reported (Table 2.2).

Curiously, when the number of liver proteins whose synthesis is under circadian control was assessed by ribosome profiling, roughly 17% of the protein coding sequences examined showed changes (Janich, Arpat et al. 2015). At least in part, the difference between measurements of protein levels and of protein synthesis rates may lie in the fact that changes in protein levels do not directly result from changes in protein synthesis rates alone. A theoretical explanation for this has been proposed to involve rhythmic degradation and depends strongly on the half-life of the protein (Luck, Thurley et al. 2014). This has experimental support in *Lingulodinium*, where circadian changes in the synthesis rates of both Rubisco and PCP have been observed (Markovic, Roenneberg et al. 1996) yet do not result in measurable changes in

protein amounts (Le, Jovine et al. 2001, Nassoury, Fritz et al. 2001). In these cases rhythmic synthesis may increase the total protein level by a fraction that is too small to reliably measure for a protein with a low turn-over rate. We suggest it will be of interest to examine dinoflagellate protein synthesis rates more globally using ribosome profiling (Ingolia 2010) to assess the number of proteins whose translation is rhythmic.

Several lines of evidence support the idea that the changes in protein levels observed here are real. First, the variable proteins were selected based on having  $p < 0.05$  and a fold difference  $> 1.5$  between any two points among the ten samples. However, when the protein levels are followed over time we observe that the variations between successive times points are gradual and progressive. This seems more biologically plausible than a single large change in a single time point. Second, proteins that we expected to be found in our dataset, such as LBP and luciferase, were indeed found. To further confirm the variations observed in LBP levels, one of the samples used in the MS analysis was also analysed by Western blotting using an anti-LBP (Fig. 2.5), and the two show good agreement. A good agreement was also observed between TPI levels measured in the MS analysis and by coupled enzyme assays in crude extracts (Fig. 2.6). These two examples support the contention that the MS analysis is reporting real changes in the protein complement of the cells over the diurnal cycle.

Are low magnitude protein level changes, such as 1.5 or 2 fold, actually meaningful? It must be borne in mind that, all other factors being equal, doubling enzyme levels would double reaction rates. We also note that other studies also report similar ranges of protein level changes. For example, concerning the rhythmic proteins in mouse liver, one study reported a mean fold-change of 1.38 (Robles, Cox et al. 2014), while another reported that only 2.6% of the rhythmic proteins had a fold-change greater than 2 (Mauvoisin, Wang et al. 2014).

It might also be considered peculiar that a given metabolic pathway has only a single oscillating protein rather than a large number of similarly regulated enzymes. For example, in transcriptomic studies there are often changes in many transcripts whose products function together in a given pathway (Harmer, Hogenesch et al. 2000, Panda, Antoch et al. 2002). The implication from these proteomic studies is that it may not be necessary for a large number of proteins in a given pathway to vary, and indeed, this would be consistent with the rate limiting steps concept. A rate limiting step is the slowest in a given metabolic pathway, and thus

represents the only step in which increases in enzyme levels would be expected to increase metabolic flux through the pathway.

It must be noted that our procedure would not detect changes resulting from post-translational modification as described above. Thus, any modified peptides would remain undetected while the quantities of other, unmodified peptides, remained the same. The Progenesis quantitation protocol selected uses the three most abundant peptides, so phosphorylation of any other peptide would have had no effect on the measured protein levels. It is thus possible that significant changes in protein phosphorylation might have occurred and yet calculated proteins levels simply have been missed. It will also be of great interest to examine the effect of phosphorylation, either by purification of phosphoproteins prior to tryptic digestion or by purification of phosphorylated peptides following tryptic digestion.

## 2.5 Methods

### *Cell Culture and proteomic analyses*

Culture growth conditions and protein extraction methods for MS analysis from the dinoflagellate *Lingulodinium polyedra* (strain 1936 from the National Center for Marine Algae, East Boothbay, Maine) have been reported previously (Tse, Beauchemin et al. 2018). Briefly, cells were grown under a 12:12 light:dark cycle at 18°C and harvested by filtration on Whatman 541 paper. Samples were taken every three hours during the light period (ZT 0, 3, 6, 9 and 12) and every two hours during the dark period (ZT 14, 16, 18, 20 and 22) in each of three different



experiments. ZT refers to Zeitgeber time, with ZT 0 corresponding to lights on and ZT 12 to lights off. Filtered cells resuspended in extraction buffer (25 mM MES pH 6, 0.25% CHAPS, 1 mM PMSF, 2 mM NaF and 20 mM DTT) and lysed by vigorous shaking for two minutes at 4°C in a beadbeater (BioSpec Products). The extract was clarified by centrifugation, first at 13,000 x g for 5 minutes at 4°C, then at 13,000 x g for 15 minutes at 4°C. Protein concentration was measured by the Bradford assay, and 0.8 mg protein was allowed to precipitate overnight in 80% acetone at -20°C. Protein was recovered by centrifugation at 13,000 x g for 20 minutes in the cold, and the pellets washed twice with 80% cold acetone and air dried for 20 min. Subsequently, the protein pellets were resuspended in lysis buffer (6 M Urea, 50 mM DTT, 10 mM Tris) to a final concentration of roughly 10 mg/ml, and the protein concentration measured using the Bradford reagent. For each sample, 100 µg protein in a final volume of 10 µL was reduced by addition of 20 µL 1.5 mg/ml fresh dithiothreitol (DTT) at 60°C for 45 minutes, then alkylated by addition of 20 µL 10 mg/ml fresh iodoacetic acid (IAA) for 30 min in the dark at room temperature. The proteins were again precipitated overnight at -20°C by addition of 250 µl cold acetone, recovered by centrifugation as above, and the pellet air-dried before resuspension in 10 µL lysis buffer. The sample was then diluted with 200 µL 25 mM NH<sub>4</sub>HCO<sub>3</sub> and digested overnight with 5 µg trypsin at 37°C. Peptides were purified from the mixture using a Pierce C18 spin columns (Thermo Fisher Scientific USA) after acidification to pH < 4 by addition of 5% trifluoroacetic acid (TFA) following the manufacturer's instructions. Samples were dried in a Speedvac (Labconco USA), resuspended in 20 µL 0.1% formic acid (FA), and transferred to an HPLC vial for injection into the mass spectrometer.

LC-tandem-mass spectroscopy was performed using an LTQ-Fusion Lumos (Thermo Scientific, USA). Peptides were first fractionated by a 15 cm nanoflow C18 column (LC packings, Netherland) using an Ultimate 3000 nanoflow liquid chromatography (Thermo Scientific, USA) using a gradient from 6% to 30% of ACN in 0.1% of FA for 90 minutes. Eluents were introduced into an electrospray (ESI) where peptides were ionized by a nozzle potential of 2550V in positive mode. The temperature of the ESI was kept at 150°C. The mass spectrometer was operated in data-independent acquisition (DIA) mode. Precursor ions were first introduced to an Orbitrap mass analyzer for precursor mass acquisition (MS<sup>1</sup>) with the mass-per-charge range of 350-1500 and a resolution of 60,000. Fragment mass acquisitions

(MS<sup>2</sup>) were then performed in a linear iontrap mass analyzer. DIA parameters were : 20 spectra per DIA; a  $4 \times 10^5$  AGC target for the MS scan; 20 ms maximal IT per scan; 30,000 DIA scans resolution; 21m/z isolation width; 32% collision energy;  $5 \times 10^5$  AGC target of DIA; 50 ms maximal IT of DIA.

### *Data analysis*

The raw data files obtained from the tandem mass spectrometer were first used for peak picking by Progenesis QI for proteomics (Nonlinear Dynamics, Waters) using the default settings. Peak processing was performed for the triplicate time series simultaneously and the alignment quality reviewed manually for the different spectra. MS/MS peak lists were exported from Progenesis QI as an .mgf file and were imported into Mascot Server 2.5 (Matrix Science UK) for peak identification using a previously described *Lingulodinium* transcriptome with 74,655 entries >300 bases long assembled from 312 million reads (RNA was extracted from cells at two different times of day under normal culture conditions as well as from nitrate depleted cultures) (Beauchemin, Roy et al. 2012). Carbamidomethylation on the cysteine residue was set as the fixed modification, whereas oxidation on the methionine residue was set as the variable modification. One missed cleavage was permitted. Precursor ion tolerance and fragment ion tolerance were set at 10 ppm and 0.5 Da, respectively. Global identity false discovery rate (FDR) was kept under 1% by searching the data against a decoy database made by reversing the sequence of the *Lingulodinium* transcriptome library. The identifications were exported from Mascot server as an .xml file and reimported into Progenesis QI for matching to the peptide ions. For each protein, the default setting averaged the abundance of the three most abundant non-conflicting peptides to provide an expression reading for the protein. A total of 5,899 peptides and 1,701 quantifiable proteins were identified in the combined experiment (three independent time series), with 99.5% of the proteins identified by three or more peptides. Of these, a total of thirteen proteins were identified as having a significant daily variation in levels (max fold change  $\geq 1.5$  and an Anova  $p$  value  $\leq 0.05$ ). The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Vizcaino, Deutsch et al. 2014) partner repository with the dataset identifier PXD006994 and 10.6019/PXD006994.

Sequences with significant variations over the course of the daily cycle were identified using BLASTx against the non-redundant protein database. The putative locations of the proteins sequences were verified using SignalP 4.1 whenever possible. Heat maps were constructed using the heatmap.2 package in R, and plots with standard deviations were made using the ggplot2 package.

Tableau 2. 1. Proteins with significant changes in levels over the LD cycle

Accession Number	Identity	BLAST e-value	Subcellular location	Role
JO692620	Luciferase (LCF)	0	Scintillon	Bioluminescence
JO692704	Luciferin-binding protein (LBP)	0	Scintillon	Bioluminescence
JO692624	Cell surface protein (p43)	0	Plasmalemma	Unknown
JO714504	Triose-phosphate isomerase (TPI)	6e <sup>-96</sup>	Cytoplasm	Glycolysis/ Gluconeogenesis
JO762431	Vitamin B6 synthesis protein (PDX1)	3e <sup>-158</sup>	Cytoplasm	Pyridoxal phosphate biosynthesis
JO721472	Aconitase	3e <sup>-119</sup>	Mitochondrion	TCA cycle
JO693385	Proliferating cell nuclear antigen (PCNA)	6e <sup>-94</sup>	Nucleus	DNA replication
JO740110	Pyrophosphate--fructose-6-phosphate 1-phosphotransferase (PFP)	0	Cytosol	Glycolysis/Gluconeogenesis
JO721688	2-epi-5-epi-valiolone synthase (EVS)	6e <sup>-99</sup>	Cytosol	C7 cyclitol synthesis
JO702949	Elongation factor 3 (EF3)	6e <sup>-41</sup>	Cytoplasm	Protein synthesis
JO740066	Sar1	5e <sup>-121</sup>	ER	Vesicle formation
JO710064	Lactate oxidoreductase (Cyb2)	5e <sup>-170</sup>	Mitochondrion	Lactate/pyruvate conver.
JO707433	acetyl-CoA:oxalate CoA transferase	1e <sup>-165</sup>	Mitochondrion	Oxalate metabolism

Tableau 2. 2. Proportion of rhythmic proteins in a range of systems.

Sample	# Proteins	# Rhythmic proteins	FDR	Reference
<i>Lingulodinium</i>	1701	13 (0.8%)	1%	This study
<i>Ostreococcus</i>	653	16 (2.4%)	1%	(Le Bihan, Martin et al. 2011)
<i>Arthrospira</i>	1306	30 (2.3%)	<2%	(Matallana-Surget, Derock et al. 2014)
<i>Synechococcus</i>	1537	77 (5%)	1%	(Guerreiro, Benevento et al. 2014)
<i>Cyanothece</i>	1232	250 (20%)	-	(Stockel, Jacobs et al. 2011)
<i>Prochlorococcus</i>	548	312 (77%)	3.2%	(Waldbauer, Rodrigue et al. 2012)
Mouse liver	3132	186 (6%)	1%	(Robles, Cox et al. 2014)
Mouse liver	5827	195 (3.3%)	1%	(Mauvoisin, Wang et al. 2014)
Mouse SCN	2112	421 (20%)	1%	(Chiang, Mehta et al. 2014)
<i>Neurospora</i>	4742	1273 (27%)	-	(Hurley, Jankowski et al. 2018)
<i>Drosophila</i> heads	3154	45 (1.4%)	1%	(Du, Zhang et al. 2018)

Figure 2. 1. Range of protein abundance.

The average levels of the 13 proteins whose levels vary over the diurnal cycle are spread out over more than two orders of magnitude. The two most abundant proteins are LBP and luciferase.

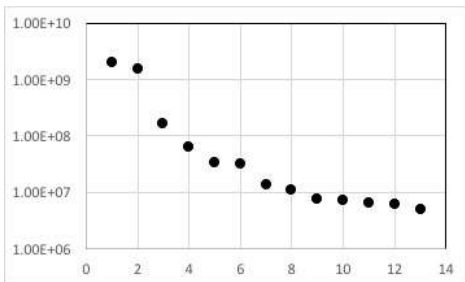


Figure 2. 2. Proteins with diurnal changes fall into two groups.

A heat map showing the levels of the 13 proteins with significant changes in abundance clustered into two main groups, those with higher levels during the night and those with higher levels during the day (upper panel). Values are normalized so that for each protein the mean is zero and the standard deviation is 1 (z-score). The means and standard deviation (grey areas) of the two different groups, normalized to a mean of 1, are plotted separately (bottom panels).

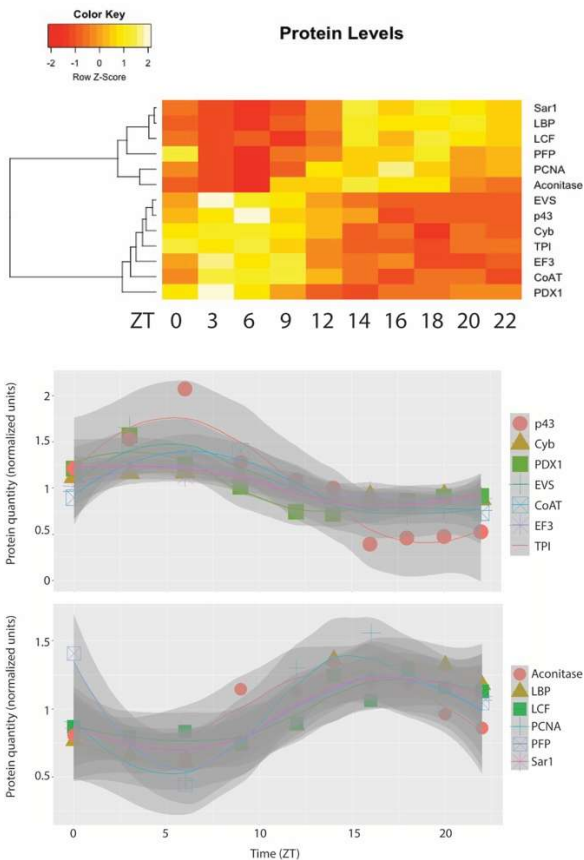


Figure 2. 3. The diurnally varying proteins plotted by abundance.

The mean (line) and standard variation (error bars) of proteins whose levels are higher during the day (A, C, E) and the night (B, D, F).

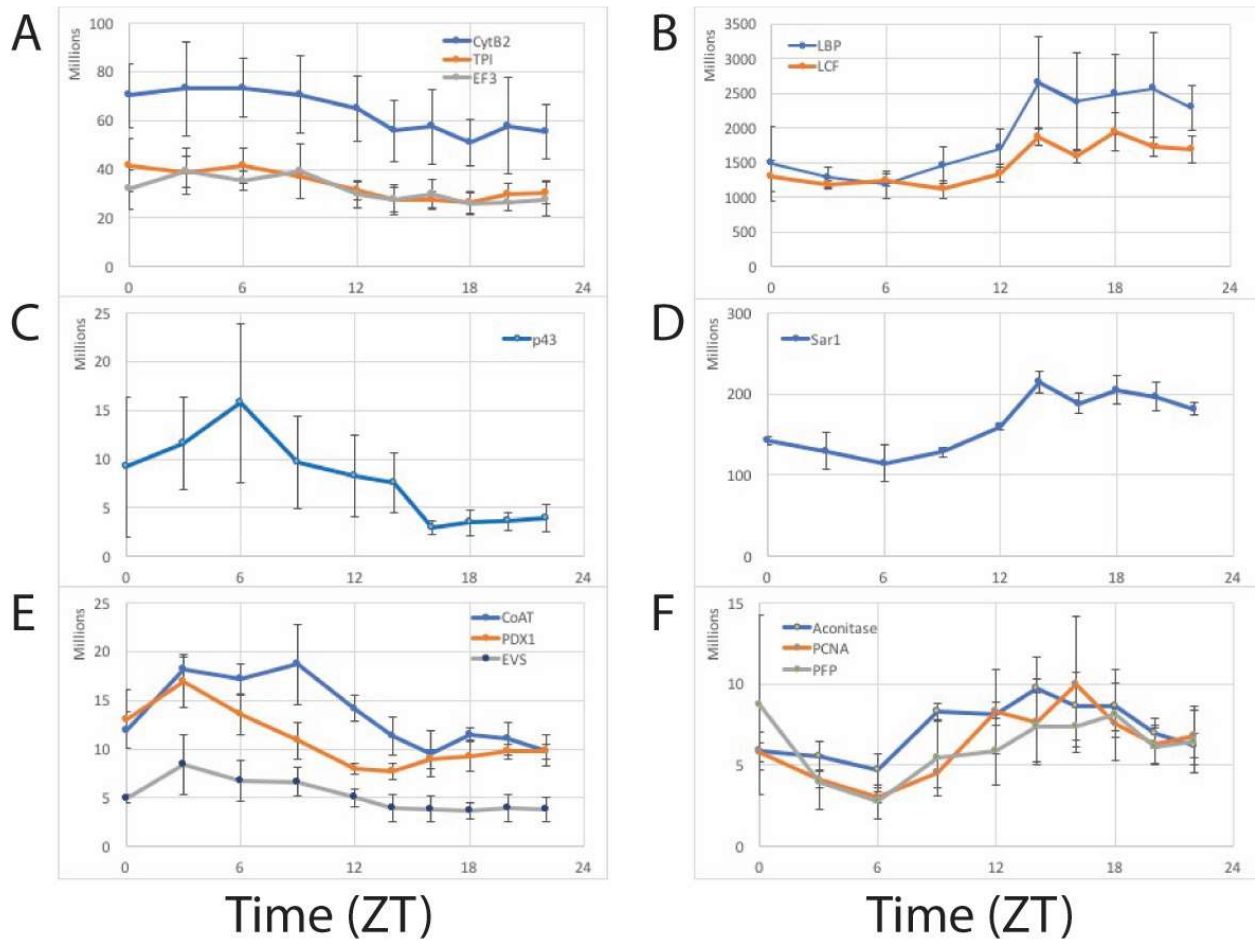




Figure 2. 4. Schematic overview of the cell localisation for diurnally varying proteins.

Proteins more abundant during the night (blue) or the day (red) are shown in their most likely subcellular location. Relevant metabolic intermediates are shown in pale grey.

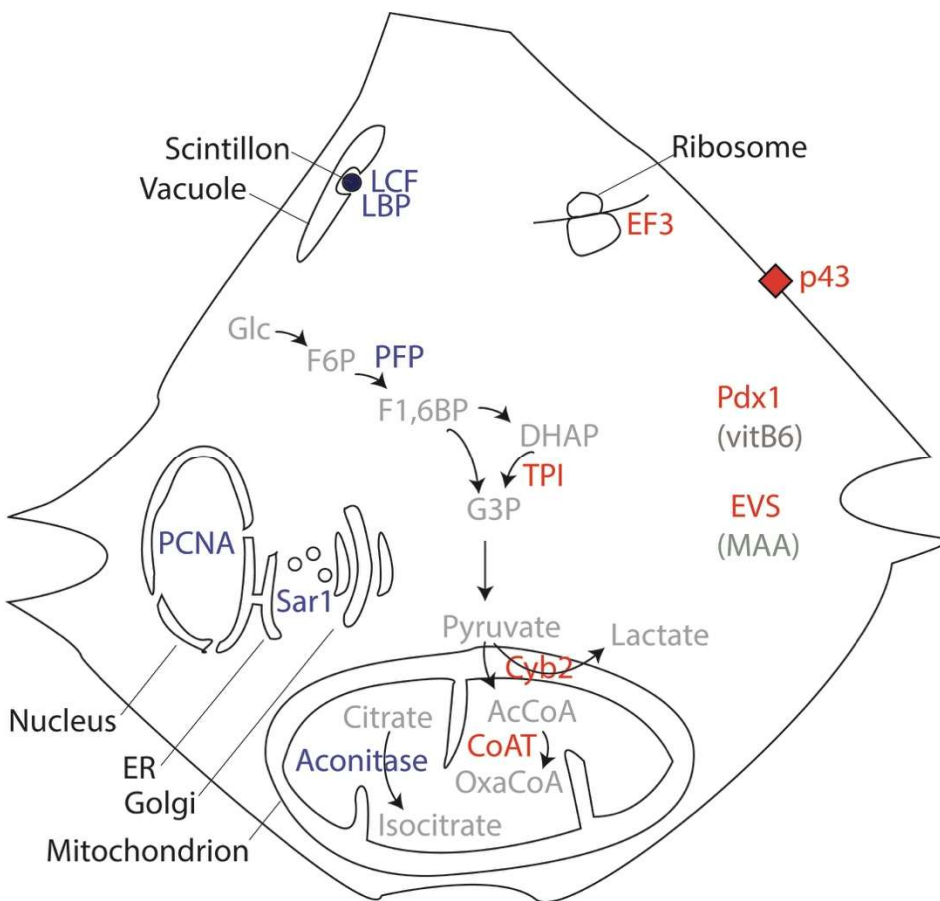


Figure 2. 5. (Supplementary) Comparison of LBP levels measured by Western blotting and by MS/MS.

A sample set probed with an anti-LBP antibody (upper panel) and stained with Ponceau (lower panel). Quantitation of the Western blots, corrected for protein load using the Ponceau stained membrane, is compared to the MS quantifications (orange bars) normalized to an average value of 1.

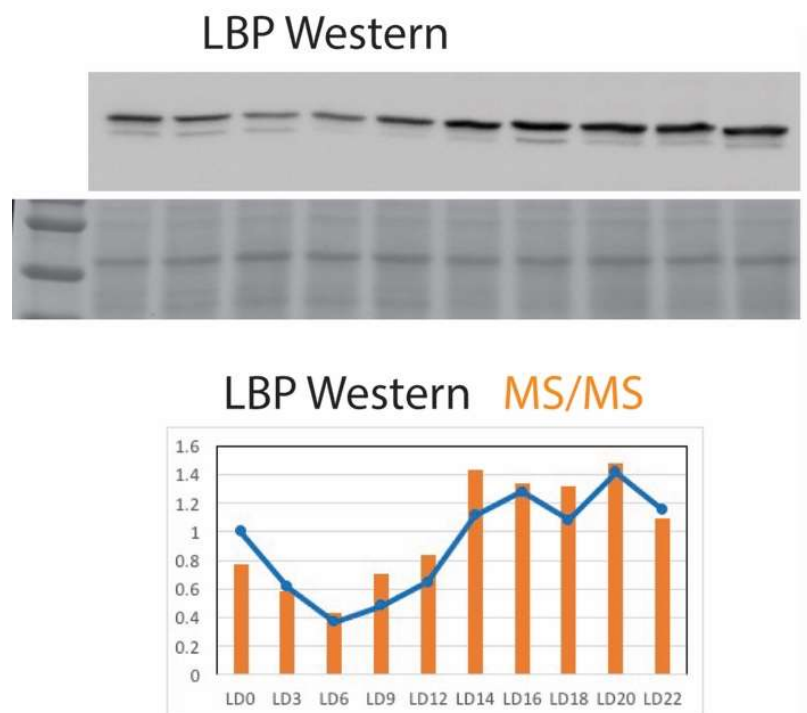
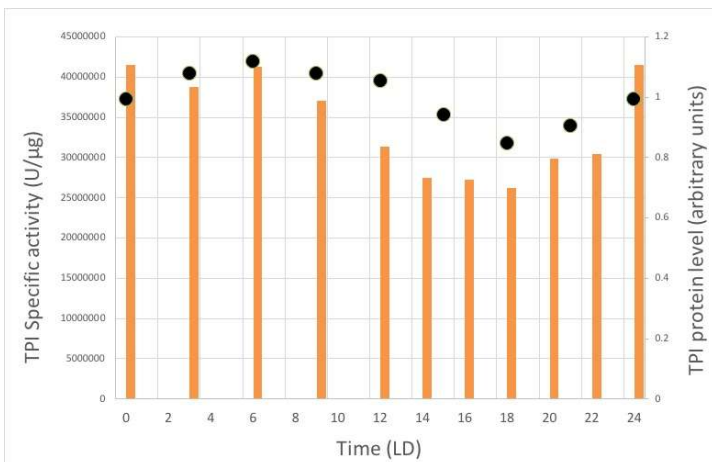


Figure 2. 6. (Supplementary) Comparison of TPI levels measured by coupled enzyme assays and by MS/MS.

TPI enzyme activity levels were averaged from a triplicate set of cell cultures (black circles) and compared to MS quantifications (orange bars) normalized to an average value of 1.





**Chapitre 3: Orchestrated translation specializes dinoflagellate metabolism three times a day.**

## **Orchestrated translation specializes dinoflagellate metabolism three times a day.**

Carl Bowazolo, Bo Song, Mathieu Beauchemin, Sonia Dorion, Jean Rivoal and David Morse

Article soumis

Contribution des auteurs:

Les tests préliminaires de la préparation des échantillons ont été réalisés avec l'aide de Mathieu Beauchemin. Pour les expériences rapportées, j'ai préparé la totalité des échantillons. Bo Song a contribué à l'analyse bio-informatique, notamment pour le comptage des fragments et la recherche des motifs dans les transcrits réglés au même moment. Avec David Morse j'ai fait l'analyse des résultats et fait plusieurs des figures. J'ai écrit la première version de l'article qui a été revue et corrigée par David. Les analyses des acides aminés ont été faites dans le laboratoire de Jean Rivoal par Sonia Dorion.

### **3.1 Abstract**

Most cells specialize for different metabolic tasks at different times of day by changes in gene expression. However, in many organisms, the number of genes affected, and the identity of the metabolic states found throughout the diel cycle, are mostly unknown. Here, we use ribosome profiling to identify thousands of transcripts showing coordinated translation of multiple enzymes in relevant biosynthetic pathways. In our dinoflagellate model, cells become specialized for daily photosynthesis around dawn (LD0), for nightly DNA synthesis around dusk (LD12) and for late night gene expression with synthesis around LD16 of proteins involved in transcription, translation and amino acid biosynthesis. The latter has physiological consequences, as measured free amino acid levels increase at night and represents a previously unknown rhythm in this model.

Keywords: dinoflagellates, ribosome profiling, diurnal rhythms, cell metabolism

### **3.2. Introduction**

Daily rhythms in physiology, behavior and biochemistry are an almost ubiquitous feature of organisms adapted to a rotating planet. Many daily rhythms in bacteria, plants and animals are regulated by endogenous circadian clocks and are thought to provide a selective advantage (Ouyang, Andersson et al. 1998, Dodd, Salathia et al. 2005, Spoelstra, Wikelski et al. 2016) by

allowing an organism to adapt to daily environmental changes even before these changes occur. The daily metabolic changes are brought about by changes in gene expression, generally transcriptional (Harmer, Hogenesch et al. 2000, Zhang, Lahens et al. 2014), that specialize cells for different tasks at different times. Dinoflagellates are a useful unicellular model to study diel metabolic changes as, like many members of the marine phytoplankton, they specialize for photosynthesis during the day and for DNA replication at night. They are also of ecological interest as important contributors to net oceanic primary production, the symbiotes of coral reefs, and the causative agents of some harmful algal blooms referred to as red tides.

The biochemical basis for some rhythms in the dinoflagellate *Lingulodinium* (formerly *Gonyaulax*) *polyedra* has been examined in some detail. Photosynthesis, a day phase rhythm, is controlled in part by rhythmic translation of ribulose biphosphate carboxylase/oxygenase (Rubisco) which results in formation of Rubisco-enriched structures called pyrenoids inside chloroplasts (Nassoury, Fritz et al. 2001, Nassoury, Wang et al. 2005). Pyrenoids exclude the light-harvesting peridinin-chlorophyll a-protein (PCP), thus increasing carbon fixation efficiency by separating the oxygen sensitive Rubisco from the source of oxygen generation. Bioluminescence, a night phase rhythm, results from translation of two essential bioluminescence proteins, luciferase and luciferin binding protein (LBP) (Morse, Milos et al. 1989, Mittag, Li et al. 1998), and levels of the two proteins vary in phase with the nightly light production (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989). Under 12:12 Light:Dark conditions (LD0 is dawn and LD12 dusk), the cell cycle is timed by a circadian clock with S phase at LD18 and M phase at LD1 (Dagenais-Bellefeuille, Bertomeu et al. 2008), but the mechanisms regulating these events are still unclear. These few specific examples suggest translation of proteins underlies changes in dinoflagellate metabolism, but do not address the extent to which changes in gene expression occur, what aspects of cell metabolism are altered over a daily period, and how many different types of specialized cell metabolism may be found over the diel cycle.

Despite their many observed rhythms, *Lingulodinium* and other dinoflagellates (Roy, Beauchemin et al. 2014, Zaheri, Dagenais-Bellefeuille et al. 2019), show essentially constant transcript levels over a diel period, so transcriptomics provides no clues as to the metabolic capacity of cells at different times of day. Here, we have used ribosome profiling to gauge genome-wide changes in RNA translation. This technique counts the number of ribosome



protected fragments (RPF) after nuclease digestion as a surrogate for genome-wide protein synthesis rate measurements (Ingolia 2010, Ingolia 2016), and has been successfully employed to assess daily rhythms in translation in animal models (Jang, Lahens et al. 2015, Janich, Arpat et al. 2015). Comparisons of transcript levels and RNA translation rates show important differences between the two, so we hypothesized that ribosome profiling might provide a comprehensive picture as to the number, timing and type of genes whose expression varies over the diel period in dinoflagellates.

### 3.3 Results and Discussion

To assess the ability of ribosome profiling to provide reliable information on changes in gene expression occurring at a translational level, we first examined three times in a 12:12 light-dark cycle (Fig. 3.1.A) where previous work using a different strain of the same species had identified high synthesis rates for four proteins (PCP at LD6; LBP at LD14; Rubisco and glyceraldehyde-3-phosphate dehydrogenase at LD20) using *in vivo* <sup>35</sup>S-methionine labeling (Markovic, Roenneberg et al. 1996). We chose a light-dark cycle (LD, where LD0 is lights-on and LD12 is lights-off) in order to identify all changes that might help an organism perform better at a particular time of day, so changes in synthesis rates will include both circadian and direct responses to changes in light intensity.

All samples contained the expected ~30 base ribosome-protected RNA fragments (RPF) as determined both by gel electrophoresis and by sequence analysis (Figs. 3.5 and 3.6). Triplicate samples from any one time were generally more similar to one another than to samples from the two other times (Fig. 3.7 and 3.1.B), and RPF were distributed along the length of the LBP transcript at its peak synthesis time (LD14), as expected (Fig. 3.8). Furthermore, RPF counts at the three times agreed with what was expected for the four control proteins (Table 3.4). Significant differences in RPF were calculated for all pairwise comparisons of two times, then combined to yield a total of 3324 transcripts with significant differences in RPF at one of the three sampling times. Roughly half the transcripts at LD6 and LD14 had high RPF at only one time (Fig. 3.1.B), while most transcripts at LD20 had high RPF at two of the three times.

Since the overall correlation between triplicate samples was generally good (Fig. 3.7), in a second experiment a single sample was taken every two h (Fig. 3.1.A) to assess the daily protein synthesis landscape at higher temporal resolution. Here, 3327 transcripts were found to have significant differences in RPF levels with a 24 h periodicity. Of these, 1020 transcripts were also present in the 3 times dataset; differences between the two experiments is likely due to the greater statistical power of triplicate samples in the first experiment and to imposition of a 24 h periodicity criterion in the second experiment. Interestingly, peak RPF times for the different transcripts are not evenly distributed across the diel cycle, but are instead concentrated around three times, (i) a transition from dark to light (LD0), (ii) a transition from light to dark (LD12) and (iii) a time (LD16) close to the middle of the night phase (Fig. 3.1.C). These patterns are readily observed on a heat map after clustering (Fig. 3.1.D).

Although the higher temporal resolution experiment finds three times where large numbers of transcripts are translated, these times do not directly correspond to the times selected for the first experiment, so this may also contribute to the differences in the sequences identified by the two experiments. However, changes in RPF values again compare favorably to *in vivo* <sup>35</sup>S-methionine incorporation for the four previously characterized proteins (Markovic, Roenneberg et al. 1996) (Fig. 3.9) as well as to the time of increasing protein levels for 13 proteins recently identified by LC-MS/MS to vary in amount over the LD cycle (Fig. 3.10) (Bowazolo, Tse et al. 2020). We conclude that RPF data is a good surrogate for changing protein synthesis rates. Interestingly, a test for enrichment of potential regulatory sequences in transcripts with peak synthesis around lights-on revealed an enrichment of GC-rich motifs not observed in transcripts synthesized at other times (Table 3.5) suggesting the possibility of common regulatory mechanisms for at least some groups of transcripts.

During the light period, cells have specialized for photosynthesis, and we hypothesized that proteins involved in this process might be among those whose peak synthesis occurred around LD0. Indeed, photosynthesis, carbon fixation and the pentose phosphate shunt are among the principal GO categories significantly enriched at LD6 in the 3 times experiment (Fig. 3.11). In the GO category for photosynthetic light reactions, including photosystem proteins, light harvesting proteins, and ferredoxin, all proteins with significant differences show their peak synthesis rates beginning just before lights on and continuing through the day phase (Fig. 3.12). The peak time for synthesis of all 13 enzymes involved in carbon fixation (Fig. 3.2.A) is more

circumscribed than observed for the photosynthesis proteins, with a more sharply defined peak of translation centered around LD0. Synthesis rate changes occur concurrently for almost all transcripts, and increased translation occurs just prior to lights on suggesting a potential for circadian control over this process. Circadian control has been documented for one of these enzymes, ribulose biphosphate carboxylase/oxygenase (Fig. 3.9) (Nassoury, Fritz et al. 2001). A similar pattern of synthesis rate changes was seen for eight of twelve enzymes in the glycolysis/gluconeogenesis pathway (Figs. 3.2.C and 3.2.D). Three of the glycolytic enzymes were not regulated, and the last, a pyrophosphate fructose-6-phosphate phosphotransferase (PFP; red asterisk), was regulated but with peak synthesis times around the onset of darkness. Interestingly, PFP uses pyrophosphate instead of ATP as a substrate, so the PFP-catalyzed reaction may support metabolic flux through glycolysis at night due to pyrophosphate production by the nightly DNA synthesis (see below). PFP is also used in the pentose phosphate pathway, which generates NADPH in a parallel pathway to glycolysis. However, the six other regulated enzymes in this pathway show peak synthesis around LD0, similar to carbon fixation (Fig. 3.13). Increased glycolytic flux during the day will produce both reducing power and pyruvate which can fuel fatty acid synthesis, the latter by conversion to acetyl-CoA. In agreement with this, fatty acid biosynthetic enzymes also show peaks in translation around LD0 (Fig. 3.14), suggesting accumulation of fatty acids during the day may store energy produced by photosynthesis. This suggestion is supported by increased translation of two enzymes involved in fatty acid degradation around LD12 (Fig. 3.14). While daily rhythms in fatty acid levels have not been reported, extensive sequestration of fixed carbon in triglycerides has been observed after nitrate deprivation, indicating cells can store photosynthetically fixed carbon in this form (Dagenais Bellefeuille, Dorion et al. 2014). Starch, another potential reserve of fixed carbon, has been observed to accumulate during the day and to disappear during the night (Dagenais Bellefeuille, Dorion et al. 2014). There are few regulated enzymes in the starch metabolic pathway, but two enzymes for starch production (starch synthase and starch branching enzyme) are synthesized at dawn while starch phosphorylase, catalyzing starch degradation, is synthesized at dusk (Fig. 3.15). Enzymes involved in biosynthesis of cellulose are also rhythmic, but this storage form of fixed carbon is unlikely to act as an energy reserve, as translation of biosynthetic enzymes is restricted to LD22 (Fig. 3.17). It seems more likely that cellulose biosynthesis will be used for formation of new cellulosic thecal plates (Chan, Kwok et al. 2019)

which surround the cell and must be created during cytokinesis at LD1. In addition to glucose, several enzymes involved in fructose and mannose production also appear regulated, with all enzymes (excepting PFP, as noted above) preferentially translated around LD0 (Fig. 3.18). Thus, during the day, specialization for photosynthesis and carbon metabolism occurs by increased translation around LD0 of multiple enzymes in a range of metabolic pathways.

In contrast to day phase cells, night phase cells specialize for bioluminescence (Morse, Milos et al. 1989) and DNA replication (Dagenais-Bellefeuille, Bertomeu et al. 2008). Bioluminescence peaks at midnight, and the only two known proteins involved (luciferase and LBP) are concurrently translated with peak synthesis rates centered around LD12 (Figs. 3.3.A-B and 3.10). DNA replication begins at midnight, and is among the principal GO categories significantly enriched at LD14 using Fisher's exact test in the 3-times experiment (Fig. 3.19). The peak translation time for the majority of regulated proteins involved in DNA replication and repair is around LD12 in the 12-times experiment (Figs. 3.3.D-E). Interestingly, the proteins involved include not only the DNA polymerase, primase and the MCM proteins, but also a dinoflagellate/viral nucleoprotein (DVNP) (Gornik, Ford et al. 2012), a major basic nuclear protein that has replaced histones in compacting DNA and is responsible for the unusual dinoflagellate chromatin structure (Irwin, Martin et al. 2018). Translation rates of the DNA clamp protein, proliferating cell nuclear antigen (PCNA), also change and agree with previously reported changes in protein levels (Bowazolo, Tse et al. 2020) (Fig. 3.10). Curiously, few enzymes involved in purine and pyrimidine biosynthesis were found to be regulated (Figs. 3.21 et 3.22) with most regulated enzymes in these pathways involved in interconverting the various phosphorylated forms of the different nucleotides. The TCA cycle also appears as an GO category enriched at LD14 in the three times experiment (Fig. 3.19). All steps catalyzed by enzymes found in the transcriptome are regulated (Fig. 3.3.F), and in both experiments (Figs. 3.3.G-H) the pattern of protein synthesis is remarkably similar to what was observed for proteins involved in DNA replication. Nightly increases in the levels of two TCA cycle enzymes have been previously reported, with at least one regulated by the circadian clock at a translational level (Akimoto, Kinumi et al. 2005). It is thus possible that metabolic flux through the TCA cycle increases during the night phase, and may provide energy for DNA synthesis by aerobic metabolism. Starch degrading enzymes show increased translation rates during night phase (Fig. 3.15), suggesting starch may be the carbon storage used to fuel aerobic metabolism.

Pyruvate is the direct source of acetyl-CoA for the TCA cycle, and pyruvate production by glycolysis at night may be aided by PFP, a glycolytic enzyme using pyrophosphate as a phosphate donor (Fig. 3.2.F). Examination of the pyruvate metabolic pathway indicated that enzymes playing a role only in pyruvate metabolism showed peak synthesis at LD12, while enzymes also playing a role in other pathways had peak synthesis rates at LD22 (Fig. 3.20).

Thus, the known day and night phase rhythms correlate well with identified metabolic pathways showing increased protein synthesis rates at LD0 and LD12, respectively. However, the third identified peak in the number of regulated transcripts seen around LD16 (Fig. 3.1.C) was unexpected. To assess what metabolic processes might be specifically occurring at LD16 in correlation with this peak, we first examined the GO categories enriched at LD20 (Fig. 3.23). We eliminated categories such as sugar metabolism and photosynthesis, clearly associated with LD0, reasoning that LD20 (midway between LD16 and LD0) might contain categories from both LD16 and LD0. Remaining GO categories for specific enrichment at LD16 thus include regulation of gene expression and biosynthetic/metabolic processes for 16 amino acids. We thus examined RPF values in the 12-times experiment for all sequences annotated as involved in transcription, translation, ribosome formation, aminoacyl-tRNA synthesis and amino acid metabolism. The protein synthesis patterns for all these categories were generally similar, with high translation rates between LD14 and LD16 (Fig. 3.4). This suggests that in addition to the previously characterized night phase rhythms, the cells may also have a late night phase rhythm in gene expression. This would be consistent with ribosome profiling in mouse liver showing increased translation of sequences associated with the protein biosynthesis machinery at LD14-16 (Janich, Arpat et al. 2015) and increased ribosomal protein synthesis in night phase *Arabidopsis* (Missra, Ernest et al. 2015).

Our current strain of *L. polyedra* does not take up radiolabeled amino acids from the external medium, precluding facile measurement of global protein synthesis rates, although a higher rate of global protein synthesis during the night has been reported for the species (Cornelius, Schroeder-Lorenz et al. 1985, Donner, Helmboldt-Caesar et al. 1985). Instead, we examined levels of free amino acids to determine if changing levels of enzymes involved in amino acid synthesis might impact cellular levels of these metabolites. Almost half (26 of 56) of the amino acid biosynthetic enzymes present in the transcriptome are regulated in both experiments, as seen when individual sequences are superimposed on a schematic illustrating

the various amino acid biosynthesis pathways (Fig. 3.24), suggesting many amino acids would show increased levels during the night. When measured at four times over the LD cycle, the levels of most individual amino acids measured as well as the total sum of all free amino acids have their lowest values during mid to late day (Fig. 3.25 and Table 3.6), with the increase in amino acid pool levels mirrored by a general increase in RPF for amino acid biosynthetic enzymes (Fig. 3.25). RPF measurements have thus correctly predicted a new physiological rhythm. Interestingly, preferential amino acid synthesis during the dark period would represent a previously unexpected aspect to N-metabolism. This process may be restricted to late night phase to partition it temporally away from carbon metabolism during the light phase or to align amino acid production with increased gene expression. In any event, our data supports the idea of a third metabolic specialization occurring the diel cycle, with increased levels of gene expression occurring between DNA synthesis and carbon metabolism.

Our results using RPF as a surrogate for gene expression in dinoflagellates indicate this method is likely to provide a more comprehensive picture of changes in metabolic activity than changes in transcript levels. Indeed, there is little correlation between RNA levels in the general transcriptome and RPF counts as a measure of translational efficiency (Fig. 3.26). Ribosome profiling may be advantageous in identifying transcripts in harmful algal blooms whose expression is increased during times of toxin production. It may also reveal the regulatory genes involved in regulating cell proliferation, or genes whose expression is induced during the formation of endosymbiotic relationships with coral. Importantly, our findings reveal that metabolic specialization of dinoflagellates during a light dark cycle is more sophisticated than a simple toggling between day and night phase states, an observation which should have important repercussions in analysis of rhythmic phenomena in other systems.

### **3.4. Methods**

#### **Cell culture and harvesting**

*Lingulodinium polyedra* (previously *Gonyaulax polyedra*) was obtained from the Provasoli-Guillard National Center for Marine Algae and Microbiota (East Boothbay, ME,

USA) as culture CCMP1936. Cell cultures were grown in normal f/2 medium (Guillard REF) prepared using Instant Ocean under a 12-hour light (40  $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$  cool white fluorescent light) and 12-hour dark regime at a temperature of  $18 \pm 1^\circ\text{C}$ . This light/dark regime, termed ZT, has ZT0 at lights on and ZT12 at lights off. After 2 minutes of pre-treatment with cycloheximide (0.10 mM) 300 mL of cells were harvested by filtration on Whatman 541 paper at ZT6, ZT14, and ZT20. Cells in 1.5 mL Eppendorf tubes containing 0.1 mm Zirconium beads for cell disruption were immediately frozen in liquid nitrogen and stored at  $-80$  until use.

### **Cell lysis and nuclease digestion**

The protocol used was adapted from previously described protocols (Ingolia, Brar et al. 2012, Reid, Shenolikar et al. 2015) and is described below. Eppendorf tubes containing cells and beads were placed on ice and resuspended in 200  $\mu\text{l}$  of lysis buffer (1% NP-40, 200 mM KOAc, 25 mM K-HEPES pH 7.2, 10 mM MgCl, 4 mM CaCl<sub>2</sub> and 3.5 $\mu\text{M}$  cycloheximide) and lysed for 2 minutes in a bead beater (BioSpec products). The supernatant, after 5 minutes centrifugation at 9000 rpm (at  $4^\circ\text{C}$ ), was taken up in a new tube and KOAc concentration adjusted to 100 mM by 1:1 with water before addition of 5  $\mu\text{l}$  RNase If (NEB). Tubes were incubated for 45 minutes at room temperature.

### **RNA extraction**

Tubes were put back on ice and then, sequentially, 1 mL of Trizol and 200  $\mu\text{l}$  of chloroform were added. After homogenization, the tube were returned to room temperature for 2-3 minutes. The tube was then centrifuged for 15 minutes at 12000 rpm ( $4^\circ\text{C}$ ). The aqueous supernatant was recovered in a new tube and incubated with 500  $\mu\text{l}$  of isopropanol for 10 minutes at room temperature, then centrifuged for 10 minutes at 12000 g ( $4^\circ\text{C}$ ). The RNA pellet was resuspended with 10  $\mu\text{l}$  of water for electrophoresis.

### **Gel migration and RPF excision on gel**

RNA loading buffer (5  $\mu\text{l}$  of 96% formamide, 2% 1 mg/mL bromophenol blue and 2% 10 mM EDTA) was added to the resuspended RNA pellet, incubated at  $70^\circ\text{C}$  for 5 minutes, then loaded on a 10% acrylamide (29:1) 1X TBE gel with 0.3  $\mu\text{g}$  of 10 bp ladder (Thermoscientific), Samples were electrophoresed at 200 volts for about 22 minutes, after which the gel was colored

with ethidium bromide (0.5 µg/ml). Bands at 30 nucleotides visualized under UV light (Figure 1A) were cut with a clean razor blade and the piece of gel put in 1.5 mL tube containing 400 µl of 400 mM NaOAc (pH 5.2). The tube was then frozen at -80°C for 10 minutes.

### **RNA extraction from gels**

The tube was thawed at 95°C for 7 minutes and the piece of gel totally crushed with a sterile pipet tip. This was followed by two cycles of 7 min at 95°C and a few moments of vortex, then the tube was centrifuged at maximum speed for 10 minutes. The supernatant was taken up in a tube containing 1 ml of ethanol, incubated for 1 hour at -20°C, centrifuged at 4°C for 15 minutes at maximum speed. The pellet containing the RPFs was washed with 80% ethanol and centrifuged again for 5 minutes. The supernatant was discarded and the pellet dried for the PNK modification.

### **PNK modification and gel migration**

The pellet was resuspended in 144 µl of water, 18 µl of buffer A (10x), 18 µl ATP (100 mM) and 2 µl PNK, and the tube incubated with shaking at 37°C for 45 minutes. After the reaction, the RNA was precipitated using ethanol as described (Walker and Lorsch 2013). The RNA pellet was resuspended in 30 µl of DEPC water and the quantity determined using a NanoDrop Spectrophotometer (MBI Lab equipment). 20 µg were electrophoresed (Figure 1B) under the same conditions as the first gel while the rest was stored at -80°C. At this stage the sample was sent to Génome Québec Innovation Centre (Montreal, Canada) for library construction and sequencing (Reid, Shenolikar et al. 2015). Sequencing was performed using an Illumina HiSeq deep sequencing protocol (Podnar, Deiderick et al. 2014) (Quail, Kozarewa et al. 2008).

### **Bioinformatics analysis**

The sequences in fasta format produced by the sequencing reaction were trimmed in Galaxy to remove the adapters using TrimGalore. The length of the sequences was determined in Geneious, and read distributions were made by mapping reads to selected reference sequences to verify the distribution. Read counts were determined using Salmon in Galaxy with a previously described Velvet assembly containing 74,655 sequences (Beauchemin, Roy et al.



2012) or a Trinity assembly containing 114,492 sequences (Roy, Letourneau et al. 2014). In one experiment, triplicate samples were taken at each of three times (LD 6, LD14, and LD 20). Pairwise significant differences between all combinations of these three times were determined using DESeq running in R, and all sequences with a  $p_{\text{adj}} < 0.05$  were combined to form a single file containing 3324 entries with at least one time showing a significant difference. In a second experiment, a single sample was taken every two hours. These data were tested for significant rhythmicity using JTK\_cycle (Hughes, Hogenesch et al. 2010) running in R, and all sequences with a 24 hour period, an average read count greater than 1 per million for all samples, and a Benjamini-Hochberg q-value less than 0.05 were selected.

### **Amino acid analysis**

Cells were disrupted in 80% (v/v) ethanol using a bead beater followed by a 30 min extraction at 70°C using a Büchi Rotavapor (Dagenais Bellefeuille, Dorion et al. 2014). The extract was brought to dryness and redissolved in 5 mL H<sub>2</sub>O. This extract was then separated into neutral, cationic and anionic fractions using tandem ion exchange chromatography as previously described (Rivoal and Hanson 1993). One mL aliquots of the basic fractions were evaporated to dryness before resuspension in 100 µL of 20% (v/v) acetonitrile in 20 mM HCl. Twenty µL of this solution were analyzed by HPLC after derivatization with the AccQ reagent (Waters). The Waters HPLC system was controlled by the Empower Pro software and equipped with a 600 controller. Detection was done using a Waters 2996 Diode Array Detector. Amino acids were separated on a Waters AccQ•Tag Amino Acid Analysis column (3.9 mm x 150 mm) equipped with a Nova-Pak C18 pre-column cartridge (3.9 mm x 20 mm). Quantification was performed using calibration curves generated from commercial standards (Thermo Scientific (Mississauga, ON, Canada) and Sigma Aldrich (St Louis, MO)). Results were corrected for recovery using L-norvaline as an external standard.

Figure 3. 1. A complex temporal landscape of protein synthesis rates with three peak times.

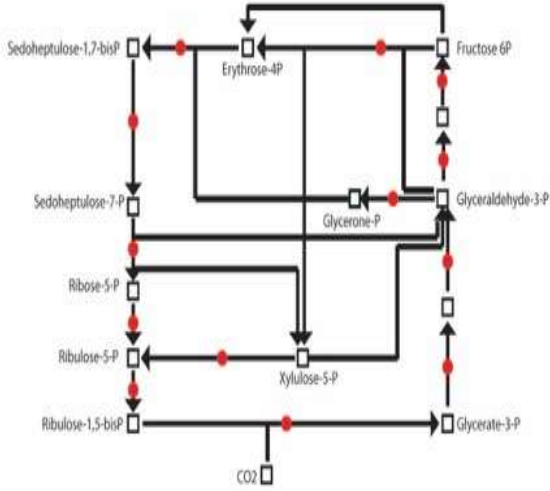
Two experiments with different temporal resolutions were tested, one at roughly 8 h intervals (top) the other at 2 h intervals. LD0 corresponds to dawn (A). A total of 3324 significant ( $p_{\text{adj}} < 0.05$ ) differences (as determined by pairwise comparisons of time points using DESeq2) between triplicate samples at the three times are shown as a clustered heat map, with transcriptome entries represented by horizontal lines and RPF read counts on a color scale (B). For the 2 h time resolution experiment, the number of significant transcripts (BH.Q value  $< 0.05$ ) with a 24 h rhythmic period peaking at each time (h, as determined by JTK cycle) is shown (C). Double-plotted read counts for 3327 transcripts indicate three major phases of protein synthesis of which two roughly coincide with transitions between light and dark (D).



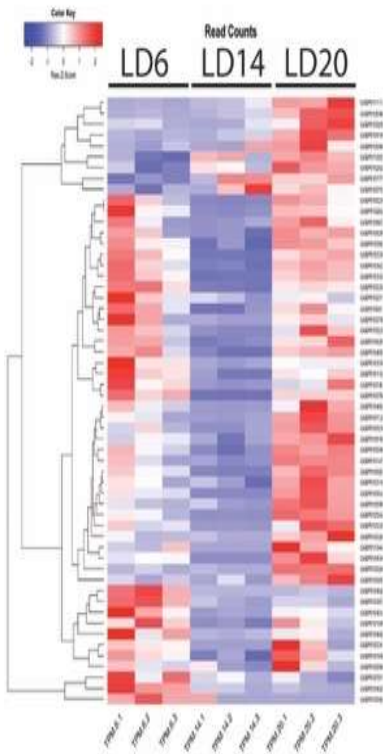
Figure 3. 2. Enzymes involved in carbon metabolism are regulated around LD0.

A schematic view of carbon fixation (from KEGG map 00710) showing enzymes whose translation is regulated as red circles, and metabolites as grey squares (A). Heat maps show that all enzymes are concurrently regulated in the three time (B) and in the 12 times experiment (C) and that increased synthesis precedes lights-on. A schematic view of glycolysis/gluconeogenesis (from KEGG map 00010) shows enzymes whose translation is regulated in red and enzymes found in the transcriptome but whose synthesis is unregulated in yellow (D). A reversible reaction catalyzed by a pyrophosphate-fructose-6-phosphate phosphotransferase is marked by a red asterisk in both the schema and the heat map at bottom. The heat maps for the 3 times (E) and 12 times (F) experiments show the same timing patterns as for photosynthesis.

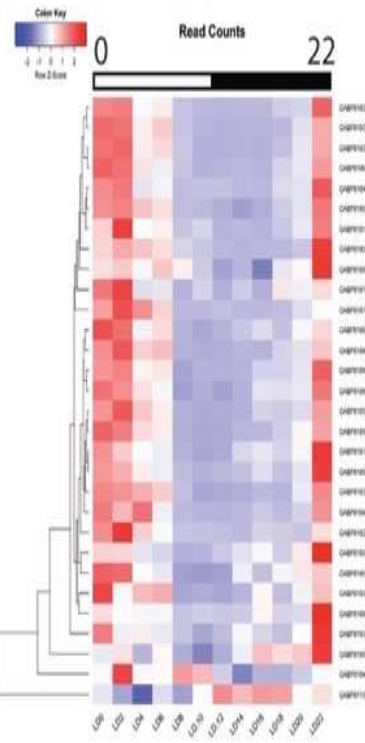
A



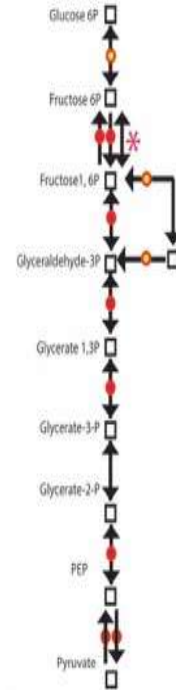
B



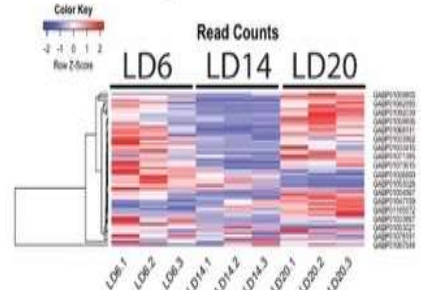
C



D



E



F

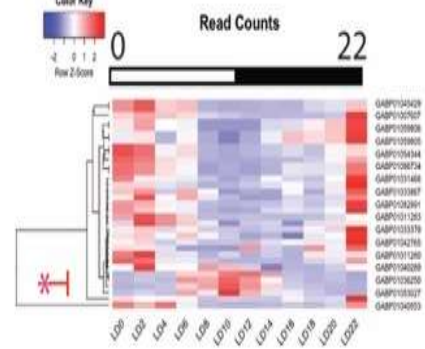


Figure 3. 3. Known night phase rhythms are concurrently regulated around LD12.

Synthesis rates for all transcripts encoding the two bioluminescence proteins are maximal at the onset of the night phase (A, B). A schematic view of DNA replication shows enzymes found in the transcriptome as rectangles, with colored rectangles representing regulated sequences (C). The timing of translation for all sequences with significant differences involved in replication and repair is similar in both the three times and twelve times experiments (D, E). A schematic view of the TCA cycle (from KEGG map 00020) shows metabolites as white squares and regulated enzymes as red circles (F). The timing of transcript translation for both experiments (G, H) is similar to DNA replication.

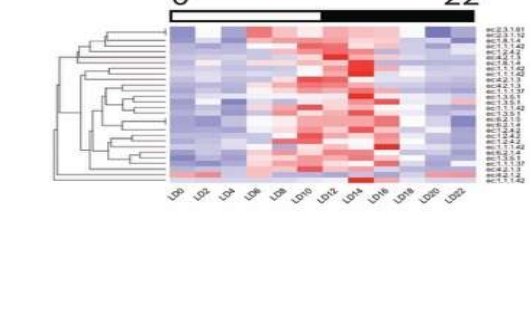
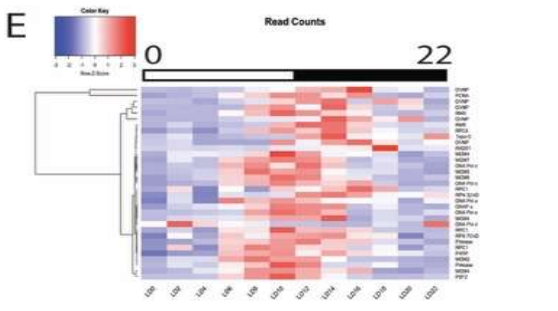
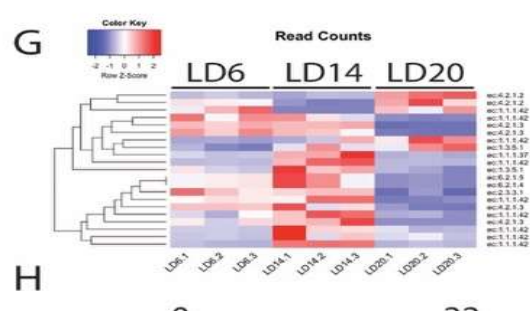
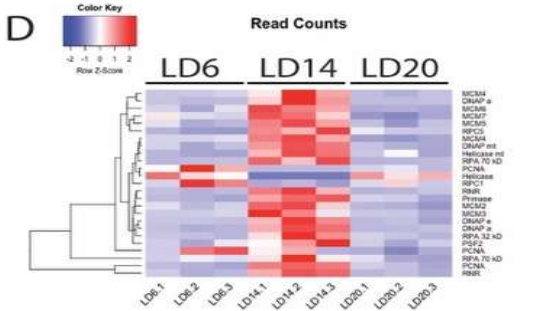
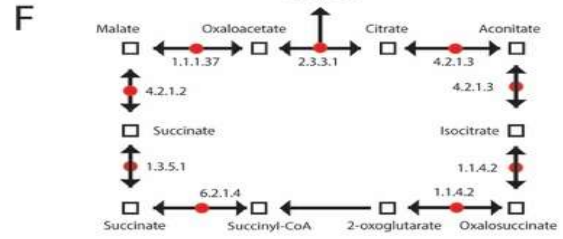
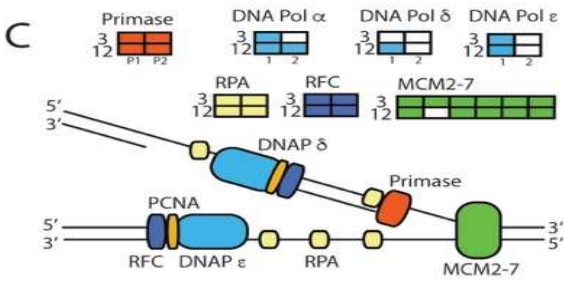
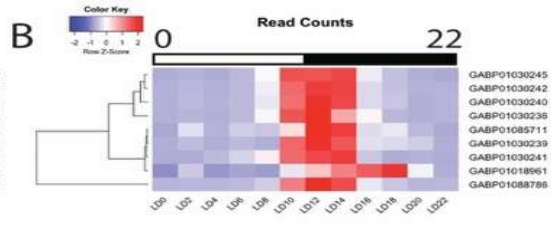
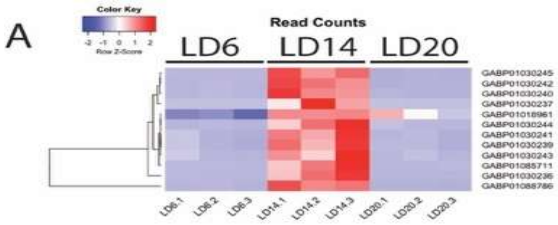


Figure 3. 4. Enzymes involved in regulation of gene expression and amino acid biosynthesis are translated around LD16.

Enzymes involved in different aspects of gene expression were identified among regulated enzymes by GO annotations in both experiments. All appear regulated at a similar time close to the peak of proteins regulated at LD16.

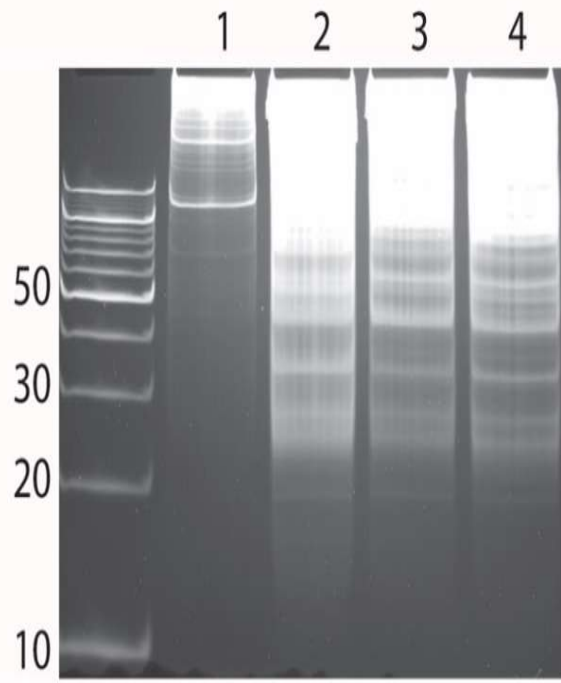




Figure 3. 5. (Supplementary) Electrophoretic analysis of RPF.

A) RNA 10% acrylamide gel of an RNA sample before (lane 1) and after (lanes 2 -4) RNase If digestion. B) Bands at 30 nucleotides were excised from gels (as in A) for each of the nine samples (ZT6, lanes 1-3; ZT14, lanes 4-6; ZT20, lanes 7-9) and rerun to validate sample integrity before preparing samples for sequencing. A sample before digestion is shown in lane 10. Molecular weight markers (Lane 1 in A, lane 11 in B) are in 10 base steps from bottom.

A



B

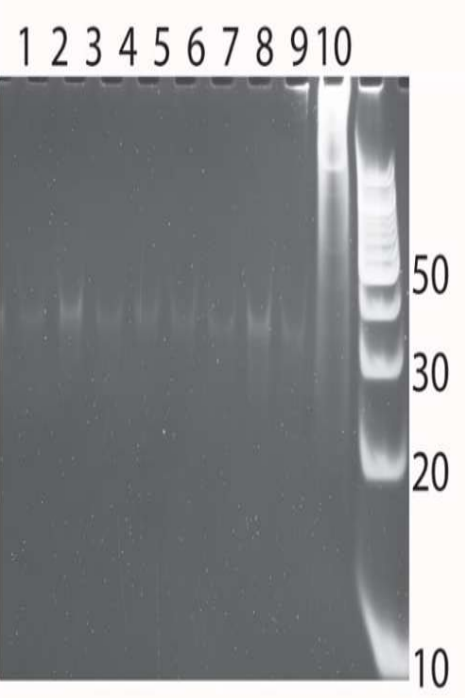


Figure 3. 6. (Supplementary) Size distribution of Ribosome Protected Fragments (RPF).

Sequences were trimmed to remove sequencing adaptors, and the proportion of sequences of different lengths recorded in each of the nine samples.

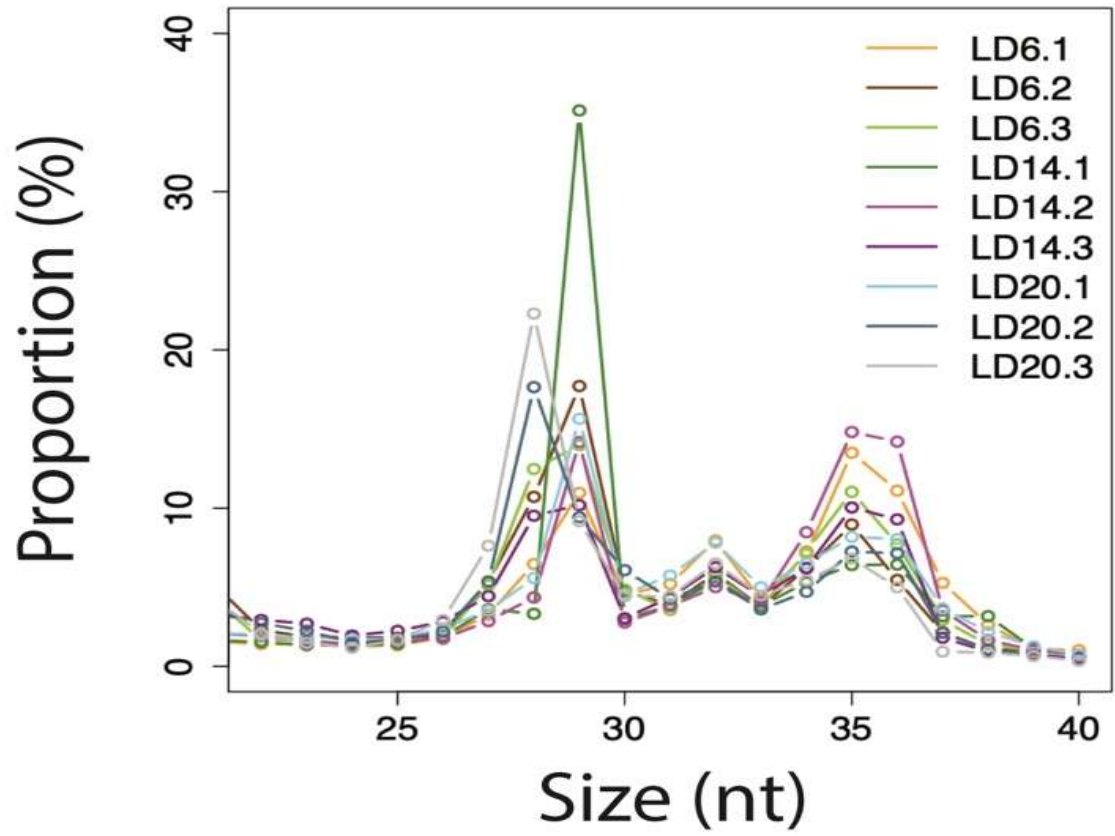


Figure 3. 7. (Supplementary) Similarity of triplicate samples to one another.

RPF read counts obtained after mapping to the transcriptome were tested for similarity to all other samples using correlation coefficients (left) and principal components analysis (right).

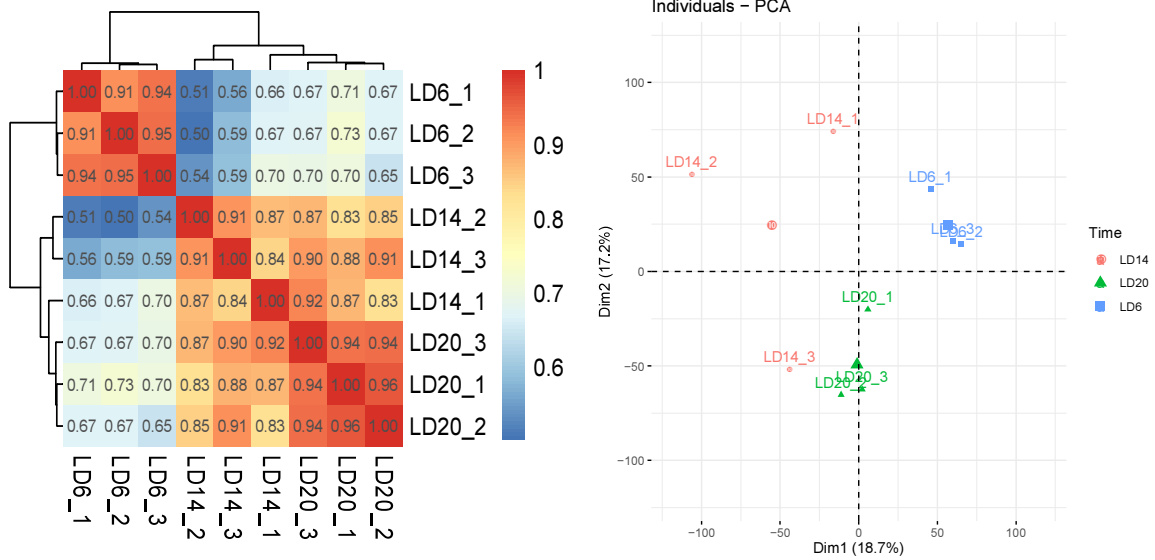


Figure 3. 8. (Supplementary) Read distribution for LBP.

Reads corresponding to triplicate LBP samples (top, LD6; middle, LD14; bottom, LD20). At LD6 and LD20, when LBP synthesis is proscribed, most of the read counts correspond to a single area within the coding sequence, whereas at LD14, when there is active translation, reads are spread out over the length of the coding sequence. Dark green indicates sequence identity among all reads. Total reads are indicated at right.

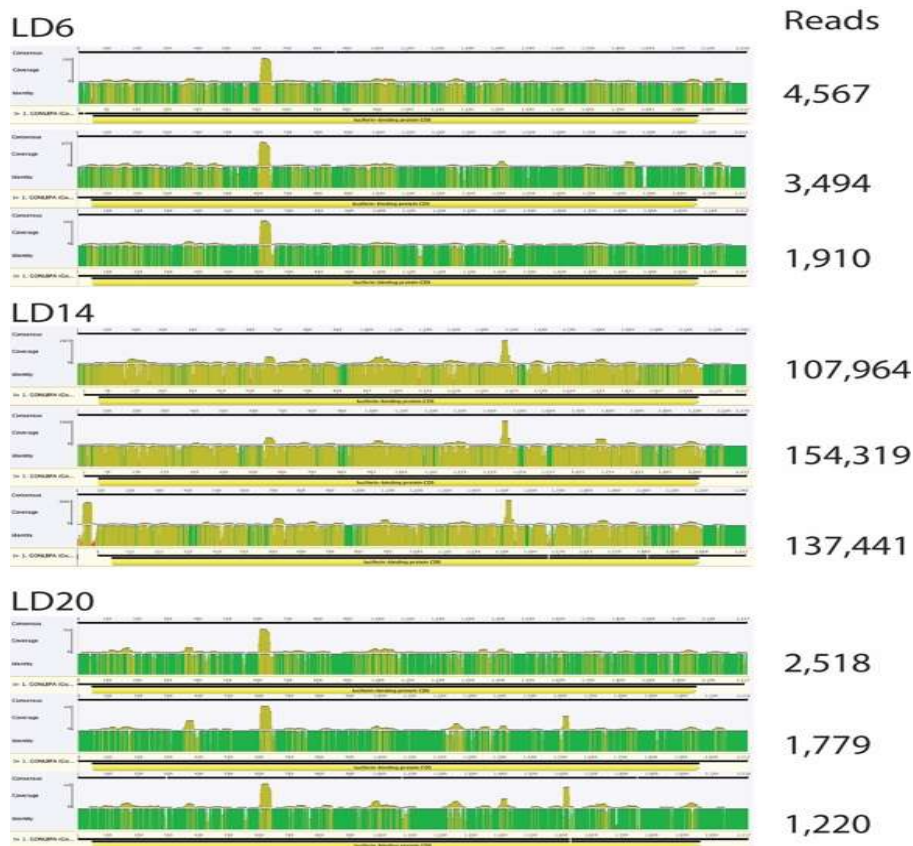


Figure 3. 9. (Supplementary) Comparison of RPF read counts and  $^{35}\text{S}$ -methionine *in vivo* labeling.

Protein synthesis rates for four different proteins, measured by incorporation of radiolabeled methionine *in vivo* (green lines) in a different *Lingulodinium* strain under LL conditions (Markovic, Roenneberg et al. 1996), compared to RPF read counts measured under LD (orange lines).



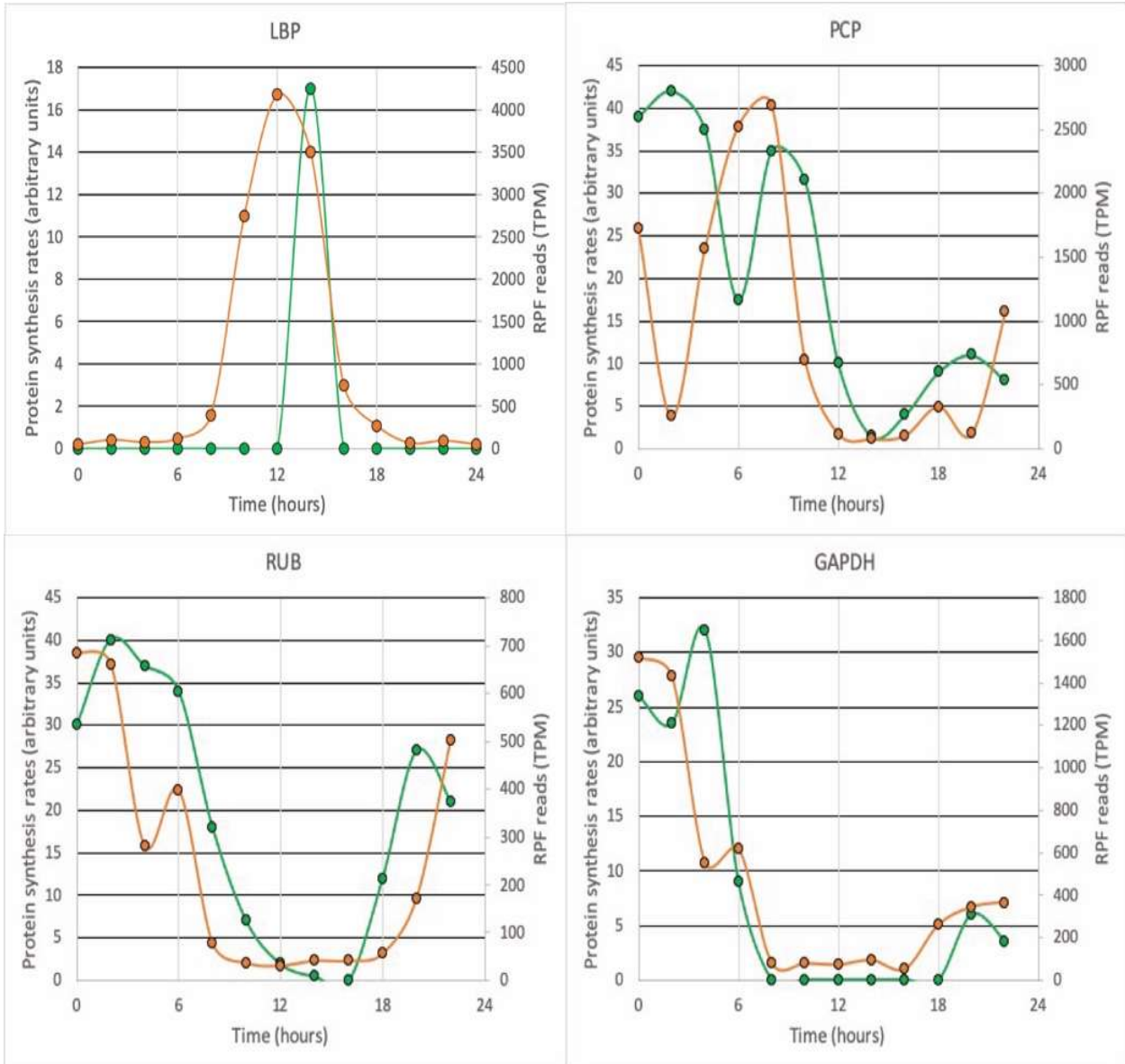


Figure 3. 10. (Supplementary) Comparison of RPF read counts with diurnal variations in protein levels.

RPF read counts (orange bars, 3 times; orange lines, 12 times) were superimposed on protein levels previously determined by LS-MS/MS (blue lines) for thirteen proteins previously shown to vary over the daily cycle (Bowazolo, Tse et al. 2020).

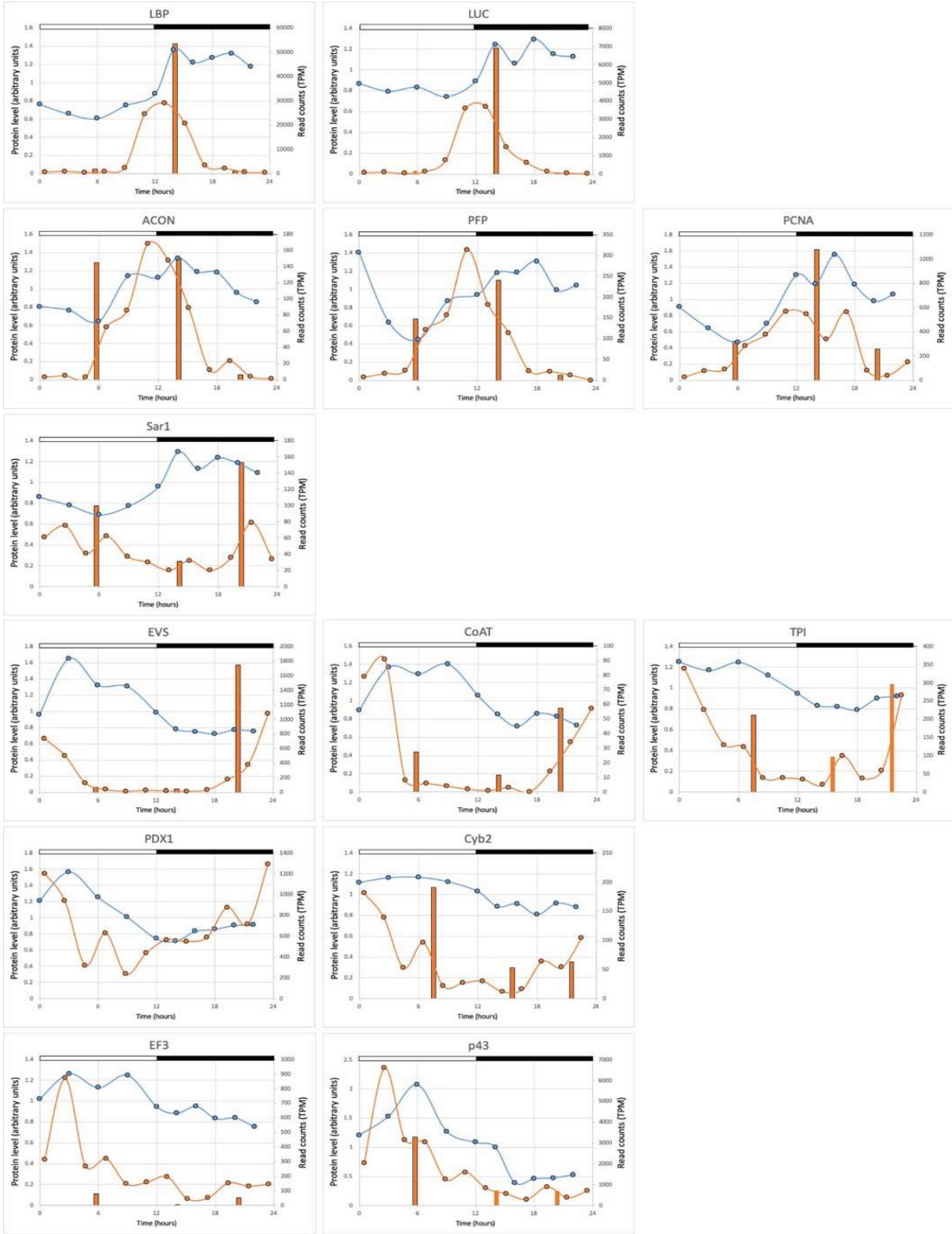


Figure 3. 11. (Supplementary) Principal GO categories for enriched RPF at LD6.

Sequences found to be significantly enriched using Fisher's exact test for the ZT6 samples were placed into GO categories. An arbitrary cutoff of 5 sequences was applied to simplify the presentation.

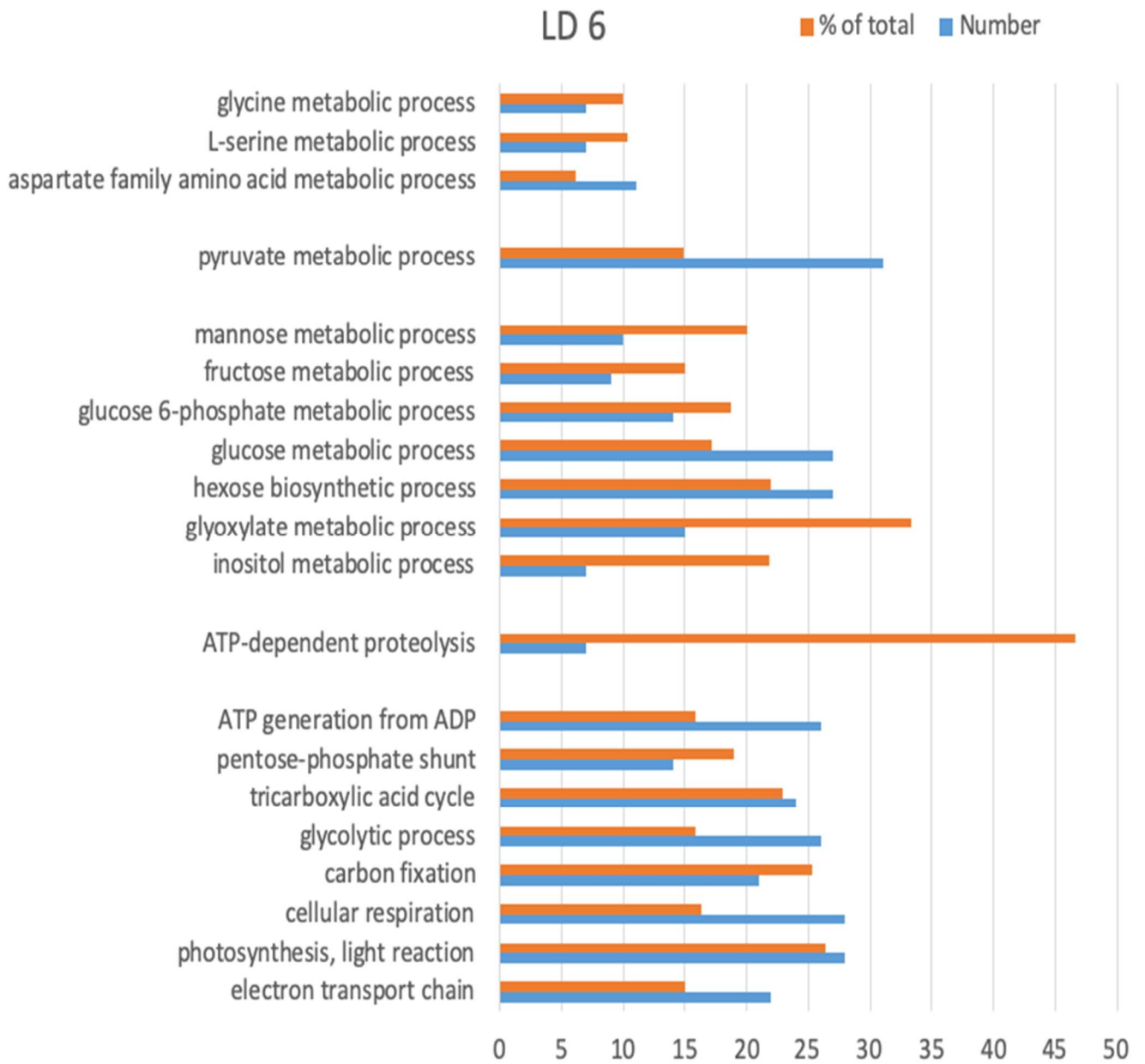


Figure 3. 12. (Supplementary) Heat maps for proteins related to the light reaction of photosynthesis.

Heat maps showing combined photosystem proteins, light harvesting proteins, and ferredoxin in the three times (left) and twelve times (right) experiments.

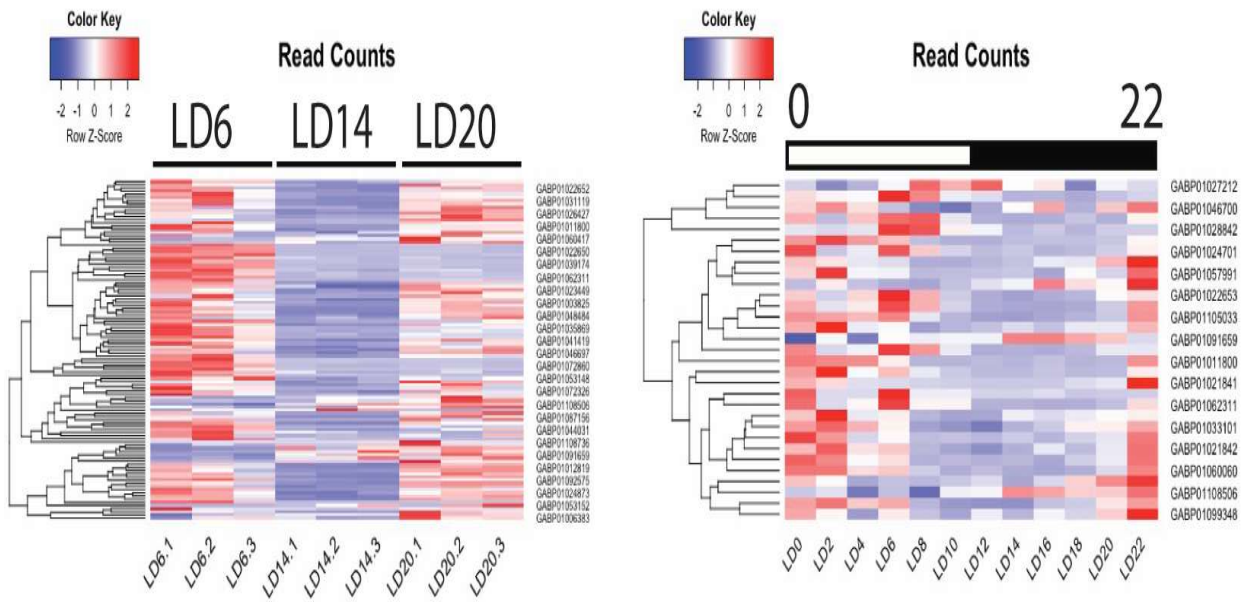


Figure 3. 13. (Supplementary) Pentose phosphate shunt.

Heat maps of pentose phosphate enzyme expression show two different patterns in both the three times (bottom left) and the 12 times (bottom right) experiment. The enzyme with peak expression around LD22 are only involved in the pentose phosphate pathway (red circles in upper and bottom right). The enzymes interconverting fructose-6-P to fructose -1,6-diP (1-phosphotransferase, EC 2.7.1.90) shows peak expression around LD10.

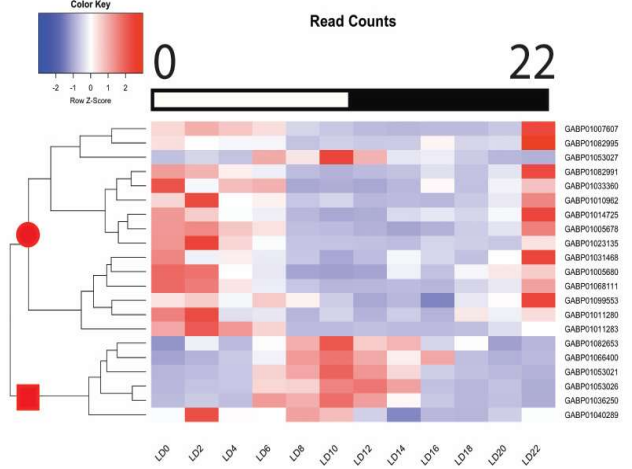
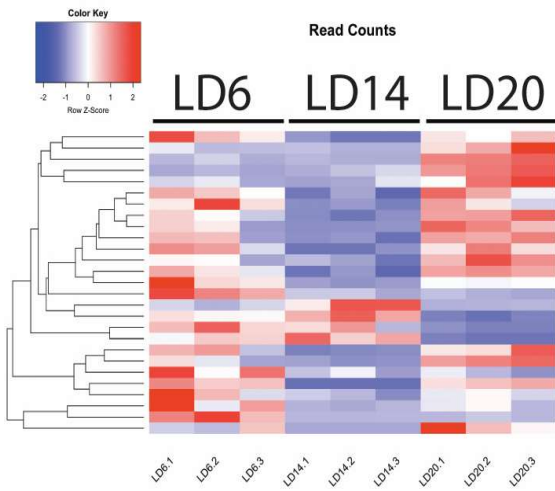
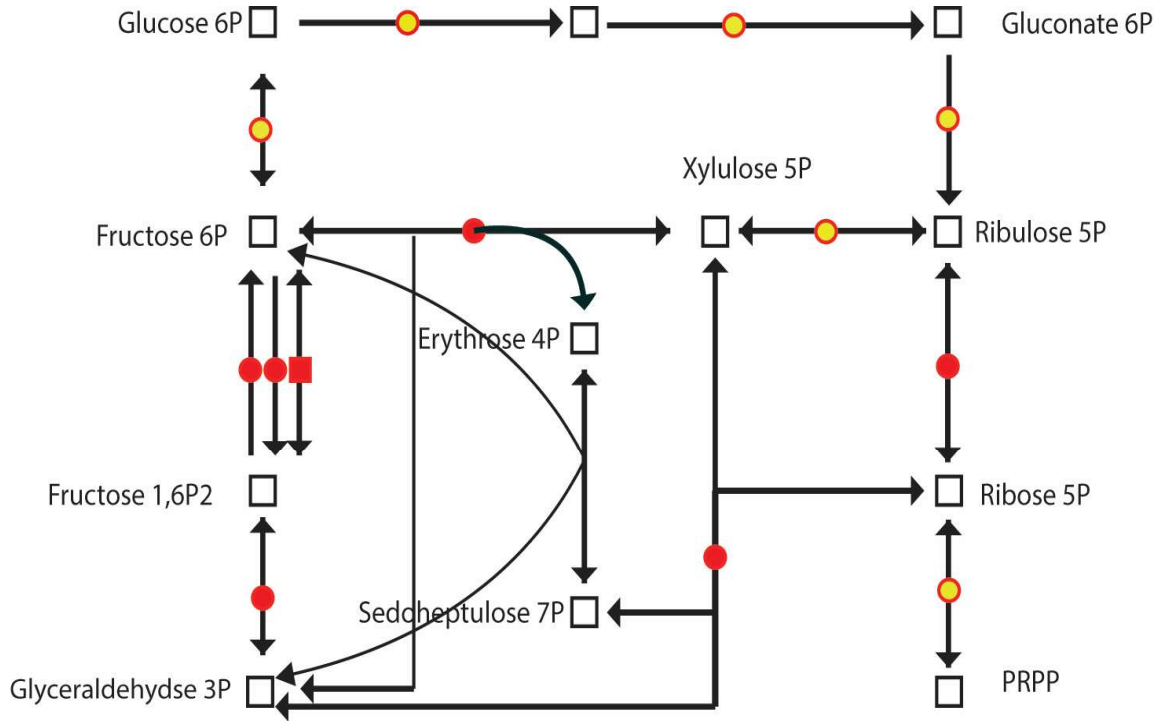






Figure 3. 15. (Supplementary) Starch metabolism.

Heat maps of expression in the three times (bottom left) and the twelve times (bottom right) experiments show two groups of enzymes. One group, whose synthesis occurs around midday, contains two enzymes used for starch biosynthesis. The other group, with peak expression later during the day, contains enzymes used for starch degradation and conversion of UDP-glucose to glucose-1-P.

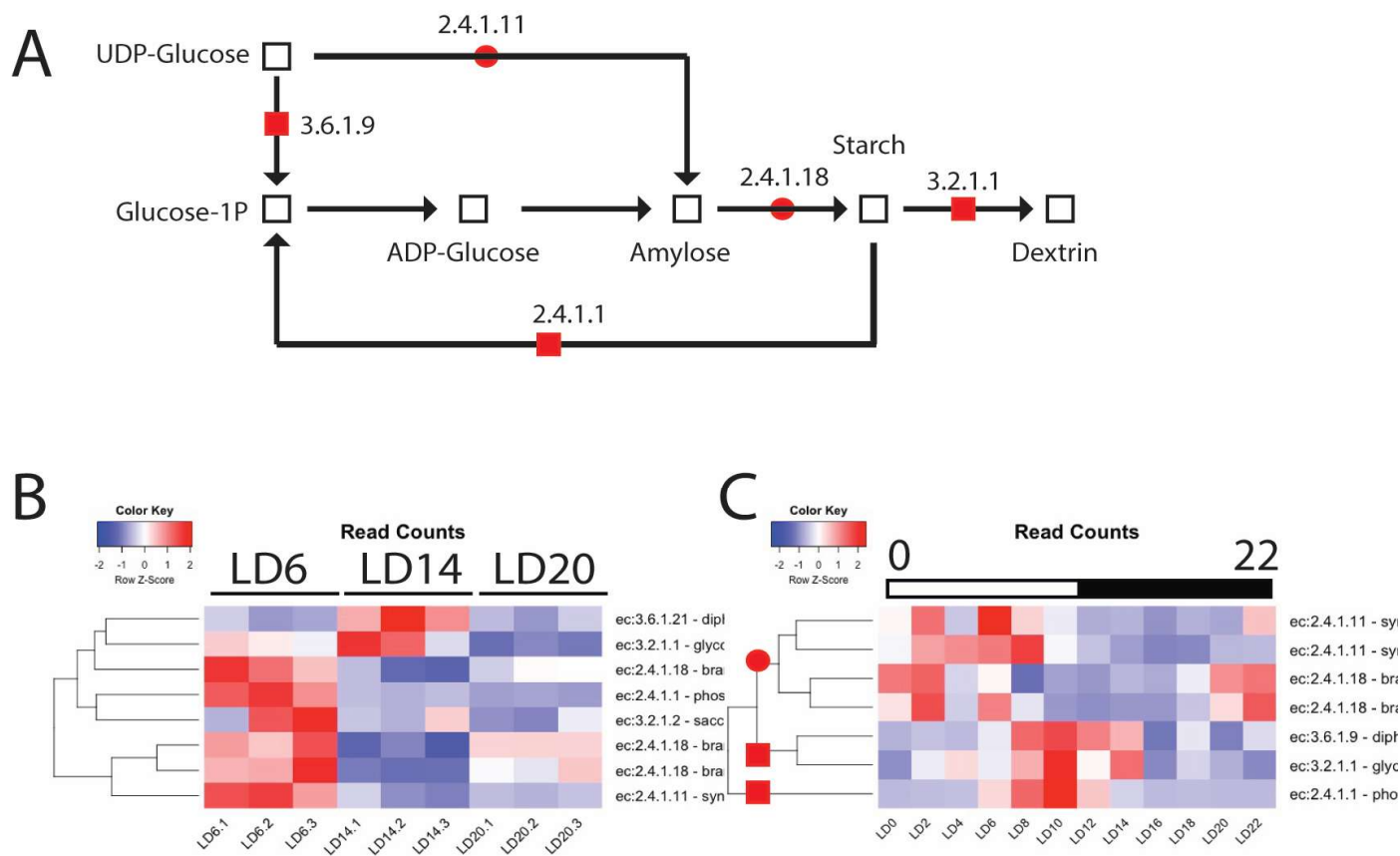


Figure 3. 16. (Supplementary) Biosynthetic pathways for valine, leucine and isoleucine.

Schema of the biosynthetic pathways, with intermediates marked in grey and reactions catalyzed by enzymes annotated by EC nomenclature. Boxed enzymes are found in the transcriptome, shaded enzymes have higher read counts at LD20 than at least one of the other two times. (B) Average read counts (n=3) for all regulated enzymes.

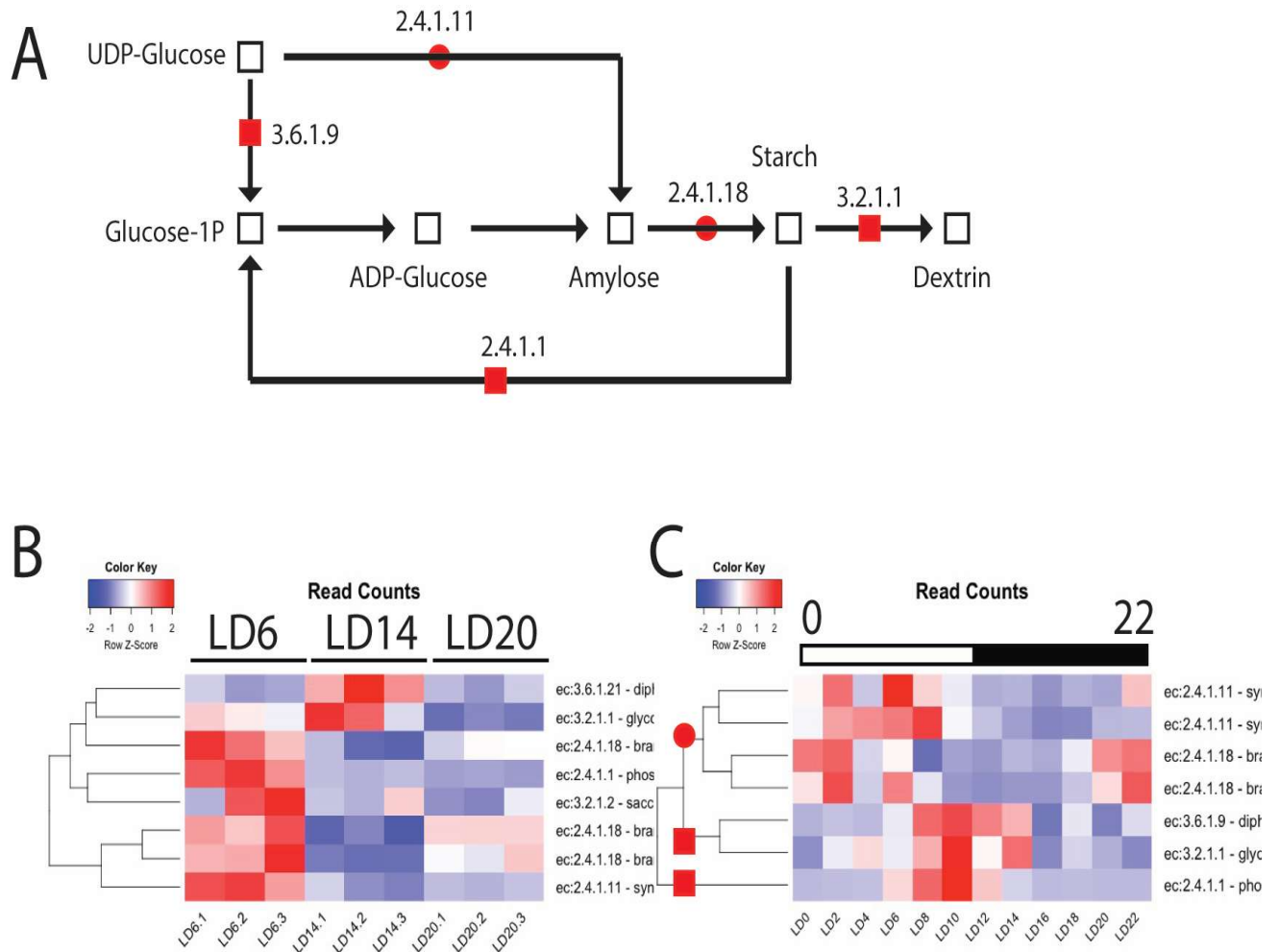


Figure 3. 17. (Supplementary) Cellulose metabolism.

Heat maps of expression in the three times (left) and the twelve times (right) experiments show most proteins have peak expression late during the night. This pattern is consistent with a requirement for forming the cellulosic thecal plates which are constructed during mitosis around LD1.

## Cellulose metabolism

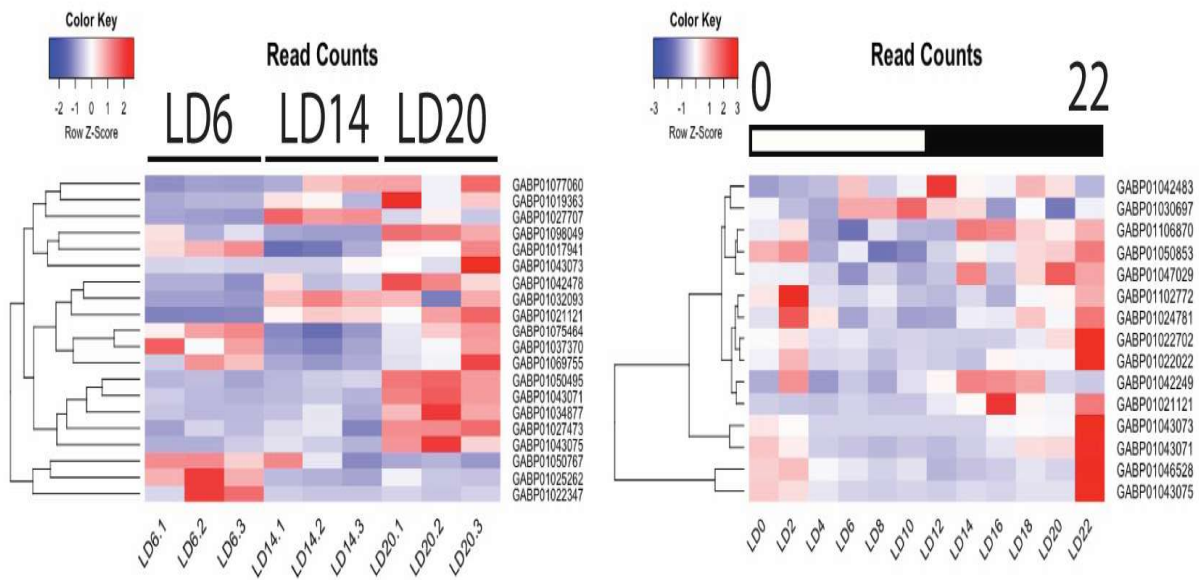


Figure 3. 18. (Supplementary) Fructose/Mannose Metabolism.

Heat maps of pentose phosphate enzyme expression show two different patterns in both the three times (bottom left) and the 12 times (bottom right) experiment. The enzyme with peak expression around LD22 are only involved in the pentose phosphate pathway (red circles in upper and bottom right). The enzyme interconverting fructose-6-P to fructose -1,6-diP (PFK, EC 2.7.1.90, red square) has a different pattern of protein synthesis in both experiments (as in figures 3, S9).

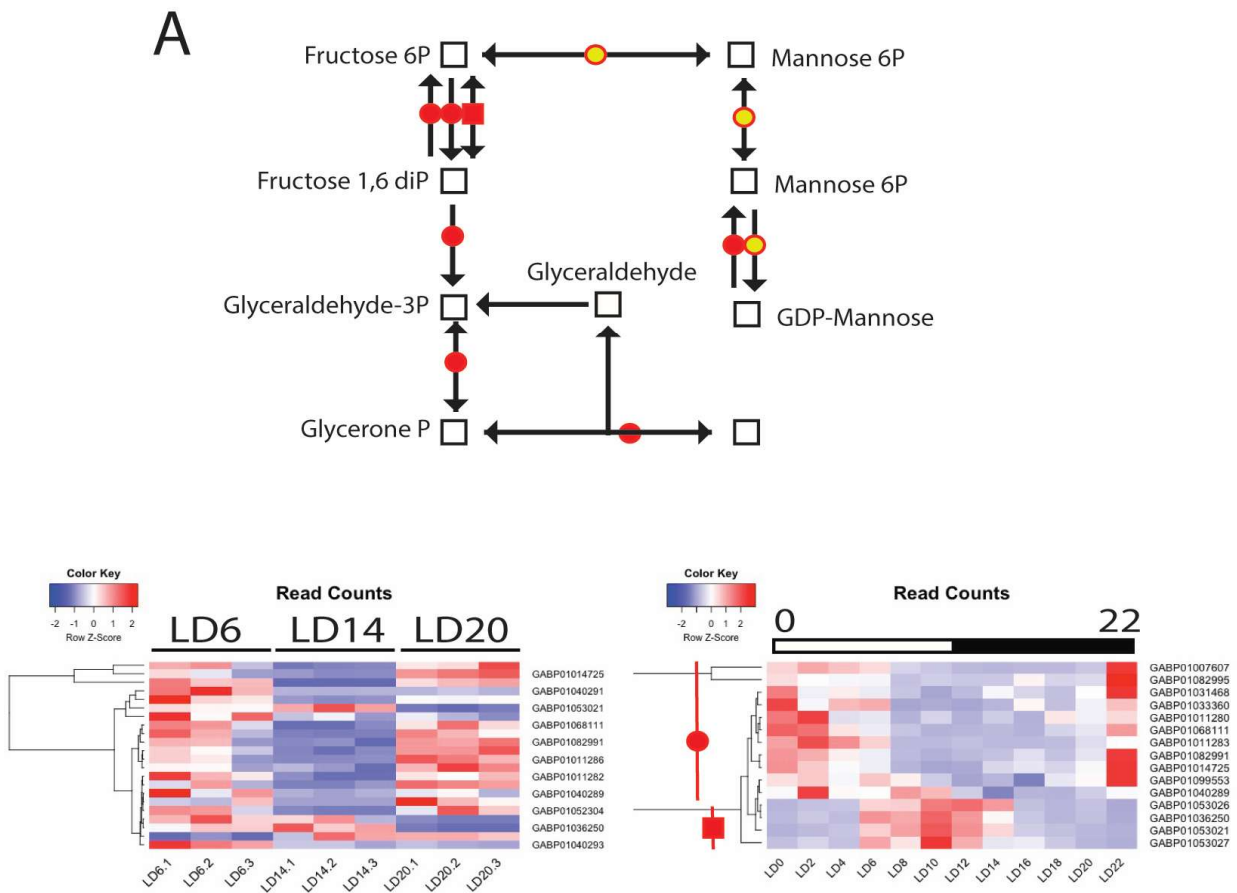


Figure 3. 19. (Supplementary) Principal GO categories for enriched RPF at LD14.

Legend as for Figure 3 S11.

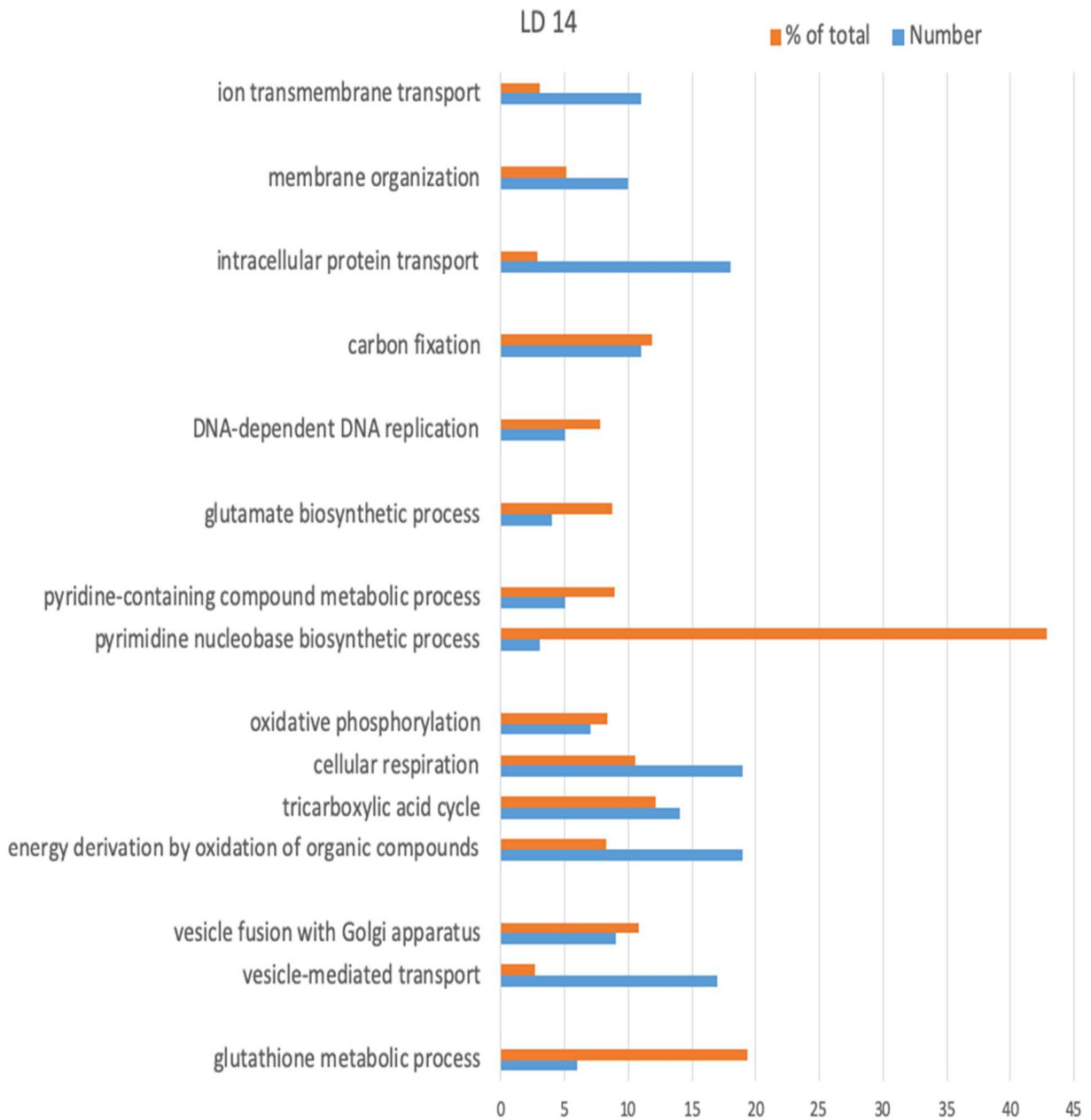


Figure 3. 20. (Supplementary) Pyruvate metabolism.

Heat maps for the 3 times (bottom left) and 12 times experiment (bottom right) show two patterns of protein synthesis, one occurring around LD10 and the other around LD22. While all enzymes found are involved in pyruvate metabolism, only those uniquely involved in pyruvate metabolism show a peak at LD10. Enzyme with a peak expression around LD22 can also be involved in glycolysis, the TCA cycle, glyoxylate metabolism or fatty acid biosynthesis.

	Pyruvate	Glycolysis	TCA cycle	Glyoxylate	FA Synth
<b>LD10</b>					
Maltate dehydrogenase	+				
Pyruvate carboxylase	+				
Dihydrolipoate transferase	+				
Dihydrolipoate Dehyd	+				
<b>LD22</b>					
Pyruvate kinase	+	+			
PEP synthase	+	+			
Fumarate hydratase	+		+		
Acyl CoA synthase	+			+	
Acetyl CoA carboxylase	+				+

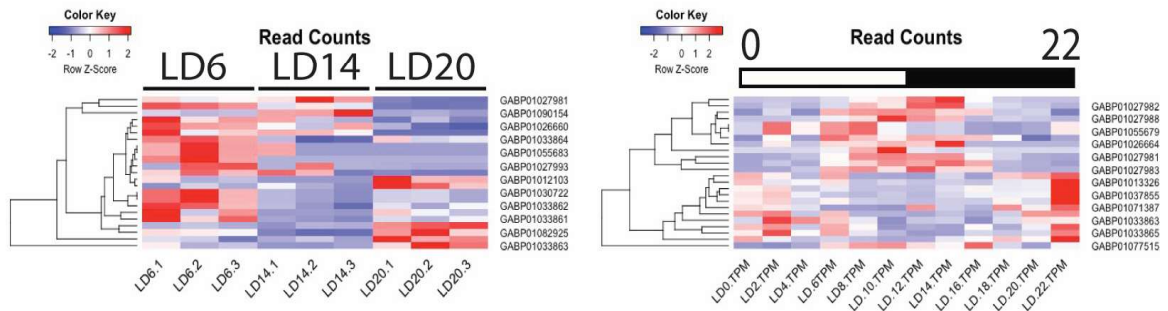


Figure 3. 21. (Supplementary) Purine Metabolism

A schematic view of the enzymes involved in purine biosynthesis shows both unregulated enzymes that are found in the transcriptome (yellow circles) and those whose synthesis is regulated (red circles) (A). The three times experiment (B) contains many more sequence synthesized during the day, while the twelve times experiment (C) contains more transcripts translated at the end of the day.

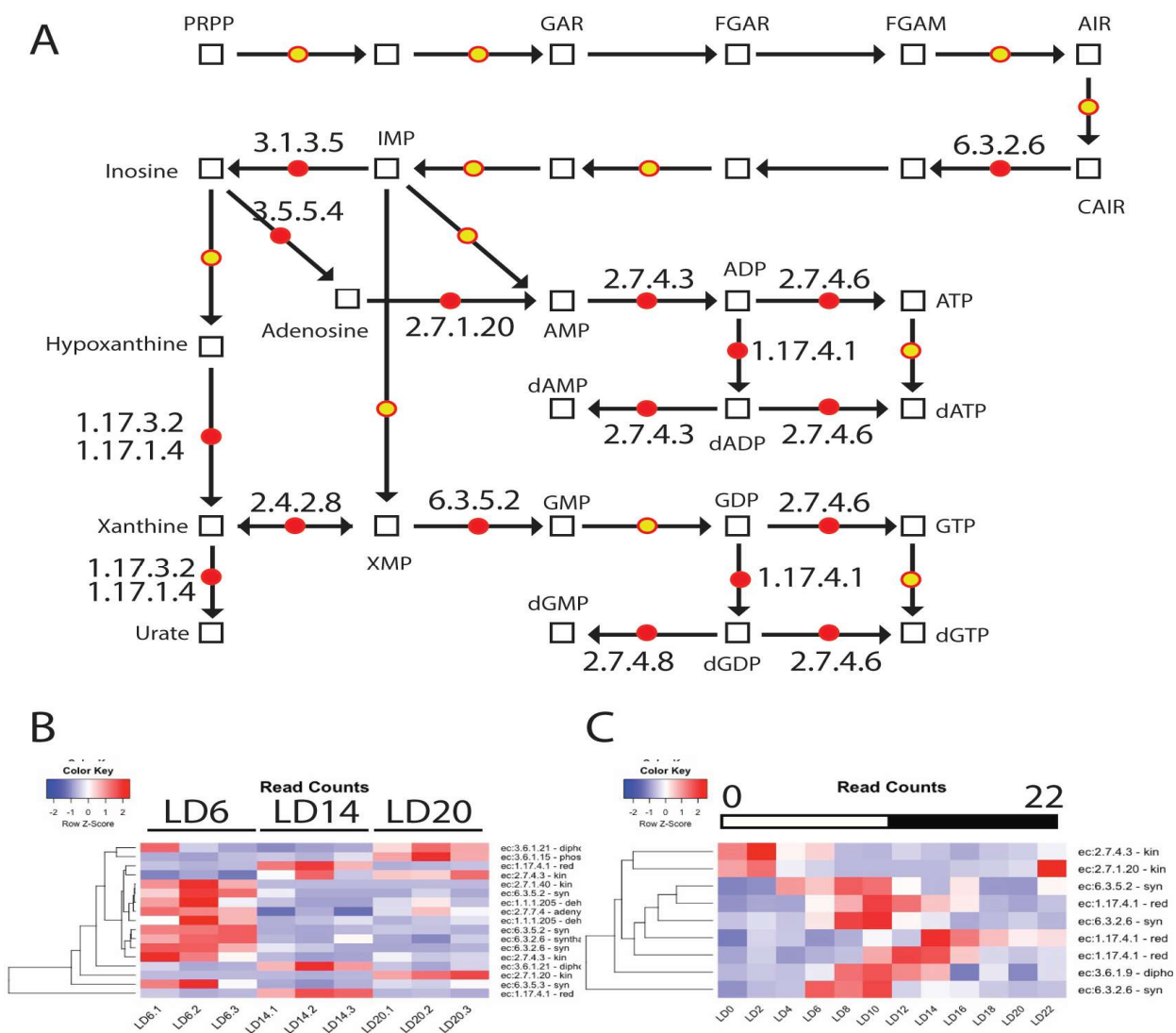






Figure 3. 23. (Supplementary). Principal GO categories for enriched RPF at LD20.

Legend as for Figure 3 S11.

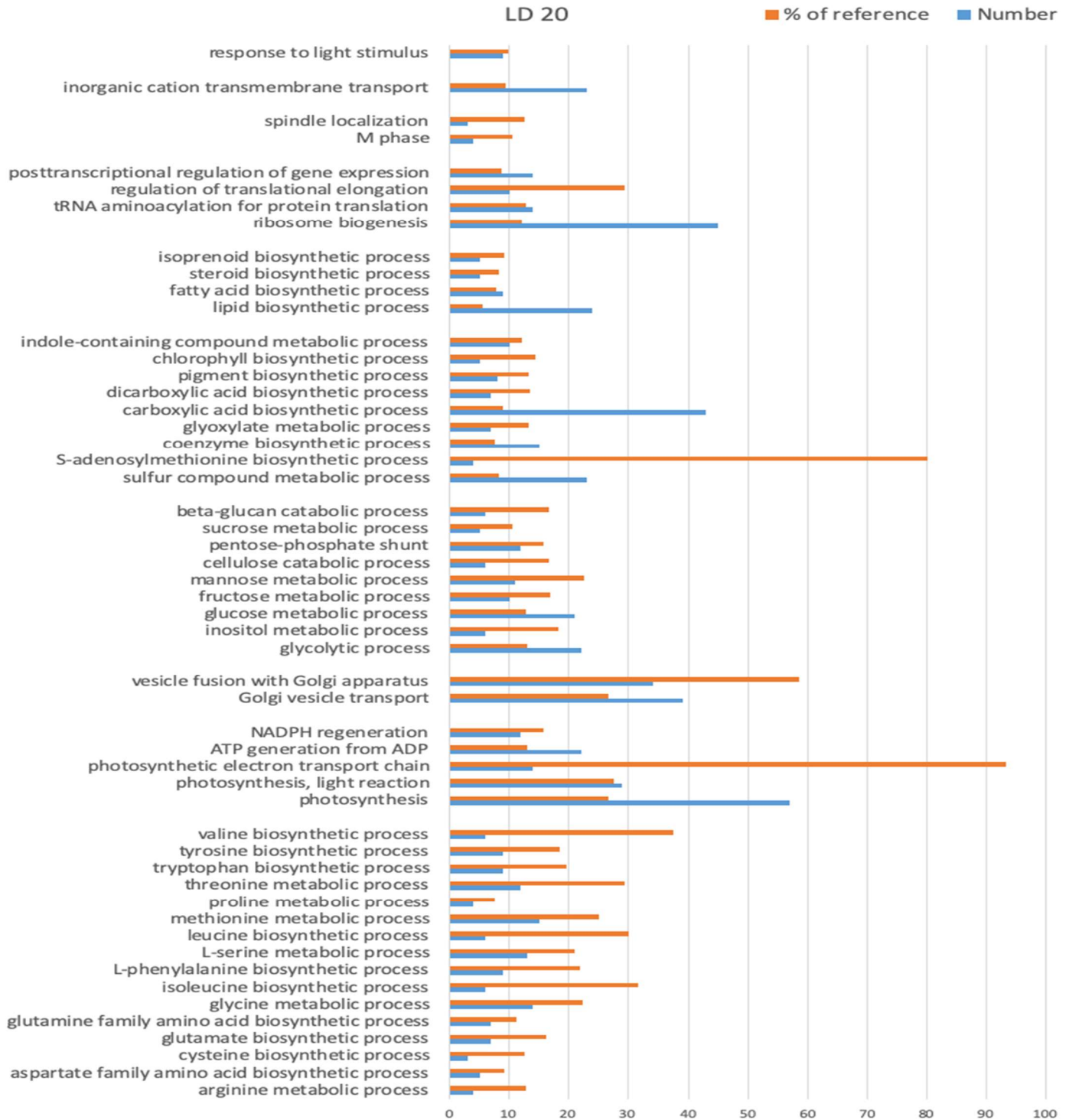


Figure 3. 24. (Supplementary) Amino acid biosynthetic enzymes regulated in the 3 times experiment.

A schematic view of amino acid biosynthesis (redrawn from KEGG map 01230). Non-regulated enzymes that are found in the transcriptome are shown as yellow circles, while regulated enzymes in the transcriptome are shown as red circles. Amino acids are shown as gray squares and other intermediates as white squares. A virtually identical pattern of regulated enzymes is seen in the 12 times experiment.

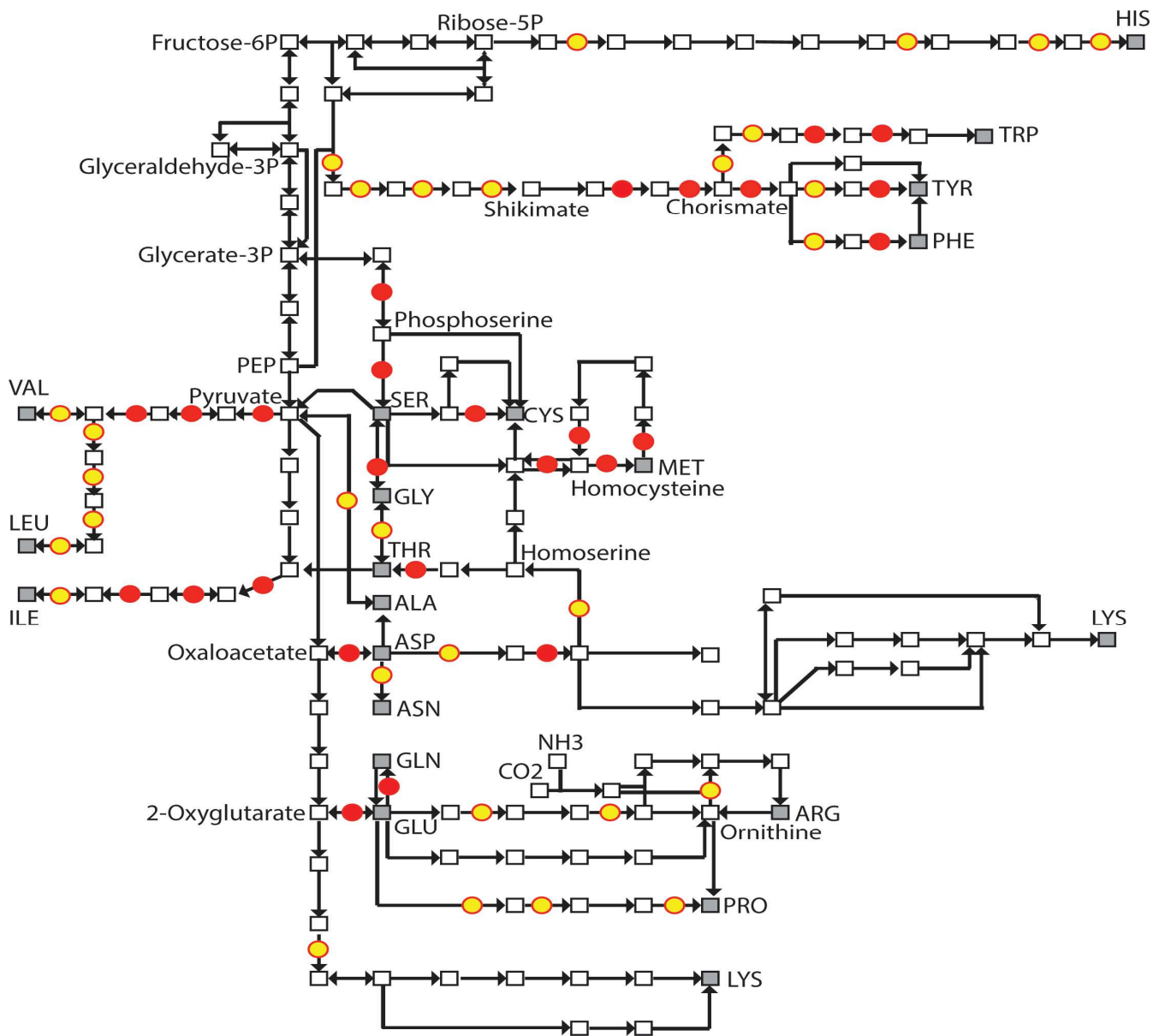


Figure 3. 25. (Supplementary) Amino acid profiles over a daily cycle shows increases during the dark phase.

Average levels of amino acids (blue curves) were measured at four times over the daily cycle (n=5, average error 15%) with LD0 values duplicated at LD24. Average read counts (TPM, red curves) were determined from the three time experiment (n=3, average error 20%) for KEGG pathway enzymes involved in biosynthesis of the amino acids present in the transcriptome. Amino acid levels marked with an asterisk are significantly different from the levels at LD0 (p<0.05, Student's t-test).

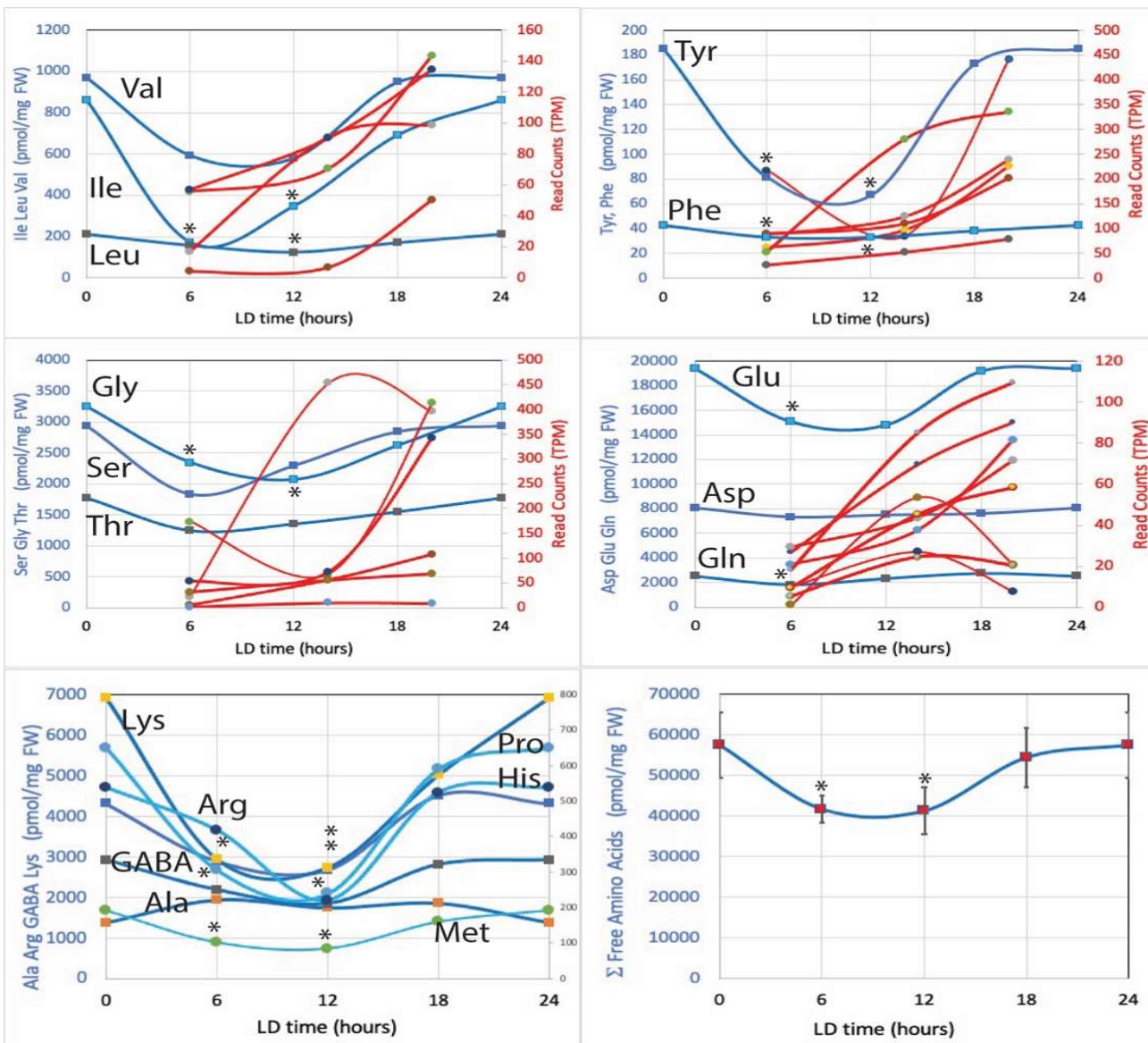


Figure 3. 26. (Supplementary) Correlation between RPF read counts with transcript abundance.

RPF read counts per million as a function of read counts in the transcriptome are shown for the entire transcriptome (blue) or 3324 sequences with significant differences in RPF counts (orange) at one of the three times. All RPF counts are the average of the three samples.

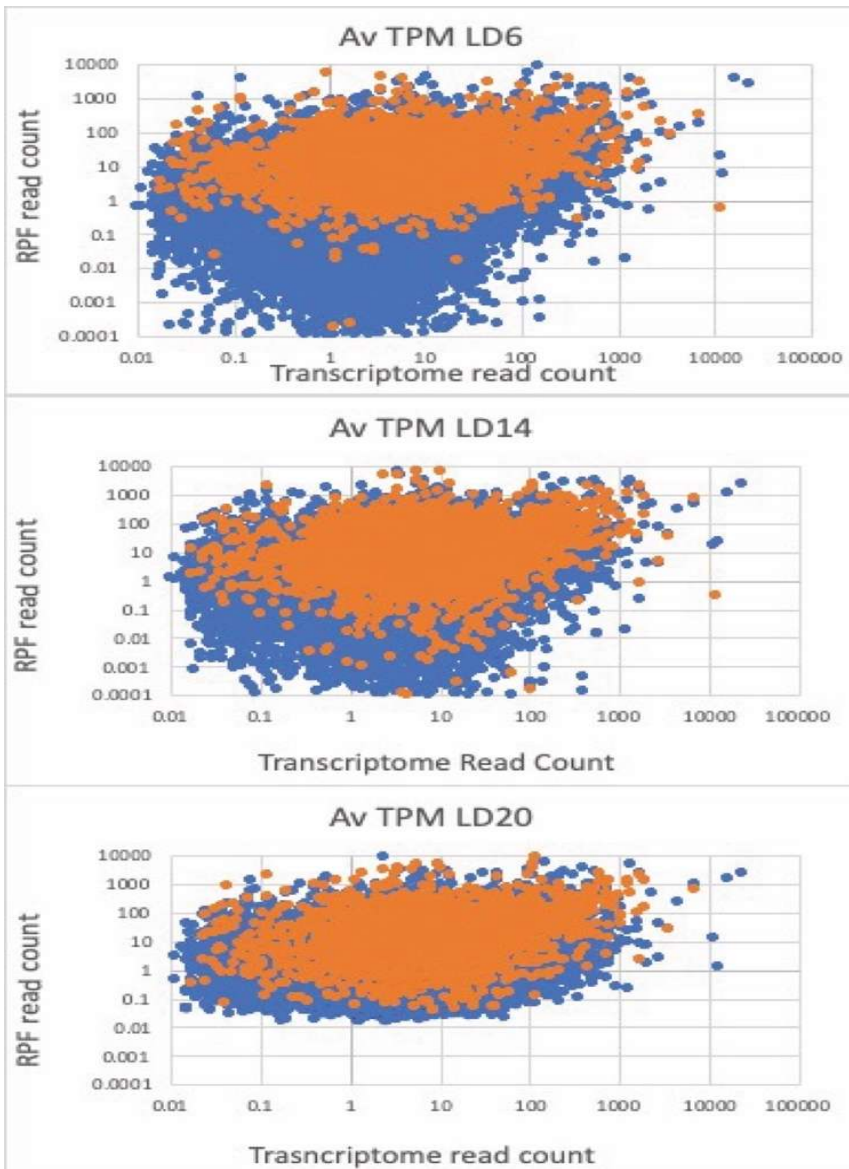


Tableau 3. 1. Number of reads before and after trimming (test run) with an LD14 sample.

	RNase If	RNase T1
Clipped reads	89,513	85,088
Hits (Trinity)	38,437	60,210
hits (LBP)	568	122
Hits (LUC)	101	28
hits (PCP)	3	5
Hits (Rub)	5	45
Hits (rRNA)	13,210	12,480

Tableau 3. 2. Number of reads before and after trimming (3 times).

	# reads	Trimmed reads	Hits (Trinity)	rRNA	mRNA	% mRNA reads
ZT6	23,381,155	22,232,069	7,410,028	5,531,668	1,878,359	8.4%
	22,215,960	20,938,357	8,252,582	6,234,990	2,017,592	9.6%
	20,843,391	19,572,370	8,573,293	6,786,326	1,786,967	9.1%
ZT14	23,231,492	21,007,233	6,483,441	4,587,817	1,895,624	9.0%
	34,922,729	32,012,000	8,425,324	6,640,948	1,784,376	5.6%
	27,072,018	19,661,514	7,434,870	5,643,890	1,790,979	9.1%
ZT20	26,856,028	24,064,518	7,089,700	5,177,901	1,911,799	7.9%
	23,526,421	17,501,882	8,124,526	6,217,352	1,907,173	10.9%
	24,572,042	21,477,100	7,865,513	5,962,238	1,903,275	8.9%



Tableau 3. 3. Number of reads before and after trimming (12 times).

	Reads	Trimmed Reads	Hits (Trinity)	rRNA	mRNA	mRNA (%)
LD0	21,553,584	19,163,757	5,047,852	3,746,476	1,301,365	6.8%
LD2	23,718,663	21,210,042	5,879,254	4,628,771	1,250,475	5.9%
LD4	16,899,542	14,367,203	3,835,450	3,218,182	617,261	4.3%
LD6	21,340,133	18,566,160	4,145,142	3,210,507	934,627	5.0%
LD8	18,819,673	17,315,642	3,874,231	3,090,970	783,255	4.5%
LD10	22,829,402	20,361,492	4,272,749	3,096,062	1,176,683	5.8%
LD12	23,617,205	20,366,949	3,606,551	2,717,638	888,910	4.4%
LD14	22,839,328	18,594,219	4,678,217	3,819,666	858,546	4.6%
LD16	20,519,326	18,375,565	3,478,858	2,848,854	629,996	3.4%
LD18	26,738,718	23,408,848	5,222,948	4,127,609	1,095,305	4.7%
LD20	21,986,161	19,308,487	7,006,718	5,987,154	1,019,543	5.3%
LD22	17,493,489	14,748,216	3,439,016	2,781,260	657,756	4.5%

Tableau 3. 4. Average RPF counts for diurnally regulated proteins at three times.

	ZT6	ZT14	ZT20
Luciferase Binding Protein	1944	53630	986
Luciferase	185	6912	95
Proliferating cell nuclear antigen	322	1077	257
Aconitase	145	151	7
Pyrophosphate F6P phosphotransferase	148	240	12
Small GTP-binding protein SAR1	99	31	153
p43	3287	730	708
Cytochrome b2	95	29	153
2-epi-5-epi-valiolone (EVS)	69	48	1746
Acetyl-CoA transferase	27	11	57
Elongation factor 3 (EF3)	79	6	52
Triose phosphate isomerase (TPI)	211	97	296
PDX1	1188	1825	2204
Peridinin-Chlorophyll a-binding Protein (PCP)	123770	775	1494
Rubisco	24906	1331	10789
Glyceraldehyde-3-phosphate dehydrogenase (G3PDH)	6513	653	1336
Oxygen-evolving enhancer protein 1 (OEE)	5599	342	1944

Tableau 3. 5. Consensus motifs detected in transcripts whose synthesis is highest near LD0, LD12, or LD16.

Motif	Consensus	Length	# Sites	Pvalue	Zscore
LD23_Motif-6	GCCGCC	6	2932	3.33e-48	3.18
LD23_Motif-13	GGCCGC	6	2461	4.35e-45	1.50
LD23_Motif-4	GCGGCG	6	774	3.11e-26	3.61
LD23_Motif-1	CGCCGG	6	1644	1.76e-23	6.50
LD23_Motif-2	CCGCCG	6	609	1.13e-23	3.66
LD23_Motif-9	CGGCCG	6	953	5.60e-22	2.34
LD23_Motif-14	GGCGCC	6	1518	1.58e-18	1.39
LD23_Motif-17	CGCGGC	6	5144	7.32e-18	3.81
LD23_Motif-27	CCGCCG	6	9397	3.04e-14	2.53
LD23_Motif-7	CCGGCC	6	496	5.86e-10	2.93
LD0_Motif-1	RTGGCC	6	1881	3.16e-29	5.19
LD0_Motif-2	TGGCCA	6	2502	2.14e-17	4.42
LD0_Motif-3	CCATGG	6	1338	3.46e-16	4.62
LD0_Motif-12	CGCCGC	6	1016	7.29e-12	0.43
LD0_Motif-11	GGCGCC	6	2405	1.33e-06	-3.06
LD2_Motif-19	AGGAGG	6	4533	1.36e-126	3.09
LD2_Motif-7	CCGCCG	6	7119	1.71e-28	1.61
LD2_Motif-6	CGGCGG	6	841	6.31e-26	1.61
LD2_Motif-1	CTTGAG	6	657	1.71e-14	4.69
LD2_Motif-10	GCCATG	6	2815	1.69e-21	1.26
LD2_Motif-16	CGCCGC	6	3570	1.62e-08	-1.03
LD2_Motif-3	GCCACG	6	3908	9.67e-08	0.34
LD12_Motif-4	GCCATG	6	4694	2.53e-44	4.80
LD16_Motif-4	GCCATG	6	9521	1.88e-34	1.68
LD16_Motif-1	CCGCCG	6	13131	2.84e-08	-1.94

Tableau 3. 6. Amino acid levels (pmol/mg FW, n=5) at four times in an LD cycle.

<b>Amino Acid</b>	<b>LD0</b>	<b>LD6</b>	<b>LD12</b>	<b>LD18</b>
Alanine	1380 ± 273	1933 ± 163	1741 ± 213	1852 ± 327
Arginine	4305 ± 559	2881 ± 227	2665 ± 341	4518 ± 746
Aspartate	8058 ± 1053	7351 ± 527	7513 ± 1135	7623 ± 845
GABA	2913 ± 575	2178 ± 643	1842 ± 153	2801 ± 339
Glutamate	19385 ± 2401	15070 ± 761	14789 ± 2210	19173 ± 2136
Glutamine	2522 ± 262	1801 ± 165	2329 ± 378	2764 ± 514
Glycine	3246 ± 376	2345 ± 166	2068 ± 277	2621 ± 352
Histidine	537 ± 181	417 ± 117	219 ± 84	523 ± 135
Isoleucine	859 ± 182	171 ± 12	348 ± 38	691 ± 181
Leucine	213 ± 28	156 ± 8	122 ± 9	170 ± 21
Lysine	6908 ± 847	2938 ± 133	2731 ± 186	5012 ± 489
Methionine	191 ± 41	101 ± 16	84 ± 21	159 ± 24
Phenylalanine	107 ± 12	82 ± 6	82 ± 9	95 ± 17
Proline	649 ± 103	306 ± 57	239 ± 66	589 ± 97
Serine	2940 ± 587	1833 ± 118	2299 ± 416	2849 ± 568
Threonine	1766 ± 251	1247 ± 114	1354 ± 55	1547 ± 254
Tyrosine	463 ± 96	204 ± 44	167 ± 46	432 ± 85
Valine	970 ± 226	592 ± 107	580 ± 132	949 ± 131

Tableau 3. 7. Number of sequences and enzymes in various KEGG annotated pathways.

	Transcriptome		3 Times		12 Times	
	Sequences	Enzymes	Sequences	Enzymes	Sequences	Enzymes
Purine Metabolism	340	37	97	14	53	11
Carbon fixation	129	18	58	15	33	12
Glycolysis	137	23	50	13	34	13
Cys/Met metabolism	102	28	27	13	14	5
Glyoxylate	89	17	30	10	19	6
TCA cycle	80	16	33	11	26	11
Pentose phosphate	78	15	27	7	21	6
Pyruvate Metabolism	76	22	22	9	23	10
Fructose/Mannose	64	13	25	8	17	5
Gly Ser Thr Metabolism	60	26	16	10	5	4
Starch Metabolism	54	18	8	6	7	5
Pyrimidine Metabolism	54	17	5	4	5	4
Glutathione	49	13	11	2		
AA-tRNA syn	42	19	6	6	2	2
Tyr Metabolism	38	11	5	4		
Trp Metabolism	36	14	9	6	8	4
Ala Asp Glu Metabolism	35	18			4	3
PheTyr Trp Biosynthesis	31	15			3	3
Arg Pro Metabolism	27	13	3	3	2	2
Fatty acid synthesis	26	7	16	4	11	3
Phe Metabolism	20	7	2	3		
Fatty acid Degradation	19	9	4	4	1	1
Arg Biosynthesis	19	9	4	3		
His metab	18	8			1	1

Val Ile Leu Biosynthesis	16	7	3	3		
Lys biosynthesis	10	4	3	2	2	1
FA elongation	9	4	3	3	1	1

## **Chapitre 4 : discussion générale**

Longtemps la régulation des rythmes circadiens de *Lingulodinium* fut obscure, et dans le but d’y faire la lumière, plusieurs hypothèses ont été testées au fil des années. Une des premières était celle liée à une variation quantitative transcriptionnelle (Roy, Beauchemin et al. 2014). Cette expérience a indiqué que le contrôle ne serait probablement pas à un niveau transcriptionnel. Par la suite l’hypothèse d’une régulation liée aux microARN (miRNA) a été examinée et aussi écartée (Dagenais-Bellefeuille, Beauchemin et al. 2017).

Une hypothèse testée ici était que plusieurs rythmes pourraient être réglés par le niveau des protéines clés. L’identification des protéines dont les niveaux changent (Chapitre 2) nous a permis de comprendre un peu plus la régulation des rythmes importants. Par exemple, la réplication de l’ADN nécessaire à la division cellulaire de cet organisme asexué a montré une régulation quantitative journalière d’une de ces protéines clés à savoir la PCNA. La PCNA a des quantités diurnes basses et des quantités nocturnes élevées, avec un pic à ZT16 juste avant le moment de la division cellulaire (ZT18). Ces variations quantitatives protéiques seraient liées aux changements de vitesse de synthèse de cette protéine, car le ribosome profiling a confirmé un rythme dans le nombre de copies en cours de synthèse. La synthèse se passe principalement en phase nocturne avec un maximum à ZT14, et un minimum observé à ZT20. Ces rythmes de synthèse sont en accord avec des observations sur une autre espèce de dinoflagellé *Prorocentrum donghaiense* (Liu, Jiao et al. 2005).

Une autre hypothèse testée ici était que les changements dans les niveaux de protéines clés pourraient être un résultat d’un contrôle par l’horloge sur la traduction des ARNm. Pour vérifier cette nouvelle hypothèse, nous avons essayé de déterminer l’étendue des protéines dont la vitesse de synthèse change (Chapitre 3). Cette étude nous a permis d’apporter un support sur la régulation des rythmes circadiens du dinoflagellé *Lingulodinium* à un niveau traductionnel.

Pour le rythme de bioluminescence, qui semble nécessaire dans la lutte contre les prédateurs, nos études par LC-MS/MS ont montré que les deux protéines clés la LBP et la LUC ont montré des quantités variantes circadiennement avec des quantités prépondérantes en phase



de nuit et des quantités faibles de jour. Les oscillations quantitatives de ces deux protéines seraient également dues aux changements de vitesse de synthèse avec des cadences soutenues en début de la phase nocturne tandis que celles-ci sont beaucoup plus faibles avant et après. Ainsi, d'après le ribosome profiling pour les deux protéines, le pic de synthèse se voit au moment même où les quantités protéiques augment plus rapidement (Johnson, Roeber et al. 1984, Morse, Milos et al. 1989). Ceci est donc entièrement en accord avec l'énoncé de Dunlap et Hastings (Dunlap and Hastings 1981) "*We therefore postulate that the circadian rhythm of luciferase activity is a result of biological clock-controlled synthesis and/or degradation of the luciferase polypeptide.*" Ici nous pouvons affirmer qu'il y a un rythme circadien dans la vitesse de synthèse de plusieurs protéines, pas seulement ceux de la bioluminescence. De plus, la conclusion de Morse (Morse, Milos et al. 1989) selon laquelle: "*The LBP mRNA levels, as determined by in vitro translations and by RNA (Northern) hybridizations, do not vary over the daily cycle, indicating that circadian control of bioluminescence in this species is mediated by translation*" semble aussi être applicable à plusieurs rythmes autres que la bioluminescence.

La principale source d'énergie de *Lingulodinium*, comme pour d'autres organismes autotrophes, est la photosynthèse nécessitant deux protéines clés à savoir d'une part les collecteurs de lumière PCP et d'autre part l'enzyme pivot du cycle de Calvin, la Rubisco. Malgré le fait que les niveaux de protéines elles-mêmes n'ont montré aucune variation quantitative circadienne (Nassoury, Fritz et al. 2001) (Bowazolo, Tse et al. 2020), nos études confirment des changements de vitesse de synthèse pendant le cycle jour nuit. Si l'augmentation de la vitesse de synthèse n'est pas cause d'un accroissement des quantités protéiques on peut supposer des mécanismes de dégradation gardant à l'équilibre les niveaux protéiques de Rubisco et de PCP. Ceci pourrait permettre une rotation des protéines individuelles dans la population (le « turnover ») et lutter contre une usure des protéines. Alternativement, la régulation de la vitesse de synthèse pourrait être reliée avec le changement de distribution des protéines dans le chloroplaste, décrit par Nassoury (Nassoury, Wang et al. 2005): "*In both cases, changes in levels of gene expression (at the translational level) initiate changes in protein localization that ultimately are observed as the rhythm.*"

Il est aussi possible que la régulation du rythme photosynthétique chez *Lingulodinium* soit influencée par l'organisation du photosystème II (PSII). Une forme de PSII associée avec une plus grande quantité de pigments photosynthétiques a été observée pendant la phase de jour (Knoetzel and Rensing 1990). Par rapport aux changements dans les quantités de pigment, nos données révèlent une variation dans le nombre de RPF pour la chlorophylle synthase (enzyme clé de la synthèse de la chlorophylle a) et la phytoène désaturase (enzyme clé de la synthèse de la xanthophylle). Toutes les deux semblent être globalement plus élevées vers la fin de la nuit (Tableau 4.1) mais plus précisément le rapport du nombre de copies en cours de synthèse entre la phytoène désaturase et la chlorophylle synthase est plus petit de jour que de nuit (Figure 4.1).

Tirer énergétiquement profit du soleil ne se fait sans méfaits, et pour faire face aux désagréments des dérivés réactifs de l'oxygène (ROS), le dinoflagellé *Lingulodinium* s'est doté des mécanismes photoprotecteurs. Un mécanisme implique la protéine PDX1 lui permettant de synthétiser la vitamine B6, un antioxydant d'une grande efficacité. Un autre mécanisme implique l'EVS, un enzyme clé de la biosynthèse des acides aminés mycosporines (MAA) lui servant « d'écran solaire ». Ces deux principaux acteurs de la photoprotection ont montré des oscillations quantitatives de leurs niveaux protéiques (Chapitre 2), avec des quantités supérieures durant les temps de jour et des quantités inférieures pendant les heures de nuit. Nous avons pu confirmer par ribosome profiling des changements dans la vitesse de synthèse qui pourraient expliquer l'accroissement massif de ces deux protéines photoprotectrices (Chapitre 3). Notre étude, ayant permis de caractériser les changements de vitesses de synthèse de l'enzyme clé de la biosynthèse des MAA, confirme les conclusions de Sinha (Sinha, Klisch et al. 2001) selon laquelle: "*There was a circadian induction in the synthesis of MAAs when the cultures were exposed to mid-latitude solar radiation... Shinorine was found to be induced mostly during the light period. During the dark period the concentration stayed almost constant.*"

On peut établir un schéma de régulation des rythmes biologiques (circadiens ou photodépendants): l'horloge endogène et/ou environnementale influence la vitesse de synthèse permettant des variations quantitatives, des protéines clés, favorables à l'accomplissement des rythmes (Figure 4.2). Désormais nous savons que l'accomplissement des principaux rythmes

circadiens du dinoflagellé *Lingulodinium* dépendrait des variations de vitesse de synthèse des protéines clés impliquées, l'une des interrogations prioritaires est celle menant à connaître les protéines initiant l'horloge et donc responsables de dicter les rythmes circadiens. Les niveaux de transcrits étant constants le long de la journée alors l'hypothèse d'une régulation par une boucle de rétrocontrôle transcriptionnelle comme chez la drosophile (Hardin, Hall et al. 1990) est écartée. Une fois ces protéines fondatrices de l'horloge, entraînant les changements de vitesses de synthèse, identifiées il restera à déterminer leur régulation afin d'élucider le mécanisme régulateur de l'horloge biologique de *Lingulodinium* qui jusqu'à présent demeure bien encore un mystère malgré les nombreuses années de recherche écoulées depuis que ce dinoflagellé a été fait modèle d'étude des rythmes circadiens. L'étude du phosphoprotéome, sur un cycle de 24h, que nous avons initiée pourrait être salvatrice si ces protéines fondatrices sont régulées par phosphorylation comme chez certains organismes (Vanselow and Kramer 2007). Sinon il faudra élargir les recherches vers d'autres modifications post-traductionnelles comme la glycolysation en analysant le glycoprotéome, bien que cette modification soit mineure chez *Lingulodinium* elle pourrait tout de même réserver de l'inattendu... Comment fonctionne l'horloge? Voilà à présent ce qui nous interroge! L'exploration future devrait conduire d'une part à la découverte de la protéine jouant chez le dinoflagellé *Lingulodinium* le rôle de CLOCK: BMAL1 qui chez les mammifères coordonne l'expression rythmique de 15% du transcriptome et contrôle la régulation quotidienne des fonctions biologiques (Trott and Menet 2018), et d'autre part à l'identification (si cela existe chez *Lingulodinium*) des protéines responsables de restaurer l'horloge à l'image de CIRCADIAN CLOCK ASSOCIATED1 (CCA1) et de LATE ELONGATED HYPOCOTYLS (LHY) décrit chez *Arabidopsis* (Yakir, Hilman et al. 2009).

Par ailleurs, au regard du fait que seules 13 protéines ont eu des niveaux variables, ceci représente effectivement peu comparé au millier ayant montré des changements de vitesse de synthèse. Il est possible que les protéines non détectées par LC-MS/MS comme variantes aient eu une différence, mais pas 1.5 fois utilisé comme seuil. Il est aussi possible qu'une augmentation dans la vitesse de dégradation permette d'annihiler les variations quantitatives qui seraient normalement provoquées par les changements de vitesses de synthèse. Tout dernièrement, il est aussi possible que les niveaux de ces protéines étaient en dessous de la quantité détectable par la spectroscopie de masse.

Tableau 4. 1. Évolution quantitative de la vitesse de synthèse de la chlorophylle synthase et de la phytoène désaturase à ZT6, 14 et 20 chez le dinoflagellé *Lingulodinium* par ribosome profiling.

	ZT6	ZT14	ZT20
Chlorophylle synthase	21	73	115
phytoène désaturase	48	355	695

Figure 4. 1. Rapport journalier du nombre de copies en cours de synthèse entre la phytoène désaturase et la chlorophylle synthase chez *Lingulodinium*.

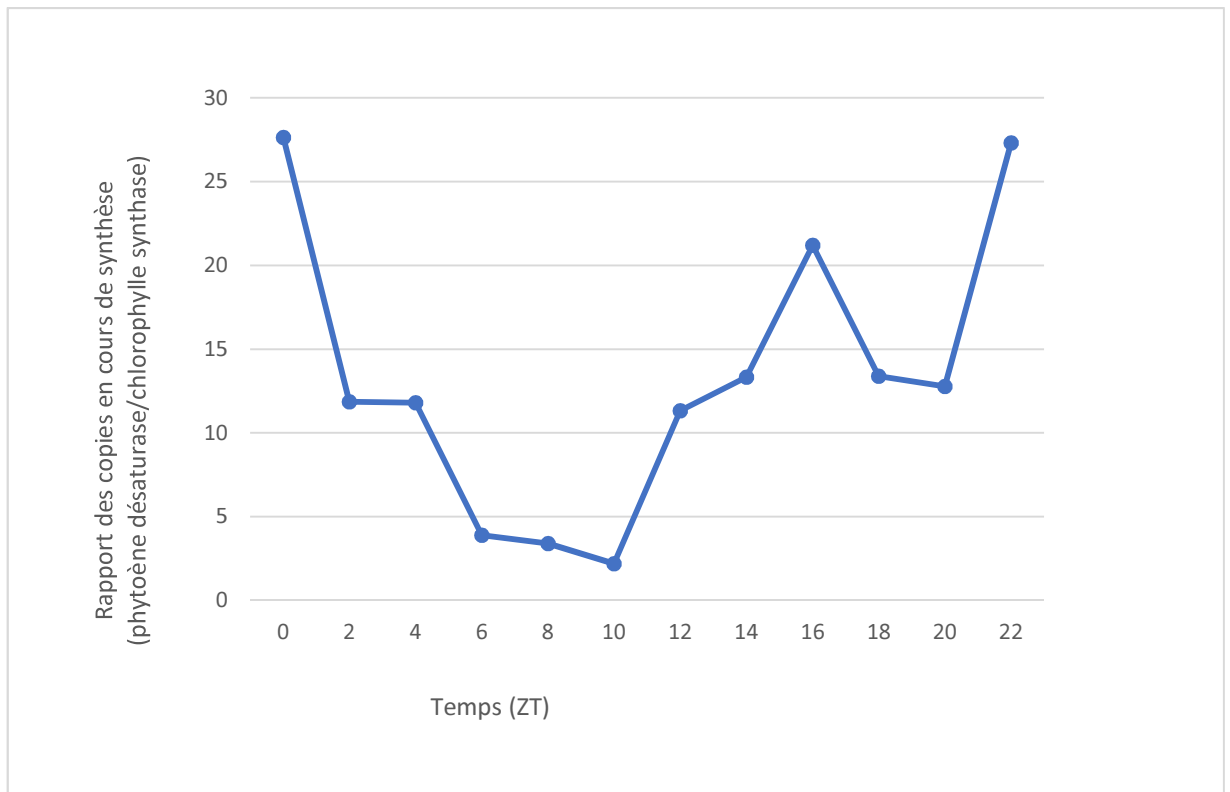
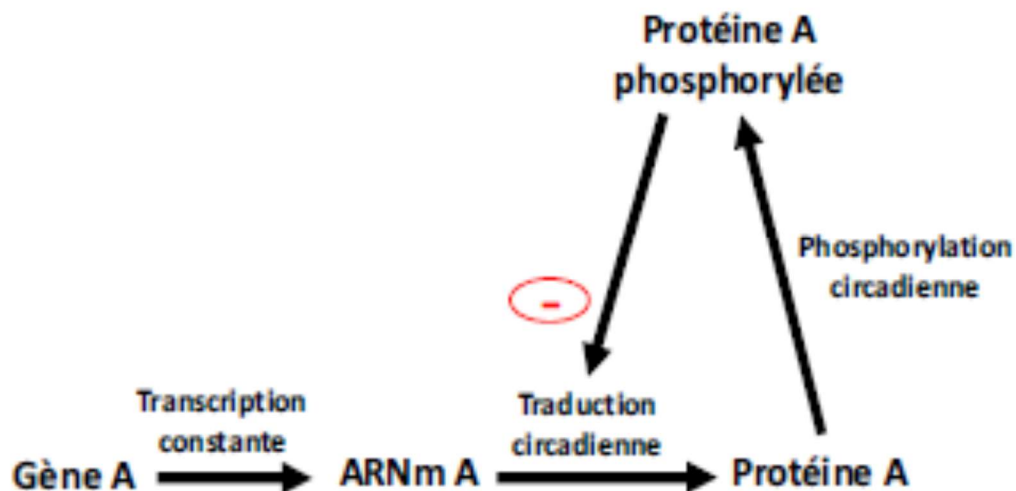
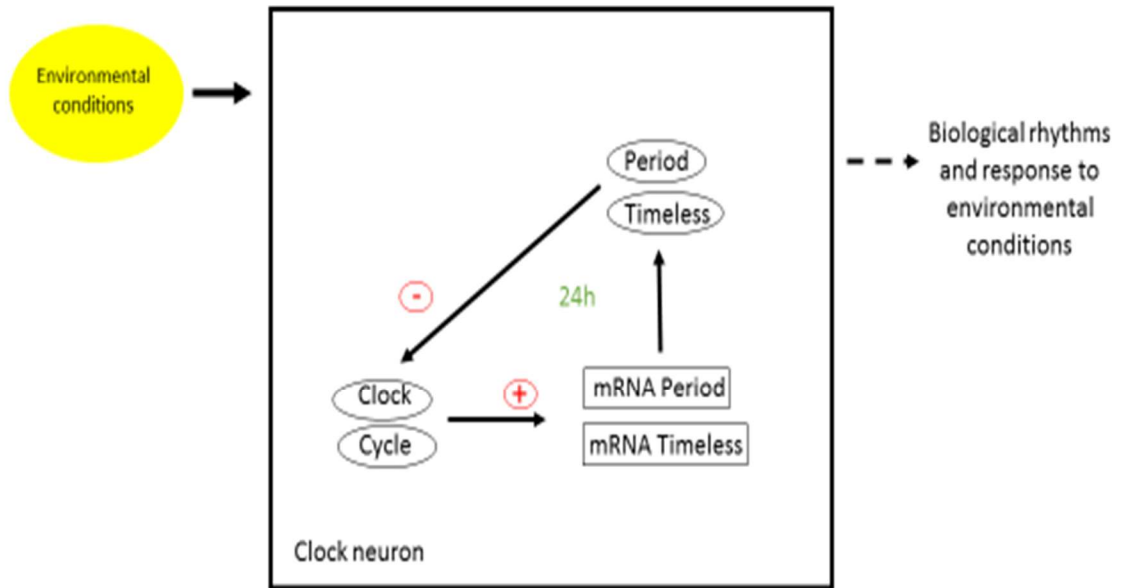


Figure 4. 2. A. Schéma proposé de la régulation des rythmes biologiques chez le dinoflagellé *Lingulodinium*. B. Modèle simplifié de boucle de rétrocontrôle dans une cellule de l'horloge de la drosophile.

Les protéines CLOCK et CYCLE activent la transcription des protéines PERIOD et TIMELESS qui à leur tour d'une part régulent les rythmes biologiques en répondant aux conditions environnementales et d'autre part inhibent CLOCK et CYCLE. Les rythmes d'activité comportementale ne dépendent pas de ces neurones réglés, mais des neurones ayant une activité cyclique. Inspiré de (Dubowy and Sehgal 2017). Chez *Lingulodinium*, on n'a pas pu démontrer des ARNm rythmiques, alors la quantité de protéine et la quantité de protéine ayant subi une modification post-traductionnelle pourraient remplacer l'ARNm et la protéine, respectivement.



A



B

## CHAPITRE 5. Proposition d'éventuels prochains projets à prioriser

Lorsque les échantillons de phosphoprotéines, toujours en cours d'analyse, auront livré leurs résultats, nous saurons à quel degré la phosphorylation chez le dinoflagellé *Lingulodinium* est impliquée dans la régulation des rythmes biologiques et/ou de l'horloge. En fonction de cette influence, il serait judicieux et pertinent dans une étape suivante d'investiguer sur les protéines, phosphorylées ou non, qui sont chargées de réguler le niveau des ARNm des protéines clés contrôlant les différents rythmes biologiques et l'horloge, afin de connaître les mécanismes permettant la stabilité quantitative des niveaux des ARNm malgré les quotidiennes variations d'une part de la vitesse de synthèse et d'autre part des niveaux protéiques. Ceci pourrait se faire par la récente méthode intitulée interactome capture.

L'interactome capture, mise au point par Hentze (Castello, Horos et al. 2013) et Landthaler (Kastelic and Landthaler 2017), est une technique permettant l'étude des interactions protéines-ARN en identifiant l'ensemble des protéines associées aux ARN messagers polyadénylés. Ainsi par cette procédure il est possible de cartographier les interactions RNA Binding Protein (RBP)-ARNm poly A dans des cultures de cellules, ce qui donne le protéome lié aux ARNm (Ryder 2016). La méthode, assez abordable, se compose de quatre principales étapes: tout d'abord la réticulation (UV conventionnelle CrossLinking (cCL) ou Photo-Activable-Ribonucléoside CrossLinking (PAR-CL)) des interactions ARNm poly A et RBP dans le but de former des liaisons covalentes entre les ARNm poly A et leurs RBP situées à moins d'un angström, ensuite la capture des ARNm polyadénylés (par oligo-désoxythymine) afin de purifier les interactions RBP-ARNm polyadénylés, après la digestion enzymatique (par ARNase) de ces interactions pour séparer les ARNm polyadénylés des protéines et enfin



l'identification des protéines (par LC-MS/MS) permettant de connaître l'ensemble des protéines en interaction avec les ARNm poly A (Bao, Guo et al. 2018) (Beckmann 2017) (Koster, Maronedze et al. 2017).

Les RBP pourraient permettre la coordination de la traduction ou de la répression des ARNm cibles, probablement par interaction avec les régions non-traduites (UTR) en 3' ou 5'. En caractérisant le protéome impliqué dans la régulation des niveaux d'ARNm chez *Lingulodinium* l'interactome capture permettrait de répertorier les RBP impliquées dans la régulation des niveaux des ARNm chez *Lingulodinium* aux moments différents de la journée. On peut émettre l'hypothèse selon laquelle la cartographie de l'ensemble des protéines interagissant avec les ARNm en cours de synthèse permettra d'identifier les protéines régulant les niveaux d'expression des ARNm dont la traduction est spécifique à un certain temps.

Pour vérifier cette hypothèse l'étude des interactions protéines-ARNm poly A par interactome capture sera faite prioritairement aux différents temps de synthèse en cycle LD chez *Lingulodinium* à savoir LD2, LD14 et LD20 et selon la pertinence des résultats ou la nécessité de compléter l'étude pourra être prolongée sur un cycle LD complet. Par cette méthode, grâce à la cartographie de l'interactome à LD2, LD14 et LD20, on pourra éventuellement résoudre l'énigme entourant la stabilité quantitative des ARNm malgré les oscillations des vitesses de synthèse. En supposant que les variations quantitatives des ARNm sont liées à l'environnement protéique immédiat de ces ARNm.

## CHAPITRE 6. Conclusion

Les résultats obtenus suite à l'accomplissement des deux projets présentés ici nous éclairent sur la régulation des rythmes biologiques du dinoflagellé *Lingulodinium*. Chez ce phytoplancton on peut maintenant affirmer qu'il y a beaucoup de variations rythmiques des vitesses de synthèse des protéines. Certaines sont sûrement des protéines clés impliquées dans les processus cellulaires, et ces oscillations de vitesse de synthèse pourraient influencer leurs niveaux journaliers. Les prochains projets à planifier auront donc pour but et défi d'expliquer comment *Lingulodinium* fait des changements de vitesse de synthèse protéique quotidiens malgré ces niveaux d'ARNm stables. Il serait aussi intéressant de commencer la description de la boucle régulatrice constituant l'horloge biologique de ce protiste marin.

La régulation des rythmes biologiques étant commune à presque toutes les espèces du monde aquatique comme terrestre soumis aux exigences diurnes et nocturnes, ces résultats apporteront des réponses pertinentes permettant d'avancer dans la compréhension de la régulation des rythmes biochimiques et physiologiques des autres espèces y compris les rythmes infradiens, circadiens et ultradiens de l'Homme.

## Références

- Adamich, M. and B. M. Sweeney (1976). "The preparation and characterization of Gonyaulax spheroplasts." Planta **130**(1): 1-6.
- Agrawal, G. K. and J. J. Thelen (2006). "Large scale identification and quantitative profiling of phosphoproteins expressed during seed filling in oilseed rape." Mol Cell Proteomics **5**(11): 2044-2059.
- Akhtar, R. A., A. B. Reddy, E. S. Maywood, J. D. Clayton, V. M. King, A. G. Smith, T. W. Gant, M. H. Hastings and C. P. Kyriacou (2002). "Circadian cycling of the mouse liver transcriptome, as revealed by cDNA microarray, is driven by the suprachiasmatic nucleus." Curr Biol **12**(7): 540-550.
- Akimoto, H., T. Kinumi and Y. Ohmiya (2005). "Circadian Rhythm of a TCA Cycle Enzyme Is Apparently Regulated at the Translational Level in the Dinoflagellate *Lingulodinium polyedrum*." Journal of Biological Rhythms **20**(6): 479-489.
- Akimoto, H., T. Kinumi and Y. Ohmiya (2005). "Circadian rhythm of a TCA cycle enzyme is apparently regulated at the translational level in the dinoflagellate *Lingulodinium polyedrum*." J Biol Rhythms **20**(6): 479-489.
- Akimoto, H., C. Wu, T. Kinumi and Y. Ohmiya (2004). "Biological rhythmicity in expressed proteins of the marine dinoflagellate *Lingulodinium polyedrum* demonstrated by chronological proteomics." Biochem Biophys Res Commun **315**(2): 306-312.
- Alabadi, D., T. Oyama, M. J. Yanovsky, F. G. Harmon, P. Mas and S. A. Kay (2001). "Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock." Science **293**(5531): 880-883.
- Allada, R., N. E. White, W. V. So, J. C. Hall and M. Rosbash (1998). "A mutant *Drosophila* homolog of mammalian Clock disrupts circadian rhythms and transcription of period and timeless." Cell **93**(5): 791-804.
- Allen, J. F. and J. Forsberg (2001). "Molecular recognition in thylakoid structure and function." Trends Plant Sci **6**(7): 317-326.
- Balskus, E. P. and C. T. Walsh (2010). "The genetic and molecular basis for sunscreen biosynthesis in cyanobacteria." Science **329**(5999): 1653-1656.
- Bao, X., X. Guo, M. Yin, M. Tariq, Y. Lai, S. Kanwal, J. Zhou, N. Li, Y. Lv, C. Pulido-Quetglas, X. Wang, L. Ji, M. J. Khan, X. Zhu, Z. Luo, C. Shao, D. H. Lim, X. Liu, N. Li, W. Wang, M. He, Y. L. Liu, C. Ward, T. Wang, G. Zhang, D. Wang, J. Yang, Y. Chen, C. Zhang, R. Jauch, Y. G. Yang, Y. Wang, B. Qin, M. L. Anko, A. P. Hutchins, H. Sun, H. Wang, X. D. Fu, B. Zhang and M. A. Esteban (2018). "Capturing the interactome of newly transcribed RNA." Nat Methods **15**(3): 213-220.
- Beauchemin, M. and D. Morse (2018). "A proteomic portrait of dinoflagellate chromatin reveals abundant RNA-binding proteins." Chromosoma **127**(1): 29-43.
- Beauchemin, M., S. Roy, P. Daoust, S. Dagenais-Bellefeuille, T. Bertomeu, L. Letourneau, B. F. Lang and D. Morse (2012). "Dinoflagellate tandem array gene transcripts are highly conserved and not polycistronic." Proc Natl Acad Sci U S A **109**(39): 15793-15798.
- Beckmann, B. M. (2017). "RNA interactome capture in yeast." Methods **118-119**: 82-92.
- Bennouna, A., Berland, B., Assobhei, O. (2002). "Eau colorée à *Lingulodinium polyedrum* (Stein) Dodge, dans une zone aquacole du littoral du Doukkala (Atlantique

marocain)Lingulodinium polyedrum (Stein) Dodge red tide in shellfish areas along Doukkala coast (Moroccan Atlantic)." Oceanologica Acta **25**(3-4): 159-170.

Bertomeu, T., J. W. Hastings and D. Morse (2003). "Vectorial labeling of dinoflagellate cell surface proteins." J. Phycol. **39**: 1254-1260.

Bonardi, V., P. Pesaresi, T. Becker, E. Schleiff, R. Wagner, T. Pfannschmidt, P. Jahns and D. Leister (2005). "Photosystem II core phosphorylation and photosynthetic acclimation require two different protein kinases." Nature **437**(7062): 1179-1182.

Bowazolo, C., S. P. K. Tse, M. Beauchemin, S. C. Lo, J. Rivoal and D. Morse (2020). "Label-free MS/MS analyses of the dinoflagellate *Lingulodinium* identifies rhythmic proteins facilitating adaptation to a diurnal LD cycle." Sci Total Environ **704**: 135430.

Bravo, R. and J. E. Celis (1980). "A search for differential polypeptide synthesis throughout the cell cycle of HeLa cells." J Cell Biol **84**(3): 795-802.

Brunelle, S. A. and F. M. van Dolah (2011). "Post-transcriptional regulation of S-phase genes in the dinoflagellate, *Karenia brevis*." J Eukaryot Microbiol **58**(4): 373-382.

Calvin, M. and A. A. Benson (1948). "The Path of Carbon in Photosynthesis." Science **107**(2784): 476-480.

Cao, R., C. G. Gkogkas, N. de Zavalía, I. D. Blum, A. Yanagiya, Y. Tsukumo, H. Xu, C. Lee, K. F. Storch, A. C. Liu, S. Amir and N. Sonenberg (2015). "Light-regulated translational control of circadian behavior by eIF4E phosphorylation." Nat Neurosci **18**(6): 855-862.

Castello, A., R. Horos, C. Strein, B. Fischer, K. Eichelbaum, L. M. Steinmetz, J. Krijgsveld and M. W. Hentze (2013). "System-wide identification of RNA-binding proteins by interactome capture." Nat Protoc **8**(3): 491-500.

Caster, S. Z., K. Castillo, M. S. Sachs and D. Bell-Pedersen (2016). "Circadian clock regulation of mRNA translation through eukaryotic elongation factor eEF-2." Proc Natl Acad Sci U S A **113**(34): 9605-9610.

Chan, W. S., A. C. M. Kwok and J. T. Y. Wong (2019). "Knockdown of Dinoflagellate Cellulose Synthase Cesa1 Resulted in Malformed Intracellular Cellulosic Thecal Plates and Severely Impeded Cyst-to-Swarmer Transition." Front Microbiol **10**: 546.

Chan, Y.-H. and J. T. Y. Wong (2007). "Concentration-dependent organization of DNA by the dinoflagellate histone-like protein HCc3." Nucleic acids research **35**(8): 2573-2583.

Cheng, P., Y. Yang, K. H. Gardner and Y. Liu (2002). "PAS domain-mediated WC-1/WC-2 interaction is essential for maintaining the steady-state level of WC-1 and the function of both proteins in circadian clock and light responses of *Neurospora*." Mol Cell Biol **22**(2): 517-524.

Chiang, C. K., N. Mehta, A. Patel, P. Zhang, Z. Ning, J. Mayne, W. Y. Sun, H. Y. Cheng and D. Figeys (2014). "The proteomic landscape of the suprachiasmatic nucleus clock reveals large-scale coordination of key biological processes." PLoS Genet **10**(10): e1004695.

Chiang, C. K., B. Xu, N. Mehta, J. Mayne, W. Y. Sun, K. Cheng, Z. Ning, J. Dong, H. Zou, H. M. Cheng and D. Figeys (2017). "Phosphoproteome Profiling Reveals Circadian Clock Regulation of Posttranslational Modifications in the Murine Hippocampus." Front Neurol **8**: 110.

Chow, M. H., K. T. H. Yan, M. J. Bennett and J. T. Y. Wong (2010). "Birefringence and DNA condensation of liquid crystalline chromosomes." Eukaryotic cell **9**(10): 1577-1587.

Chudnovsky, Y., J. F. Li, P. J. Rizzo, J. W. Hastings and T. F. Fagan (2002). "CLONING, EXPRESSION, AND CHARACTERIZATION OF A HISTONE-LIKE PROTEIN FROM THE MARINE DINOFLAGELLATE LINGULODINIUM POLYEDRUM (DINOPHYCEAE)1." Journal of Phycology **38**(3): 543-550.

Comolli, J., W. Taylor and J. W. Hastings (1994). "An Inhibitor of Protein Phosphorylation Stops the Circadian Oscillator and Blocks Light-Induced Phase Shifting in *Gonyaulax polyedra*." Journal of Biological Rhythms **9**(1): 13-26.

Comolli, J., W. Taylor, J. Rehman and J. W. Hastings (1996). "Inhibitors of serine/threonine phosphoprotein phosphatases alter circadian properties in *Gonyaulax polyedra*." Plant physiology **111**(1): 285-291.

Comolli, J. C., T. Fagan and J. W. Hastings (2003). "A type-1 phosphoprotein phosphatase from a dinoflagellate as a possible component of the circadian mechanism." J Biol Rhythms **18**(5): 367-376.

Comolli, J. C. and J. W. Hastings (1999). "Novel Effects on the *Gonyaulax* Circadian System Produced by the Protein Kinase Inhibitor Staurosporine." Journal of Biological Rhythms **14**(1): 11-19.

Cornelius, G., A. Schroeder-Lorenz and L. Rensing (1985). "Circadian-clock control of protein synthesis and degradation in *Gonyaulax polyedra*." Planta **166**: 365-370.

Correa, A., Z. A. Lewis, A. V. Greene, I. J. March, R. H. Gomer and D. Bell-Pedersen (2003). "Multiple oscillators regulate circadian gene expression in *Neurospora*." Proc Natl Acad Sci U S A **100**(23): 13597-13602.

Covington, M. F. and S. L. Harmer (2007). "The circadian clock regulates auxin signaling and responses in *Arabidopsis*." PLoS Biol **5**(8): e222.

Cullen, J. J. (1985). "Diel vertical migration by dinoflagellates: roles of carbohydrate metabolism and behavioral flexibility." Contributions in Marine Science **27**: 135-152.

Cullen, J. J. and S. G. Horrigan (1981). "Effects of nitrate on the diurnal vertical migration, carbon to nitrogen ratio, and the photosynthetic capacity of the dinoflagellate *Gymnodinium splendens*." Marine Biology **62**(2): 81-89.

Dagenais-Bellefeuille, S., M. Beauchemin and D. Morse (2017). "miRNAs Do Not Regulate Circadian Protein Synthesis in the Dinoflagellate *Lingulodinium polyedrum*." PLOS ONE **12**(1): e0168817.

Dagenais-Bellefeuille, S., T. Bertomeu and D. Morse (2008). "S-phase and M-phase timing are under independent circadian control in the dinoflagellate *Lingulodinium*." J Biol Rhythms **23**(5): 400-408.

Dagenais Bellefeuille, S., S. Dorion, J. Rivoal and D. Morse (2014). "The dinoflagellate *Lingulodinium polyedrum* responds to N depletion by a polarized deposition of starch and lipid bodies." PLoS One **9**(11): e111067.

Desjardins, M. and D. Morse (1993). "The polypeptide components of scintillons, the bioluminescence organelles of the dinoflagellate *Gonyaulax polyedra*." Biochem. Cell. Biol. **71**(3-4): 176-182.

Dobson, L., I. Remenyi and G. E. Tusnady (2015). "CCTOP: a Consensus Constrained TOPology prediction web server." Nucleic Acids Res **43**(W1): W408-412.

Dodd, A. N., N. Salathia, A. Hall, E. Kevei, R. Toth, F. Nagy, J. M. Hibberd, A. J. Millar and A. A. Webb (2005). "Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage." Science **309**(5734): 630-633.

Dodge, J. D. (1989). "Some Revisions of the Family Gonyaulacaceae (Dinophyceae) Based on Scanning Electron Microscope Study." Botanica Marina **32**(4): 275-298.

Dong, W., X. Tang, Y. Yu, R. Nilsen, R. Kim, J. Griffith, J. Arnold and H. B. Schuttler (2008). "Systems biology of the clock in *Neurospora crassa*." PLoS One **3**(8): e3105.

Donner, B., U. Helmboldt-Caesar and L. Rensing (1985). "Circadian rhythm of total protein synthesis in the cytoplasm and chloroplasts of *Gonyaulax polyedra*." Chronobiol Int **2**: 1-9.

Du, J., Y. Zhang, Y. Xue, X. Zhao, X. Zhao, Y. Wei, Z. Li, Y. Zhang and Z. Zhao (2018). "Diurnal protein oscillation profiles in *Drosophila* head." FEBS Lett **592**(22): 3736-3749.

Dubowy, C. and A. Sehgal (2017). "Circadian Rhythms and Sleep in *Drosophila melanogaster*." Genetics **205**(4): 1373-1397.

Dunlap, J. C. and J. W. Hastings (1981). "The biological clock in *Gonyaulax* controls luciferase activity by regulating turnover." J Biol Chem **256**(20): 10509-10518.

Dunlap, J. C., W. Taylor and J. W. Hastings (1980). "The effects of protein synthesis inhibitors on the *Gonyaulax* clock." Journal of comparative physiology **138**(1): 1-8.

Dunlap, J. C., W. R. Taylor and J. W. Hastings (1981). Bioluminescence: Current Perspectives, Burgess Intl Group.

Edwards, K. D., P. E. Anderson, A. Hall, N. S. Salathia, J. C. Locke, J. R. Lynn, M. Straume, J. Q. Smith and A. J. Millar (2006). "FLOWERING LOCUS C mediates natural variation in the high-temperature response of the *Arabidopsis* circadian clock." Plant Cell **18**(3): 639-650.

Ehrenschaft, M., P. Bilski, M. Y. Li, C. F. Chignell and M. E. Daub (1999). "A highly conserved sequence is a novel gene involved in de novo vitamin B6 biosynthesis." Proc Natl Acad Sci U S A **96**(16): 9374-9378.

Eisensamer, B. and T. Roenneberg (2004). "Extracellular pH is under circadian control in *Gonyaulax polyedra* and forms a metabolic feedback loop." Chronobiol Int **21**(1): 27-41.

Eppley, R. W., O. Holm-Hansen and J. D. H. Strickland (1968). "Some observations on the vertical migration of dinoflagellates." J. Phycol. **4**: 333-340.

Fagan, T., D. Morse and J. Hastings (1999). "Circadian synthesis of a nuclear encoded chloroplast Glyceraldehyde-3-phosphate dehydrogenase in the dinoflagellate *Gonyaulax polyedra* is translationally controlled." Biochemistry **38**: 7689-7695.

Field, C. B., M. J. Behrenfeld, J. T. Randerson and P. Falkowski (1998). "Primary production of the biosphere: integrating terrestrial and oceanic components." Science **281**(5374): 237-240.

Foster, J., B. Luo and P. A. Nakata (2016). "An Oxalyl-CoA Dependent Pathway of Oxalate Catabolism Plays a Role in Regulating Calcium Oxalate Crystal Accumulation and Defending against Oxalate-Secreting Phytopathogens in *Medicago truncatula*." PLoS One **11**(2): e0149850.

Franceschi, V. R. and P. A. Nakata (2005). "Calcium oxalate in plants: formation and function." Annu Rev Plant Biol **56**: 41-71.

Fritz, L., D. Morse and J. W. Hastings (1990). "The circadian bioluminescence rhythm of *Gonyaulax* is related to daily variations in the number of light-emitting organelles." J. Cell Sci. **95**(Pt 2): 321-328.

Froehlich, A. C., J. J. Loros and J. C. Dunlap (2003). "Rhythmic binding of a WHITE COLLAR-containing complex to the frequency promoter is inhibited by FREQUENCY." Proc Natl Acad Sci U S A **100**(10): 5914-5919.

Gengenbacher, M., T. B. Fitzpatrick, T. Raschle, K. Flicker, I. Sinning, S. Muller, P. Macheroux, I. Tews and B. Kappes (2006). "Vitamin B6 biosynthesis by the malaria parasite *Plasmodium falciparum*: biochemical and structural insights." J Biol Chem **281**(6): 3633-3641.

Gornik, S. G., K. L. Ford, T. D. Mulhern, A. Bacic, G. I. McFadden and R. F. Waller (2012). "Loss of nucleosomal DNA condensation coincides with appearance of a novel nuclear protein in dinoflagellates." Curr Biol **22**(24): 2303-2312.

Guerreiro, A. C., M. Benevento, R. Lehmann, B. van Breukelen, H. Post, P. Giansanti, A. F. Maarten Altelaar, I. M. Axmann and A. J. Heck (2014). "Daily rhythms in the cyanobacterium *synechococcus elongatus* probed by high-resolution mass spectrometry-based proteomics reveals a small defined set of cyclic proteins." Mol Cell Proteomics **13**(8): 2042-2055.

Guillard, R. R. and J. H. Ryther (1962). "Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt, and *Detonula confervacea* (Cleve) Gran." Can J Microbiol **8**: 229-239.

Hardin, P. E., J. C. Hall and M. Rosbash (1990). "Feedback of the *Drosophila* period gene product on circadian cycling of its messenger RNA levels." Nature **343**(6258): 536-540.

Harmer, S. L., J. B. Hogenesch, M. Straume, H. S. Chang, B. Han, T. Zhu, X. Wang, J. A. Kreps and S. A. Kay (2000). "Orchestrated transcription of key pathways in *Arabidopsis* by the circadian clock." Science **290**(5499): 2110-2113.

Harvey, E. N. (1952). Bioluminescence, Academic Press, New York.

Hastings, J. W. (1989). "Chemistry, clones, and circadian control of the dinoflagellate bioluminescent system. The Marlene DeLuca memorial lecture." J Biolumin Chemilumin **4**(1): 12-19.

Hastings, J. W. (2013). "Circadian Rhythms in Dinoflagellates: What Is the Purpose of Synthesis and Destruction of Proteins?" Microorganisms **1**(1): 26–32.

Hastings, J. W., L. Astrachan and B. M. Sweeney (1961). "A persistent daily rhythm in photosynthesis." J Gen Physiol **45**: 69-76.

Hastings, J. W. and B. M. Sweeney (1958). "A persistent diurnal rhythm of luminescence in *Gonyaulax polyedra*." Biol. Bull. **115**: 440-458.

Hays, G. C., H. Kennedy and B. W. Frost (2001). "Individual variability in diel vertical migration of a marine copepod: Why some individuals remain at depth when others migrate." Limnology and Oceanography **46**(8): 2050-2054.

Herzog, M. and M. O. Soyer (1983). "The native structure of dinoflagellate chromosomes and their stabilization by Ca<sup>2+</sup> and Mg<sup>2+</sup> cations." Eur J Cell Biol **30**(1): 33-41.

Hollnagel, H. C., E. Pinto, D. Morse and P. Colepicolo (2002). "The oscillation of photosynthetic capacity in *Gonyaulax polyedra* is not related to differences in Rubisco, peridinin or chlorophyll *a* amounts." Biol. Rhythms Res. **33**(4): 443-458.

Honner, S., R. M. Kudela and E. Handler (2012). "Bilateral mastoiditis from red tide exposure." J Emerg Med **43**(4): 663-666.

Hosokawa, N., T. S. Hatakeyama, T. Kojima, Y. Kikuchi, H. Ito and H. Iwasaki (2011). "Circadian transcriptional regulation by the posttranslational oscillator without de novo clock gene expression in *Synechococcus*." Proc Natl Acad Sci U S A **108**(37): 15396-15401.

Hou, Y. and S. Lin (2009). "Distinct gene number-genome size relationships for eukaryotes and non-eukaryotes: gene content estimation for dinoflagellate genomes." PloS one **4**(9): e6978-e6978.

Hughes, M. E., J. B. Hogenesch and K. Kornacker (2010). "JTK\_CYCLE: an efficient nonparametric algorithm for detecting rhythmic components in genome-scale data sets." J Biol Rhythms **25**(5): 372-380.

Hui, S., J. M. Ghergurovich, R. J. Morscher, C. Jang, X. Teng, W. Lu, L. A. Esparza, T. Reya, Z. Le, J. Yanxiang Guo, E. White and J. D. Rabinowitz (2017). "Glucose feeds the TCA cycle via circulating lactate." Nature **551**(7678): 115-118.

Hurley, J. M., A. Dasgupta, J. M. Emerson, X. Zhou, C. S. Ringelberg, N. Knabe, A. M. Lipzen, E. A. Lindquist, C. G. Daum, K. W. Barry, I. V. Grigoriev, K. M. Smith, J. E. Galagan, D. Bell-Pedersen, M. Freitag, C. Cheng, J. J. Loros and J. C. Dunlap (2014). "Analysis of clock-

regulated genes in *Neurospora* reveals widespread posttranscriptional control of metabolic potential." Proceedings of the National Academy of Sciences **111**(48): 16995-17002.

Hurley, J. M., M. S. Jankowski, H. De Los Santos, A. M. Crowell, S. B. Fordyce, J. D. Zucker, N. Kumar, S. O. Purvine, E. W. Robinson, A. Shukla, E. Zink, W. R. Cannon, S. E. Baker, J. J. Loros and J. C. Dunlap (2018). "Circadian Proteomic Analysis Uncovers Mechanisms of Post-Transcriptional Regulation in Metabolic Pathways." Cell Syst **7**(6): 613-626 e615.

Ingolia, N. T. (2010). "Genome-wide translational profiling by ribosome footprinting." Methods Enzymol **470**: 119-142.

Ingolia, N. T. (2016). "Ribosome Footprint Profiling of Translation throughout the Genome." Cell **165**(1): 22-33.

Ingolia, N. T., G. A. Brar, S. Rouskin, A. M. McGeachy and J. S. Weissman (2012). "The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments." Nat Protoc **7**(8): 1534-1550.

Irwin, N. A. T., B. J. E. Martin, B. P. Young, M. J. G. Browne, A. Flaus, C. J. R. Loewen, P. J. Keeling and L. J. Howe (2018). "Viral proteins as a potential driver of histone depletion in dinoflagellates." Nat Commun **9**(1): 1535.

Ishiura, M., S. Kutsuna, S. Aoki, H. Iwasaki, C. R. Andersson, A. Tanabe, S. S. Golden, C. H. Johnson and T. Kondo (1998). "Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria." Science **281**(5382): 1519-1523.

Ito, H., M. Mutsuda, Y. Murayama, J. Tomita, N. Hosokawa, K. Terauchi, C. Sugita, M. Sugita, T. Kondo and H. Iwasaki (2009). "Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in *Synechococcus elongatus*." Proc Natl Acad Sci U S A **106**(33): 14168-14173.

Jang, C., N. F. Lahens, J. B. Hogenesch and A. Sehgal (2015). "Ribosome profiling reveals an important role for translational control in circadian gene expression." Genome Res **25**(12): 1836-1847.

Janich, P., A. B. Arpat, V. Castelo-Szekely, M. Lopes and D. Gatfield (2015). "Ribosome profiling reveals the rhythmic liver transcriptome and circadian clock regulation by upstream open reading frames." Genome Res **25**(12): 1848-1859.

Johnson, C. H., S. Inoué, A. Flint and J. W. Hastings (1985). "Compartmentalization of algal bioluminescence: autofluorescence of bioluminescent particles in the dinoflagellate *Gonyaulax* as studied with image-intensified video microscopy and flow cytometry." J Cell Biol **100**(5): 1435-1446.

Johnson, C. H., J. F. Roeber and J. W. Hastings (1984). "Circadian changes in enzyme concentration account for rhythm of enzyme activity in *Gonyaulax*." Science **223**: 1428-1430.

Juntawong, P., T. Girke, J. Bazin and J. Bailey-Serres (2014). "Translational dynamics revealed by genome-wide profiling of ribosome footprints in *Arabidopsis*." Proc Natl Acad Sci U S A **111**(1): E203-212.

Kageyama, H., T. Kondo and H. Iwasaki (2003). "Circadian formation of clock protein complexes by KaiA, KaiB, KaiC, and SasA in cyanobacteria." J Biol Chem **278**(4): 2388-2395.

Karakashian, M. W. and J. W. Hastings (1963). "The Effects of Inhibitors of Macromolecular Biosynthesis Upon the Persistent Rhythm of Luminescence in *Gonyaulax*." J Gen Physiol **47**: 1-12.

Kastelic, N. and M. Landthaler (2017). "mRNA interactome capture in mammalian cells." Methods **126**: 38-43.



Knoetzel, J. and L. Rensing (1990). "Characterization of the Photosynthetic Apparatus from the Marine Dinoflagellate *Gonyaulax polyedra*: II. Circadian Rhythmicity of Photosynthesis and the Supramolecular Organization of Pigment-Protein Complexes." Journal of Plant Physiology **136**(3): 280-288.

Kombrink, E., N. J. Kruger and H. Beevers (1984). "Kinetic properties of pyrophosphate:fructose-6-phosphate phosphotransferase from germinating castor bean endosperm." Plant Physiol **74**(2): 395-401.

Koster, T., C. Marondedze, K. Meyer and D. Staiger (2017). "RNA-Binding Proteins Revisited - The Emerging Arabidopsis mRNA Interactome." Trends Plant Sci **22**(6): 512-526.

Lakin-Thomas, P. L. (2006). "Transcriptional feedback oscillators: maybe, maybe not." J Biol Rhythms **21**(2): 83-92.

Lapointe, M. and D. Morse (2008). "Reassessing the role of a 3'-UTR-binding translational inhibitor in regulation of circadian bioluminescence rhythm in the dinoflagellate *Gonyaulax*." Biol Chem **389**(1): 13-19.

Larsen, M. R., T. E. Thingholm, O. N. Jensen, P. Roepstorff and T. J. Jorgensen (2005). "Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns." Mol Cell Proteomics **4**(7): 873-886.

Le Bihan, T., S. F. Martin, E. S. Chirnside, G. van Ooijen, M. E. Barrios-Llerena, J. S. O'Neill, P. V. Shliaha, L. E. Kerr and A. J. Millar (2011). "Shotgun proteomic analysis of the unicellular alga *Ostreococcus tauri*." J Proteomics **74**(10): 2060-2070.

Le, Q., R. Jovine, P. Markovic and D. Morse (2001). "PCP is not implicated in the photosynthesis rhythm of *Gonyaulax* despite circadian regulation of its translation." Biol. Rhythm Research **32**(5): 579-594.

Li, S. and L. Zhang (2015). "Circadian Control of Global Transcription." Biomed Res Int **2015**: 187809.

Lin, S. and E. J. Carpenter (1998). "Identification and preliminary characterization of PCNA gene in the marine phytoplankton *Dunaliella tertiolecta* and *Isochrysis galbana*." Mol Mar Biol Biotechnol **7**(1): 62-71.

Lin, S., S. Cheng, B. Song, X. Zhong, X. Lin, W. Li, L. Li, Y. Zhang, H. Zhang, Z. Ji, M. Cai, Y. Zhuang, X. Shi, L. Lin, L. Wang, Z. Wang, X. Liu, S. Yu, P. Zeng, H. Hao, Q. Zou, C. Chen, Y. Li, Y. Wang, C. Xu, S. Meng, X. Xu, J. Wang, H. Yang, D. A. Campbell, N. R. Sturm, S. Dagenais-Bellefeuille and D. Morse (2015). "The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis." Science **350**(6261): 691-694.

Lindstrom, J., W. Grebner, K. Rigby and E. Selander (2017). "Effects of predator lipids on dinoflagellate defence mechanisms - increased bioluminescence capacity." Sci Rep **7**(1): 13104.

Liu, B., S. C. Lo, D. P. Matton, B. F. Lang and D. Morse (2012). "Daily changes in the phosphoproteome of the dinoflagellate *Lingulodinium*." Protist **163**(5): 746-754.

Liu, J., N. Jiao, H. Hong, T. Luo and H. Cai (2005). "Proliferating cell nuclear antigen (PCNA) as a marker of cell proliferation in the marine dinoflagellate *Prorocentrum donghaiense* Lu and the green alga *Dunaliella salina* Teodoresco." Journal of Applied Phycology **17**(4): 323-330.

Lorimer, G. H., T. J. Andrews and N. E. Tolbert (1973). "Ribulose diphosphate oxygenase. II. Further proof of reaction products and mechanism of action." Biochemistry **12**(1): 18-23.

Luck, S., K. Thurley, P. F. Thaben and P. O. Westermark (2014). "Rhythmic degradation explains and unifies circadian transcriptome and proteome data." Cell Rep **9**(2): 741-751.

Mackenzie, T. D. and D. Morse (2011). "Circadian photosynthetic reductant flow in the dinoflagellate *Lingulodinium* is limited by carbon availability." Plant Cell Environ.

Markovic, P., T. Roenneberg and D. Morse (1996). "Phased protein synthesis at several circadian times does not change protein levels in *Gonyaulax*." J Biol Rhythms **11**(1): 57-67.

Matallana-Surget, S., J. Derock, B. Leroy, H. Badri, F. Deschoenmaecker and R. Wattiez (2014). "Proteome-wide analysis and diel proteomic profiling of the cyanobacterium *Arthrospira platensis* PCC 8005." PLoS One **9**(6): e99076.

Mateyak, M. K., J. K. Pupek, A. E. Garino, M. C. Knapp, S. F. Colmer, T. G. Kinzy and S. Dunaway (2018). "Demonstration of translation elongation factor 3 activity from a non-fungal species, *Phytophthora infestans*." PLoS One **13**(1): e0190524.

Matsushika, A., S. Makino, M. Kojima and T. Mizuno (2000). "Circadian waves of expression of the APRR1/TOC1 family of pseudo-response regulators in *Arabidopsis thaliana*: insight into the plant circadian clock." Plant Cell Physiol **41**(9): 1002-1012.

Mauvoisin, D., J. Wang, C. Jouffe, E. Martin, F. Atger, P. Waridel, M. Quadroni, F. Gachon and F. Naef (2014). "Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal functions in mouse liver." Proc Natl Acad Sci U S A **111**(1): 167-172.

Milos, P., D. Morse and J. W. Hastings (1990). "Circadian control over synthesis of many *Gonyaulax* proteins is at a translational level." Naturwissenschaften **77**(2): 87-89.

Missra, A., B. Ernest, T. Lohoff, Q. Jia, J. Satterlee, K. Ke and A. G. von Arnim (2015). "The Circadian Clock Modulates Global Daily Cycles of mRNA Ribosome Loading." Plant Cell **27**(9): 2582-2599.

Mittag, M., D. H. Lee and J. W. Hastings (1994). "Circadian expression of the luciferin-binding protein correlates with the binding of a protein to the 3' untranslated region of its mRNA." Proc Natl Acad Sci U S A **91**(12): 5257-5261.

Mittag, M., L. Li and J. W. Hastings (1998). "The mRNA level of the circadian regulated *Gonyaulax* luciferase remains constant over the cycle." Chronobiol. Int. **15**(1): 93-98.

Moorthi, S. D., P. D. Countway, B. A. Stauffer and D. A. Caron (2006). "Use of quantitative real-time PCR to investigate the dynamics of the red tide dinoflagellate *Lingulodinium polyedrum*." Microb Ecol **52**(1): 136-150.

Morales-Ramirez, A., R. Viquez, K. Rodriguez and M. Vargas (2001). "[Red tide bloom produced by *Lingulodinium polyedrum* (Peridinales, Dinophyceae) in Bahia Culebra, Papagayo Gulf, Costa Rica]." Rev Biol Trop **49 Suppl 2**: 19-23.

Morey, J. S. and F. M. van Dolah (2013). "Global analysis of mRNA half-lives and de novo transcription in a dinoflagellate, *Karenia brevis*." PLoS One **8**(6): e66347.

Morse, D., P. M. Milos, E. Roux and J. W. Hastings (1989). "Circadian regulation of bioluminescence in *Gonyaulax* involves translational control." Proc. Natl. Acad. Sci. U.S.A. **86**(1): 172-176.

Morse, D., A. M. Pappenheimer, Jr. and J. W. Hastings (1989). "Role of a luciferin-binding protein in the circadian bioluminescent reaction of *Gonyaulax polyedra*." J. Biol. Chem. **264**(20): 11822-11826.

Morse, D., P. Salois, P. Markovic and J. W. Hastings (1995). "A nuclear encoded form II rubisco in dinoflagellates." Science **268**: 1622-1624.

Morse, D., S. P. K. Tse and S. C. L. Lo (2018). "Exploring dinoflagellate biology with high-throughput proteomics." Harmful Algae **75**: 16-26.

Nakajima, M., K. Imai, H. Ito, T. Nishiwaki, Y. Murayama, H. Iwasaki, T. Oyama and T. Kondo (2005). "Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro." Science **308**(5720): 414-415.

Nakamura, H., Y. Kishi, O. Shimomura, D. Morse and J. W. Hastings (1989). "Structure of dinoflagellate luciferin and its enzymic and nonenzymic air-oxidation products." Journal of the American Chemical Society **111**(19): 7607-7611.

Nakano, A. and M. Muramatsu (1989). "A novel GTP-binding protein, Sar1p, is involved in transport from the endoplasmic reticulum to the Golgi apparatus." J Cell Biol **109**(6 Pt 1): 2677-2691.

Nassoury, N., M. Cappadocia and D. Morse (2003). "Plastid ultrastructure defines the protein import pathway in dinoflagellates." J Cell Sci **116**(Pt 14): 2867-2874.

Nassoury, N., L. Fritz and D. Morse (2001). "Circadian changes in ribulose-1,5-bisphosphate carboxylase/oxygenase distribution inside individual chloroplasts can account for the rhythm in dinoflagellate carbon fixation." Plant Cell **13**(4): 923-934.

Nassoury, N., Y. Wang and D. Morse (2005). "Brefeldin A inhibits circadian remodeling of chloroplast structure in the dinoflagellate *Gonyaulax*." Traffic **6**(7): 548-561.

Nicolas, M. T., D. Morse, J. M. Bassot and J. Woodland Hastings (1991). "Colocalization of luciferin binding protein and luciferase to the scintillons of *Gonyaulax polyedra* revealed by double immunolabeling after fast-freeze fixation." Protoplasma **160**(2): 159-166.

Nowrousian, M., G. E. Duffield, J. J. Loros and J. C. Dunlap (2003). "The frequency gene is required for temperature-dependent regulation of many clock-controlled genes in *Neurospora crassa*." Genetics **164**(3): 923-933.

Okamoto, O. K. and J. W. Hastings (2003). "Novel dinoflagellate clock-related genes identified through microarray analysis 1." Journal of Phycology **39**(3): 519-526.

Ouyang, Y., C. Andersson, T. Kondo, S. Golden and C. Johnson (1998). "Resonating circadian clocks enhance fitness in cyanobacteria." Proc. Natl. Acad. Sci. USA **95**: 8660-8664.

Panda, S., M. P. Antoch, B. H. Miller, A. I. Su, A. B. Schook, M. Straume, P. G. Schultz, S. A. Kay, J. S. Takahashi and J. B. Hogenesch (2002). "Coordinated transcription of key pathways in the mouse by the circadian clock." Cell **109**(3): 307-320.

Podnar, J., H. Deiderick, G. Huerta and S. Hunicke-Smith (2014). "Next-Generation Sequencing RNA-Seq Library Construction." Curr Protoc Mol Biol **106**: 4 21 21-19.

Pokhilko, A., P. Mas and A. J. Millar (2013). "Modelling the widespread effects of TOC1 signalling on the plant circadian clock and its outputs." BMC Syst Biol **7**: 23.

Prosperi, E. (1997). "Multiple roles of the proliferating cell nuclear antigen: DNA replication, repair and cell cycle control." Prog Cell Cycle Res **3**: 193-210.

Quail, M. A., I. Kozarewa, F. Smith, A. Scally, P. J. Stephens, R. Durbin, H. Swerdlow and D. J. Turner (2008). "A large genome center's improvements to the Illumina sequencing system." Nat Methods **5**(12): 1005-1010.

Raschke, M., S. Boycheva, M. Crevecoeur, A. Nunes-Nesi, S. Witt, A. R. Fernie, N. Amrhein and T. B. Fitzpatrick (2011). "Enhanced levels of vitamin B(6) increase aerial organ size and positively affect stress tolerance in *Arabidopsis*." Plant J **66**(3): 414-432.

Raschle, T., N. Amrhein and T. B. Fitzpatrick (2005). "On the two components of pyridoxal 5'-phosphate synthase from *Bacillus subtilis*." J Biol Chem **280**(37): 32291-32300.

Raven, J. (1995). "Photosynthetic and non-photosynthetic roles of carbonic anhydrase in algae and cyanobacteria." Phycologia **34**: 93-101.

Raven, J. (1997). "Putting the C in phycology." European Journal of Phycology **32**: 319-333.

Reddy, A. B., N. A. Karp, E. S. Maywood, E. A. Sage, M. Deery, J. S. O'Neill, G. K. Wong, J. Chesham, M. Odell, K. S. Lilley, C. P. Kyriacou and M. H. Hastings (2006). "Circadian orchestration of the hepatic proteome." Curr Biol **16**(11): 1107-1115.

Reid, D. W., S. Shenolikar and C. V. Nicchitta (2015). "Simple and inexpensive ribosome profiling analysis of mRNA translation." Methods **91**: 69-74.

Rensing, L., G. Cornelius and A. Schroeder-Lorenz (1985). "Circadian-clock control of protein synthesis and degradation in *Gonyaulax polyedra*." **166**: 365-370.

Rivoal, J. and A. D. Hanson (1993). "Evidence for a Large and Sustained Glycolytic Flux to Lactate in Anoxic Roots of Some Members of the Halophytic Genus *Limonium*." Plant Physiol **101**(2): 553-560.

Robles, M. S., J. Cox and M. Mann (2014). "In-vivo quantitative proteomics reveals a key contribution of post-transcriptional mechanisms to the circadian regulation of liver metabolism." PLoS Genet **10**(1): e1004047.

Robles, M. S., S. J. Humphrey and M. Mann (2017). "Phosphorylation Is a Central Mechanism for Circadian Control of Metabolism and Physiology." Cell Metab **25**(1): 118-127.

Rochaix, J. D. (2007). "Role of thylakoid protein kinases in photosynthetic acclimation." FEBS Lett **581**(15): 2768-2775.

Roenneberg, T., G. N. Colfax and J. W. Hastings (1989). "A circadian rhythm of population behavior in *Gonyaulax polyedra*." J. Biol. Rhythms **4**: 201-216.

Roenneberg, T. and D. Morse (1993). "Two circadian oscillators in one cell." Nature **362**(6418): 362-364.

Roy, S., M. Beauchemin, S. Dagenais-Bellefeuille, L. Letourneau, M. Cappadocia and D. Morse (2014). "The *Lingulodinium* circadian system lacks rhythmic changes in transcript abundance." BMC Biol **12**(1): 107.

Roy, S., L. Letourneau and D. Morse (2014). "Cold-induced cysts of the photosynthetic dinoflagellate *Lingulodinium polyedrum* have an arrested circadian bioluminescence rhythm and lower levels of protein phosphorylation." Plant Physiol **164**(2): 966-977.

Roy, S. and D. Morse (2012). "A full suite of histone and histone modifying genes are transcribed in the dinoflagellate *Lingulodinium*." PLoS One **7**(4): e34340.

Roy, S. and D. Morse (2014). "The dinoflagellate *Lingulodinium* has predicted casein kinase 2 sites in many RNA binding proteins." Protist **165**(3): 330-342.

Rutila, J. E., V. Suri, M. Le, W. V. So, M. Rosbash and J. C. Hall (1998). "CYCLE is a second bHLH-PAS clock protein essential for circadian rhythmicity and transcription of *Drosophila* period and timeless." Cell **93**(5): 805-814.

Ryder, S. P. (2016). "Protein-mRNA interactome capture: cartography of the mRNP landscape." F1000Res **5**: 2627.

Sala-Rovira, M., M. L. Geraud, D. Caput, F. Jacques, M. O. Soyer-Gobillard, G. Vernet and M. Herzog (1991). "Molecular cloning and immunolocalization of two variants of the major basic nuclear protein (HCc) from the histone-less eukaryote *Cryptothecodinium cohnii* (Pyrrhophyta)." Chromosoma **100**(8): 510-518.

Schaffer, R., J. Landgraf, M. Accerbi, V. Simon, M. Larson and E. Wisman (2001). "Microarray analysis of diurnal and circadian-regulated genes in *Arabidopsis*." Plant Cell **13**(1): 113-123.

Schaffer, R., N. Ramsay, A. Samach, S. Corden, J. Putterill, I. A. Carre and G. Coupland (1998). "The late elongated hypocotyl mutation of *Arabidopsis* disrupts circadian rhythms and the photoperiodic control of flowering." Cell **93**(7): 1219-1229.

Schmitter, R. E. (1971). "The fine structure of *Gonyaulax polyedra*, a bioluminescent marine dinoflagellate." J Cell Sci **9**(1): 147-173.

Shestov, A. A., X. Liu, Z. Ser, A. A. Cluntun, Y. P. Hung, L. Huang, D. Kim, A. Le, G. Yellen, J. G. Albeck and J. W. Locasale (2014). "Quantitative determinants of aerobic glycolysis identify flux through the enzyme GAPDH as a limiting step." Elife **3**.

Shi, X., L. Li and S. Lin (2018). "Circadian and irradiance effects on expression of antenna protein genes and pigment contents in dinoflagellate *Prorocentrum donghaiense* (Dinophyceae)." Harmful Algae **75**: 27-34.

Sigee, D. C. (1983). "Structural DNA and genetically active DNA in dinoflagellate chromosomes." Biosystems **16**(3): 203-210.

Sinha, R. P., M. Klisch, E. Walter Helbling and D.-P. Häder (2001). "Induction of mycosporine-like amino acids (MAAs) in cyanobacteria by solar ultraviolet-B radiation." Journal of Photochemistry and Photobiology B: Biology **60**(2): 129-135.

Spoelstra, K., M. Wikelski, S. Daan, A. S. Loudon and M. Hau (2016). "Natural selection against a circadian clock gene mutation in mice." Proc Natl Acad Sci U S A **113**(3): 686-691.

Stat, M., E. Morris and R. D. Gates (2008). "Functional diversity in coral-dinoflagellate symbiosis." Proc Natl Acad Sci U S A **105**(27): 9256-9261.

Stockel, J., J. M. Jacobs, T. R. Elvitigala, M. Liberton, E. A. Welsh, A. D. Polpitiya, M. A. Gritsenko, C. D. Nicora, D. W. Koppenaal, R. D. Smith and H. B. Pakrasi (2011). "Diurnal rhythms result in significant changes in the cellular protein complement in the cyanobacterium *Cyanothece* 51142." PLoS One **6**(2): e16680.

Sun, S., M. Liu, F. Dong, S. Fan and Y. Yao (2013). "A histone-like protein induces plasmid DNA to form liquid crystals in vitro and gene compaction in vivo." International journal of molecular sciences **14**(12): 23842-23857.

Sweeney, B. M. (1960). "The photosynthetic rhythm in single cells of *Gonyaulax polyedra*." Cold Spring Harbor Symp. Quant. Biol. **25**: 145-148.

Sweeney, B. M. and J. W. Hastings (1958). "Rhythmic cell division in populations of *Gonyaulax polyedra*." J. Protozool. **5**: 217-224.

Tabita, F. R. (1995). Anoxygenic Photosynthetic Bacteria, Springer Netherlands Kluwer Academic Publishers.

Taira, H. and S. Taguchi (2017). "Cellular Mycosporine-like amino acids protect photosystem II of the Dinoflagellate *Scrippsiella sweeneyae* from ultraviolet radiation damage." J Photochem Photobiol B **174**: 27-34.

Taylor, F. J. R. (1975). "Non-helical transverse flagella in dinoflagellates." Phycologia **14**: 45-47.

Taylor, W., R. Krasnow, J. C. Dunlap, H. Broda and J. W. Hastings (1982). "Critical pulses of anisomycin drive the circadian oscillator in *Gonyaulax* towards its singularity." Journal of comparative physiology **148**(1): 11-25.

Taylor, W. R., J. C. Dunlap and J. W. Hastings (1982). "Inhibitors of protein synthesis on 80S ribosomes phase shift the *Gonyaulax* clock." J Exp Biol **97**: 121-136.

Terauchi, K., Y. Kitayama, T. Nishiwaki, K. Miwa, Y. Murayama, T. Oyama and T. Kondo (2007). "ATPase activity of KaiC determines the basic timing for circadian clock of cyanobacteria." Proc Natl Acad Sci U S A **104**(41): 16377-16381.

Theng, V., S. Agarie and A. Nose (2007). "Regulatory Properties of Phosphoenolpyruvate Carboxylase in Crassulacean Acid Metabolism Plants: Diurnal Changes in Phosphorylation State and Regulation of Gene Expression." Plant Production Science **10**(2): 171-181.

Titiz, O., M. Tambasco-Studart, E. Warzych, K. Apel, N. Amrhein, C. Laloi and T. B. Fitzpatrick (2006). "PDX1 is essential for vitamin B6 biosynthesis, development and stress tolerance in Arabidopsis." Plant J **48**(6): 933-946.

Tomita, J., M. Nakajima, T. Kondo and H. Iwasaki (2005). "No transcription-translation feedback in circadian rhythm of KaiC phosphorylation." Science **307**(5707): 251-254.

Trott, A. J. and J. S. Menet (2018). "Regulation of circadian clock transcriptional output by CLOCK:BMAL1." PLoS genetics **14**(1): e1007156-e1007156.

Tse, S. P. K., M. Beauchemin, D. Morse and S. C. L. Lo (2018). "Refining transcriptome gene catalogs by MS-validation of expressed proteins." Proteomics **18**(1): 201700271.

Ueda, H. R., A. Matsumoto, M. Kawamura, M. Iino, T. Tanimura and S. Hashimoto (2002). "Genome-wide transcriptional orchestration of circadian rhythms in Drosophila." J Biol Chem **277**(16): 14048-14052.

Vanselow, K. and A. Kramer (2007). "Role of phosphorylation in the mammalian circadian clock." Cold Spring Harb Symp Quant Biol **72**: 167-176.

Veldhuis, M. J. W., T. L. Cucci and M. E. Sieracki (1997). "Cellular DNA content of marine phytoplankton using two new fluorochromes: taxonomic and ecological implications." Journal of Phycology **33**(3): 527-541.

Vijayan, V., R. Zuzow and E. K. O'Shea (2009). "Oscillations in supercoiling drive circadian gene expression in cyanobacteria." Proc Natl Acad Sci U S A **106**(52): 22564-22568.

Vizcaino, J. A., E. W. Deutsch, R. Wang, A. Csordas, F. Reisinger, D. Rios, J. A. Dianes, Z. Sun, T. Farrah, N. Bandeira, P. A. Binz, I. Xenarios, M. Eisenacher, G. Mayer, L. Gatto, A. Campos, R. J. Chalkley, H. J. Kraus, J. P. Albar, S. Martinez-Bartolome, R. Apweiler, G. S. Omenn, L. Martens, A. R. Jones and H. Hermjakob (2014). "ProteomeXchange provides globally coordinated proteomics data submission and dissemination." Nat Biotechnol **32**(3): 223-226.

Waldbauer, J. R., S. Rodrigue, M. L. Coleman and S. W. Chisholm (2012). "Transcriptome and proteome dynamics of a light-dark synchronized bacterial cell cycle." PLoS One **7**(8): e43432.

Walker, S. E. and J. Lorsch (2013). "RNA purification--precipitation methods." Methods Enzymol **530**: 337-343.

Wang, D. Z., Y. J. Zhang, S. F. Zhang, L. Lin and H. S. Hong (2013). "Quantitative proteomic analysis of cell cycle of the dinoflagellate *Prorocentrum donghaiense* (Dinophyceae)." PLoS One **8**(5): e63659.

Wang, Z. Y. and E. M. Tobin (1998). "Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression." Cell **93**(7): 1207-1217.

Wargo, M. J. and P. J. Rizzo (2001). "Exception to eukaryotic rules." Science **294**(5551): 2477.

Woelfle, M. A. and C. H. Johnson (2006). "No promoter left behind: global circadian gene expression in cyanobacteria." J Biol Rhythms **21**(6): 419-431.

Wong, J. T. Y., D. C. New, J. C. W. Wong and V. K. L. Hung (2003). "Histone-like proteins of the dinoflagellate *Cryptocodinium cohnii* have homologies to bacterial DNA-binding proteins." Eukaryotic cell **2**(3): 646-650.

Xu, Y., T. Mori and C. H. Johnson (2003). "Cyanobacterial circadian clockwork: roles of KaiA, KaiB and the kaiBC promoter in regulating KaiC." EMBO J **22**(9): 2117-2126.

Yakir, E., D. Hilman, I. Kron, M. Hassidim, N. Melamed-Book and R. M. Green (2009). "Posttranslational regulation of CIRCADIAN CLOCK ASSOCIATED1 in the circadian oscillator of Arabidopsis." Plant physiology **150**(2): 844-857.

- Yu, Y., L. Bai, K. Minagawa, X. Jian, L. Li, J. Li, S. Chen, E. Cao, T. Mahmud, H. G. Floss, X. Zhou and Z. Deng (2005). "Gene cluster responsible for validamycin biosynthesis in *Streptomyces hygrosopicus* subsp. *jinggangensis* 5008." *Appl Environ Microbiol* **71**(9): 5066-5076.
- Zaheri, B., S. Dagenais-Bellefeuille, B. Song and D. Morse (2019). "Assessing Transcriptional Responses to Light by the Dinoflagellate *Symbiodinium*." *Microorganisms* **7**(8).
- Zhang, R., N. F. Lahens, H. I. Ballance, M. E. Hughes and J. B. Hogenesch (2014). "A circadian gene expression atlas in mammals: implications for biology and medicine." *Proc Natl Acad Sci U S A* **111**(45): 16219-16224.
- Zonneveld, K., et al. (2013). "Atlas of modern dinoflagellate cyst distribution based on 2405 datapoints." *Palaeobotany and Palynology* **191**: p.1-197.

