Université de Montréal

# Étude de la transcription par des techniques à haut-débit chez le dinoflagellé *Lingulodinium polyedrum*

par Mathieu Beauchemin

Département de Sciences Biologiques
Faculté des Arts et des Sciences

Thèse présentée
en vue de l'obtention du grade de Ph. D.
en Sciences Biologiques

Mars 2017

© Mathieu Beauchemin, 2017

# Résumé

Les dinoflagellés sont un groupe très varié d'eucaryotes unicellulaires principalement retrouvés en milieu marin où ils ont un rôle écologique majeur dans la production primaire des océans et dans la formation des récifs coralliens. Ils sont aussi responsables de phénomènes néfastes comme les marées rouges et la production de toxines qui contaminent les crustacées et les poissons lorsque les conditions sont propices à leur multiplication rapide. Ces organismes possèdent des caractéristiques moléculaires uniques chez les eucaryotes, particulièrement en ce qui a trait au noyau. Le génome des dinoflagellés s'y retrouve condensé sous forme de cristal liquide stabilisé par des cations métalliques plutôt que sous forme de nucléosomes organisés par des histones. Cette forme condensée persiste tout au long du cycle cellulaire et limite la transcription des gènes à des boucles d'ADN qui se retrouvent à la périphérie des chromosomes. Ces caractéristiques inhabituelles, jumelées à une taille souvent imposante de leur génome, ont grandement limité l'étude des mécanismes moléculaires de base comme la régulation de la transcription. Afin de mieux caractériser ce processus chez ces organismes, le dinoflagellé *Lingulodinium polyedrum*, qui est étudié majoritairement pour ses multiples rythmes circadiens, a été utilisé.

Tout d'abord, le séquençage à haut-débit a été utilisé afin de caractériser le transcriptome complet de l'organisme à quatre temps différents au cours d'une journée. De façon surprenante, ce séquençage a montré que très peu des transcrits varient entre les temps et que ces variations sont de faible amplitude, ce qui contraste fortement avec les résultats obtenus chez d'autres groupes d'organismes. Une inhibition pharmacologique de la transcription a aussi été effectuée et montre que les rythmes de photosynthèse et de bioluminescence continuent d'osciller de façon circadienne en absence de transcription. Ces résultats suggèrent que le modèle traditionnel de génération des rythmes circadiens par des boucles de rétroaction transcription/traduction est probablement absent chez *Lingulodinium.*

La deuxième approche consiste en la caractérisation de deux protéines à domaine *cold-shock* (CSD), un domaine d'origine bactérienne qui est fortement surreprésenté chez les dinoflagellés et annoté comme potentiel facteur de transcription. Ces deux protéines ont une préférence pour la liaison de l'ADN simple brin versus l'ADN double brin tout en étant

capable de lier de l'ARN. Par ailleurs, ces liaisons aux acides nucléiques ne sont pas spécifiques à une séquence particulière. L'analyse de leur abondance après un passage prolongé en condition froide n'a pas permis de déceler une augmentation importante de l'abondance des deux protéines et celles-ci sont également incapable de complémenter un quadruple mutant des protéines *cold-shock* chez *E. coli*. Ces données suggèrent que les CSDs des dinoflagellés ne sont probablement pas des facteurs de transcription séquence-spécifique et qu'ils ont également une fonction différente des protéines bactériennes.

La troisième approche consiste en la réticulation *in vivo* des protéines interagissant avec la chromatine. Après réticulation, la chromatine a été purifiée et les protéines associées à ceux-ci ont été extraites et identifiées par spectrométrie de masse. Parmi les protéines identifiées, il y a un nombre réduit de protéines de liaison à l'ADN en comparaison avec des données similaires chez les animaux. Des peptides provenant d'une histone H4 ont aussi été identifiés, ce qui consiste en une des premières identifications de ce type de protéine chez les dinoflagellés. Plusieurs protéines de liaison à l'ARN ont été identifiées et pourraient permettre de réguler diverses étapes de la biologie des ARNm et ainsi moduler la traduction. Quelques protéines reliées au contrôle du cycle cellulaire et à la réparation de l'ADN ont aussi été identifiées. Mises ensemble, ces trois approches permettent de suggérer que la régulation de la transcription et de l'abondance des ARNm semblent avoir une importance moindre chez les dinoflagellés que chez les autres eucaryotes.

**Mots-clés** : Dinoflagellés, *Lingulodinium polyedrum*, transcription, expression génique, protéine à domaine Cold-Shock, rythmes circadiens, chromatine

# Abstract

Dinoflagellates are a diverse group of unicellular eukaryotes found in marine habitat where they have a major role in the primary production of the ocean and in the formation of coral reefs. They are also responsible for some of the harmful algal blooms whose toxin production can contaminate crustacean and fish. Dinoflagellates possess unique molecular characteristics in eukaryotes, especially for their nucleus. For example, their genome is found in a permanently condensed liquid crystalline state stabilized by metallic cations instead of the nucleosome organized by histones. This condensed form persists throughout the cell cycle and limits transcription to DNA loops localized at the periphery of the chromosomes. These unusual characteristics, coupled with a generally large genome size, have greatly limited the study of basic molecular mechanisms such as transcription regulation. In order to better characterize this process for the dinoflagellates, I used *Lingulodinium polyedrum*, a dinoflagellate studied extensively for its circadian rhythms.

As a first approach, high-throughput sequencing was used to characterize the complete transcriptome of the organism at four different times throughout the day. Surprisingly, this sequencing revealed that an extremely low number of transcripts vary in abundance between time points and that those variations are of low amplitude, a result in stark contrast with what has been observed in other organisms. A pharmacological inhibition of transcription was also done and shows that bioluminescence and photosynthesis rhythms persist in absence of transcription, suggesting that the classical transcription/translation feedback loop used to generate rhythmic timing in eukaryotes is probably absent in *Lingulodinium*.

The second approach was the characterization of two proteins with a cold-shock domain (CSD), a type of domain strongly overrepresented in dinoflagellates and predicted to be a potential transcription factor. Those two proteins showed a preferential binding to single stranded DNA versus double stranded DNA while also being able to bind RNA, and were not specific to a particular sequence. Protein abundance analysis after a prolonged cold shock did not yield a massive increase in the abundance of those two proteins, as seen in *E. coli*. Furthermore, neither protein was able to complement a quadruple cold shock protein (CSP) mutant in *E. coli*. Data gathered here suggest that CSD proteins in dinoflagellates are probably

not sequence-specific transcription factors and may also have a function different from bacterial CSP.

The third approach consisted of the *in vivo* cross-linking of chromatin-interacting proteins. After cross-linking, the chromatin was purified and the proteins associated with it extracted and identified by mass spectrometry. Of the identified proteins, few DNA binding proteins were found, unlike similar studies done in animals. Peptides derived from a histone H4 were discovered, one of the first instances of histone identification in dinoflagellates. Multiple proteins able to bind RNA have been identified and could be used to regulate multiple steps of RNA biology and therefore modulate RNA translation. Some proteins related to cell cycle control and DNA repair were also identified. Taken together, these three approaches support the view that the transcriptional regulation and control over mRNAs abundance seem to have a lower importance in dinoflagellate than in other eukaryotes.

**Keywords** : Dinoflagellates, *Lingulodinium polyedrum*, transcription, gene expression, Cold-Shock domain protein, circadian rhythms, chromatin

# Table des matières

Résumé ........................................ i

Abstract ........................................ iii

Table des matières ........................................ v

Liste des tableaux ........................................ ix

Liste des figures ........................................ x

Liste des abréviations ........................................ xi

Remerciements ........................................ xv

Chapitre 1 : Introduction et revue de littérature ........................................ 1

1.1. Introduction générale sur les dinoflagellés ........................................ 2
1.1.1. Rôle écologique ........................................ 2
1.1.2. Histoire évolutive et phylogénie ........................................ 2
1.2. Particularités moléculaires des dinoflagellés ........................................ 5
1.2.1. Chloroplastes et mitochondries ........................................ 5
1.2.2. Cycle cellulaire ........................................ 7
1.2.3. Régulation génique ........................................ 8
1.2.3.1. Régulation de la transcription ........................................ 8
1.2.3.1.1. Organisation génomique ........................................ 8
1.2.3.1.2. Machinerie transcriptionnelle et facteurs de transcription ........................................ 11
1.2.3.1.3. Études à large échelle ........................................ 13
1.2.3.2. Régulation post-transcriptionnelle ........................................ 14
1.3. Rythmes circadiens ........................................ 17
1.3.1. Modèles de régulation circadienne ........................................ 18
1.3.1.1. Modèle classique de boucle de transcription/traduction : exemple chez *O. tauri* et *A. thaliana* ........................................ 21
1.3.1.2. Émergence de modèles à oscillateurs post-traductionnel ........................................ 25
1.3.1.2.1. L'oscillateur biochimique des cyanobactéries ........................................ 26

1.3.1.2.2. Oscillations post-transcriptionnelles chez les eucaryotes ........................ 34
1.3.2. Rythmes circadiens chez *Lingulodinium* ........................................................ 36
1.3.2.1. Mécanisme de l'horloge chez *Lingulodinium* ............................................. 37
1.3.2.2. Intrants .......................................................................................................... 40
1.3.2.3. Sorties ............................................................................................................ 41
1.4. Présentation générale du projet ............................................................................ 45
Chapitre 2: Publication # 1 ......................................................................................... 47
The Lingulodinium circadian system lacks rhythmic changes in transcript abundance ........... 48
2.1. Abstract .................................................................................................................. 49
2.2. Background ............................................................................................................ 50
2.3. Results .................................................................................................................... 52
2.3. Discussion .............................................................................................................. 59
2.5. Conclusion ............................................................................................................. 62
2.6. Methods .................................................................................................................. 63
2.6.1. Cell Culture ......................................................................................................... 63
2.6.2. RNA extraction and RNA-Seq ............................................................................ 63
2.6.3. Transcription Inhibition ...................................................................................... 64
2.6.4. Rhythm Measurements ....................................................................................... 65
2.7. Acknowledgements ................................................................................................ 66
Chapitre 3: Publication # 2 ......................................................................................... 70
Characterization of two dinoflagellate cold shock domain proteins ........................................ 71
3.1. Abstract .................................................................................................................. 72
3.2. Introduction ............................................................................................................ 73
3.3. Results .................................................................................................................... 75
3.3.1. *Lp*CSP1 and *Lp*CSP2 are part of a distinct clade in the eukaryotic CSP family. 75
3.3.2. *Lp*CSP1 and *Lp*CSP2 are not functional complements of *E. coli* CSPs ........... 78
3.3.3. *Lp*CSP1 and *Lp*CSP2 expression are not detectable after cold shock ............. 79
3.3.4. *Lp*CSPs bind nucleic acids *in vitro* ................................................................ 80
3.4. Discussion .............................................................................................................. 83

3.5. Methods ........ 86
3.5.1. Cell Culture ........ 86
3.5.2. Sequence alignment and Phylogenetic Analyses ........ 86
3.5.3. CSPs cloning, expression and purification ........ 87
3.5.4. Bacterial complementation assay ........ 88
3.5.5. 2D PAGE of protein from Cysts, LD12 cells and thrombin-cleaved LpCSP1 and LpCSP2 ........ 89
3.5.6. Electrophoretic mobility shift assays ........ 89
3.6. Acknowledgements ........ 91
Chapitre 4: Publication # 3 ........ 96
Abundant RNA binding proteins in the dinoflagellate nuclear proteome ........ 97
4.1. Abstract ........ 98
4.2. Introduction ........ 99
4.3. Methods ........ 101
4.3.1. Algal culture ........ 101
4.3.2. Purification of nuclei and protein extraction ........ 101
4.3.3. Western blotting ........ 102
4.3.4. Microscopy ........ 103
4.3.5. Mass spectrometry ........ 103
4.3.6. Data analysis ........ 104
4.4. Results ........ 106
4.4.1. Validation of the chromatin extraction method ........ 106
4.4.2. MS results and general protein classification ........ 108
4.4.3. DNA replication, transcription and mRNA processing ........ 111
4.4.4. Nucleic acid binding proteins ........ 113
4.4.5. Other proteins of interest ........ 116
4.4.6. Differential expression between midday and midnight ........ 117
4.5. Discussion ........ 119
4.6. Acknowledgments ........ 123
Chapitre 5 : Discussion générale ........ 126

Bibliographie........................................................................................................................ i

Annexe .............................................................................................................................. xxxi

# Liste des tableaux

Tableau 3.S1. Primers used for amplifying and cloning *Lp*CSP1 and *Lp*CSP2 in the pGEX4T2 and pINIII plasmids. .......... 92
Tableau 4.1. Number of peptides and protein identified in nine chromatin-enriched samples .......... 109
Tableau 4.2. Proteins involved in DNA replication, transcription and RNA splicing found in chromatin-enriched samples .......... 112
Tableau 4.3. Proteins involved in DNA repair and processing.......... 113
Tableau 4.4. Nucleic acid binding domain present in the chromatin-associated-enriched samples.......... 115
Tableau 4.S1. List of 1245 proteins identified in the chromatin enriched samples with BLAST description, GO terms and manually classified functional characterization.......... 124
Tableau 4.S2. Gene ontology terms significantly enriched and depleted in the chromatin enriched samples.......... 124
Tableau 4.S3. Differentially expressed proteins between LD6 and LD18 .......... 124

# Liste des figures

Figure 1.1. Position phylogénétique des dinoflagellés ........................................................ 4
Figure 1.2. Organisation génomique chez les dinoflagellés ................................................... 9
Figure 1.3. Modèles classiques d'oscillateurs circadiens ...................................................... 20
Figure 1.4. Oscillateur circadien post-traductionnel ............................................................ 29
Figure 1.5. Rythmes circadiens observés chez le dinoflagellé *Lingulodinium polyedrum* ....... 37
Figure 2.1. Transcript abundance does not change between midday (ZT 6) and midnight (ZT 18). ........................................................................................................................ 53
Figure 2.2. *Lingulodinium* does not have rhythmic transcripts over a 24 hour cycle. ............. 56
Figure 2.3. Rhythms continue in the presence of transcription inhibitors. ............................. 58
Figure 2.S1. Analysis using the Velvet assembly ................................................................. 67
Figure 2.S2. Analysis using duplicate ZT6 and ZT/CT18 samples ........................................ 68
Figure 2.S3. Individual bioluminescence traces. .................................................................. 69
Figure 3.1. Cold Shock Proteins (CSP) contain a conserved cold shock domain (CSD). ........ 76
Figure 3.2. Dinoflagellate CSPs are found in two distinct clades. ......................................... 77
Figure 3.3. Dinoflagellate CSPs do not complement a quadruple bacterial CSP mutant. ........ 78
Figure 3.4. Cold treatment does not induce detectable levels of *Lp*CSP1 and *Lp*CSP2. .......... 80
Figure 3.5. LpCSP1 and LpCSP2 show similar nucleic acid binding activity in EMSA. ........ 82
Figure 3.S1. Sequence alignment of CSD protein ................................................................ 93
Figure 3.S2. LpCSP1, LpCSP2 and CspA purification ......................................................... 94
Figure 3.S3. *Lp*CSP1 shows a preference for ssDNA over dsDNA ....................................... 95
Figure 4.1. Outline and validation of the chromatin purification procedure ......................... 107
Figure 4.2. General classification of proteins retrieved in chromatin sample ....................... 110
Figure 4.3. Differential protein abundance between LD6 and LD18 .................................... 118
Figure 4.S1. Gene ontology enrichment analysis ................................................................ 125

# Liste des abréviations

3’ UTR : 3’ untranslated region

5’ UTR: 5’ untranslated region

ATX: Ataxin

BLAST: Basic local alignment tool

CCA1 : Circadian clock associated 1

CDK: Cyclin dependent kinase

CHEP: Chromatin Enrichment for Proteomics

CikA: Circadian output kinase A

CK1: Casein kinase 1

CK2: Casein kinase 2

CSD : Cold Shock Domain

CSP : Cold Shock Protein

CT : Circadian time

DAPI: 4',6-diamidino-2-phénylindole

DD: Dark/Dark

Dinap1: Dinoflagellate nuclear associated protein 1

dsDNA: Double stranded DNA

DTT: Dithiothreitol

DVNP: Dinoflagellate viral nucleoprotein

EDTA: Ethylenediaminetetraacetic

ELF: Early flowering

EMSA: Electrophoretic mobility shift assay

GO: Gene Ontology

GSK: Glycogen synthase kinase 3

GST: Gluthatione S-transferase

HLP: Histone-like protein

HMG: High-mobility group

HSP: Heat shock protein

KEGG: Kyoto Encyclopedia of genes and genomes

KH: K Homology domain

LBP : Luciferin binding protein

LC: Liquid chromatography

LCF : Luciferase

LD: Light/Dark

LdpA: Light dependant protein A

LHY: Late elongated hypocotyl

LL: Light/Light

LOV-HK: Light, Oxygen or Voltage sensitive-Histidine Kinase

LUCA: Last Universal Common Ancestor

LUX: Lux arythmo

MCM complex: mini-chromosome maintenance protein complex

MS: Mass spectrometry

NagC: N-acetylglucosamine repressor

ORF: Open reading frame

PBS: Phosphate-buffered saline

PCNA: Proliferating Cell nuclear antigen

PCP : Peridinin-Chlorophyll a binding protein

Per: Period

Pex: Period extender

PMSF: Phenylmethylsulfonyl fluoride

PPIase: Peptidyl-prolyl isomerase

PRP19: pre-mRNA processing factor

PRR: Pseudo response regulator

PsR: Pseudo-Receiver

RCC1: Regulator of chromosome condensation 1

RNP: Ribonucleoprotein particule

RPA: replication protein A

RPKM: Reads per Kilobase per Million Reads

RRM: RNA recognition motif

RubisCO: Ribulose 1,5-bisphosphate carboxylase/oxygenase

SAP: SAF-A/B, Acinus and PIAS

SAR: Straménopile-Alvéolé-Rhizaria

SL: Spliced leader

ssDNA: Single stranded DNA

TBP: TATA-box binding protein

TBS: Tris-buffered saline

TBST: Tris-buffered saline plus Tween 20

TBSTM: Tris-buffered saline plus Tween 20 plus Milk

TF: Transcription factor

TLF: TBP-like factor

TOC1: Timing of Cab expression

TREX: transcription/export complex

TTFL: Transcriptional-translational feedback loop

TYF: Twenty four

ZT: Zeitgeiber time

ZTL: Zeitlupe

# Remerciements

Un premier mot de remerciement va à David Morse, mon directeur de recherche. Sa grande curiosité et son enthousiasme pour les nouvelles idées expérimentales, ses extraordinaires talents de communicateur, tant à l'oral qu'à l'écrit, et la grande liberté intellectuelle qu'il laisse à ses étudiants font de lui un directeur exceptionnel. Je n'aurais pu envisager passer les premières années de ma carrière scientifiques sous de meilleurs auspices.

J'ai eu l'immense chance d'avoir travaillé avec deux extraordinaires collègues en Steve Dagenais-Bellefeuille et Sougata Roy pour la majeure partie de mes études au laboratoire. Leur talent, leur créativité et leur ardeur au travail ont contribué de façon importante à cette thèse et à ma progression scientifique. Merci également à Philippe Daoust qui m'a formé lors de mon arrivée au laboratoire.

Un gros merci à mes collègues Caroline, Audrey, Sébastien, Yang et Jonathan qui ont fait de l'IRBV un milieu d'étude et de vie des plus intéressant. J'aimerais aussi souligner les contributions, tant scientifiques que sociales, de plusieurs stagiaires qui sont passés au laboratoire au fil des ans, en particulier Sofia, Alex, Frédéric, Sarah, Claire et Josianne.

Au point de vue plus personnel, merci à mes parents pour leur soutien indéfectible et leurs encouragements tout au long de mes études. J'aimerais également souligner l'apport de plusieurs de mes amis, Julien et Laurence, Olivier et Émilie, Audrey et Sébastien, et Alex et Mathieu pour leur soutien et leur présence durant ces longues années d'études.

J'aimerais finalement souligner le financement reçu durant mon doctorat sous la forme d'une bourse doctorale du CRSNG ainsi qu'une bourse Pehr-Kalm des Amis du Jardin Botanique et une bourse du Fonds de bourses en Sciences Biologiques.

# Chapitre 1 : Introduction et revue de littérature

# 1.1. Introduction générale sur les dinoflagellés

## 1.1.1. Rôle écologique

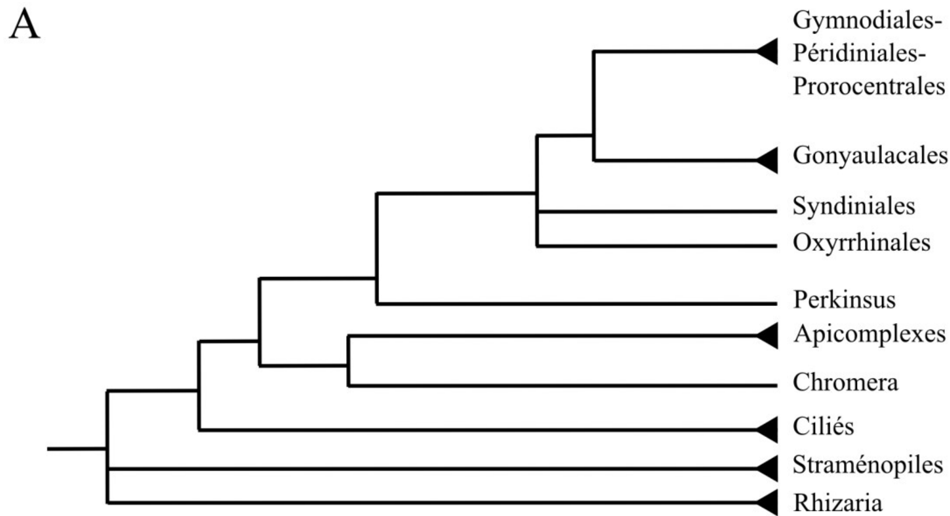
Les dinoflagellés sont un groupe d'eucaryotes unicellulaires extrêmement diversifié et se retrouvent dans une multitude de niches écologiques, surtout en milieu marin. Ils forment un groupe contenant plus de 2000 espèces, dont environ la moitié sont photosynthétiques et l'autre moitié hétérotrophes, avec quelques espèces parasitiques [1]. Une partie des espèces photosynthétiques, en particulier celles faisant partie du genre *Symbiodinium*, sont retrouvées en symbiose avec les coraux, leur fournissant des sucres en échange de certains nutriments [2]. Les dinoflagellés, avec les diatomées, sont également responsables d'une part importante de la production primaire des océans par le biais de leur activité photosynthétique et jouent ainsi un rôle majeur dans le cycle géochimique du carbone à l'échelle de la planète [3].

Les dinoflagellés sont également associés à des phénomènes plus néfastes. Lorsque les conditions propices sont en place, certaines espèces peuvent se diviser très rapidement et former des efflorescences (ou marées rouges). Celles-ci sont associées à une coloration rouge de la surface de la mer due à la présence de la péridinine, un pigment photosynthétique unique aux dinoflagellés. La fréquence des marrés rouges est en hausse depuis quelques décennies, probablement dû à un plus grand apport de nutriments comme le phosphate et le nitrate en milieu marin [4]. La haute concentration en dinoflagellés est également responsable de l'asphyxie de la mer ainsi que d'une consommation importante de nutriments, ce qui laisse peu de place à d'autres organismes vivants pour proliférer [5]. Finalement, certains dinoflagellés sont capables de produire des phycotoxines, comme des toxines paralysantes ou des toxines diarrhéiques, qui peuvent contaminer des crustacés et des poissons, les rendant ainsi impropres à la consommation par les populations humaines [6].

## 1.1.2. Histoire évolutive et phylogénie

Les dinoflagellés font partie du super phylum Straménopile-Alvéolé-Rhizaria (SAR) qui contient une grande variété d'organismes unicellulaires et ayant divergé des autres eucaryotes il y a environ 1,5 milliard d'années (Figure 1.1. A) [7, 8]. Les straménopiles

contiennent notamment les diatomées, un groupe de protistes photosynthétiques possédant un exosquelette composé de silice et vivant principalement en milieu marin. Quant à eux, les organismes du groupe rhizaria contiennent entre autres les radiolaires et les foraminifères. Les alvéolés doivent leur nom à la présence d'alvéoles corticales qui sont des vésicules aplaties supportant la membrane plasmique. Ils forment trois principaux groupes, soit les ciliés, les apicomplexes et les dinoflagellés qui possèdent des caractéristiques écologiques et moléculaires fort divergentes [8]. Le premier groupe, les ciliés, dont font partie les genres *Paramecium* et *Tetrahymena*, sont des organismes hétérotrophes qui se retrouvent généralement en milieu aquatique. Le deuxième groupe, celui des apicomplexes, contient exclusivement des parasites des animaux comme l'agent causant la malaria, *Plasmodium falciparum* et l'agent causant la toxoplasmose, *Toxoplasma gondii*. À la base de ce groupe se retrouve *Chromera velia*, un protiste photosynthétique perçu comme l'ancêtre des apicomplexes parasitiques [9].



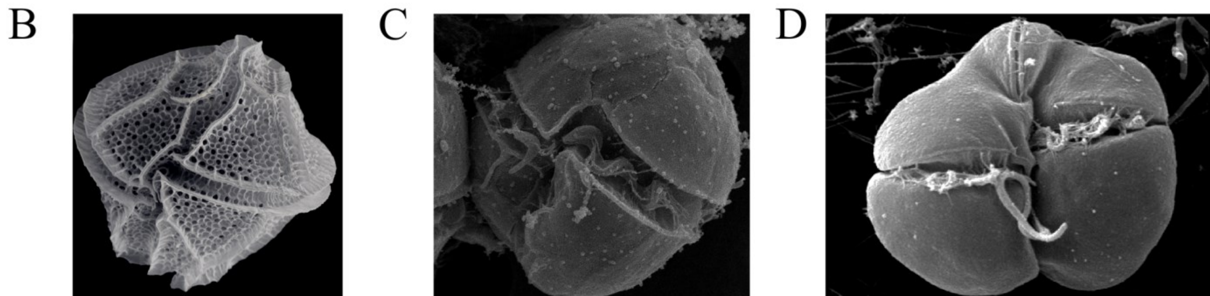




**Figure 1.1. Position phylogénétique des dinoflagellés**

A. Phylogénie du super-groupe des Straménopiles-Alvéolés-Rhizaria (SAR), adapté de [10]. B-D. Photographies prises au microscope électronique à balayage de trois dinoflagellés : B. *Lingulodinium polyedrum* C. *Alexandrium minutum* D. *Karenia brevis*. Crédit photo : Florida Fish and Wildlife Conservation Commision.

Le troisième groupe des alvéolés, les dinoflagellés, a eu son dernier ancêtre commun avec les apicomplexes il y a environ 1 milliard d'années [8]. Ils ont comme principales caractéristiques morphologiques la présence de deux flagelles orientés de façon perpendiculaire qui leur permettent de nager en tourbillonnant ainsi que, pour plusieurs espèces, une thèque formée d'un assemblage de plusieurs plaques cellulosiques qui sont entourées par les alvéoles (Figure 1.1. B-D). Les dinoflagellés peuvent être séparés en deux grands groupes, soit un groupe basal comprenant les Syndiniales, composé d'espèces

parasitiques et les Oxyrrhinales, et le groupe des dinoflagellés typiques, lui-même divisé en deux groupes, soit les Gonyaucales et le groupe des Gymnodiales-Peridiniales-Prorocentrales [10]. Comme dans le cas des apicomplexes, un ancêtre des dinoflagellés, *Perkinsus marinus*, est présent à la base de la phylogénie et possède certaines caractéristiques similaires aux dinoflagellés, comme la présence de transcrits épissés en trans [11] et certaines caractéristiques antérieures à leur divergence, comme l'utilisation d'histones pour organiser la chromatine [12].

## 1.2. Particularités moléculaires des dinoflagellés

Les dinoflagellés sont également reconnus pour posséder une multitude de caractéristiques moléculaires uniques parmi les eucaryotes, à tel point qu'ils ont été considérés comme étant des mésocaryote, un intermédiaire entre les procaryotes et les eucaryotes, avant l'avènement des phylogénies moléculaires [13]. Parmi les caractéristiques les plus inhabituelles, on peut noter, entre autres, la présence d'un noyau condensé en tout temps pendant le cycle cellulaire et la persistance de l'enveloppe nucléaire durant la division, l'absence d'histones et l'empaquetage du matériel génétique sous forme de cristal liquide, la présence de plusieurs gènes organisés en tandem ainsi que des génomes chloroplastiques et mitochondriaux organisés de façon unique. Malgré ces caractéristiques fascinantes et leur rôle écologique important, la biologie moléculaire des dinoflagellés n'a été qu'assez peu étudiée dans une poignée d'espèces modèles.

### 1.2.1. Chloroplastes et mitochondries

Les dinoflagellés ont plusieurs caractéristiques uniques quant à la photosynthèse qui s'explique par l'acquisition par endosymbiose secondaire d'un chloroplaste provenant d'une algue rouge [14, 15]. Ce chloroplaste est bordé par trois membranes et possède un mécanisme d'import protéique nécessitant le passage des protéines chloroplastiques à travers le Golgi par le biais d'un peptide signal à deux parties dont la première partie permet le passage dans le réticulum endoplasmique et la deuxième permet le passage des deux membranes

chloroplastiques restantes [16, 17]. La machinerie enzymatique responsable de la photosynthèse chez les dinoflagellés se distingue de plusieurs façons des autres protistes photosynthétiques. Ils possèdent un pigment qui leur est unique et qui est responsable de leur coloration rouge, la péridinine, une xanthophylle qui est liée par PCP (*Peridinin Chlorophyll a binding Protein*), le complexe collecteur de lumière le plus abondant. Ce pigment permet d'élargir considérablement le spectre d'absorption pour la lumière bleu vert, qui est plus abondante sous l'eau [18]. Les dinoflagellés sont également les seuls eucaryotes à posséder une RubisCO de type II provenant d'une α-protéobactérie et maintenant encodée dans le noyau [19, 20]. Malgré une fonction enzymatique identique, cette enzyme ne possède que 25 % de similarité avec la RubisCO de type I présente chez les algues vertes et les plantes supérieures et se retrouve sous forme d'un homodimère sans petites sous-unités plutôt que d'un hétérooctamère [21, 22]. Comme chez la majorité des organismes photosynthétiques unicellulaires, cette enzyme peut parfois être séquestrée dans une région du chloroplaste spécialisée dans la fixation du carbone, le pyrénoïde [23]. La RubisCO est alimentée par un système de concentration de carbone comprenant une anhydrase carbonique de type delta située sur la face externe de la membrane plasmique et qui permet de convertir le bicarbonate en $CO_2$ pour le faire ensuite entrer de façon passive dans la cellule [24].

Ce chloroplaste possède également le génome le plus fortement réduit parmi tous les eucaryotes photosynthétiques. Contrairement au génome chloroplastique des plantes et des algues vertes, qui code pour environ 120 protéines sur un génome circulaire de 120 à 200 kb, celui des dinoflagellés code pour moins d'une vingtaine de gènes et possède une organisation tout à fait différente [25, 26]. Ces gènes codent principalement pour des sous-unités des photosystèmes I et II et du complexe cytochrome $b_6f$ [27]. Ils sont encodés sur des molécules d'ADN circulaire ressemblant à des plasmides et appelées « *minicircle* » [28] codant généralement pour 1 à 3 gènes, toujours orientés de la même façon par rapport à une séquence assez bien conservée dans chaque mini-cercle et pouvant potentiellement servir de promoteur [25]. Les transcrits matures de ces gènes possèdent également une queue poly-U à la place d'une queue poly-A [29, 30].

Les mitochondries des dinoflagellés ont également un génome inusité. Celui-ci est très fortement réduit en termes de nombre de séquences codantes, qui sont limités à trois gènes,

soit *cob*, *cox1* et *cox3*, et plusieurs fragments codant pour l'ARN ribosomal, une situation similaire à ce qui est observé chez les apicomplexes [31] et qui est le plus faible nombre de gènes observé dans un génome mitochondrial [32]. En opposition au faible nombre de gènes, le génome mitochondrial des dinoflagellés a subi une très grande expansion par rapport aux apicomplexes, qui ont des génomes d'environ 6 à 8 kb [31]. Le génome mitochondrial de *Symbiodinium minutum* a été complètement séquencé et contient 326 kb de matériel génétique [33]. Son organisation est très complexe, contenant plusieurs copies de chacun des 3 gènes, 27 fragments potentiels d'ARN ribosomal ainsi que 99% de séquences non codantes. Une fois transcrits, les ARN codants pour les gènes mitochondriaux sont fortement édités alors que plus de 5% des nucléotides subissent des substitutions après la transcription [33, 34]. Finalement, chez *Karlodinium micrum*, le gène *cox3* est encodé par deux transcrits différents, l'un codant pour les 6 premières hélices et un autre qui encode la dernière hélice et qui sont joints ensemble par épissage après leur transcription [35].

### 1.2.2. Cycle cellulaire

Les dinoflagellés possèdent un cycle cellulaire qui a certaines caractéristiques uniques parmi les eucaryotes, comme la condensation permanente de l'ADN tout au long du cycle ainsi que le maintien de l'enveloppe nucléaire lors de la mitose [36]. Le cycle cellulaire suit une progression normale pour un eucaryote, soit la succession des phases G1, S, G2 et M. La régulation de ces évènements au niveau moléculaire a été quelque peu étudiée. Des régulateurs potentiels de la division cellulaire ont été identifiés chez *Lingulodinium*, comme une cycline associée à la phase M [37] ainsi qu'une CDK (*Cyclin Dependant Kinase*) [38]. Il a de plus été montré, chez le dinoflagellé *Karenia brevis*, que PCNA (*Proliferating Cell Nuclear Antigen*, un facteur de processivité de l'ADN polymérase) subi une modification post-traductionnelle qui corrèle avec sa localisation au noyau lors de la réplication de l'ADN [39]. Finalement, chez le dinoflagellé hétérotrophe *Crypthecodinium cohnii*, l'activité d'une cellulase dégradant la paroi cellulaire est requise afin de permettre la progression de la division cellulaire [40].

### 1.2.3. Régulation génique

La régulation génique chez les dinoflagellés, que ce soit au niveau transcriptionnel ou post-transcriptionnel, est encore plutôt mal comprise. L'architecture du génome, comprenant souvent des gènes présents en plusieurs copies organisées en tandem, sa taille souvent importante, l'absence d'histones, un nombre réduit de facteurs généraux de la transcription et une faible diversité des facteurs spécifiques de transcription contribuent à faire des dinoflagellés un cas unique chez les eucaryotes.

#### 1.2.3.1. Régulation de la transcription

##### 1.2.3.1.1. Organisation génomique

Les dinoflagellés sont reconnus pour avoir des génomes parmi les plus imposants chez les eucaryotes. Des estimations de taille effectuées à l'aide de marqueur fluorescent pour l'ADN montrent une variation importante de la taille des génomes qui se situent entre 1 et 230 pg (entre 0,98 et 225 Gbp) de matériel génétique haploïde par cellule [41] pouvant être réparti sur plus d'une centaine de chromosomes linéaires possédant des télomères similaires à ceux des plantes [42, 43]. Ces grandes tailles ainsi que l'abondante présence de séquences répétées a retardé le séquençage d'un premier génome de dinoflagellé par rapport à plusieurs organismes phylogénétiquement proche comme les ciliés et les diatomées. Ce n'est qu'en 2015 qu'un des plus petits génomes (1,18 milliard de paires de bases), celui du dinoflagellé symbiotique *Symbiodinium kawagutii*, a été séquencé [44]. En plus d'avoir une taille qui peut être impressionnante, le génome des dinoflagellés se retrouve sous forme d'un cristal liquide cholestérique stabilisé par des ARNs et des cations métalliques au lieu de passer par un empaquetage par des histones [42, 45]. Les chromosomes des dinoflagellés seraient ainsi formés de différentes couches d'ADN superposées l'une sur l'autre desquelles sortent des boucles d'ADN périchromosomales, là où se déroule la transcription (Figure 1.2. A-B) [42, 46, 47].

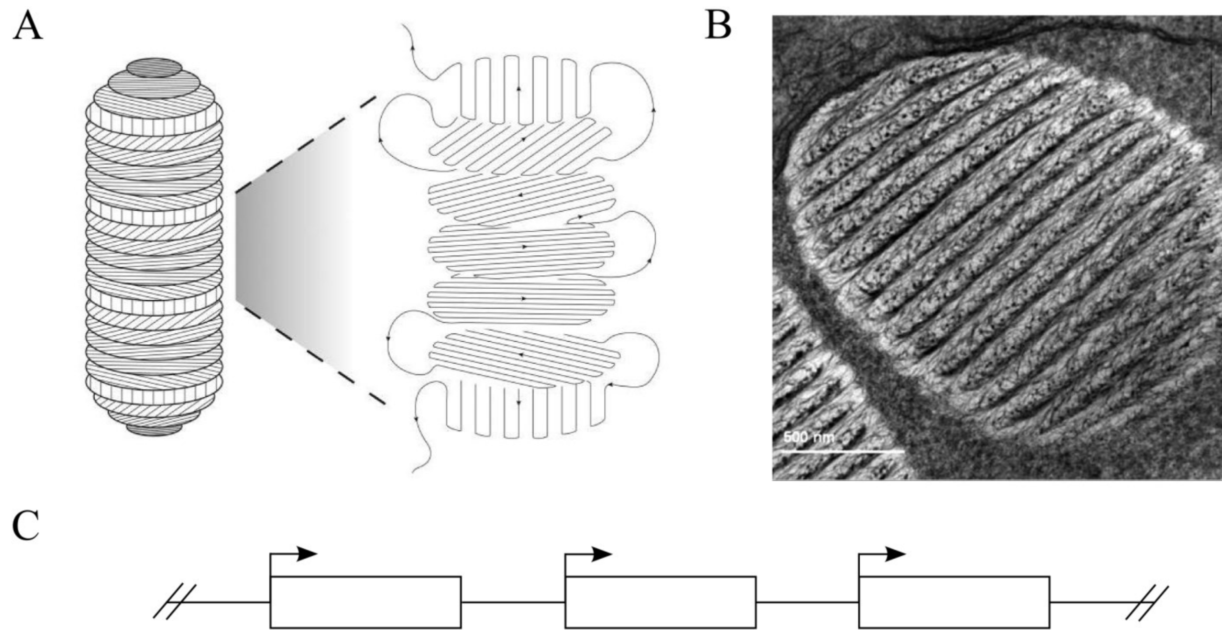



**Figure 1.2. Organisation génomique chez les dinoflagellés**

A. Modèle d'un chromosome de dinoflagellé montrant l'empilement en succession de disques contenant l'ADN génomique. B. Photographie prise au microscope électronique à transmission montrant l'empilement des disques d'ADN d'un chromosome de *Prorocentrum micans*. (A et B tiré de [42]). C. Modèle d'un gène organisé en tandem. Les lignes pleines correspondent aux séquences intergéniques qui ne sont pas transcrites et les boîtes aux séquences qui codent pour un même gène. Les flèches indiquent le lieu du départ de la transcription pour chaque unité codante.

Ce mode de compaction du génome se reflète également dans le ratio protéines:ADN qui, chez les eucaryotes, est d'environ 1:1 tandis qu'il est d'environ 1:10 chez les dinoflagellés [48]. Les histones ont longtemps été considérées comme absentes chez les dinoflagellés et ce n'est qu'avec l'avènement du séquençage à haut débit que des transcrits d'histones et un éventail complet de protéines de modifications post-traductionnelles des histones ont pu être observées [49, 50]. Toutefois, les histones ne semblent pas être traduites à un niveau suffisamment important pour être détectées, ce qui soulève un doute quant à leur utilité réelle dans la régulation de l'expression génique [50, 51]. Des observations microscopiques

complémentaires montrent également que l'organisation classique de l'ADN en une suite de nucléosomes, soit celui d'un collier de perles, est absente chez les dinoflagellés [52].

La perte des nucléosomes et des histones est corrélée avec l'apparition de deux autres types de protéines capables d'organiser l'ADN, soit les *Dinoflagellate/Viral Nucleoproteins* (DVNPs) [12] et les *Histone-like proteins* (HLPs), qui sont similaires à des protéines bactériennes [53-55]. Une HLP provenant de chez *Crypthecodinium cohnii* a été caractérisée comme étant capable d'induire la condensation de l'ADN sous forme de cristal liquide autant *in vitro* qu'*in vivo* [56]. Les autres rôles possibles de ces protéines n'ont toujours pas été étudiés en détail, mais la modification post-traductionnelle des DVNPs par phosphorylation [12] et des HLPs par acétylation [57] ainsi que la localisation des HLPs à la périphérie des chromosomes [54], là où se déroulerait la transcription chez les dinoflagellés [46], suggère un rôle potentiellement important dans la régulation de celle-ci.

L'organisation des gènes dans le génome est aussi particulière chez les dinoflagellés. Certains gènes, comme ceux codant pour LCF (Luciférase), LBP (*Luciferin-binding protein*) et PCP sont organisés en tandem (Figure 1.2. C), c'est-à-dire que de multiples unités codantes (environ 5000 pour PCP) très bien conservées et séparées par une séquence intergénique, sont placées une à la suite de l'autre dans le génome chez *Lingulodinium* [58-60]. On ne sait par contre pas si toutes ces copies sont regroupées à un seul endroit du génome ou bien si elles sont dispersées en plus petits groupes. Les données actuellement disponibles nous laissent croire que les séquences régulatrices pourraient se retrouver dans les séquences intergéniques séparant les unités codantes. Par exemple, pour PCP, un long ARN polycistronique n'a jamais été observé, les séquences intergéniques sont en très faible abondance par rapport à la séquence codante dans les données transcriptomiques [61] et la région codante pour le 3'UTR contient un motif utilisé par les dinoflagellés pour signaler la polyadénylation. Chacune des unités de PCP semble donc être transcrite individuellement, ce qui nécessiterait un nouvel assemblage de la machinerie transcriptionnelle pour chaque séquence codante. Il est aussi intéressant de noter que chez *Amphidinium carterae* les gènes les plus fortement exprimés semblent être en grande partie organisés en tandem et présents en de multiples copies dans le génome tandis que les gènes moins exprimés sont en général présents en une seule copie [62], corrélation qui n'est par contre pas observée chez *Lingulodinium* [61]. Chez *S. kawagutii*, le

séquençage du génome a permis de montrer qu'environ 9 % des gènes sont organisés en tandem et que le nombre de copies de ces gènes est faible, avec au maximum une dizaine de copies [44]. Une situation similaire a été observée lors du séquençage partiel du génome de *Symbiodinium minutum* [63]. Il est possible que la petite taille des génomes du genre *Symbiodinium* ait limité le nombre et l'étendue des gènes organisés en tandem en comparaison du très grand nombre de copies estimé pour *Lingulodinium*. Le séquençage du génome d'autres dinoflagellés dans les prochaines années devrait permettre d'avoir une meilleure idée quant à l'étendue de ce phénomène.

Finalement, le génome des dinoflagellés contient une forte proportion de nucléotides méthylés (par exemple, des thymines sont remplacées par des 5-hydroxyméthyluracils, des cytosines par des 5-méthylcytosines et des adénines par des N6-méthyladénines), et ce de façon non aléatoire, ce qui suggère un rôle potentiel de ces nucléotides dans l'organisation du génome et dans la régulation de la transcription [64]. La transcription des gènes qui codent pour PCP et pour LHC (*Light harvesting complex*) chez *A. carterae* semble d'ailleurs être affectée par la méthylation d'îlots CpG en amont de ces gènes [65] sans toutefois que des protéines ayant une préférence pour ces nucléotides méthylés aient été identifiées à ce jour.

##### 1.2.3.1.2. Machinerie transcriptionnelle et facteurs de transcription

Il n'y a que très peu de détails connus sur de potentiels facteurs de transcription chez les dinoflagellés. En ce qui a trait aux facteurs généraux d'initiation la transcription (TFIIA, B, D, E, F et H), les dinoflagellés montrent une réduction assez importante dans le nombre de séquences qui codent pour ces protéines, une situation qui est répandue chez les chromoalvéolés [66, 67]. TFIID, le facteur qui permet de lier la boîte TATA via la sous-unité TBP (*TATA-box binding protein*) est un cas particulier chez les dinoflagellés. TBP se retrouve sous une forme très divergente nommée *TBP-like factor* (TLF), d'abord découverte chez le dinoflagellé hétérotrophe *C. cohnii* [68]. TLF présente des mutations dans quatre phénylalanines nécessaires pour lier la séquence TATA qui sont probablement responsables du fait que cette protéine est capable de lier une séquence TTTT au lieu d'une séquence TATA lors d'essai *in vitro* [68]. Cette séquence TTTT a d'ailleurs été retrouvée une trentaine de nucléotides en amont du site de départ de la transcription de 94% des gènes dans le génome de

*Symbiodinium*, suggérant un rôle important dans la régulation de l'initiation de la transcription [44]. L'analyse des promoteurs des gènes a également montré un certain enrichissement pour un motif TATGTATG dans les promoteurs de 257 gènes, ce qui suggère un potentiel de régulation spécifique pour les gènes possédant cette séquence dans leur promoteur [44]. Finalement, un motif de 13 nucléotides est conservé dans les séquences intergéniques de LCF et de PCP chez *Lingulodinium*, sans toutefois qu'une analyse plus poussée de la fonction de cette séquence ait été effectuée [59]

Les facteurs de transcription responsables de l'activation ou la répression de gènes précis semblent également être présents en plus faible nombre et en moins grande diversité chez les dinoflagellés. L'analyse des transcriptomes de *Symbiodinium sp.* et de *Lingulodinium* a en effet montré qu'une plus faible partie des gènes exprimés par ces deux dinoflagellés codent pour des facteurs de transcription en comparaison avec d'autres eucaryotes unicellulaires [61, 69]. Plusieurs types de facteurs de transcription retrouvés chez une vaste gamme de protistes sont également absents de ces deux transcriptomes. Par contre, les deux espèces montrent une expansion de la famille des protéines à domaine *cold shock* (CSD), prédites comme étant des facteurs de transcription. Ces protéines, d'abord identifiées chez *E. coli* comme étant capables de lier les ARNs afin de permettre leur traduction lors d'un choc thermique [70], ont ensuite été retrouvées chez tous les grands groupes taxonomiques [71]. Elles sont en général capables de lier tout aussi bien l'ARN que l'ADN et peuvent prendre part à des fonctions très diverses comme la régulation de la transcription [72, 73], l'épissage alternatif [74] et la régulation de la génération des miRNAs [75]. La fonction précise de ces protéines reste donc à élucider chez les dinoflagellés.

Une seule autre protéine associée à la régulation de la transcription a été caractérisée. Il s'agit d'une protéine ayant deux domaines doigt de zinc de type C2H2 et qui est retrouvée dans les zones où il y a de la transcription lors de la phase G1 chez *C. cohnii* [76]. Cette protéine, nommée Dinap1 (*Dinoflagellate nuclear associated protein 1*) est capable d'interagir avec deux protéines, Dip1 et DapC, ayant une certaine homologie avec des facteurs d'épissage [77]. Dinap1 et DapC sont également capables de moduler de façon positive ou négative, respectivement, le niveau général de la transcription lorsqu'ajoutée à des extraits de noyau, suggérant que ces protéines fassent partie d'un complexe associé à l'ARN polymérase II [68].

Finalement, le processus de maturation des ARN en ARNm est complexifié par la présence d'un mécanisme de *trans-splicing*, où certains ARNs nouvellement transcrits sont modifiés en 5' par l'ajout d'une séquence conservée de 22 nucléotides, le *Spliced Leader* (SL) [78]. Ce mécanisme semble surtout présent pour les gènes ayant plusieurs copies arrangées en tandem dans le génome [62] et la proportion et l'identité des ARNm ayant un SL sont potentiellement différentes d'une espèce à l'autre [79-81]. Ce mécanisme, couplé au fait que les séquences intergéniques ne comportent pas de séquences promotrices connues et sont assez variables [82], a ouvert la porte à la suggestion que de longs ARNs polycistroniques provenant des régions où les gènes sont organisés en tandem pourraient être transcrit pour être subséquemment épissés en ARNm fonctionnel ne codant que pour un seul gène, suivant un mécanisme similaire à ce qui est observé chez les euglènes [78, 83]. Cette hypothèse n'a par contre jamais pu être prouvée de façon expérimentale vu l'absence de détection de longs ARNs et des séquences intergéniques [61].

#### 1.2.3.1.3. Études à large échelle

L'étendue et l'amplitude des changements transcriptionnels chez les dinoflagellés ont été relativement peu étudiées par des méthodes à plus haut débit. Parmi celles-ci, notons la mesure des réponses transcriptionnelles lors de changements de salinité [84], de changements dans les concentrations de phosphate et de nitrate [85-87], de changements de densité cellulaire lors de *blooms* [88] et de changements de conditions lumineuses [89]. Les résultats produits par ces études ne permettent pas de déterminer de patron clair quant aux changements d'expression génique. Par exemple, chez *Karenia brevis*, un ajout de nitrate ou de phosphate après une carence pour ces nutriments induit d'importants changements transcriptionnels seulement pour l'ajout de nitrate [86]. À l'inverse, une carence en phosphate chez *Alexandrium tamarense* induit de plus importants changements qu'une carence en nitrate [87] et qui ne sont significatifs que pour moins de 5% des séquences analysées, alors qu'une étude similaire chez une espèce du même genre, *Alexandrium fundyense,* a décelé des changements d'abondance pour plus de 10% des gènes testés dans des conditions similaires [85]. De plus, l'amplitude des changements d'expression observée dans les études ci-haut est en général assez faible pour les transcrits identifiés comme exprimés de façon différentielle. Finalement, la demi-vie globale des ARNm est très longue chez *K. brevis* (médiane de 33 heures),

suggérant une quantité relativement stable des transcrits en tout temps [90]. En jumelant la faible proportion et la faible amplitude des changements transcriptionnels observés au fait que les dinoflagellés ont une portion réduite de leur transcriptome dédiée au contrôle transcriptionnel, il semble que ce type de régulation pourrait jouer un rôle moins important chez ces organismes que chez les autres eucaryotes dans le contrôle de l'expression génique de la cellule.

#### 1.2.3.2. Régulation post-transcriptionnelle

Une grande partie des connaissances sur la régulation traductionnelle chez les dinoflagellés provient de l'étude des rythmes circadiens chez *Lingulodinium*. En effet, la régulation de ces rythmes semble passer surtout par la régulation du moment de la traduction de certains gènes plutôt que par la régulation de la transcription comme cela est le cas chez les autres eucaryotes. Deux études ont montré que le moment de traduction des protéines peut être regroupé en trois groupes différents chez *Lingulodinium*, soit au début du jour, au début de la nuit et au milieu de la nuit [91, 92], bien que le niveau de leurs transcrits soit stable tout au long de la journée. Même si le moment de traduction de certains gènes est connu, très peu d'études se sont penchées sur les acteurs moléculaires permettant de réguler la traduction. Une seule étude a précédemment montré, par des essais de retard sur gel, qu'une protéine serait capable de lier le 3'UTR de LBP pendant le jour, moment où ce gène n'est pas traduit [93]. Une tentative ultérieure d'identification de cette protéine de liaison à l'ARN n'a toutefois pas été fructueuse [94]. Aucun autre facteur modulant la traduction n'a été spécifiquement étudié. Une étude à large échelle a par ailleurs montré que des protéines de liaison à l'ARN étaient phosphorylées de façon différente entre le début du jour et le début de la nuit et que plusieurs des sites de phosphorylation étaient prédits comme étant phosphorylés par la Caséine Kinase 2, une kinase impliquée dans la régulation des rythmes circadiens chez les eucaryotes [95].

À l'opposé de la réduction dans le nombre et la variété des facteurs généraux de transcription, les dinoflagellés montrent une expansion importante de la famille du facteur d'initiation de la traduction eIF4E, possédant environ le double du nombre de séquences qui codent pour ceux-ci lorsque comparé à d'autres protistes [96, 97]. Ces facteurs se retrouvent

principalement sous deux classes (eIF4E-1 et eIF4E-2) séparées en plusieurs sous-classes. Une partie de ceux-ci, comme différentes formes de eIF4E-1a, permettent de réguler l'initiation de la traduction par le ribosome en s'associant à d'autres facteurs d'initiation de la traduction et en liant la coiffe de l'ARNm. Certaines variantes de ce facteur sont aussi impliquées dans la séquestration des ARNm, leur export nucléaire ainsi que l'interaction avec différentes protéines de liaison à l'ARN [98]. En plus de posséder plus de séquences codantes pour les deux premières classes de eIF4E, les dinoflagellés et quelques hétérokontes possèdent des séquences pour une troisième classe, eIF4E-3 [96]. Il a donc été proposé que cette plus grande variation des séquences qui codent pour eIF4E puisse permettre de réguler finement la traduction chez les dinoflagellés et ainsi pallier au faible niveau de régulation transcriptionnelle. Des analyses similaires n'ont par contre pas été effectuées pour les autres facteurs d'initiation, d'élongation et de terminaison de la traduction ainsi que pour les protéines de liaison à eIF4E, qui sont reconnues pour en réguler l'activité.

La régulation de la traduction chez les eucaryotes multicellulaires est également influencée par la présence de micro ARNs (miRNA) qui permettent de moduler la traduction et la stabilité des transcrits. Chez les organismes unicellulaires, la présence et l'utilisation de miRNAs pour réguler la traduction n'ont été que peu étudiées vu que ce mécanisme de régulation était vu comme étant spécifique aux organismes multicellulaires. Les premières évidences de la présence de miRNAs chez ces organismes proviennent de l'algue verte *Chlamydomonas reinhardtii* [99, 100], chez qui une preuve fonctionnelle a également été produite [101]. Le séquençage de plusieurs génomes d'organismes unicellulaires a ensuite permis l'identification de miRNAs chez plusieurs lignées, comme chez les diatomées [102, 103], les apicomplexes [104] et les amibes [105]. Une analyse plus stringente de ces données, notamment quant à la présence de structures secondaires dans les ARNs précurseurs des miRNAs, a par contre remis en doute la présence réelle d'une grande partie de ces miRNAs [106]. L'étendue du rôle que pourrait jouer ce mode de régulation n'est pas encore très bien comprise chez les unicellulaires.

Chez les dinoflagellés, quelques publications mettent de l'avant la possibilité que les miRNAs soient impliqués dans la régulation traductionnelle, en se basant surtout sur des prédictions bio-informatiques. Chez *Symbiodinium kawagutii,* la présence d'une centaine de

miRNAs a été prédite par bio-informatique et par séquençage de petits ARNs [44]. Les cibles potentielles de ces miRNAs incluent beaucoup de gènes chez les coraux, les symbiontes de *Symbiodinium*, suggérant ainsi que le système d'ARN d'interférence pourrait jouer un rôle dans l'établissement et le maintien de leur relation symbiotique [44]. De plus, une vingtaine de nouveaux miRNAs ont été identifié chez *Symbiodinium microadriaticum* par séquençage de miRNAs et ayant potentiellement des cibles assez diverses [107]. Des résultats similaires ont été observés chez *Alexandrium catenella* et *Alexandrium tamarense*, deux espèces non symbiotiques [108, 109]. Chez *Lingulodinium*, le séquençage à haut débit de petits ARNs n'a pas permis d'identifier de miRNAs qui auraient comme cible des transcrits dont la traduction est régulée de façon circadienne. De plus, l'introduction par biolistique de séquences précurseures de miRNAs n'a pas permis d'inhiber la traduction de LBP, un des composants du système de bioluminescence, ce qui, couplé à la très faible abondance de petits ARNs, suggère que le système d'ARN d'interférence n'est peut-être pas fonctionnel chez ce dinoflagellé [110].

Suite à leur traduction, les protéines peuvent être modifiées de plusieurs façons par plusieurs modifications post-traductionnelles, tels la phosphorylation, l'acétylation, l'ubiquitination et la SUMOylation. Toutes ces modifications peuvent moduler l'activité, la localisation, les interactions entre protéines et la stabilité des protéines. De façon similaire à la régulation de la traduction, les modifications post-traductionnelles des protéines chez les dinoflagellés ont surtout été étudiées dans un contexte de régulation circadienne. Ces études ont surtout mis l'accent sur le rôle de la phosphorylation et de la déphosphorylation chez *Lingulodinium*. Une série de publications utilisant des inhibiteurs pour différentes protéines kinase [111, 112] et phosphatase [113, 114] a en effet montré un effet important de ces inhibiteurs sur la rythmicité des cellules et leur capacité à répondre à des changements de conditions lumineuses. Un séquençage de protéine montrant des variations importantes quant à leur niveau de phosphorylation entre le milieu du jour et le milieu de la nuit a permis d'identifier trois protéines de liaison à l'ARN qui sont phosphorylées le jour, mais pas la nuit [115]. De plus, une étude à plus large échelle a permis d'identifier 74 protéines phosphorylées qui avaient un domaine de liaison à l'ARN [95]. Ces évènements de phosphorylation

pourraient potentiellement expliquer les variations du niveau de traduction qui sont observées chez *Lingulodinium* à différent temps de la journée.

Quant à l'ubiquitination, elle est probablement impliquée dans la régulation de PCNA chez *K. brevis*. En effet, une hausse d'environ 9 kDa dans le poids de PCNA, similaire au poids de l'ubiquitine, est observée au moment où il y a réplication de l'ADN, ce qui est également corrélé avec un changement de distribution dans le noyau [39]. Chez plusieurs autres eucaryotes, PCNA est ubiquitiné lorsque le processus de réparation de l'ADN redémarre en présence de traitements qui inhibent la réplication [116], ce qui suggère que le rôle de cette modification post-traductionnelle pourrait être différent chez les dinoflagellés.

## 1.3. Rythmes circadiens

La grande majorité des organismes vivants possèdent des horloges moléculaires leur permettant de synchroniser leur biochimie et leur physiologie avec la période de rotation de la Terre. L'anticipation des changements de luminosité et de température associés à la rotation de la Terre permet à ces organismes d'avoir un avantage compétitif. Par exemple, des mutants pour des gènes de l'horloge chez la plante modèle *Arabidopsis thaliana* ont une croissance et une survie plus faibles lorsqu'elles sont incapables de s'acclimater à des changements dans la durée du jour [117, 118]. Des observations similaires ont été faites chez des cyanobactéries, où la compétition directe entre des souches d'une même espèce ayant des mutations permettant d'avoir des périodes différentes a permis de montrer que plus la période interne était similaire avec la période des conditions de cultures, plus l'avantage compétitif augmentait [119, 120].

Ces horloges internes sont capables de générer des rythmes circadiens qui sont définis par trois caractéristiques principales. Tout d'abord, l`origine endogène de ces rythmes est étudiée en plaçant les organismes en condition de culture constante (appelé LL (*Light-Light)* ou DD (*Dark-Dark*)), en opposition à LD (*Light-Dark*)) en ce qui a trait à la luminosité et à la température. Ces conditions peuvent également être nommées CT (continuous time) ou ZT (Zeitgeber time). Les bases de ce type d'expérience ont été jetées il y a près de 300 ans par le

scientifique français de Mairan qui a observé que les feuilles de la plante *Mimosa pudica* montrent les mêmes mouvements en obscurité qu'en alternance jour/nuit (revu dans [121]). Un phénomène physiologique ou biochimique qui présente un rythme en condition d'alternance jour/nuit doit présenter un rythme d'une période similaire (environ 24h) en condition constante pour être circadien, ce qui implique un contrôle direct par l'horloge circadienne [122]. Ces oscillations d'environ 24h peuvent persister pendant plusieurs semaines chez *Lingulodinium* en LL ou plusieurs mois chez certains rongeurs en DD. Un phénomène qui perdrait sa rythmicité lorsque les conditions sont constantes est un rythme journalier, qui est à la remorque de la perception des changements de luminosité ou de température par différents systèmes de signalisation. Une deuxième caractéristique importante est le phénomène de compensation à la température. Ce phénomène stipule que l'augmentation de la température externe ne devrait pas avoir d'effet sur la vitesse des réactions de l'horloge [123]. Contrairement aux réactions biochimiques classiques qui ont des valeurs de Q10 (augmentation de l'activité suite à une augmentation de la température de 10 °C) de deux à trois fois, les oscillateurs circadiens montrent des valeurs de Q10 proche de 1 [124-126], montrant ainsi une grande stabilité face à des changements importants de température. Cette stabilité est importante vu que les changements de température d'une journée à l'autre n'impliquent en aucun cas un changement dans la durée de rotation de la Terre. Finalement, l'oscillateur circadien doit être capable de se synchroniser aux conditions de culture pour être en phase avec son environnement [122]. Cette synchronisation peut passer par une perception directe de la lumière, de la température ou bien par la perception de changements métaboliques intracellulaires, et est importante étant donné que les cycles circadiens endogènes ne sont pas exactement de 24 heures.

### 1.3.1. Modèles de régulation circadienne

La régulation circadienne contrôle une part importante de la physiologie et du métabolisme de la grande majorité des êtres vivants. Depuis le début des années 80, les composants génétiques formant la base de l'horloge circadienne ont été découverts chez une

grande variété d'eucaryotes [123]. Bien que les gènes codant pour les protéines de l'horloge ne montrent pas d'homologie de séquence entre les grands groupes taxonomiques, l'architecture de l'horloge est similaire [127]. Le modèle généralement accepté chez les eucaryotes stipule que ce contrôle s'effectue par une boucle de rétroaction transcriptionnelle/traductionnelle (*transcription-translation feedback loop, (TTFL)*) où le produit d'un gène est capable, à terme, de réguler sa propre transcription selon une période d'environ 24h. L'horloge du picoeucaryote photosynthétique *Ostreococcus tauri* en est un exemple simple, où deux facteurs de transcription suffisent pour générer les rythmes circadiens [128]. Ce mécanisme est capable d'intégrer des changements dans la durée du jour et dans la température (les intrants de l'horloge) afin de réguler, généralement par transcription, les sorties de l'horloge (Figure 1.3. A). Chez les eucaryotes multicellulaires, entre 1 et 15% des gènes montrent une variation circadienne dans leur abondance [129-132], proportion qui peut aller jusqu'à 80% chez des organismes unicellulaires [133], et qui est corrélé avec le fait qu'une part importante des génomes montre des changements circadiens dans l'occupation des promoteurs [134, 135]. Les horloges peuvent être plus complexes, notamment par l'intégration de multiples signaux et l'interconnexion de plusieurs horloges dans un même système. Par exemple, chez la plante modèle *Arabidopsis thaliana*, trois boucles de rétroaction interdépendantes sont reliées ensemble et régulent séquentiellement une multitude de processus tout au long de la journée [136]. D'autres modèles de régulation circadienne ont émergé dans la dernière décennie. Par exemple, chez les cyanobactéries, l'horloge moléculaire peut être entièrement régulée par phosphorylation et déphosphorylation de la protéine KaiC. De nouvelles évidences métaboliques, notamment quant à la régulation redox, mettent également le rôle des TTFL en perspective chez les eucaryotes et suggère une intégration importante de plusieurs processus distincts par l'horloge circadienne pour mieux réguler par la suite une très large part de la physiologie cellulaire.

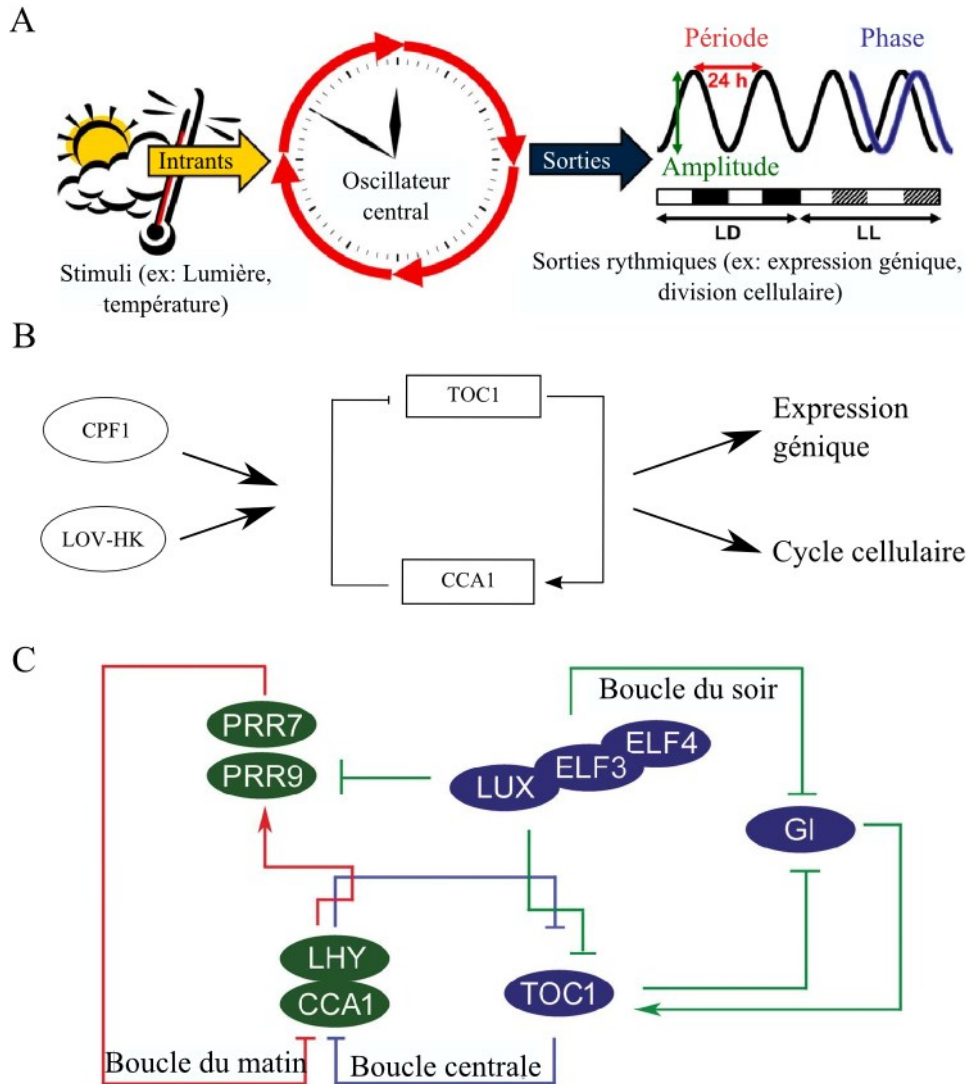



**Figure 1.3. Modèles classiques d'oscillateurs circadiens**

A. Schéma général de l'architecture des horloges circadiennes. Les intrants, comme la lumière et la température, sont perçus par l'oscillateur central, permettant ensuite de synchroniser l'organisme à son environnement en influençant de multiples processus biologiques. Tiré de [137]. B. Horloge transcriptionnel chez l'algue verte *Ostreococcus tauri*. L'oscillateur central, formé de deux facteurs de transcription CCA1 et TOC1, intègre les signaux lumineux par le biais de deux récepteurs, LOVHK et CPF1 et régule, entre autres, la division cellulaire et l'expression génique. C. Oscillateur circadien chez *Arabidopsis thaliana*, qui est formé d'une boucle centrale, composée de TOC1 et LHY-CCA1, et qui est elle-même régulée par la boucle du matin et la boucle du soir. Tiré de [138]

#### 1.3.1.1. Modèle classique de boucle de transcription/traduction : exemple chez *O. tauri* et *A. thaliana*

L'algue verte *O. tauri* a émergé dans les dernières années comme un modèle fort intéressant pour l'étude moléculaire des organismes photosynthétiques. Ce picoeucaryote possède un minuscule génome de seulement 12,5 millions de paires de bases qui est extrêmement compact [139]. Cette relative simplicité ainsi qu'une très faible redondance dans les familles de gènes ont mené à l'utilisation de cet organisme comme un modèle minimal dans l'étude des rythmes circadiens chez la lignée des plantes vertes, qui ont auparavant été très largement étudiés chez *Arabidopsis thaliana*.

**Oscillateur central**

Le système circadien chez *O. tauri* semble être une version simplifiée de celle présente chez *A. thaliana* (Figure 1.3. B-C). En effet, chez cette dernière, trois boucles de rétroaction sont à l'œuvre afin de permettre la régulation circadienne de l'organisme [140]. La boucle centrale est formée de trois gènes codant pour des facteurs de transcription, soit TOC1 (*TIMING OF CAB EXPRESSION*, un membre de la famille des *Pseudo Response Regulators*, ou *PRR*) qui a un pic d'expression au crépuscule et qui contrôle négativement l'expression de deux facteurs à domaine Myb formant un hétérodimère, soit CCA1 (*CIRCADIAN CLOCK-ASSOCIATED 1*) et LHY (*LATE ELONGATED HYPOCOTYL*) [141, 142]. Ceux-ci ont un pic d'expression à l'aube [143] et contrôlent positivement l'expression de deux autres PRR, PRR7 et PRR9 [144], qui font partie de la boucle du matin et qui régulent négativement l'expression de CCA1/LHY [145]. PRR7 et PRR9 sont à leur tour régulés négativement par les membres de la boucle de soir, soit LUX (*LUX ARRYTHMO*), EFL3 (*EARLY FLOWERING 3*) et EFL4, qui sont tous des facteurs de transcription à domaine Myb, ce qui permet de lever l'inhibition de la transcription de CCA1/LHY [146-148].

Probablement en conséquence de la grande diminution de la taille du génome et du nombre de gènes, il n'a été possible d'identifier que deux membres de la boucle centrale de rétroaction chez *O. tauri*, soit CCA1 et TOC1 [149]. LHY et les autres éléments des boucles du matin et du soir ne semblent pas être présents. De façon similaire à *A. thaliana*, TOC1

montre un maximum d'expression au crépuscule, suggérant un rôle important dans la détermination de la durée du jour [150]. CCA1 montre par contre un patron différent en étant exprimé fortement toute la nuit tandis qu'il a un pic à l'aube chez *A. thaliana*. CCA1 est capable de réprimer la transcription de TOC1 en liant un motif conservé appelé la boîte EE (E*vening Element)* présente dans son promoteur [149]. TOC1 semble être le joueur le plus important pour la régulation circadienne vu que des changements dans son niveau d'expression provoquent de l'arythmie alors que la sous-expression de CCA1 ne semble pas avoir d'effet [149, 151].

La dégradation des deux protéines de l'horloge par le protéasome joue également un rôle majeur dans le maintien des rythmes. L'inhibition pharmacologique du protéasome permet en effet d'abolir tout rythme chez *O. tauri* et d'arrêter l'horloge au temps où l'inhibiteur est ajouté [150, 152]. La dégradation de CCA1 est régulée de façon circadienne et montre un maximum pendant le jour tandis que TOC1 est dégradé en réponse à la noirceur au début de la nuit [152]. La dégradation de TOC1 à la noirceur permettrait donc, en plus de l'inhibition de sa transcription par CCA1, de bien marquer la transition du jour vers la nuit. De plus, la régulation au crépuscule est également influencée par la CK2 (*Casein Kinase 2*), une kinase ayant un rôle important dans l'horloge d'une multitude d'organismes. L'addition d'un inhibiteur de CK2 a pour effet d'avancer l'horloge d'environ quatre heures, tandis que la surexpression de CK2 permet d'allonger la période [153]. Un effet similaire est observé lors de la manipulation de la CK1 [154, 155].

Bien que l'horloge de *O. tauri* semble minimaliste par rapport à celle de *A. thaliana*, des analyses mathématiques montrent qu'elle possède une robustesse surprenante face aux changements lumineux qu'elle rencontre durant la journée [156]. Cette robustesse est probablement reliée à l'utilisation de plusieurs voies de perception de la lumière ainsi qu'à différentes fenêtres d'opportunité fonctionnelles pour ces récepteurs [157] et serait une alternative génétiquement moins complexe que les multiples boucles de rétroaction présentes chez *A. thaliana* [158].

**Intrants**

La synchronisation de l'horloge de *O. tauri* à son environnement a surtout été étudiée en regard de la perception de la qualité lumineuse. Celle-ci passe principalement par deux

récepteurs distincts, soit un photorécepteur de type cryptochrome/photolyase (OtCPF1) [159] et un de type LOV-HK (*Light, Oxygen or Voltage sensitive-Histidine Kinase*) [160]. Les deux récepteurs sont capables d'absorber de la lumière bleue qui représente une plus grande proportion de la lumière visible plus on descend dans la colonne d'eau. Le promoteur de OtCPF1 permet une expression circadienne en début de journée et une baisse de son expression par ARN antisens a permis d'allonger la période tout en diminuant l'amplitude des variations circadiennes de l'expression de CCA1 fusionné à un rapporteur luciférase [159]. Un autre récepteur de la même famille, OtCPF2, montre aussi une variation d'expression en alternance jour/nuit, mais pas en condition constante, suggérant potentiellement un autre rôle pour ce récepteur. De façon très intéressante, les deux récepteurs sont capables, dans un système hétérologue, d'inhiber la transcription de la luciférase sous contrôle d'un promoteur lié par CLOCK:BMAL1, des facteurs de transcription de type bHLH de l'horloge des mammifères [159]. Une observation similaire a d'ailleurs été effectuée chez la diatomée *P. tricornutum* avec le récepteur PtCPF1 [161]. Ceci suggère que le rôle de OtCPF1 pourrait être similaire à celui des cryptochromes des mammifères qui inhibent directement les protéines de l'oscillateur central.

Le deuxième récepteur est de type LOV-HK, un gène retrouvé majoritairement chez les cyanobactéries et observé dans les génomes des diatomées [162] et de certains prasinophytes comme *O. tauri* mais pas chez les plantes terrestres [160], suggérant un transfert horizontal de gènes entre microorganismes marins. Le domaine LOV est par contre présent chez des protéines qui sont impliquées dans la régulation circadienne comme ZTL (Zeitlupe) qui régule la stabilité de TOC1 chez *A. thaliana* [163]. Le patron d'expression de LOV-HK est assez similaire à celui de OtCPF1, avec une plus grande abondance de l'ARNm à l'aube et un pic d'abondance de la protéine au milieu de la journée, et ce tout aussi bien en condition d'alternance jour/nuit qu'en condition constante [160]. La traduction du récepteur peut être régulée par la lumière alors que la quantité du récepteur fusionné à la luciférase augmente en quelques minutes après l'exposition des algues à la lumière bleue. La sous et la surexpression de LOV-HK permettent d'abolir la rythmicité du rapporteur CCA1-luciférase, suggérant un rôle important dans la régulation de l'oscillateur circadien [160]. Une modélisation mathématique a également montré que LOV-HK ainsi qu'une autre histidine kinase, couplée

cette fois-ci avec une rhodopsine, sont probablement à la base de la synchronisation de l'horloge par la lumière bleue et verte chez *O. tauri* [164].

Chez *Arabidospis*, la perception de la lumière rouge se fait principalement par cinq gènes codant pour des phytochromes [165]. Aucun phytochrome n'a pu être identifié dans le génome de *O. tauri*, bien que *Micromonas pusilla*¸ un genre proche, en possède un, suggérant une perte de ce gène lors de la réduction du génome de *O. tauri* [166]. Chez *Micromonas*, le phytochrome est exprimé tout juste avant l'aube et pourrait réguler la transcription de gènes reliés à la photosynthèse. Des observations similaires ont été effectuées chez les diatomées [167]. Le mécanisme de perception de longueur d'ondes plus courtes par les phytochromes ne semble donc pas être présent chez *O. tauri*, bien que la surexpression de LOV-HK soit capable de raccourcir la période de l'horloge dans la lumière rouge [160].

L'horloge d'*Arabidopsis* est également capable d'intégrer des signaux métaboliques, plus particulièrement les sucres provenant de la photosynthèse, et ce par le biais de PRR7 [168]. La possibilité de ce type de régulation n'a pas encore été testée chez *O. tauri* mais comme PRR7 ne semble pas être présent, l'intégration de ce signal vers l'horloge passerait par un autre mécanisme. De plus, un rythme circadien dans l'oxydation des peroxyrédoxines, qui se déroule de façon indépendante de la transcription [169] (voir section 1.3.1.2.2.), pourrait être un indicateur de la perception du statut métabolique de la cellule et son utilisation par l'horloge.

**Sorties**

Les sorties de l'horloge de *O. tauri* qui ont été analysées sont peu nombreuses, se limitant à l'analyse des changements transcriptionnels et le contrôle du cycle cellulaire. L'analyse à large échelle de l'abondance des transcrits par micropuce à ADN a permis de montrer que près de 80% des gènes de cette algue ont des variations circadiennes dans leur abondance [133]. De plus, l'expression des gènes reliés à des processus précis comme la photosynthèse et le contrôle du cycle cellulaire sont généralement regroupés ensemble. Par exemple, les gènes de la photosynthèse sont exprimés à la fin de la nuit et au début du jour, probablement de façon à ce que la cellule soit préparée à effectuer la photosynthèse dès l'apparition de la lumière. Les régulateurs du cycle cellulaire ainsi que des protéines de

réparation de l'ADN sont quant à eux exprimés vers le milieu de la journée, au moment de la réplication de l'ADN [133]. Des observations similaires ont été effectuées chez *A. thaliana*, où entre 5 et 15% des gènes sont sous contrôle circadien [132] et où la transcription des gènes de la photosynthèse, de réponse aux stress biotiques et abiotiques et de la croissance et du développement de la plante sont regroupés [170-172]. De plus, de multiples étapes régulant la transcription et la traduction ont été liées à l'horloge circadienne dans les dernières années, comme la modification de la chromatine, l'épissage alternatif et la production de miRNA [173]. Ces différents niveaux de régulation n'ont pas encore été étudiés chez *O. tauri*.

L'horloge de *O. tauri* contrôle aussi partiellement le cycle de division cellulaire. Les algues montrent clairement un rythme dans le moment de division, qui se déroule au début de la phase de nuit subjective en condition constante [174]. Ce processus est corrélé avec l'expression circadienne de quelques gènes régulant la division cellulaire, comme des cyclines et des kinases cycline-dépendantes, qui est indépendante du passage ou non de la cellule en phase S [175]. Il a été montré que la prise de décision de la cellule à passer en phase S pour répliquer son ADN est régulée par la durée et l'intensité lumineuse d'un *pulse* de lumière lorsque les cellules sont placées en condition de noirceur [175]. Celle-ci requiert un *pulse* d'au moins quatre heures d'une lumière intense pour passer en phase S et permettrait ainsi d'assurer à la cellule qu'elle a les ressources énergétiques suffisantes pour effectuer la division cellulaire. Il a d'ailleurs été montré qu'un certain niveau d'AMP cyclique est nécessaire pour permettre le début de la réplication de l'ADN [175].

#### 1.3.1.2. Émergence de modèles à oscillateurs post-traductionnel

Jusqu'à récemment, la régulation circadienne de la transcription par les horloges et l'accumulation rythmique des transcrits qui s'ensuit était vue comme la cause majeure de la présence de rythmes chez la majorité des organismes. Les dernières années ont vu plusieurs publications mettre de l'avant le rôle important de la régulation traductionnelle et post-traductionnelle pour générer et maintenir des rythmes circadiens. Par exemple, l'observation qu'une protéine de l'horloge de la cyanobactérie *Synechococcus elongatus*, KaiC, était phosphorylée selon un rythme d'environ 24h dans une réaction *in vitro* en présence des deux autres protéines de l'horloge (KaiA et KaiB) et d'ATP a montré la possibilité d'obtenir des

rythmes circadiens en absence totale de transcription [176]. Un rythme circadien dans l'oxydation des peroxyrédoxines a également été observé chez une vaste gamme d'organismes, suggérant une apparition antérieure de ce mécanisme circadien par rapport aux TTFL [177]. Ces observations remettent en question la centralité du modèle classique de l'horloge moléculaire classique à base de TTFL et font que l'on parle maintenant d'un réseau de régulation circadien plutôt que d'un modèle linéaire.

#### 1.3.1.2.1. L'oscillateur biochimique des cyanobactéries

Les rythmes circadiens chez les procaryotes ont longtemps été considérés comme absents par le fait que ces cellules se divisent à une fréquence très élevée, ce qui limiterait l'utilité d'avoir des rythmes circadiens, et par leur relative simplicité, qui était vue comme un empêchement à la présence d'un oscillateur circadien. Ce n'est que vers la fin des années 80 que les premiers rythmes circadiens dans la fixation du nitrate [178] et l'acquisition d'acides aminés [179] ont été formellement identifiés chez *Synechococcus* sp. RF-1, une cyanobactérie. Une autre espèce du même genre, *S. elongatus* PCC 7942, a ensuite été très largement utilisée afin de disséquer l'horloge moléculaire chez ces organismes, qui est très différente de celle des eucaryotes. Il est par contre important de noter que les cyanobactéries sont les seules bactéries pour lesquelles des rythmes circadiens ont été observés et étudiés.

##### Oscillateur central

L'oscillateur moléculaire des cyanobactéries comprend trois gènes, *KaiA*, *KaiB* et *KaiC* qui sont organisés en régulon et encodées par deux transcrits, un codant pour *KaiA* et un autre pour *KaiB* et *KaiC*. L'activité des promoteurs de ces gènes et l'accumulation des transcrits se font de façon rythmique et ont un pic vers le début de la nuit [180]. Une variation d'environ trois fois dans la quantité de protéine n'est par contre observée que pour KaiB et KaiC, avec un pic d'abondance quelques heures après le début de la phase de nuit [181, 182]. L'inactivation de l'un ou l'autre de ces trois gènes est suffisante pour abolir tout rythme circadien chez *Synechoccocus* [180].

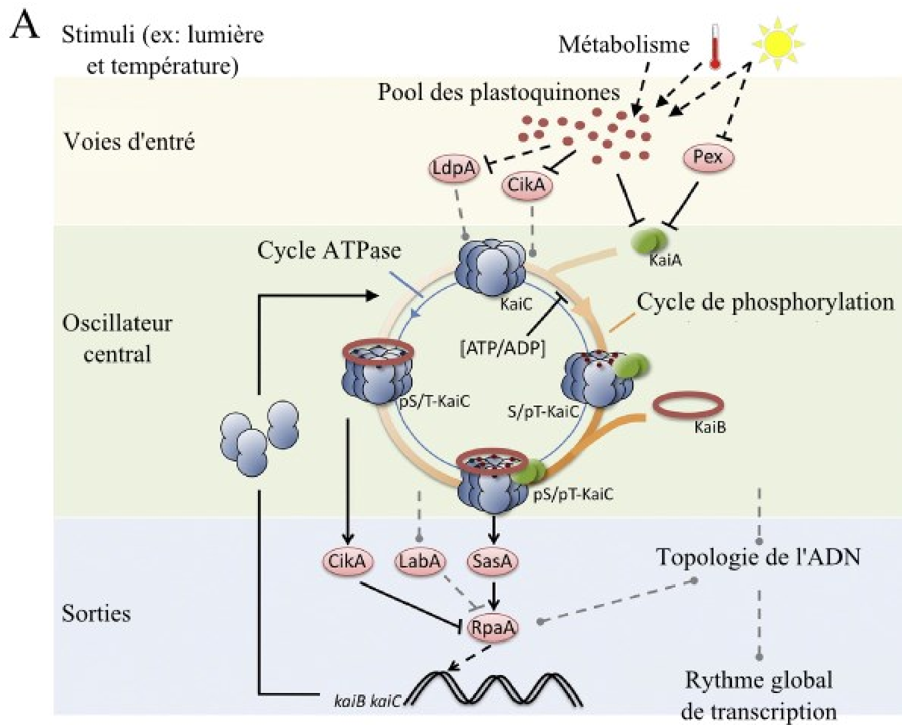
Une propriété exceptionnelle de l'oscillateur de *Synechococcus* est le fait qu'il est fonctionnel sans l'apport de la transcription et la traduction des gènes de l'oscillateur. En effet, les cycles de phosphorylation-déphosphorylation sont observables *in vivo* chez des

cyanobactéries mises en condition de noirceur constante, ce qui cause l'arrêt de la transcription de *KaiBC* [183]. Cette transcription est par contre importante afin de donner une certaine robustesse à l'horloge dans des conditions où les bactéries se divisent rapidement et où le contenu cellulaire est dilué plusieurs fois en un court laps de temps [184, 185]. Une expérience encore plus remarquable a montré qu'il est possible de faire fonctionner l'oscillateur *in vitro* avec une période d'environ 24h, et ce, pendant plusieurs cycles en mettant tout simplement les trois protéines purifiées en présence d'ATP [176]. Cette propriété a permis une caractérisation intensive des différentes protéines composant l'oscillateur central, leur rôle dans le maintien d'une oscillation d'environ 24h ainsi que des mutations affectant la vitesse et l'amplitude de l'horloge.

Des trois protéines composant l'oscillateur moléculaire, KaiC est la pièce maitresse, KaiA et KaiB jouant un rôle de modulateur de KaiC. Celle-ci est composée de deux domaines dupliqués en tandem (domaine CI et CII) et d'une queue C-terminale (A-loop) [180]. Elle possède des activités autokinase, autophosphatase et ATPase et s'assemble en homohexamère pour former une structure cylindrique [186, 187]. Bien que les deux domaines ont des activités kinase et phosphatase fonctionnelles, la phosphorylation qui est capable de réguler la fonction de la protéine s'effectue uniquement sur des résidus du domaine CII tandis que l'ATP est lié par le domaine CI [188].

La phosphorylation et la déphosphorylation séquentielle de KaiC sont responsables de son activité circadienne (Figure 1.4. A). L'autophosphorylation de KaiC sur deux résidus se fait de façon séquentielle à partir du début du cycle circadien [189, 190]. Tout d'abord, le résidu T432 est phosphorylé, suivi par la phosphorylation du résidu adjacent, S431. Ces deux étapes de phosphorylation sont stimulées par la présence de dimères de KaiA, qui, en interagissant avec la A-loop de KaiC, favorise son activité autokinase [191]. Lorsque la protéine KaiC est dans un état hyperphosphorylé au début de la nuit, un changement dans la conformation des domaines CI et CII permet l'exposition d'un autre motif, le B-loop, qui permet à KaiB d'interagir avec KaiC [192]. Cet évènement est concomitant avec un changement dans le repliement de KaiB, qui passe d'un état tétramérique à un état monomérique [193]. Ceci lui permet non seulement d'interagir avec le domaine CI de KaiC sous forme de monomère, mais aussi de séquestrer KaiA et ainsi limiter sa fonction

stimulatrice de l’activité kinase de KaiC [193]. Cette forme de KaiB est également capable de déplacer SasA, un régulateur des sorties de l’horloge, de son site de liaison sur le domaine CI [194]. La déphosphorylation de KaiC peut alors être amorcée et s’effectue en sens inverse de la phosphorylation (T432 suivi de S431), ce qui permet de donner une direction au mécanisme de l’horloge [189, 190]. Une fois KaiC complètement déphosphorylé vers la fin de la nuit, le cycle peut recommencer par le biais de la liaison de KaiA sur la A-loop.



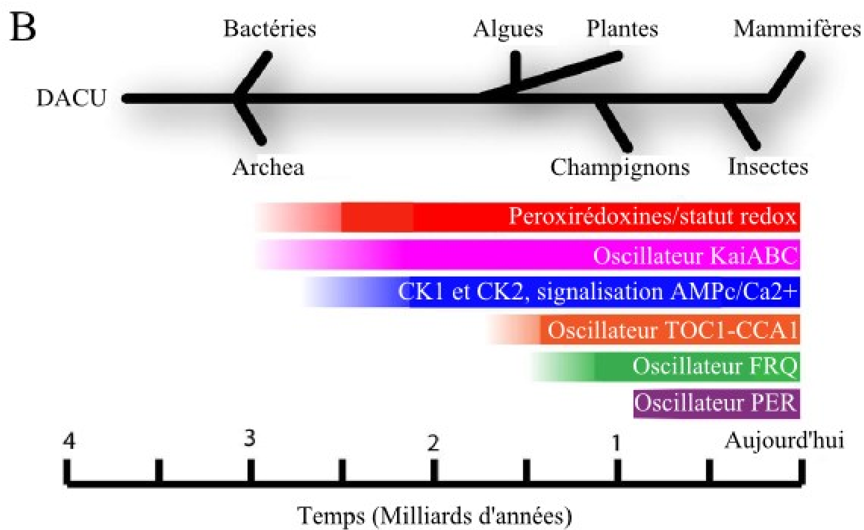




**Figure 1.4. Oscillateur circadien post-traductionnel**

A. Schéma de l’oscillateur KaiABC chez *Synechococcus elongatus* PCC 7942. Tiré de [195]. B. Évolution des différents types d’oscillateurs circadiens. CK1/2 : Caséine kinase 1 et 2; Oscillateur des plantes : TOC1-CCA1; Oscillateur des champignons : FRQ; Oscillateur des métazoaires : PER; DACU Dernier ancêtre commun universel. Tiré de [196].

L'état hyperphosphorylé de KaiC enclenche également un élégant système d'échange des monomères de KaiC entre les hexamères présents dans la cellule vers le début de la nuit [197]. Ces échanges permettent de synchroniser l'oscillateur à l'échelle de la cellule et de tenir compte des effets de dilutions amenés par la division cellulaire ainsi que la synthèse de nouvelles unités qui ne sont pas au préalable phosphorylées [198, 199], participant ainsi à la robustesse du mécanisme circadien.

La présence des trois gènes codants pour l'oscillateur est relativement bien conservée chez d'autres cyanobactéries que *S. elongatus*. L'origine de ces gènes est par ailleurs très ancienne pour *KaiC* (environ 3,5 milliards d'années) et *KaiB* (environ 2,3 milliards d'années) par rapport à *KaiA* (environ 1 milliard d'années), ce qui explique la présence de *KaiC* dans la grande majorité des clades de cyanobactéries (Figure 1.4. B) [200]. Certains genres montrent également des changements dans l'organisation de l'oscillateur. Par exemple, KaiA a été éliminée du génome des cyanobactéries du genre *Prochlorococcus* [201] tandis que KaiB et KaiC sont parfois présents en de multiples copies, comme chez le genre *Synechocystis* [200]. L'absence de *KaiA* est par ailleurs corrélée avec une incapacité à maintenir, en condition constante, des rythmes dans l'abondance des transcrits de certains gènes et dans le contrôle du cycle cellulaire [201].

**Intrants**

La perception du temps et la synchronisation de l'oscillateur central chez les cyanobactéries passent en grande partie par la perception de changements métaboliques reliés à la photosynthèse et au statut redox de la cellule plutôt que par la perception directe des changements de luminosité par des photo-récepteurs comme chez les eucaryotes. Parmi ces changements métaboliques, le ratio ATP:ADP et la présence de plastoquinone oxydée sont deux marqueurs importants pour la synchronisation directe ou indirecte de l'oscillateur central (figure 1.4. A). Tout d'abord, le passage des cellules de la lumière à la noirceur réduit graduellement le ratio ATP:ADP, signe d'un changement dans la balance énergétique de la cellule. Celui-ci peut être directement perçu par KaiC, qui voit son activité autokinase être réduite plus la quantité d'ADP augmente [202]. Le passage du jour vers la nuit produit également une hausse abrupte de la quantité de plastoquinone oxydée due à l'arrêt de la photosynthèse. L'addition de ce métabolite, que ce soit à des cellules *in vivo* ou directement à

l'oscillateur *in vitro*, permet d'ailleurs de reproduire les effets d'un passage vers la noirceur [203]. Ce métabolite peut directement être lié par le domaine PsR (*Pseudo-receiver*) de KaiA, ce qui cause son oligomérisation et sa dégradation, réduisant ainsi la stimulation de l'activité autokinase de KaiC [204]. La transcription de *KaiA* est par ailleurs réprimée vers la fin de la nuit par Pex (*Period extender*) [205], une protéine de liaison à l'ADN qui semble être capable de finement réguler le moment de transition entre la nuit et le jour, probablement par la perception de changements dans la qualité de la lumière [206].

Deux gènes ont également été identifiés comme étant capables de relier le statut du pool de plastoquinone à l'oscillateur central et ainsi complémenter le rôle de KaiA. Le premier, *cikA* (*circadian input kinase*) a été identifié dans un mutant ayant une période d'environ 22h en condition constante et étant incapable de réinitialiser l'horloge après un passage de 5h à la noirceur [207]. Cette protéine possède trois domaines distincts : un domaine GAF, un domaine histidine kinase et un domaine PsR. Les deux domaines situés aux extrémités de la protéine servent à réguler l'activité kinase du domaine central. Le domaine GAF favorise l'autophosphorylation du domaine histidine kinase, permettant ainsi d'augmenter son activité tandis que le domaine PsR inhibe son autophosphorylation [208, 209]. Le domaine PsR permet à CikA, de façon similaire à KaiA, de lier la plastoquinone oxydée, mais pas la plastoquinone réduite [210]. Cette liaison permet à CikA de former un complexe avec KaiA au pôle de la cellule, ce qui provoque la dégradation des deux protéines et par le fait même la baisse de la stimulation de l'activité autokinase de KaiC [204]. CikA est également impliqué dans la régulation des sorties de l'horloge (voir section ci-dessous).

Un deuxième gène, *LdpA* (*light dependent period*) a été identifié lors d'un criblage génétique à la suite duquel les cellules n'arrivaient pas à parfaitement synchroniser l'horloge à de nouvelles conditions lumineuses [211]. Ce gène code pour une protéine de la superfamille des ferrédoxines possédant deux clusters 4Fe-4S. Il a été montré que LdpA est capable d'interagir avec KaiA ainsi qu'avec deux protéines impliquées dans les sorties de l'horloge, CikA et SasA [212]. Ces interactions et l'abondance de LdpA sont modulées par la quantité de plastoquinone réduite présente dans la cellule, ce qui permet de lier l'activité photosynthétique à l'oscillateur central et aux sorties de celui-ci [212]. La perception de la baisse graduelle du ratio ATP:ADP durant la nuit ainsi que de l'apparition abrupte de plastoquinone oxydée au

tout début de la nuit permettent donc à l'oscillateur central des cyanobactéries de mesurer de façon élégante la durée de la phase de nuit ainsi que le moment précis du début de celle-ci pour se synchroniser à son environnement.

**Sorties**

La régulation circadienne exercée par l'oscillateur affecte plusieurs processus physiologiques et métaboliques chez *S. elongatus*, comme le contrôle de la transcription d'une grande majorité des gènes et le contrôle du cycle cellulaire (Figure 1.4. A). Tout d'abord, il y a un contrôle important dans le moment de la transcription d'une grande partie du génome. L'insertion aléatoire dans le génome d'un gène codant pour une luciférase bactérienne, luxAB, a permis de montrer que les gènes sont exprimés majoritairement en deux classes d'expression, soit lors des deux transitions entre le jour et la nuit [213]. Les gènes dits de classe 1, dont font partie *kaiBC*, sont exprimés lors du crépuscule alors que les gènes de classe 2 sont exprimés à l'aube. De façon complémentaire, l'utilisation de micropuce a montré qu'entre 35 et 65% des gènes avaient une abondance significativement différente à certains moments du cycle circadien [214, 215].

Deux mécanismes sont responsables du contrôle de la transcription, soit la régulation d'une centaine de gènes par le facteur de transcription RpaA [216] et le contrôle subséquent de la transcription du reste du génome par des changements de topologie des chromosomes [215]. RpaA est le régulateur le plus important des sorties transcriptionnelles de l'horloge. Sa délétion mène à une arythmie complète de l'expression des gènes, et ce même si l'oscillateur central est toujours fonctionnel [216, 217]. Il est régulé de façon positive par SasA, qui est capable de stimuler son autophosphorylation après avoir interagi avec KaiC hyperphosphorylé [218, 219]. Cette stimulation se fait graduellement tout au long de la phase de jour pour atteindre un maximum de RpaA phosphorylé à la transition jour/nuit [217]. À ce moment, RpaA est capable à la fois d'activer la transcription des gènes de classe 1 et de réprimer les gènes de classe 2. La désactivation de RpaA se fait par le biais de l'activité phosphatase de CikA, qui est activée par son interaction avec KaiB lorsque celui-ci est lié à KaiC lors de sa phase de déphosphorylation [218]. Le facteur de transcription RpaB, un paralogue de rpaA, semble avoir un rôle d'antagoniste à RpaA bien qu'il ne soit pas régulé directement par l'horloge circadienne [220]. En effet, il est capable de lier un élément présent dans le

promoteur de plusieurs gènes régulés par RpaA, dont *KaiBC*, au début de la phase de jour et servirait à réguler négativement la transcription de ces gènes [221]. Il permet également de stimuler la déphosphorylation de RpaA afin de le désactiver [220]. La centaine de promoteurs liés par RpaA incluent *kaiBC*, des régulateurs globaux de la transcription comme des facteurs sigma de groupe 2 ainsi que des régulateurs de la division cellulaire, dont ftsZ et sepF. De façon intéressante, des mutants pour ce type de facteur sigma montrent des changements dans la phase et l'amplitude de l'expression de plusieurs gènes [222], suggérant un rôle potentiel en aval de RpaA.

En ce qui a trait au changement circadien dans la topologie de l'ADN, l'observation microscopique des chromosomes a permis de montrer que ceux-ci se compactent graduellement durant la phase de jour pour atteindre une compaction maximale au crépuscule, phénomène qui s'inverse complètement au fil de la nuit [219]. Le mode de compaction du génome est encore plutôt mal compris. L'observation que KaiC, qu'il soit phosphorylé ou non, a la capacité de lier l'ADN *in* vitro [187] a mené à la suggestion qu'il soit directement responsable de la compaction/décompaction du génome sans toutefois que son implication directe dans le phénomène n'ait pu être démontré.

L'horloge contrôle également le moment de la division cellulaire chez *Synecoccocus* par le biais de la fermeture de la fenêtre d'opportunité pour la cellule d'effectuer la cytocinèse au début de la nuit [223, 224]. Ce phénomène de *gating* n'est d'ailleurs pas observable dans des mutants *cikA* et *kaiB* placé en condition constante [225]. Ceux-ci montrent également un défaut dans la localisation de ftsZ, une protéine essentielle à la division cellulaire qui se retrouve normalement au milieu de la cellule en division, et qui, dans ces mutants, se trouve à être distribuée de façon plus ou moins aléatoire tout le long de la cellule [225]. L'expression du régulateur RpaA muté pour mimer un état phosphorylé permet également de simuler le phénomène de *gating* de la division cellulaire alors que la surexpression de RpaA non phosphorylé abolie le phénomène [216]. Bien que le promoteur de *ftsZ* et celui de son régulateur *sepF* soient des cibles de RpaA, le niveau de transcrits de ces deux gènes ne change pas de façon significative, ce qui suggère qu'un mécanisme encore inconnu relie RpaA au contrôle du cycle cellulaire [216].

#### 1.3.1.2.2. Oscillations post-transcriptionnelles chez les eucaryotes

Plusieurs observations faites au cours des dernières décennies montraient que des rythmes circadiens chez certains organismes ne passaient pas nécessairement par le modèle des TTFL, nommément un contrôle des rythmes de la cellule par une transcription rythmique d'une part plus ou moins importante des gènes. Ces observations, comme le fait que plusieurs rythmes chez *Lingulodinium* passent entièrement par le contrôle traductionnel ou post-traductionnel [226], qu'un groupe de gènes restaient rythmique dans un mutant de l'horloge chez la drosophile [227] et que les rythmes se poursuivent pendant plusieurs jours dans des cellules énuclées de l'algue verte macroscopique *Acetabularia* [228] étaient vu comme des curiosités plus ou moins importantes [229]. La découverte de l'oscillateur *KaiABC* chez les cyanobactéries et de la présence de rythmes d'oxydation des péroxyredoxines chez une très grande diversité d'organisme a ramené à l'avant-plan l'apport de mécanismes post-transcriptionnels dans l'horloge et à l'émergence de modèle d'oscillateur non transcriptionnel. Ces mécanismes sont également bien mieux conservés entre des organismes distants que les acteurs des TTFL, suggérant ainsi qu'ils sont les mécanismes originaux de la régulation circadienne (Figure 1.4. B).

Une découverte majeure des dernières années est l'observation d'un rythme circadien dans l'oxydation et la réduction des peroxyrédoxines. Ces protéines sont impliquées dans la détoxification et la signalisation du $H_2O_2$, un marqueur du statut redox de la cellule [230]. Le groupement thiol d'un résidu cystéine de ces protéines, qui forment des homodimères, peut être oxydé pour former de l'acide sulfénique. À ce stade, ce groupement peut former un pont disulfure avec une autre cystéine et être ensuite recyclé par une glutharédoxine. Une petite fraction des peroxyrédoxines peut passer à un état hyperoxydé et former des groupements d'acide sulphinique [230]. Ceux-ci peuvent à leur tour être réduits de façon ATP-dépendante par des sulphirédoxines pour retourner à un état oxydé [231]. Un rythme circadien dans l'accumulation des peroxyrédoxines hyperoxydées a tout d'abord été observé dans des érythrocytes humains [232]. Étant donnée l'absence de noyau, le maintien de ce rythme pendant quelques jours et sa synchronisation avec des changements de température se fait de façon totalement indépendante de la transcription. Une observation similaire a été faite chez *O. tauri*, alors que le même rythme dans l'hyperoxydation des peroxyrédoxines est capable de

persister pendant cinq jours en présence d'inhibiteurs de la transcription et de la traduction ainsi que dans des conditions de noirceur constante, ce qui provoque l'arrêt complet de la transcription [169]. Un rythme dans l'import mitochondrial redox-dépendant des sulphirédoxine a été montré chez la souris, permettant ainsi de recycler les peroxyrédoxines selon un cycle de 24 heures [233]. Ce rythme dans l'hyperoxydation des peroxyrédoxines a également été observé chez plusieurs organismes provenant des trois grands groupes d'organismes vivants et chez des mutants arythmiques de l'horloge chez la drosophile, l'ascomycète *Neurospora crassa* et *Synechoccocus*, suggérant qu'il pourrait être le mécanisme de contrôle circadien le plus ancestral [177]. Il n'est par contre pas encore connu si ce rythme dans l'hyperoxidation des peroxyrédoxines est réellement capable de contrôler d'autres rythmes ou bien s'il s'agit seulement d'un marqueur du statut redox de la cellule [234]. De façon intéressante, un rythme dans l'expression et l'activité de la superoxyde dismutase, une enzyme qui transforme le superoxyde en $O_2$ et en $H_2O_2$, a aussi été observé chez plusieurs organismes [235-237] et pourrait potentiellement jouer un rôle dans les rythmes redox. D'autres marqueurs redox, comme le ratio $NAD(P)^+$:NAD(P)H et la concentration en glutathion montrent des variations circadiennes chez les plantes et les animaux et pourrait potentiellement réguler des composants de l'horloge [238]. De plus, l'oscillation circadienne de certains métabolites, comme le magnésium, permet de réguler la balance énergétique de la cellule par son association avec l'ATP et par le fait même une grande partie des processus cellulaires [239].

Finalement, d'autres mécanismes de régulation post-traductionnels de l'horloge, comme la phosphorylation, la SUMOylation et la dégradation des protéines sont bien conservés chez une grande variété d'organismes. Par exemple, les *casein kinase* 1 [155, 240, 241] et 2 [242-244] ainsi que la *glycogen synthase kinase 3* (GSK3) [245, 246] sont impliquées dans la phosphorylation de protéines de l'horloge chez plusieurs organismes, même si l'identité de celle-ci n'est pas conservée entre les grands groupes. L'inhibition ou l'activation pharmacologique de ces protéines chez différents organismes mène d'ailleurs à des changements similaires dans la période de l'horloge, suggérant que ces mécanismes sont bien conservés [169]. Il faut toutefois noter que la nature des signaux qui régule l'activité de ces protéines dans un contexte circadien n'est pas encore très bien comprise [247, 248].

### 1.3.2. Rythmes circadiens chez *Lingulodinium*

Le dinoflagellé *Lingulodinium polyedrum*, précédemment connu sous le nom de *Gonyaulax polyedra* et membre du groupe des Gonyaulaucales (Figure 1.1. A-B), a été un des premiers modèles utilisés extensivement en biologie circadienne dû en partie à sa production de bioluminescence servant de rapporteur facilement observable des rythmes circadiens. En plus de ce phénomène, quelques autres phénomènes rythmiques ont été étudiés comme la photosynthèse, la division cellulaire et la migration verticale dans la colonne d'eau (Figure 1.5.). *Lingulodinium* est également un des seuls exemples où la régulation des rythmes circadiens passe par la traduction des gènes à des temps précis plutôt que par l'accumulation et la dégradation des ARNm [226].

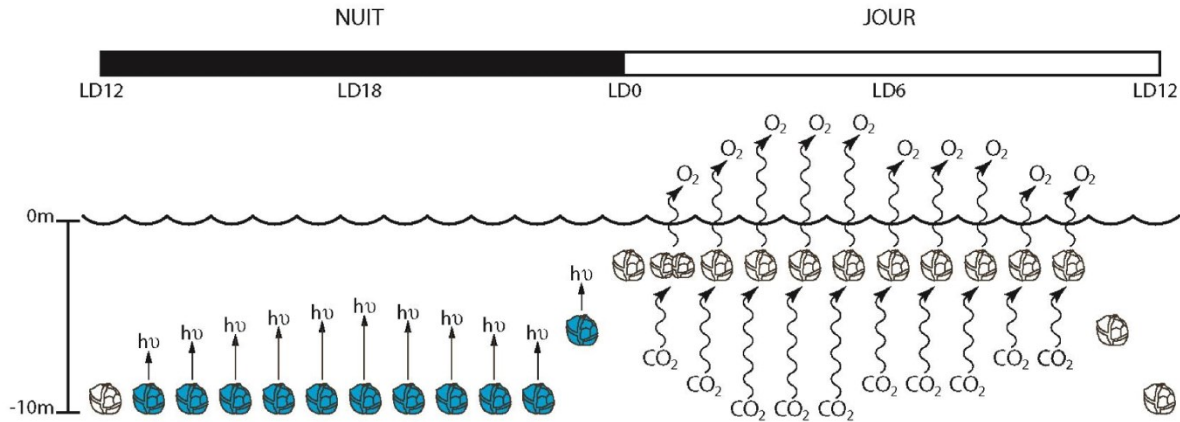



**Figure 1.5. Rythmes circadiens observés chez le dinoflagellé *Lingulodinium polyedrum***

Schémas des différents rythmes observables chez *Lingulodinium* pendant la nuit (barre noire, de LD12 à LD0) et pendant le jour (barre blanche, LD0 à LD12). L'émission de bioluminescence se fait par l'organisme pendant la nuit. À la fin de la nuit, l'algue entreprend une migration vers le haut de la colonne d'eau afin de maximiser la quantité de lumière reçue pour effectuer la photosynthèse pendant le jour. C'est également pendant cette transition que s'effectue la division cellulaire. Tiré de [249].

#### 1.3.2.1. Mécanisme de l'horloge chez *Lingulodinium*

Bien qu'ayant servi de modèle circadien pendant plus de six décennies le mécanisme central de l'horloge circadienne chez *Lingulodinium* n'est toujours pas connu. Quelques études ont montré que *Lingulodinium* pourrait posséder au moins deux horloges circadiennes différentes [250]. Tout d'abord, lorsque l'algue est placée en condition constante, les pics intenses de bioluminescence (le total des éclairs produits par des cellules individuelles) conservent un rythme d'environ 24h tandis que la lueur faible émise par les algues, possiblement un résultat de la dégradation des scintillons, avait une période plus courte d'une heure en condition constante [250, 251]. L'augmentation de la température modifie également de façon différente les périodes de ces deux phénomènes. Le traitement des cultures en condition constante avec de la lumière rouge a de plus permis de montrer que les rythmes d'agrégation et de bioluminescence ne possèdent pas la même période et réagissent

différemment aux deux couleurs de lumières [250, 252]. Ces données montrent que différents rythmes circadiens chez *Lingulodinium* sont probablement régulés par au moins deux voies différentes.

Malgré l'absence de régulateurs moléculaires identifiés jusqu'à maintenant, quelques points ressortent de la régulation des différents rythmes les plus étudiés. Cette régulation semble passer principalement par un contrôle traductionnel et post-traductionnel plutôt que par un contrôle transcriptionnel comme c'est le cas en général pour les eucaryotes [226]. En effet, les transcrits de chacun des gènes étudiés dans un contexte circadien chez *Lingulodinium* restent stable pendant la journée tandis qu'ils sont traduits à des temps bien précis et que les protéines peuvent être modifiées de façon post-traductionnelle ou bien voir leur localisation changer. De plus, la traduction *in vitro* des ARNm de *Lingulodinium* montre des patrons très similaires de protéines traduites entre des ARNm extraits tant en phase de jour qu'en phase de nuit, suggérant une certaine stabilité dans la quantité de transcrits entre différents temps [253]. Il est également intéressant de mentionner que chez d'autres dinoflagellés, les changements circadiens dans l'expression des gènes ont été investigués et sont minimes chez *Pyrocystis lunula* et chez *K. brevis*. Environ 3% des 3500 transcrits analysés par des expériences de micropuces à ADN varie en fonction du temps circadien chez *P. lunula* mais aucun ne voit son niveau augmenter ou diminuer de plus de 3 fois [254]. Des résultats similaires ont également été obtenus pour l'analyse de 4629 gènes par micropuce à ADN chez *K. brevis*, où environ 10% des transcrits varient en alternance jour/nuit alors que 3% des gènes continuent à osciller en conditions constantes [255]. Comme mentionné plus haut, les variations circadiennes dans l'abondance des transcrits chez d'autres organismes unicellulaires est beaucoup plus important que ces deux études, alors qu'environ le tiers des transcrits chez la diatomée *Phaeodactylum tricornutum* montrent des variations significatives sur un cycle jour/nuit [256] et que près de 80% des transcrits de *O. tauri* en font de même [133].

Plusieurs études se sont penchées sur le rôle de la traduction dans la génération et le maintien des rythmes circadiens chez *Lingulodinium*. Tout d'abord, il a été montré, par marquage des protéines nouvellement traduites à la leucine tritiée, que la synthèse de plusieurs protéines était contrôlée de façon circadienne, soit au début de la phase de nuit (LBP, Luciferase), au milieu de la phase de nuit (RubisCO) et à la transition entre la fin de la nuit et

le début de la journée (PCP, GAPDH) [91, 92, 253]. De façon intéressante, ces moments de synthèse de protéines corrèlent très bien avec les maximums des rythmes physiologiques dans lesquels les protéines sont impliquées chez *Lingulodinium*. L'inhibition de la synthèse protéique par différents inhibiteurs chimiques comme l'anisomycine et le cycloheximide a aussi d'importants effets sur l'horloge de *Lingulodinium*, permettant d'avancer ou de reculer la phase de l'horloge dépendant du temps d'administration des inhibiteurs [257-259]. Peu de détails sont par contre connus sur les mécanismes permettant de réguler de façon circadienne la synthèse protéique chez *Lingulodinium*. Une protéine pouvant lier le 3'-UTR de l'ARNm de LBP dans la journée, mais pas pendant la nuit a été identifiée [93] sans que ces résultats puissent faire l'objet d'une confirmation lors d'une étude plus approfondie [94].

Le rôle de la régulation post-traductionnelle dans les rythmes circadiens a aussi été étudié par le biais d'études d'inhibition chimique de kinase et de phosphatase. Par exemple, l'utilisation d'inhibiteur non spécifique des kinases comme la staurosporine et le 6-dimethylaminopurine permettent d'allonger de façon concentration-dépendante la période de l'horloge et de la stopper à forte concentration [111, 112]. Ces inhibiteurs permettent aussi de bloquer la synchronisation de l'horloge à des changements de phase lumineuse. De plus, l'inhibition spécifique des protéines phosphatase 1 et 2A permettent également d'allonger la période de l'horloge, mais pas de bloquer la synchronisation avec son environnement [113, 114]. Le rôle de la phosphorylation a été plus récemment étudié par des études de phosphoprotéomique. La comparaison par électrophorèse bidimensionnelle d'échantillons protéiques colorés au *ProQ diamond*, un colorant des phosphoprotéines, provenant du milieu du jour (LD6) et du milieu de la nuit (LD18) a montré que la phosphorylation de 47 protéines était plus importante à LD6 et 34 protéines l'étaient plus à LD18 [115]. Deux protéines phosphorylées à LD18 ont pu être identifiées par spectrométrie de masse comme étant un complexe collecteur de lumière et RAD24, une protéine impliquée dans le contrôle de la réplication de l'ADN. Comme cette étape du cycle cellulaire s'effectue vers le milieu de la nuit chez *Lingulodinium* [260], la phosphorylation de RAD24 peut potentiellement être relié au contrôle de la réplication. Une étude à plus large échelle par le biais de la purification de phosphoprotéines suivie d'un séquençage par spectrométrie de masse a permis d'identifier des protéines contenant des domaines de liaison à l'ARN dont plusieurs pourraient être des cibles

de la *Casein kinase 2* [95]. Comme cette kinase est impliquée dans l'horloge de plusieurs organismes [261], il est possible qu'elle joue un rôle dans la régulation circadienne de la traduction chez *Lingulodinium*. Finalement, des changements circadiens dans la structure de la cellule ont également été observés. Par exemple, les chloroplastes prennent une forme allongée pendant le jour, avec les thylakoïdes, où la lumière est captée, se retrouvant près de la surface de la cellule et le pyrénoïde, où se fait la fixation du carbone, se retrouvant plus près du centre. La nuit, les chloroplastes sont moins allongés et se retrouvent au centre de la cellule tandis que les scintillons se retrouvent à la surface [23].

#### 1.3.2.2. Intrants

L'observation que *Lingulodinium* est capable de réagir différemment lorsque soumis à des couleurs de lumière différentes indiquent qu'il pourrait avoir au moins deux récepteurs de lumière différents, sans toutefois qu'un de ceux-ci n'ait été identifié et caractérisé. Deux potentiels récepteurs lumineux ont par contre été identifiés récemment chez d'autres dinoflagellés. Un transcrit codant pour un cryptochrome a été identifié chez *Symbiodinium* (souche SSB01) et montre une diminution de son abondance lors d'une exposition à de fortes intensités lumineuses [89]. Également, certains dinoflagellés possèdent une protéorhodopsine, une pompe à proton activé par la lumière et qui a probablement été acquise par un transfert latéral de gènes. Ces gènes ont été partiellement caractérisés chez deux dinoflagellés, soit *Prorocentrum donghaiense* [262] et *Oxyrrhis marina* [263, 264]. De façon intéressante, les transcrits des deux espèces montrent une plus grande abondance pendant le jour que pendant la nuit, rythme qui est aboli lorsque les cultures sont placées en conditions constantes. Le rôle de ces récepteurs de lumière dans la transduction de signaux lumineux vers l'horloge circadienne n'a par contre pas encore été étudié.

Il est également possible qu'un ou des oscillateurs soient capables d'intégrer des signaux métaboliques, un peu comme le fait l'oscillateur de *Synechoccocus*. Le pH du milieu de culture de *Lingulodinium* va s'alcaliniser durant la journée dû à la consommation de $CO_2$ par la photosynthèse et va s'acidifier pendant la nuit alors que la respiration remet du $CO_2$ en solution. L'ajout d'acide dans le milieu de culture lorsque les cultures sont en conditions

constantes permet de changer la phase du rythme de bioluminescence, particulièrement lorsqu'il se fait au temps subjectif où le pH est au maximum [265]. Par ailleurs, l'ajout de nitrate à des cultures en phase subjective de jour permet de devancer la phase de l'horloge tandis que son ajout pendant la nuit subjective retarde l'horloge [265, 266].

#### 1.3.2.3. Sorties

**Bioluminescence**

Le phénomène de bioluminescence se retrouve chez quelques espèces de dinoflagellés et servirait de système de défense afin de faire fuir les copépodes qui s'en nourrissent. Ces prédateurs adoptent une nage plus rapide et moins propice au broutage des algues lorsque confrontés à des éclairs de bioluminescence [267]. Ceux-ci permettent également d'attirer et de faciliter la tâche des prédateurs des copépodes, agissant ainsi comme un système d'alarme [268, 269].

Le système de bioluminescence chez *Lingulodinium* est assez simple et ne comporte que deux protéines, la *Luciferin binding protein* (LBP) et la *Luciférase* (LCF) ainsi que la luciférine, un tétrapyrrole dérivé de la chlorophylle [270]. Ces composants se retrouvent dans de petites organelles associées avec la membrane vacuolaire nommée scintillons [271, 272]. LBP est responsable de lier la luciférine pour la protéger de la dégradation lorsque le pH est basique et la luciférase est l'enzyme qui dégrade par oxydation la luciférine à pH acide et permet ainsi l'émission de lumière [273]. Cette enzyme est composée d'une répétition de trois domaines fonctionnels chez plusieurs espèces de dinoflagellés [274, 275]. Une organisation particulière de ces deux gènes a par ailleurs été observée chez un dinoflagellé non photosynthétique, *Noctiluca scintillans*, qui possède un transcrit qui code à la fois pour LCF en 5' et LBP en 3' et un transcrit codant seulement pour LBP [276, 277].

La bioluminescence chez *Lingulodinium* prend deux formes bien distinctes : une lueur faible émise lors de la dégradation des scintillons à la fin de la nuit et des éclairs qui sont une importante émission de lumière visible à l'œil nu et dont le maximum se produit vers le milieu de la nuit [251]. Les éclairs peuvent avoir lieu de façon spontanée ou bien lorsqu'il y a un mouvement important dans le milieu de culture de l'organisme, provoquant un changement de

pH dans les scintillons pour favoriser la dégradation de la luciférine. Un canal à proton voltage-dépendant a d'ailleurs été identifié chez le dinoflagellé *Karlodinium veneficum* et a été suggéré comme pouvant potentiellement être apte à réguler le pH des scintillons, sans toutefois qu'une localisation à la membrane de ceux-ci n'ait été testée [278].

La régulation circadienne de la bioluminescence passe surtout par une traduction et une accumulation rythmique de LBP et LCF. La traduction de ces deux protéines se fait intensément pendant quelques heures au début de la phase de nuit [92, 279] et corrèle avec le moment de la formation des scintillons [280]. Par contre, le niveau des transcrits de ces deux gènes reste stable tout au long du cycle, ce qui implique que le contrôle de ce rythme passe par le contrôle du moment de la traduction de ces protéines [279, 281]. Les scintillons sont environ 10 fois plus nombreux durant la nuit, suivant ainsi la quantité de LBP et luciférase [280].

Une protéine capable d'interagir avec une courte séquence riche en GU située dans le 3'UTR de LBP a été identifiée par une expérience de retard sur gel, sans toutefois que cette interaction ait pu être confirmée [93, 94]. De façon intéressante, Chlamy1, une protéine capable de lier une séquence similaire, a été identifiée chez l'algue verte *Chlamydomonas reinhardtii* [282] et semble être au cœur de la régulation circadienne de la traduction chez cet organisme [283, 284]. Il est également possible que le transcrit de LBP change d'endroit dans la cellule au cours de la journée, permettant ainsi sa synthèse au moment propice [285, 286].

**Photosynthèse**

La photosynthèse chez *Lingulodinium* est également contrôlée de façon circadienne et est rythmique quant à la fixation du carbone et l'évolution de l'$O_2$. Ces deux phénomènes mesurés de façon indépendante en conditions constantes ont de façon surprenante des phases différentes. L'évolution d'$O_2$, générée par la photolyse de l'eau par le photosystème II et nécessitant absolument de la lumière, commence au tout début de la phase de jour subjectif et atteint un maximum au milieu de la journée. Comme la quantité de PCP est assez stable pendant tout le cycle, il a été suggéré que l'organisation du photosystème II serait responsable du rythme de l'évolution d'$O_2$ [287]. Quant à elle, la fixation du $CO_2$ par la RubisCO débute, en condition constante, vers le milieu de la nuit subjective pour atteindre un maximum au

début de la journée. L'abondance de la RubisCO, tant au niveau de l'ARNm qu'au niveau protéique, est également similaire tout au long d'un cycle circadien [23, 288]. Par contre, une variation circadienne dans l'enrichissement en RubisCO dans une région du chloroplaste appelée le pyrénoïde est en grande partie responsable de la capacité de *L. polyedrum* à fixer efficacement le carbone [23]. Un blocage de l'import protéique au chloroplaste empêche d'ailleurs l'enrichissement de la RubisCO et la formation du pyrénoïde, suggérant que la synthèse d'une protéine particulière pourrait réguler cette étape [289, 290]. De plus, contrairement à la RubisCO de type I, qui doit être activée par la RubisCO activase avant d'être fonctionnel, la RubisCO de type II de *Lingulodinium* semble être activée en tout temps, ne montrant pas de variation dans son activité *in vitro*, qu'elle provienne d'extraits de cellule provenant de la phase de jour comme de la phase de nuit [291].

Les chloroplastes de *Lingulodinium* prennent également des formes différentes entre le jour et la nuit. Le jour, les chloroplastes sont allongés et plus près de la surface de la cellule. Les complexes collecteurs de lumières situés dans le lumen du thylakoïde, se retrouve surtout en périphérie de la surface de la cellule pour permettre de maximiser la quantité de lumière captée [290]. Quant à elle, la RubisCO, se trouvant dans le stroma, se retrouve plus au centre de la cellule, au pyrénoïde [23]. Ceci permet de séparer les réactions photochimiques, qui génèrent beaucoup d'$O_2$, de la réaction de fixation du $CO_2$ par la RubisCO qui est limitée par la compétition avec l'$O_2$. Cette ségrégation spatiale de la RubisCO est abolie vers la fin de la journée, où ces deux protéines montrent une séparation beaucoup moins marquée, ce qui corrèle avec un changement dans la forme des chloroplastes qui deviennent moins allongés et occupent une position plus centrale dans la cellule [290].

### Cycle cellulaire

Le cycle cellulaire chez *Lingulodinium* montre une synchronisation de la phase S, qui se déroule vers le milieu de la nuit et de la phase M, qui se déroule vers la transition nuit/jour [292]. Ce contrôle temporel dans le moment de réplication de l'ADN est aussi observé chez *Chlamydomonas* [293] et *O. tauri* [174]. La réplication du matériel génétique se déroulant pendant la nuit, il est suggéré que ce moment permet de limiter les dommages à l'ADN liés à la présence de rayons ultraviolets. De façon intéressante, la régulation de la phase S et de la phase M semble être sous contrôle de deux oscillateurs différents [260]. Comme mentionné

plus haut, quelques régulateurs potentiels du cycle cellulaire ont été caractérisés chez *Lingulodinium* sans toutefois qu'un lien potentiel avec le contrôle circadien du cycle cellulaire n'ait été étudié plus en détail.

### Migration verticale et acquisition du nitrate

*Lingulodinium* montre une migration verticale d'environ dix mètres dans la colonne d'eau entre le jour et la nuit [294, 295]. Il a été suggéré que cette migration verticale permet, pendant la journée, d'être près de la surface de l'océan afin de maximiser la quantité et la qualité de lumière captée par les systèmes photosynthétiques tandis que la descente dans la colonne d'eau vers la fin de la journée favoriserait la prise de nitrate, qui s'y retrouve en quantité plus abondante [296]. Par contre, l'acquisition et l'assimilation du nitrate montrent des variations journalières chez *Lingulodinium*, qui ne concordent pas avec le rôle présumé de la migration verticale. En effet, l'assimilation du nitrate, mesuré par le biais de l'activité et la quantité de l'enzyme nitrate réductase a été montré comme étant circadienne, avec un maximum se trouvant au milieu de la journée [297]. L'acquisition du nitrate semble quant à elle montrer un rythme journalier, avec un maximum également retrouvé au milieu de la journée, mais ce rythme n'est pas observé lors d'un cycle en lumière constante, contrairement au rythme de la nitrate réductase [298].

### Autres rythmes

Quelques autres rythmes ont été sommairement étudiés chez *Lingulodinium* comme l'activité de la superoxyde dismutase [235, 299] et l'activité d'enzymes du cycle de l'acide citrique. Ce dernier voit deux enzymes, l'isocitrate déshydrogénase NADP-dépendante et la succinate déshydrogénase, avoir une grande variation dans la quantité d'enzyme au cours du cycle. La première a un pic de quantité de protéine et d'activité au milieu de la nuit, sans toutefois que le niveau de transcrits change de façon significative. La deuxième a un maximum de protéine en deuxième moitié de journée et en début de nuit, tout en ayant également un niveau de transcrit égale tout au long du cycle [300].

## 1.4. Présentation générale du projet

La biologie moléculaire chez les dinoflagellés n'est que peu caractérisée par rapport aux autres eucaryotes. L'absence de mode de transformation et de criblage des mutants de ces organismes, couplés à des génomes de grandes tailles, fortement complexes et présentant des caractéristiques inusuelles a limité grandement leur étude moléculaire. Comme mentionné dans la revue de littérature, les modes de régulation génique, qu'ils soient transcriptionnels ou post-transcriptionnels sont encore très mal compris et ont surtout été étudiés en lien avec la régulation circadienne de la biochimie et de la physiologie de ces organismes.

L'objectif principal de ma thèse est de mieux caractériser la régulation transcriptionnelle chez les dinoflagellés en tirant profit des nouvelles techniques de séquençage à très haut-débit et des avancées en spectrométrie de masse. Pour ce faire, trois approches différentes seront utilisées. La première consiste en le séquençage à très haut débit du transcriptome de *Lingulodinium* provenant de quatre temps différents dans un cycle journalier. L'objectif de ce projet est double en ce sens qu'il pourrait permettre d'identifier des gènes faisant partie de l'horloge circadienne tout en permettant d'obtenir le profil d'expression journalier de plusieurs dizaines de millier de gènes. Comme les gènes codants pour les oscillateurs chez les eucaryotes montrent en général des variations circadiennes robustes dans leur abondance et permettent par le fait même de générer et contrôler les rythmes de l'organisme, j'ai fait l'hypothèse que ces gènes pourraient être identifiés par le séquençage massif des ARNm à différents temps. De plus, jusqu'à maintenant, l'étude circadienne d'un très petit nombre de gènes a été effectuée chez *Lingulodinium* et n'ont montré que de très faibles variations dans l'abondance des transcrits [226]. La progression des techniques de séquençage à très haut débit, comme le séquençage par synthèse de type Illumina, permet maintenant de séquencer l'entièreté d'un transcriptome en plus d'obtenir une estimation de l'abondance relative des gènes entre eux dans un même échantillon et entre plusieurs échantillons. Elles permettent également de pallier au manque de disponibilité commerciale des puces à ADN pour des organismes peu étudiés. Cette expérience permettra donc d'augmenter grandement le nombre de gènes pour lesquelles des informations sur l'expression

génique journalière seront disponibles et pourraient également permettre d'identifier certains transcrits rythmiques.

La deuxième approche découle du récent séquençage de transcriptomes complets de dinoflagellés qui ont permis d'obtenir des catalogues de gènes d'une grande ampleur et ont pu pallier à l'absence de génome séquencés [61, 69]. Parmi les informations tirées de ces projets, l'identification de potentiels facteurs de transcription chez les dinoflagellés a révélé une situation inusuelle parmi les eucaryotes, à savoir un très faible nombre de ce type de protéines couplé à une grande expansion de la famille des protéines à domaine *cold shock* (CSD) qui peuvent potentiellement servir à réguler la transcription. Comme ces domaines se retrouvent dans des protéines ayant une grande diversité de fonctions et qu'ils peuvent lier l'ARN et l'ADN, leur rôle dans la régulation de la transcription chez les dinoflagellés n'est pas très clair. Afin de mieux comprendre leurs rôles potentiels, deux de ces protéines seront caractérisées pour évaluer différents aspects de leurs fonctionnalités comme leur spécificité de liaison aux acides nucléiques par le biais d'essais de retard sur gel et leur implication dans la réponse à des chocs thermiques par le biais d'essais de complémentation de mutants chez *E. coli* et par la mesure de leur induction par le froid.

La dernière approche vise à identifier les protéines qui sont associées à la chromatine des dinoflagellés. Celle-ci est compacté sous forme de cristal liquide stabilisée par de l'ARN et des cations métalliques plutôt que sous forme de nucléosome organisé par des histones et très peu de protéines y ont été localisées. Afin d'identifier les protéines qui sont associés aux chromosomes, une réticulation *in vitro* suivi d'une purification de la chromatine et un séquençage des protéines qui y sont associées par spectrométrie de masse seront effectués. L'identité et la fonction des protéines qui y sont associés pourraient permettre de trouver de potentiels régulateurs de la transcription et du contrôle du cycle cellulaire. Cette approche pourrait aussi permettre d'identifier des protéines impliquées dans divers processus reliés aux ARNm comme la régulation de leur stabilité et leur traduction.

# Chapitre 2: Publication # 1

# The Lingulodinium circadian system lacks rhythmic changes in transcript abundance

Sougata Roy*, Mathieu Beauchemin*, Steve Dagenais-Bellefeuille*, Louis Létourneau, Mario Cappadocia et David Morse

*Contribution égale des auteurs

Publié dans BMC Biology (2014) 12:107 **DOI:** 10.1186/s12915-014-0107-z

Contribution des auteurs :

Pour cet article, j'ai participé à la conception et à la réalisation d'une partie des expériences de biologie moléculaire tout en fournissant une large part des analyses bio-informatiques. J'ai également participé à la rédaction de l'article. S. Roy et S. Dagenais-Bellefeuille ont contribué de façon égale à cet article en effectuant plusieurs des expériences de biologie moléculaire et en participant à la rédaction. L. Létourneau a fourni quelques analyses bio-informatiques supplémentaires et M. Cappadocia a participé à la rédaction de l'article. D. Morse a supervisé le projet tout en effectuant des analyses bio-informatique et en participant à la rédaction et la révision de l'article.

## 2.1. Abstract

**Background**

Almost all cells display circadian rhythms, ~24 hour period changes in their biochemistry, physiology or behavior. These rhythms are orchestrated by an endogenous circadian clock whose mechanism is based on transcription-translation feedback loops (TTFL) where the translated products of clock genes act to inhibit their own transcription.

**Methods**

We have used RNA-Seq to measure the abundance of all transcripts in a RNA-Seq-derived *de novo* gene catalog in two different experiments. One compared midday and midnight in a light-dark cycle (ZT6 and ZT18) with similar times under constant light (CT6 and CT18). The second compared four times (ZT2, ZT6, ZT14 and ZT18) under a light-dark cycle. The timing for the bioluminescence and photosynthesis rhythms were also determined in the presence of the transcription inhibitors actinomycin D and cordycepin.

**Results**

We show here that despite an elaborate repertoire of biological rhythms, the unicellular dinoflagellate *Lingulodinium* has no detectable daily variation in the abundance of any transcript in an RNA-Seq-derived *de novo* gene catalog. Furthermore, notwithstanding a marked decrease in rhythm amplitude, the timing of the two rhythms is unchanged in the presence of transcription inhibitors.

**Conclusions**

The lack of detectable daily variation in transcript levels indicates that the endogenous circadian timer of *Lingulodinium* does not require rhythmic RNA. If the circadian timer is considered as a limit cycle oscillator, then cellular time in this organism must be defined by variations in state variables that do not include the amount of a clock gene transcript.

## 2.2. Background

Circadian rhythms are changes in a cell's biochemistry, physiology or behavior that occur with a roughly 24-hour period. The rhythms are the observable outputs of a cell-autonomous and resettable timekeeper [122] called a circadian clock. This endogenous circadian clock is ubiquitous and organized as a transcriptional-translational feedback loop (TTFL) in eukaryotes, and although the genes that constitute the clock differ between plant, animal, and fungal model systems [301], transcriptional control is thought to be an integral component of the clock mechanism. Even in cyanobacteria, where a daily rhythm in the phosphorylation state of the clock protein KaiC has been proposed to be the pacemaker [176], a KaiC TTFL may be required for generating robust rhythms [184].

Recent observations, however, have indicated that transcription is not always required to produce circadian oscillations, as evidenced by the daily oscillation of redox cycles of peroxiredoxin in red blood cells [232] and the green algae *Ostreococcus tauri* [169]. Interestingly, when *O. tauri* is placed in prolonged darkness, transcription rates fall below detectable limits and the rhythm in luciferase fused with the clock component CCA dampens after one day. However, when rhythmicity is re-initiated by transfer to constant light, the phase of the rhythm varies depending on the time when cells are exposed to light [169]. Thus, either a timer driving the observed rhythm of this translational reporter continues even though the overt rhythm itself is undetectable, or the TTFL of *O. tauri* may be influenced by cross-talk with the non-transcriptional peroxiredoxin rhythm which has been shown to continue unabated in darkness.

The marine dinoflagellate *Lingulodinium* displays a large variety of overt rhythms, and has been a model for study of the mechanisms linking the clock with these rhythms for many years [296]. For example, the bioluminescence rhythm is correlated with rhythmic changes in the amount of the reaction catalyst (dinoflagellate luciferase) [302] and of a luciferin binding protein (LBP) [279] that protects the bioluminescence substrate luciferin from non-bioluminescent oxidation. In addition, the sequestration of the key carbon-fixing enzyme Rubisco within the pyrenoid of the chloroplast is correlated with the capacity of the cell to efficiently fix carbon [23]. Both these different rhythms correlate with rhythms in the rate of

protein synthesis *in vivo*, indicating that clock control over gene expression may regulate the timing of the rhythms. Importantly, in these and other examples, the control over protein synthesis occurs at a translational level since levels of the corresponding mRNAs do not change over the daily cycle [253, 279, 288, 303].

In contrast to the depth of knowledge concerning the rhythms, the mechanism of the circadian clock in *Lingulodinium* remains unknown. To compound the difficulty in characterizing the central timer, physiological studies have shown that these single celled organisms actually contain two different endogenous clocks, as the rhythms of bioluminescence and swimming behavior can run with different periods [250] and show different phase resetting behavior [252]. In the present study we sought to identify rhythmic transcripts in *Lingulodinium* in order to identify potential TTFL components. We used RNA-Seq to assess levels of all RNA species in a *Lingulodiunium* transcriptome [61] over both diurnal and circadian cycles. Surprisingly, our analyses indicate that *Lingulodinium* does not express any detectable rhythmic transcripts. This suggests that the mechanism of the endogenous timers in this organism will instead involve translational and post-translational mechanisms.

## 2.3. Results

To assess the possibility of isolating components of a transcription-based oscillator in a dinoflagellate, RNA-Seq was used to globally quantitate transcript levels at different times. Two different RNA-Seq experiments were performed, the first of which generated 252 million 76 bp paired end reads (using Zeitgeber times ZT 6, ZT 18, and circadian times CT 6 and CT 18) while the second generated 545 million 100 bp paired end reads (taken at times ZT 2, ZT 6, ZT 14 and ZT 18), of which 51% and 92%, respectively, mapped to our *de novo* gene assemblies. We first compared ZT 6 and ZT 18 by mapping the 100 bp reads to a 103,266 contig Trinity assembly [304] (Figure 2.1. A), expecting to find both light-induced and circadian differences between the two times. Instead, read counts from the two times, normalized as reads per kilobase per million reads (RPKM) [305] show surprisingly few differences in mRNA levels. DESeq [306] analysis indicated only five contigs showed significantly different levels between the two times ($p_{adj} < 0.001$). A similar result was obtained when the reads were mapped to a 74,655 contig Velvet assembly [61] for which DESeq identified 13 significant differences ($p_{adj} < 0.001$) (Figure 2.S1. A). Of the five significant differences uncovered using the Trinity assembly, three were also identified as significant using the Velvet assembly, while the remaining two do not have corresponding contigs in the Velvet assembly. The increased number of significant differences in the Velvet assembly is primarily due to slight differences in the read counts that move either the fold difference or the mean read counts above the significance threshold. Similar results were obtained using EdgeR [307] to assess significant differences, so DESeq was used for all further analyses.

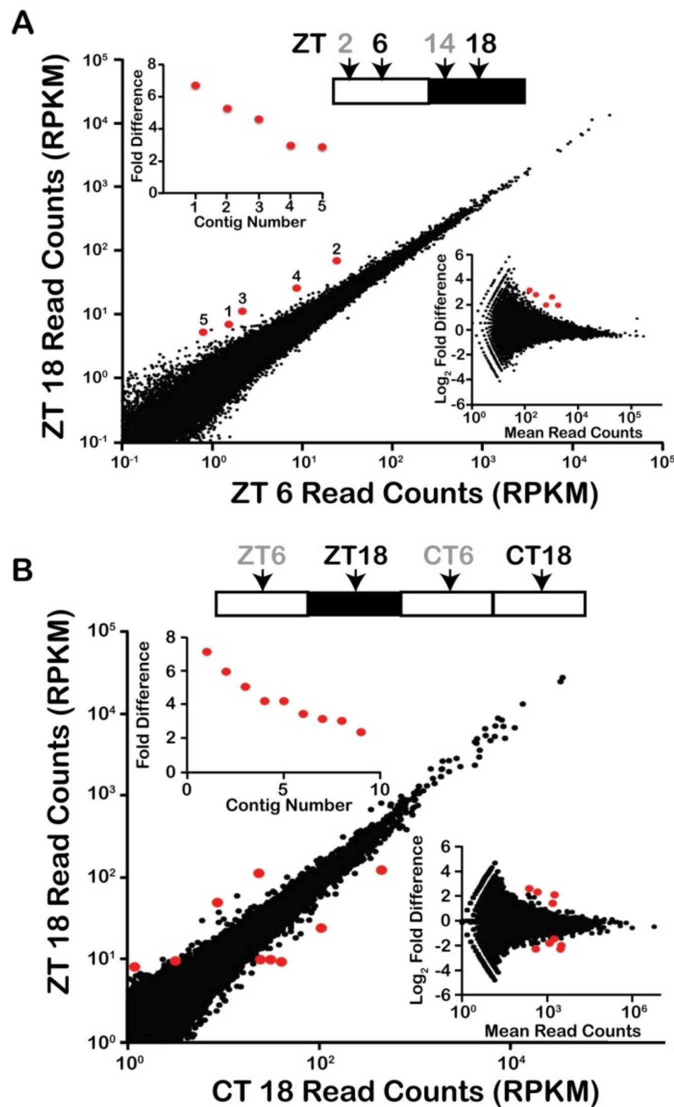



**Figure 2.1. Transcript abundance does not change between midday (ZT 6) and midnight (ZT 18).**

A. Read counts from the two samples (as reads per kilobase per million, RPKM), obtained by mapping raw read data to 103,266 contig Trinity assembly, were plotted against each other. Contigs corresponding to rRNA were removed. Insets show (upper left) the fold difference for DESeq statistically significant differences (padj $< 0.001$) and (lower right) the MA plot used to determine statistically valid differences. B. Transcript abundance does not change between true (ZT 18) and subjective (CT 18) midnight.

To confirm the similarity in transcript levels between midday and midnight, the midday/midnight comparison was repeated using the 76 bp paired-end read experiment. This experiment included samples taken at both true (ZT 18) and at subjective (CT 18) midnight, so we first assessed the similarity between these two. DESeq analysis identified 9 and 6 significant differences when read counts were mapped to the Trinity (Figure 2.1. B) and Velvet (Figure 2.S1. B) assemblies, respectively, with only one contig common to both. We conclude from this that constant light during the night phase does not cause a major alteration in gene expression, and thus the two midnight samples are essentially duplicates. We therefore compared the duplicate midday samples with these two midnight samples using DESeq (Figure 2.S2.). This analysis revealed no significant differences using the Trinity assembly and only three significant differences with the Velvet assembly. Since none of the differences found in this second experiment were in common with those found in the first experiment, we conclude that the mRNA complement of the cells at midday is the same as at midnight.

To test the possibility that cells might express rhythmic RNA with maxima lying between midday and midnight, we next compared the two additional times (ZT 2 and ZT 14) in the 100 bp read experiment. These times, immediately after the light/dark transitions, were chosen to maximize the chances of finding differentially accumulated RNAs. We determined significant differences from all possible pairwise comparisons of read counts mapped to the Trinity assembly using DESeq ($p_{adj}<0.001$) and combined all to yield a final list of 131 non-rRNA sequences (Figure 2.2. A). Those showing the greatest fold-differences were typically lower in abundance (Figure 2.2. B). Hierarchical clustering identified four main groups of sequences, with those peaking at ZT 14 containing the greatest number of significantly different RNAs (Figure 2.2. C). Surprisingly, however, among the 42 sequences that could be identified by sequence homology were found 2 of the 3 known mitochondrial protein-coding genes [308] and 9 of the 10 known plastid protein-coding genes [29], and all these organelle-encoded transcripts were among the group with peak expression at ZT 14. The plastid-encoded sequences included the *psbA* transcript that had been previously shown by Northern analyses to be arrhythmic [309], suggesting that at least some of the transcripts identified as significantly different in our high throughput approach might in fact be false positives. To assess this possibility, Northern analyses were performed using a random selection of 11

sequences (Figure 2.2. D). No rhythmicity was observed in the levels of any of these RNAs, as confirmed by densitometric scans (Figure 2.2. E). Taken together, the two RNA-Seq experiments thus indicate that the entire mRNA complement of the cell is maintained at constant levels over the 24-hour cycle, both under light/dark cycles or in constant light.

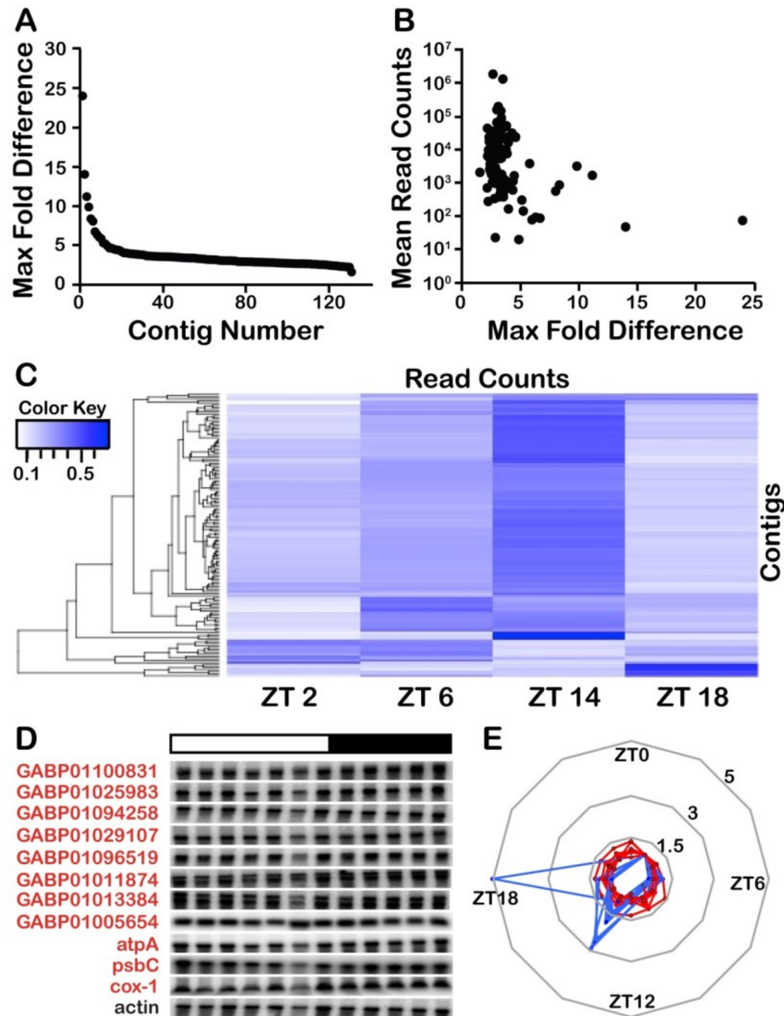



**Figure 2.2.** ***Lingulodinium*** **does not have rhythmic transcripts over a 24 hour cycle.**

A. DESeq significant differences for all pairs of samples taken LD 2, 6, 14 and 18 under a light dark cycle were combined to a single list of 131 contigs. B. Contigs with more than five-fold maximum difference are typically those with lower mean read counts. C. A heat map of all the significant differences detected shows a preponderance of contigs whose abundance appears greatest at ZT 14. Hierarchical clustering was performed only with rows to preserve the order of the time points. D. Northern blots using probes prepared from a random selection of contigs indicate the RNA-Seq predicted differences are false positives. E. Densitometric scans of Northerns relative to actin as a loading control (red) compared to normalized RNA-Seq data (blue).

The absence of rhythmic RNAs suggested that the circadian clock would continue unabated in the presence of transcription inhibitors. To test this, we exposed cell cultures to a combination of the transcription inhibitors actinomycin D and cordycepin (3-doxyadenosine). Actinomycin D binds to DNA and interferes with transcription and replication [310] while cordycepin interferes with polyadenylation of mRNAs [311]. This combined treatment reduced RNA synthesis to levels roughly 5% of vehicle-treated cultures (Figure 2.3. A) and is lethal after three days. Importantly, within this three-day time frame, the presence of the inhibitors does not change the timing of the bioluminescence rhythm in constant darkness (Figure 2.3. B) although the amplitude of the rhythm shows a marked decrease. Furthermore, the timing of the photosynthesis rhythm in constant light (Figure 2.3. C) appears similar in the presence of the drugs, although again we note a marked decrease in the rhythmic amplitude, especially during the third day (the second day under constant conditions). We conclude from these experiments that, despite a strong effect on the amplitude of the two rhythms, robust transcription is not required for determining their circadian timing.

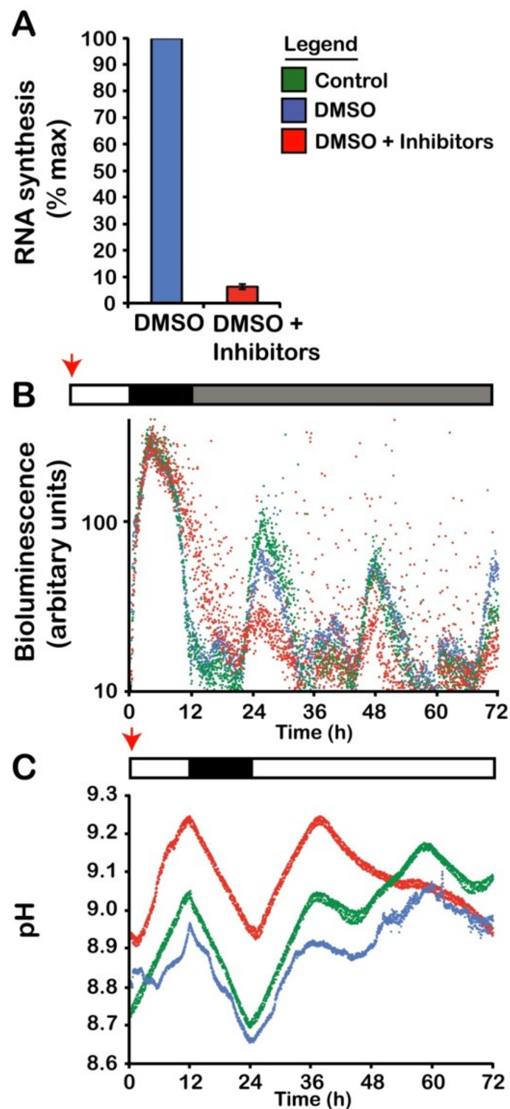



**Figure 2.3. Rhythms continue in the presence of transcription inhibitors.**

A. A combination of actinomycin D and cordycepin reduces RNA synthesis rates to around 5% (mean ± SD, n=4) of the values of vehicle treated cultures. Cultures survive 3-5 days at these concentrations of inhibitors. B. There is no change in the timing of the bioluminescence rhythm in constant dark. Horizontal bars indicate the light regime, with drugs were added to the cultures at the indicated times (red arrows). Each data point is the average of 3 samples. C. There is no change in the timing of the pH rhythm used to monitor CO2 fixation. A representative example of three different experiments is shown for each treatment. The amplitude of the rhythm is severely diminished by the third day of treatment.

## 2.3. Discussion

The results described here indicate that none of the sequences in our *Lingulodinium* transcriptomes show circadian or diurnal variations in transcript levels (Figure 2.1. and 2.2.). These analyses rely on assemblies, derived from high throughput Illumina RNA-Seq experiments that have been constructed using either Velvet or Trinity. The Velvet assembly has been most extensively characterized, with the 74,655 contigs estimated to represent ~94% of the total transcripts expressed by the cells [61]. The 103,266 contig Trinity transcriptome contains all the Velvet contigs as well as roughly 28,000 additional contigs [304]. Thus, it seems unlikely that *Lingulodinium* will express a rhythmic transcript in an RNA species that is not represented in either of the two assemblies. This conclusion is also supported by the finding that reducing the transcription rate to 5% of control levels does not affect the period of either the bioluminescence or the photosynthesis rhythms (Figure 2.3.).

The manifestation of eukaryotic circadian rhythms in the absence of transcription was first suggested by observations of the photosynthesis rhythm in enucleated *Acetabularia* [228]. More recently, the finding of a circadian rhythm in peroxiredoxin redox state in red blood cells [232], the demonstration of a FRQ-less oscillator in *Neurospora* [312] and the ability of transcription-incompetent *O. tauri* [169] to keep time in constant darkness have also provided support for this view. However, since these latter two organisms also have a canonical TTFL, the significance of a non-transcriptional oscillator operating in parallel remains unclear. In contrast, our findings indicate that the entire *Lingulodinium* circadian system has evolved to function without a TTFL, a particularly remarkable finding considering that there are at least two independent circadian timers in these cells [250]. It is unknown at this time if *Lingulodinium* will also show a circadian rhythm in peroxiredoxin redox state.

One possible explanation for the absence of transcriptional control in the *Lingulodinium* clock is the potential difficulty in regulating transcription in the dinoflagellates. Indeed, these organisms do not contain detectable levels of histone proteins [50], their chromosomes are permanently condensed [313], and the number and diversity of transcription factors is also much reduced compared to other eukaryotes [61, 69]. In addition, a global analysis of mRNAs in the dinoflagellate *Karenia* has shown their half-lives to be substantially

longer than reported in other organisms [90], as might be expected if little control was exerted at the level of RNA synthesis rates. *Lingulodinium* also appears to have quite stable RNAs, at least for the few specific cases that have been examined [314].

How might timing signals be generated in *Lingulodinium* in the absence of rhythmic transcripts? Previous work with the system has shown that inhibitors of translation (such as anisomycin, puromycin and cycloheximide) have major phase-shifting effects on the bioluminescence rhythm [257, 259, 315]. In particular, a specific dose of the inhibitor anisomycin given at a specific time, but not lower or higher doses, can induce arrhythmicity [257]. This has been interpreted by a limit cycle model to mean that the clock has been driven to a singularity and provides strong presumptive evidence that translation is a state variable in the clock mechanism. In addition to inhibitors of translation, the kinase inhibitors staurosporine and 6-dimethyl amino purine (6-DMAP) affect the period of the bioluminescence rhythm when administered chronically, and strongly affect light-induced phase shifts [112]. Similarly, a range of protein phosphatase inhibitors (okadaic acid, calyculin A and cantharidin) also affect the rhythmic period when given chronically, although they appear unable to block light-induced phase shifts [114]. We suggest that, in the absence of rhythmic transcripts, *Lingulodinium* may have evolved translational/posttranslational feedback loops with 24-hour rhythmicity that can act as the central timing mechanism. One intriguing possibility is that protein phosphorylation may be used to control translation, as *Lingulodinium* has casein kinase 2 (CK2) sites in a large number of RNA binding proteins [95]. CK2 is a kinase that has been implicated in animal, plant and fungal clock mechanisms [316].

There is growing awareness that RNA-binding proteins may play an important role in clock function. For example, the *Chlamydomonas* protein CHLAMY1 is involved in controlling both the period and phase of the circadian rhythms [284]. In addition, a cold-induced RNA-binding protein appears to be required for efficient translation of the clock gene *Clock* transcripts in mammalian cells [317]. Lastly, ATX (Ataxin-2) in *Drosophila* is an RNA-associated protein that, together with TYF (Twenty-four), is required for translating the transcripts from the clock gene *Period* (*Per*) [318, 319]. TYF binds both to PABP, a factor binding the polyadenylated tail at the 3’ end of mRNA, and to the eukaryotic translation initiation factor eIF4E, which binds the cap at the 5’ end of the mRNA [319]. This

circularization is thought to increase translation [320] which in this case would be targeted to the *Per* transcript. In mouse liver cells, changes in the length of the poly(A) tail occur independently from the steady state level of the transcripts, and tail length is associated with circadian control over protein synthesis [321].

The lack of rhythmic transcripts over the daily cycle in *Lingulodinium* provides new impetus for an examination of translation and post-translational mechanisms in the clock. The former may be addressed by methods such as ribosome profiling, which has been used to demonstrate rhythmic translation from constant levels of mRNA in drosophila [322]. The latter, at least for modifications involving phosphorylation, is now becoming accessible through phosphoprotein purification and MS/MS de novo sequencing [95]. In any event, it is clear that the long-studied *Lingulodinium* circadian system has still some surprises in reserve.

## 2.5. Conclusion

The lack of oscillating RNAs over the circadian cycle of the dinoflagellate *Lingulodinium* indicates that the clock mechanism in this organism does not use or require rhythmic changes in RNA amounts.

## 2.6. Methods

### 2.6.1. Cell Culture

Unialgal but not axenic *Lingulodinium polyedrum* (CCMP 1936, previously *Gonyaulax polyedra*) was obtained from the Provasoli-Guillard National Center for Marine Algae and Microbiota (East Boothbay, ME, USA). Cell cultures were grown in normal f/2 medium prepared using Instant Ocean under a 12 h light (40 μmol photons $m^{-2}$ $s^{-1}$ cool white fluorescent light) and 12 h darkness at a temperature of 18 ± 1 °C. This light/dark regime is termed LD, with LD 0 corresponding to lights on and LD 12 to lights off. Cells were harvested by filtration on Whatman 541 paper and washed with 200 mL of sterile seawater to reduce bacterial contamination.

### 2.6.2. RNA extraction and RNA-Seq

RNA-Seq was performed for two different experiments. In one experiment, *Lingulodinium polyedrum* cells grown in f/2 seawater medium at 18°C under a 12:12 light:dark cycle [279] were isolated by centrifugation (500 g for 1 min) at two times (Zeitgeber Time ZT 6, ZT 18) and two times under constant light (Circadian Time CT 6, CT 18). The onset of the light phase in a light dark cycle or under continuous light is termed ZT 0 or CT 0, respectively. In another experiment, cells were taken at ZT 2, ZT 6, ZT 14 and ZT 18. All cell pellets were washed twice with fresh seawater and recentrifuged, and the cell pellets were resuspended in 1 mL of Trizol reagent (Life Technologies) and total RNA was extracted as per the manufacturers' protocol. For the first experiment, RNA quality was assessed by Northern blots before and after purification of poly(A) RNA using the poly(A) tract mRNA isolation kit (Promega) and sequenced using a Genome Analyzer IIX (Illumina) at the McGill University and Génome Québec Innovation Center. For the second experiment, RNA pellets were dissolved in 50 μL DEPC treated $H_2O$, and a Bioanalyzer (Agilent) test performed to assess the quality of the extracted RNA samples. An mRNA-Seq Sample Preparation Kit (Illumina) was used prior to sequencing of the mRNAs using a HiSeq platform

(Illumina) at the Institut de Recherche en Immunologie et Cancérologie (Université de Montréal).

Reads were mapped to both a Velvet assembly (accession nos. JO692619-JO767447) [61] and to a Trinity assembly (accession nos. GABP01000001-GABP01114492) [304] using BWA [323]. Sequencing produced 252 million 76 bp paired end reads (Illumina GAII, accession numbers SRR330443-6), of which 49.8% and 51.8% mapped to the Velvet and Trinity assemblies respectively, and 545 million 100 bp paired end reads (HiSeq, accession numbers SRR1184543, 1184608, 1184657, 1184666) of which 95.5 and 88.3% mapped to the Velvet and Trinity assemblies, respectively. Read counts were analyzed by DESeq [306] and EdgeR [307] to uncover statistically significant differences. All reads and assemblies are available as a part of the NCBI BioProject PRJNA69549.

Northern blots were performed using total RNA isolated by Trizol extraction as described by the manufacturer. From the 131 sequences with DESeq significant differences, 67 pairs of oligonucleotides were designed from the Trinity assembly sequences and tested by PCR, with 27 pairs producing sequence-validated amplicons. The majority of these were from the same group (maximal levels at ZT 14) so only 17 were tested by Northern blotting, and of these 9 (including a mitochondrial *cox1* probe) yielded a well-defined signal after exposure. An additional two probes were prepared from previously cloned plastid-encoded sequences *atpA* and *psbC* [29]. An actin probe was used to control for RNA loading.

### 2.6.3. Transcription Inhibition

All the inhibitors were dissolved in 100% DMSO to prepare the stock solution. To inhibit transcription, 20 µL from a stock of 10 mg/mL Cordycepin (final concentration 20 µg/mL) and 5 µL from a stock of 1 mM Actinomycin D (final concentration 0.5 µM) were added to a 10 mL culture of *Lingulodinium* cells. To monitor the inhibition of RNA synthesis, 5 µCi of $^{32}P$ radionuclide was added after 8 hours incubation with inhibitors. Total RNA was extracted using Trizol reagent (Life Technologies) after 8 hours of incubation with the label. RNA pellets were dissolved in 50 µL of DEPC treated water and 2 µL aliquots mixed with 2

mL of scintillation fluid and counted using a scintillation counter (Perkin Elmer TriCarb 2800TR). The control was prepared by adding only the vehicle (DMSO) to the culture.

### 2.6.4. Rhythm Measurements

To monitor the bioluminescence rhythm after inhibition of transcription, cordycepin and actinomycin D was added to the same concentrations as above. After 4 hours of incubation of the cultures with inhibitors, six 280 μL aliquots of cell culture were loaded in to a 96-well microtiter plate. Control cells in seawater or with the vehicle alone (DMSO) were also included in the plate. The bioluminescence rhythm was monitored in constant darkness as described [304] using a Spectramax M5 (Molecular Devices) microplate reader kept in the culture room at constant temperature. Bioluminescence output was recorded for one second every 2 min for the next 72 hours. Wells containing cultures were surrounded by wells filled with only seawater to limit evaporation. The pH rhythm was measured as described [265, 324] using a pH electrode in a culture flask kept under constant light and temperature in the culture room. The treatments with DMSO alone or DMSO containing the inhibitors were repeated three times.

## 2.7. Acknowledgements

We thank the Canadian National Science and Engineering Research Council for financial assistance in the form of a discovery grant (D.M.) and a research studentship (M.B.), and to the Fonds de Recherche du Quebec- Nature et Technologies for a research studentship (S.D.B.). We thank N. Cermakian and N. Luckhurst for review and comments on the manuscript. We dedicate this manuscript to the memory of J. Woodland Hastings, a pioneer of circadian biology and bioluminescence.

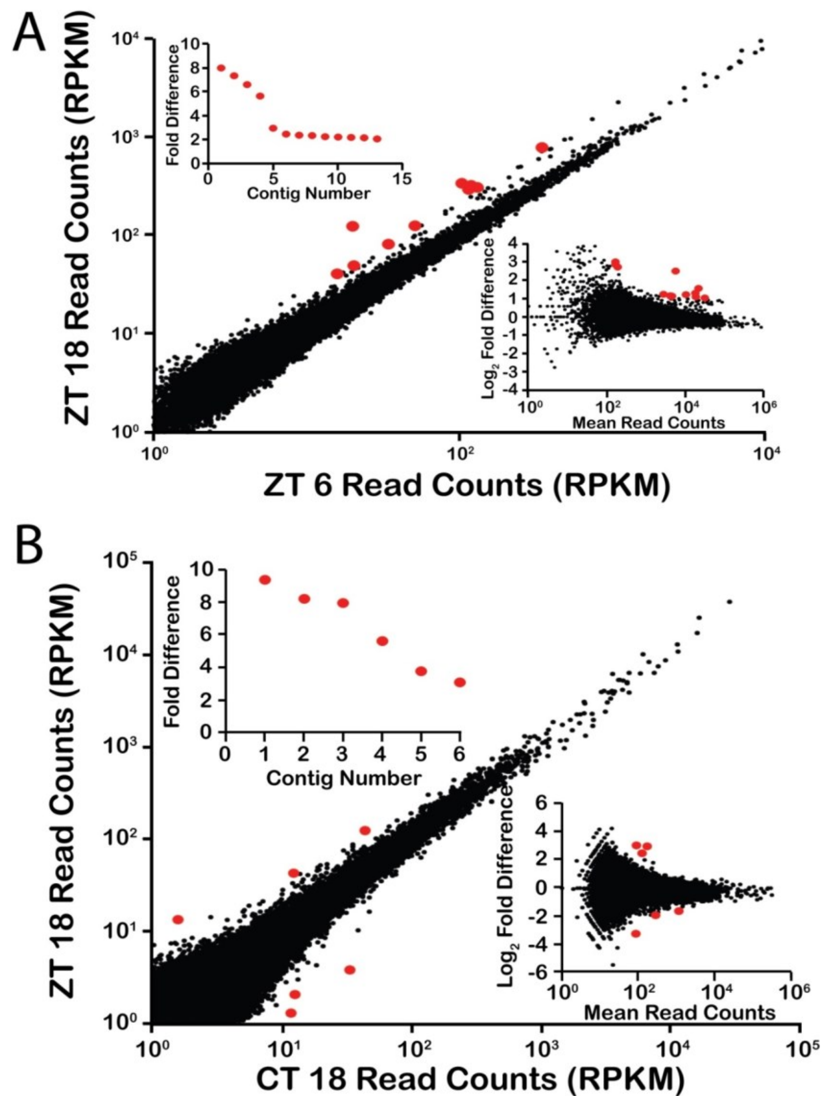



**Figure 2.S1. Analysis using the Velvet assembly**

A. Transcript abundance for ZT6 and ZT18 samples (as reads per kilobase per million, RPKM) obtained by mapping raw read data from the 100 bp read experiment to a 74,655 contig Velvet assembly. The fold difference for the contigs that showed a significant difference is shown as an inset (upper left) while the MA plot used to determine significant differences is at lower right. Contigs corresponding to rRNA have been removed. B. Transcript abundance for ZT 18 and CT 18 analyzed as above for the 76 bp read experiment.

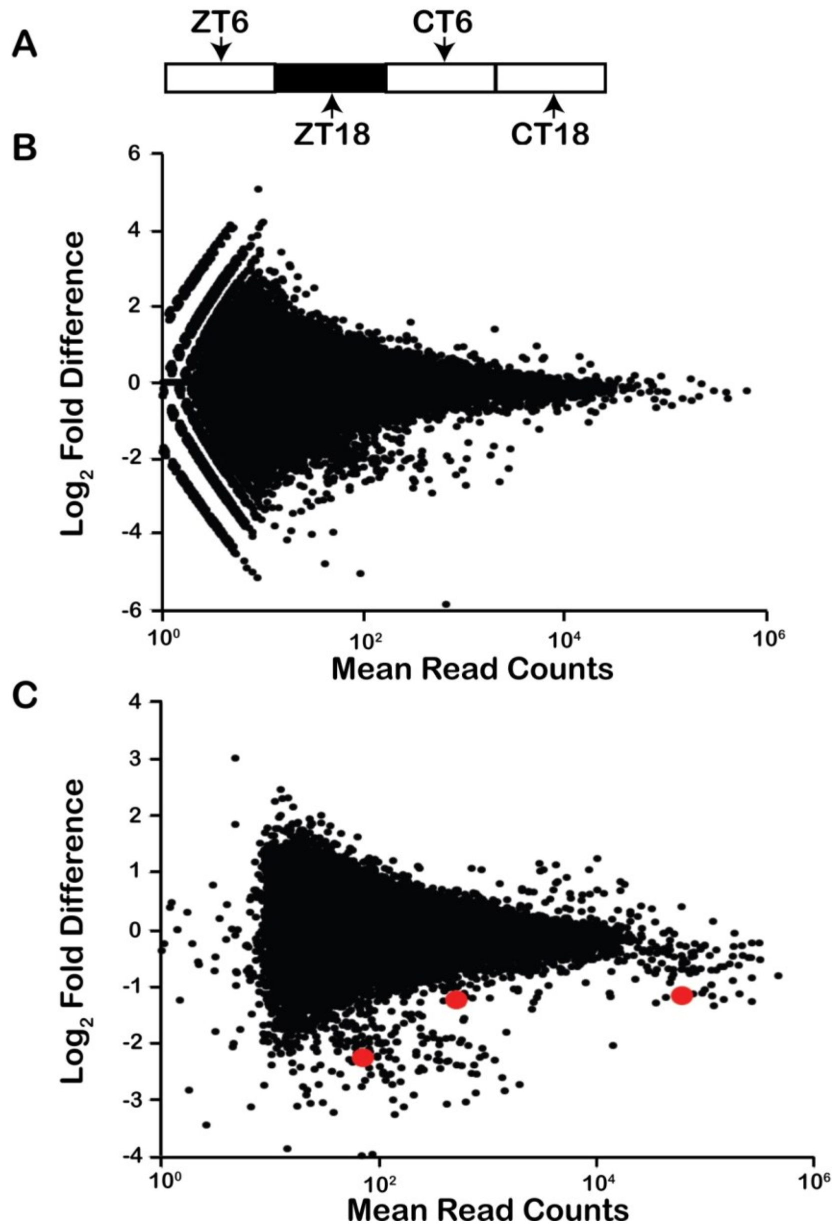



**Figure 2.S2. Analysis using duplicate ZT6 and ZT/CT18 samples**

A. Time of sampling in a light/dark cycle and constant light. B. MA plots using ZT 6/CT 6 samples and ZT 18/CT 18 pairs as duplicates were analyzed by DESeq. There are no significant differences (padj $< 0.001$) when mapped to the Trinity assembly. C. Three significant differences are identified when mapped to the Velvet assembly.

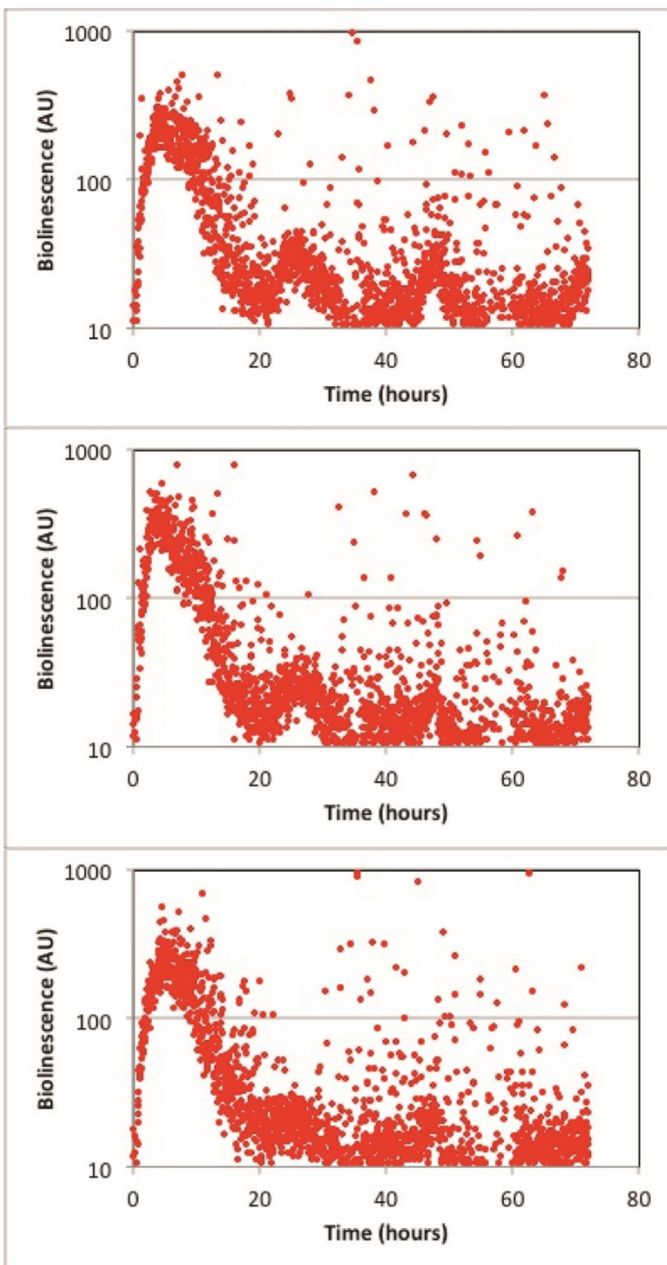



**Figure 2.S3. Individual bioluminescence traces.**

Three individual bioluminescence traces used to determine the average of inhibitor-treated cultures shown in Figure 2.3B.

# Chapitre 3: Publication # 2

# Characterization of two dinoflagellate cold shock domain proteins

Mathieu Beauchemin*, Sougata Roy*, Sarah Pelletier, Alex Averback, Frédéric Lanthier et David Morse

*Contribution égale des auteurs

Publié dans mSphere 1(1):e00034-15. doi:10.1128/mSphere.00034-15.

Contribution des auteurs :

Pour cet article, j'ai effectué une grande part des expériences avec S. Pelletier, préparé les figures et rédigé une partie de l'article. S. Roy a planifié les expériences, supervisé A. Averback et F. Lanthier, qui ont fait des expériences préliminaires, et rédigé une partie de l'article. D. Morse a supervisé le projet et participé à la rédaction et la révision de l’article.

## 3.1. Abstract

Roughly two thirds of the proteins annotated as transcription factors in dinoflagellate transcriptomes are cold shock domain-containing proteins (CSPs), an uncommon condition in eukaryotic organisms. However, no functional analysis has ever been reported for a dinoflagellate CSP so it is not known if they do in fact act as transcription factors. We describe here some of the properties of two CSPs from the dinoflagellate *Lingulodinium polyedrum*, LpCSP1 and LpCSP2, which contain a glycine-rich C-terminal domain and a N-terminal cold shock domain phylogenetically related to those in bacteria. However, neither of the two LpCSPs act as does the bacterial CSP since they do not functionally complement the *Escherichia. coli* quadruple cold shock domain protein mutant BX04 and cold shock does not induce LpCSP1 and LpCSP2 to detectable levels based on two-dimensional gel electrophoresis. Both CSPs bind to RNA and ssDNA in a non-specific manner in electrophoretic mobility shift assays (EMSA), and both proteins also bind double stranded DNA non-specifically albeit more weakly. These CSPs are thus unlikely to act alone as sequence specific transcription factors.

## 3.2. Introduction

Cold shock domains (CSD) are an ancient and conserved nucleic acid binding module [325]. They are small, roughly 70 amino acids in length, contain two amino acid motifs that are shared by RNA recognition motif domains, and have been shown to bind both DNA [326, 327] and RNA [328]. They are found in some archaea, in eubacteria, plants, animals and some fungal lineages. However, the roles played by the CSD depend on the organism and on specific domains associated with it. In bacteria, the cold shock response includes a transient and global translation arrest [329], during which time the cells synthesize a small number of cold-inducible proteins and begin reprogramming their translational machinery to accommodate growth at lower temperatures [330]. Cold shock proteins (CSPs), which in bacteria consist only of a CSD, are among the few proteins whose synthesis increases after a 15°C cold shock [331]. Many of these bacterial CSPs appear to be functionally redundant, as it is necessary to delete four of the nine *E. coli csp* genes before growth at cold temperatures is impaired [332].

The nucleic acid binding properties of CSPs are an integral part of the bacterial cold shock response. For example, *E. coli* CspA stimulates transcription of the DNA gyrase *gyrA* [333], an effect attributable to binding of specific DNA sequences in the *gyrA* promoter [327]. In addition, CspA can aid transcription by acting as an anti-terminator, an effect suggested to rely on CspA binding to single stranded regions of the newly synthesized RNA [334]. CSP binding to RNA has also been proposed to melt RNA secondary structures that form due to reduced temperature [70]. This may improve translation by eliminating inhibitory secondary structures and may also relieve transcriptional stalling due to the formation of stem-loop secondary structures in the nascent RNA. Lastly, CspA can bind its own transcript and protect it from degradation, as the half-life of the message increases roughly 100-fold at 15°C [335]. Thus in *E. coli*, most of the effects produced by the CSP appear to rely on a capacity for binding RNA.

Plant CSPs also contain a CSD, but in addition include a C terminal glycine-rich domain interspersed with a variable number of CCHC-type zinc finger DNA-binding domains [336, 337]. The model plant *Arabidopsis thaliana* has four CSD-containing proteins, some of

which are up-regulated by exposure to cold [336]. *Arabidopsis* overexpressing selected CSPs showed increased tolerance to freezing [338] and drought [339] through an mRNA chaperone activity [340]. Some (but not all) *Arabidopsis* CSPs also complement the cold-sensitive *E. coli* quadruple *csp* mutant [341]. However, these CSPs seem to play a broader role as regulators of embryonic development, seed germination and flowering [342, 343]. A similar function may be performed by two CSPs from *Oriza sativa*, as the highest expression is found in flowers and seeds [344].

The core CSD in animal CSPs is surrounded by an N-terminal alanine-proline rich domain and a C-terminal domain with alternating acidic and basic regions. The first example, a protein called YB-1, was identified due to its ability to bind the Y-box (CTGATTGGCCAA) in promoters of major histocompatibility locus genes [326], yet was subsequently revealed to have potent mRNA stabilizing activity [345]. Furthermore, YB-1 has also been shown to regulate translation of specific mRNAs [346], while the frog Y-box protein (FRGY2) was found in ribonucleoproteins particles (RNPs) that sequester maternal mRNAs in oocytes [347].

Recent transcriptomic analyses of dinoflagellates have shown an abundance of CSD proteins [61, 69]. These domains are classified as DNA binding by gene ontology (GO categories), and in dinoflagellate transcriptomes they constitute roughly two-thirds of all the potential transcription factors identified. However, given the many examples of RNA binding by CSPs, it is not clear if the transcription factor classification of dinoflagellate CSPs is accurate. To assess the function of these dinoflagellate proteins, we have expressed and purified two dinoflagellate CSPs, which we named *Lp*CSP1 and *Lp*CSP2, and compared their nucleic acid binding properties with those of the bacterial CSPs to which they are most closely related. Curiously, while able to bind both RNA and DNA in EMSA assays, the dinoflagellate proteins do not complement a bacterial CSP mutant and are not induced in cold conditions.

## 3.3. Results

### 3.3.1. *Lp*CSP1 and *Lp*CSP2 are part of a distinct clade in the eukaryotic CSP family.

Roughly two-thirds of dinoflagellate proteins classified as transcription factors by gene ontology (GO) are cold shock domain (CSD) containing proteins (CSPs) [61, 69]. The over-representation of this class of protein thus suggests they might play an important role in the regulation of gene expression. To begin to address the role of these CSPs, we first determined the types of CSPs expressed by dinoflagellates. We used the CSD in BLAST searches of the transcriptomes of *Lingulodinium* and *Symbiodinium* as well as ESTs from *Alexandrium tamarense* and *Karenia brevis*. We recovered a total of 23 full-length sequences as defined by the presence of a single ORF with in-frame start and stop codons encompassing the CSD. All dinoflagellate CSDs contain the two characteristic RNA binding motifs (KGFGFI and VFVHF) (Figure 3.1.). However, the CSD itself is found in at least four different domain architectures. The vast majority of the sequences recovered contained a CSD either alone or with a C-terminal G-rich domain (Figure 3.1.). Smaller numbers of representatives contained a Zn-finger domain following the G-rich domain, and even fewer examples were found of sequences with multiple CSDs and one or more RNA recognition motifs (RRM) (pie chart distribution of sequence numbers in Figure 3.1.). The domain structure of most dinoflagellate CSPs is thus closest to that found in bacteria or in plants, although there is no sequence similarity between the G-rich domains of plant [348] and dinoflagellate CSPs.

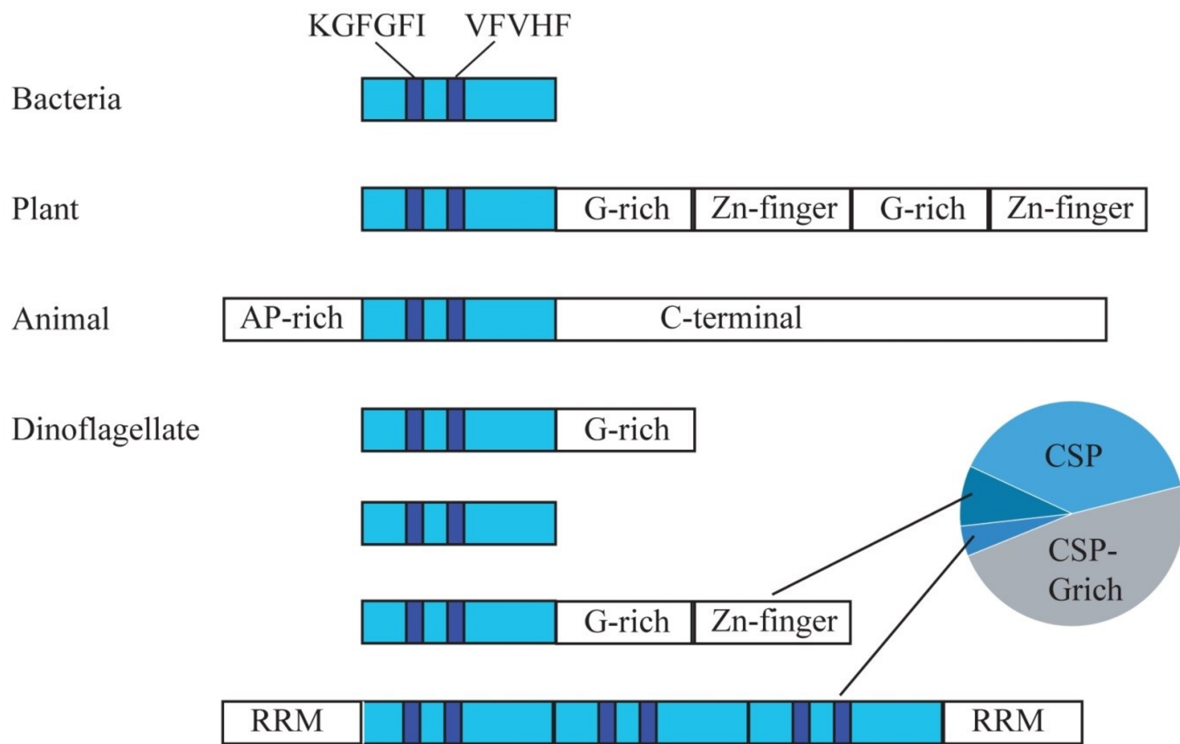



**Figure 3.1. Cold Shock Proteins (CSP) contain a conserved cold shock domain (CSD).**

The ~70 amino acid Cold Shock Domain (CSD, light blue) is highly conserved in bacteria, plants, animals and dinoflagellates and contains two RNA recognition motifs (dark blue, KGFGFI and VFVHF). Bacterial CSPs consist only of a CSD, while the other three classes contain C-terminal extensions. The C-terminal of plant CSPs typically contain two to seven repeats of a glycine rich (G-rich) – Zn finger region, a pattern observed in a few dinoflagellate proteins. More typically, when present, the C-terminal extensions of dinoflagellate are G-rich only. An atypical CSD containing protein found in *Lingulodinium* has three CSD repeat flanked by RNA recognition module (RRM). Finally, animal CSDs are unique in that they contain an AP-rich N-terminal extension. The pie chart shows the relative abundance of the four different dinoflagellate CSP architectures in the transcriptome of *Lingulodinium* and *Symbiodinium sp.* and in the ESTs of *Alexandrium tamarense* and *Karenia brevis*.

As a complement to the domain structure analysis, a phylogenetic analysis of the CSDs from ten *Lingulodinium* sequences was also carried out using sequences from a wide array of organisms. In general, support for the different clades is poor, as the CSD is short and the sequence quite conserved. However, there is strong support for grouping the dinoflagellate sequences into two different clades (Figure 3.2.). One of these two clades also contains both animal and bacterial CSPs, and since the bacterial CSPs are among the best characterized, we elected to begin our analysis with *Lingulodinium* representatives from this clade.

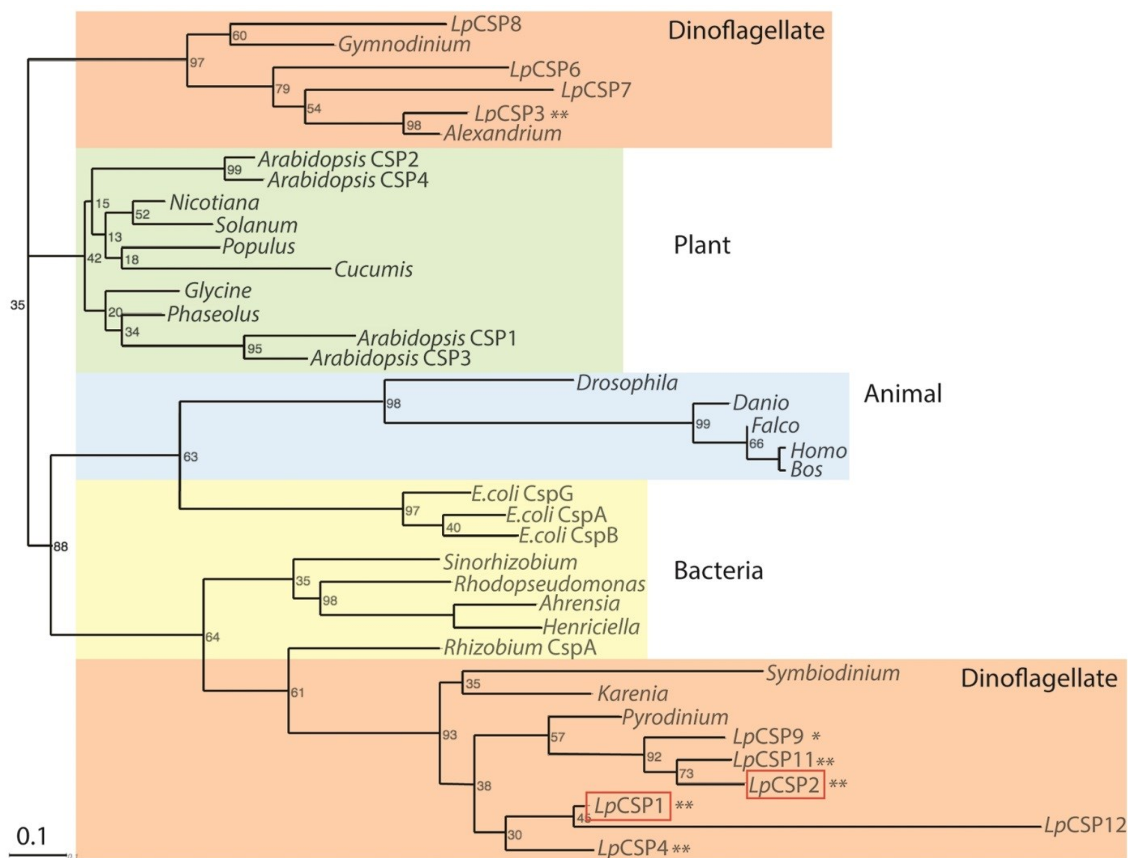



**Figure 3.2. Dinoflagellate CSPs are found in two distinct clades.**

Phylogenetic reconstructions place dinoflagellate CSPs in two clades, one clustered with bacterial and animal sequences and the second in a separate clade unique to dinoflagellates. The *Lingulodinium* CSPs analyzed here are boxed. Sequences marked with a single asterisk contain only a CSD, while those marked with two asterisks contain a CSD and a G-rich domain. The accession numbers for the sequences used are listed in the methods.

### 3.3.2. *Lp*CSP1 and *Lp*CSP2 are not functional complements of *E. coli* CSPs

Two dinoflagellate CSPs, which we have termed *Lp*CSP1 and *Lp*CSP2, were cloned and expressed in *E. coli*. Both are small proteins (113 and 110 residues, respectively) whose domain structure is similar to that found in the largest class of CSD proteins (an N terminal CSD followed by a Glycine-rich domain). In both, the CSD contains the two expected RNA-binding motifs (Figure 3.1. and 3.S1.). Bacterial CSPs are required for cell growth at low temperature [349], and a bacterial strain harboring a mutation in four different CSP genes is unable to grow at 17°C [332]. However, while over-expression of the bacterial CspA gene into the quadruple mutant allows for growth at lower temperatures (Figure 3.3.), neither an empty vector (pINIII) nor the *Lingulodinium* CSPs in pINIII are able to fulfill this role. These dinoflagellate CSPs thus differ from the bacterial CSPs.

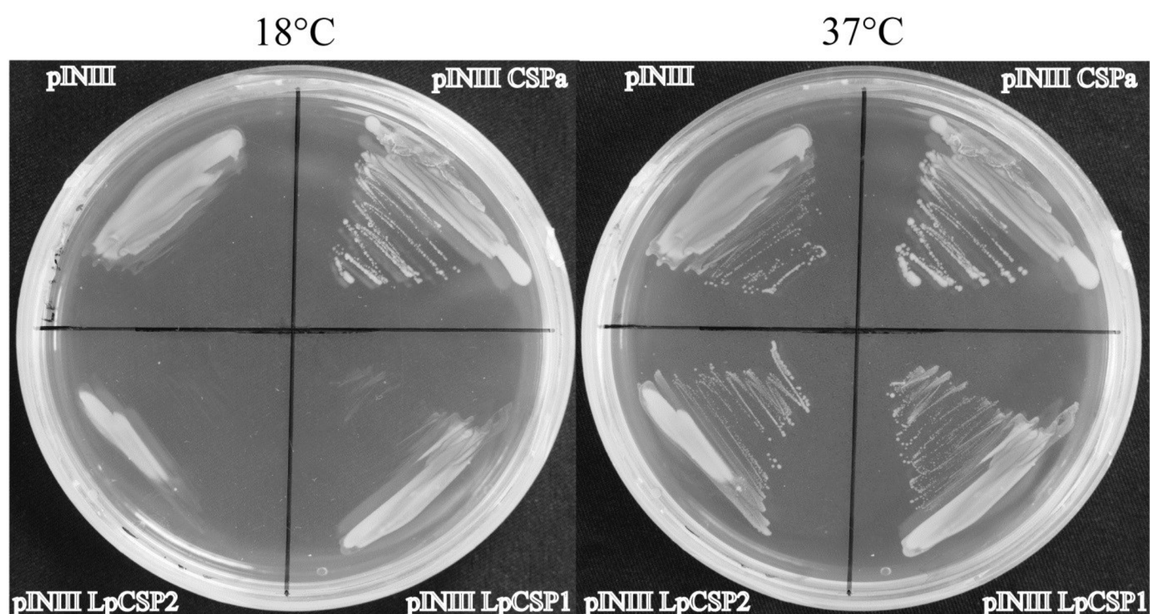



**Figure 3.3. Dinoflagellate CSPs do not complement a quadruple bacterial CSP mutant.**

Plates were streaked with a quadruple CSP mutant strain BX04 transformed with an empty vector (pINIII), an *E. coli* CspA, *Lp*CSP1 or *Lp*CSP2 and grown first at 18°C for five days (left panel) then transferred to 37°C overnight (right panel).

### 3.3.3. *Lp*CSP1 and *Lp*CSP2 expression are not detectable after cold shock

Since bacterial CSPs are strongly induced by cold temperature, we also tested if *Lp*CSP1 and *Lp*CSP2 induction could be detected by 2D-PAGE, as seen in *E. coli* [350]. The encystment of *Lingulodinium* as a cellular response to cold (8°C) temperatures has previously been reported [304], and we therefore compared the protein profile of cells grown at normal culture room temperatures (Figure 3.4. A) with those of cysts at 8°C (Figure 3.4. B). To determine the expected position of *Lp*CSP1 and *Lp*CSP2, 150 µg of *Lingulodinium* protein samples were spiked with 150 ng of purified LpCSP1 or LpCSP2 before electrophoresis (Figure 3.4. C-D). No detectable induction of *Lingulodinium* CSPs was observed after 24 h at 8°C.

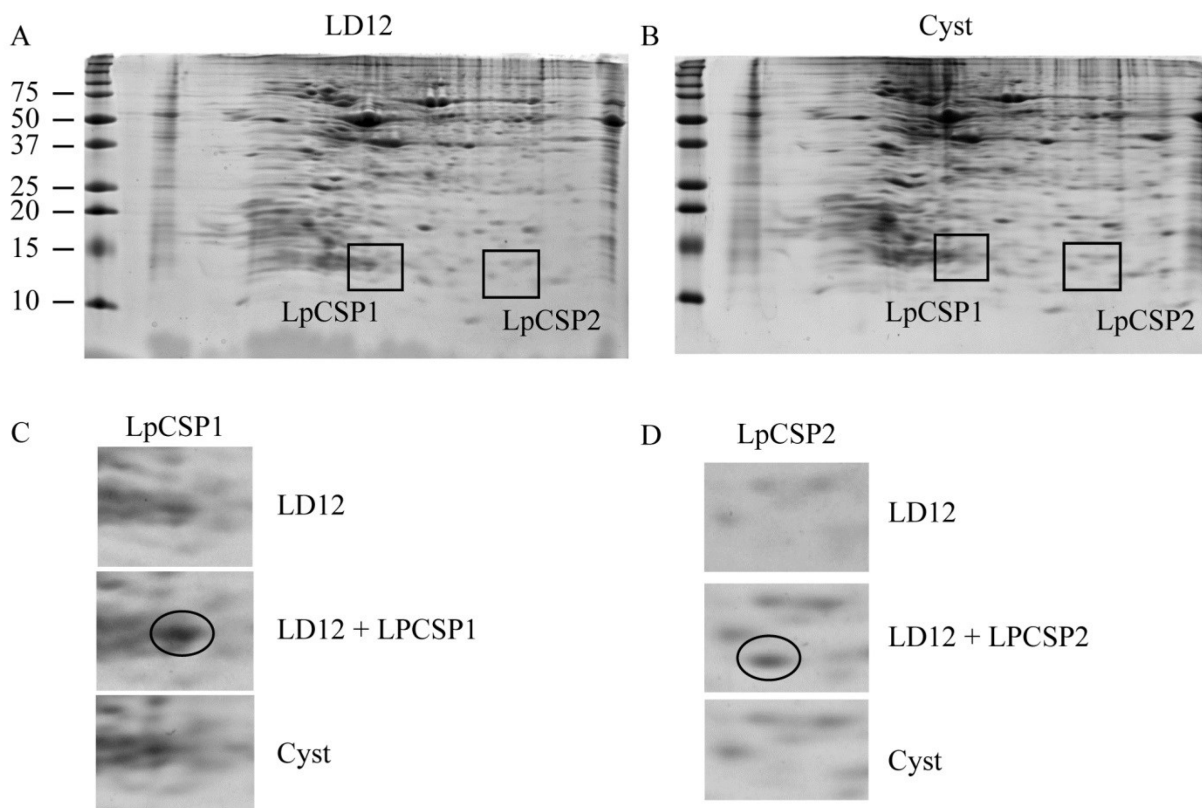



**Figure 3.4. Cold treatment does not induce detectable levels of *Lp*CSP1 and *Lp*CSP2.**

2D PAGE analysis of *Lingulodinium* total proteins from cells grown at 18°C and collected at LD12 (A) and from cysts harvested after 24h at 8°C (B). Boxed area in A and B corresponding to the position of LpCSP1 (C) and LpCSP2 (D) are shown from both gels and from a gel (middle panel) in which 150 ng of recombinant protein was added before isoelectric focusing (circled area). Molecular weights in kDa were estimated using Precision Plus Protein Standards (Bio-Rad).

### 3.3.4. *Lp*CSPs bind nucleic acids *in vitro*

Binding of DNA and RNA was evaluated for *Lp*CSP1 and *Lp*CSP2 using EMSA. Both proteins were expressed as GST fusions, purified, and the GST domain removed by thrombin digestion prior to use. However, since some residual GST remained in the purified protein fraction (Figure 3.S2.), GST alone was also tested for binding under the same conditions. When single stranded DNA (ssDNA) was used as a probe, both *Lp*CSP1 and *Lp*CSP2 showed

the concentration dependent slower migrating bands on EMSA, indicative of nucleic acid binding (Figure 3.5. A). However, a considerable amount of radiolabeled DNA remained in the loading wells, suggestive of binding by a multimeric CSP to several probe molecules. Both proteins were also able to bind double stranded DNA (dsDNA), although in this case all bound radiolabel was observed to migrate into the gel (Figure 3.5. B). Lastly, binding to radiolabelled RNA was also detected for both proteins, with binding similar to that observed for an authentic *E. coli* CSP (Figure 3.5. C). Both the two *Lingulodinium* proteins thus display a broad nucleic acid binding capacity.

To assess the possibility that *Lp*CSP1 might display a capacity for sequence specific binding, different competitors were added to *Lp*CSP1 in the presence of a ssDNA probe (Figure 3.S3. A). These competition experiments indicate that dsDNA competes poorly for binding to the ssDNA, and that altering the sequence of the competing ssDNA does not affect its ability to compete. In contrast, ssDNA is an effective competitor of the binding between *Lp*CSP1 and a dsDNA probe (Figure 3.S3. B). *Lp*CSP1 thus prefers ssDNA substrates.

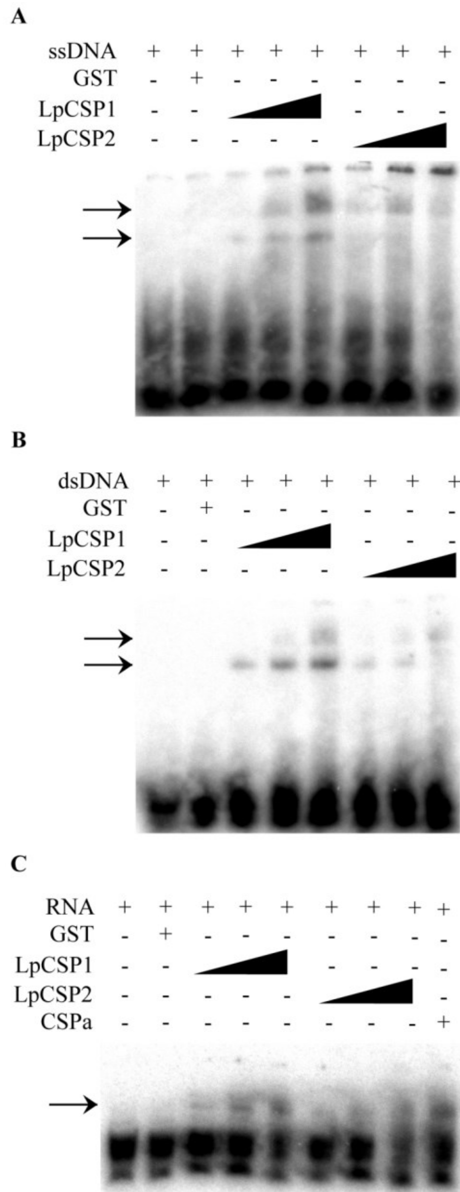



**Figure 3.5. LpCSP1 and LpCSP2 show similar nucleic acid binding activity in EMSA.**

Electrophoretic Mobility Shift Assays (EMSA) was carried out using a ssDNA probe (A), a dsDNA probe (B) and a RNA probe (C). The black triangle above the autoradiograms denotes the different concentrations of the CSPs recombinant protein used for the assays (100, 300 and 1000 ng in all the assays; 1000 ng of CSPa in C). Arrows show the bands corresponding to shifted nucleic acid.

## 3.4. Discussion

In this report we have assessed the role of two dinoflagellate cold shock domain-containing proteins (CSPs). CSPs are potentially important in dinoflagellates since high-throughput transcriptome studies in *Lingulodinium* [61] and *Symbiodinium* [69] have shown that the majority of the proteins annotated as transcription factors have cold shock domains (CSDs). Despite this annotation, however, it is not clear whether any of the dinoflagellate CSPs actually play a role in transcription. To date, with the sole exception of a TBP-like protein, which replaces the TATA binding protein usually found in eukaryotic TFIID complexes [68], no transcription factor has been described and characterized experimentally in dinoflagellates.

Two types of CSPs are likely predominant in dinoflagellates as based on the protein domain structure. One form contains only the CSD, while the other contains the CSD and a C terminal G-rich domain (Figure 3.1.). Although this analysis is based on the frequency of finding different types of sequence within a transcriptome, it seems likely that the more frequently found forms would result in a greater amount of protein. The cellular *Lingulodinium* CSP pool may thus be dominated by these two forms, although we cannot rule out the possibility that a protein with a different domain structure could be highly expressed.

Molecular phylogeny of the *L. polyedrum* CSD sequences indicates they are found in two well-supported clades (Figure 3.2.), perhaps indicative of a functional diversity among dinoflagellate CSPs. We have examined the properties of two members of the group most closely related to bacterial CSPs, *Lp*CSP1 and *Lp*CSP2 (Figure 3.2.). Despite this relationship with bacterial sequences, however, neither protein appears functionally equivalent to those in bacteria. This is most clearly shown through its inability to complement the growth of the mutant *E. coli* BX04 strain at low temperatures (Figure 3.3.). The non-specific mRNA binding activity of bacterial CSPs is the key to their chaperone activity during cold stress [325], so the lack of complementation by *Lp*CSP1 and *Lp*CSP2 is puzzling given that both dinoflagellate proteins are able to bind a RNA probe *in vitro* (Figure 3.5. C). However, comparison of two similar *Arabidopsis* CSPs, each containing an N terminal CSD and a C terminal G-rich domain, showed that only one was able to complement the BX04 mutant [341]. It is thus

possible that small differences between sequences are sufficient to disrupt the RNA chaperone activity, which may in turn suggest that other dinoflagellate CSD containing proteins could show RNA chaperone activity and potentially rescue the BX04 strain. However, it is also possible that the lack of complementation in the bacterial assay may reflect the amount of the *Lingulodinium* CSPs actually expressed in the bacteria or an element of sequence specificity that remained undetected in our *in vitro* assays.

We find no evidence that *Lp*CSP1 and *Lp*CSP2 are induced by exposure to cold temperatures. Transcription of both *Arabidopsis* and *E. coli* CSPs can be induced during cold shock [341, 351], but a recent comparison of the transcriptomes of normally growing and cold-shocked *L. polyedrum* cells showed no induction of CSP transcripts by a cold treatment [304]. In addition, exposure to cold induces CSP at the protein level in both prokaryotes and eukaryotes, in keeping with their ability to prevent cells from freezing, but a global analysis of the proteome of cold shocked *L. polyedrum* showed no difference when compared to normally growing cells [304]. This agrees with the 2D-gel analysis shown here where we have specifically looked for *Lp*CSP1 and *Lp*CSP2 induction at low temperatures (Figure 3.4.). Thus, unlike the bacterial CSPs [331, 350], there is no induction of *Lp*CSP1 or *Lp*CSP2 mRNAs nor are more proteins synthesized, although we cannot rule out a potential post-translational modification that could modify the position of CSPs spots on a 2D-gel. Taken together, a role in cold tolerance seems unlikely for these dinoflagellate CSPs.

Unlike the bacterial CSPs, whose cellular role seems primarily to block formation of extensive RNA secondary structure at cold temperatures [334, 352], CSPs in vertebrates and plants display a number of different roles. The mammalian Y-box binding protein YB1, which shares 40% amino acid identity with the bacterial CSPs [353, 354], has diverse physiological roles apart from cold stress response [337]. YB1 is known to regulate transcription [355, 356] by binding to duplex DNA containing a Y-box (CTGATTGGCT) [326]. Similarly, the frog Y-box protein FRGY1 has been shown to stimulate transcription from a promoter containing a Y-box [357]. YB-1 has been shown to bind ssDNA with greater affinity than dsDNA [73], and the resulting destabilization of the DNA double helix has been proposed as the transcriptional activation mechanism. However, *Lp*CSP1 and LpCSP2 bind a Y-box dsDNA at very high protein concentrations (300 and 1000 ng, Figure 3.5. B) compared to binding of Y-box DNA

sequence by YB1 where only 30 ng is sufficient [358]. The physiological levels of the *Lp*CSP1 and *Lp*CSP2 protein appear much lower than this, as indicated by our 2D-gels profiles of normally growing cells (Figure 3.4.). This binding also does not appear to be sequence specific (Figure 3.S3.), indicating that specific targeting of Y-box sequence is doubtful. However, it must be noted that we have not yet tested if other DNA sequences might show preferential binding enabling lower concentrations of *Lp*CSPs to be effective. The stronger binding to ssDNA leave the door open to a potential role of *Lp*CSPs as destabilizer of the DNA double helix, potentially in cooperation with other, more specific, transcription factors.

Plants CSD-containing proteins also have a greater role than a simple response to cold shock. The precise molecular mechanism that enables plant CSPs to regulate processes other than acclimation to abiotic stress, such as seed and flower development, is still unknown. However, the interactions of AtCSP3 with diverse proteins such as poly-A binding proteins, ribosomal proteins and mRNA decapping protein suggest an involvement in multiple RNA processing steps [359]. A similar role might be envisioned for *Lp*CSP1 and *Lp*CSP2 as they are able to bind RNA (Figure 3.5. C). Furthermore, the non-specific RNA-binding by *Lp*CSPs is consistent with a role in mRNA packaging and stability [360], an intriguing prospect in dinoflagellates where particularly long mRNA half-lives have been documented [90].

Taken together, our results are most consistent with the idea that *Lp*CSP1 and *Lp*CSP2 are not sequence specific transcription factors, contrary to the impression left by the Gene Ontology assignment. This further reduces the already scarce number of potential transcription factors in dinoflagellate, a fact that fits well with the limited scale of transcriptional variation seen during the circadian cycle [361], as well as in response to nutrient limitation [86, 87] or abiotic stress [304]. We suggest that *Lp*CSPs may still be able to participate in the transcription process, potentially by unwinding the DNA helix due to their capacity to bind ssDNA. However, further studies, such as protein localization and identification of binding partners will be required to define more precisely the role of these proteins in dinoflagellates. It will also be of interest to examine members of the second dinoflagellate clade in case these CSPs have substantially different properties.

## 3.5. Methods

### 3.5.1. Cell Culture

Unialgal but not axenic cultures of *Lingulodinium polyedrum* (formerly *Gonyaulax polyedra*; strain CCMP1936) were obtained from the Provasoli-Guillard National Center for Marine Algae and Microbiota (Boothbay Harbor, Maine) and grown in modified seawater medium (f/2) [362] at constant temperature (18 ± 1°C). The culture room's light cycle was 12h with cool white fluorescent light at an intensity of 50 μmol photons $m^{-2}$ $s^{-2}$ followed by 12h dark. Under these conditions, the beginning of the light period is termed LD 0 and the beginning of the dark period LD 12. Cultures were typically grown to a cell density of 12–14,000 cells/mL before cell collection by filtration on Whatman 541 paper supported by a Buchner funnel. Cysts were obtained by placing the cultures at 8°C for 24 h as described [304]. All cells were either used immediately or frozen in liquid nitrogen and stored at -80°C until further use.

### 3.5.2. Sequence alignment and Phylogenetic Analyses

The CSP sequences used for sequence alignment and phylogenetic analyses were obtained from the *Lingulodinium* transcriptome deposited at NCBI. A search of the *Lingolodinium* transcriptome deposited at The Marine Microbial Eukaryote Transcriptome Sequencing Project [363] was also done to retrieve complete sequences when needed. Trees were constructed using an online tool available at www.phylogeny.fr [364]. In the workflow, multiple sequences were aligned using MUSCLE and curated using GBlocks. Phylogenetic reconstructions were made using PhyML and visualized using TreeDyn. Sequences used for phylogenetic reconstructions were from *Ahrensia* (WP_018687722), *Alexandrium* (GAIT01073544), *Arabidopsis* CSP1 (AEE86603), *Arabidopsis* CSP2 (NP_195580), *Arabidopsis* CSP3 (NP_565427), *Arabidopsis* CSP4 (Q38896), *Bos* (DAA26237), *Cucumis* (XP_004140332), *Danio* (XP_001340141), *Drosophila* (NP_647983), *E. coli* CspA (AAN82813), *E. coli* CspB (AAB61739), *E. coli* CspG (NP_309172), *Falco* (XP_005244100), *Glycine* (XP_003540832), *Gymnodinium* (GAIL01018775), *Henricella* (WP_018146825),

*Homo* (EAW48433), *Karenia* (FK848095), *Lingulodinium* 1 (JO733348), *Lingulodinium* 2 (JO729000 and CAMNT_0033837443), *Lingulodinium* 3 (JO730956 and CAMNT_0033776061), *Lingulodinium* 4 (JO734870 and CAMNT_0033828465), *Lingulodinium* 6 (JO720996 and CAMNT_0033712139), *Lingulodinium* 7 (JO766444 and CAMNT_0033724295), *Lingulodinium* 8 (JO761018 and CAMNT_0033635737), *Lingulodinium* 9 (JO730992 and CAMNT_0033829387), *Lingulodinium* 11 (JO736519), *Lingulodinium* 12 (JO732587), *Nicotiana* (P27484), *Pyrodinium* (GAIO01020278), *Rhizobium* (YP_770349), *Phaseolus* (ESW08176), *Populus* (XP_002313723), *Rhodopseudomonas* (NP_948738), *Sinorhizobium* (AAC64672), *Solanum* (XP_006359670) and *Symbiodinium* 1 (GAFO01002801).

### 3.5.3. CSPs cloning, expression and purification

Primers designed from the *Lingulodinium* transcriptome sequence JO733348 and JO729000 (Table 3.S1.) [61], were used to amplify what was termed *Lp*CSP1 and *Lp*CSP2, respectively, from a first strand cDNA reaction product prepared from total RNA extracted from *L. polyedrum* cells using Trizol (Invitrogen). The reverse transcription reaction was performed with a MMLV RT (Clontech) and the 5′ CDS primer A of the SMARTer RACE cDNA Amplification kit (Clontech). The sequences were cloned in the pGEM-T vector (Promega) and verified by sequencing. To allow directional cloning into the multiple cloning site of the bacterial expression vectors pGEX-4T2 (GE Healthcare) and pINIII, a second PCR was performed on the pGEM-T plasmid containing *Lp*CSP1 or *Lp*CSP2 to add proper restriction sites (Table 3.S1.). *E. coli* CSPa was amplified from a pINIII-CspA plasmid [332] and cloned in pGEX4T2 using similar procedures. The correct frame was verified by sequencing and the size of GST-*Lp*CSP1, GST-*Lp*CSP2 and GST-CspA fusion protein was verified on SDS PAGE. The pGEX4T2-*LpCSP1*, pGEX4T2-*Lp*CSP2 and pGEX4T2-CspA vector were used to transform chemically competent BL21 host cells (Life Technologies).

A single colony of BL21 *E. coli* containing either pGEX4T2-*Lp*CSP1, pGEX4T2-*Lp*CSP2 or pGEX4T2-CspA was inoculated in 5 mL of Luria Bertani (LB) broth and cultured overnight at 37°C with vigorous shaking. The overnight cultures were transferred to 250 mL

of LB medium supplemented with ampicillin (100 μg/ml) and grown at 37°C with vigorous shaking to an optical density of 0.5 at 600 nm. At this point, protein expression was induced for 2h by addition of Isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 0.2 mM. After harvesting by centrifugation, the bacterial pellets were resuspended in cold phosphate buffered saline (PBS) (137 mM NaCl, 2.7 mM KCl, 4.3 mM $Na_2HPO_4$, 1.47 mM $KH_2PO_4$) containing 1 mM phenylmethylsulfonyl fluoride (PMSF), 1 mM Dithiothreitol (DTT), 1 mM Ethylenediaminetetraacetic acid (EDTA) and 0.25% Triton-X100 and broken with a French pressure cell press (Fisher Scientific). The cell lysates were then centrifuged at 10,000 g for 10 min and the supernatants were incubated with 100 μL of Gluathione Sepharose 4B beads (Promega) for 45 min at room temperature with end-over-end agitation. Beads were washed 4 times in PBS and resuspended in 300 μL of PBS supplemented with 2 units of thrombin at room temperature for 2h to cleave the GST tag. Supernatants containing the cleaved CSPs were then electrophoresed by SDS-PAGE on an 18% acrylamide gel to assess purity and protein concentration was estimated using the Bradford assay (BioRad). Aliquots of purified protein were frozen in liquid nitrogen and stored at -80°C until further use.

### 3.5.4. Bacterial complementation assay

BX04, a quadruple deletion *E. coli* mutant lacking four CSPs, was used to assess the role of *Lp*CSP1 and *Lp*CSP2 in allowing growth at low temperature (18°C) [332]. The *E. coli* CspA in pINIII and the empty pINIII vector were used as positive and negative controls, respectively. The plasmids were transformed into chemo-competent BX04 cells. A single transformed BX04 colony with the respective plasmids was inoculated in 5 ml of LB medium and then cultured overnight at 37°C with vigorous shaking. The cultures were streaked on a single LB plate with ampicillin (100 μg/ml) and IPTG (0.2 mM) and grown at 18°C for 120h. Growth was monitored every 24h and after 120h the same plate was incubated overnight at 37°C.

### 3.5.5. 2D PAGE of protein from Cysts, LD12 cells and thrombin-cleaved LpCSP1 and LpCSP2

Total protein was extracted from cysts and LD12 cells using Trizol (Invitrogen) as described [304]. A total of 150 µg of LD12 or cyst proteins, or 150 µg of LD12 proteins spiked with 150 ng of either *Lp*CSP1 or *Lp*CSP2, were loaded on separate 7 cm immobilized pH gradient strips (pH 3-10, Bio-Rad) as the first dimension. Gels containing 15% Acrylamide SDS-PAGE were used for the second dimension and were stained overnight with Coomassie Brilliant Blue G-250 [365].

### 3.5.6. Electrophoretic mobility shift assays

Double-stranded (ds) and single-stranded (ss) oligonucleotides were designed as described [358] and used after HPLC purification. The ssDNA was a 32-mer 5′-TCGATCGGGGCGGGGCGATCGGGGCGGGGCGA-3′ and the dsDNA was prepared by mixing equimolar amount of the 25-mer 5′-GGTGAGG<u>CTGATTGGCT</u>GGGCAGGA-3′ (the Y-box is underlined) and its reverse complement. All oligonucleotides were end-labeled with [γ-$^{32}$P]-ATP (Perkin Elmer) using Polynucleotide Kinase (NEB) and the conditions described by the manufacturer. Labeled probes were purified using the QIAquick Nucleotide Removal kit (QIAGEN). Typically, 1 ng radiolabeled ssDNA or dsDNA probes were incubated with the purified proteins for 30 min in 1X binding buffer (2X binding buffer contained 20 mM Tris-Cl pH 7.0, 50 mM KCl, 1 mM DTT, 20 mM $MgCl_2$ and 10% glycerol) at room temperature. Binding was assessed by electrophoresis on 5% native polyacrylamide gels in 0.5X TBE buffer at 70V for 1h at room temperature followed by an overnight autoradiography on a phosphorimager screen (Amersham) and subsequent imaging with a Typhoon Trio+ (Amersham). For competition assays, the binding buffer was supplemented with a 50X excess of unlabeled ssDNA or dsDNA probe for specific competition and a 50X excess of random single stranded oligonucleotides or salmon sperm DNA for non-specific competition.

RNA probes were prepared by *in vitro* transcription using the T7 RiboMAX RNA production system (Promega) from a 120 base pair dsDNA template that included the 5’UTR

of the Peridinin Chlorophyll-a binding protein (PCP, accession number U93077) preceded by the spliced-leader sequence [78] and a T7 promoter. The ds-DNA templates were degraded after completion of the reaction using an RQI RNase-free DNase (Promega). The purified RNAs were end-labeled by the same procedure used for ssDNA and dsDNA probes, and unincorporated nucleotides were removed by chromatography on a Bio-Gel P10 column (Bio-Rad). Binding assays were performed as described above.

## 3.6. Acknowledgements

We gratefully acknowledge the generous gift of the BX04 mutant E. coli, CspA in pINIII and the empty pINIII vectors from Drs. M. Inouye and S. Phadtare. M. B. is the recipient of a doctoral scholarship and S. P. and F. L. are recipients of an undergraduate student research award from the Natural Sciences and Engineering Research Council (NSERC). Funding for this work is provided through a NSERC research grant to D. M.

Tableau 3.S1. Primers used for amplifying and cloning *Lp*CSP1 and *Lp*CSP2 in the pGEX4T2 and pINIII plasmids.

| | **Accession number** | **Primers name** | **Primers sequence (5' to 3')** |
|---|---|---|---|
| LpCsp1 | JO733348 | CSP1F | GCAGCAATGCCTTCCGGCACTGTGAAGAAG |
| | | CSP1F BamH1 | TGACACGGATCCATGCCTTCCGGCACTGTGAAGAAG |
| | | CSP1F Nde1 | TGACCATATGCCTTCCGGCACTGTGAAGAAG |
| | | CSP1R | ACCCTCAGCTCAGAAACCTGAGGAGGGTCC |
| LpCsp2 | JO729000 | CSP2F | ATGGCCCAGAGCGGCACGG |
| | | CSP2R Xho1 | AACCCTCGAGTCACCAGTAGTCGCCCTTCCC |
| | | CSP2F Nde1 | AACCCATATGGCCCAGAGCGGCACGG |
| | | CSP2R | TCACCAGTAGTCGCCCTTCCC |
| CspA | NP_418012 | CSPaF | ATGTCCGGTAAAATGACTGGTATC |
| | | CSPaR | TTACAGGCTGGTTACGTTACCAG |
| | | CSPaR XhoI | AACCCTCGAGTTACAGGCTGGTTACGTTACCAG |

Restriction sites used for cloning are underlined

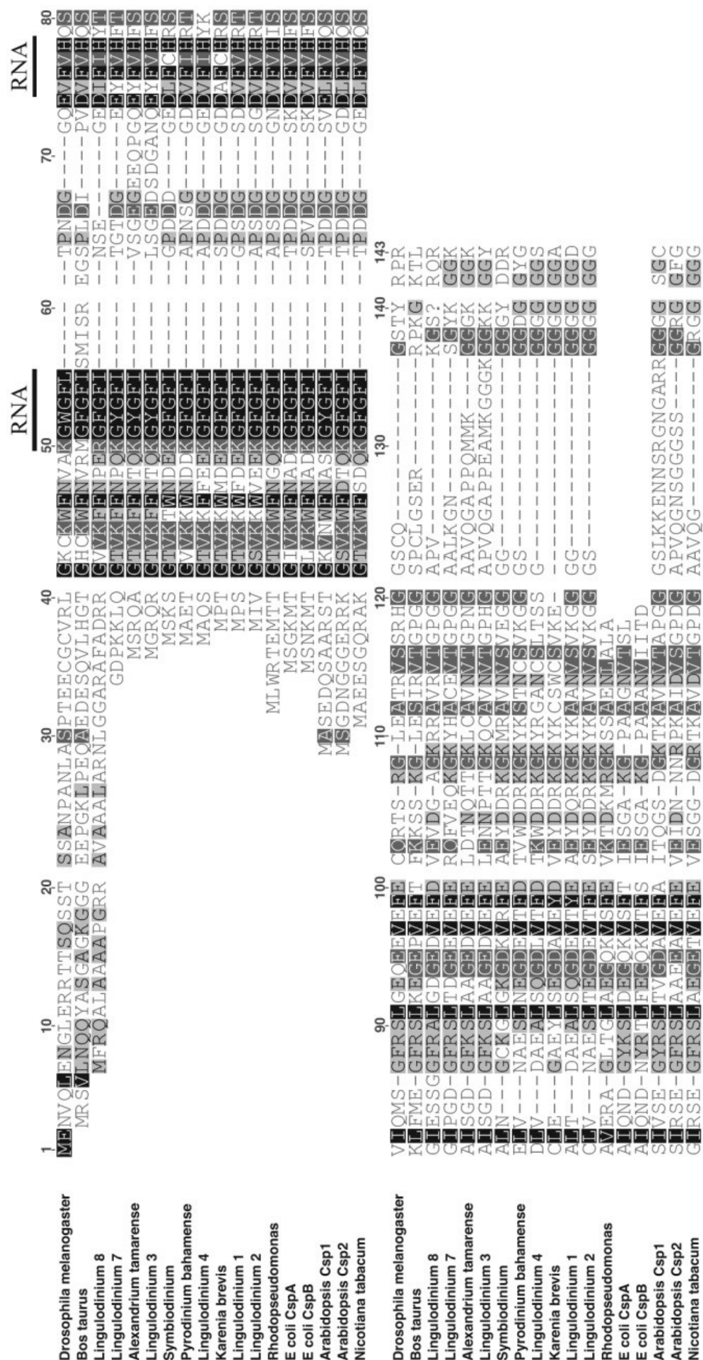



**Figure 3.S1. Sequence alignment of CSD protein**

Sequences of 18 different CSD proteins from a wide diversity of organisms were aligned using MUSCLE. Conserved residues important for RNA binding are highlighted. Accession numbers are listed in the methods section.

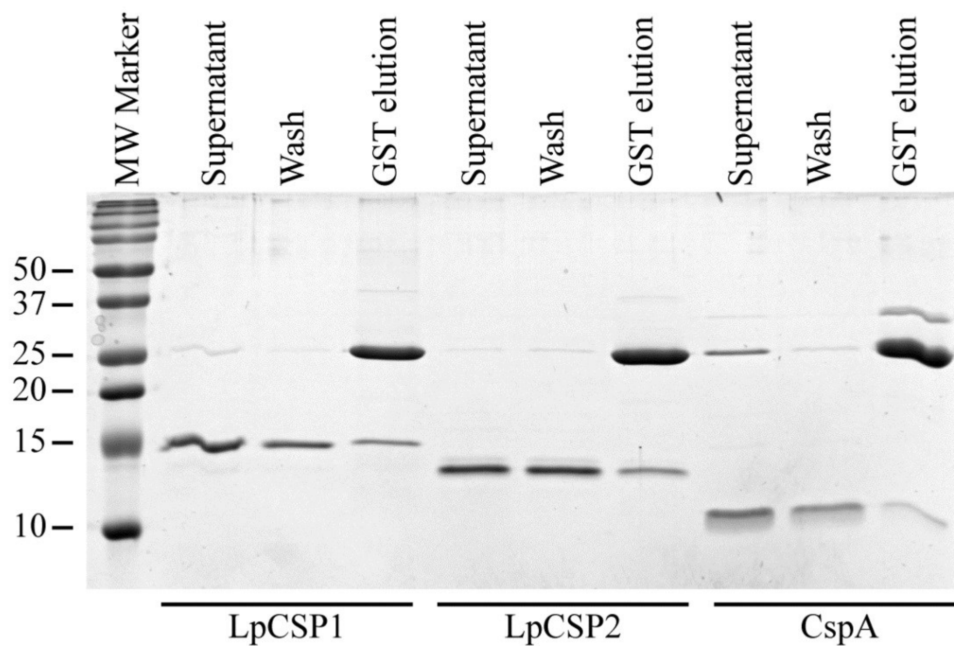



**Figure 3.S2. LpCSP1, LpCSP2 and CspA purification**

Recombinant protein were analysed on a 18% acrylamide SDS-PAGE gel to assess the purity of CSP proteins after GST-tag cleavage by thrombin (supernatant lane). Subsequent wash and elution of the Gluathione Sepharose 4B beads are also shown to assess cleavage efficiency. Protein size in kDa was estimated by comparison to a molecular weight marker (Precision Plus Protein Standards, Bio-Rad)

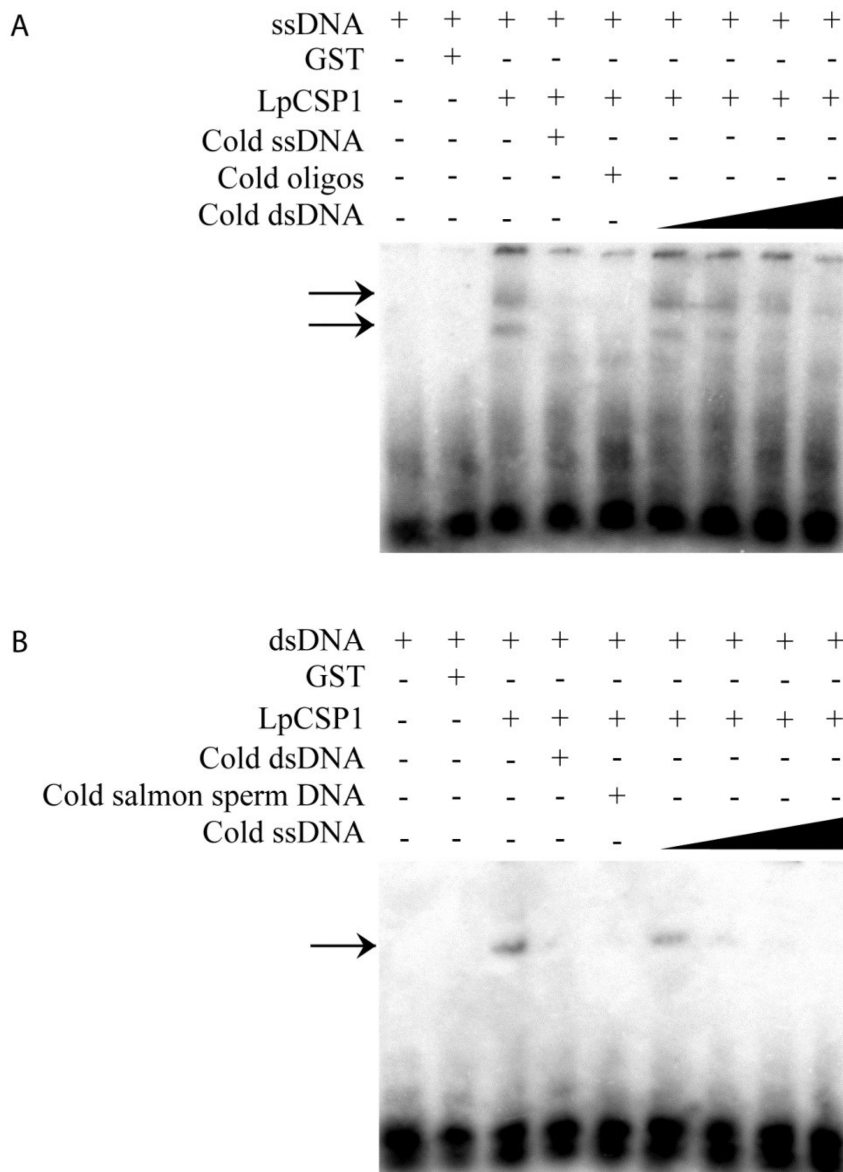



**Figure 3.S3. *Lp*CSP1 shows a preference for ssDNA over dsDNA**

Electrophoretic Mobility Shift Assays (EMSA) was carried out using a ssDNA probe (A) or a dsDNA probe (B) and 600 ng of *Lp*CSP1. Specific and non-specific competition were assessed by adding 50X unlabelled probe or 50X of a source of unlabelled ssDNA (A) or dsDNA (B). The black triangle above the autoradiograms denotes the different fold excess of cold probe added to the assay (1X, 10X, 30X and 100X). Arrows show the bands corresponding to shifted nucleic acid.

# Chapitre 4: Publication # 3

# Abundant RNA binding proteins in the dinoflagellate nuclear proteome

Mathieu Beauchemin et David Morse

Article en préparation

Contribution des auteurs :

Pour cet article, j'ai fait l'entièreté des expériences, des analyses bio-informatiques et la rédaction de l'article, qui a été révisé par D. Morse.

## 4.1. Abstract

The chromatin of dinoflagellates is unique among eukaryotes, as their chromosomes are permanently condensed in a liquid crystal state instead of being packed in nucleosomes. However, how it is organised is still an unsolved mystery, in part due to the lack of a comprehensive catalog of nuclear proteins in dinoflagellates. Here, we report the results of CHromatin Enrichment for Proteomics (CHEP) followed by shotgun mass spectrometry sequencing of the chromatin-associated proteins from the dinoflagellate *Lingulodinum polyedrum*. Our analysis identifies proteins involved in DNA replication and repair, transcription and mRNA splicing with a low level of external contamination from other organelles. We identified a limited number of proteins containing DNA-binding domain which is consistent with the lack of diversity of these proteins in dinoflagellate transcriptomes. However, the number of proteins containing RNA-binding domain was unexpectedly high and this supports the potential roles for this type of protein in mediating circadian gene expression and in chromatin organization. We also identified a number of proteins involved in chromosome condensation and cell-cycle progression as well as a single histone protein (H4). Our results provide the first detailed look at the nuclear proteins associated with the unusual chromatin structure of dinoflagellate nuclei and provides important insights into the biochemical basis of its structure and function.

Keywords: Dinoflagellate, nucleic acid binding protein, chromatin, proteomics

## 4.2. Introduction

Dinoflagellates are a diverse group of unicellular protists, mostly found in marine environments. They play a major role in the global carbon cycle since about half the 2000 known species are photosynthetic [1], and the role of some of them as coral symbionts is vital for maintaining marine biodiversity. They also possess some of the most unusual molecular characteristics of all eukaryotes. A distinctive feature of dinoflagellates is the presence of a highly unusual nucleus termed the "dinokaryon". The defining characteristics of the dinokaryon are the permanently condensed chromosomes, which are packed in a liquid crystalline state and stabilized by RNA and metal cations [366, 367] and by a very low protein–to-DNA ratio (1:10 vs 1:1 in other eukaryotes) [48]. In addition, they are characterized by the absence of histones and nucleosomes [50, 52], the presence of a subset of genes organized in tandem [44, 62] and a closed mitosis, where the nuclei membrane stays intact throughout cell division. Some dinoflagellates also harbor a gigantic genome which can exceed 200 pg per haploid cell [41] which can also contain a large proportion of methylated nucleotides [368]. These unusual characteristics, coupled with a lack of stable genetic transformation methods, have hampered the molecular studies of dinoflagellate over the last few decades, leaving such basic processes as gene regulation and cell cycle control mostly unexplored.

Only a few nuclear proteins have been identified and studied at the biochemical level in dinoflagellate. Canonical histones are replaced by histone-like proteins (HLPs) [54, 57] which are probably inherited from proteobacteria [369]. They are located on external loops of de-condensed chromatin at the periphery of the chromosomes, where transcription is thought to occur [46]. A second type of chromosomal protein, the Dinoflagellate/viral nucleoprotein (DVNP), has recently been identified in *Hematodinium sp.*, a basal dinoflagellate, and is likely derived from a marine DNA virus. Phylogenetic analyses have correlated the acquisition of DVNPs with the loss of the nucleosome, which predates the acquisition of HLPs [12, 370]. Curiously, although nucleosomes are absents, sequences encoding all the canonical histones and many histone modifying enzymes are found in the transcriptome of multiple dinoflagellates [49, 50, 371], but their detection at the protein level has proven problematic

[12, 50]. Other nuclear proteins studied include the TATA-box binding protein-like factor (TLF), a dinoflagellate specific TATA-box binding protein that binds to a TTTT box instead of a TATA box [68], and three proteins able to modulate the global transcription rate when added to isolated nuclei from *Crypthecodinium cohnii* [77].

Considering the unusual features stated above, it is thus not surprising that transcriptional regulation in dinoflagellates is poorly understood. The recent observation that the mRNA abundance in *Lingulodinium* is basically constant throughout a 24h cycle [361], coupled with long median half-life for mRNA in the dinoflagellate *Karenia brevis* [90], raised interesting questions about the extent and the mode of transcriptional regulation in dinoflagellates. Analysis of the transcriptome of two dinoflagellates revealed a paucity of TF, with cold-shock domain (CSD) containing proteins as the most abundant predicted TF [61, 69]. Characterization of two of those proteins in *Lingulodinum polyedrum* showed that they can bind both DNA and RNA and that this binding is unlikely to be sequence-specific, casting doubt on their role as *bone fide* transcriptional regulators. The low number and diversity of potential transcriptional regulator may also suggest a more prominent role of translational and post-translational regulation for controlling gene expression than what is found in other eukaryotes.

Here we performed a CHromatin Enrichment for Proteomics (CHEP) procedure followed by mass spectrometry (MS) sequencing of the enriched proteins [372] to explore the protein components of chromatin in *Lingulodinium polyedrum*, an organism widely studied for its circadian biology [226]. We identified only a few potential transcription factors and, surprisingly, some peptides corresponding to a histone H4. We also uncovered a large amount of proteins having a RNA binding domain, supporting the view that RNA is playing an important role in controlling chromatin structure in dinoflagellates.

## 4.3. Methods

### 4.3.1. Algal culture

Unialgal clonal cultures of the dinoflagellate *Lingulodinium polyedrum* (strain CCMP1936, formerly known as *Gonyaulax polyedra*) were obtained from the Provasoli-Guillard National Center for Marine Algae and Microbiota (Boothbay Harbor, Maine, USA). Cells were grown in a modified seawater medium (f/2 minus silica) [362] at a constant temperature of 18 ± 1°C and under a 12h:12h light:dark (LD) cycle supplied by cool fluorescent lighting at an intensity of 50 µmol photon $m^{-2}$ $s^{-2}$. Cells were grown until late log phase (density of 12000 to 14000 cells $ml^{-1}$) before harvesting and fixation.

### 4.3.2. Purification of nuclei and protein extraction

The protocol for cross-linking and purification of protein-DNA complexes was adapted from [372]. Additional purification steps required to remove the theca and chloroplast fragments were inspired from a plant nuclei extraction protocol [373]. Typically, 3 L of culture were harvested by filtration on a Whatman 541 filter paper and resuspended in 1/50th volume of f/2 medium. Fixation was performed by adding fresh formaldehyde to a final concentration of 1% v/v followed by 10 min of end over end mixing. Formaldehyde quenching was achieved by addition of glycine to a final concentration of 250 mM for 10 min with end-over-end mixing. Cells were harvested again by filtration on a Whatman 541 filter paper, washed with fresh sea water and resuspended in cold lysis buffer (0.25 M sucrose, 0.1 M Tris-HCl pH 7.5, 5 mM $MgCl_2$, 5 mM $CaCl_2$, 1 mM DTT and 1 mM PMSF). Cells were physically disrupted in 1 ml aliquots using a mini bead beater (Biospec) at maximum power for 2 min. The lysed cells were filtered on a 25 µm nylon filter to remove unbroken cells and large thecal fragments. The filtered lysate was centrifuged at 200g for 2 min and the pellet containing the nuclei was then washed twice with lysis buffer supplemented with 0.25% Triton X-100. The nuclei were further purified by layering on a 2 step gradient consisting of 60% Percoll equilibrated in lysis buffer placed on top of a 2.5 M sucrose cushion and centrifuged for 30 min at 1000g. Nuclei were retrieved at the interface between Percoll and

sucrose and washed once with 3 volumes of lysis buffer. Stringent washes to remove non-covalently bound proteins were performed by resuspending the purified nuclei in 300 µl of SDS buffer (50 mM Tris-HCl pH 7.5, 10 mM EDTA, 4% SDS (w/v) and 1 mM DTT) for 10 min with frequent inversion followed by the addition of 3 volumes of urea buffer (10 mM Tris-HCl pH 7,5, 1 mM EDTA and 8 M urea). Nuclei were centrifuged at 16,000g for 20 min and the washing solution was removed and replaced with 1.2 ml of SDS buffer to remove urea. Nuclei were again centrifuged for 20 min at 16,000g and finally washed by the same procedure with the elution buffer (10 mM Tris-HCl pH 7.5, 1 mM EDTA, 25 mM NaCl and 1 mM DTT). Four hundred µl of fresh elution buffer were added to the nuclei followed by 20 min of sonication with a Beckman sonicator to break the chromatin. Protein-DNA cross-linking was reversed by placing the sonicated chromatin at 95 °C for 30 min with occasional mixing followed by centrifugation at 16,000g for 20 min. Supernatants containing the eluted chromatin-interacting proteins were concentrated four to five fold by using a microcon centrifugal filter (10 kDa molecular weight cut-off, EMD Milipore) and analysed by Western blotting to assess potential contamination.

### 4.3.3. Western blotting

Biochemical purity of the purified DNA-protein complexes was assessed by Western blotting using antibodies generated in rabbit against *L. polyedrum* type II RubisCO [23] and Luciferin binding protein (LBP). Briefly, protein fractions from different purification steps were electrophoresed on a 12.5 % SDS-PAGE and transferred to a nitrocellulose membrane using a BioRad wet transfer cell. Ponceau staining was used to assess transfer quality and membranes were blocked for 1 h in a Tris-buffered saline solution (20 mM Tris pH 7.6, 150 mM NaCl) supplemented with 0.05% Triton X-100 and 5% w/v non-fat dry milk (TBSTM) at room temperature under gentle agitation. Primary antibodies were added at 1:4000 for α-RubisCO and 1:2000 for α-LBP in fresh TBSTM buffer and kept under light agitation at 4 °C overnight. After three washes of 5 min with fresh TBSTM, horseradish peroxidase conjugated goat anti-rabbit (Sigma) was added at 1:20000 in fresh TBST and kept under gentle agitation for 1 h at room temperature. Three other washes of 5 min were done with TBST followed by a

chemiluminescent detection using HyGLO (Denville Scientific, Saint-Laurent, QC, Canada) and imaged using a GE ImageQuant LAS 4000.

### 4.3.4. Microscopy

Physical integrity of purified nuclei was assessed by light and fluorescence microscopy. Aliquots of samples were taken at different steps of the purification, stained by adding DAPI to a final concentration of 0.1 ng/µl and observed on a Zeiss Axio Imager M1 fluorescence microscope equipped with a Zeiss AxioCam HRc color digital camera.

### 4.3.5. Mass spectrometry

Twenty micrograms of protein from each of the biological triplicates from each of three different times over the LD cycle (LD6, LD18 and LD23) were precipitated by addition of 4 volumes of cold 100% acetone, kept overnight at -20 °C and centrifuged for 20 min at 16,000g. Protein pellets were washed twice with 500 µl of cold 100% acetone to remove residual SDS from the last purification step. Those protein pellets were used for trypsin digestion and LC-MS/MS analysis at the proteomic facility of l'Institut de Recherche en Immunologie et en Cancérologie (IRIC, Université de Montréal, Canada) as described previously [304]. Briefly, protein pellets were resuspended in 50 µl of 50 mM ammonium bicarbonate and 5 mM TCEP (tris (2-carboxyethyl) phosphine) and incubated at 37 °C for 30 min. This was followed by the addition of 30 µl of 55 mM chloroacetamide and incubation for an additional 30 min at 37 °C. Samples were digested overnight at 37 °C with 1 µg of trypsin, dried in a Speed-Vac and resolubilized in 50 µl of 5% acetonitrile/ 0.2% formic acid. Twenty microliters of each sample was injected on a C-18 precolumn (0,3 mm i.d. x 5 mm) and peptides were separated on a C-18 analytical column (150 µm i.d. x 100 mm) using an Eksigent nanoLC-2D system. A 56-min gradient from A to B (10–60% (A: formic acid 0.2 %, B: acetonitrile/0.2% formic acid)) was used to elute peptides with a flow rate set at 600 nanoliter/min. The LC system was coupled to a LTQ-Orbitrap Velos mass spectrometer (Thermo Fisher). Each full MS spectrum was followed by 12 MS/MS spectra (thirteen scan

events), and the 12 most abundant multiply charged ions selected for MS/MS sequencing. Tandem MS experiments were performed using collision-induced dissociation in the linear ion trap. Spectra obtained from MS were mapped against the six-frame translation of a *L. polyedrum* transcriptome containing 114,492 sequences (Bioproject accession number PRJNA69549; sequence accession numbers GAPB01000001 to GABP01114492 [304]) using PEAKS Studio version 7.0 (Bioinformatics Solutions Inc.). Amino acid modifications such as deamination, carbamidomethyl (C) and oxidation (M) were included in the analysis. The precursor ion and fragment tolerances were set at 10 ppm and 0.5 Da, respectively. Mass spectrometry results were visualized and further analyzed using the Scaffold 4 software (Proteome Software Inc.). Proteins for which peptides (identified at a 1% False Discovery Rate (FDR)) were present in at least two out of three replicates for each time point and for which $\geq 2$ peptides were present in one of the triplicate samples were chosen for further analysis.

### 4.3.6. Data analysis

Sequences from the transcriptome corresponding to the peptides as described above were used in BLASTx searches [374] conducted from a local BLAST instance on a Calcul Québec high performance computing cluster against the nr database (release 75) to identify the different proteins present in the sample. Further annotation by Gene Ontology (GO) terms [375] were executed using the Blast2GO software [376]. A manual classification of the dataset was performed to categorize the protein set according to the BLAST results, domains identified in the NCBI Conserved Domain Database [377] and GO classifications. A GO enrichment analysis was conducted using Blast2GO by comparing the chromatin enriched protein to the previously mentioned transcriptome. Enriched and depleted GO terms were identified by applying Fisher's exact test with a p-value threshold of $< 0,01$ and further trimmed to remove redundant terms using GO Trimming [378]. Contamination from mitochondria, chloroplasts, cell membranes and scintillons were identified using BLAST results, GO terms and complementary BLAST searches against UniProtKB/SwissProt-reviewed proteins for those cell subcellular localizations [379]. The Kyoto Encyclopedia of

Genes and Genomes (KEGG) pathways [380] for DNA replication, RNA polymerase and the spliceosome were used to identify components of these process using the dinoflagellate *Symbiodinium minutum* as a reference with additional data from previous analysis in *Lingulodinium* [61, 66].

To estimate protein expression differences between LD6 and LD18, the abundance of each protein was evaluated by doing the addition of the peak intensity of each peptide (above a peak intensity threshold of 10000) identified for a particular protein using ProteoProfile (http://www.thibault.iric.ca/proteoprofile/files/TechnicalGuide.pdf). Protein intensities for both set of triplicates were then compared with EdgeR [307] and significant changes between time were identified using a p-value of 0.05. In order to represent graphically the log10 intensity of all the proteins analyzed, we added one spectrum count to the triplicates average of all proteins before the log calculation.

## 4.4. Results

### 4.4.1. Validation of the chromatin extraction method

As to identify the proteins present in the dinoflagellate chromatin, we purified and analyzed the proteinsassociated with a chromatin fraction after *in vivo* protein cross-linking (Figure 4.1. A). Different methods for cell lysis were tested, including sonication, French pressure cell, bead beater and grinding with a mortar and pestle in liquid nitrogen, with the bead beater providing the highest and most consistent number of broken cells with intact nuclei. The other methods were either less efficient at breaking cells or resulted in fragmented nuclei not suitable for purification. Contamination of the nuclear fraction by other cellular components was assessed by Western blotting and microscopic observation. For Western blotting, we used antibodies against LBP (localized in scintillons, bioluminescent organelles in the cytoplasm adjacent to the vacuolar membrane) and against RubisCO (localized in the chloroplast) (Figure 4.1. B). Contamination from these two proteins was low after the physical purification of nuclei by step gradient centrifugation and was below the detection level after urea and SDS washes, which yielded the final purified sample. We also tested an antibody raised against PCNA (Proliferating Cell Nuclear Antigen) from the dinoflagellate *Karenia brevis* [39] as a positive control, but this antibody cross-reacted poorly with PCNA from *Lingulodinium* extract. Microscopic observations of the nuclear fraction after density gradient centrifugation show that this fraction contains mainly intact nuclei with a small amount of thecal plates (Figure 4.1. C).

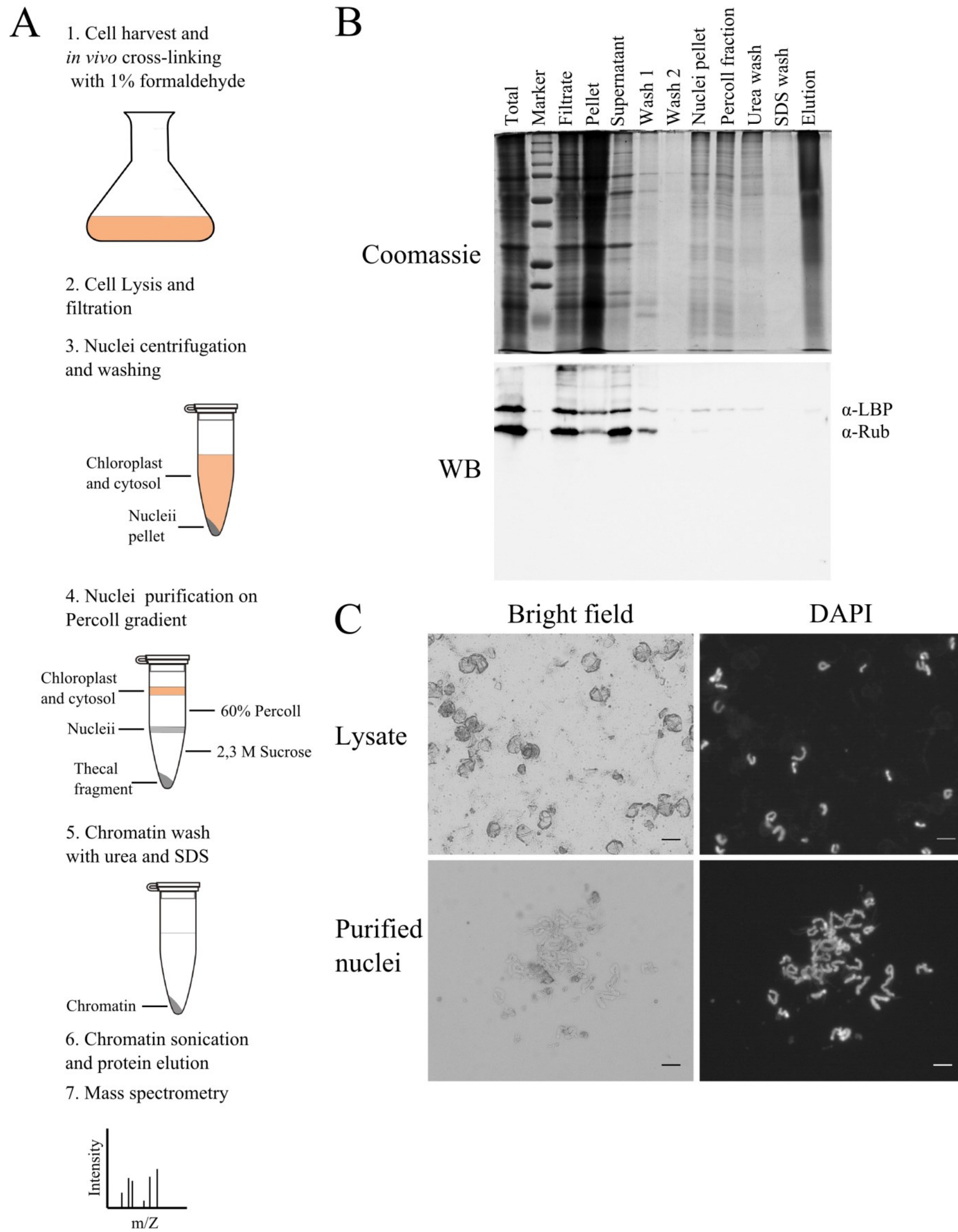



**Figure 4.1. Outline and validation of the chromatin purification procedure**

A. Graphical summary of the chromatin purification procedure from *in vivo* protein crosslinking to nuclei purification and protein isolation. B. Typical protein profile (top panel) and western blot against RubisCO and LBP (bottom panel) obtained from the purification procedure. C. Microscopic examination of cell lysate (top panel) and nuclei after gradient purification (bottom panel) in bright field (left side) and in fluorescence (right side) after DAPI staining. The typical C-shaped nucleus of *Lingulodinium* is conserved after cell lysis. Scale bar: 40 μM

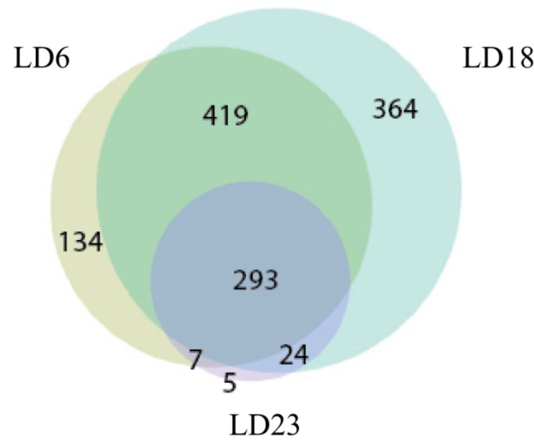
### 4.4.2. MS results and general protein classification

We extracted chromatin-associated proteins at three different time points (LD6, midday; LD18, midnight; and LD23, one hour before dawn) to obtain a global portrait of the DNA interactome over the *Lingulodinium* cell cycle. At LD6, all cells are in interphase, whereas at LD18 and LD23 roughly one third to one fourth of the cell population are either in S phase or in M phase, respectively [260, 292]. Three biological replicates for each time were analyzed, yielding a total of 28403 identified peptides coming from 2714 proteins (Tableau 4.1.). To enhance confidence in the identified proteins for downstream analysis, we required proteins to be identified by two or more peptides in at least one of the three replicates (1727 proteins) and for the peptides to be present in at least two of the three replicates (1245 proteins) (Figure. 4.2. A, Tableau 4.1.). A classification of the final protein dataset using BLAST results and GO terms with additional domain identification when needed (Figure 4.2. B, Tableau 4.S1.) shows that around 35% of the proteins analyzed have a function related to chromatin and the nucleus, and only 97 were identified as contaminants from other cellular compartments such as the endomembrane systems, mitochondria, chloroplasts and scintillons. GO terms corresponding to nuclear localization, functions, and biological processes are also significantly enriched in our sample (Figure 4.1. and 4.S1., Tableau 4.S2.).

Tableau 4.1. Number of peptides and protein identified in nine chromatin-enriched samples

| Samples | Identified peptides | Identified proteins |
|---|---|---|
| LD6-1 | 4110 | 1036 |
| LD6-2 | 4684 | 1053 |
| LD6-3 | 2072 | 567 |
| LD18-1 | 6226 | 1414 |
| LD18-2 | 4867 | 1084 |
| LD18-3 | 2795 | 837 |
| LD23-1 | 206 | 112 |
| LD23-2 | 1058 | 365 |
| LD23-3 | 2385 | 613 |
| Total number of identified peptides | 28403 | |
| Total number of identified proteins | | 1727 |
| Total number of proteins in at least 2 replicates | | 1245 |
| Organelle and membrane contamination | 1900 | 97 |

A




B

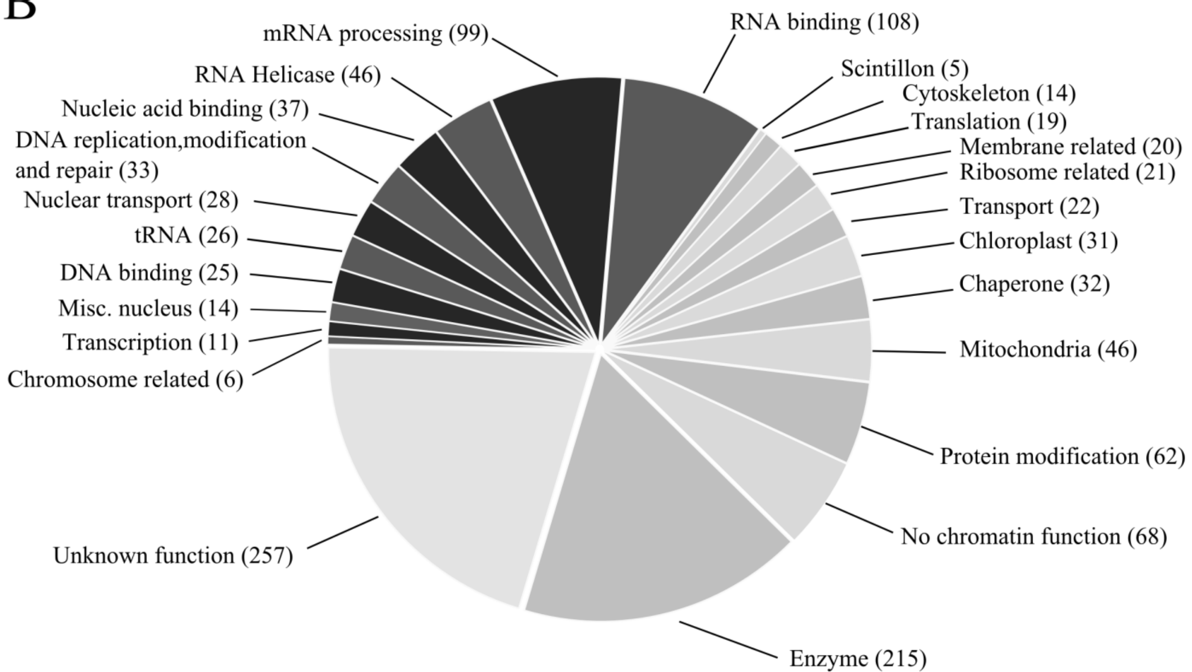



**Figure 4.2. General classification of proteins retrieved in chromatin sample**

A. Venn diagram of the protein distribution throughout the three sampled times (LD6, LD18 and LD23). The cut off requirement for 2 peptides per protein in two of three samples has contributed to the lower number of proteins estimated at LD23. B. Functional classification of the 1245 proteins identified. Functions related to nucleus are in the darker shades of grey.

### 4.4.3. DNA replication, transcription and mRNA processing

To assess the coverage of the nuclear protein complement, we searched for proteins implicated in DNA replication, transcription and mRNA splicing that were previously identified in the *Lingulodinium* transcriptome [61] (Tableau 4.2.). For DNA replication, we were able to retrieve only the large subunit of DNA polymerase delta. However, 9 out of 16 components associated with the DNA polymerase were identified, including four of the six Mini Chromosome Maintenance (MCM) complex proteins. For transcription, we identified the two largest subunits of RNA polymerase II and the B9 specific component along with TFIIB and one of the six proteins from the TFIIH complex. The TATA-box binding protein like factor (TLF), a dinoflagellate specific transcription factor binding to a TTTT sequence [68] and the only known member of the TFIID complex [61], was not found in our sample. For RNA splicing, a nearly complete spliceosome was recovered (Tableau 4.2.). Finally, we found several proteins implicated in a wide array of DNA repair mechanisms (Tableau 4.3.). Overall, 50% of the proteins expected to be associated to chromatin were recovered by our proteomic analysis.

Tableau 4.2. Proteins involved in DNA replication, transcription and RNA splicing found in chromatin-enriched samples

| | | Components | Lingulodinium transcriptome | Chromatin-enriched samples |
|---|---|---|---|---|
| DNA replication | DNA polymerase | Alpha | 4 | 0 |
| | | Delta | 2 | 1 |
| | | Epsilon | 2 | 0 |
| | Other components | MCM complex | 6 | 4 |
| | | RPA | 1 | 1 |
| | | Clamp | 4 | 2 |
| | | Other | 5 | 3 |
| Transcription | RNA Pol II | Core | 4 | 2 |
| | | Specific | 3 | 1 |
| | | Common | 5 | 0 |
| | General transcription factor | TFIID and TLF | 2 | 0 |
| | | Holo-TFIIH | 6 | 1 |
| Splicing and mRNA processing | Splicing components | U1 | 5 | 5 |
| | | U2 | 8 | 6 |
| | | U4/U6 | 7 | 4 |
| | | U5 | 7 | 6 |
| | | U4/U5/U6 | 2 | 2 |
| | | PRP19-related | 12 | 4 |
| | | TREX | 3 | 3 |
| | | Common | 3 | 2 |
| | 3' end processing | | 11 | 4 |

Lingulodinium transcriptome: Total number of sequences found for each of those components in the *Lingulodinium* transcriptome.
MCM complex: mini-chromosome maintenance protein complex; RPA: replication protein A; Clamp: PCNA and Replication Factor C; TLF: TATA-box binding protein-like factor; PRP19: pre-mRNA processing factor; TREX: transcription/export complex

Tableau 4.3. Proteins involved in DNA repair and processing

| Process | Number of sequences found | Protein identities |
|---|---|---|
| Base excision repair | 4 | poly [ADP-ribose] polymerase |
| Damage checkpoint | 3 | RAD24 |
| Double strand break repair | 1 | RAD50 |
| Non-homologous end-joining | 2 | DNA ligase |
| Nucleotide excision repair | 2 | XPG domain N-terminal, RPA |
| Other | 2 | Endonuclease, DNA helicase |
| Topoisomerase | 11 | Multiple DNA topoisomerase |
| UV excision repair | 2 | RAD23 |

RPA: replication protein A; XPG: Xeroderma Pigmentosum Complementation Group G

### 4.4.4. Nucleic acid binding proteins

Chromosome organization in dinoflagellates is unique among eukaryotes as DNA is not compacted by histones into nucleosomes. Instead, a liquid crystal arrangement stabilized by RNA and metal cations has been proposed [381]. Although transcriptomes of dinoflagellates contain sequences encoding classical histones as well as many histone modifying enzymes, the levels of histone proteins are typically below current levels of detection. Here, we were able to identify two histone H4 peptides in one of our replicates (LD6-2), as well as a histone deacetylase, which supports the view that histones may have a functional role in the dinoflagellate nucleus. Canonical histones are thought to be replaced, in the periphery of chromosomes where transcription is active, by histone-like proteins and DVNP (Dinoflagellate/viral nucleoprotein), both of which have been placed in the nucleus by biochemical studies. We found both of these proteins in our samples, with 60 peptides identified for HLPs and a total of 15 peptides for multiple isoforms of DVNPs. Finally, we found two proteins, a Condensin 2 subunit and a protein harboring multiple repeats of the Regulator of Chromosome Condensation domain (RCC1) which might be implicated in chromosome condensation.

Transcripts encoding DNA-binding proteins are scarce in transcriptomes of dinoflagellates, and there is generally a low diversity of DNA-binding domains [61, 69]. Here,

we identified 26 proteins that contain a DNA-binding domain and/or match a DNA-binding protein based on BLAST analysis (Tableau 4.4.). As expected, we found a large number of cold-shock domain (CSD) proteins, which is the most abundant DNA-binding domain present in the transcriptome. In particular, we retrieved LpCSP1, a *Lingulodinium* CSD-containing protein that has been previously characterized biochemically [382]. We also uncovered 9 proteins containing a High-mobility group box (HMG-box) domain, which is the second most abundant DNA-binding domain in the *Lingulodinium* transcriptome [61]. Finally, we found seven proteins with similarity to a known bacterial transcriptional regulator of the NagC family (N-acetylglucosamine repressor) but for which the DNA-binding domain, a helix-turn-helix normally located at the N-terminal end, could not be identified, probably because the 5' end of the transcript appears to be missing in our transcriptome.

Tableau 4.4. Nucleic acid binding domain present in the chromatin-associated-enriched samples

| | **Domain** | **Number of sequences** |
|---|---|---|
| Nucleic Acid Binding | Cold shock | 25 |
| | Cold shock/ RRM | 6 |
| | Other | 2 |
| | Unknown | 2 |
| | Cold shock/ KH | 1 |
| | KH/ Forkhead | 1 |
| | ***Total*** | ***37*** |
| DNA Binding | High Mobility Group box | 9 |
| | Similar to NagC (bacterial TF) | 7 |
| | Other (AP2, Leucine Zipper, DEK-C) | 4 |
| | Zinc finger | 3 |
| | SAP | 3 |
| | ***Total*** | ***26*** |
| RNA binding | RRM | 44 |
| | KH | 28 |
| | Ribosomal S1-like | 11 |
| | Unknown (similar to diverse RNA binding protein) | 11 |
| | Zinc finger CCCH | 4 |
| | KH and RRM | 4 |
| | LA motif and Zn finger CCCH | 2 |
| | Zinc finger and RRM | 2 |
| | Pumilio | 1 |
| | PWI | 1 |
| | ***Total*** | ***108*** |

KH: K Homology domain; NagC: N-acetylglucosamine repressor; RRM: RNA Recognition Motif; SAP: SAF-A/B, Acinus and PIAS

We also found that RNA-binding proteins (RBPs), RNA helicases and proteins involved in splicing and mRNA processing to be highly represented, with nearly 250 of these proteins identified. Indeed, RBPs were the most abundant type of nucleic acid binding protein present (Tableau 4.4.), and represent a far greater relative abundance than the RBPs found in human interphase chromatin (about 5% [383]). The vast majority of the RBPs we identified

harbor either a RRM and/or a KH domain. Among the RRM domain containing proteins, we identified three proteins similar to the *Arabidopsis thaliana* Flowering time control protein FCA, which is involved in various co-transcriptional processes and RNA maturation [384]. Among the proteins with a KH domain are found 12 proteins with various degrees of homology to diverse isoforms of the far upstream binding element protein, which binds both RNA and ssDNA and can regulate transcription and translation [385]. There are also 11 sequences with a Ribosomal S1-like domain, a large family of nucleic acid binding domain that includes the cold-shock domain. Lastly, we recovered a large number of proteins classified as RNA Helicases due to the presence of DEAD-box domains. The surprising abundance of the RBPs lends credence to the view that RNA is involved in stabilising the structure of the chromatin in dinoflagellates [367].

### 4.4.5. Other proteins of interest

A large number of enzymes and proteins involved in post-translational modifications were also recovered (Figure 4.2. B) and these may have potential roles in chromatin biology. In particular, we note the presence of multiple peptidyl-prolyl cis-trans isomerases (PPIases), which are known to alter transcription rates by modifying transcription factors and the C-terminal domain of RNA polymerase II [386]. Multiple proteins corresponding to S-adenosylmethionine synthetase, an enzyme involved in DNA methylation, were also identified [387] and are of interest as a high proportion of dinoflagellate nucleotides are in a methylated form [368]. Some glycolytic enzymes for which potential roles in the nucleus have been demonstrated were also found in our dataset. For example, we identified an enolase that plays a role in DNA binding and transcription regulation of a large group of genes in *Toxoplasma gondii*, a close relative to the dinoflagellates [388]. Other glycolytic enzymes found include the cytosolic isoform of glyceraldehyde 3-phosphate dehydrogenase [389] and pyruvate kinase [390], both of which are potentially localized in the nucleus of other organisms. We also note the presence of multiple types of HSP70 proteins, which along with their role as molecular chaperones, have been implicated in nuclear transport and in the protection of DNA from ssDNA breaks [391, 392].

### 4.4.6. Differential expression between midday and midnight

We were also interested in protein components that might be associated with chromatin at a specific time of the day. As shown in Tableau 4.1. and Figure 4.2. A, the sequencing of the LD23 samples yielded a much lower number of identified peptides and proteins when compared either to LD6 or LD18, despite a similar number of spectra sequenced by MS. We thus removed this sample from our quantitative analysis so as to not bias the analysis toward technical disparities. We compared the peak intensity of the proteins from the LD6 and LD18 samples using EdgeR to identify proteins for which abundance was significantly different (at a p-value < 0,05) between the two times (Figure 4.3. A and B). Five proteins overrepresented at LD6 and 29 proteins overrepresented at LD18 were identified (Tableau 4.S3.). Among these, 4 are RNA-binding proteins and 2 possessed a nucleic acid-binding domain.

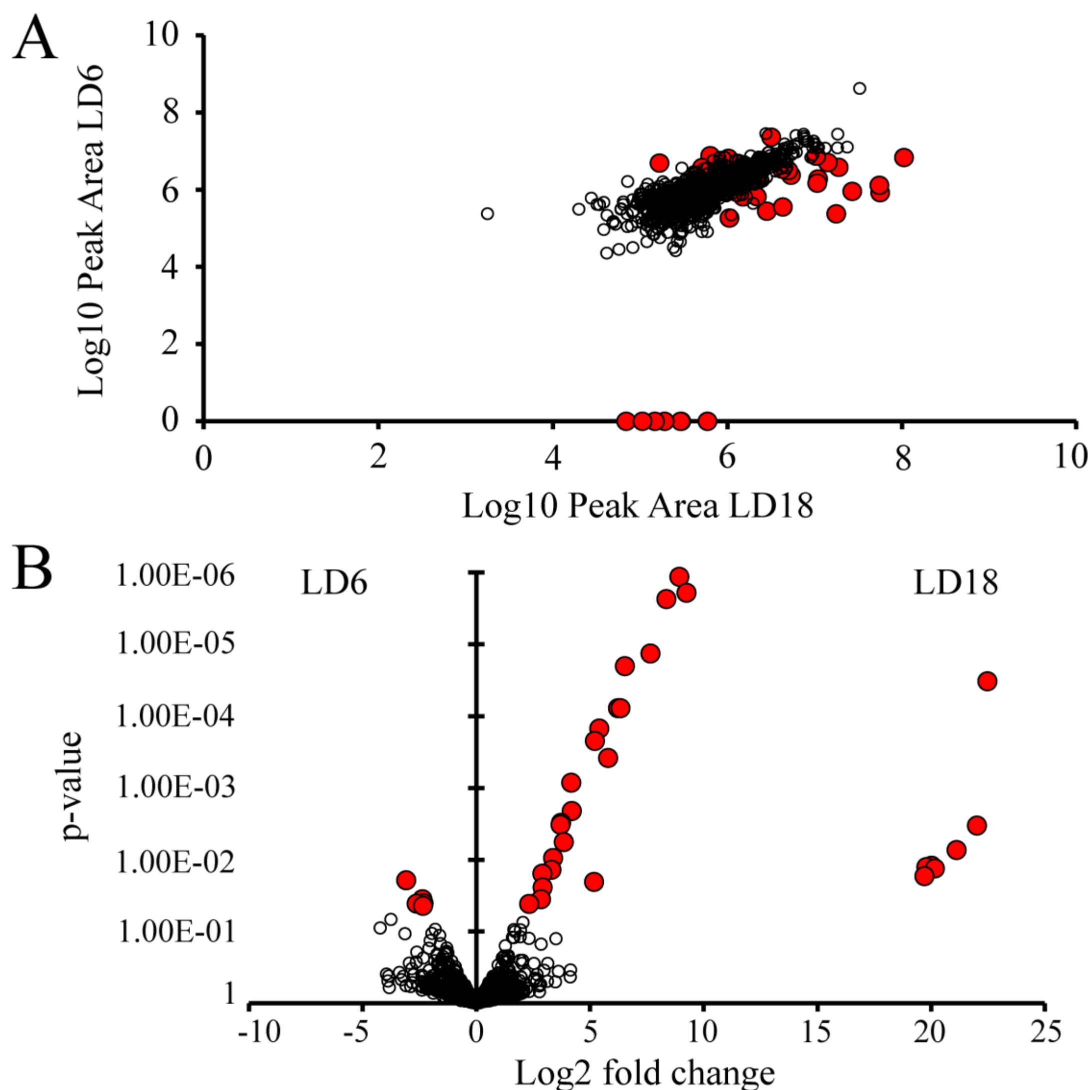



**Figure 4.3. Differential protein abundance between LD6 and LD18**

A. and B. Average log10 peak area (A) and log2 fold change (B) between LD6 and LD18. White circles: no statistical difference; red circles: statistically different at a p-value of < 0,05. Proteins not detected at one particular time (as in Figure 2) may not be considered statistically significant when their spectral counts are too low or when variation between replicates is too high.

## 4.5. Discussion

The goal of this work was to establish a catalog of chromatin-associated proteins in a dinoflagellate to help understand their unique chromosome organization and gain insights about their transcriptional and translational regulation. We adapted a previously described chromatin purification protocol (Figure 4.1. A) [372], which was particularly effective for dinoflagellates and their permanently condensed chromosomes were purified with a high level of physical integrity (Figure 4.1. C). In addition, stringent urea and SDS washes successfully reduce contamination of the chromatin by proteins from other organelles (Figure 4.1. B, Tableau 4.1.). The procedure was validated by the fact that many of the proteins expected to be involved in basic processes related to DNA replication, transcription and mRNA splicing were found (Tableau 4.2.). Analysis of the 1245 proteins retrieved in our samples shows that approximately 35% of proteins are assigned to categories related to chromatin, such as mRNA processing, RNA and DNA binding, transcription and DNA replication. This is a similar proportion to what was previously reported for similar experiments using animal cells [383, 393] (Figure 4.2. B). Concomitantly, GO enrichment analysis shows that nuclear proteins are the most highly enriched class of proteins in our samples (Figure 4.S1.). Thus, these results confirm that the adapted CHEP procedure performed here is suitable for the enrichment of proteins bound to dinoflagellate chromatin.

The absence of histones and nucleosomes has long been recognized as one of the most distinctive features of dinoflagellates. However, recent transcriptomic sequencing yielded a major surprise, as a full set of core histones (although H1 is absent in some species) and multiple histone modifying enzymes were identified in dinoflagellates [371]. Here, we identified a histone deacetylase and a histone H4 in our samples, which, combined with a previous detection of H2A from acid extracted samples [12], confirm that dinoflagellates are also expressing histones at the protein level. Immuno-localization of histones in the chromatin at different times will be needed to provide a better idea of their role in dinoflagellates.

We found a limited number of proteins containing DNA-binding domains (26 proteins; 2,1% of total protein) in our experiment (Figure 4.2. B, Tableau 4.4.), which is a lower proportion than in similar studies done in animal cells, for which between six and 14% of all

proteins fall in this category [383, 393, 394]. This lower proportion should not come as a surprise as the amount and diversity of DNA binding domain in dinoflagellates is also highly reduced at the transcriptome level. We indeed found a relatively good representation of the DNA-binding domain present in the *Lingulodinium* transcriptome [61] in our CHEP experiment (Tableau 4.4.). Of interest, we found seven proteins with varying degrees of similarity with NagC, a bacterial transcription factor of the ROK family implicated in the regulation of the N-Acetylglucosamine-specific transporter (NagE) [395]. All seven coding sequences have a GC content above 60%, typical of the *Lingulodinium* transcriptome [61]. Thus, while these genes may have been acquired by lateral transfer from bacteria, a process known to occur frequently in marine microbial eukaryotes [396, 397], this is unlikely to have been a recent occurrence as most marine bacterial genomes have low GC content. Furthermore, we found multiple CSD proteins in the chromatin-enriched fraction. This domain is the most abundant DNA-binding domain in dinoflagellates [61, 69], yet a role in transcription has yet to be established. We have previously shown that two *Lingulodinium* CSD proteins bind ssDNA and RNA in a non-specific manner [382], suggesting that, by binding RNA, they may play a role in determining chromatin structure. However, it is still possible they play a role in transcription, either by destabilizing the double helix to promote transcription or by melting DNA duplexes, as shown for the frog CSD protein YB1 [398, 399].

Surprisingly, the ratio of RNA to DNA-binding proteins in our samples is over tenfold greater than the ratio recently reported for the wheat nuclear proteome [400]. One possible explanation for the high abundance of RBPs is that they are involved in regulating gene expression. When nascent mRNAs are transcribed, they are simultaneously coated with a wide array of proteins that affect multiple aspects of their biology including their ability to be translated, their localization within the cell and their stability [401]. As multiple processes, and especially circadian rhythms regulation in *Lingulodinium* seems to be at the translational and post-translational steps rather than at the transcriptional step [226], important regulators for circadian control of translation might interact with nascent mRNAs. Six of the RBPs identified here were previously identified among proteins potentially phosphorylated by Casein Kinase 2 [95], an important and widely conserved kinase for regulation of circadian rhythms [316]. In addition to RBPs, we found a large number of RNA helicases, mostly from the DEAD/DEAH-

box family. Although these proteins are classified as helicases, many of them play other roles in RNA biology in the nucleus [402]. For example, a subset of DEAD box RNA helicases in animals is able to co-activate or co-repress transcription of multiple genes by their association with components of the transcriptional machinery [403]. Other functions include rRNA biogenesis [404, 405] and mRNA export from the nucleus [406]. A nuclear localized DEAD box RNA helicase, PRD-1, has recently been linked to the FRQ-less circadian oscillator of the fungi *Neurospora crassa* [407, 408]. Although no transcripts from *Lingulodinium* share homology to PRD-1, a potential role for DEAD box RNA helicases in circadian biology in this organism is intriguing. In this regard, it is interesting that one RNA helicase was more abundant at LD6 whereas three different RNA-binding proteins were more abundant at LD18 (Figure 4.3., Tableau 4.S3.). Another potential role for RBPs in the dinoflagellate nucleus may be to mediate the role of RNA in chromatin structure. RNA was implicated in the organisation of the chromosomes by studies showing RNase A, but not proteinase K, produced chromatin stretching and unwinding [409]. Furthermore, some of the RNA was observed to be distributed regularly with respect to the organisation of the chromatin by electron microscopy [410].

The *Lingulodinium* chromatin also contains proteins related to cell-cycle progression, as several proteins related to DNA replication and repair were identified (Tableau 4.2. and Tableau 4.3.). Among these, we note the presence of RAD24, previously identified as a highly phosphorylated protein at LD18 in *Lingulodinium* [115], two Replication Factor C proteins, which act as a clamp helping to load PCNA on DNA [411], and Replication Factor A, which binds ssDNA to prevent the formation of secondary structure and participates in DNA repair [412]. Both replication factor proteins showed a small increase in abundance at the moment of DNA replication in *Karenia brevis* [39]. Unfortunately, the poor spectra identification for the LD23 samples may have hampered the discovery of proteins implicated in M-phase (Tableau 4.1., Figure 4.2. A). However, as *Lingulodinium* populations double every three to five days and cannot be synchronized, the low proportion of dividing cells in our sample may also have hindered discovery of other cell cycle regulators.

Few studies have attempted to identify proteins present in the unusual nuclei of dinoflagellates [48, 413]. These earlier studies were seminal as they documented the very low

protein:DNA ratio of dinoflagellate chromatin and highlighted the lack of histones and nucleosomes. Here, we provide the first glimpse of the repertoire of chromatin associated proteins in dinoflagellates. Our results confirm a low diversity and abundance of DNA-binding proteins and demonstrate that a large number of RNA-binding proteins could play a role in translational regulation and, in conjunction with structural RNAs, in chromatin organization. Our procedure can be readily adapted to the study of other dinoflagellates in diverse environmental conditions, such as nutrient starvation and response to temperature changes, providing a new avenue to explore transcriptional and translational regulation in dinoflagellates.

## 4.6. Acknowledgments

Proteomics analyses were performed by the Center for Advanced Proteomics Analyses, a Node of the Canadian Genomic Innovation Network that is supported by the Canadian Government through Genome Canada. Computer-intensive analyses were made on the supercomputer Guillimin at McGill University, managed by Calcul Québec and Compute Canada. The operation of this supercomputer is funded by the Canada Foundation for Innovation (CFI), the ministère de l'Économie, de la science et de l'innovation du Québec (MESI) and the Fonds de recherche du Québec - Nature et technologies (FRQNT). Funding of this work was provided by National Science and Engineering Research Council of Canada (NSERC) through an Alexander-Graham-Bell Canada Doctoral Scholarship awarded to M. B. and an NSERC Discovery research grant awarded to D. M. (number 171382-03).

The following Supplementary tables for this chapter can be at the end of this thesis, in the Annexe section

Tableau 4.S1. List of 1245 proteins identified in the chromatin enriched samples with BLAST description, GO terms and manually classified functional characterization

Tableau 4.S2. Gene ontology terms significantly enriched and depleted in the chromatin enriched samples

Tableau 4.S3. Differentially expressed proteins between LD6 and LD18

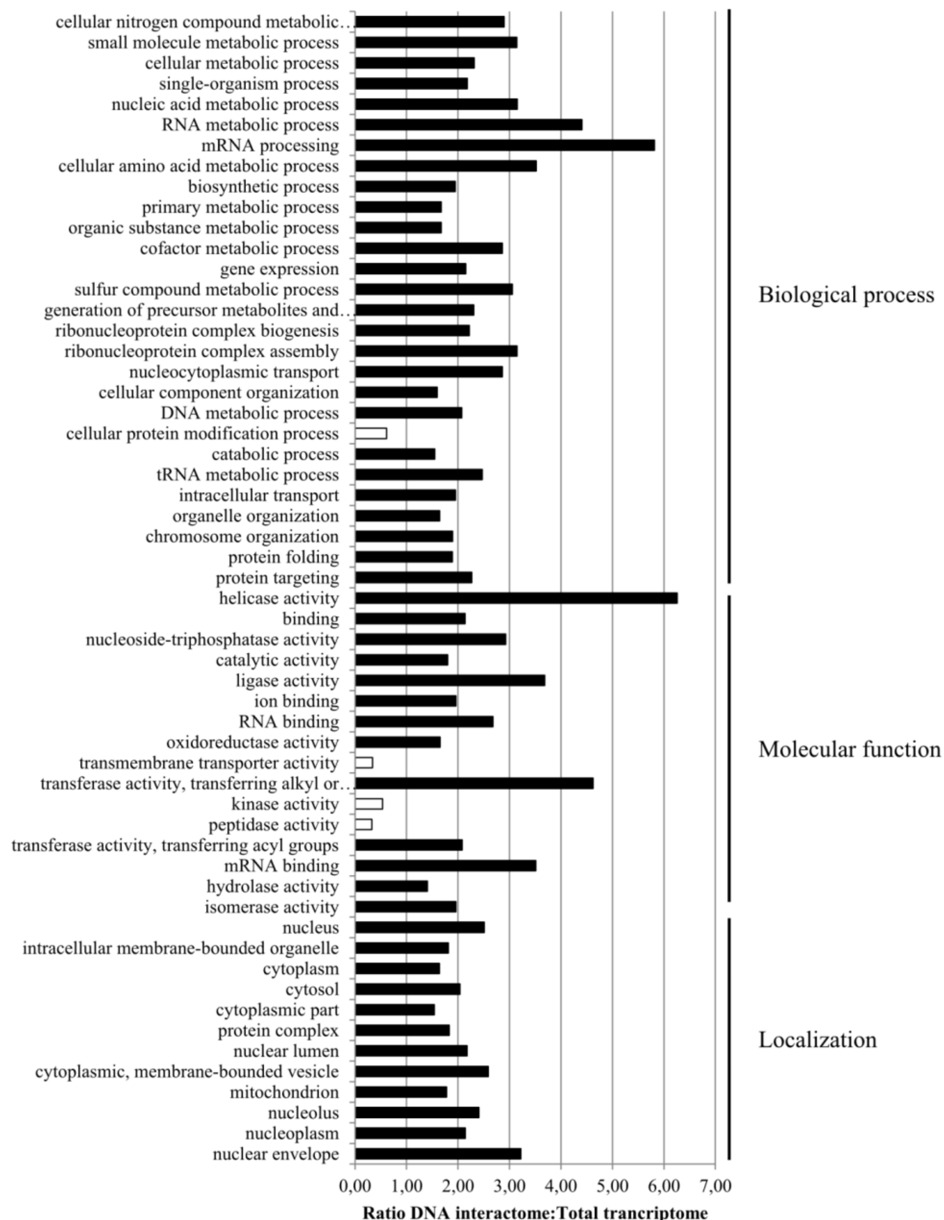



**Figure 4.S1. Gene ontology enrichment analysis**

Gene ontology terms over and under-represented (p-value < 0,05) in the chromatin-enriched sample for term related to Biological process, Molecular function and Cell Localization compared to *Lingulodinium* complete transcriptome. For each category, terms are classified by increasing p-value, with lower p-value at the top. White bar: under-represented terms, Solid bar: over represented terms

# Chapitre 5 : Discussion générale

Bien que faisant partie des eucaryotes, les dinoflagellés ont des caractéristiques très inusuelles quant à plusieurs aspects de leur biologie cellulaire et moléculaire. La présence d'un génome condensé en tout temps et l'absence d'histones permettant d'organiser l'ADN en nucléosome a même généré le terme de mésocaryote, ou noyau intermédiaire situé entre les procaryotes et les eucaryotes, pour décrire ces organismes, avant que les phylogénies moléculaires permettent de les situer adéquatement dans le groupe des alvéolés [13, 52]. Ces caractéristiques, couplées à l'absence de méthodes de transformation stable et de criblage génétique, ont fortement ralenti, pendant les deux dernières décennies, l'exploration des mécanismes de régulation génique. L'avancement des techniques de séquençage à haut-débit permet maintenant d'envisager de contourner ces problèmes et d'obtenir des informations globales sur la biologie de ces organismes. L'objectif général de ma thèse de doctorat était donc d'utiliser des techniques modernes de séquençages d'ARNs et de protéines afin de caractériser divers aspects de la régulation de la transcription chez le dinoflagellé *Lingulodinium polyedrum*.

Le premier objectif était de caractériser les changements d'abondance dans les transcrits d'ARNm au cours d'un cycle jour/nuit. J'ai fait l'hypothèse que, comme la grande majorité des organismes montrent des variations journalières dans l'accumulation d'une plus ou moins grande partie des transcrits, dont ceux de l'horloge, l'étude à large échelle de l'abondance des transcrits chez *Lingulodinium* pourrait permettre d'identifier des gènes directement impliqués dans l'horloge circadienne ainsi que des groupes de gènes régulés au niveau transcriptionnel par cette horloge. De façon surprenante, aucun des transcrits du transcriptome complet de *Lingulodinium* ne montre de variations significatives lorsqu'échantillonné à quatre temps différents lors d'un cycle journalier (Figure 2.1.), suggérant que le mécanisme de génération des rythmes circadiens ne passe pas par l'accumulation et la dégradation de transcrits à des temps précis comme chez les autres eucaryotes. De plus, l'inhibition chimique de la transcription ne permet pas d'abolir les rythmes de photosynthèse et de bioluminescence, et ce sur quelques cycles malgré une diminution de l'amplitude de ces phénomènes (Figure 2.3.). Ces résultats semblent indiquer que le modèle classique de boucle de transcription/traduction n'est pas utilisé chez les dinoflagellés pour générer des rythmes circadiens.

Le deuxième objectif consistait en la caractérisation de deux protéines ayant un CSD, domaine capable de lier les acides nucléiques et présent chez une grande variété d'organismes. Les protéines ayant ce domaine peuvent avoir plusieurs fonctions différentes, dont celle de réguler diverses étapes de la transcription. L'expansion de cette famille de gènes chez les dinoflagellés et la relative pauvreté des autres types de protéines de liaison a l'ADN, tous deux révélé par le séquençage récent à haut-débit des transcriptomes de *Lingulodinum* et de *Symbiodinium*, soulève l'hypothèse que ces protéines pourraient avoir un rôle majeur dans la régulation de l'expression des gènes [61, 69]. Ces protéines forment deux groupes distincts chez les dinoflagellés, dont un qui est proche des CSPs bactériens (Figure 3.2.). J'ai donc testé les propriétés de liaisons aux acides nucléiques de deux de ces protéines qui possèdent un CSD. Bien que celles-ci soient capables de lier l'ADN double brin, cette liaison ne semble pas spécifique à une séquence particulière et est plus faible que la liaison à l'ADN simple brin et à l'ARN (Figure 3.5. et 3.S3.). De plus, ces deux protéines, contrairement aux protéines bactériennes, ne sont pas induites lorsque *Lingulodinium* est mis en condition froide (Figure 3.4.) et ne sont pas capables de complémenter des protéines *cold-shock* chez *E. coli* (Figure 3.3.). Ces données suggèrent que ces protéines ne sont probablement pas des régulateurs transcriptionnels spécifiques à des séquences particulières, mais pourraient participer à certains processus généraux de la transcription.

Finalement, le troisième projet consistait en la caractérisation à large échelle des protéines associées à la chromatine chez *Lingulodinium*. Les caractéristiques inusuelles du génome des dinoflagellés, comme son empaquetage sous forme de cristal liquide stabilisé par des cations métalliques et de l'ARN, et l'absence d'histones classiques en font un modèle intriguant pour l'étude d'une organisation nucléaire alternative. Afin de parvenir à identifier les protéines qui y sont associées, j'ai adapté un protocole permettant de récupérer les protéines faisant partie de la chromatine après réticulation *in vivo* et d'identifier ces protéines par spectrométrie de masse (Figure 4.1.). La validation des données obtenues, notamment par l'identification de plusieurs protéines ayant des domaines de liaisons aux acides nucléiques et la grande quantité de protéines impliqués dans des processus comme la réplication de l'ADN et l'épissage des ARNm, jumelé à un bas niveau de contamination par les protéines d'autres organelles, permet de proposer un catalogue de protéines faisant partie de la chromatine chez

les dinoflagellés. Parmi ces protéines, on peut noter un faible nombre de protéines avec des domaines de liaison à l'ADN, avec le domaine CSD comme le plus abondant. À l'inverse, une grande quantité de protéines avec des domaines RRM, KH et DEAD box ont été retrouvées et pourraient potentiellement être impliquées dans l'organisation de la chromatine par le biais de leur interaction avec l'ARN qui stabilise celle-ci (Tableau 4.4.). De façon surprenante, des peptides correspondant à l'histone H4 ont été identifié, suggérant que ces protéines sont retrouvées dans la chromatine des dinoflagellés, quoiqu'en très faible abondance.

Un thème commun soulevé par les trois chapitres est la faible importance que semble prendre la régulation de la transcription dans le contrôle de l'expression génique chez les dinoflagellés alors que l'abondance des transcrits montre une stabilité remarquable sur un cycle journalier (Chapitre 2). La découverte et la caractérisation de potentiels facteurs de transcription séquence-spécifique chez les dinoflagellés sont encore élusives. Les protéines à domaine CSD, bien que prédite par bio-informatique comme étant des facteurs de transcription, ne semblent pas être des candidats très intéressants pour remplir cette tâche vu leur faible spécificité de liaison aux acides nucléiques (Chapitre 3). L'expérience de purification de chromatine a également permis de trouver quelques motifs, comme les boîtes HMG et le domaine AP2, qui ont déjà été caractérisés chez les apicomplexes, les plus proches parents des dinoflagellés (Chapitre 4). Les protéines à boîte HMG y jouent un rôle de régulation transcriptionnelle durant certains stades de vie des parasites en liant les structures cruciformes de l'ADN et en changeant la courbure de l'ADN pour faciliter la transcription [414-416]. Bien que la présence du domaine AP2 soit intéressante alors qu'il est particulièrement abondant chez les membres du groupe des apicomplexes [417, 418], le faible nombre de ces séquences chez les dinoflagellés suggèrent un rôle spécifique à ces parasites vu que ces protéines sont des régulateurs majeurs du développement chez *Plasmodium spp.* [419], plutôt qu'un rôle important chez tous les alvéolés.

Le faible nombre de facteurs de transcriptions potentiels suggèrent donc qu'un petit nombre de ces protéines pourraient contrôler la transcription chez les dinoflagellés. Par contre, l'analyse des séquences en amont des gènes, effectuée avec le récent séquençage du génome de *Symbiodinium kawagutii*, n'a pas permis de trouver beaucoup de séquences potentiellement promotrices [44]. La seule séquence qui a été retrouvée en abondance est une boîte

TTTT/TTTG qui se retrouve tout juste en amont du site de départ de la transcription de la quasi-entièreté des gènes. L'expérience de purification de la chromatine (Chapitre 4) n'a malheureusement pas permis de repêcher la protéine TLF qui lie la boîte TTTT et qui pourrait potentiellement remplacer TBP [68]. Si la boîte TTTT remplace bel et bien la boîte TATA au niveau fonctionnel, sa présence dans le promoteur proximal d'une si grande proportion de gènes diffère beaucoup de ce qui est observé chez la levure et chez *Arabidopsis*, où moins de 30% des gènes possèdent cette séquence régulatrice [420, 421]. Très peu d'autres séquences promotrices potentielles ont été identifiées dans le génome de *S. kawagutii*, que ce soit d'autres éléments de promoteur basal comme le BRE (*B recognition element)* ou bien des promoteurs plus distants. L'avancement des techniques de séquençage devrait mener au séquençage d'autres génomes de dinoflagellés dans un avenir relativement proche, ce qui pourrait permettre d'obtenir un meilleur portrait des potentiels séquences promotrices chez différentes espèces. Les dinoflagellés se retrouvent donc dans une situation assez inusitée où très peu de régulateurs potentiels de la transcription ont été identifiés et où ces régulateurs n'auraient de toute façon que très peu de séquences régulatrices à lier. Ces observations corrèlent bien avec l'absence de grandes variations dans l'expression des gènes observés au chapitre 2 et aux données récoltées lors de la réponse à plusieurs perturbations des milieux de cultures de ces algues (section 1.2.3.1.3.).

Le très faible nombre de facteurs de transcription et de séquences régulatrices impliquent que la transcription ne serait que très peu régulée chez les dinoflagellés. Par contre, le niveau d'expression des gènes varie beaucoup en termes d'abondance entre les différents gènes. Comme montré au chapitre 2 (Figure 2.1. et 2.S1.), les données d'expression montrent une différence dans l'abondance des transcrits de plus de 10000 fois alors qu'un petit nombre de gènes ont une très forte expression tandis que la majorité des gènes ont une expression assez faible. Ceci implique donc que certains gènes sont très fortement exprimés, et potentiellement transcrit de façon très active, alors qu'une majorité d'autres le sont beaucoup moins. Une hypothèse sur la corrélation entre le nombre de copies des gènes dans le génome et leur niveau d'expression a déjà été formulée avec des observations effectuées chez *A. carterae* [62]. Les gènes organisés en tandem y seraient fortement exprimés alors que ceux présents en une seule copie auraient un niveau d'expression beaucoup plus faible. Par contre,

le niveau d'expression chez *Lingulodinium* ne corrèle pas très bien avec le nombre de copies, qui n'est cependant disponible que pour quelques gènes [61]. Un peu plus de 1000 groupes de gènes sont organisé en tandem dans le génome de *S. kawagutii* [44] mais l'absence de données transcriptomiques de qualité pour cet organisme ne permet malheureusement pas de tester cette hypothèse. Le niveau de méthylation des nucléotides faisant parties des séquences en amont des gènes pourrait également jouer un rôle important dans l'activité des promoteurs. Il a déjà été observé chez *A. carterae* que le niveau de méthylation des séquences intergéniques de PCP changeait en fonction de l'intensité lumineuse et corrélait avec les changements du niveau d'expression de ce gène [65]. Comme les techniques de séquençage permettent maintenant d'identifier le niveau de méthylation des cytosines sur l'entièreté d'un génome, il pourrait être intéressant de les corréler avec le niveau d'expression des gènes chez les dinoflagellés.

Une autre possibilité pouvant influencer l'expression des gènes sans recourir à des facteurs de transcription spécifiques seraient un positionnement différent des gènes sur les chromosomes. Comme la transcription semble se faire préférentiellement sur les boucles périchromosomales qui ne sont pas condensées [46], les gènes qui se retrouvent près de la jonction entre ces boucles et la portion condensée des chromosomes ou même dans la partie condensée de ceux-ci pourraient avoir une accessibilité réduite à l'ARN polymérase et ainsi limiter leur niveau de transcription. La localisation dans les génomes, par le biais d'hybridation *in situ*, de gènes fortement et faiblement exprimés, pourraient être effectuée et corrélée avec leur niveau d'expression. En conclusion, les données amassées dans cette thèse, combinées avec celles précédemment connues sur la transcription, pointent toutes vers une expansion du rôle de la régulation traductionnelle et post-traductionnelle de l'expression des gènes au détriment de la régulation transcriptionnelle chez les dinoflagellés.

Ces données soulèvent également d'intéressantes questions sur les mécanismes de régulation circadienne chez *Lingulodinium*. Comme le modèle traditionnelle de boucle de rétroaction transcription/traduction présent chez les eucaryotes stipule qu'une variation rythmique dans l'abondance des transcrits permet de générer des rythmes, l'absence de changements rythmiques observé au chapitre 2 implique qu'un mécanisme différent se retrouve chez les *Lingulodinium*. Les données de transcriptomiques obtenues ici sont en phase

avec les données obtenues par divers moyens lors de l'étude de quelques gènes comme RubisCO, PCP et LBP. De plus, le fait que l'inhibition de la transcription n'affecte pas la phase et la période des rythmes soutient le rôle peu important que la transcription peut jouer dans la génération et le maintien des rythmes circadiens chez *Lingulodinium*. Bien que l'accumulation des transcrits ne changent pas en fonction du moment de la journée, il est toujours possible que le moment de la transcription d'un groupe plus ou moins important de gènes pourrait être régulé de façon circadienne. Afin de tester cette hypothèse, il pourrait être envisagé d'utiliser un marquage par du 4-thiouracil des ARNs nouvellement transcrits à différents temps, suivi par une biotinylation de ces nucléotides, une technique qui a déjà été utilisée chez le dinoflagellé *K. brevis* [90]. Une purification de ces ARNs suivi d'un séquençage à haut débit permettrait de connaître le moment de transcription d'une grande quantité de gènes. Une étude de la demi-vie des ARNs peut aussi être envisagée avec cette technique. Ces deux types d'expériences permettraient de boucler la boucle et ainsi confirmer la faible importance de la régulation transcriptionnelle dans la génération des rythmes chez *Lingulodinium*.

Comme plusieurs essais d'inhibition de la traduction et de la phosphorylation et déphosphorylation ont montré des perturbations majeures dans les rythmes, il est fort probable que le système circadien de *Lingulodinium* soit à base d'une boucle de rétroaction traduction/post-traduction. Afin de tester cette hypothèse, l'avancée dans les techniques à haut-débit pourrait permettre d'obtenir des informations cruciales sur les protéines impliquées dans ce type de régulation. En ce qui a trait à la traduction, comme il a déjà été montré que les protéines sont majoritairement synthétisées à trois moments différents dans la journée [92, 253], il serait envisageable d'identifier ces protéines à l'échelle de la cellule. La technique de profilage de ribosome (*ribosome profiling*), qui permet de séquencer à haut-débit les ARNs en cours de traduction par les ribosomes [422], permettrait d'obtenir une idée globale de l'identité des gènes qui sont traduits à des moments précis. Cette technique a d'ailleurs été utilisée récemment chez les animaux [423, 424] et les plantes [425] afin d'identifier les gènes traduits de façon rythmique. Elle pourrait être utilisée en conjonction avec l'identification et l'abondance des protéines présentes à différents temps dans la cellule. De façon intéressante, ce type d'expérience a été fait chez les souris et a montré qu'une part importante du protéome

total du foie changeait de façon circadienne, et ce pour plusieurs gènes dont les transcrits ne montraient pas de variations d'abondance [426, 427].

Une deuxième utilisation des techniques à haut-débit concernerait l'identification des modifications post-traductionnelles, en particulier la phosphorylation, à différents temps. La purification de phosphoprotéines suivi par l'identification des protéines par spectrométrie de masse a déjà été effectué à petite échelle chez *Lingulodinium* [95] et pourrait être exploitée afin d'identifier à plusieurs temps les protéines qui y sont phosphorylées. Chez la souris, des études de ce type ont montré qu'environ 25% des sites de phosphorylation montraient des changements journalier important dans leur niveau de phosphorylation et que ces changements avaient une influence majeure sur les réseaux de signalisation et le métabolisme dans le foie [428]. Il serait également possible de combiner cette approche avec les différents inhibiteurs de kinases et de phosphatases qui sont connues pour affecter les rythmes circadiens [111-114] en plus d'inhibiteurs de certaines kinases, comme les CK1 et 2, qui sont des régulateurs conservés des rythmes circadiens chez une grande variété de groupes phylogénétiques. Les différences de phosphorylation entre les cellules traitées avec ces inhibiteurs et les cellules non-traitées pourraient permettre l'identification de plusieurs cibles de l'oscillateur circadien. Finalement, comme le statut redox des cellules, en particulier le niveau d'oxydation des péroxyrédoxines chez les eucaryotes [177] et le pool de plastoquinone [203] chez les cyanobactéries, joue un rôle important dans la régulation circadienne, il pourrait être intéressant de s'attarder à ces mécanismes chez *Lingulodinium* afin de déterminer s'ils sont conservés et s'ils jouent un rôle comme potentiels régulateurs circadiens. La combinaison de ces différentes approches devrait permettre, dans les prochaines années, d'identifier plusieurs des cibles de l'horloge circadienne, et peut-être même l'oscillateur lui-même, chez *Lingulodinium*.

# Bibliographie


1. Taylor F. J. R., Hoppenrath M., Saldarriaga J. F. (2008) **Dinoflagellate diversity and distribution**. *Biodiversity and Conservation*, **17**(2):407-418.
2. Stat M., Morris E., Gates R. D. (2008) **Functional diversity in coral–dinoflagellate symbiosis**. *Proceedings of the National Academy of Sciences*, **105**(27):9256-9261.
3. Field C. B., Behrenfeld M. J., Randerson J. T., Falkowski P. (1998) **Primary production of the biosphere: integrating terrestrial and oceanic components**. *Science (New York, NY)*, **281**(5374):237-240.
4. Anderson D. M., Burkholder J. M., Cochlan W. P., Glibert P. M., Gobler C. J., Heil C. A., Kudela R. M., Parsons M. L., Rensel J. E. J., Townsend D. W., Trainer V. L., Vargo G. A. (2008) **Harmful algal blooms and eutrophication: Examining linkages from selected coastal regions of the United States**. *Harmful Algae*, **8**(1):39-53.
5. Smayda T. J. (1997) **What is a bloom? A commentary**. *Limnology and Oceanography*, **42**(5part2):1132-1136.
6. Wang D.-Z. (2008) **Neurotoxins from Marine Dinoflagellates: A Brief Review**. *Marine Drugs*, **6**(2):349-371.
7. Burki F., Kaplan M., Tikhonenkov D. V., Zlatogursky V., Minh B. Q., Radaykina L. V., Smirnov A., Mylnikov A. P., Keeling P. J. (2016) **Untangling the early diversification of eukaryotes: a phylogenomic study of the evolutionary origins of Centrohelida, Haptophyta and Cryptista**. *Proceedings of the Royal Society B: Biological Sciences*, **283**(1823).
8. Parfrey L. W., Lahr D. J. G., Knoll A. H., Katz L. A. (2011) **Estimating the timing of early eukaryotic diversification with multigene molecular clocks**. *Proceedings of the National Academy of Sciences*, **108**(33):13624-13629.
9. Moore R. B., Obornik M., Janouskovec J., Chrudimsky T., Vancova M., Green D. H., Wright S. W., Davies N. W., Bolch C. J. S., Heimann K., Slapeta J., Hoegh-Guldberg O., Logsdon J. M., Carter D. A. (2008) **A photosynthetic alveolate closely related to apicomplexan parasites**. *Nature*, **451**(7181):959-963.
10. Wisecaver J. H., Hackett J. D. (2011) **Dinoflagellate genome evolution**. *Annual review of microbiology*, **65**:369-387.
11. Joseph S. J., Fernández-Robledo J. A., Gardner M. J., El-Sayed N. M., Kuo C.-H., Schott E. J., Wang H., Kissinger J. C., Vasta G. R. (2010) **The Alveolate Perkinsus marinus: Biological Insights from EST Gene Discovery**. *BMC genomics*, **11**(1):1-21.
12. Gornik Sebastian G., Ford Kristina L., Mulhern Terrence D., Bacic A., McFadden Geoffrey I., Waller Ross F. (2012) **Loss of Nucleosomal DNA Condensation Coincides with Appearance of a Novel Nuclear Protein in Dinoflagellates**. *Current Biology*, **22**(24):2303-2312.
13. Soyer-Gobillard M. O., Dolan M. F. (2015) **Chromosomes of Protists: The crucible of evolution**. *International microbiology : the official journal of the Spanish Society for Microbiology*, **18**(4):209-216.

14. Fast N. M., Kissinger J. C., Roos D. S., Keeling P. J. (2001) **Nuclear-encoded, plastid-targeted genes suggest a single common origin for apicomplexan and dinoflagellate plastids**. *Mol Biol Evol*, **18**(3):418-426.
15. Janouskovec J., Horak A., Obornik M., Lukes J., Keeling P. J. (2010) **A common red algal origin of the apicomplexan, dinoflagellate, and heterokont plastids**. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(24):10949-10954.
16. Nassoury N., Morse D. (2005) **Protein targeting to the chloroplasts of photosynthetic eukaryotes: getting there is half the fun**. *Biochimica et biophysica acta*, **1743**(1-2):5-19.
17. Nassoury N., Cappadocia M., Morse D. (2003) **Plastid ultrastructure defines the protein import pathway in dinoflagellates**. *J Cell Sci*, **116**(Pt 14):2867-2874.
18. Polivka T., van Stokkum I. H., Zigmantas D., van Grondelle R., Sundstrom V., Hiller R. G. (2006) **Energy transfer in the major intrinsic light-harvesting complex from Amphidinium carterae**. *Biochemistry*, **45**(28):8516-8526.
19. Morse D., Salois P., Markovic P., Hastings J. W. (1995) **A nuclear-encoded form II RuBisCO in dinoflagellates**. *Science (New York, NY)*, **268**:1622-1624.
20. Rowan R., Whitney S. M., Fowler A., Yellowlees D. (1996) **Rubisco in marine symbiotic dinoflagellates: form II enzymes in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family**. *Plant Cell*, **8**(3):539-553.
21. Tabita F. R., Hanson T. E., Satagopan S., Witte B. H., Kreel N. E. (2008) **Phylogenetic and evolutionary relationships of RubisCO and the RubisCO-like proteins and the functional lessons provided by diverse molecular forms**. *Philosophical transactions of the Royal Society of London Series B, Biological sciences*, **363**(1504):2629-2640.
22. Schneider G., Lindqvist Y., Branden C. I., Lorimer G. (1986) **Three-dimensional structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from Rhodospirillum rubrum at 2.9 A resolution**. *The EMBO journal*, **5**(13):3409-3415.
23. Nassoury N., Fritz L., Morse D. (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.
24. Lapointe M., Mackenzie T. D., Morse D. (2008) **An external delta-carbonic anhydrase in a free-living marine dinoflagellate may circumvent diffusion-limited carbon acquisition**. *Plant physiology*, **147**(3):1427-1436.
25. Howe C. J., Nisbet R. E., Barbrook A. C. (2008) **The remarkable chloroplast genome of dinoflagellates**. *Journal of experimental botany*, **59**(5):1035-1045.
26. Mungpakdee S., Shinzato C., Takeuchi T., Kawashima T., Koyanagi R., Hisata K., Tanaka M., Goto H., Fujie M., Lin S., Satoh N., Shoguchi E. (2014) **Massive Gene Transfer and Extensive RNA Editing of a Symbiotic Dinoflagellate Plastid Genome**. *Genome Biology and Evolution*, **6**(6):1408-1422.
27. Koumandou V. L., Nisbet R. E., Barbrook A. C., Howe C. J. (2004) **Dinoflagellate chloroplasts--where have all the genes gone?** *Trends Genet*, **20**(5):261-267.
28. Zhang Z., Green B. R., Cavalier-Smith T. (1999) **Single gene circles in dinoflagellate chloroplast genomes**. *Nature*, **400**(6740):155-159.

29. Wang Y., Morse D. (2006) **Rampant polyuridylylation of plastid gene transcripts in the dinoflagellate Lingulodinium**. *Nucleic acids research*, **34**(2):613-619.
30. Dorrell R. G., Hinksman G. A., Howe C. J. (2016) **Diversity of transcripts and transcript processing forms in plastids of the dinoflagellate alga Karenia mikimotoi**. *Plant molecular biology*, **90**(3):233-247.
31. Hikosaka K., Kita K., Tanabe K. (2013) **Diversity of mitochondrial genome structure in the phylum Apicomplexa**. *Molecular and Biochemical Parasitology*, **188**(1):26-33.
32. Waller R. F., Jackson C. J. (2009) **Dinoflagellate mitochondrial genomes: stretching the rules of molecular biology**. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **31**(2):237-245.
33. Shoguchi E., Shinzato C., Hisata K., Satoh N., Mungpakdee S. (2015) **The Large Mitochondrial Genome of Symbiodinium minutum Reveals Conserved Noncoding Sequences between Dinoflagellates and Apicomplexans**. *Genome Biology and Evolution*, **7**(8):2237-2244.
34. Lin S., Zhang H., Spencer D. F., Norman J. E., Gray M. W. (2002) **Widespread and extensive editing of mitochondrial mRNAS in dinoflagellates**. *Journal of molecular biology*, **320**(4):727-739.
35. Jackson C. J., Norman J. E., Schnare M. N., Gray M. W., Keeling P. J., Waller R. F. (2007) **Broad genomic and transcriptional analysis reveals a highly derived genome in dinoflagellate mitochondria**. *BMC biology*, **5**(1):1-17.
36. Soyer-Gobillard M. O., Gillet B., Géraud M. L., Bhaud Y. (1999) **Dinoflagellate chromosome behaviour during stages of replication**. *International Microbiology*, **2**(2):93-102.
37. Bertomeu T., Morse D. (2004) **Isolation of a dinoflagellate mitotic cyclin by functional complementation in yeast**. *Biochemical and biophysical research communications*, **323**(4):1172-1183.
38. Bertomeu T., Rivoal J., Morse D. (2007) **A dinoflagellate CDK5-like cyclin-dependent kinase**. *Biology of the Cell*, **99**(9):531-540.
39. Brunelle S. A., Van Dolah F. M. (2011) **Post-transcriptional regulation of S-phase genes in the dinoflagellate, Karenia brevis**. *The Journal of eukaryotic microbiology*, **58**(4):373-382.
40. Kwok A. C., Wong J. T. (2010) **The activity of a wall-bound cellulase is required for and is coupled to cell cycle progression in the dinoflagellate Crypthecodinium cohnii**. *Plant Cell*, **22**(4):1281-1298.
41. Veldhuis M. J. W., Cucci T. L., Sieracki M. E. (1997) **Cellular DNA content of marine phytoplankton using two new fluorochromes: Taxonomic and ecological implications**. *Journal of Phycology*, **33**(3):527-541.
42. Chow M. H., Yan K. T., Bennett M. J., Wong J. T. (2010) **Birefringence and DNA condensation of liquid crystalline chromosomes**. *Eukaryotic cell*, **9**(10):1577-1587.
43. Fojtova M., Wong J. T., Dvorackova M., Yan K. T., Sykorova E., Fajkus J. (2010) **Telomere maintenance in liquid crystalline chromosomes of dinoflagellates**. *Chromosoma*, **119**(5):485-493.
44. Lin S., Cheng S., Song B., Zhong X., Lin X., Li W., Li L., Zhang Y., Zhang H., Ji Z., Cai M., Zhuang Y., Shi X., Lin L., Wang L., Wang Z., Liu X., Yu S., Zeng P., Hao H., Zou Q., Chen C., Li Y., Wang Y., Xu C., Meng S., Xu X., Wang J., Yang H.,

Campbell D. A., Sturm N. R., Dagenais-Bellefeuille S., Morse D. (2015) **The Symbiodinium kawagutii genome illuminates dinoflagellate gene expression and coral symbiosis**. *Science (New York, NY)*, **350**(6261):691-694.

45. Herzog M., Soyer M. O. (1983) **The native structure of dinoflagellate chromosomes and their stabilization by Ca2+ and Mg2+ cations**. *European journal of cell biology*, **30**(1):33-41.
46. Sigee D. C. (1983) **Structural DNA and genetically active DNA in dinoflagellate chromosomes**. *Bio Systems*, **16**(3-4):203-210.
47. Soyer-Gobillard M. O., Geraud M. L., Coulaud D., Barray M., Theveny B., Revet B., Delain E. (1990) **Location of B- and Z-DNA in the chromosomes of a primitive eukaryote dinoflagellate**. *The Journal of cell biology*, **111**(2):293-304.
48. Rizzo P. J., Jones M., Ray S. M. (1982) **Isolation and properties of isolated nuclei from the Florida red tide dinoflagellate Gymnodinium breve (Davis)**. *The Journal of protozoology*, **29**(2):217-222.
49. Lin S., Zhang H., Zhuang Y., Tran B., Gill J. (2010) **Spliced leader-based metatranscriptomic analyses lead to recognition of hidden genomic features in dinoflagellates**. *Proceedings of the National Academy of Sciences of the United States of America*, **107**(46):20033-20038.
50. Roy S., Morse D. (2012) **A full suite of histone and histone modifying genes are transcribed in the dinoflagellate Lingulodinium**. *PLoS One*, **7**(4):e34340.
51. Rizzo P. J. (1981) **Comparative aspects of basic chromatin proteins in dinoflagellates**. *Bio Systems*, **14**(3-4):433-443.
52. Bodansky S., Mintz L. B., Holmes D. S. (1979) **The mesokaryote Gyrodinium cohnii lacks nucleosomes**. *Biochemical and biophysical research communications*, **88**(4):1329-1336.
53. Chan Y. H., Wong J. T. (2007) **Concentration-dependent organization of DNA by the dinoflagellate histone-like protein HCc3**. *Nucleic acids research*, **35**(8):2573-2583.
54. Sala-Rovira M., Geraud M. L., Caput D., Jacques F., Soyer-Gobillard M. O., Vernet G., Herzog M. (1991) **Molecular cloning and immunolocalization of two variants of the major basic nuclear protein (HCc) from the histone-less eukaryote Crypthecodinium cohnii (Pyrrhophyta)**. *Chromosoma*, **100**(8):510-518.
55. Wong J. T., New D. C., Wong J. C., Hung V. K. (2003) **Histone-like proteins of the dinoflagellate Crypthecodinium cohnii have homologies to bacterial DNA-binding proteins**. *Eukaryotic cell*, **2**(3):646-650.
56. Sun S., Liu M., Dong F., Fan S., Yao Y. (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.
57. Chudnovsky Y., Li J. F., Rizzo P. J., Hastings J. W., Fagan T. F. (2002) **Cloning, expression, and characterization of a histone-like protein from the marine dinoflagellate Lingulodinium polyedrum (dinophyceae)**. *Journal of Phycology*, **38**(3):543-550.
58. Le Q. H., Markovic P., Hastings J. W., Jovine R. V., Morse D. (1997) **Structure and organization of the peridinin-chlorophyll a-binding protein gene in Gonyaulax polyedra**. *Mol Gen Genet*, **255**(6):595-604.

59. Li L., Hastings J. W. (1998) **The structure and organization of the luciferase gene in the photosynthetic dinoflagellate Gonyaulax polyedra**. *Plant molecular biology*, **36**(2):275-284.
60. Machabée S., Wall L., Morse D. (1994) **Expression and genomic organization of a dinoflagellate gene family**. *Plant molecular biology*, **25**(1):23-31.
61. Beauchemin M., Roy S., Daoust P., Dagenais-Bellefeuille S., Bertomeu T., Letourneau L., Lang B. F., Morse D. (2012) **Dinoflagellate tandem array gene transcripts are highly conserved and not polycistronic**. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(39):15793-15798.
62. Bachvaroff T. R., Place A. R. (2008) **From stop to start: tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate Amphidinium carterae**. *PLoS One*, **3**(8):e2929.
63. Shoguchi E., Shinzato C., Kawashima T., Gyoja F., Mungpakdee S., Koyanagi R., Takeuchi T., Hisata K., Tanaka M., Fujiwara M., Hamada M., Seidi A., Fujie M., Usami T., Goto H., Yamasaki S., Arakaki N., Suzuki Y., Sugano S., Toyoda A., Kuroki Y., Fujiyama A., Medina M., Coffroth M. A., Bhattacharya D., Satoh N. (2013) **Draft assembly of the Symbiodinium minutum nuclear genome reveals dinoflagellate gene structure**. *Current biology : CB*, **23**(15):1399-1408.
64. Steele R. E., Rae P. M. M. (1980) **Ordered distribution of modified bases in the DNA of a dinoflagellate**. *Nucleic acids research*, **8**(20):4709-4726.
65. ten Lohuis M. R., Miller D. J. (1998) **Light-regulated transcription of genes encoding peridinin chlorophyll a proteins and the major intrinsic light-harvesting complex proteins in the dinoflagellate amphidinium carterae hulburt (Dinophycae). Changes In cytosine methylation accompany photoadaptation**. *Plant physiology*, **117**(1):189-196.
66. Roy S., Morse D. (2013) **Transcription and Maturation of mRNA in Dinoflagellates**. *Microorganisms*, **1**(1).
67. Meissner M., Soldati D. (2005) **The transcription machinery and the molecular toolbox to control gene expression in Toxoplasma gondii and other protozoan parasites**. *Microbes and infection / Institut Pasteur*, **7**(13):1376-1384.
68. Guillebault D., Sasorith S., Derelle E., Wurtz J. M., Lozano J. C., Bingham S., Tora L., Moreau H. (2002) **A new class of transcription initiation factors, intermediate between TATA box-binding proteins (TBPs) and TBP-like factors (TLFs), is present in the marine unicellular organism, the dinoflagellate Crypthecodinium cohnii**. *The Journal of biological chemistry*, **277**(43):40881-40886.
69. Bayer T., Aranda M., Sunagawa S., Yum L. K., Desalvo M. K., Lindquist E., Coffroth M. A., Voolstra C. R., Medina M. (2012) **Symbiodinium transcriptomes: genome insights into the dinoflagellate symbionts of reef-building corals**. *PLoS One*, **7**(4):e35269.
70. Jiang W., Hou Y., Inouye M. (1997) **CspA, the major cold-shock protein of Escherichia coli, is an RNA chaperone**. *The Journal of biological chemistry*, **272**(1):196-202.
71. Mihailovich M., Militti C., Gabaldon T., Gebauer F. (2010) **Eukaryotic cold shock domain proteins: highly versatile regulators of gene expression**. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **32**(2):109-118.

72. Ting J. P., Painter A., Zeleznik-Le N. J., MacDonald G., Moore T. M., Brown A., Schwartz B. D. (1994) **YB-1 DNA-binding protein represses interferon gamma activation of class II major histocompatibility complex genes**. *The Journal of experimental medicine*, **179**(5):1605-1611.
73. MacDonald G. H., Itoh-Lindstrom Y., Ting J. P. (1995) **The transcriptional regulatory protein, YB-1, promotes single-stranded regions in the DRA promoter**. *The Journal of biological chemistry*, **270**(8):3527-3533.
74. Stickeler E., Fraser S. D., Honig A., Chen A. L., Berget S. M., Cooper T. A. (2001) **The RNA binding protein YB-1 binds A/C-rich exon enhancers and stimulates splicing of the CD44 alternative exon v4**. *The EMBO journal*, **20**(14):3821-3830.
75. Viswanathan S. R., Daley G. Q., Gregory R. I. (2008) **Selective blockade of microRNA processing by Lin28**. *Science (New York, NY)*, **320**(5872):97-100.
76. Bhaud Y., Geraud M. L., Ausseil J., Soyer-Gobillard M. O., Moreau H. (1999) **Cyclic expression of a nuclear protein in a dinoflagellate**. *The Journal of eukaryotic microbiology*, **46**(3):259-267.
77. Guillebault D., Derelle E., Bhaud Y., Moreau H. (2001) **Role of nuclear WW domains and proline-rich proteins in dinoflagellate transcription**. *Protist*, **152**(2):127-138.
78. Zhang H., Hou Y., Miranda L., Campbell D. A., Sturm N. R., Gaasterland T., Lin S. (2007) **Spliced leader RNA trans-splicing in dinoflagellates**. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(11):4618-4623.
79. Guo R., Ki J.-S. (2011) **Spliced leader sequences detected in EST data of the dinoflagellates Cochlodinium polykrikoides and Prorocentrum minimum**. *Algae*, **26**(3):229-235.
80. Zhang S., Sui Z., Chang L., Kang K., Ma J., Kong F., Zhou W., Wang J., Guo L., Geng H., Zhong J., Ma Q. (2014) **Transcriptome de novo assembly sequencing and analysis of the toxic dinoflagellate Alexandrium catenella using the Illumina platform**. *Gene*, **537**(2):285-293.
81. Zhang H., Campbell D. A., Sturm N. R., Lin S. (2009) **Dinoflagellate spliced leader RNA genes display a variety of sequences and genomic arrangements**. *Mol Biol Evol*, **26**(8):1757-1771.
82. Liu L., Hastings J. W. (2006) **Novel and rapidly diverging intergenic sequences between tandem repeats of the luciferase genes in seven dinoflagellate species1**. *Journal of Phycology*, **42**(1):96-103.
83. Lukes J., Leander B. S., Keeling P. J. (2009) **Cascades of convergent evolution: the corresponding evolutionary histories of euglenozoans and dinoflagellates**. *Proceedings of the National Academy of Sciences of the United States of America*, **106 Suppl 1**:9963-9970.
84. Lowe C. D., Mello L. V., Samatar N., Martin L. E., Montagnes D. J., Watts P. C. (2011) **The transcriptome of the novel dinoflagellate Oxyrrhis marina (Alveolata: Dinophyceae): response to salinity examined by 454 sequencing**. *BMC genomics*, **12**:519.
85. Erdner D. L., Anderson D. M. (2006) **Global transcriptional profiling of the toxic dinoflagellate Alexandrium fundyense using Massively Parallel Signature Sequencing**. *BMC genomics*, **7**:88.

86. Morey J. S., Monroe E. A., Kinney A. L., Beal M., Johnson J. G., Hitchcock G. L., Van Dolah F. M. (2011) **Transcriptomic response of the red tide dinoflagellate, Karenia brevis, to nitrogen and phosphorus depletion and addition**. *BMC genomics*, **12**:346.
87. Moustafa A., Evans A. N., Kulis D. M., Hackett J. D., Erdner D. L., Anderson D. M., Bhattacharya D. (2010) **Transcriptome profiling of a toxic dinoflagellate reveals a gene-rich protist and a potential impact on gene expression due to bacterial presence**. *PLoS One*, **5**(3):e9688.
88. Johnson J. G., Morey J. S., Neely M. G., Ryan J. C., Van Dolah F. M. (2012) **Transcriptome remodeling associated with chronological aging in the dinoflagellate, Karenia brevis**. *Marine genomics*, **5**:15-25.
89. Xiang T., Nelson W., Rodriguez J., Tolleter D., Grossman A. R. (2015) **Symbiodinium transcriptome and global responses of cells to immediate changes in light intensity when grown under autotrophic or mixotrophic conditions**. *The Plant journal : for cell and molecular biology*, **82**(1):67-80.
90. Morey J. S., Van Dolah F. M. (2013) **Global analysis of mRNA half-lives and de novo transcription in a dinoflagellate, Karenia brevis**. *PLoS One*, **8**(6):e66347.
91. Akimoto H., Wu C., Kinumi T., Ohmiya Y. (2004) **Biological rhythmicity in expressed proteins of the marine dinoflagellate Lingulodinium polyedrum demonstrated by chronological proteomics**. *Biochemical and biophysical research communications*, **315**(2):306-312.
92. Markovic P., Roenneberg T., Morse D. (1996) **Phased protein synthesis at several circadian times does not change protein levels in Gonyaulax**. *Journal of biological rhythms*, **11**(1):57-67.
93. Mittag M., Lee D. H., Hastings J. W. (1994) **Circadian expression of the luciferin-binding protein correlates with the binding of a protein to the 3' untranslated region of its mRNA**. *Proceedings of the National Academy of Sciences of the United States of America*, **91**(12):5257-5261.
94. Lapointe M., Morse D. (2008) **Reassessing the role of a 3'-UTR-binding translational inhibitor in regulation of circadian bioluminescence rhythm in the dinoflagellate Gonyaulax**. *Biological chemistry*, **389**(1):13-19.
95. Roy S., Morse D. (2014) **The dinoflagellate Lingulodinium has predicted casein kinase 2 sites in many RNA binding proteins**. *Protist*, **165**(3):330-342.
96. Jones G. D., Williams E. P., Place A. R., Jagus R., Bachvaroff T. R. (2015) **The alveolate translation initiation factor 4E family reveals a custom toolkit for translational control in core dinoflagellates**. *BMC evolutionary biology*, **15**:14.
97. Jagus R., Bachvaroff T. R., Joshi B., Place A. R. (2012) **Diversity of Eukaryotic Translational Initiation Factor eIF4E in Protists**. *Comparative and functional genomics*, **2012**:134839.
98. Rhoads R. E. (2009) **eIF4E: new family members, new binding partners, new roles**. *The Journal of biological chemistry*, **284**(25):16711-16715.
99. Zhao T., Li G., Mi S., Li S., Hannon G. J., Wang X. J., Qi Y. (2007) **A complex system of small RNAs in the unicellular green alga Chlamydomonas reinhardtii**. *Genes & development*, **21**(10):1190-1203.

100. Molnar A., Schwach F., Studholme D. J., Thuenemann E. C., Baulcombe D. C. (2007) **miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii**. *Nature*, **447**(7148):1126-1129.
101. Yamasaki T., Voshall A., Kim E. J., Moriyama E., Cerutti H., Ohama T. (2013) **Complementarity to an miRNA seed region is sufficient to induce moderate repression of a target transcript in the unicellular green alga Chlamydomonas reinhardtii**. *The Plant journal : for cell and molecular biology*, **76**(6):1045-1056.
102. Norden-Krichmar T. M., Allen A. E., Gaasterland T., Hildebrand M. (2011) **Characterization of the small RNA transcriptome of the diatom, Thalassiosira pseudonana**. *PLoS One*, **6**(8):e22870.
103. Huang A., He L., Wang G. (2011) **Identification and characterization of microRNAs from Phaeodactylum tricornutum by high-throughput sequencing and bioinformatics analysis**. *BMC genomics*, **12**:337.
104. Braun L., Cannella D., Ortet P., Barakat M., Sautel C. F., Kieffer S., Garin J., Bastien O., Voinnet O., Hakimi M. A. (2010) **A complex small RNA repertoire is generated by a plant/fungal-like machinery and effected by a metazoan-like Argonaute in the single-cell human parasite Toxoplasma gondii**. *PLoS pathogens*, **6**(5):e1000920.
105. Hinas A., Reimegard J., Wagner E. G., Nellen W., Ambros V. R., Soderbom F. (2007) **The small RNA repertoire of Dictyostelium discoideum and its regulation by components of the RNAi pathway**. *Nucleic acids research*, **35**(20):6714-6726.
106. Tarver J. E., Donoghue P. C., Peterson K. J. (2012) **Do miRNAs have a deep evolutionary history?** *BioEssays : news and reviews in molecular, cellular and developmental biology*, **34**(10):857-866.
107. Baumgarten S., Bayer T., Aranda M., Liew Y. J., Carr A., Micklem G., Voolstra C. R. (2013) **Integrating microRNA and mRNA expression profiling in Symbiodinium microadriaticum, a dinoflagellate symbiont of reef-building corals**. *BMC genomics*, **14**:704.
108. Gao D., Qiu L., Hou Z., Zhang Q., Wu J., Gao Q., Song L. (2013) **Computational Identification of MicroRNAs from the Expressed Sequence Tags of Toxic Dinoflagellate Alexandrium Tamarense**. *Evolutionary bioinformatics online*, **9**:479-485.
109. Geng H., Sui Z., Zhang S., Du Q., Ren Y., Liu Y., Kong F., Zhong J., Ma Q. (2015) **Identification of microRNAs in the Toxigenic Dinoflagellate Alexandrium catenella by High-Throughput Illumina Sequencing and Bioinformatic Analysis**. *PLoS One*, **10**(9):e0138709.
110. Dagenais-Bellefeuille S., Beauchemin M., Morse D. (2017) **miRNAs Do Not Regulate Circadian Protein Synthesis in the Dinoflagellate Lingulodinium polyedrum**. *PLOS ONE*, **12**(1):e0168817.
111. Comolli J. C., Hastings J. W. (1999) **Novel effects on the Gonyaulax circadian system produced by the protein kinase inhibitor staurosporine**. *Journal of biological rhythms*, **14**(1):11-19.
112. Comolli J., Taylor W., Hastings J. W. (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.

113. Comolli J. C., Fagan T., Hastings J. W. (2003) **A type-1 phosphoprotein phosphatase from a dinoflagellate as a possible component of the circadian mechanism**. *Journal of biological rhythms*, **18**(5):367-376.
114. Comolli J., Taylor W., Rehman J., Hastings J. W. (1996) **Inhibitors of serine/threonine phosphoprotein phosphatases alter circadian properties in Gonyaulax polyedra**. *Plant physiology*, **111**(1):285-291.
115. Liu B., Lo S. C., Matton D. P., Lang B. F., Morse D. (2012) **Daily changes in the phosphoproteome of the dinoflagellate Lingulodinium**. *Protist*, **163**(5):746-754.
116. Zhang W., Qin Z., Zhang X., Xiao W. (2011) **Roles of sequential ubiquitination of PCNA in DNA-damage tolerance**. *FEBS letters*, **585**(18):2786-2794.
117. Green R. M., Tingay S., Wang Z.-Y., Tobin E. M. (2002) **Circadian Rhythms Confer a Higher Level of Fitness to Arabidopsis Plants**. *Plant physiology*, **129**(2):576-584.
118. Dodd A. N., Salathia N., Hall A., Kévei E., Tóth R., Nagy F., Hibberd J. M., Millar A. J., Webb A. A. R. (2005) **Plant Circadian Clocks Increase Photosynthesis, Growth, Survival, and Competitive Advantage**. *Science (New York, NY)*, **309**(5734):630-633.
119. Ouyang Y., Andersson C. R., Kondo T., Golden S. S., Johnson C. H. (1998) **Resonating circadian clocks enhance fitness in cyanobacteria**. *Proceedings of the National Academy of Sciences of the United States of America*, **95**(15):8660-8664.
120. Woelfle M. A., Ouyang Y., Phanvijhitsiri K., Johnson C. H. (2004) **The Adaptive Value of Circadian Clocks: An Experimental Assessment in Cyanobacteria**. *Current Biology*, **14**(16):1481-1486.
121. McClung C. R. (2006) **Plant Circadian Rhythms**. *The Plant Cell*, **18**(4):792-803.
122. Dunlap J. C., Loros J. J., DeCoursey P. J. (2004) **Chronobiology : Biological Timekeeping**. Sunderland MA: Sinauer Associates.
123. Brody S. (2011) **The Genetics of Circadian Rhythms**, vol. 74: Academic Press.
124. Hastings J. W., Sweeney B. M. (1957) **On the mechanism of temperature independence in a biological clock**. *Proceedings of the National Academy of Sciences of the United States of America*, **43**(9):804-811.
125. Reyes B. A., Pendergast J. S., Yamazaki S. (2008) **Mammalian Peripheral Circadian Oscillators Are Temperature Compensated**. *Journal of biological rhythms*, **23**(1):95-98.
126. Kusakina J., Gould P. D., Hall A. (2014) **A fast circadian clock at high temperatures is a conserved feature across Arabidopsis accessions and likely to be important for vegetative yield**. *Plant Cell Environ*, **37**(2):327-340.
127. Young M. W., Kay S. A. (2001) **Time zones: a comparative genetics of circadian clocks**. *Nat Rev Genet*, **2**(9):702-715.
128. Bouget F. Y., Lefranc M., Thommen Q., Pfeuty B., Lozano J. C., Schatt P., Botebol H., Verge V. (2014) **Transcriptional versus non-transcriptional clocks: A case study in Ostreococcus**. *Marine genomics*, **14c**:17-22.
129. Panda S., Antoch M. P., Miller B. H., Su A. I., Schook A. B., Straume M., Schultz P. G., Kay S. A., Takahashi J. S., Hogenesch J. B. (2002) **Coordinated transcription of key pathways in the mouse by the circadian clock**. *Cell*, **109**(3):307-320.
130. Harmer S. L., Hogenesch J. B., Straume M., Chang H. S., Han B., Zhu T., Wang X., Kreps J. A., Kay S. A. (2000) **Orchestrated transcription of key pathways in Arabidopsis by the circadian clock**. *Science (New York, NY)*, **290**(5499):2110-2113.

131. Duffield G. E., Best J. D., Meurers B. H., Bittner A., Loros J. J., Dunlap J. C. (2002) **Circadian programs of transcriptional activation, signaling, and protein turnover revealed by microarray analysis of mammalian cells**. *Current biology : CB*, **12**(7):551-557.
132. Covington M., Maloof J., Straume M., Kay S., Harmer S. (2008) **Global transcriptome analysis reveals circadian regulation of key pathways in plant growth and development**. *Genome biology*, **9**(8):R130.
133. Monnier A., Liverani S., Bouvet R., Jesson B., Smith J. Q., Mosser J., Corellou F., Bouget F. Y. (2010) **Orchestrated transcription of biological processes in the marine picoeukaryote Ostreococcus exposed to light/dark cycles**. *BMC genomics*, **11**:192.
134. Le Martelot G., Canella D., Symul L., Migliavacca E., Gilardi F., Liechti R., Martin O., Harshman K., Delorenzi M., Desvergne B., Herr W., Deplancke B., Schibler U., Rougemont J., Guex N., Hernandez N., Naef F., the Cycli X. c. (2012) **Genome-Wide RNA Polymerase II Profiles and RNA Accumulation Reveal Kinetics of Transcription and Associated Epigenetic Changes During Diurnal Cycles**. *PLoS biology*, **10**(11):e1001442.
135. Michael T. P., McClung C. R. (2003) **Enhancer trapping reveals widespread circadian clock transcriptional control in Arabidopsis**. *Plant physiology*, **132**(2):629-639.
136. Harmer S. L. (2009) **The circadian system in higher plants**. *Annual review of plant biology*, **60**:357-377.
137. Noordally Z. B., Millar A. J. (2015) **Clocks in algae**. *Biochemistry*, **54**(2):171-183.
138. Cui Z., Xu Q., Wang X. (2014) **Regulation of the circadian clock through pre-mRNA splicing in Arabidopsis**. *Journal of experimental botany*, **65**(8):1973-1980.
139. Derelle E., Ferraz C., Rombauts S., Rouze P., Worden A. Z., Robbens S., Partensky F., Degroeve S., Echeynie S., Cooke R., Saeys Y., Wuyts J., Jabbari K., Bowler C., Panaud O., Piegu B., Ball S. G., Ral J. P., Bouget F. Y., Piganeau G., De Baets B., Picard A., Delseny M., Demaille J., Van de Peer Y., Moreau H. (2006) **Genome analysis of the smallest free-living eukaryote Ostreococcus tauri unveils many unique features**. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(31):11647-11652.
140. Hsu P. Y., Harmer S. L. (2014) **Wheels within wheels: the plant circadian system**. *Trends in plant science*, **19**(4):240-249.
141. Huang W., Perez-Garcia P., Pokhilko A., Millar A. J., Antoshechkin I., Riechmann J. L., Mas P. (2012) **Mapping the core of the Arabidopsis circadian clock defines the network structure of the oscillator**. *Science (New York, NY)*, **336**(6077):75-79.
142. Pokhilko A., Fernandez A. P., Edwards K. D., Southern M. M., Halliday K. J., Millar A. J. (2012) **The clock gene circuit in Arabidopsis includes a repressilator with additional feedback loops**. *Molecular systems biology*, **8**:574.
143. Lu S. X., Knowles S. M., Andronis C., Ong M. S., Tobin E. M. (2009) **CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL function synergistically in the circadian clock of Arabidopsis**. *Plant physiology*, **150**(2):834-843.

144. Farre E. M., Harmer S. L., Harmon F. G., Yanovsky M. J., Kay S. A. (2005) **Overlapping and distinct roles of PRR7 and PRR9 in the Arabidopsis circadian clock**. *Current biology : CB*, **15**(1):47-54.
145. Nakamichi N., Kiba T., Henriques R., Mizuno T., Chua N. H., Sakakibara H. (2010) **PSEUDO-RESPONSE REGULATORS 9, 7, and 5 are transcriptional repressors in the Arabidopsis circadian clock**. *Plant Cell*, **22**(3):594-605.
146. Herrero E., Kolmos E., Bujdoso N., Yuan Y., Wang M., Berns M. C., Uhlworm H., Coupland G., Saini R., Jaskolski M., Webb A., Gonçalves J., Davis S. J. (2012) **EARLY FLOWERING4 Recruitment of EARLY FLOWERING3 in the Nucleus Sustains the Arabidopsis Circadian Clock**. *The Plant Cell*, **24**(2):428-443.
147. Helfer A., Nusinow D. A., Chow B. Y., Gehrke A. R., Bulyk M. L., Kay S. A. (2011) **LUX ARRHYTHMO Encodes a Nighttime Repressor of Circadian Gene Expression in the Arabidopsis Core Clock**. *Current Biology*, **21**(2):126-133.
148. Dixon L. E., Knox K., Kozma-Bognar L., Southern M. M., Pokhilko A., Millar A. J. (2011) **Temporal Repression of Core Circadian Genes Is Mediated through EARLY FLOWERING 3 in Arabidopsis**. *Current Biology*, **21**(2):120-125.
149. Corellou F., Schwartz C., Motta J. P., Djouani-Tahri E. B., Sanchez F., Bouget Fç Y. (2009) **Clocks in the Green Lineage: Comparative Functional Analysis of the Circadian Architecture of the Picoeukaryote Ostreococcus**. *Plant Cell*, **21**(11):3436-3449.
150. Djouani-Tahri E. B., Motta J. P., Bouget Fç Y., Corellou F. (2010) **Insights into the regulation of the core clock component TOC1 in the green picoeukaryote Ostreococcus**. *Plant Signaling & Behavior*, **5**(3):332-335.
151. Djouani-Tahri el B., Sanchez F., Lozano J. C., Bouget F. Y. (2011) **A phosphate-regulated promoter for fine-tuned and reversible overexpression in Ostreococcus: application to circadian clock functional analysis**. *PLoS One*, **6**(12):e28471.
152. van Ooijen G., Dixon L. E., Troein C., Millar A. J. (2011) **Proteasome function is required for biological timing throughout the twenty-four hour cycle**. *Current biology : CB*, **21**(10):869-875.
153. Le Bihan T., Hindle M., Martin S. F., Barrios-Llerena M. E., Krahmer J., Kis K., Millar A. J., van Ooijen G. (2015) **Label-free quantitative analysis of the casein kinase 2-responsive phosphoproteome of the marine minimal model species Ostreococcus tauri**. *Proteomics*, **15**(23-24):4135-4144.
154. van Ooijen G., Hindle M., Martin S. F., Barrios-Llerena M., Sanchez F., Bouget F. Y., O'Neill J. S., Le Bihan T., Millar A. J. (2013) **Functional analysis of Casein Kinase 1 in a minimal circadian system**. *PLoS One*, **8**(7):e70021.
155. van Ooijen G., Martin S. F., Barrios-Llerena M. E., Hindle M., Le Bihan T., O'Neill J. S., Millar A. J. (2013) **Functional analysis of the rodent CK1tau mutation in the circadian clock of a marine unicellular alga**. *BMC cell biology*, **14**:46.
156. Morant P. E., Thommen Q., Pfeuty B., Vandermoere C., Corellou F., Bouget F. Y., Lefranc M. (2010) **A robust two-gene oscillator at the core of Ostreococcus tauri circadian clock**. *Chaos (Woodbury, NY)*, **20**(4):045108.
157. Troein C., Corellou F., Dixon L. E., van Ooijen G., O'Neill J. S., Bouget F. Y., Millar A. J. (2011) **Multiple light inputs to a simple clock circuit allow complex biological rhythms**. *The Plant journal : for cell and molecular biology*, **66**(2):375-385.

158. Dixon L. E., Hodge S. K., van Ooijen G., Troein C., Akman O. E., Millar A. J. (2014) **Light and circadian regulation of clock components aids flexible responses to environmental signals**. *The New phytologist*, **203**(2):568-577.
159. Heijde M., Zabulon G., Corellou F., Ishikawa T., Brazard J., Usman A., Sanchez F., Plaza P., Martin M., Falciatore A., Todo T., Bouget F. Y., Bowler C. (2010) **Characterization of two members of the cryptochrome/photolyase family from Ostreococcus tauri provides insights into the origin and evolution of cryptochromes**. *Plant Cell Environ*, **33**(10):1614-1626.
160. Djouani-Tahri el B., Christie J. M., Sanchez-Ferandin S., Sanchez F., Bouget F. Y., Corellou F. (2011) **A eukaryotic LOV-histidine kinase with circadian clock function in the picoalga Ostreococcus**. *The Plant journal : for cell and molecular biology*, **65**(4):578-588.
161. Coesel S., Mangogna M., Ishikawa T., Heijde M., Rogato A., Finazzi G., Todo T., Bowler C., Falciatore A. (2009) **Diatom PtCPF1 is a new cryptochrome/photolyase family member with DNA repair and transcription regulation activity**. *EMBO reports*, **10**(6):655-661.
162. Bowler C., Allen A. E., Badger J. H., Grimwood J., Jabbari K., Kuo A., Maheswari U., Martens C., Maumus F., Otillar R. P., Rayko E., Salamov A., Vandepoele K., Beszteri B., Gruber A., Heijde M., Katinka M., Mock T., Valentin K., Verret F., Berges J. A., Brownlee C., Cadoret J.-P., Chiovitti A., Choi C. J., Coesel S., De Martino A., Detter J. C., Durkin C., Falciatore A., Fournet J., Haruta M., Huysman M. J. J., Jenkins B. D., Jiroutova K., Jorgensen R. E., Joubert Y., Kaplan A., Kroger N., Kroth P. G., La Roche J., Lindquist E., Lommer M., Martin-Jezequel V., Lopez P. J., Lucas S., Mangogna M., McGinnis K., Medlin L. K., Montsant A., Secq M.-P. O.-L., Napoli C., Obornik M., Parker M. S., Petit J.-L., Porcel B. M., Poulsen N., Robison M., Rychlewski L., Rynearson T. A., Schmutz J., Shapiro H., Siaut M., Stanley M., Sussman M. R., Taylor A. R., Vardi A., von Dassow P., Vyverman W., Willis A., Wyrwicz L. S., Rokhsar D. S., Weissenbach J., Armbrust E. V., Green B. R., Van de Peer Y., Grigoriev I. V. (2008) **The Phaeodactylum genome reveals the evolutionary history of diatom genomes**. *Nature*, **456**(7219):239-244.
163. Mas P., Kim W.-Y., Somers D. E., Kay S. A. (2003) **Targeted degradation of TOC1 by ZTL modulates circadian function in Arabidopsis thaliana**. *Nature*, **426**(6966):567-570.
164. Thommen Q., Pfeuty B., Schatt P., Bijoux A., Bouget F Y., Lefranc M. (2015) **Probing entrainment of Ostreococcus tauri circadian clock by green and blue light through a mathematical modeling approach**. *Frontiers in Genetics*, **6**.
165. Wang H., Wang H. (2015) **Phytochrome signaling: time to tighten up the loose ends**. *Molecular plant*, **8**(4):540-551.
166. Duanmu D., Bachy C., Sudek S., Wong C. H., Jiménez V., Rockwell N. C., Martin S. S., Ngan C. Y., Reistetter E. N., van Baren M. J., Price D. C., Wei C. L., Reyes-Prieto A., Lagarias J. C., Worden A. Z. (2014) **Marine algae and land plants share conserved phytochrome signaling systems**. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(44):15827-15832.
167. Fortunato A. E., Jaubert M., Enomoto G., Bouly J. P., Raniello R., Thaler M., Malviya S., Bernardes J. S., Rappaport F., Gentili B., Huysman M. J. J., Carbone A., Bowler C.,

d'Alcalà M. R., Ikeuchi M., Falciatore A. (2016) **Diatom Phytochromes Reveal the Existence of Far-Red-Light-Based Sensing in the Ocean**. *Plant Cell*, **28**(3):616-628.

168. Haydon M. J., Mielczarek O., Robertson F. C., Hubbard K. E., Webb A. A. (2013) **Photosynthetic entrainment of the Arabidopsis thaliana circadian clock**. *Nature*, **502**(7473):689-692.
169. O'Neill J. S., van Ooijen G., Dixon L. E., Troein C., Corellou F., Bouget F. Y., Reddy A. B., Millar A. J. (2011) **Circadian rhythms persist without transcription in a eukaryote**. *Nature*, **469**(7331):554-558.
170. Nakamichi N., Kiba T., Kamioka M., Suzuki T., Yamashino T., Higashiyama T., Sakakibara H., Mizuno T. (2012) **Transcriptional repressor PRR5 directly regulates clock-output pathways**. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(42):17123-17128.
171. Noordally Z. B., Ishii K., Atkins K. A., Wetherill S. J., Kusakina J., Walton E. J., Kato M., Azuma M., Tanaka K., Hanaoka M., Dodd A. N. (2013) **Circadian Control of Chloroplast Transcription by a Nuclear-Encoded Timing Signal**. *Science (New York, NY)*, **339**(6125):1316-1319.
172. Nusinow D. A., Helfer A., Hamilton E. E., King J. J., Imaizumi T., Schultz T. F., Farre E. M., Kay S. A. (2011) **The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth**. *Nature*, **475**(7356):398-402.
173. Nolte C., Staiger D. (2015) **RNA around the clock - regulation at the RNA level in biological timing**. *Front Plant Sci*, **6**:311.
174. Moulager M., Monnier A., Jesson B., Bouvet R., Mosser J., Schwartz C., Garnier L., Corellou F., Bouget Fç Y. (2007) **Light-Dependent Regulation of Cell Division in Ostreococcus: Evidence for a Major Transcriptional Input**. *Plant physiology*, **144**(3):1360-1369.
175. Moulager M., Corellou F., Verge V., Escande M. L., Bouget F. Y. (2010) **Integration of light signals by the retinoblastoma pathway in the control of S phase entry in the picophytoplanktonic cell Ostreococcus**. *PLoS genetics*, **6**(5):e1000957.
176. Nakajima M., Imai K., Ito H., Nishiwaki T., Murayama Y., Iwasaki H., Oyama T., Kondo T. (2005) **Reconstitution of circadian oscillation of cyanobacterial KaiC phosphorylation in vitro**. *Science (New York, NY)*, **308**(5720):414-415.
177. Edgar R. S., Green E. W., Zhao Y., van Ooijen G., Olmedo M., Qin X., Xu Y., Pan M., Valekunja U. K., Feeney K. A., Maywood E. S., Hastings M. H., Baliga N. S., Merrow M., Millar A. J., Johnson C. H., Kyriacou C. P., O'Neill J. S., Reddy A. B. (2012) **Peroxiredoxins are conserved markers of circadian rhythms**. *Nature*, **485**(7399):459-464.
178. Grobbelaar N., Huang T. C., Lin H. Y., Chow T. J. (1986) **Dinitrogen-fixing endogenous rhythm in Synechococcus RF-1**. *FEMS Microbiology Letters*, **37**(2):173-177.
179. Chen T. H., Chen T. L., Hung L. M., Huang T. C. (1991) **Circadian Rhythm in Amino Acid Uptake by Synechococcus RF-1**. *Plant physiology*, **97**(1):55-59.
180. Ishiura M., Kutsuna S., Aoki S., Iwasaki H., Andersson C. R., Tanabe A., Golden S. S., Johnson C. H., Kondo T. (1998) **Expression of a gene cluster kaiABC as a circadian feedback process in cyanobacteria**. *Science (New York, NY)*, **281**(5382):1519-1523.

181. Kitayama Y., Iwasaki H., Nishiwaki T., Kondo T. (2003) **KaiB functions as an attenuator of KaiC phosphorylation in the cyanobacterial circadian clock system**. *The EMBO journal*, **22**(9):2127-2134.
182. Xu Y., Mori T., Johnson C. H. (2000) **Circadian clock-protein expression in cyanobacteria: rhythms and phase setting**. *The EMBO journal*, **19**(13):3349-3357.
183. Tomita J., Nakajima M., Kondo T., Iwasaki H. (2005) **No Transcription-Translation Feedback in Circadian Rhythm of KaiC Phosphorylation**. *Science (New York, NY)*, **307**(5707):251-254.
184. Teng S. W., Mukherji S., Moffitt J. R., de Buyl S., O'Shea E. K. (2013) **Robust circadian oscillations in growing cyanobacteria require transcriptional feedback**. *Science (New York, NY)*, **340**(6133):737-740.
185. Zwicker D., Lubensky D. K., ten Wolde P. R. (2010) **Robust circadian clocks from coupled protein-modification and transcription–translation cycles**. *Proceedings of the National Academy of Sciences*, **107**(52):22540-22545.
186. Pattanayek R., Wang J., Mori T., Xu Y., Johnson C. H., Egli M. (2004) **Visualizing a Circadian Clock Protein: Crystal Structure of KaiC and Functional Insights**. *Molecular Cell*, **15**(3):375-388.
187. Mori T., Saveliev S. V., Xu Y., Stafford W. F., Cox M. M., Inman R. B., Johnson C. H. (2002) **Circadian clock protein KaiC forms ATP-dependent hexameric rings and binds DNA**. *Proceedings of the National Academy of Sciences of the United States of America*, **99**(26):17203-17208.
188. Nishiwaki T., Iwasaki H., Ishiura M., Kondo T. (2000) **Nucleotide binding and autophosphorylation of the clock protein KaiC as a circadian timing process of cyanobacteria**. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(1):495-499.
189. Nishiwaki T., Satomi Y., Kitayama Y., Terauchi K., Kiyohara R., Takao T., Kondo T. (2007) **A sequential program of dual phosphorylation of KaiC as a basis for circadian rhythm in cyanobacteria**. *The EMBO journal*, **26**(17):4029-4037.
190. Rust M. J., Markson J. S., Lane W. S., Fisher D. S., O'Shea E. K. (2007) **Ordered phosphorylation governs oscillation of a three-protein circadian clock**. *Science (New York, NY)*, **318**(5851):809-812.
191. Kim Y. I., Dong G., Carruthers C. W., Jr., Golden S. S., LiWang A. (2008) **The day/night switch in KaiC, a central oscillator component of the circadian clock of cyanobacteria**. *Proceedings of the National Academy of Sciences of the United States of America*, **105**(35):12825-12830.
192. Chang Y. G., Tseng R., Kuo N. W., LiWang A. (2012) **Rhythmic ring-ring stacking drives the circadian oscillator clockwise**. *Proceedings of the National Academy of Sciences of the United States of America*, **109**(42):16847-16851.
193. Chang Y. G., Cohen S. E., Phong C., Myers W. K., Kim Y. I., Tseng R., Lin J., Zhang L., Boyd J. S., Lee Y., Kang S., Lee D., Li S., Britt R. D., Rust M. J., Golden S. S., LiWang A. (2015) **A protein fold switch joins the circadian oscillator to clock output in cyanobacteria**. *Science (New York, NY)*, **349**(6245):324-328.
194. Tseng R., Chang Y. G., Bravo I., Latham R., Chaudhary A., Kuo N. W., Liwang A. (2014) **Cooperative KaiA-KaiB-KaiC interactions affect KaiB/SasA competition in the circadian clock of cyanobacteria**. *Journal of molecular biology*, **426**(2):389-402.

195. Axmann I. M., Hertel S., Wiegard A., Dorrich A. K., Wilde A. (2014) **Diversity of KaiC-based timing systems in marine Cyanobacteria**. *Marine genomics*, **14**:3-16.
196. Hoyle N. P., O'Neill J. S. (2015) **Oxidation-reduction cycles of peroxiredoxin proteins and nontranscriptional aspects of timekeeping**. *Biochemistry*, **54**(2):184-193.
197. Kageyama H., Nishiwaki T., Nakajima M., Iwasaki H., Oyama T., Kondo T. (2006) **Cyanobacterial circadian pacemaker: Kai protein complex dynamics in the KaiC phosphorylation cycle in vitro**. *Mol Cell*, **23**(2):161-171.
198. Ito H., Kageyama H., Mutsuda M., Nakajima M., Oyama T., Kondo T. (2007) **Autonomous synchronization of the circadian KaiC phosphorylation rhythm**. *Nature structural & molecular biology*, **14**(11):1084-1088.
199. Mori T., Williams D. R., Byrne M. O., Qin X., Egli M., McHaourab H. S., Stewart P. L., Johnson C. H. (2007) **Elucidating the Ticking of an In Vitro Circadian Clockwork**. *PLoS biology*, **5**(4):e93.
200. Dvornyk V., Vinogradova O., Nevo E. (2003) **Origin and evolution of circadian clock genes in prokaryotes**. *Proceedings of the National Academy of Sciences of the United States of America*, **100**(5):2495-2500.
201. Holtzendorff J., Partensky F., Mella D., Lennon J. F., Hess W. R., Garczarek L. (2008) **Genome streamlining results in loss of robustness of the circadian clock in the marine cyanobacterium Prochlorococcus marinus PCC 9511**. *Journal of biological rhythms*, **23**(3):187-199.
202. Rust M. J., Golden S. S., O'Shea E. K. (2011) **Light-Driven Changes in Energy Metabolism Directly Entrain the Cyanobacterial Circadian Oscillator**. *Science (New York, NY)*, **331**(6014):220-223.
203. Kim Y.-I., Vinyard D. J., Ananyev G. M., Dismukes G. C., Golden S. S. (2012) **Oxidized quinones signal onset of darkness directly to the cyanobacterial circadian oscillator**. *Proceedings of the National Academy of Sciences*, **109**(44):17765-17769.
204. Wood T. L., Bridwell-Rabb J., Kim Y.-I., Gao T., Chang Y.-G., LiWang A., Barondeau D. P., Golden S. S. (2010) **The KaiA protein of the cyanobacterial circadian oscillator is modulated by a redox-active cofactor**. *Proceedings of the National Academy of Sciences*, **107**(13):5804-5809.
205. Takai N., Ikeuchi S., Manabe K., Kutsuna S. (2006) **Expression of the circadian clock-related gene pex in cyanobacteria increases in darkness and is required to delay the clock**. *Journal of biological rhythms*, **21**(4):235-244.
206. Kobayashi T., Obana Y., Kuboi N., Kitayama Y., Hayashi S., Oka M., Wada N., Arita K., Shimizu T., Sato M., Kanaly R. A., Kutsuna S. (2016) **Analysis of the Fine-Tuning of Cyanobacterial Circadian Phase by Monochromatic Light and Long-Day Conditions**. *Plant and Cell Physiology*, **57**(1):105-114.
207. Schmitz O., Katayama M., Williams S. B., Kondo T., Golden S. S. (2000) **CikA, a bacteriophytochrome that resets the cyanobacterial circadian clock**. *Science (New York, NY)*, **289**(5480):765-768.
208. Mutsuda M., Michel K. P., Zhang X., Montgomery B. L., Golden S. S. (2003) **Biochemical properties of CikA, an unusual phytochrome-like histidine protein kinase that resets the circadian clock in Synechococcus elongatus PCC 7942**. *The Journal of biological chemistry*, **278**(21):19102-19110.

209. Zhang X., Dong G., Golden S. S. (2006) **The pseudo-receiver domain of CikA regulates the cyanobacterial circadian input pathway**. *Molecular microbiology*, **60**(3):658-668.
210. Ivleva N. B., Gao T., LiWang A. C., Golden S. S. (2006) **Quinone sensing by the circadian input kinase of the cyanobacterial circadian clock**. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(46):17468-17473.
211. Katayama M., Kondo T., Xiong J., Golden S. S. (2003) **ldpA Encodes an Iron-Sulfur Protein Involved in Light-Dependent Modulation of the Circadian Period in the Cyanobacterium Synechococcus elongatus PCC 7942**. *Journal of bacteriology*, **185**(4):1415-1422.
212. Ivleva N. B., Bramlett M. R., Lindahl P. A., Golden S. S. (2005) **LdpA: a component of the circadian clock senses redox state of the cell**. *The EMBO journal*, **24**(6):1202-1210.
213. Liu Y., Golden S. S., Kondo T., Ishiura M., Johnson C. H. (1995) **Bacterial luciferase as a reporter of circadian gene expression in cyanobacteria**. *Journal of bacteriology*, **177**(8):2080-2086.
214. Ito H., Mutsuda M., Murayama Y., Tomita J., Hosokawa N., Terauchi K., Sugita C., Sugita M., Kondo T., Iwasaki H. (2009) **Cyanobacterial daily life with Kai-based circadian and diurnal genome-wide transcriptional control in Synechococcus elongatus**. *Proceedings of the National Academy of Sciences*, **106**(33):14168-14173.
215. Vijayan V., Zuzow R., O'Shea E. K. (2009) **Oscillations in supercoiling drive circadian gene expression in cyanobacteria**. *Proceedings of the National Academy of Sciences of the United States of America*, **106**(52):22564-22568.
216. Markson J. S., Piechura J. R., Puszynska A. M., O'Shea E. K. (2013) **Circadian control of global gene expression by the cyanobacterial master regulator RpaA**. *Cell*, **155**(6):1396-1408.
217. Takai N., Nakajima M., Oyama T., Kito R., Sugita C., Sugita M., Kondo T., Iwasaki H. (2006) **A KaiC-associating SasA-RpaA two-component regulatory system as a major circadian timing mediator in cyanobacteria**. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(32):12109-12114.
218. Gutu A., O'Shea E. K. (2013) **Two antagonistic clock-regulated histidine kinases time the activation of circadian gene expression**. *Mol Cell*, **50**(2):288-294.
219. Smith R. M., Williams S. B. (2006) **Circadian rhythms in gene transcription imparted by chromosome compaction in the cyanobacterium Synechococcus elongatus**. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(22):8564-8569.
220. Espinosa J., Boyd J. S., Cantos R., Salinas P., Golden S. S., Contreras A. (2015) **Cross-talk and regulatory interactions between the essential response regulator RpaB and cyanobacterial circadian clock output**. *Proceedings of the National Academy of Sciences of the United States of America*, **112**(7):2198-2203.
221. Hanaoka M., Takai N., Hosokawa N., Fujiwara M., Akimoto Y., Kobori N., Iwasaki H., Kondo T., Tanaka K. (2012) **RpaB, another response regulator operating circadian clock-dependent transcriptional regulation in Synechococcus elongatus PCC 7942**. *The Journal of biological chemistry*, **287**(31):26321-26327.

222. Nair U., Ditty J. L., Min H., Golden S. S. (2002) **Roles for sigma factors in global circadian regulation of the cyanobacterial genome**. *Journal of bacteriology*, **184**(13):3530-3538.
223. Mori T., Binder B., Johnson C. H. (1996) **Circadian gating of cell division in cyanobacteria growing with average doubling times of less than 24 hours**. *Proceedings of the National Academy of Sciences of the United States of America*, **93**(19):10183-10188.
224. Yang Q., Pando B. F., Dong G., Golden S. S., van Oudenaarden A. (2010) **Circadian gating of the cell cycle revealed in single cyanobacterial cells**. *Science (New York, NY)*, **327**(5972):1522-1526.
225. Dong G., Yang Q., Wang Q., Kim Y. I., Wood T. L., Osteryoung K. W., van Oudenaarden A., Golden S. S. (2010) **Elevated ATPase activity of KaiC applies a circadian checkpoint on cell division in Synechococcus elongatus**. *Cell*, **140**(4):529-539.
226. Hastings J. W. (2007) **The Gonyaulax clock at 50: translational control of circadian expression**. *Cold Spring Harb Symp Quant Biol*, **72**:141-144.
227. Ueda H. R., Matsumoto A., Kawamura M., Iino M., Tanimura T., Hashimoto S. (2002) **Genome-wide transcriptional orchestration of circadian rhythms in Drosophila**. *The Journal of biological chemistry*, **277**(16):14048-14052.
228. Woolum J. C. (1991) **A re-examination of the role of the nucleus in generating the circadian rhythm in Acetabularia**. *Journal of biological rhythms*, **6**(2):129-136.
229. Lakin-Thomas P. L. (2006) **Transcriptional feedback oscillators: maybe, maybe not**. *Journal of biological rhythms*, **21**(2):83-92.
230. Poynton R. A., Hampton M. B. (2014) **Peroxiredoxins as biomarkers of oxidative stress**. *Biochimica et biophysica acta*, **1840**(2):906-912.
231. Jeong W., Bae S. H., Toledano M. B., Rhee S. G. (2012) **Role of sulfiredoxin as a regulator of peroxiredoxin function and regulation of its expression**. *Free radical biology & medicine*, **53**(3):447-456.
232. O'Neill J. S., Reddy A. B. (2011) **Circadian clocks in human red blood cells**. *Nature*, **469**(7331):498-503.
233. Kil I. S., Ryu K. W., Lee S. K., Kim J. Y., Chu S. Y., Kim J. H., Park S., Rhee S. G. (2015) **Circadian Oscillation of Sulfiredoxin in the Mitochondria**. *Mol Cell*, **59**(4):651-663.
234. Reddy A. B., Rey G. (2014) **Metabolic and nontranscriptional circadian clocks: eukaryotes**. *Annual review of biochemistry*, **83**:165-189.
235. Okamoto O. K., Robertson D. L., Fagan T. F., Hastings J. W., Colepicolo P. (2001) **Different regulatory mechanisms modulate the expression of a dinoflagellate iron-superoxide dismutase**. *The Journal of biological chemistry*, **276**(23):19989-19993.
236. Yoshida Y., Maeda T., Lee B., Hasunuma K. (2008) **Conidiation rhythm and light entrainment in superoxide dismutase mutant in Neurospora crassa**. *Molecular genetics and genomics : MGG*, **279**(2):193-202.
237. Martin V., Sainz R. M., Mayo J. C., Antolin I., Herrera F., Rodriguez C. (2003) **Daily rhythm of gene expression in rat superoxide dismutases**. *Endocrine research*, **29**(1):83-95.

238. Rutter J., Reick M., Wu L. C., McKnight S. L. (2001) **Regulation of clock and NPAS2 DNA binding by the redox state of NAD cofactors**. *Science (New York, NY)*, **293**(5529):510-514.
239. Feeney K. A., Hansen L. L., Putker M., Olivares-Yañez C., Day J., Eades L. J., Larrondo L. F., Hoyle N. P., O'Neill J. S., van Ooijen G. (2016) **Daily magnesium fluxes regulate cellular timekeeping and energy balance**. *Nature*, **532**(7599):375-379.
240. Querfurth C., Diernfellner A., Heise F., Lauinger L., Neiss A., Tataroglu O., Brunner M., Schafmeier T. (2007) **Posttranslational regulation of Neurospora circadian clock by CK1a-dependent phosphorylation**. *Cold Spring Harb Symp Quant Biol*, **72**:177-183.
241. Maywood E. S., Chesham J. E., Smyllie N. J., Hastings M. H. (2014) **The Tau mutation of casein kinase 1epsilon sets the period of the mammalian pacemaker via regulation of Period1 or Period2 clock proteins**. *Journal of biological rhythms*, **29**(2):110-118.
242. Tamaru T., Hirayama J., Isojima Y., Nagai K., Norioka S., Takamatsu K., Sassone-Corsi P. (2009) **CK2alpha phosphorylates BMAL1 to regulate the mammalian clock**. *Nature structural & molecular biology*, **16**(4):446-448.
243. Szabo A., Papin C., Zorn D., Ponien P., Weber F., Raabe T., Rouyer F. (2013) **The CK2 kinase stabilizes CLOCK and represses its activity in the Drosophila circadian oscillator**. *PLoS biology*, **11**(8):e1001645.
244. Portoles S., Mas P. (2010) **The functional interplay between protein kinase CK2 and CCA1 transcriptional activity is essential for clock temperature compensation in Arabidopsis**. *PLoS genetics*, **6**(11):e1001201.
245. Sahar S., Zocchi L., Kinoshita C., Borrelli E., Sassone-Corsi P. (2010) **Regulation of BMAL1 protein stability and circadian function by GSK3beta-mediated phosphorylation**. *PLoS One*, **5**(1):e8561.
246. Top D., Harms E., Syed S., Adams E. L., Saez L. (2016) **GSK-3 and CK2 Kinases Converge on Timeless to Regulate the Master Clock**. *Cell reports*, **16**(2):357-367.
247. Litchfield D. W. (2003) **Protein kinase CK2: structure, regulation and role in cellular decisions of life and death**. *Biochemical Journal*, **369**(Pt 1):1-15.
248. Schittek B., Sinnberg T. (2014) **Biological functions of casein kinase 1 isoforms and putative roles in tumorigenesis**. *Molecular Cancer*, **13**.
249. Bertomeu T. (2007) **Identification et caractérisation de candidats régulateurs du cycle cellulaire chez le dinoflagellé Lingulodinium polyedrum**. Montréal: Unversité de Montréal.
250. Roenneberg T., Morse D. (1993) **Two circadian oscillators in one cell**. *Nature*, **362**(6418):362-364.
251. Von der Heyde F., Wilkens A., Rensing L. (1992) **The Effects of Temperature on the Circadian Rhythms of Flashing and Glow in Gonyaulax polyedra: Are the Two Rhythms Controlled by Two Oscillators?** *Journal of biological rhythms*, **7**(2):115-123.
252. Morse D., Hastings J. W., Roenneberg T. (1994) **Different phase responses of the two circadian oscillators in Gonyaulax**. *Journal of biological rhythms*, **9**(3-4):263-274.

253. Milos P., Morse D., Hastings J. W. (1990) **Circadian control over synthesis of many *Gonyaulax* proteins is at a translational level**. *Naturwissenschaften*, **77**:87-89.
254. Okamoto O. K., Hastings J. W. (2003) **Novel Dinoflagellate Clock-Related Genes Identified Through Microarray Analysis**. *Journal of Phycology*, **39**(3):519-526.
255. Van Dolah F. M., Lidie K. B., Morey J. S., Brunelle S. A., Ryan J. C., Monroe E. A., Haynes B. L. (2007) **Microarray analysis of diurnal- and circadian-regulated in the Florida red-tide dinoflagellate Karenia brevis (Dinophyceae)**. *Journal of Phycology*, **43**(4):741-752.
256. Chauton M. S., Winge P., Brembu T., Vadstein O., Bones A. M. (2013) **Gene regulation of carbon fixation, storage, and utilization in the diatom Phaeodactylum tricornutum acclimated to light/dark cycles**. *Plant physiology*, **161**(2):1034-1048.
257. Taylor W., Krasnow R., Dunlap J. C., Broda H., Hastings J. W. (1982) **Critical pulses of anisomycin drive the circadian oscillator in Gonyaulax towards its singularity**. *Journal of comparative physiology*, **148**(1):11-25.
258. Taylor W. R., Dunlap J. C., Hastings J. W. (1982) **Inhibitors of protein synthesis on 80S ribosomes phase shift the Gonyaulax clock**. *The Journal of experimental biology*, **97**:121-136.
259. Dunlap J. C., Taylor W. R., Hastings J. W. (1980) **The effects of protein synthesis inhibitors on the Gonyaulax clock I: Phase shifting effects of cycloheximide**. *J Comp Physiol*, **138**:1-8.
260. Dagenais-Bellefeuille S., Bertomeu T., Morse D. (2008) **S-phase and M-phase timing are under independent circadian control in the dinoflagellate Lingulodinium**. *Journal of biological rhythms*, **23**(5):400-408.
261. Mizoguchi T., Putterill J., Ohkoshi Y.: **Kinase and Phosphatase: The Cog and Spring of the Circadian Clock**. In: *International Review of Cytology.* Edited by Kwang WJ, vol. Volume 250: Academic Press; 2006: 47-72.
262. Shi X., Li L., Guo C., Lin X., Li M., Lin S. (2015) **Rhodopsin gene expression regulated by the light dark cycle, light spectrum and light intensity in the dinoflagellate Prorocentrum**. *Frontiers in microbiology*, **6**:555.
263. Guo Z., Zhang H., Lin S. (2014) **Light-promoted rhodopsin expression and starvation survival in the marine dinoflagellate Oxyrrhis marina**. *PLoS One*, **9**(12):e114941.
264. Slamovits C. H., Okamoto N., Burri L., James E. R., Keeling P. J. (2011) **A bacterial proteorhodopsin proton pump in marine eukaryotes**. *Nature communications*, **2**:183.
265. Eisensamer B., Roenneberg T. (2004) **Extracellular pH is under circadian control in Gonyaulax polyedra and forms a metabolic feedback loop**. *Chronobiology international*, **21**(1):27-41.
266. Roenneberg T., Rehman J. (1996) **Nitrate, a nonphotic signal for the circadian system**. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology*, **10**(12):1443-1447.
267. Buskey E., Mills L., Swift E. (1983) **The effects of dinoflagellate bioluminescence on the swimming behavior of a marine copepod**. *Limnology and Oceanography*, **28**(3):575-579.

268. Mesinger A. F., Case J. F. (1992) **Dinoflagellate luminescence increases susceptibility of zooplankton to teleost predation**. *Marine Biology*, **112**(2):207-210.
269. Fleisher K. J., Case J. F. (1995) **Cephalopod Predation Facilitated by Dinoflagellate Luminescence**. *The Biological Bulletin*, **189**(3):263-271.
270. Dunlap J. C., Hastings J. W., Shimomura O. (1980) **Crossreactivity between the light-emitting systems of distantly related organisms: Novel type of light-emitting compound**. *Proceedings of the National Academy of Sciences of the United States of America*, **77**(3):1394-1397.
271. Nicolas M. T., Johnson C. H., Bassot J. M., Hastings J. W. (1985) **Immunogold labeling of organelles in the bioluminescent dinoflagellate Gonyaulax polyedra with anti-luciferase antibody**. *Cell biology international reports*, **9**(9):797-802.
272. Desjardins M., Morse D. (1993) **The polypeptide components of scintillons, the bioluminescence organelles of the dinoflagellate Gonyaulax polyedra**. *Biochem Cell Biol*, **71**(3-4):176-182.
273. Morse D., Pappenheimer A. M., Jr., Hastings J. W. (1989) **Role of a luciferin-binding protein in the circadian bioluminescent reaction of Gonyaulax polyedra**. *The Journal of biological chemistry*, **264**(20):11822-11826.
274. Li L., Hong R., Hastings J. W. (1997) **Three functional luciferase domains in a single polypeptide chain**. *Proceedings of the National Academy of Sciences of the United States of America*, **94**(17):8954-8958.
275. Valiadi M., Debora Iglesias-Rodriguez M., Amorim A. (2012) **Distribution and genetic diversity of the luciferase gene within marine dinoflagellates**. *Journal of Phycology*, **48**(3):826-836.
276. Liu L., Hastings J. W. (2007) **Two different domains of the luciferase gene in the heterotrophic dinoflagellate Noctiluca scintillans occur as two separate genes in photosynthetic species**. *Proceedings of the National Academy of Sciences of the United States of America*, **104**(3):696-701.
277. Valiadi M., Iglesias-Rodriguez M. D. (2014) **Diversity of the Luciferin Binding Protein Gene in Bioluminescent Dinoflagellates – Insights from a New Gene in Noctiluca scintillans and Sequences from Gonyaulacoid Genera**. *Journal of Eukaryotic Microbiology*, **61**(2):134-145.
278. Smith S. M. E., Morgan D., Musset B., Cherny V. V., Place A. R., Hastings J. W., DeCoursey T. E. (2011) **Voltage-gated proton channel in a dinoflagellate**. *Proceedings of the National Academy of Sciences*, **108**(44):18162-18167.
279. Morse D., Milos P. M., Roux E., Hastings J. W. (1989) **Circadian regulation of bioluminescence in *Gonyaulax* involves translational control**. *Proc Natl Acad Sci USA*, **86**(1):172-176.
280. Fritz L., Morse D., Hastings J. W. (1990) **The circadian bioluminescence rhythm of Gonyaulax is related to daily variations in the number of light-emitting organelles**. *Journal of Cell Science*, **95**(2):321-328.
281. Mittag M., Li L., Hastings J. W. (1998) **The mRNA level of the circadian regulated Gonyaulax luciferase remains constant over the cycle**. *Chronobiology international*, **15**(1):93-98.
282. Mittag M. (1996) **Conserved circadian elements in phylogenetically diverse algae**. *Proceedings of the National Academy of Sciences*, **93**(25):14401-14404.

283. Dathe H., Prager K., Mittag M. (2012) **Novel interaction of two clock-relevant RNA-binding proteins C3 and XRN1 in Chlamydomonas reinhardtii**. *FEBS letters*, **586**(22):3969-3973.
284. Iliev D., Voytsekh O., Schmidt E. M., Fiedler M., Nykytenko A., Mittag M. (2006) **A heteromeric RNA-binding protein is involved in maintaining acrophase and period of the circadian clock**. *Plant Physiol*, **142**(2):797-806.
285. Fritz L., Milos P., Morse D., Hastings J. W. (1991) **In situ hybridization of luciferin-binding protein anti-sense rna to thin sections of the bioluminescent dinoflagellate gonyaulax polyedra**. *Journal of Phycology*, **27**(3):436-441.
286. Techel D., Chuang E., Lee D. H., Comolli J., Hastings J. W. (1996) **Changes in the ultrastructural localization of mRNA for a circadian regulated protein**. *Molecular marine biology and biotechnology*, **5**(4):241-248.
287. Samuelsson G., Sweeney B. M., Matlick H. A., Prézelin B. B. (1983) **Changes in Photosystem II Account for the Circadian Rhythm in Photosynthesis in Gonyaulax polyedra**. *Plant physiology*, **73**(2):329-331.
288. Hollnagel H. C., Pinto E., Morse D., Colepicolo P. (2002) **The Oscillation of Photosynthetic Capacity in Lingulodinium polyedrum is not related to differences in RuBisCo, Peridinin or Chlorophyll a Amounts**. *Biological Rhythm Research*, **33**(4):443-458.
289. Nassoury N., Wang Y., Morse D. (2005) **Brefeldin a inhibits circadian remodeling of chloroplast structure in the dinoflagellate gonyaulax**. *Traffic (Copenhagen, Denmark)*, **6**(7):548-561.
290. Rensing L., Taylor W., Dunlap J., Hastings J. W. (1980) **The effects of protein synthesis inhibitors on the Gonyaulax clock**. *Journal of comparative physiology*, **138**(1):9-18.
291. Bush K. J., Sweeney B. M. (1972) **The Activity of Ribulose Diphosphate Carboxylase in Extracts of Gonyaulax polyedra in the Day and the Night Phases of the Circadian Rhythm of Photosynthesis**. *Plant physiology*, **50**(4):446-451.
292. Homma K., Hastings J. W. (1989) **The s phase is discrete and is controlled by the circadian clock in the marine dinoflagellate Gonyaulax polyedra**. *Experimental cell research*, **182**(2):635-644.
293. Goto K., Johnson C. H. (1995) **Is the cell division cycle gated by a circadian clock? The case of Chlamydomonas reinhardtii**. *The Journal of cell biology*, **129**(4):1061-1069.
294. Eppley R. W., Holm-Harisen O., Strickland J. D. H. (1968) **Some observations on the vertical migration ofdinoflagellates**. *Journal of Phycology*, **4**(4):333-340.
295. Roenneberg T., Colfax G. N., Hastings J. W. (1989) **A circadian rhythm of population behavior in Gonyaulax polyedra**. *Journal of biological rhythms*, **4**(2):201-216.
296. Nassoury N., Hastings J. W., Morse D.: **Biochemistry and circadian regulation of output from the *Gonyaulax* clock: are there many clocks or simply many hands?** In: *The circadian clock in eukaryotic microbes.* Edited by Kippert F: Kluwer Academic and Plenum; 2005.
297. Ramalho C. B., Hastings J. W., Colepicolo P. (1995) **Circadian oscillation of nitrate reductase activity in Gonyaulax polyedra is due to changes in cellular protein levels**. *Plant physiology*, **107**(1):225-231.

298. Dagenais Bellefeuille S., Morse D. (2016) **The main nitrate transporter of the dinoflagellate Lingulodinium polyedrum is constitutively expressed and not responsible for daily variations in nitrate uptake rates**. *Harmful Algae*, **55**:272-281.
299. Colepicolo P., Camarero V. C. P. C., Hastings J. W. (1992) **A Circadian Rhythm in the Activity of Superoxide Dismutase in the Photosynthetic Alga Gonyaulax polyedra**. *Chronobiology international*, **9**(4):266-268.
300. Akimoto H., Kinumi T., Ohmiya Y. (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.
301. Bell-Pedersen D., Cassone V. M., Earnest D. J., Golden S. S., Hardin P. E., Thomas T. L., Zoran M. J. (2005) **Circadian rhythms from multiple oscillators: lessons from diverse organisms**. *Nat Rev Genet*, **6**(7):544-556.
302. Johnson C. H., Roeber J. F., Hastings J. W. (1984) **Circadian changes in enzyme concentration account for rhythm of enzyme activity in *Gonyaulax***. *Science (New York, NY)*, **223**:1428-1430.
303. Fagan T., Morse D., Hastings J. (1999) **Circadian synthesis of a nuclear encoded chloroplast Glyceraldehyde-3-phosphate dehydrogenase in the dinoflagellate *Gonyaulax polyedra* is translationally controlled.** *Biochemistry*, **38**:7689-7695.
304. Roy S., Letourneau L., Morse D. (2014) **Cold-induced cysts of the photosynthetic dinoflagellate Lingulodinium polyedrum have an arrested circadian bioluminescence rhythm and lower levels of protein phosphorylation**. *Plant physiology*, **164**(2):966-977.
305. Mortazavi A., Williams B. A., McCue K., Schaeffer L., Wold B. (2008) **Mapping and quantifying mammalian transcriptomes by RNA-Seq**. *Nat Methods*, **5**(7):621-628.
306. Anders S., Huber W. (2010) **Differential expression analysis for sequence count data**. *Genome biology*, **11**(10):R106.
307. Robinson M. D., McCarthy D. J., Smyth G. K. (2010) **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data**. *Bioinformatics (Oxford, England)*, **26**(1):139-140.
308. Chaput H., Wang Y., Morse D. (2002) **Polyadenylated transcripts containing random gene fragments are expressed in dinoflagellate mitochondria**. *Protist*, **153**(2):111-122.
309. Wang Y., Jensen L., Hojrup P., Morse D. (2005) **Synthesis and degradation of dinoflagellate plastid-encoded psbA proteins are light-regulated, not circadian-regulated**. *Proc Natl Acad Sci USA*, **102**(8):2844-2849.
310. Sobell H. M. (1985) **Actinomycin and DNA transcription**. *Proceedings of the National Academy of Sciences of the United States of America*, **82**(16):5328-5331.
311. Muller W. E., Seibert G., Beyer R., Breter H. J., Maidhof A., Zahn R. K. (1977) **Effect of cordycepin on nucleic acid metabolism in L5178Y cells and on nucleic acid-synthesizing enzyme systems**. *Cancer research*, **37**(10):3824-3833.
312. Lakin-Thomas P. L. (2006) **Circadian clock genes frequency and white collar-1 are not essential for entrainment to temperature cycles in Neurospora crassa**. *Proceedings of the National Academy of Sciences of the United States of America*, **103**(12):4469-4474.
313. Rizzo P. (1991) **The enigma of the dinoflagellate chromosome.** *J Protozool*, **38**:246-252.

314. Rossini C., Taylor W., Fagan T., Hastings J. W. (2003) **Lifetimes of mRNAs for clock regulated proteins in a dinoflagellate.** *Chronobiol Int*, **20**:963-976.
315. Karakashian M. W., Hastings J. W. (1963) **The effects of inhibitors of macromolecular biosynthesis upon the persistent rhythm of luminescence in *Gonyaulax***. *J Gen Physiol*, **47**:1-12.
316. Allada R., Meissner R. A. (2005) **Casein kinase 2, circadian clocks, and the flight from mutagenic light**. *Molecular and cellular biochemistry*, **274**(1-2):141-149.
317. Morf J., Rey G., Schneider K., Stratmann M., Fujita J., Naef F., Schibler U. (2012) **Cold-inducible RNA-binding protein modulates circadian gene expression posttranscriptionally**. *Science (New York, NY)*, **338**(6105):379-383.
318. Zhang Y., Ling J., Yuan C., Dubruille R., Emery P. (2013) **A role for Drosophila ATX2 in activation of PER translation and circadian behavior**. *Science (New York, NY)*, **340**(6134):879-882.
319. Lim C., Allada R. (2013) **ATAXIN-2 activates PERIOD translation to sustain circadian rhythms in Drosophila**. *Science (New York, NY)*, **340**(6134):875-879.
320. Tarun S. J., Sachs A. (1996) **Association of the yeast poly(A) tail binding protein with translation initiation factor eIF-4G.** *EMBO J*, **15**:7168-7177.
321. Kojima S., Sher-Chen E. L., Green C. B. (2012) **Circadian control of mRNA polyadenylation dynamics regulates rhythmic protein expression**. *Genes & development*, **26**(24):2724-2736.
322. Huang Y., Ainsley J. A., Reijmers L. G., Jackson F. R. (2013) **Translational profiling of clock cells reveals circadianly synchronized protein synthesis**. *PLoS biology*, **11**(11):e1001703.
323. Li H., Durbin R. (2009) **Fast and accurate short read alignment with Burrows-Wheeler transform**. *Bioinformatics (Oxford, England)*, **25**(14):1754-1760.
324. Mackenzie T. D., Morse D. (2011) **Circadian photosynthetic reductant flow in the dinoflagellate Lingulodinium is limited by carbon availability**. *Plant Cell Environ*.
325. Graumann P. L., Marahiel M. A. (1998) **A superfamily of proteins that contain the cold-shock domain**. *Trends in biochemical sciences*, **23**(8):286-290.
326. Didier D. K., Schiffenbauer J., Woulfe S. L., Zacheis M., Schwartz B. D. (1988) **Characterization of the cDNA encoding a protein binding to the major histocompatibility complex class II Y box**. *Proceedings of the National Academy of Sciences of the United States of America*, **85**(19):7322-7326.
327. Brandi A., Pon C. L., Gualerzi C. O. (1994) **Interaction of the main cold shock protein CS7.4 (CspA) of Escherichia coli with the promoter region of hns**. *Biochimie*, **76**(10-11):1090-1098.
328. Phadtare S., Inouye M. (1999) **Sequence-selective interactions with RNA by CspB, CspC and CspE, members of the CspA family of Escherichia coli**. *Molecular microbiology*, **33**(5):1004-1014.
329. Jones P. G., VanBogelen R. A., Neidhardt F. C. (1987) **Induction of proteins in response to low temperature in Escherichia coli**. *Journal of bacteriology*, **169**(5):2092-2095.
330. Gualerzi C. O., Giuliodori A. M., Pon C. L. (2003) **Transcriptional and post-transcriptional control of cold-shock genes**. *Journal of molecular biology*, **331**(3):527-539.

331. Goldstein J., Pollitt N. S., Inouye M. (1990) **Major cold shock protein of Escherichia coli**. *Proceedings of the National Academy of Sciences of the United States of America*, **87**(1):283-287.
332. Xia B., Ke H., Inouye M. (2001) **Acquirement of cold sensitivity by quadruple deletion of the cspA family and its suppression by PNPase S1 domain in Escherichia coli**. *Molecular microbiology*, **40**(1):179-188.
333. Jones P. G., Krah R., Tafuri S. R., Wolffe A. P. (1992) **DNA gyrase, CS7.4, and the cold shock response in Escherichia coli**. *Journal of bacteriology*, **174**(18):5798-5802.
334. Bae W., Xia B., Inouye M., Severinov K. (2000) **Escherichia coli CspA-family RNA chaperones are transcription antiterminators**. *Proceedings of the National Academy of Sciences of the United States of America*, **97**(14):7784-7789.
335. Goldenberg D., Azar I., Oppenheim A. B. (1996) **Differential mRNA stability of the cspA gene in the cold-shock response of Escherichia coli**. *Molecular microbiology*, **19**(2):241-248.
336. Karlson D., Imai R. (2003) **Conservation of the cold shock domain protein family in plants**. *Plant physiology*, **131**(1):12-15.
337. Sasaki K., Imai R. (2011) **Pleiotropic roles of cold shock domain proteins in plants**. *Front Plant Sci*, **2**:116.
338. Kim M. H., Sasaki K., Imai R. (2009) **Cold shock domain protein 3 regulates freezing tolerance in Arabidopsis thaliana**. *The Journal of biological chemistry*, **284**(35):23454-23460.
339. Park S. J., Kwak K. J., Oh T. R., Kim Y. O., Kang H. (2009) **Cold shock domain proteins affect seed germination and growth of Arabidopsis thaliana under abiotic stress conditions**. *Plant & cell physiology*, **50**(4):869-878.
340. Juntawong P., Sorenson R., Bailey-Serres J. (2013) **Cold shock protein 1 chaperones mRNAs during translation in Arabidopsis thaliana**. *The Plant journal : for cell and molecular biology*, **74**(6):1016-1028.
341. Kim J. S., Park S. J., Kwak K. J., Kim Y. O., Kim J. Y., Song J., Jang B., Jung C. H., Kang H. (2007) **Cold shock domain proteins and glycine-rich RNA-binding proteins from Arabidopsis thaliana can promote the cold adaptation process in Escherichia coli**. *Nucleic acids research*, **35**(2):506-516.
342. Yang Y., Karlson D. T. (2011) **Overexpression of AtCSP4 affects late stages of embryo development in Arabidopsis**. *Journal of experimental botany*, **62**(6):2079-2091.
343. Fusaro A. F., Bocca S. N., Ramos R. L., Barroco R. M., Magioli C., Jorge V. C., Coutinho T. C., Rangel-Lima C. M., De Rycke R., Inze D., Engler G., Sachetto-Martins G. (2007) **AtGRP2, a cold-induced nucleo-cytoplasmic RNA-binding protein, has a role in flower and seed development**. *Planta*, **225**(6):1339-1351.
344. Chaikam V., Karlson D. (2008) **Functional characterization of two cold shock domain proteins from Oryza sativa**. *Plant Cell Environ*, **31**(7):995-1006.
345. Evdokimova V., Ruzanov P., Imataka H., Raught B., Svitkin Y., Ovchinnikov L. P., Sonenberg N. (2001) **The major mRNA-associated protein YB-1 is a potent 5' cap-dependent mRNA stabilizer**. *The EMBO journal*, **20**(19):5491-5502.
346. Evdokimova V., Tognon C., Ng T., Ruzanov P., Melnyk N., Fink D., Sorokin A., Ovchinnikov L. P., Davicioni E., Triche T. J., Sorensen P. H. (2009) **Translational**

**activation of snail1 and other developmentally regulated transcription factors by YB-1 promotes an epithelial-mesenchymal transition**. *Cancer cell*, **15**(5):402-415.
347. Bouvet P., Wolffe A. P. (1994) **A role for transcription and FRGY2 in masking maternal mRNA within Xenopus oocytes**. *Cell*, **77**(6):931-941.
348. Karlson D., Nakaminami K., Toyomasu T., Imai R. (2002) **A cold-regulated nucleic acid-binding protein of winter wheat shares a domain with bacterial cold shock proteins**. *The Journal of biological chemistry*, **277**(38):35248-35256.
349. Phadtare S. (2004) **Recent developments in bacterial cold-shock response**. *Current issues in molecular biology*, **6**(2):125-136.
350. Etchegaray J. P., Inouye M. (1999) **CspA, CspB, and CspG, major cold shock proteins of Escherichia coli, are induced at low temperature under conditions that completely block protein synthesis**. *Journal of bacteriology*, **181**(6):1827-1830.
351. Phadtare S., Inouye M. (2004) **Genome-wide transcriptional analysis of the cold shock response in wild-type and cold-sensitive, quadruple-csp-deletion strains of Escherichia coli**. *Journal of bacteriology*, **186**(20):7007-7014.
352. Sachs R., Max K. E., Heinemann U., Balbach J. (2012) **RNA single strands bind to a conserved surface of the major cold shock protein in crystals and solution**. *Rna*, **18**(1):65-76.
353. Wistow G. (1990) **Cold shock and DNA binding**. *Nature*, **344**(6269):823-824.
354. Wolffe A. P., Tafuri S., Ranjan M., Familari M. (1992) **The Y-box factors: a family of nucleic acid binding proteins conserved from Escherichia coli to man**. *The New biologist*, **4**(4):290-298.
355. Wang S., Hirschberg R. (2011) **Y-box protein-1 is a transcriptional regulator of BMP7**. *Journal of cellular biochemistry*, **112**(4):1130-1137.
356. Lasham A., Moloney S., Hale T., Homer C., Zhang Y. F., Murison J. G., Braithwaite A. W., Watson J. (2003) **The Y-box-binding protein, YB1, is a potential negative regulator of the p53 tumor suppressor**. *The Journal of biological chemistry*, **278**(37):35516-35523.
357. Tafuri S. R., Wolffe A. P. (1990) **Xenopus Y-box transcription factors: molecular cloning, functional analysis and developmental regulation**. *Proceedings of the National Academy of Sciences of the United States of America*, **87**(22):9028-9032.
358. Izumi H., Imamura T., Nagatani G., Ise T., Murakami T., Uramoto H., Torigoe T., Ishiguchi H., Yoshida Y., Nomoto M., Okamoto T., Uchiumi T., Kuwano M., Funa K., Kohno K. (2001) **Y box-binding protein-1 binds preferentially to single-stranded nucleic acids and exhibits 3'-->5' exonuclease activity**. *Nucleic acids research*, **29**(5):1200-1207.
359. Kim M. H., Sonoda Y., Sasaki K., Kaminaka H., Imai R. (2013) **Interactome analysis reveals versatile functions of Arabidopsis COLD SHOCK DOMAIN PROTEIN 3 in RNA processing within the nucleus and cytoplasm**. *Cell stress & chaperones*, **18**(4):517-525.
360. Sommerville J. (1999) **Activities of cold-shock domain proteins in translation control**. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **21**(4):319-325.
361. Roy S., Beauchemin M., Dagenais-Bellefeuille S., Letourneau L., Cappadocia M., Morse D. (2014) **The Lingulodinium circadian system lacks rhythmic changes in transcript abundance**. *BMC biology*, **12**:107.

362. Guillard R. R. L., Ryther J. H. (1962) **Studies on marine planktonic diatoms: *Cyclotella nana* Hufstedt and *Denotula confervacea* (Cleve) Gran.** *Canadian Journal Microbiology*, **8**:229-239.
363. Keeling P. J., Burki F., Wilcox H. M., Allam B., Allen E. E., Amaral-Zettler L. A., Armbrust E. V., Archibald J. M., Bharti A. K., Bell C. J., Beszteri B., Bidle K. D., Cameron C. T., Campbell L., Caron D. A., Cattolico R. A., Collier J. L., Coyne K., Davy S. K., Deschamps P., Dyhrman S. T., Edvardsen B., Gates R. D., Gobler C. J., Greenwood S. J., Guida S. M., Jacobi J. L., Jakobsen K. S., James E. R., Jenkins B., John U., Johnson M. D., Juhl A. R., Kamp A., Katz L. A., Kiene R., Kudryavtsev A., Leander B. S., Lin S., Lovejoy C., Lynn D., Marchetti A., McManus G., Nedelcu A. M., Menden-Deuer S., Miceli C., Mock T., Montresor M., Moran M. A., Murray S., Nadathur G., Nagai S., Ngam P. B., Palenik B., Pawlowski J., Petroni G., Piganeau G., Posewitz M. C., Rengefors K., Romano G., Rumpho M. E., Rynearson T., Schilling K. B., Schroeder D. C., Simpson A. G. B., Slamovits C. H., Smith D. R., Smith G. J., Smith S. R., Sosik H. M., Stief P., Theriot E., Twary S. N., Umale P. E., Vaulot D., Wawrik B., Wheeler G. L., Wilson W. H., Xu Y., Zingone A., Worden A. Z. (2014) **The Marine Microbial Eukaryote Transcriptome Sequencing Project (MMETSP): Illuminating the Functional Diversity of Eukaryotic Life in the Oceans through Transcriptome Sequencing**. *PLoS biology*, **12**(6):e1001889.
364. Dereeper A., Guignon V., Blanc G., Audic S., Buffet S., Chevenet F., Dufayard J. F., Guindon S., Lefort V., Lescot M., Claverie J. M., Gascuel O. (2008) **Phylogeny.fr: robust phylogenetic analysis for the non-specialist**. *Nucleic acids research*, **36**(Web Server issue):W465-469.
365. Candiano G., Bruschi M., Musante L., Santucci L., Ghiggeri G. M., Carnemolla B., Orecchia P., Zardi L., Righetti P. G. (2004) **Blue silver: A very sensitive colloidal Coomassie G-250 staining for proteome analysis**. *ELECTROPHORESIS*, **25**(9):1327-1333.
366. Levi-Setti R., Gavrilov K. L., Rizzo P. J. (2008) **Divalent cation distribution in dinoflagellate chromosomes imaged by high-resolution ion probe mass spectrometry**. *European journal of cell biology*, **87**(12):963-976.
367. Soyer M.-O., Haapala O. K. (1974) **Structural changes of dinoflagellate chromosomes by pronase and ribonuclease**. *Chromosoma*, **47**(2):179-192.
368. Herzog M., Soyer M. O., Daney de Marcillac G. (1982) **A high level of thymine replacement by 5-hydroxymethyluracil in nuclear DNA of the primitive dinoflagellate Prorocentrum micans E**. *European journal of cell biology*, **27**(2):151-155.
369. Chan Y. H., Kwok A. C., Tsang J. S., Wong J. T. (2006) **Alveolata histone-like proteins have different evolutionary origins**. *Journal of evolutionary biology*, **19**(5):1717-1721.
370. Janouskovec J., Gavelis G. S., Burki F., Dinh D., Bachvaroff T. R., Gornik S. G., Bright K. J., Imanian B., Strom S. L., Delwiche C. F., Waller R. F., Fensome R. A., Leander B. S., Rohwer F. L., Saldarriaga J. F. (2017) **Major transitions in dinoflagellate evolution unveiled by phylotranscriptomics**. *Proceedings of the National Academy of Sciences of the United States of America*, **114**(2):E171-e180.
371. Marinov G. K., Lynch M. (2016) **Diversity and Divergence of Dinoflagellate Histone Proteins**. *G3: Genes|Genomes|Genetics*, **6**(2):397-422.

372. Kustatscher G., Wills K. L., Furlan C., Rappsilber J. (2014) **Chromatin enrichment for proteomics**. *Nature protocols*, **9**(9):2090-2099.
373. Sikorskaite S., Rajamaki M. L., Baniulis D., Stanys V., Valkonen J. P. (2013) **Protocol: Optimised methodology for isolation of nuclei from leaves of species in the Solanaceae and Rosaceae families**. *Plant methods*, **9**:31.
374. Camacho C., Coulouris G., Avagyan V., Ma N., Papadopoulos J., Bealer K., Madden T. L. (2009) **BLAST+: architecture and applications**. *BMC bioinformatics*, **10**:421.
375. The-Gene-Ontology-Consortium (2015) **Gene Ontology Consortium: going forward**. *Nucleic acids research*, **43**(Database issue):D1049-1056.
376. Gotz S., Garcia-Gomez J. M., Terol J., Williams T. D., Nagaraj S. H., Nueda M. J., Robles M., Talon M., Dopazo J., Conesa A. (2008) **High-throughput functional annotation and data mining with the Blast2GO suite**. *Nucleic acids research*, **36**(10):3420-3435.
377. Marchler-Bauer A., Derbyshire M. K., Gonzales N. R., Lu S., Chitsaz F., Geer L. Y., Geer R. C., He J., Gwadz M., Hurwitz D. I., Lanczycki C. J., Lu F., Marchler G. H., Song J. S., Thanki N., Wang Z., Yamashita R. A., Zhang D., Zheng C., Bryant S. H. (2015) **CDD: NCBI's conserved domain database**. *Nucleic acids research*, **43**(Database issue):D222-226.
378. Jantzen S. G., Sutherland B. J. G., Minkley D. R., Koop B. F. (2011) **GO Trimming: Systematically reducing redundancy in large Gene Ontology datasets**. *BMC Research Notes*, **4**(1):267.
379. Boutet E., Lieberherr D., Tognolli M., Schneider M., Bansal P., Bridge A. J., Poux S., Bougueleret L., Xenarios I. (2016) **UniProtKB/Swiss-Prot, the Manually Annotated Section of the UniProt KnowledgeBase: How to Use the Entry View**. *Methods in molecular biology (Clifton, NJ)*, **1374**:23-54.
380. Kanehisa M., Sato Y., Kawashima M., Furumichi M., Tanabe M. (2015) **KEGG as a reference resource for gene and protein annotation**. *Nucleic acids research*.
381. Costas E., Goyanes V. (2005) **Architecture and evolution of dinoflagellate chromosomes: an enigmatic origin**. *Cytogenetic and genome research*, **109**(1-3):268-275.
382. Beauchemin M., Roy S., Pelletier S., Averback A., Lanthier F., Morse D. (2016) **Characterization of Two Dinoflagellate Cold Shock Domain Proteins**. *mSphere*, **1**(1).
383. Kustatscher G., Hegarat N., Wills K. L., Furlan C., Bukowski-Wills J. C., Hochegger H., Rappsilber J. (2014) **Proteomics of a fuzzy organelle: interphase chromatin**. *The EMBO journal*, **33**(6):648-664.
384. Sonmez C., Baurle I., Magusin A., Dreos R., Laubinger S., Weigel D., Dean C. (2011) **RNA 3' processing functions of Arabidopsis FCA and FPA limit intergenic transcription**. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(20):8508-8513.
385. Zhang J., Chen Q. M. (2013) **Far upstream element binding protein 1: a commander of transcription, translation and beyond**. *Oncogene*, **32**(24):2907-2916.
386. Hanes S. D. (2015) **Prolyl isomerases in gene transcription**. *Biochimica et biophysica acta*, **1850**(10):2017-2034.

387. Ho P., Kong K., Chan Y., Tsang J. S. H., Wong J. T. Y. (2007) **An unusual S-adenosylmethionine synthetase gene from dinoflagellate is methylated**. *BMC Molecular Biology*, **8**:87.
388. Mouveaux T., Oria G., Werkmeister E., Slomianny C., Fox B. A., Bzik D. J., Tomavo S. (2014) **Nuclear Glycolytic Enzyme Enolase of Toxoplasma gondii Functions as a Transcriptional Regulator**. *PLoS ONE*, **9**(8):e105820.
389. Ventura M., Mateo F., Serratosa J., Salaet I., Carujo S., Bachs O., Pujol M. J. (2010) **Nuclear translocation of glyceraldehyde-3-phosphate dehydrogenase is regulated by acetylation**. *The international journal of biochemistry & cell biology*, **42**(10):1672-1680.
390. Matsuda S., Adachi J., Ihara M., Tanuma N., Shima H., Kakizuka A., Ikura M., Ikura T., Matsuda T. (2016) **Nuclear pyruvate kinase M2 complex serves as a transcriptional coactivator of arylhydrocarbon receptor**. *Nucleic acids research*, **44**(2):636-647.
391. Kotoglou P., Kalaitzakis A., Vezyraki P., Tzavaras T., Michalis L. K., Dantzer F., Jung J. U., Angelidis C. (2009) **Hsp70 translocates to the nuclei and nucleoli, binds to XRCC1 and PARP-1, and protects HeLa cells from single-strand DNA breaks**. *Cell stress & chaperones*, **14**(4):391-406.
392. Michels A. A., Kanon B., Konings A. W. T., Ohtsuka K., Bensaude O., Kampinga H. H. (1997) **Hsp70 and Hsp40 Chaperone Activities in the Cytoplasm and the Nucleus of Mammalian Cells**. *Journal of Biological Chemistry*, **272**(52):33283-33289.
393. Ohta S., Bukowski-Wills J. C., Sanchez-Pulido L., Alves Fde L., Wood L., Chen Z. A., Platani M., Fischer L., Hudson D. F., Ponting C. P., Fukagawa T., Earnshaw W. C., Rappsilber J. (2010) **The protein composition of mitotic chromosomes determined using multiclassifier combinatorial proteomics**. *Cell*, **142**(5):810-821.
394. Qin H., Wang Y. (2009) **Exploring DNA-binding proteins with in vivo chemical cross-linking and mass spectrometry**. *Journal of proteome research*, **8**(4):1983-1991.
395. Plumbridge J. (2001) **DNA binding sites for the Mlc and NagC proteins: regulation of nagE, encoding the N-acetylglucosamine-specific transporter in Escherichia coli**. *Nucleic acids research*, **29**(2):506-514.
396. Nosenko T., Bhattacharya D. (2007) **Horizontal gene transfer in chromalveolates**. *BMC evolutionary biology*, **7**:173.
397. Keeling P. J., Palmer J. D. (2008) **Horizontal gene transfer in eukaryotic evolution**. *Nat Rev Genet*, **9**(8):605-618.
398. Lyabin D. N., Eliseeva I. A., Ovchinnikov L. P. (2014) **YB-1 protein: functions and regulation**. *Wiley interdisciplinary reviews RNA*, **5**(1):95-110.
399. Skabkin M. A., Evdokimova V., Thomas A. A., Ovchinnikov L. P. (2001) **The major messenger ribonucleoprotein particle protein p50 (YB-1) promotes nucleic acid strand annealing**. *The Journal of biological chemistry*, **276**(48):44841-44847.
400. Bancel E., Bonnot T., Davanture M., Branlard G., Zivy M., Martre P. (2015) **Proteomic Approach to Identify Nuclear Proteins in Wheat Grain**. *Journal of proteome research*, **14**(10):4432-4439.

401. Singh G., Pratt G., Yeo G. W., Moore M. J. (2015) **The Clothes Make the mRNA: Past and Present Trends in mRNP Fashion**. *Annual review of biochemistry*, **84**:325-354.
402. Linder P., Jankowsky E. (2011) **From unwinding to clamping — the DEAD box RNA helicase family**. *Nat Rev Mol Cell Biol*, **12**(8):505-516.
403. Xiang N., He M., Ishaq M., Gao Y., Song F., Guo L., Ma L., Sun G., Liu D., Guo D., Chen Y. (2016) **The DEAD-Box RNA Helicase DDX3 Interacts with NF-kappaB Subunit p65 and Suppresses p65-Mediated Transcription**. *PLoS One*, **11**(10):e0164471.
404. Huang C. K., Shen Y. L., Huang L. F., Wu S. J., Yeh C. H., Lu C. A. (2016) **The DEAD-Box RNA Helicase AtRH7/PRH75 Participates in Pre-rRNA Processing, Plant Development and Cold Tolerance in Arabidopsis**. *Plant & cell physiology*, **57**(1):174-191.
405. Calo E., Flynn R. A., Martin L., Spitale R. C., Chang H. Y., Wysocka J. (2015) **RNA helicase DDX21 coordinates transcription and ribosomal RNA processing**. *Nature*, **518**(7538):249-253.
406. Montpetit B., Thomsen N. D., Helmke K. J., Seeliger M. A., Berger J. M., Weis K. (2011) **A conserved mechanism of DEAD-box ATPase activation by nucleoporins and InsP6 in mRNA export**. *Nature*, **472**(7342):238-242.
407. Adhvaryu K., Firoozi G., Motavaze K., Lakin-Thomas P. (2016) **PRD-1, a Component of the Circadian System of Neurospora crassa, Is a Member of the DEAD-box RNA Helicase Family**. *Journal of biological rhythms*, **31**(3):258-271.
408. Emerson J. M., Bartholomai B. M., Ringelberg C. S., Baker S. E., Loros J. J., Dunlap J. C. (2015) **period-1 encodes an ATP-dependent RNA helicase that influences nutritional compensation of the Neurospora circadian clock**. *Proceedings of the National Academy of Sciences*, **112**(51):15707-15712.
409. Soyer-Gobillard M. O., Herzog M. (1985) **The native structure of dinoflagellate chromosomes - involvement of structural RNA**. *European journal of cell biology*, **36**(2):334-342.
410. Soyer M.-O., Haapala O. K. (1974) **Electron microscopy of RNA in dinoflagellate chromosomes**. *Histochemistry*, **42**(3):239-246.
411. Indiani C., O'Donnell M. (2006) **The replication clamp-loading machine at work in the three domains of life**. *Nat Rev Mol Cell Biol*, **7**(10):751-761.
412. Oakley G. G., Patrick S. M. (2010) **Replication protein A: directing traffic at the intersection of replication and repair**. *Frontiers in bioscience (Landmark edition)*, **15**:883-900.
413. Rizzo P. J., Nooden L. D. (1972) **Chromosomal proteins in the dinoflagellate alga Gyrodinium cohnii**. *Science (New York, NY)*, **176**(4036):796-797.
414. Briquet S., Boschet C., Gissot M., Tissandié E., Sevilla E., Franetich J.-F., Thiery I., Hamid Z., Bourgouin C., Vaquero C. (2006) **High-Mobility-Group Box Nuclear Factors of Plasmodium falciparum**. *Eukaryotic cell*, **5**(4):672-682.
415. Gissot M., Ting L.-M., Daly T. M., Bergman L. W., Sinnis P., Kim K. (2008) **High Mobility Group Protein HMGB2 Is a Critical Regulator of Plasmodium Oocyst Development**. *The Journal of biological chemistry*, **283**(25):17030-17038.
416. Wang H., Lei T., Liu J., Li M., Nan H., Liu Q. (2014) **A Nuclear Factor of High Mobility Group Box Protein in Toxoplasma gondii**. *PLoS ONE*, **9**(11):e111993.

417. de Mendoza A., Sebé-Pedrós A., Šestak M. S., Matejčić M., Torruella G., Domazet-Lošo T., Ruiz-Trillo I. (2013) **Transcription factor evolution in eukaryotes and the assembly of the regulatory toolkit in multicellular lineages**. *Proceedings of the National Academy of Sciences*, **110**(50):E4858-E4866.
418. Oberstaller J., Pumpalova Y., Schieler A., Llinas M., Kissinger J. C. (2014) **The Cryptosporidium parvum ApiAP2 gene family: insights into the evolution of apicomplexan AP2 regulatory systems**. *Nucleic acids research*, **42**(13):8271-8284.
419. Painter H. J., Campbell T. L., Llinás M. (2011) **The Apicomplexan AP2 family: Integral factors regulating Plasmodium development**. *Molecular and biochemical parasitology*, **176**(1):1-7.
420. Molina C., Grotewold E. (2005) **Genome wide analysis of Arabidopsis core promoters**. *BMC genomics*, **6**(1):25.
421. Basehoar A. D., Zanton S. J., Pugh B. F. (2004) **Identification and distinct regulation of yeast TATA box-containing genes**. *Cell*, **116**(5):699-709.
422. Ingolia N. T., Brar G. A., Rouskin S., McGeachy A. M., Weissman J. S. (2012) **The ribosome profiling strategy for monitoring translation in vivo by deep sequencing of ribosome-protected mRNA fragments**. *Nature protocols*, **7**(8):1534-1550.
423. Janich P., Arpat A. B., Castelo-Szekely V., Lopes M., Gatfield D. (2015) **Ribosome profiling reveals the rhythmic liver translatome and circadian clock regulation by upstream open reading frames**. *Genome Research*, **25**(12):1848-1859.
424. Jang C., Lahens N. F., Hogenesch J. B., Sehgal A. (2015) **Ribosome profiling reveals an important role for translational control in circadian gene expression**. *Genome Research*, **25**(12):1836-1847.
425. Juntawong P., Girke T., Bazin J., Bailey-Serres J. (2014) **Translational dynamics revealed by genome-wide profiling of ribosome footprints in Arabidopsis**. *Proceedings of the National Academy of Sciences*, **111**(1):E203-E212.
426. Robles M. S., Cox J., Mann M. (2014) **In-Vivo Quantitative Proteomics Reveals a Key Contribution of Post-Transcriptional Mechanisms to the Circadian Regulation of Liver Metabolism**. *PLoS genetics*, **10**(1):e1004047.
427. Mauvoisin D., Wang J., Jouffe C., Martin E., Atger F., Waridel P., Quadroni M., Gachon F., Naef F. (2014) **Circadian clock-dependent and -independent rhythmic proteomes implement distinct diurnal functions in mouse liver**. *Proceedings of the National Academy of Sciences*, **111**(1):167-172.
428. Robles M. S., Humphrey S. J., Mann M. (2016) **Phosphorylation Is a Central Mechanism for Circadian Control of Metabolism and Physiology**. *Cell metabolism*.

# Annexe

Tableau 4.S1. List of 1245 proteins identified in the chromatin enriched sample with BLAST description, GO terms and manually classified functional characterization

| Accession | Blast Description | GO Names list | Functional characterization |
|---|---|---|---|
| GABP01003893.1 | heat shock 90 | F:ATPase activity; F:ion binding; F:unfolded protein binding; P:protein folding; P:response to stress | Chaperone |
| GABP01018324.1 | molecular chaperone | F:heat shock protein binding; F:zinc ion binding; P:DNA replication; P:cell redox homeostasis; F:unfolded protein binding; P:protein folding; F:ATP binding; C:cell; P:response to heat; C:cytoplasm; F:metal ion binding; C:membrane; C:integral component of membrane | Chaperone |
| GABP01018766.1 | heat shock | F:ATPase activity; F:oxidoreductase activity; F:ion binding; C:endoplasmic reticulum | Chaperone |
| GABP01020124.1 | dnaJ homolog | P:biological_process; F:ion binding; C:cytosol | Chaperone |
| GABP01025865.1 | heat shock 70 kDa 15-like | C:nucleus; C:cell wall; C:cytosol; C:plasma membrane | Chaperone |
| GABP01025866.1 | heat shock 70 | C:nucleus; C:cell wall; C:cytosol; C:plasma membrane; P:response to stress | Chaperone |
| GABP01028133.1 | heat shock HSP60 | F:ATPase activity; P:protein targeting; C:mitochondrion; F:ion binding; P:protein folding; C:vacuole; C:plasma membrane; C:plastid; P:mitochondrion organization; P:response to stress | Chaperone |
| GABP01034472.1 | heat shock 70 kDa 15-like | F:nucleotide binding; F:ATP binding; C:nucleus; F:2-alkenal reductase [NAD(P)] activity; F:oxidoreductase activity; C:cytosol; P:oxidation-reduction process | Chaperone |
| GABP01034474.1 | heat shock 70 kDa 4 | P:membrane organization; P:protein targeting; P:transmembrane transport; P:cell death; P:protein complex assembly; C:cytosol; F:transcription factor binding; C:organelle; C:extracellular region; P:response to stress; P:mitochondrion organization | Chaperone |
| GABP01038010.1 | DnaJ and TPR domain protein | P:biological_process; F:zinc ion binding; P:regulation of transcription, DNA-templated; F:sequence-specific DNA binding; F:molecular_function; F:DNA binding; C:nucleus; F:transcription factor activity, sequence-specific DNA binding; C:viral envelope; F:metal ion binding; C:membrane; C:integral component of membrane; C:cytosol | Chaperone |
| GABP01040516.1 | heat shock 70 | C:nucleus; C:cytoplasm; F:molecular_function; P:response to stress; P:reproduction | Chaperone |
| GABP01050876.1 | sacsin isoform X1, | F:molecular_function | Chaperone |
| GABP01057428.1 | Hsc70-interacting | F:ATPase activity; P:membrane organization; P:signal transduction; P:cellular component assembly; F:ion binding; P:cellular protein modification process; F:unfolded protein binding; P:biosynthetic process; C:protein complex; P:mRNA processing; P:protein folding; P:transport; P:response to stress; P:immune system process; P:mitochondrion organization; C:nucleoplasm; P:small molecule metabolic process; C:extracellular space; P:cell death; C:microtubule organizing center; C:mitochondrion; C:cytosol; F:enzyme binding | Chaperone |
| GABP01065568.1 | domain containing | F:heat shock protein binding; F:ATP binding; P:response to heat; F:zinc ion binding; C:cytoplasm; F:oxidoreductase activity; F:metal ion binding; P:DNA replication; P:oxidation-reduction process; F:unfolded protein binding; P:protein folding | Chaperone |

| | | | |
|---|---|---|---|
| GABP01066848.1 | Heat shock 70 kDa protein 17 | F:molecular_function | Chaperone |
| GABP01071377.1 | molecular chaperone | F:nucleotide binding; F:ATP binding; C:membrane; C:integral component of membrane; F:unfolded protein binding; P:protein folding | Chaperone |
| GABP01077791.1 | heat shock cognate 70 kDa 2-like | F:ion binding; P:anatomical structure development; C:Golgi apparatus; P:biosynthetic process; C:nucleolus; C:ribosome; C:extracellular region; P:protein folding; C:vacuole; C:plasma membrane; C:plastid; P:response to stress; C:cell wall; C:cytosol; P:cellular nitrogen compound metabolic process; F:enzyme binding | Chaperone |
| GABP01078643.1 | chaperone dnaJ 10-like | P:protein import into peroxisome matrix; C:cytoplasm; C:cytosol; P:protein folding | Chaperone |
| GABP01078701.1 | HSP20-like chaperone | F:protein homodimerization activity; P:negative regulation of cell death; P:response to growth hormone; P:positive regulation of DNA replication; P:aging; C:nuclear envelope lumen; C:neuron projection; C:nucleoplasm; C:beta-catenin destruction complex; P:heart development; C:nucleus; C:extracellular exosome; P:cellular response to calcium ion; C:cytoplasm; C:cell body; P:cardiac muscle cell differentiation | Chaperone |
| GABP01080878.1 | heat shock | F:molecular_function | Chaperone |
| GABP01088774.1 | heat shock | F:ATPase activity; P:biological_process; F:ion binding; C:cellular_component | Chaperone |
| GABP01088775.1 | heat shock | C:nucleus; P:protein targeting; P:membrane organization; P:transmembrane transport; F:oxidoreductase activity; C:cell wall; F:ion binding; F:unfolded protein binding; P:protein folding; C:plasma membrane; C:plastid | Chaperone |
| GABP01088778.1 | Heat shock partial | F:ATPase activity; P:biological_process; F:ion binding; C:cellular_component | Chaperone |
| GABP01088781.1 | heat shock partial | F:ATPase activity; P:biological_process; F:ion binding; C:cellular_component | Chaperone |
| GABP01102469.1 | heat shock Hsp88 | F:nucleotide binding; F:ATPase activity; P:biological_process; F:ATP binding; C:nucleus; P:metabolic process; C:cytosol; F:hydrolase activity | Chaperone |
| GABP01102486.1 | heat shock | F:ATPase activity; C:mitochondrion; F:ion binding; F:unfolded protein binding; P:protein folding; P:response to stress | Chaperone |
| GABP01102501.1 | heat shock | F:ATPase activity; F:oxidoreductase activity; C:mitochondrion; F:ion binding; C:Golgi apparatus; F:unfolded protein binding; P:protein folding; P:transport; P:response to stress | Chaperone |
| GABP01104643.1 | molecular chaperone | P:biological_process; C:cytoplasm; F:ion binding | Chaperone |
| GABP01110608.1 | calnexin | C:cellular_component; F:molecular_function | Chaperone |
| GABP01112238.1 | heat shock | C:cell; P:biological_process | Chaperone |
| GABP01113947.1 | molecular chaperone | P:protein targeting; F:oxidoreductase activity; C:mitochondrion; F:ion binding; P:anatomical structure development; F:unfolded protein binding; P:protein folding; P:mitochondrion organization; P:response to stress; P:immune system process; P:growth | Chaperone |
| GABP01114474.1 | heat shock 90 | F:ATPase activity; F:ATP binding; P:response to heat; C:cytoplasm; F:unfolded protein binding; P:protein folding; P:response to unfolded protein | Chaperone |
| GABP01053147.1 | light harvesting complex | P:generation of precursor metabolites and energy; P:cellular protein modification process; P:photosynthesis; C:protein complex; C:plastid; F:molecular_function; C:thylakoid | Chloroplast |

| | | | |
|---|---|---|---|
| GABP01054344.1 | glyceraldehyde-3-phosphate dehydrogenase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; F:oxidoreductase activity; P:cofactor metabolic process; P:photosynthesis; P:biosynthetic process; P:cellular nitrogen compound metabolic process; C:plastid | Chloroplast |
| GABP01000395.1 | NADPH--cytochrome P450 reductase | F:oxidoreductase activity | Chloroplast |
| GABP01000957.1 | ascorbate peroxidase | F:oxidoreductase activity; C:plastid | Chloroplast |
| GABP01005025.1 | photosynthetic NDH subunit of lumenal location chloroplastic | P:cellular protein modification process; P:protein folding; C:plastid; F:isomerase activity | Chloroplast |
| GABP01006383.1 | CP43 chlorophyll apo of photosystem II (chloroplast) | P:generation of precursor metabolites and energy; F:ion binding; P:cellular protein modification process; P:photosynthesis; C:protein complex; C:plastid; C:thylakoid | Chloroplast |
| GABP01006569.1 | probable L-ascorbate peroxidase chloroplastic | F:oxidoreductase activity; C:plastid; P:response to stress; C:thylakoid | Chloroplast |
| GABP01019002.1 | ATP synthase beta- partial (chloroplast) | F:ATPase activity; P:small molecule metabolic process; F:transmembrane transporter activity; F:ion binding; C:protein complex; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:plastid; C:thylakoid | Chloroplast |
| GABP01019003.1 | ATP synthase beta partial (chloroplast) | F:ATPase activity; P:small molecule metabolic process; F:transmembrane transporter activity; F:ion binding; C:protein complex; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:plasma membrane; C:plastid; C:thylakoid | Chloroplast |
| GABP01019992.1 | phosphoribosylaminoimidazolesuccinocarboxamide synthase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:plastid; F:ligase activity | Chloroplast |
| GABP01019993.1 | phosphoribosylaminoimidazolesuccinocarboxamide synthase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Chloroplast |
| GABP01020546.1 | LOW PSII ACCUMULATION chloroplastic | C:chloroplast | Chloroplast |
| GABP01020785.1 | glutamate synthase 1 [NADH] chloroplastic isoform X1 | F:oxidoreductase activity; F:ion binding; P:nitrogen cycle metabolic process; P:cellular amino acid metabolic process; P:biosynthetic process; C:plastid; P:growth | Chloroplast |
| GABP01022650.1 | chloroplast light harvesting complex | P:generation of precursor metabolites and energy; P:cellular protein modification process; P:photosynthesis; C:protein complex; C:plastid; F:molecular_function; C:thylakoid | Chloroplast |
| GABP01023135.1 | transketolase | P:biological_process; C:plastid; F:molecular_function | Chloroplast |
| GABP01031253.1 | elongation factor Tu | F:RNA binding; C:plastid; P:translation | Chloroplast |
| GABP01033302.1 | ribulose 1,5-bisphosphate carboxylase oxygenase form II | P:carbohydrate metabolic process; F:oxidoreductase activity; F:ion binding; F:lyase activity; P:photosynthesis; P:biosynthetic process; C:plastid | Chloroplast |
| GABP01044079.1 | dihydrodipicolinate synthase | F:lyase activity; P:cellular amino acid metabolic process; P:biosynthetic process; C:plastid | Chloroplast |

| | | | |
|---|---|---|---|
| GABP01045429.1 | glyceraldehyde-3-phosphate dehydrogenase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; F:oxidoreductase activity; P:cofactor metabolic process; P:photosynthesis; P:biosynthetic process; P:cellular nitrogen compound metabolic process; C:plastid | Chloroplast |
| GABP01050590.1 | 50S ribosomal L5 | C:cytoplasm; C:organelle; C:plasma membrane; P:growth | Chloroplast |
| GABP01059090.1 | cyclase | F:transferase activity, transferring glycosyl groups; F:lyase activity; P:cellular amino acid metabolic process; P:biosynthetic process; C:plastid | Chloroplast |
| GABP01067325.1 | chloroplast soluble peridinin-chlorophyll a-binding precursor | P:cellular protein modification process; C:protein complex; C:plastid; F:molecular_function | Chloroplast |
| GABP01068391.1 | beta-amylase chloroplastic | P:carbohydrate metabolic process; C:nucleus; P:polysaccharide catabolic process; F:beta-amylase activity; P:metabolic process; C:cytosol; P:starch catabolic process; F:hydrolase activity; F:hydrolase activity, acting on glycosyl bonds | Chloroplast |
| GABP01073221.1 | bifunctional chloroplastic | P:small molecule metabolic process; C:plastid; F:molecular_function | Chloroplast |
| GABP01076222.1 | plastid NAP50 | | Chloroplast |
| GABP01092571.1 | Light harvesting complex | P:photosynthesis; C:plastid; C:thylakoid | Chloroplast |
| GABP01093502.1 | oxygen-evolving enhancer | P:generation of precursor metabolites and energy; F:ion binding; P:photosynthesis; C:protein complex; C:thylakoid | Chloroplast |
| GABP01096872.1 | arogenate dehydrogenase chloroplastic-like | F:oxidoreductase activity; P:cellular amino acid metabolic process; P:biosynthetic process; C:plastid | Chloroplast |
| GABP01102472.1 | molecular chaperone | F:ion binding; F:unfolded protein binding; P:protein folding; C:plastid | Chloroplast |
| GABP01105033.1 | light-harvesting partial | P:generation of precursor metabolites and energy; P:cellular protein modification process; P:photosynthesis; C:plastid; F:molecular_function; C:thylakoid | Chloroplast |
| GABP01109850.1 | nitrate reductase | C:cellular_component | Chloroplast |
| GABP01008093.1 | centrin | P:cell division; C:cytoskeleton; F:oxidoreductase activity; C:cilium; F:ion binding; P:mitotic nuclear division; C:protein complex | Chromosome related |
| GABP01037410.1 | dinoflagellate viral nucleo | F:DNA binding; C:motile cilium; P:transcription, DNA-templated; C:viral nucleocapsid; P:regulation of transcription, DNA-templated | Chromosome related |
| GABP01047187.1 | regulator of chromosome condensation | P:cell division; C:nucleus; P:mitotic nuclear division | Chromosome related |
| GABP01079327.1 | major basic nuclear, histone like protein | F:DNA binding | Chromosome related |
| GABP01098070.1 | major basic nuclear, histone like protein | F:DNA binding | Chromosome related |
| GABP01103763.1 | Condensin-2 complex subunit partial | P:meiotic chromosome condensation; P:mitotic chromosome condensation; F:histone binding; F:methylated histone binding; C:nuclear pericentric heterochromatin; C:membrane; C:condensed chromosome, centromeric region; C:nuclear condensin complex; F:chromatin binding; P:chromosome separation | Chromosome related |
| GABP01018538.1 | beta-tubulin | P:biological_process; C:cytoskeleton; C:cytoplasm; F:GTPase activity; F:structural molecule activity; F:ion binding | Cytoskeleton |

| | | | |
|---|---|---|---|
| GABP01018540.1 | beta-tubulin | P:biological_process; C:cytoskeleton; C:cytoplasm; F:GTPase activity; F:structural molecule activity; F:ion binding | Cytoskeleton |
| GABP01029270.1 | actin | F:ion binding | Cytoskeleton |
| GABP01034007.1 | profilin | F:actin binding; P:actin cytoskeleton organization | Cytoskeleton |
| GABP01045095.1 | muscular 20 | F:actin binding; P:actomyosin structure organization | Cytoskeleton |
| GABP01060297.1 | EF hand domain containing | P:cytoskeleton organization; C:mating projection tip; F:calcium-dependent protein binding; P:transcription factor import into nucleus; P:cell budding; P:vacuole fusion, non-autophagic; P:phosphatidylinositol biosynthetic process; P:karyogamy involved in conjugation with cellular fusion; P:receptor-mediated endocytosis; C:nucleus; C:cellular bud tip; C:incipient cellular bud site; C:central plaque of spindle pole body; P:spindle pole body organization; C:cellular bud neck; C:cytoplasm; F:metal ion binding; F:calcium ion binding; P:NLS-bearing protein import into nucleus; P:intracellular signal transduction; P:lysosomal microautophagy | Cytoskeleton |
| GABP01065064.1 | actin | F:ion binding | Cytoskeleton |
| GABP01065068.1 | actin | F:ion binding | Cytoskeleton |
| GABP01080648.1 | microtubule-associated partial | C:membrane; C:integral component of membrane; P:microtubule cytoskeleton organization; F:microtubule binding; C:microtubule | Cytoskeleton |
| GABP01083604.1 | tubulin alpha chain | P:biological_process; C:cytoskeleton; C:cytoplasm; F:GTPase activity; F:structural molecule activity; F:ion binding | Cytoskeleton |
| GABP01083606.1 | alpha- partial | P:biological_process; C:cytoskeleton; C:cytoplasm; F:GTPase activity; F:structural molecule activity; F:ion binding | Cytoskeleton |
| GABP01086023.1 | actin-depolymerizing factor 7-like | F:actin binding; P:actin filament depolymerization; C:actin cytoskeleton; C:intracellular | Cytoskeleton |
| GABP01096343.1 | myosin F | F:molecular_function | Cytoskeleton |
| GABP01113353.1 | actin cortical patch | | Cytoskeleton |
| GABP01040802.1 | transcriptional regulator HMG domain | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01105259.1 | transcriptional regulator | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01003768.1 | SAP domain-containing partial | C:nucleus; C:membrane; C:integral component of membrane; P:DNA-templated transcription, termination; P:poly(A)+ mRNA export from nucleus | DNA binding |
| GABP01009729.1 | transcriptional regulator | F:DNA binding; F:ATP binding; P:transcription, DNA-templated; P:phosphorelay signal transduction system; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated; F:protein kinase activity; F:sequence-specific DNA binding; F:ADP binding; P:protein phosphorylation; C:intracellular | DNA binding |
| GABP01019396.1 | unnamed protein product | P:nucleic acid phosphodiester bond hydrolysis; F:DNA binding; F:transcription factor activity, sequence-specific DNA binding; F:endonuclease activity; P:regulation of transcription, DNA-templated | DNA binding |
| GABP01026117.1 | HET-R | F:GTP binding; F:nucleotide binding; F:ATP binding; P:signal transduction; P:small GTPase mediated signal transduction; C:membrane; P:protein transport; C:integral component of membrane; F:ADP binding; C:intracellular | DNA binding |

| | | | |
|---|---|---|---|
| GABP01030993.1 | zinc finger chloroplastic | F:nucleotide binding; F:nucleic acid binding; F:nucleotidyltransferase activity; F:DNA-directed DNA polymerase activity; F:zinc ion binding; C:cytoplasmic stress granule; C:cytoplasm; P:metabolic process; P:DNA biosynthetic process; F:transferase activity; C:nucleolus | DNA binding |
| GABP01035565.1 | Leucine zipper transcription factor 1 | C:motile cilium; P:metabolic process; F:transferase activity; F:N-acetyltransferase activity | DNA binding |
| GABP01040801.1 | transcriptional regulator HMG domain | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01040803.1 | high mobility group B2 | P:male gonad development; P:base-excision repair; P:positive regulation of interleukin-1 beta secretion; F:double-stranded DNA binding; P:positive regulation of NIK/NF-kappaB signaling; P:chromatin organization; P:positive regulation of erythrocyte differentiation; F:single-stranded DNA binding; P:positive regulation of interferon-alpha production; C:cell; P:inflammatory response; F:RNA binding; P:negative regulation of transcription, DNA-templated; F:double-stranded RNA binding; P:positive regulation of transcription, DNA-templated; P:positive regulation of interferon-beta production; F:single-stranded RNA binding; P:positive regulation of interleukin-1 secretion; P:defense response to Gram-positive bacterium; P:negative regulation of extrinsic apoptotic signaling pathway via death domain receptors; F:protein domain specific binding; P:positive regulation of glycogen catabolic process; P:positive regulation of megakaryocyte differentiation; C:protein complex; F:cytokine activity; F:enhancer sequence-specific DNA binding; P:nucleosome assembly; P:immune system process; C:plasma membrane; P:regulation of tolerance induction; C:nucleus; F:protein binding; P:tumor necrosis factor secretion; F:histone binding; C:transcriptional repressor complex; P:negative regulation of RNA polymerase II transcriptional preinitiation complex assembly; P:axon regeneration; P:DNA geometric change; F:transcription regulatory region DNA binding; P:regulation of stem cell proliferation; C:chromosome; P:DNA topological change; P:regulation of nucleotide-excision repair; P:positive regulation of transcription from RNA polymerase II promoter; C:extracellular region; C:nucleolus; C:neuron projection; F:lipopolysaccharide binding; F:supercoiled DNA binding; P:chemotaxis; P:response to cold; F:transcription factor activity, sequence-specific DNA binding; F:DNA polymerase binding; P:positive regulation of interleukin-6 secretion; P:activation of innate immune response; P:response to glucocorticoid; P:activation of protein kinase activity; P:positive regulation of mismatch repair; P:negative regulation of blood vessel endothelial cell migration; P:positive regulation of sprouting angiogenesis; P:DNA recombination; P:response to steroid hormone; P:positive regulation of MAPK cascade; F:damaged DNA binding; P:myeloid dendritic cell activation; F:DNA binding; P:eye development; C:extracellular space; P:negative regulation of apoptotic cell clearance; C:cytoplasm; P:apoptotic DNA fragmentation; P:positive regulation of protein phosphorylation; P:endothelial cell proliferation; P:endothelial cell chemotaxis; P:response to lipopolysaccharide; P:positive regulation of endothelial cell proliferation; F:C-X-C chemokine binding; F:drug binding; P:positive chemotaxis; P:macrophage activation involved in immune response; F:chemoattractant activity; F:DNA binding, bending; P:positive regulation of DNA binding; P:cell chemotaxis; P:forebrain development; P:inflammatory response to antigenic stimulus; F:calcium-dependent protein kinase regulator activity; C:cytosol; P:positive regulation of tumor necrosis factor production; P:positive regulation of wound healing; P:regulation of protein kinase activity; P:positive regulation of toll-like receptor 4 signaling pathway; P:positive regulation of monocyte chemotaxis; P:negative regulation of B cell differentiation; C:condensed chromosome; P:regulation of T cell mediated immune response to tumor cell; P:positive regulation of cysteine-type endopeptidase activity involved in apoptotic process; P:negative regulation of gene expression; C:nuclear chromatin; P:positive regulation of gene expression; P:regulation of autophagy; P:positive regulation of toll-like receptor 2 signaling pathway; | DNA binding |

| | | | |
|---|---|---|---|
| GABP01040806.1 | Non-histone chromosomal 6 | F:DNA binding; P:DNA repair; C:nucleus; P:inflammatory response; P:transcription, DNA-templated; P:negative regulation of cell differentiation; P:regulation of transcription, DNA-templated; C:chromosome; F:cytokine activity; P:cellular response to DNA damage stimulus | DNA binding |
| GABP01040807.1 | nucleolar transcription factor 1 isoform X1 | F:DNA binding; C:nucleus; P:negative regulation of cell differentiation | DNA binding |
| GABP01041390.1 | SAP domain-containing | F:DNA binding; P:biological_process; C:nucleus; C:chloroplast; C:intracellular ribonucleoprotein complex; F:catalytic activity; P:metabolic process; C:viral nucleocapsid; P:DNA-templated transcription, termination; F:phosphopantetheine binding; F:transferase activity | DNA binding |
| GABP01045692.1 | high mobility group B2-like | F:DNA binding; C:nucleus; P:inflammatory response; P:regulation of cell differentiation; P:defense response to Gram-negative bacterium; P:spermatogenesis; P:DNA integration; P:regulation of transcription, DNA-templated; C:chromosome; F:cytokine activity; P:transposition, DNA-mediated | DNA binding |
| GABP01047847.1 | transcriptional regulator | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01047994.1 | zf-PARP type zinc finger | C:organelle; C:intracellular | DNA binding |
| GABP01050709.1 | SAP DNA-binding domain-containing | F:DNA binding; C:membrane; C:integral component of membrane | DNA binding |
| GABP01065827.1 | transcriptional regulator | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01105255.1 | transcriptional regulator | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01105256.1 | transcriptional regulator | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01105257.1 | transcriptional regulator | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01105260.1 | transcriptional regulator | P:phosphorylation; F:kinase activity | DNA binding |
| GABP01108721.1 | HMG box | P:mitochondrion inheritance; P:mitochondrial DNA packaging; F:DNA binding; C:mitochondrial chromosome; C:mitochondrion; C:membrane; C:integral component of membrane; F:DNA binding, bending; P:mitochondrial genome maintenance; P:protein insertion into membrane; C:mitochondrial nucleoid | DNA binding |
| GABP01111203.1 | high mobility group B3 | C:nucleus; F:DNA binding; F:RNA binding; P:chromosome organization; P:anatomical structure development; P:cell differentiation; P:immune system process | DNA binding |
| GABP01111204.1 | high mobility group B3 | F:DNA binding; F:RNA binding; P:cell differentiation; P:anatomical structure development; C:organelle; C:intracellular; P:immune system process | DNA binding |
| GABP01064866.1 | DNA topoisomerase 2 | F:ATPase activity; C:nucleus; P:chromosome organization; P:chromosome segregation; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:cell cycle; P:reproduction; F:isomerase activity | DNA replication, modification and repair |
| GABP01004150.1 | DNA ligase (NAD(+)) | F:DNA binding; P:DNA repair; F:DNA ligase (NAD+) activity; F:metal ion binding; P:DNA replication; P:DNA ligation; P:DNA metabolic process; P:cellular response to DNA damage stimulus; F:ligase activity | DNA replication, modification and repair |
| GABP01018058.1 | DNA topoisomerase 2 | F:ATPase activity; P:cytoskeleton organization; C:nuclear chromosome; P:chromosome segregation; P:mitotic nuclear division; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:reproduction; P:response to stress; F:DNA binding; P:chromosome organization; F:mRNA binding; F:isomerase activity | DNA replication, modification and repair |

| | | | |
|---|---|---|---|
| GABP01020034.1 | DNA replication licensing factor mcm7 | C:nucleus; P:chromosome organization; F:helicase activity; C:protein complex; P:biosynthetic process; P:DNA metabolic process | DNA replication, modification and repair |
| GABP01021323.1 | Poly [ADP-ribose] polymerase 1 | P:DNA metabolic process; C:intracellular; F:molecular_function; P:response to stress | DNA replication, modification and repair |
| GABP01022103.1 | DNA topoisomerase 3-beta-1 | P:biological_process; F:isomerase activity | DNA replication, modification and repair |
| GABP01023412.1 | DNA gyrase subunit A | F:2-alkenal reductase [NAD(P)] activity; P:DNA-dependent DNA replication; P:regulation of transcription, DNA-templated; C:chromosome; P:DNA topological change; P:DNA metabolic process; C:chromatin; F:nucleotide binding; F:DNA binding; F:ATP binding; C:nucleus; C:cytoplasm; F:oxidoreductase activity; F:DNA topoisomerase activity; P:oxidation-reduction process; F:DNA topoisomerase type II (ATP-hydrolyzing) activity; F:isomerase activity | DNA replication, modification and repair |
| GABP01026450.1 | poly [ADP-ribose] polymerase 2 | C:nucleus; F:transferase activity, transferring glycosyl groups; C:cytoplasm; P:biosynthetic process; P:DNA metabolic process; F:ligase activity; P:response to stress | DNA replication, modification and repair |
| GABP01026919.1 | DNA replication licensing factor Mcm5 | C:nucleus; F:DNA binding; C:cytoplasm; P:chromosome organization; C:protein complex; P:biosynthetic process; P:DNA metabolic process | DNA replication, modification and repair |
| GABP01027019.1 | DNA repair RAD50 | F:ATPase activity; C:nuclear chromosome; F:helicase activity; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:response to stress; P:cell cycle; P:reproduction; P:small molecule metabolic process; F:DNA binding; P:chromosome organization; P:homeostatic process; F:nuclease activity; F:kinase activity | DNA replication, modification and repair |
| GABP01028444.1 | DNA replication licensing factor mcm6 | F:ATPase activity; C:nucleus; C:cytoplasm; F:helicase activity; C:protein complex; P:biosynthetic process; P:DNA metabolic process | DNA replication, modification and repair |
| GABP01031995.1 | DNA replication licensing factor Mcm2 | F:DNA binding; C:nuclear chromosome; F:helicase activity; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:transport; C:nucleoplasm | DNA replication, modification and repair |
| GABP01033002.1 | DNA topoisomerase 3-alpha | C:nucleus; P:negative regulation of catalytic activity; C:cytoplasm; F:ATPase inhibitor activity; C:cytosol; F:mRNA binding; F:Hsp90 protein binding; P:protein folding; F:Hsp70 protein binding | DNA replication, modification and repair |
| GABP01034470.1 | DnaK family protein | F:nucleotide binding; F:ATPase activity; F:ATP binding; C:nucleus; P:metabolic process; C:cytosol; F:hydrolase activity | DNA replication, modification and repair |
| GABP01034529.1 | DNA topoisomerase II | F:ATPase activity; C:nucleus; P:chromosome organization; P:chromosome segregation; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:cell cycle; P:reproduction; F:isomerase activity | DNA replication, modification and repair |
| GABP01039371.1 | DNA polymerase delta catalytic subunit | F:nuclease activity; P:DNA metabolic process | DNA replication, modification and repair |
| GABP01050104.1 | DNA damage checkpoint RAD24 14-3-3 sigma | P:membrane organization; P:signal transduction; P:anatomical structure development; P:cellular protein modification process; P:transport; P:mitochondrion organization; P:response to stress; P:cell cycle; C:extracellular space; P:cell death; C:mitochondrion; P:cell differentiation; C:cytosol; C:cytoplasmic, membrane-bounded vesicle; F:enzyme regulator activity | DNA replication, modification and repair |
| GABP01051921.1 | DNA damage checkpoint RAD24 14-3-3 sigma | F:molecular_function | DNA replication, modification and repair |
| GABP01053777.1 | UV excision repair Rad23 | P:biological_process; C:intracellular; F:molecular_function | DNA replication, modification and repair |

| | | | |
|---|---|---|---|
| GABP01053778.1 | UV excision repair Rad23 | P:biological_process; C:intracellular; F:molecular_function | DNA replication, modification and repair |
| GABP01060268.1 | Apurinic endonuclease-redox | F:nuclease activity; F:lyase activity; C:organelle; C:intracellular | DNA replication, modification and repair |
| GABP01064874.1 | DNA topoisomerase 2 | F:ATPase activity; F:ion binding; P:anatomical structure development; P:chromosome segregation; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:cell cycle; P:reproduction; C:nucleus; F:DNA binding; F:oxidoreductase activity; P:chromosome organization; F:isomerase activity | DNA replication, modification and repair |
| GABP01066334.1 | poly [ADP-ribose] polymerase 3 | F:transferase activity, transferring glycosyl groups; P:DNA metabolic process | DNA replication, modification and repair |
| GABP01078391.1 | DNA topoisomerase I | C:nucleus; P:chromosome organization; C:chromosome; P:DNA metabolic process; F:isomerase activity | DNA replication, modification and repair |
| GABP01079451.1 | replication factor-a 1 | C:nucleus; P:chromosome organization; P:DNA metabolic process; F:molecular_function | DNA replication, modification and repair |
| GABP01087174.1 | DNA topoisomerase 2 | F:ATPase activity; C:nucleus; P:chromosome organization; P:chromosome segregation; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:cell cycle; P:reproduction; F:isomerase activity | DNA replication, modification and repair |
| GABP01087179.1 | DNA topoisomerase 2 | F:ATPase activity; C:nucleus; P:chromosome organization; P:chromosome segregation; C:protein complex; P:biosynthetic process; P:DNA metabolic process; P:cell cycle; P:reproduction; F:isomerase activity | DNA replication, modification and repair |
| GABP01093499.1 | dNA ligase | P:metabolic process; P:maintenance of stationary phase; P:chaperone mediated protein folding requiring cofactor; F:peptide binding; F:unfolded protein binding; C:periplasmic space; P:protein peptidyl-prolyl isomerization; P:protein folding; F:peptidyl-prolyl cis-trans isomerase activity; P:Gram-negative-bacterium-type cell outer membrane assembly; C:outer membrane-bounded periplasmic space; P:protein stabilization; F:isomerase activity | DNA replication, modification and repair |
| GABP01106746.1 | XPG N-terminal domain-containing | P:biological_process | DNA replication, modification and repair |
| GABP01108818.1 | ATP-dependent DNA helicase | F:nucleotide binding; F:nucleic acid binding; F:ATP binding; F:catalytic activity; P:DNA recombination; F:ATP-dependent helicase activity; F:helicase activity; P:cellular metabolic process; F:hydrolase activity | DNA replication, modification and repair |
| GABP01109762.1 | DNA topoisomerase II | P:chromosome organization; P:chromosome segregation; P:DNA metabolic process; C:intracellular; P:cell cycle; F:isomerase activity | DNA replication, modification and repair |
| GABP01110104.1 | DNA damage checkpoint RAD24 14-3-3 sigma | F:molecular_function | DNA replication, modification and repair |
| GABP01113325.1 | Poly [ADP-ribose] polymerase 1 | P:DNA metabolic process; F:molecular_function | DNA replication, modification and repair |
| GABP01033237.1 | GMP synthase [glutamine-hydrolyzing] | P:small molecule metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01033248.1 | GMP synthetase | F:ion binding; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01044637.1 | carbonate dehydratase | F:zinc ion binding; F:metal ion binding; F:lyase activity; F:carbonate dehydratase activity; P:one-carbon metabolic process | Enzyme |
| GABP01062994.1 | glutamine amidotransferase subunit pdxt | F:glutaminase activity; P:pyridoxal phosphate biosynthetic process; P:vitamin B6 biosynthetic process; P:glutamine metabolic process; F:transferase activity | Enzyme |

| | | | |
|---|---|---|---|
| GABP01109680.1 | bifunctional aconitate hydratase 2 2-methylisocitrate dehydratase | P:small molecule metabolic process; P:generation of precursor metabolites and energy; F:lyase activity; F:ion binding; C:cytosol | Enzyme |
| GABP01000275.1 | type I polyketide synthase | P:biological_process; F:transferase activity, transferring acyl groups | Enzyme |
| GABP01001178.1 | glutathione S-transferase Mu 3-like | P:biological_process | Enzyme |
| GABP01001349.1 | glycine dehydrogenase (aminomethyl-transferring) | P:catabolic process; F:oxidoreductase activity; C:mitochondrion; F:ion binding; P:cellular amino acid metabolic process; C:cytosol; C:protein complex; C:extracellular region; C:plastid; C:thylakoid | Enzyme |
| GABP01001468.1 | GMP synthase (glutamine-hydrolyzing) domain-containing | P:small molecule metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01002799.1 | inosine-5 -monophosphate dehydrogenase | P:small molecule metabolic process; P:vesicle-mediated transport; F:oxidoreductase activity; P:locomotion; P:cell morphogenesis; P:cell differentiation; P:cellular nitrogen compound metabolic process; P:biosynthetic process; P:reproduction | Enzyme |
| GABP01002801.1 | inosine-5 -monophosphate dehydrogenase | P:cell proliferation; P:small molecule metabolic process; C:nucleus; C:peroxisome; F:oxidoreductase activity; F:ion binding; P:biosynthetic process; P:cellular nitrogen compound metabolic process; C:extracellular region; P:immune system process | Enzyme |
| GABP01003248.1 | asparagine synthase B | F:ligase activity | Enzyme |
| GABP01004213.1 | methionine synthase | P:sulfur compound metabolic process; F:ion binding; P:cellular amino acid metabolic process; F:methyltransferase activity; P:biosynthetic process; C:intracellular | Enzyme |
| GABP01004840.1 | glutamine synthetase | P:biological_process; F:ligase activity | Enzyme |
| GABP01005456.1 | adenosylhomocysteinase | P:sulfur compound metabolic process; P:cofactor metabolic process; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; F:molecular_function | Enzyme |
| GABP01005507.1 | squalene synthase | P:lipid metabolic process; P:biosynthetic process; C:cellular_component; F:transferase activity, transferring alkyl or aryl (other than methyl) groups; F:isomerase activity | Enzyme |
| GABP01005969.1 | Nucleoside diphosphate kinase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; C:cellular_component; F:kinase activity | Enzyme |
| GABP01005972.1 | phosphoglycerate kinase 1 | P:generation of precursor metabolites and energy; P:catabolic process; F:ion binding; P:anatomical structure development; P:biosynthetic process; C:organelle; C:extracellular region; P:small molecule metabolic process; P:carbohydrate metabolic process; P:cofactor metabolic process; P:cell differentiation; C:cytosol; P:cellular nitrogen compound metabolic process; F:kinase activity | Enzyme |
| GABP01006382.1 | adenosylhomocysteinase | P:sulfur compound metabolic process; P:cofactor metabolic process; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; F:molecular_function | Enzyme |

| | | | |
|---|---|---|---|
| GABP01007544.1 | glyceraldehyde-3-phosphate dehydrogenase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; F:oxidoreductase activity; P:cofactor metabolic process; F:ion binding; F:lyase activity; C:cytosol; C:protein complex; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01008622.1 | hypothetical protein | F:hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds, in linear amides; P:metabolic process; C:membrane; C:integral component of membrane | Enzyme |
| GABP01009094.1 | microcystin synthetase-associated partial | F:molecular_function | Enzyme |
| GABP01012557.1 | Ankyrin repeat domain-containing 27 | F:ATP binding; P:signal transduction; F:2-alkenal reductase [NAD(P)] activity; F:oxidoreductase activity; F:fatty-acyl-CoA binding; P:metabolic process; P:oxidation-reduction process; F:protein kinase activity; P:protein phosphorylation; F:hydrolase activity | Enzyme |
| GABP01015402.1 | glutaredoxin-C4 | C:cytoplasm; F:oxidoreductase activity; P:anatomical structure development; C:organelle | Enzyme |
| GABP01017484.1 | Phosphoribosylamine--glycine ligase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01018345.1 | aspartate aminotransferase | C:peroxisome; P:generation of precursor metabolites and energy; C:mitochondrion; F:ion binding; P:anatomical structure development; P:cellular amino acid metabolic process; C:cytosol; P:biosynthetic process; P:aging; C:plastid | Enzyme |
| GABP01018477.1 | carbonate dehydratase | F:nucleotide binding; F:acyl-CoA dehydrogenase activity; F:ATP binding; F:oxidoreductase activity, acting on the CH-CH group of donors; F:flavin adenine dinucleotide binding; F:oxidoreductase activity; F:metal ion binding; P:metabolic process; C:membrane; C:integral component of membrane; P:oxidation-reduction process; F:hydrolase activity | Enzyme |
| GABP01018919.1 | enolase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; P:cofactor metabolic process; F:lyase activity; F:ion binding; C:cytosol; C:protein complex; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01019343.1 | fumarate reductase | F:metal ion binding; F:oxidoreductase activity; C:membrane; C:integral component of membrane; P:oxidation-reduction process; F:succinate dehydrogenase activity; F:heme binding | Enzyme |
| GABP01019988.1 | phosphoribosylaminoimidazole-succinocarboxamide synthase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01019990.1 | multifunctional ADE2 | P:small molecule metabolic process; P:embryo development; P:pigmentation; F:ion binding; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01020047.1 | glycosyltransferase | F:transferase activity, transferring glycosyl groups; P:biosynthetic process; C:plastid | Enzyme |
| GABP01020442.1 | oxidoreductase | P:metabolic process; F:isomerase activity | Enzyme |
| GABP01020980.1 | argininosuccinate lyase | P:cellular amino acid metabolic process; P:biosynthetic process; F:molecular_function | Enzyme |

| | | | |
|---|---|---|---|
| GABP01021487.1 | bifunctional diaminohydroxyphosphoribosylaminopyrimidine deaminase 5-amino-6-(5-phosphoribosylamino)uracil reductase | P:biological_process; F:molecular_function | Enzyme |
| GABP01022835.1 | nitrite reductase | F:molecular_function | Enzyme |
| GABP01023168.1 | UDP-glucose 4-epimerase | P:carbohydrate metabolic process; C:Golgi apparatus; F:isomerase activity | Enzyme |
| GABP01023490.1 | glu leu phe val dehydrogenase | F:oxidoreductase activity; P:glutamate catabolic process to 2-oxoglutarate; P:cellular amino acid metabolic process; P:oxidation-reduction process; F:glutamate dehydrogenase (NAD+) activity; F:hydrolase activity | Enzyme |
| GABP01023633.1 | phospholipase A-2-activating | C:plasmodesma; F:oxidoreductase activity; F:aromatase activity; F:myosin heavy chain kinase activity; C:membrane; C:integral component of membrane; C:cytosol; P:oxidation-reduction process; F:transferase activity; P:protein phosphorylation | Enzyme |
| GABP01023944.1 | fumarate reductase | F:oxidoreductase activity; C:cellular_component | Enzyme |
| GABP01023947.1 | fumarate reductase | F:oxidoreductase activity; C:cellular_component | Enzyme |
| GABP01023988.1 | polyketide synthase | F:transferase activity, transferring acyl groups | Enzyme |
| GABP01024114.1 | d-alanine--D-alanine ligase | F:ion binding | Enzyme |
| GABP01024115.1 | D-alanine--D-alanine ligase | F:D-alanine-D-alanine ligase activity; F:ATP binding; F:metal ion binding; F:catalytic activity; P:metabolic process; F:ligase activity | Enzyme |
| GABP01024118.1 | D-alanine--D-alanine ligase | P:biological_process; F:ion binding; F:ligase activity | Enzyme |
| GABP01025581.1 | isochorismatase | C:cellular_component; F:molecular_function | Enzyme |
| GABP01025657.1 | D-galacturonic acid reductase | F:oxidoreductase activity | Enzyme |
| GABP01026612.1 | partial | P:biological_process; F:transferase activity, transferring acyl groups; F:ion binding | Enzyme |
| GABP01027990.1 | pyruvate carboxylase | P:biological_process; F:ligase activity | Enzyme |
| GABP01029576.1 | aldehyde dehydrogenase | F:succinate-semialdehyde dehydrogenase (NAD+) activity; F:aldehyde dehydrogenase (NAD) activity; F:ATPase activity; F:ATP binding; P:cellular aldehyde metabolic process; F:succinate-semialdehyde dehydrogenase [NAD(P)+] activity; F:oxidoreductase activity; F:aldehyde dehydrogenase [NAD(P)+] activity; P:metabolic process; P:oxidation-reduction process; F:oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor | Enzyme |

| | | | |
|---|---|---|---|
| GABP01029769.1 | malic enzyme | F:transferase activity, transferring acyl groups; F:oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor; P:metabolic process; F:malate dehydrogenase (decarboxylating) (NAD+) activity; F:malate dehydrogenase (decarboxylating) (NADP+) activity; F:hydrolase activity, hydrolyzing O-glycosyl compounds; F:phosphate acetyltransferase activity; F:oxaloacetate decarboxylase activity; P:carbohydrate metabolic process; P:malate metabolic process; F:metal ion binding; F:oxidoreductase activity; F:NAD binding; C:cytosol; P:oxidation-reduction process; F:manganese ion binding; F:transferase activity; F:hydrolase activity, acting on glycosyl bonds; F:hydrolase activity | Enzyme |
| GABP01030370.1 | cyclopropane-fatty-acyl-phospholipid synthase | C:cytoplasm; C:membrane; C:integral component of membrane; P:lipid biosynthetic process; F:methyltransferase activity; F:transferase activity; F:S-adenosylmethionine-dependent methyltransferase activity; P:methylation | Enzyme |
| GABP01030531.1 | enolase 3 | P:generation of precursor metabolites and energy; P:catabolic process; F:ion binding; F:lyase activity; C:protein complex; C:organelle; C:plasma membrane; P:small molecule metabolic process; P:carbohydrate metabolic process; C:extracellular space; F:oxidoreductase activity; P:cofactor metabolic process; C:cytosol; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01030788.1 | AMP-dependent synthetase | F:zinc ion binding; F:transferase activity, transferring acyl groups other than amino-acyl groups; P:metabolic process; P:ribosome biogenesis; P:biosynthetic process; F:GTP binding; F:nucleotide binding; F:acetate-CoA ligase activity; F:oxidoreductase activity; F:catalytic activity; P:oxidation-reduction process; F:ligase activity; F:hydrolase activity | Enzyme |
| GABP01031273.1 | Nucleoside diphosphate kinase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; C:cellular_component | Enzyme |
| GABP01031274.1 | Nucleoside diphosphate kinase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; C:cellular_component | Enzyme |
| GABP01032183.1 | malate dehydrogenase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; C:cell wall; C:mitochondrion; F:ion binding; C:extracellular region; C:plastid; P:response to stress | Enzyme |
| GABP01032894.1 | S-adenosylmethionine synthetase | P:small molecule metabolic process; P:sulfur compound metabolic process; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01033238.1 | GMP synthase | P:small molecule metabolic process; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01033240.1 | GMP synthetase | F:ion binding; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01033242.1 | GMP synthase [glutamine-hydrolyzing] | P:small molecule metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01033246.1 | GMP synthase (glutamine-hydrolyzing) domain-containing | F:ion binding; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01033247.1 | gMP synthase [glutamine-hydrolyzing] | P:small molecule metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |

| GABP01033249.1 | GMP synthetase | F:ion binding; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
|---|---|---|---|
| GABP01033486.1 | allophanate hydrolase | F:molecular_function | Enzyme |
| GABP01033558.1 | guanylate cyclase | P:small molecule metabolic process; P:signal transduction; F:lyase activity; P:cellular nitrogen compound metabolic process; P:biosynthetic process | Enzyme |
| GABP01033863.1 | phosphoenolpyruvate synthase | P:small molecule metabolic process; F:ion binding; F:kinase activity | Enzyme |
| GABP01033880.1 | phosphoglycerate mutase | P:small molecule metabolic process; P:secondary metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; P:cofactor metabolic process; C:cytosol; P:biosynthetic process; P:cellular nitrogen compound metabolic process; F:isomerase activity | Enzyme |
| GABP01034178.1 | methionine adenosyltransferase | P:small molecule metabolic process; P:sulfur compound metabolic process; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01034313.1 | carbon-nitrogen hydrolase | F:molecular_function | Enzyme |
| GABP01035448.1 | malonyl -acyl carrier transacylase | P:biological_process; F:molecular_function | Enzyme |
| GABP01035471.1 | short-chain dehydrogenase | F:oxidoreductase activity | Enzyme |
| GABP01036104.1 | tryptophan synthase subunit beta | C:cytoplasm; F:lyase activity; F:ion binding; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process | Enzyme |
| GABP01036105.1 | tryptophan synthase | C:cytoplasm; F:lyase activity; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process | Enzyme |
| GABP01036106.1 | tryptophan synthase | F:lyase activity; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:cellular_component | Enzyme |
| GABP01036269.1 | glutamine synthetase | F:ligase activity | Enzyme |
| GABP01036908.1 | argininosuccinate synthase |  | Enzyme |
| GABP01038195.1 | phosphoribosylformylglycinamidine cyclo-ligase | P:biological_process; F:ligase activity | Enzyme |
| GABP01038337.1 | glu leu phe val dehydrogenase | F:oxidoreductase activity; P:cellular amino acid metabolic process | Enzyme |
| GABP01038338.1 | glu leu phe val dehydrogenase | F:oxidoreductase activity; P:cellular amino acid metabolic process | Enzyme |
| GABP01038648.1 | s-adenosyl methionine partial | P:small molecule metabolic process; P:sulfur compound metabolic process; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01039841.1 | octopine partial | F:oxidoreductase activity | Enzyme |
| GABP01039844.1 | tauropine dehydrogenase-like | F:oxidoreductase activity | Enzyme |
| GABP01039929.1 | bifunctional 5,10-methylene-tetrahydrofolate dehydrogenase 5,10-methylene-tetrahydrofolate cyclohydrolase | P:cellular amino acid metabolic process; P:biosynthetic process; P:cellular nitrogen compound metabolic process; F:molecular_function | Enzyme |

| | | | |
|---|---|---|---|
| GABP01040247.1 | cellulose synthase | P:biological_process; F:UDP-N-acetylmuramate-L-alanine ligase activity; P:UDP-glucose metabolic process; P:biosynthetic process; C:cellular_component; F:molecular_function; F:ATP binding; F:oxidoreductase activity; C:membrane; P:oxidation-reduction process; P:cellulose biosynthetic process; F:ligase activity; C:outer membrane | Enzyme |
| GABP01040552.1 | phosphoglycerate kinase | P:generation of precursor metabolites and energy; P:catabolic process; F:ion binding; C:extracellular region; C:vacuole; C:plasma membrane; C:plastid; P:small molecule metabolic process; C:nucleus; P:carbohydrate metabolic process; P:cofactor metabolic process; C:cytosol; P:cellular nitrogen compound metabolic process; F:kinase activity | Enzyme |
| GABP01040553.1 | phosphoglycerate kinase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; C:cytoplasm; F:oxidoreductase activity; P:cofactor metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; C:organelle; F:kinase activity | Enzyme |
| GABP01041118.1 | N-acetylneuraminate synthase | P:carbohydrate metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01042467.1 | amidohydrolase | P:metabolic process; F:hydrolase activity | Enzyme |
| GABP01042482.1 | beta-galactosidase | P:carbohydrate metabolic process; F:hydrolase activity, acting on glycosyl bonds | Enzyme |
| GABP01042483.1 | beta-galactosidase | P:carbohydrate metabolic process; F:hydrolase activity, acting on glycosyl bonds | Enzyme |
| GABP01043143.1 | cystathionine beta-synthase | P:sulfur compound metabolic process; F:lyase activity; P:cellular amino acid metabolic process; P:biosynthetic process | Enzyme |
| GABP01043533.1 | 3-oxoacyl-ACP reductase | P:biological_process | Enzyme |
| GABP01043568.1 | cyclopropane fatty acid synthase methyltransferase | P:methionine biosynthetic process; F:transferase activity, transferring acyl groups other than amino-acyl groups; P:L-methionine biosynthetic process from methylthioadenosine; P:metabolic process; P:regulation of transcription, DNA-templated; F:methyltransferase activity; P:cellular amino acid biosynthetic process; F:dioxygenase activity; C:nucleus; F:histone acetyltransferase activity; C:cytoplasm; F:oxidoreductase activity; F:acireductone dioxygenase [iron(II)-requiring] activity; P:oxidation-reduction process; P:histone acetylation; F:transferase activity; P:methylation | Enzyme |
| GABP01043743.1 | phosphoribosylaminoimidazolesuccinocarboxamide synthase | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01044390.1 | glycogen phosphorylase 1 | P:carbohydrate metabolic process; F:transferase activity, transferring glycosyl groups; P:generation of precursor metabolites and energy; C:cytoplasm; P:catabolic process; F:ion binding | Enzyme |
| GABP01044898.1 | adenylyl cyclase-associated 1 isoform X1 | P:biological_process | Enzyme |
| GABP01044945.1 | glu leu phe val dehydrogenase | P:biological_process; F:molecular_function | Enzyme |
| GABP01044986.1 | type I polyketide synthase | F:molecular_function | Enzyme |
| GABP01045467.1 | aspartate-semialdehyde dehydrogenase | P:sulfur compound metabolic process; C:cytoplasm; F:oxidoreductase activity; P:cellular amino acid metabolic process; P:biosynthetic process | Enzyme |

| | | | |
|---|---|---|---|
| GABP01045562.1 | 3-ketoacyl- peroxisomal | P:small molecule metabolic process; C:peroxisome; F:transferase activity, transferring acyl groups; F:oxidoreductase activity; P:catabolic process; C:mitochondrion; P:lipid metabolic process | Enzyme |
| GABP01046242.1 | enolase | P:small molecule metabolic process; P:carbohydrate metabolic process; F:transferase activity, transferring acyl groups; P:generation of precursor metabolites and energy; P:catabolic process; F:oxidoreductase activity; P:cofactor metabolic process; F:ion binding; F:lyase activity; C:cytosol; C:protein complex; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01046258.1 | tyrosine dual specificity domain | P:metabolic process; P:proteolysis; C:membrane; C:integral component of membrane; F:peptidase activity; F:metalloendopeptidase activity; F:metallopeptidase activity; F:NAD+ ADP-ribosyltransferase activity; F:hydrolase activity | Enzyme |
| GABP01046314.1 | trans-2-enoyl- mitochondrial isoform X2 | F:zinc ion binding; F:oxidoreductase activity; F:catalytic activity; F:trans-2-enoyl-CoA reductase (NADPH) activity; P:metabolic process; C:mitochondrion; C:membrane; C:integral component of membrane; P:oxidation-reduction process; F:transferase activity; P:fatty acid metabolic process | Enzyme |
| GABP01047158.1 | alcohol dehydrogenase class-3 | P:small molecule metabolic process; F:oxidoreductase activity; F:ion binding; C:chromosome | Enzyme |
| GABP01047160.1 | S-formylglutathione hydrolase | F:molecular_function | Enzyme |
| GABP01047673.1 | succinate | F:oxidoreductase activity; C:cellular_component | Enzyme |
| GABP01047675.1 | fumarate reductase | F:oxidoreductase activity; C:cellular_component | Enzyme |
| GABP01048097.1 | S-adenosylmethionine synthetase | P:small molecule metabolic process; P:sulfur compound metabolic process; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01048099.1 | S-adenosylmethionine synthetase | P:small molecule metabolic process; P:sulfur compound metabolic process; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01048101.1 | s-adenosyl methionine partial | P:small molecule metabolic process; P:sulfur compound metabolic process; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01048242.1 | dihydrodipicolinate reductase | F:oxidoreductase activity; P:cellular amino acid metabolic process; P:biosynthetic process; C:plastid | Enzyme |
| GABP01048602.1 | argininosuccinate synthase | F:ion binding; P:cellular amino acid metabolic process; P:biosynthetic process; F:ligase activity | Enzyme |
| GABP01050032.1 | oligosaccharyl transferase | P:carbohydrate metabolic process; F:transferase activity, transferring glycosyl groups; P:cellular protein modification process; C:endoplasmic reticulum; C:protein complex; P:biosynthetic process; P:cellular nitrogen compound metabolic process; C:ribosome | Enzyme |
| GABP01050188.1 | NAD-dependent epimerase dehydratase | F:molecular_function | Enzyme |
| GABP01051910.1 | Thioesterase | F:molecular_function | Enzyme |

| | | | |
|---|---|---|---|
| GABP01052490.1 | malic enzyme | F:oxidoreductase activity; C:intracellular | Enzyme |
| GABP01052707.1 | spermine partial | C:cytoplasm; P:cell morphogenesis; P:cell differentiation; P:cellular nitrogen compound metabolic process; P:biosynthetic process; P:transport; P:growth; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01052906.1 | anthranilate phosphoribosyltransferase | P:cellular amino acid metabolic process; P:biosynthetic process; F:molecular_function | Enzyme |
| GABP01053397.1 | 3-phosphoglycerate dehydrogenase | F:oxidoreductase activity | Enzyme |
| GABP01055571.1 | adenosylhomocysteinase | P:sulfur compound metabolic process; F:oxidoreductase activity; P:cofactor metabolic process; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01055573.1 | adenosylhomocysteinase | P:sulfur compound metabolic process; P:cofactor metabolic process; C:cytosol; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; F:molecular_function | Enzyme |
| GABP01055830.1 | polyketide synthase | F:molecular_function | Enzyme |
| GABP01055850.1 | multifunctional polyketide-peptide syntase | F:transferase activity, transferring acyl groups | Enzyme |
| GABP01056151.1 | carboxyvinyl-carboxyphosphonate phosphorylmutase | P:small molecule metabolic process; P:carbohydrate metabolic process; F:transferase activity, transferring acyl groups; C:cytoplasm | Enzyme |
| GABP01057741.1 | glutamate-1-semialdehyde 2,1-aminomutase | P:biological_process; F:ion binding | Enzyme |
| GABP01057848.1 | Methylthioribose-1-phosphate isomerase | P:sulfur compound metabolic process; C:cytoplasm; P:cellular amino acid metabolic process; P:biosynthetic process; F:isomerase activity | Enzyme |
| GABP01057964.1 | Adk1p | P:small molecule metabolic process; C:cytoplasm; F:ion binding; P:cellular nitrogen compound metabolic process; C:organelle; F:kinase activity | Enzyme |
| GABP01058350.1 | methionine synthase | F:ion binding; C:cytosol; F:methyltransferase activity | Enzyme |
| GABP01058351.1 | methionine synthase | F:ion binding; F:methyltransferase activity; C:intracellular | Enzyme |
| GABP01058354.1 | 5-methyltetrahydrofolate--homocysteine methyltransferase | P:sulfur compound metabolic process; F:ion binding; C:cytosol; P:cellular amino acid metabolic process; F:methyltransferase activity; P:biosynthetic process | Enzyme |
| GABP01058356.1 | methionine synthase | F:ion binding; F:methyltransferase activity | Enzyme |
| GABP01058583.1 | glycerol-3-phosphate dehydrogenase Gpd1 | P:carbohydrate metabolic process; C:cytoplasm; F:oxidoreductase activity | Enzyme |
| GABP01058616.1 | cyclopropane fatty acid synthase | P:biological_process; F:molecular_function | Enzyme |
| GABP01059032.1 | FO synthase | F:catalytic activity; P:metabolic process; C:membrane; C:integral component of membrane | Enzyme |
| GABP01059145.1 | glutathione s-transferase | P:metabolic process; F:glutathione transferase activity; F:transferase activity; F:isomerase activity | Enzyme |
| GABP01059518.1 | atp-citrate synthase | P:small molecule metabolic process; F:transferase activity, transferring acyl groups; P:sulfur compound metabolic process; P:cofactor metabolic process; P:lipid metabolic process; F:ion binding; C:cytosol; P:biosynthetic process; C:extracellular region; C:plasma membrane; C:nucleoplasm | Enzyme |

| | | | |
|---|---|---|---|
| GABP01059519.1 | atp-citrate synthase | C:nucleus; F:transferase activity, transferring acyl groups; C:cytoplasm; P:lipid metabolic process; F:ion binding; P:biosynthetic process | Enzyme |
| GABP01059520.1 | atp-citrate synthase | F:transferase activity, transferring acyl groups; C:cytoplasm; P:lipid metabolic process; F:ion binding; P:biosynthetic process | Enzyme |
| GABP01059806.1 | phosphoglycerate kinase 1 | P:generation of precursor metabolites and energy; P:catabolic process; F:ion binding; C:extracellular region; P:mitochondrion organization; P:response to stress; P:small molecule metabolic process; P:cell motility; C:nucleus; P:carbohydrate metabolic process; C:cytoplasm; P:autophagy; P:cofactor metabolic process; C:cilium; P:cellular nitrogen compound metabolic process; F:kinase activity | Enzyme |
| GABP01061174.1 | 2-dehydropantoate 2-reductase | P:pantothenate biosynthetic process; F:NADP binding; C:cytoplasm; F:oxidoreductase activity; P:oxidation-reduction process; F:2-dehydropantoate 2-reductase activity | Enzyme |
| GABP01061589.1 | cystathionine beta-synthase | P:sulfur compound metabolic process; C:cytoplasm; F:lyase activity; F:ion binding; P:cellular amino acid metabolic process; P:biosynthetic process | Enzyme |
| GABP01061945.1 | phosphoserine aminotransferase | C:cytoplasm; P:cellular amino acid metabolic process; P:biosynthetic process; C:organelle; C:extracellular region; F:molecular_function | Enzyme |
| GABP01062204.1 | succinate | F:oxidoreductase activity; F:ion binding; C:cellular_component | Enzyme |
| GABP01062506.1 | thioesterase | F:pyrroline-5-carboxylate reductase activity; C:integral component of membrane; P:proline biosynthetic process; P:oxidation-reduction process | Enzyme |
| GABP01063373.1 | homoserine kinase | F:ion binding; P:cellular amino acid metabolic process; F:kinase activity | Enzyme |
| GABP01063523.1 | isocitrate dehydrogenase | P:small molecule metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; F:ion binding | Enzyme |
| GABP01066459.1 | acetyl propionyl- carboxylase subuit alpha | F:molecular_function | Enzyme |
| GABP01066460.1 | acetyl propionyl- carboxylase subuit alpha | F:molecular_function | Enzyme |
| GABP01066463.1 | pyruvate carboxylase | F:ligase activity | Enzyme |
| GABP01066635.1 | mannitol 1-phosphate dehydrogenase | F:mannitol-1-phosphate 5-dehydrogenase activity; P:homoiothermy; F:oxidoreductase activity; P:metabolic process; C:membrane; C:integral component of membrane; F:ice binding; P:response to freezing; P:oxidation-reduction process; F:hydrolase activity | Enzyme |
| GABP01067645.1 | nucleoside diphosphate | P:small molecule metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; C:cellular_component | Enzyme |
| GABP01069960.1 | enolase | P:generation of precursor metabolites and energy; P:catabolic process; P:cellular component assembly; F:ion binding; F:lyase activity; C:protein complex; P:biosynthetic process; C:plasma membrane; P:immune system process; C:lysosome; P:small molecule metabolic process; C:nucleus; P:carbohydrate metabolic process; C:cytoskeleton; P:autophagy; P:cofactor metabolic process; C:cytosol; C:cytoplasmic, membrane-bounded vesicle; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01070797.1 | lysine decarboxylase | P:cytokinin biosynthetic process; F:hydrolase activity, hydrolyzing N-glycosyl compounds; P:metabolic process; C:membrane; C:integral component of membrane; F:hydrolase activity | Enzyme |

| | | | |
|---|---|---|---|
| GABP01071380.1 | pyruvate kinase | P:protein complex assembly; F:ion binding | Enzyme |
| GABP01071717.1 | polysaccharide deacetylase | P:biological_process; F:molecular_function | Enzyme |
| GABP01072190.1 | oligosaccharyl transferase | P:carbohydrate metabolic process; P:cellular protein modification process; C:endoplasmic reticulum; C:protein complex; P:biosynthetic process; F:molecular_function | Enzyme |
| GABP01072598.1 | glycosyl transferase | F:transferase activity, transferring acyl groups; F:sterol 3-beta-glucosyltransferase activity; P:protein autophosphorylation; F:calmodulin-dependent protein kinase activity; P:protein phosphorylation; C:cellular_component; F:calcium-dependent protein serine/threonine kinase activity; F:nucleotide binding; F:ATP binding; F:calcium ion binding; P:phosphorylation; P:intracellular signal transduction; F:ligase activity; F:transferase activity, transferring glycosyl groups; F:transferase activity, transferring hexosyl groups; P:metabolic process; F:protein kinase activity; C:plasma membrane; P:carbohydrate metabolic process; C:nucleus; P:abscisic acid-activated signaling pathway; F:calmodulin binding; F:metal ion binding; C:cytoplasm; C:membrane; C:integral component of membrane; F:transferase activity; P:peptidyl-serine phosphorylation; F:kinase activity; P:lipid glycosylation | Enzyme |
| GABP01073312.1 | sucrose-phosphate synthase | F:transferase activity, transferring glycosyl groups | Enzyme |
| GABP01074610.1 | adenylosuccinate synthetase | P:small molecule metabolic process; P:cellular nitrogen compound metabolic process; C:intracellular; F:ligase activity | Enzyme |
| GABP01074785.1 | phosphoglyceromutase | P:small molecule metabolic process; P:secondary metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; P:cofactor metabolic process; C:cytosol; P:biosynthetic process; P:cellular nitrogen compound metabolic process; F:isomerase activity | Enzyme |
| GABP01074793.1 | glutathione S-transferase Mu 1-like | P:metabolic process; C:membrane; C:integral component of membrane; F:glutathione transferase activity; F:transferase activity | Enzyme |
| GABP01075186.1 | farnesyl pyrophosphate synthase | C:cytoplasm; P:lipid metabolic process; P:biosynthetic process; F:transferase activity, transferring alkyl or aryl (other than methyl) groups | Enzyme |
| GABP01076020.1 | malonyl -acyl carrier transacylase | F:molecular_function | Enzyme |
| GABP01076065.1 | phosphoribosylformylglycinamidine cyclo-ligase | P:biological_process; F:ligase activity | Enzyme |
| GABP01076386.1 | GDP-mannose 4,6 dehydratase | F:lyase activity; F:ion binding; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01078669.1 | polyketide synthase | F:transferase activity, transferring acyl groups | Enzyme |
| GABP01079572.1 | cysteine hydrolase | F:molecular_function | Enzyme |
| GABP01079601.1 | alcohol dehydrogenase | F:oxidoreductase activity | Enzyme |
| GABP01079700.1 | 3-deoxy-7-phosphoheptulonate synthase | P:aromatic amino acid family biosynthetic process; F:3-deoxy-7-phosphoheptulonate synthase activity; F:transferase activity | Enzyme |
| GABP01080241.1 | phosphoribosylaminoimidazolecarboxamide formyltransferase IMP cyclohydrolase | P:small molecule metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:cellular_component; F:molecular_function | Enzyme |

| | | | |
|---|---|---|---|
| GABP01080967.1 | malate synthase G | P:small molecule metabolic process; F:transferase activity, transferring acyl groups | Enzyme |
| GABP01080969.1 | malate synthase G | P:small molecule metabolic process; P:carbohydrate metabolic process; F:transferase activity, transferring acyl groups; P:generation of precursor metabolites and energy; C:cytoplasm; F:ion binding | Enzyme |
| GABP01081420.1 | aminotransferase | F:molecular_function | Enzyme |
| GABP01081620.1 | type I polyketide synthase | P:biological_process; F:transferase activity, transferring acyl groups | Enzyme |
| GABP01083301.1 | isocitrate dehydrogenase | P:small molecule metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; F:ion binding | Enzyme |
| GABP01085219.1 | lysine decarboxylase | P:cytokinin biosynthetic process; F:hydrolase activity, hydrolyzing N-glycosyl compounds; P:metabolic process; C:membrane; C:integral component of membrane; F:hydrolase activity | Enzyme |
| GABP01085237.1 | polyketide synthase | F:transferase activity, transferring acyl groups | Enzyme |
| GABP01085786.1 | type I fatty acid partial | P:biological_process; F:transferase activity, transferring acyl groups | Enzyme |
| GABP01085844.1 | oxidoreductase | F:oxidoreductase activity; P:metabolic process; P:oxidation-reduction process; F:trans-1,2-dihydrobenzene-1,2-diol dehydrogenase activity | Enzyme |
| GABP01086137.1 | UTP--glucose-1-phosphate uridylyltransferase | P:small molecule metabolic process; P:carbohydrate metabolic process; F:nucleotidyltransferase activity; P:generation of precursor metabolites and energy; P:catabolic process; C:cytosol; P:biosynthetic process; F:isomerase activity | Enzyme |
| GABP01086285.1 | anthranilate synthase | P:biological_process; F:lyase activity | Enzyme |
| GABP01086975.1 | N-acylneuraminate cytidylyltransferase | F:molecular_function | Enzyme |
| GABP01087688.1 | pyruvate kinase | F:ion binding | Enzyme |
| GABP01087789.1 | isocitrate dehydrogenase | P:small molecule metabolic process; F:oxidoreductase activity; F:ion binding | Enzyme |
| GABP01088211.1 | aminotransferase | F:molecular_function | Enzyme |
| GABP01089091.1 | type I fatty acid partial | P:biological_process; F:transferase activity, transferring acyl groups; F:ion binding | Enzyme |
| GABP01089891.1 | Amidohydrolase 1 | F:hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds; P:transmembrane transport; P:metabolic process; C:membrane; C:integral component of membrane; F:hydrolase activity | Enzyme |
| GABP01090699.1 | isocitrate dehydrogenase | P:small molecule metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; F:ion binding | Enzyme |
| GABP01092030.1 | Phosphoglycerate kinase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; C:cytoplasm; F:oxidoreductase activity; P:cofactor metabolic process; F:ion binding; P:cellular nitrogen compound metabolic process; C:organelle; F:kinase activity | Enzyme |
| GABP01092929.1 | polyketide synthase | F:transferase activity, transferring acyl groups | Enzyme |
| GABP01093321.1 | non-ribosomal peptide synthetase | F:molecular_function | Enzyme |

| | | | |
|---|---|---|---|
| GABP01094844.1 | dehydrogenase | F:inositol 2-dehydrogenase activity; F:oxidoreductase activity; P:metabolic process; C:membrane; C:integral component of membrane; P:oxidation-reduction process | Enzyme |
| GABP01095199.1 | glycoside hydrolase family 43 | P:carbohydrate metabolic process; F:catalytic activity; P:metabolic process; F:hydrolase activity, hydrolyzing O-glycosyl compounds; F:carbohydrate binding; F:hydrolase activity; F:hydrolase activity, acting on glycosyl bonds | Enzyme |
| GABP01096166.1 | non-specific lipid-transfer | F:transferase activity, transferring acyl groups; F:transferase activity, transferring acyl groups other than amino-acyl groups; P:metabolic process; F:lipid binding; P:steroid biosynthetic process; C:cellular_component; P:transport; P:lipid transport; C:peroxisome; F:propanoyl-CoA C-acyltransferase activity; F:sterol binding; C:cytoplasm; F:catalytic activity; C:mitochondrion; P:phospholipid transport; F:transferase activity; P:cholesterol transport; F:sterol transporter activity | Enzyme |
| GABP01098062.1 | succinate | F:oxidoreductase activity; F:ion binding; C:cellular_component | Enzyme |
| GABP01098066.1 | succinate | F:oxidoreductase activity | Enzyme |
| GABP01099683.1 | type I polyketide synthase | F:transferase activity, transferring acyl groups | Enzyme |
| GABP01099890.1 | lysine decarboxylase | P:cytokinin biosynthetic process; F:hydrolase activity, hydrolyzing N-glycosyl compounds; P:metabolic process; C:membrane; C:integral component of membrane; F:hydrolase activity | Enzyme |
| GABP01100510.1 | long chain acyl- synthetase peroxisomal-like | F:catalytic activity; P:metabolic process | Enzyme |
| GABP01100685.1 | type I fatty acid partial | F:transferase activity, transferring acyl groups | Enzyme |
| GABP01101449.1 | S-adenosyl-L-homocysteine NAD binding domain | F:oxidoreductase activity | Enzyme |
| GABP01101451.1 | 3-phosphoglycerate dehydrogenase | F:oxidoreductase activity | Enzyme |
| GABP01102272.1 | 3-ketoacyl- peroxisomal | P:small molecule metabolic process; C:peroxisome; F:transferase activity, transferring acyl groups; F:oxidoreductase activity; P:catabolic process; C:mitochondrion; P:lipid metabolic process | Enzyme |
| GABP01104755.1 | Pseudouridine synthase | F:molecular_function | Enzyme |
| GABP01106764.1 | acyl- dehydrogenase | F:acyl-CoA dehydrogenase activity; F:oxidoreductase activity, acting on the CH-CH group of donors; F:flavin adenine dinucleotide binding; F:oxidoreductase activity; P:metabolic process; C:membrane; C:integral component of membrane; P:oxidation-reduction process; F:ADP binding | Enzyme |
| GABP01108277.1 | glycerol-3-phosphate dehydrogenase [NAD(+)] cytoplasmic | F:oxidoreductase activity | Enzyme |

| | | | |
|---|---|---|---|
| GABP01108279.1 | glycerol-3-phosphate dehydrogenase (NAD(+)) | P:glycerol-3-phosphate metabolic process; F:protein homodimerization activity; F:oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor; C:glycerol-3-phosphate dehydrogenase complex; F:glycerol-3-phosphate dehydrogenase [NAD+] activity; P:carbohydrate metabolic process; C:cytoplasm; P:glycerol-3-phosphate catabolic process; F:oxidoreductase activity; F:NAD binding; C:membrane; C:integral component of membrane; P:oxidation-reduction process | Enzyme |
| GABP01108283.1 | glycerol-3-phosphate dehydrogenase [NAD(+)] | P:carbohydrate metabolic process; C:cytoplasm; F:oxidoreductase activity; P:catabolic process; C:protein complex | Enzyme |
| GABP01108362.1 | glyceraldehyde-3-phosphate dehydrogenase | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; P:catabolic process; F:oxidoreductase activity; P:cofactor metabolic process; F:ion binding; F:lyase activity; C:cytosol; C:protein complex; P:cellular nitrogen compound metabolic process | Enzyme |
| GABP01108783.1 | aspartate ammonia-lyase | P:small molecule metabolic process; F:lyase activity | Enzyme |
| GABP01108981.1 | Alcohol dehydrogenase class-3 | P:small molecule metabolic process; F:oxidoreductase activity; F:ion binding; C:chromosome | Enzyme |
| GABP01109287.1 | starch branching enzyme 4 | P:carbohydrate metabolic process; F:transferase activity, transferring glycosyl groups; C:plastid; F:hydrolase activity, acting on glycosyl bonds | Enzyme |
| GABP01109843.1 | type I fatty acid partial | P:biological_process; F:transferase activity, transferring acyl groups | Enzyme |
| GABP01111306.1 | isocitrate dehydrogenase | P:small molecule metabolic process; F:oxidoreductase activity; F:ion binding | Enzyme |
| GABP01038659.1 | hybrid signal transduction histidine | C:cellular_component; F:signal transducer activity; F:kinase activity | Membrane related |
| GABP01003193.1 | Tetratricopeptide TPR_1 repeat-containing | F:ATP binding; P:signal transduction; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | Membrane related |
| GABP01008743.1 | capsule biosynthesis phosphatase | F:nucleotidyltransferase activity; F:glucose-1-phosphate thymidylyltransferase activity; C:membrane; C:integral component of membrane; P:biosynthetic process; F:transferase activity | Membrane related |
| GABP01014721.1 | hybrid sensor histidine kinase response regulator | F:signal transducer activity; F:kinase activity | Membrane related |
| GABP01015257.1 | calcium-transporting | F:ATPase activity; F:transmembrane transporter activity; F:ion binding; C:endoplasmic reticulum; C:plasma membrane | Membrane related |
| GABP01017080.1 | cellulosome enzyme dockerin type I | C:membrane; C:integral component of membrane | Membrane related |
| GABP01025959.1 | hybrid signal transduction histidine | C:cellular_component; F:signal transducer activity; F:kinase activity | Membrane related |
| GABP01025961.1 | hybrid signal transduction histidine | C:cellular_component; F:signal transducer activity; F:kinase activity | Membrane related |
| GABP01025964.1 | hybrid signal transduction histidine | C:cellular_component; F:signal transducer activity; F:kinase activity | Membrane related |
| GABP01026224.1 | transmembrane partial | C:membrane; C:integral component of membrane | Membrane related |
| GABP01030583.1 | ABC transporter permease | C:membrane; C:integral component of membrane; P:transport; C:plasma membrane; F:transporter activity | Membrane related |

| GABP01039934.1 | vacuolar membrane - | C:cellular_component | Membrane related |
|---|---|---|---|
| GABP01051835.1 | ABC transporter ATP-binding | F:ATPase activity; P:metabolic process; F:protein kinase activity; F:xenobiotic-transporting ATPase activity; P:protein phosphorylation; P:transport; F:nucleotide binding; F:ATP binding; P:drug transmembrane transport; C:membrane; C:integral component of membrane; P:xenobiotic transport; P:organic phosphonate transport; P:organophosphate ester transport; F:ATPase-coupled organic phosphonate transmembrane transporter activity; F:hydrolase activity | Membrane related |
| GABP01053463.1 | hybrid signal transduction histidine | P:signal transduction; F:molecular_function | Membrane related |
| GABP01053464.1 | hybrid signal transduction histidine | P:signal transduction; F:molecular_function | Membrane related |
| GABP01053467.1 | hybrid signal transduction histidine | F:ion binding; P:cellular protein modification process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:kinase activity; F:signal transducer activity; C:intracellular | Membrane related |
| GABP01055574.1 | cell surface p43 | | Membrane related |
| GABP01055792.1 | hybrid signal transduction histidine | P:biological_process; F:molecular_function | Membrane related |
| GABP01063344.1 | hybrid signal transduction histidine | F:ion binding; P:cellular protein modification process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:kinase activity; F:signal transducer activity; C:intracellular | Membrane related |
| GABP01085658.1 | coatomer subunit beta | P:vesicle-mediated transport; C:cytoplasmic, membrane-bounded vesicle; C:Golgi apparatus; C:protein complex; F:molecular_function | Membrane related |
| GABP01012170.1 | mago nashi homolog | P:embryo development; P:locomotion; C:protein complex; P:cellular nitrogen compound metabolic process; C:nucleolus; C:plastid; P:reproduction; C:nucleoplasm | Misc. nucleus |

| | | | |
|---|---|---|---|
| GABP01019866.1 | poly-glutamine tract-binding | F:phosphothreonine binding; C:nuclear speck; P:negative regulation of protein catabolic process; P:protein peptidyl-prolyl isomerization; P:positive regulation of transcription from RNA polymerase II promoter; C:neuron projection; P:negative regulation of protein binding; P:positive regulation of neuron apoptotic process; P:positive regulation of GTPase activity; P:negative regulation of neuron apoptotic process; C:nucleoplasm; F:motor activity; P:fatty acid beta-oxidation; P:positive regulation of ubiquitin-protein transferase activity; P:regulation of cytokinesis; C:cytosol; F:phosphoserine binding; P:calcium-mediated signaling; F:mitogen-activated protein kinase kinase binding; F:isomerase activity; P:negative regulation of type I interferon production; P:protein import into peroxisome matrix; P:neuron differentiation; P:regulation of mitotic nuclear division; P:synapse organization; P:regulation of pathway-restricted SMAD protein phosphorylation; P:metabolic process; P:regulation of signal transduction by p53 class mediator; P:positive regulation of canonical Wnt signaling pathway; F:beta-catenin binding; P:cell cycle; C:intracellular; P:photoperiodism, flowering; C:nucleus; F:peptidyl-prolyl cis-trans isomerase activity; F:protein binding; C:cytoplasm; P:negative regulation of ERK1 and ERK2 cascade; C:mitochondrion; P:positive regulation of protein phosphorylation; P:negative regulation of cell motility; P:positive regulation of protein dephosphorylation; P:positive regulation of cell growth involved in cardiac muscle cell development; C:midbody; F:GTPase activating protein binding; P:protein stabilization; P:negative regulation of transforming growth factor beta receptor signaling pathway; P:regulation of protein localization to nucleus | Misc. nucleus |
| GABP01024920.1 | nuclear factor NF2 | P:Mo-molybdopterin cofactor biosynthetic process; P:molybdopterin cofactor biosynthetic process | Misc. nucleus |
| GABP01027339.1 | ankyrin repeat domain-containing 49 | C:nucleus; P:positive regulation of transcription, DNA-templated; F:fatty-acyl-CoA binding; P:spermatogenesis; C:membrane; C:integral component of membrane; P:cell differentiation; C:cellular_component; F:molecular_function | Misc. nucleus |
| GABP01038248.1 | calreticulin | P:cell proliferation; P:protein complex assembly; F:ion binding; P:cell adhesion; P:nucleocytoplasmic transport; P:vesicle-mediated transport; P:cell death; C:cytosol; F:mRNA binding; P:protein maturation; P:signal transduction; P:cell morphogenesis; C:endoplasmic reticulum; F:unfolded protein binding; C:protein complex; P:protein folding; P:DNA metabolic process; P:aging; P:immune system process; P:cell cycle; P:response to stress; P:translation; P:cell motility; C:nucleus; F:DNA binding; C:extracellular space; P:homeostatic process; P:cell differentiation; C:cytoplasmic, membrane-bounded vesicle; F:enzyme binding | Misc. nucleus |
| GABP01052767.1 | nucleolar 56-like | F:RNA binding; P:ribosome biogenesis; P:cellular nitrogen compound metabolic process; C:nucleolus | Misc. nucleus |
| GABP01062473.1 | Nucleolar | F:RNA binding; P:ribosome biogenesis; P:cellular nitrogen compound metabolic process; C:nucleolus | Misc. nucleus |
| GABP01086337.1 | casein kinase I | C:nucleus; P:vesicle-mediated transport; C:cytoplasm; P:signal transduction; P:cell morphogenesis; F:ion binding; P:cellular protein modification process; F:kinase activity | Misc. nucleus |

| | | | |
|---|---|---|---|
| GABP01094724.1 | T-complex 1 subunit delta | F:RNA binding; C:microtubule organizing center; F:ion binding; C:cytosol; F:unfolded protein binding; C:protein complex; C:extracellular region; P:protein folding; P:transport; P:response to stress; P:reproduction; C:nucleoplasm | Misc. nucleus |
| GABP01101189.1 | apoptosis-inducing factor homolog A-like | P:nucleocytoplasmic transport; C:nucleus; F:oxidoreductase activity; P:positive regulation of apoptotic process; C:mitochondrion; F:structural constituent of nuclear pore; P:oxidation-reduction process; P:response to singlet oxygen; C:nuclear pore | Misc. nucleus |
| GABP01104632.1 | T-complex 1 subunit gamma | C:nucleus; F:nucleotidyltransferase activity; P:lipid metabolic process; F:ion binding; C:cytosol; F:unfolded protein binding; C:protein complex; P:protein folding; F:kinase activity | Misc. nucleus |
| GABP01105647.1 | cullin-associated NEDD8-dissociated 1 | C:cell | Misc. nucleus |
| GABP01106344.1 | diguanylate cyclase | P:regulation of Rho protein signal transduction; F:zinc ion binding; P:metabolic process; F:guanyl-nucleotide exchange factor activity; F:methyltransferase activity; F:Rho guanyl-nucleotide exchange factor activity; F:nucleotide binding; F:GTP binding; C:nucleus; P:translational termination; F:catalytic activity; F:translation release factor activity; F:metal ion binding; F:GTPase activity; C:membrane; C:integral component of membrane; P:intracellular signal transduction; P:methylation | Misc. nucleus |
| GABP01114472.1 | Endoplasmin | P:cytoskeleton organization; P:catabolic process; P:cellular component assembly; F:ion binding; P:anatomical structure development; P:cellular protein modification process; C:Golgi apparatus; C:extracellular region; C:vacuole; C:plastid; F:RNA binding; P:vesicle-mediated transport; P:cell death; C:cytosol; P:signal transduction; C:endoplasmic reticulum; F:unfolded protein binding; P:protein folding; C:plasma membrane; P:response to stress; P:immune system process; P:growth; C:nucleus; C:mitochondrion; C:cytoplasmic, membrane-bounded vesicle; F:enzyme binding | Misc. nucleus |
| GABP01001132.1 | type I fatty acid partial | P:biological_process; F:transferase activity, transferring acyl groups | Mitochondria |
| GABP01005988.1 | succinyl- :3-ketoacid coenzyme A transferase mitochondrial | P:small molecule metabolic process; P:catabolic process; C:mitochondrion; F:molecular_function | Mitochondria |
| GABP01008987.1 | long-chain-acyl- synthetase | F:ligase activity | Mitochondria |
| GABP01009432.1 | 2-isopropylmalate synthase | F:transferase activity, transferring acyl groups; P:cellular amino acid metabolic process; P:biosynthetic process | Mitochondria |
| GABP01009663.1 | succinate-- ligase subunit beta | P:biological_process; F:ion binding; C:organelle; C:intracellular; F:ligase activity | Mitochondria |
| GABP01009670.1 | succinate-Coenzyme A beta subunit | P:small molecule metabolic process; P:generation of precursor metabolites and energy; P:sulfur compound metabolic process; C:mitochondrion; P:cofactor metabolic process; F:ion binding; F:ligase activity | Mitochondria |
| GABP01011204.1 | ATP synthase subunit mitochondrial | F:ATPase activity; P:cell proliferation; F:transmembrane transporter activity; F:ion binding; C:protein complex; P:biosynthetic process; C:extracellular region; C:plasma membrane; P:small molecule metabolic process; C:nucleus; F:RNA binding; C:mitochondrion; P:lipid metabolic process; P:cellular nitrogen compound metabolic process | Mitochondria |

| | | | |
|---|---|---|---|
| GABP01011205.1 | ATP synthase subunit alpha | F:ATPase activity; P:small molecule metabolic process; F:transmembrane transporter activity; C:mitochondrion; F:ion binding; C:protein complex; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:plasma membrane | Mitochondria |
| GABP01012079.1 | acetyl-coa carboxylase | P:biological_process; C:organelle; C:intracellular; F:ligase activity | Mitochondria |
| GABP01012084.1 | acetyl- carboxylase | P:small molecule metabolic process; P:lipid metabolic process; F:ion binding; P:biosynthetic process; C:cellular_component; F:ligase activity | Mitochondria |
| GABP01012121.1 | acetyl-coa carboxylase | F:ligase activity | Mitochondria |
| GABP01012124.1 | acetyl- carboxylase | P:biological_process; F:ion binding; C:cellular_component; F:ligase activity | Mitochondria |
| GABP01016767.1 | malate mitochondrial | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity | Mitochondria |
| GABP01017923.1 | ATP-binding cassette sub-family B member mitochondrial | P:biological_process; C:cytoplasm; C:organelle; F:molecular_function | Mitochondria |
| GABP01022200.1 | peroxiredoxin- mitochondrial | P:catabolic process; P:biosynthetic process; P:response to stress; P:small molecule metabolic process; C:nucleus; C:peroxisome; F:DNA binding; C:extracellular space; P:cell death; F:oxidoreductase activity; P:homeostatic process; C:mitochondrion; P:cofactor metabolic process; C:cytosol; P:cellular nitrogen compound metabolic process; F:enzyme regulator activity | Mitochondria |
| GABP01025042.1 | NAD(P) mitochondrial | F:oxidoreductase activity; C:mitochondrion; P:transport | Mitochondria |
| GABP01025043.1 | NAD(P) transhydrogenase subunit alpha | F:oxidoreductase activity | Mitochondria |
| GABP01028131.1 | chaperonin CPN60- mitochondrial | P:protein complex assembly; C:mitochondrion; F:ion binding; C:cytosol; C:Golgi apparatus; C:ribosome; P:protein folding; C:vacuole; C:plasma membrane; C:plastid; P:mitochondrion organization; P:response to stress | Mitochondria |
| GABP01029170.1 | Nde1p | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; P:catabolic process; C:mitochondrion; P:cofactor metabolic process; P:cellular nitrogen compound metabolic process; P:aging | Mitochondria |
| GABP01030459.1 | succinyl- ligase [GDP-forming] subunit mitochondrial | P:small molecule metabolic process; P:generation of precursor metabolites and energy; F:RNA binding; P:sulfur compound metabolic process; C:mitochondrion; P:cofactor metabolic process; C:cytosol; P:biosynthetic process; P:cellular nitrogen compound metabolic process; C:extracellular region; C:plasma membrane; F:ligase activity | Mitochondria |
| GABP01030460.1 | succinyl- -synthetase alpha SCSA | P:small molecule metabolic process; P:embryo development; P:generation of precursor metabolites and energy; P:sulfur compound metabolic process; C:mitochondrion; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:ligase activity | Mitochondria |
| GABP01031839.1 | probable citrate synthase mitochondrial | P:small molecule metabolic process; F:transferase activity, transferring acyl groups | Mitochondria |
| GABP01040271.1 | methylmalonate-semialdehyde dehydrogenase [acylating] mitochondrial | F:oxidoreductase activity | Mitochondria |

| GABP01040688.1 | atp-citrate synthase | P:small molecule metabolic process; F:transferase activity, transferring acyl groups; P:sulfur compound metabolic process; C:cytoplasm; P:cofactor metabolic process; P:lipid metabolic process; F:ion binding; P:biosynthetic process | Mitochondria |
|---|---|---|---|
| GABP01043652.1 | ketol-acid reductoisomerase | F:oxidoreductase activity; P:cellular amino acid metabolic process; P:biosynthetic process; F:isomerase activity | Mitochondria |
| GABP01047531.1 | 3-hydroxyacyl- dehydrogenase | P:lipid metabolic process; F:molecular_function | Mitochondria |
| GABP01050566.1 | Malonyl- -acyl carrier mitochondrial | P:biological_process; C:cytoplasm; C:organelle; F:molecular_function | Mitochondria |
| GABP01053511.1 | L-lactate dehydrogenase (cytochrome) | P:small molecule metabolic process; F:oxidoreductase activity; C:mitochondrion | Mitochondria |
| GABP01053518.1 | mitochondrial cytochrome | C:mitochondrial intermembrane space; F:FMN binding; F:metal ion binding; F:catalytic activity; F:oxidoreductase activity; C:mitochondrion; F:L-lactate dehydrogenase (cytochrome) activity; P:oxidation-reduction process; P:lactate metabolic process; F:heme binding | Mitochondria |
| GABP01061222.1 | isocitrate lyase | P:small molecule metabolic process; P:carbohydrate metabolic process; F:transferase activity, transferring acyl groups; C:cytoplasm; F:lyase activity | Mitochondria |
| GABP01061223.1 | isocitrate lyase | P:small molecule metabolic process; P:carbohydrate metabolic process; F:transferase activity, transferring acyl groups; C:cytoplasm; F:lyase activity | Mitochondria |
| GABP01066946.1 | UPF0160 mitochondrial | | Mitochondria |
| GABP01069560.1 | malonyl -acyl carrier transacylase | F:molecular_function | Mitochondria |
| GABP01070139.1 | cysteine desulfurase | C:nucleus; P:sulfur compound metabolic process; C:mitochondrion; P:cofactor metabolic process; P:cellular component assembly; F:ion binding; P:biosynthetic process | Mitochondria |
| GABP01074658.1 | 2Fe-2S partial | P:small molecule metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; C:mitochondrion; F:ion binding | Mitochondria |
| GABP01075027.1 | adenylate kinase | P:small molecule metabolic process; C:mitochondrion; F:ion binding; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:kinase activity | Mitochondria |
| GABP01077515.1 | malate mitochondrial | P:small molecule metabolic process; P:carbohydrate metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; C:mitochondrion; F:ion binding; C:extracellular region; P:response to stress | Mitochondria |
| GABP01082482.1 | Succinate dehydrogenase [ubiquinone] flavo subunit mitochondrial | P:small molecule metabolic process; P:generation of precursor metabolites and energy; F:oxidoreductase activity; C:mitochondrion; F:ion binding | Mitochondria |
| GABP01084230.1 | serine mitochondrial | C:mitochondrion; P:cofactor metabolic process; F:ion binding; C:cytosol; P:cellular amino acid metabolic process; F:methyltransferase activity; P:cellular nitrogen compound metabolic process; C:plastid | Mitochondria |
| GABP01090154.1 | malate mitochondrial | P:small molecule metabolic process; F:oxidoreductase activity; C:cell wall; C:mitochondrion; F:ion binding; C:extracellular region; C:plastid; P:response to stress | Mitochondria |
| GABP01091444.1 | ATP synthase subunit beta | F:ATPase activity; P:small molecule metabolic process; F:transmembrane transporter activity; C:mitochondrion; F:ion binding; C:protein complex; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:plasma membrane | Mitochondria |

| | | | |
|---|---|---|---|
| GABP01091445.1 | ATP synthase subunit mitochondrial | F:ATPase activity; F:transmembrane transporter activity; F:ion binding; P:anatomical structure formation involved in morphogenesis; P:cell adhesion; C:protein complex; P:biosynthetic process; C:extracellular region; C:plasma membrane; P:cell motility; P:small molecule metabolic process; C:nucleus; P:homeostatic process; C:mitochondrion; P:lipid metabolic process; P:cell differentiation; P:cellular nitrogen compound metabolic process | Mitochondria |
| GABP01091945.1 | mitochondrial carrier | F:structural constituent of ribosome; C:cellular_component; P:translation | Mitochondria |
| GABP01098068.1 | succinate | F:oxidoreductase activity; F:ion binding; C:cellular_component | Mitochondria |
| GABP01098642.1 | succinate | F:oxidoreductase activity; F:ion binding; C:cellular_component | Mitochondria |
| GABP01105805.1 | mitochondrial | F:methyltransferase activity; C:plastid | Mitochondria |
| GABP01024272.1 | heterogeneous nuclear ribonucleo hrp1 | F:molecular_function | mRNA processing |
| GABP01059033.1 | heterogeneous nuclear ribonucleo U 1 | C:intracellular ribonucleoprotein complex; C:membrane; C:integral component of membrane; C:viral nucleocapsid | mRNA processing |
| GABP01001943.1 | cleavage stimulation factor subunit 2-like | F:mRNA binding; C:protein complex; P:biosynthetic process; P:mRNA processing; P:growth; C:nucleoplasm | mRNA processing |
| GABP01002608.1 | heterogeneous nuclear ribonucleo U 1 | F:nucleic acid binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid; F:protein dimerization activity; P:RNA processing; P:response to virus | mRNA processing |
| GABP01003623.1 | splicing rnp complex component | P:7-methylguanosine mRNA capping; P:mRNA processing; P:transport; F:RNA polymerase II transcription cofactor activity; P:regulation of transcription from RNA polymerase II promoter; F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA 7-methylguanosine cap binding; F:RNA binding; C:cytoplasm; C:mediator complex; C:cytosol; F:mRNA binding; P:mRNA transport | mRNA processing |
| GABP01005163.1 | RNA-binding 25 | P:cell death; F:mRNA binding; P:mRNA processing; C:nucleoplasm | mRNA processing |
| GABP01005490.1 | SART-1 family | C:nucleus; P:anatomical structure development; P:mRNA processing | mRNA processing |
| GABP01005584.1 | pre-mRNA-splicing factor SYF1 | C:nucleus; P:ribonucleoprotein complex assembly; C:protein complex; P:mRNA processing | mRNA processing |
| GABP01006953.1 | u4 u6 small nuclear ribonucleo prp4 | P:embryo development | mRNA processing |
| GABP01009340.1 | serine arginine-rich splicing factor 2-like | P:nucleocytoplasmic transport; F:RNA binding; P:mRNA processing; P:biosynthetic process; C:extracellular region; C:nucleoplasm | mRNA processing |
| GABP01011746.1 | small nuclear ribonucleo Sm D3 | F:RNA binding; C:cytoplasm; P:ribonucleoprotein complex assembly; C:protein complex; P:mRNA processing; C:nucleolus; C:nucleoplasm | mRNA processing |
| GABP01012312.1 | regulator of nonsense transcripts 1 ( RENT1) | C:nucleus; F:DNA binding; F:RNA binding; C:cytoplasm; P:chromosome organization; P:homeostatic process; P:nucleobase-containing compound catabolic process; F:helicase activity; C:chromosome; P:DNA metabolic process; P:translation | mRNA processing |
| GABP01016262.1 | pre-mRNA-splicing factor CWC22 homolog | C:nucleus; F:RNA binding; P:mRNA processing | mRNA processing |
| GABP01016574.1 | Heterogeneous nuclear ribonucleo U | C:intracellular | mRNA processing |

| | | | |
|---|---|---|---|
| GABP01017436.1 | Crooked neck 1 | C:nucleus; P:ribonucleoprotein complex assembly; C:protein complex; P:mRNA processing | mRNA processing |
| GABP01017553.1 | Heterogeneous nuclear ribonucleo U 1 | F:nucleic acid binding; C:nucleus; C:intracellular ribonucleoprotein complex; P:metabolic process; C:viral nucleocapsid; F:protein dimerization activity; F:hydrolase activity | mRNA processing |
| GABP01019808.1 | heterogeneous nuclear ribonucleo U 1 | F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:membrane; C:integral component of membrane; C:viral nucleocapsid | mRNA processing |
| GABP01021492.1 | polyadenylate-binding 1 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; C:intracellular ribonucleoprotein complex; C:cytoplasm; C:virion; C:viral nucleocapsid | mRNA processing |
| GABP01024209.1 | heterogeneous nuclear ribonucleo 1-like | F:nucleotide binding; P:regulation of translation; F:nucleic acid binding; F:RNA binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid | mRNA processing |
| GABP01024273.1 | heterogeneous nuclear ribonucleo hrp1 | F:nucleotide binding; F:nucleic acid binding; P:biological_process; F:RNA binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid | mRNA processing |
| GABP01024303.1 | splicing factor u2af large subunit, putative | F:molecular_function | mRNA processing |
| GABP01026270.1 | heterogeneous nuclear ribonucleo U 1 | C:nucleus; C:intracellular ribonucleoprotein complex; C:membrane; C:integral component of membrane; C:viral nucleocapsid | mRNA processing |
| GABP01027383.1 | polyadenylate binding | F:molecular_function | mRNA processing |
| GABP01027657.1 | U5 small nuclear ribonucleo 200 kDa helicase | F:ATPase activity; F:RNA binding; P:ribonucleoprotein complex assembly; F:ion binding; P:cell differentiation; F:helicase activity; P:mRNA processing; C:nucleoplasm | mRNA processing |
| GABP01027816.1 | arginine/serine-rich splicing factor4-like | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding; F:metal ion binding; C:spliceosomal complex; F:methyltransferase activity; C:Prp19 complex; F:transferase activity; P:mRNA splicing, via spliceosome; P:methylation | mRNA processing |
| GABP01028262.1 | heterogeneous nuclear ribonucleo 1-like | P:regulation of postsynaptic membrane potential; P:transmembrane transport; P:ion transmembrane transport; F:sodium channel activity; F:voltage-gated ion channel activity; P:transport; F:voltage-gated sodium channel activity; F:nucleotide binding; F:nucleic acid binding; P:sodium ion transport; F:ion channel activity; P:ion transport; C:membrane; C:integral component of membrane; P:regulation of ion transmembrane transport; C:voltage-gated sodium channel complex; P:sodium ion transmembrane transport | mRNA processing |
| GABP01028898.1 | Heterogeneous nuclear ribonucleo U 1 | C:nucleus; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid; P:RNA processing; P:response to virus | mRNA processing |
| GABP01029384.1 | pre-mRNA-splicing factor ISY1 homolog | C:nucleus; P:ribonucleoprotein complex assembly; C:protein complex; P:mRNA processing | mRNA processing |
| GABP01030863.1 | U2 snRNP auxiliary small | C:nucleus; F:RNA binding; F:ion binding; C:protein complex; P:mRNA processing | mRNA processing |
| GABP01031268.1 | heterogeneous nuclear ribonucleo R | F:molecular_function | mRNA processing |
| GABP01031623.1 | 3 -5 exonuclease | F:nucleic acid binding; P:nucleic acid phosphodiester bond hydrolysis; P:RNA phosphodiester bond hydrolysis, exonucleolytic; F:3'-5' exonuclease activity; C:nucleus; F:exonuclease activity; C:cytoplasm; F:ribonuclease D activity; P:nucleobase-containing compound metabolic process; F:hydrolase activity | mRNA processing |

| | | | |
|---|---|---|---|
| GABP01031756.1 | heterogeneous nuclear ribonucleo U 1 | C:intracellular ribonucleoprotein complex; C:precatalytic spliceosome; P:metabolic process; C:viral nucleocapsid; F:mRNA binding; P:mRNA splicing, via spliceosome; F:hydrolase activity | mRNA processing |
| GABP01031793.1 | polyadenylate-binding 2-like | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; C:cytoplasm; F:metal ion binding; C:cytosol | mRNA processing |
| GABP01031850.1 | regulator of nonsense transcripts 1 | F:DNA binding; P:nucleic acid phosphodiester bond hydrolysis; F:ATP binding; F:zinc ion binding; P:nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; C:cytoplasm; F:endonuclease activity; P:metabolic process; F:helicase activity | mRNA processing |
| GABP01032057.1 | Polyadenylate-binding 1 | F:nucleotide binding; P:regulation of translation; F:nucleic acid binding; F:translation initiation factor activity; C:nucleus; P:regulation of nuclear-transcribed mRNA poly(A) tail shortening; F:RNA binding; P:nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; C:cytoplasm; P:mRNA processing | mRNA processing |
| GABP01032288.1 | Heterogeneous nuclear ribonucleo U 1 | C:nucleus; P:RNA processing; P:response to virus | mRNA processing |
| GABP01032289.1 | Heterogeneous nuclear ribonucleo U 1 | C:nucleus; P:RNA processing; P:response to virus | mRNA processing |
| GABP01034257.1 | THO complex subunit 1 | P:mRNA export from nucleus; C:nucleus; P:signal transduction; P:gene silencing by RNA; C:THO complex; P:defense response to fungus; C:plastid; P:production of ta-siRNAs involved in RNA interference | mRNA processing |
| GABP01034532.1 | serine arginine-rich splicing factor SR34A-like isoform X1 | F:identical protein binding; F:poly(A) binding; F:eukaryotic initiation factor 4G binding; F:RNA polymerase II core binding; F:regulatory region RNA binding; P:mRNA splicing, via spliceosome; F:nucleotide binding; P:mRNA export from nucleus; P:positive regulation of transcription elongation from RNA polymerase II promoter; F:nucleic acid binding; P:negative regulation of termination of RNA polymerase II transcription, poly(A)-coupled; C:nucleus; P:translational termination; C:cytoplasm; F:mRNA binding; P:negative regulation of translation | mRNA processing |
| GABP01035247.1 | CUGBP Elav-like family member 2 isoform X1 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; C:cytoplasm; F:metal ion binding | mRNA processing |
| GABP01036338.1 | polyadenylate binding | F:molecular_function | mRNA processing |
| GABP01038926.1 | pre-mRNA proceesing factor 40 A-like | C:nucleus; P:biological_process; F:molecular_function | mRNA processing |
| GABP01040243.1 | heterogeneous nuclear ribonucleo U 1 isoform X2 | C:intracellular | mRNA processing |
| GABP01040246.1 | heterogeneous nuclear ribonucleo U isoform X2 | C:nucleus | mRNA processing |
| GABP01041389.1 | pre-mRNA-processing factor 40 | C:intracellular | mRNA processing |
| GABP01042904.1 | splicing factor 3a | C:nucleus; F:molecular_function | mRNA processing |
| GABP01042973.1 | cleavage stimulation factor subunit 77 isoform X1 | P:cellular nitrogen compound metabolic process | mRNA processing |

| | | | |
|---|---|---|---|
| GABP01043793.1 | Heterogeneous nuclear ribonucleo U 1 | P:metabolic process; F:helicase activity; P:translational initiation; C:eukaryotic translation initiation factor 3 complex; F:ATP-dependent RNA helicase activity; P:translation; F:nucleic acid binding; F:translation initiation factor activity; F:ATP binding; C:cytoplasm; C:membrane; C:integral component of membrane; P:RNA secondary structure unwinding; F:hydrolase activity | mRNA processing |
| GABP01044635.1 | heterogeneous nuclear ribonucleo U 1 | F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:membrane; C:integral component of membrane; C:viral nucleocapsid | mRNA processing |
| GABP01044730.1 | heterogeneous nuclear ribonucleo U 1 | C:nucleus; F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:virion; C:membrane; C:integral component of membrane; C:viral nucleocapsid; P:RNA processing; P:response to virus | mRNA processing |
| GABP01045189.1 | Heterogeneous nuclear ribonucleo U 1 | C:intracellular ribonucleoprotein complex; C:membrane; C:integral component of membrane; C:viral nucleocapsid | mRNA processing |
| GABP01045190.1 | heterogeneous nuclear ribonucleo U 1 | F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:membrane; C:integral component of membrane; C:viral nucleocapsid; P:RNA processing; P:response to virus | mRNA processing |
| GABP01045391.1 | CCR4-NOT complex subunit 1 | P:cellular nitrogen compound metabolic process; C:intracellular | mRNA processing |
| GABP01047376.1 | heterogeneous nuclear ribonucleo U 1 | F:nucleic acid binding; C:intracellular ribonucleoprotein complex; P:metabolic process; C:membrane; C:integral component of membrane; C:viral nucleocapsid; F:protein dimerization activity; F:hydrolase activity | mRNA processing |
| GABP01050501.1 | CUGBP Elav-like family member 6 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding | mRNA processing |
| GABP01051435.1 | CUGBP Elav-like family member 4 isoform X15 | F:RNA binding; P:mRNA processing | mRNA processing |
| GABP01051667.1 | heterogeneous nuclear ribonucleo R isoform X1 | F:nucleotide binding; F:nucleic acid binding; F:D-alanine-D-alanine ligase activity; F:ATP binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; F:catalytic activity; P:metabolic process; C:viral nucleocapsid; F:ligase activity | mRNA processing |
| GABP01051673.1 | heterogeneous nuclear ribonucleo U 1 | F:nucleotide binding; F:nucleic acid binding; F:ATP binding; F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; P:metabolic process; C:membrane; C:integral component of membrane; C:viral nucleocapsid; F:helicase activity; F:hydrolase activity | mRNA processing |
| GABP01052101.1 | U2 small nuclear ribonucleo A | C:nucleus; F:RNA binding; C:cytoplasm | mRNA processing |
| GABP01053365.1 | heterogeneous nuclear ribonucleo 1-like | F:molecular_function | mRNA processing |
| GABP01054309.1 | 5 -3 exoribonuclease 3 isoform X1 | F:nuclease activity; F:ion binding | mRNA processing |
| GABP01054427.1 | THO complex subunit 2 | P:mRNA export from nucleus; C:chromosome, telomeric region; P:viral mRNA export from host cell nucleus; P:generation of neurons; C:THO complex; C:THO complex part of transcription export complex; C:transcription export complex; P:poly(A)+ mRNA export from nucleus | mRNA processing |
| GABP01055203.1 | THO complex subunit 2 | P:clathrin coat assembly; F:clathrin binding; C:THO complex; F:phospholipid binding; C:clathrin-coated vesicle; P:intracellular protein transport; F:1-phosphatidylinositol binding; C:intracellular | mRNA processing |
| GABP01055584.1 | splicing factor | F:RNA binding; P:cellular nitrogen compound metabolic process | mRNA processing |

| | | | |
|---|---|---|---|
| GABP01056036.1 | serine arginine-rich splicing factor 1 | F:molecular_function | mRNA processing |
| GABP01056832.1 | heterogeneous nuclear ribonucleo U | C:intracellular | mRNA processing |
| GABP01059035.1 | heterogeneous nuclear ribonucleo U 1 | C:nucleus; F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:membrane; C:integral component of membrane; C:viral nucleocapsid | mRNA processing |
| GABP01060309.1 | putative U5 snRNP-associated 102 kDa protein | C:nucleus; P:ribonucleoprotein complex assembly; P:mRNA processing; F:molecular_function | mRNA processing |
| GABP01062609.1 | heterogeneous nuclear ribonucleo U 1 | F:nucleic acid binding; C:intracellular ribonucleoprotein complex; P:metabolic process; C:viral nucleocapsid; F:protein dimerization activity; F:hydrolase activity | mRNA processing |
| GABP01062610.1 | heterogeneous nuclear ribonucleo U 1 | F:nucleic acid binding; C:intracellular ribonucleoprotein complex; P:metabolic process; C:viral nucleocapsid; F:protein dimerization activity; F:hydrolase activity | mRNA processing |
| GABP01064944.1 | U4 tri-snRNP-associated 2-like | F:ion binding; P:anatomical structure development; F:peptidase activity; P:cellular nitrogen compound metabolic process | mRNA processing |
| GABP01067818.1 | U1 small nuclear ribonucleo 70 kDa | P:catabolic process; C:cytoplasm; P:autophagy; P:anatomical structure development; P:cell differentiation; F:mRNA binding; P:mRNA processing; P:protein folding; P:immune system process; C:nucleoplasm | mRNA processing |
| GABP01069110.1 | pre-mRNA-processing-splicing factor 8 | F:ubiquitin-like protein binding; C:nucleus; F:RNA binding; P:ribonucleoprotein complex assembly; P:mRNA processing | mRNA processing |
| GABP01069581.1 | U4 U6 small nuclear ribonucleo Prp3-like | C:U4/U6 x U5 tri-snRNP complex; P:mRNA splicing, via spliceosome | mRNA processing |
| GABP01071229.1 | probable splicing factor 3A subunit 1 isoform X1 | | mRNA processing |
| GABP01075518.1 | U5 small nuclear ribonucleo 200 kDa | F:nucleotide binding; F:nucleic acid binding; F:DNA binding; F:ATP binding; F:RNA binding; F:RNA helicase activity; C:intracellular ribonucleoprotein complex; P:metabolic process; C:viral nucleocapsid; F:helicase activity; F:hydrolase activity; P:RNA catabolic process | mRNA processing |
| GABP01077166.1 | pre-mRNA-processing factor 19 | P:catabolic process; P:signal transduction; P:ribonucleoprotein complex assembly; P:cellular protein modification process; C:nuclear chromosome; C:protein complex; P:mRNA processing; P:DNA metabolic process; F:molecular_function; P:response to stress; P:cell cycle; C:nucleoplasm; C:cytoplasm | mRNA processing |
| GABP01077863.1 | THO complex subunit 3 | P:mRNA export from nucleus; C:chromosome, telomeric region; P:viral mRNA export from host cell nucleus; P:gene silencing by RNA; C:THO complex; C:THO complex part of transcription export complex; C:transcription export complex; P:production of ta-siRNAs involved in RNA interference | mRNA processing |
| GABP01078419.1 | CD2 antigen cytoplasmic tail-binding 2 | C:U5 snRNP | mRNA processing |
| GABP01079215.1 | U2 snRNP auxilliary large splicing factor subfamily | C:nucleus; F:molecular_function | mRNA processing |
| GABP01083012.1 | 116 kDa U5 small nuclear ribonucleo component | F:RNA binding; F:GTPase activity; C:nucleoplasm; P:translation | mRNA processing |

| | | | |
|---|---|---|---|
| GABP01086334.1 | ELAV 1 isoform X1 | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding; C:cytosol; P:positive regulation of angiogenesis | mRNA processing |
| GABP01086367.1 | THO complex subunit 1 | P:signal transduction; F:catalytic activity; P:nitrogen compound metabolic process; C:integral component of membrane; P:regulation of apoptotic process; F:glutamate-ammonia ligase activity | mRNA processing |
| GABP01090231.1 | polyadenylate binding | F:RNA binding; C:cytoplasm; P:anatomical structure development; P:cell differentiation; P:developmental maturation; P:mRNA processing; P:reproduction | mRNA processing |
| GABP01091749.1 | splicing factor 3 subunit 1 | F:RNA binding; P:cellular nitrogen compound metabolic process | mRNA processing |
| GABP01092259.1 | Splicing factor 3B subunit 3 | C:nucleus; F:RNA binding; P:mRNA processing | mRNA processing |
| GABP01094135.1 | arginine/serine-rich splicing factor4-like | F:DNA binding; F:transcription factor activity, sequence-specific DNA binding; P:transcription, DNA-templated; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated; F:sequence-specific DNA binding | mRNA processing |
| GABP01097040.1 | Pre-mRNA-processing factor 17 | C:catalytic step 2 spliceosome; P:mRNA splicing, via spliceosome | mRNA processing |
| GABP01097640.1 | heterogeneous nuclear ribonucleo 1-like | F:molecular_function | mRNA processing |
| GABP01097943.1 | splicing factor 3B subunit partial | C:nucleus; P:cell death; P:pigmentation; P:ribonucleoprotein complex assembly; P:cell differentiation; P:anatomical structure development; F:mRNA binding; P:mRNA processing | mRNA processing |
| GABP01098067.1 | Cleavage and polyadenylation specificity factor subunit 2 | F:oxidoreductase activity; P:cellular nitrogen compound metabolic process; C:organelle; C:intracellular | mRNA processing |
| GABP01102835.1 | pre-mRNA-splicing factor SPF27 homolog | C:nucleus; P:biological_process | mRNA processing |
| GABP01103730.1 | heterogeneous nuclear ribonucleo 1-like | F:molecular_function | mRNA processing |
| GABP01107533.1 | probable pre-mRNA-splicing factor ATP-dependent RNA helicase | F:ATPase activity; C:nucleus; F:RNA binding; C:mitochondrion; F:ion binding; C:cytosol; F:helicase activity; P:mRNA processing | mRNA processing |
| GABP01108492.1 | polyadenylate binding | F:molecular_function | mRNA processing |
| GABP01109933.1 | heterogeneous nuclear ribonucleo H-like | F:RNA binding; C:cytoplasm; P:signal transduction; P:mRNA processing; C:nucleoplasm | mRNA processing |
| GABP01113141.1 | U4 U6 small nuclear ribonucleo Prp31 | C:intracellular ribonucleoprotein complex; C:U4/U6 x U5 tri-snRNP complex; C:viral nucleocapsid; P:spliceosomal tri-snRNP complex assembly; P:mRNA splicing, via spliceosome | mRNA processing |
| GABP01113320.1 | luc7 3 isoform X3 | C:nucleus | mRNA processing |
| GABP01113784.1 | U5 snRNP-associated 102 kDa | C:nucleus; P:mRNA processing | mRNA processing |

| | | | |
|---|---|---|---|
| GABP01000445.1 | guanylyl cyclase-activating 2-like | P:protein autophosphorylation; C:Golgi apparatus; F:calmodulin-dependent protein kinase activity; P:protein phosphorylation; F:calcium-dependent protein serine/threonine kinase activity; P:response to auxin; P:response to mechanical stimulus; F:nucleotide binding; P:response to cold; F:ATP binding; P:response to heat; P:regulation of flower development; F:calcium ion binding; P:response to metal ion; P:phosphorylation; P:intracellular signal transduction; P:regulation of nitric oxide metabolic process; F:protein kinase activity; P:response to 1-aminocyclopropane-1-carboxylic acid; P:response to hydrogen peroxide; P:response to absence of light; C:plasma membrane; P:response to gibberellin; C:nucleus; P:abscisic acid-activated signaling pathway; F:calmodulin binding; P:response to abscisic acid; F:metal ion binding; C:cytoplasm; P:innate immune response; P:response to calcium ion; P:long-day photoperiodism, flowering; C:membrane; C:integral component of membrane; F:transferase activity; P:peptidyl-serine phosphorylation; F:kinase activity | No chromatin function |
| GABP01004476.1 | RING finger | F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:membrane; C:integral component of membrane; C:viral nucleocapsid | No chromatin function |
| GABP01004737.1 | flagellar associated | C:cell; P:cellular amino acid metabolic process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:molecular_function | No chromatin function |
| GABP01009392.1 | guanine nucleotide-binding -like 3 partial | C:nucleus; P:ribosome biogenesis; F:molecular_function | No chromatin function |
| GABP01012259.1 | TPR repeat-containing | F:ATP binding; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | No chromatin function |
| GABP01013110.1 | PAP OAS1 substrate-binding domain-containing | P:biological_process; F:ion binding | No chromatin function |
| GABP01014833.1 | Cullin-1 | P:catabolic process; F:molecular_function | No chromatin function |
| GABP01016489.1 | hybrid signal transduction histidine | P:signal transduction; F:molecular_function | No chromatin function |
| GABP01016490.1 | hybrid signal transduction histidine | C:cellular_component; F:signal transducer activity; F:kinase activity | No chromatin function |
| GABP01018267.1 | Cysteine Histidine-rich C1 domain family isoform 1 | F:zinc ion binding; F:metal ion binding; F:protein-disulfide reductase activity; P:proteolysis; F:serine-type peptidase activity; P:oxidation-reduction process; P:intracellular signal transduction; C:intracellular | No chromatin function |
| GABP01020349.1 | phosphatase 1 regulatory subunit 27-like | F:transferase activity, transferring acyl groups; F:protein-cysteine S-palmitoyltransferase activity; F:zinc ion binding; F:metal ion binding; F:fatty-acyl-CoA binding; P:metabolic process; C:membrane; C:integral component of membrane; F:transferase activity | No chromatin function |
| GABP01021182.1 | stage v sporulation s | | No chromatin function |

| | | | |
|---|---|---|---|
| GABP01022278.1 | poly-glutamine tract-binding | C:nuclear speck; P:negative regulation of protein catabolic process; P:protein peptidyl-prolyl isomerization; P:negative regulation of protein binding; P:positive regulation of GTPase activity; F:motor activity; P:positive regulation of ubiquitin-protein transferase activity; C:cytosol; F:mitogen-activated protein kinase kinase binding; F:isomerase activity; P:negative regulation of type I interferon production; P:neuron differentiation; P:regulation of pathway-restricted SMAD protein phosphorylation; P:synapse organization; P:metabolic process; P:positive regulation of canonical Wnt signaling pathway; P:regulation of signal transduction by p53 class mediator; C:nucleus; F:protein binding; P:positive regulation of protein dephosphorylation; P:positive regulation of cell growth involved in cardiac muscle cell development; C:midbody; F:GTPase activating protein binding; P:regulation of protein localization to nucleus; P:protein ubiquitination; F:phosphothreonine binding; P:positive regulation of transcription from RNA polymerase II promoter; C:neuron projection; P:positive regulation of neuron apoptotic process; P:negative regulation of neuron apoptotic process; C:nucleoplasm; F:nucleotide binding; F:ATP binding; P:regulation of cytokinesis; F:phosphoserine binding; F:ligase activity; C:myosin complex; P:regulation of mitotic nuclear division; P:signal transduction; F:beta-catenin binding; P:cell cycle; F:ubiquitin-protein transferase activity; F:peptidyl-prolyl cis-trans isomerase activity; C:cytoplasm; P:negative regulation of ERK1 and ERK2 cascade; C:mitochondrion; P:positive regulation of protein phosphorylation; P:negative regulation of cell motility; P:protein stabilization; P:negative regulation of transforming growth factor beta receptor signaling pathway | No chromatin function |
| GABP01023074.1 | ATPase | F:ATP binding; C:membrane; C:integral component of membrane | No chromatin function |
| GABP01023375.1 | NEDD4-binding 2-like 1 | P:biological_process; C:membrane; P:cyclic nucleotide catabolic process; F:2',3'-cyclic-nucleotide 3'-phosphodiesterase activity; C:cellular_component; F:molecular_function | No chromatin function |
| GABP01026376.1 | probable phosphatase 2C 8 isoform X1 | F:molecular_function | No chromatin function |
| GABP01027708.1 | ATPase family associated with various cellular activities (AAA) subfamily | F:molecular_function | No chromatin function |
| GABP01028114.1 | plant UBX domain-containing 8-like | C:membrane; C:integral component of membrane; P:protein processing; P:ubiquitin-dependent protein catabolic process | No chromatin function |
| GABP01028193.1 | izumo sperm-egg fusion 4 isoform X1 | P:GPI anchor metabolic process; F:hydrolase activity, acting on ester bonds; P:metabolic process; F:methyltransferase activity; F:transferase activity; F:hydrolase activity, hydrolyzing O-glycosyl compounds; P:intracellular protein transport; C:intracellular; P:methylation | No chromatin function |
| GABP01028840.1 | adenosine kinase 2 | P:small molecule metabolic process; F:ion binding; C:cytosol; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:kinase activity; C:plasma membrane | No chromatin function |
| GABP01029966.1 | glyco FP21 | P:catabolic process; F:kinase activity | No chromatin function |
| GABP01030514.1 | proteasome-associated ECM29 homolog | F:nucleotide binding; F:protein complex scaffold; F:ATP binding; C:nucleus; C:cytosol; C:proteasome complex; P:proteasome assembly | No chromatin function |

| | | | |
|---|---|---|---|
| GABP01030885.1 | conserved hypothetical protein | C:mitochondrial proton-transporting ATP synthase complex, coupling factor F(o); F:aspartic-type endopeptidase activity; F:hydrogen ion transmembrane transporter activity; C:membrane; P:proteolysis; C:integral component of membrane; F:peptidase activity; P:ATP synthesis coupled proton transport; F:hydrolase activity | No chromatin function |
| GABP01031372.1 | MULTISPECIES: alpha beta hydrolase | | No chromatin function |
| GABP01032123.1 | type I polyketide synthase | F:molecular_function | No chromatin function |
| GABP01032261.1 | TPR repeat-containing | F:ATP binding; F:identical protein binding; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | No chromatin function |
| GABP01032264.1 | Tetratricopeptide TPR_2 repeat | F:ATP binding; F:identical protein binding; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | No chromatin function |
| GABP01033006.1 | tetratricopeptide repeat domain containing | C:cytosol; F:molecular_function | No chromatin function |
| GABP01033008.1 | tetratricopeptide repeat domain containing | C:cytosol | No chromatin function |
| GABP01034306.1 | apoptosis-inducing factor 2-like | C:cell; F:oxidoreductase activity | No chromatin function |
| GABP01034308.1 | FHA domain-containing | F:ATPase activity; F:phosphorus-oxygen lyase activity; P:metabolic process; F:protein kinase activity; F:calmodulin-dependent protein kinase activity; F:protein serine/threonine kinase activity; P:protein phosphorylation; C:intracellular; F:nucleotide binding; F:ATP binding; C:membrane; C:integral component of membrane; P:phosphorylation; F:transferase activity; F:kinase activity; P:intracellular signal transduction; P:cyclic nucleotide biosynthetic process | No chromatin function |
| GABP01034852.1 | methyltransferase | F:molecular_function | No chromatin function |
| GABP01035130.1 | hybrid signal transduction histidine | C:cellular_component; F:signal transducer activity; F:kinase activity | No chromatin function |
| GABP01037506.1 | 26S proteasome non-ATPase regulatory subunit 11 homolog | C:intracellular | No chromatin function |
| GABP01038313.1 | FHA domain-containing | F:ATPase activity; F:phosphorus-oxygen lyase activity; P:metabolic process; F:protein kinase activity; F:protein serine/threonine kinase activity; P:protein phosphorylation; C:intracellular; F:nucleotide binding; P:DNA damage checkpoint; F:ATP binding; P:cell cycle checkpoint; P:replication fork protection; C:membrane; C:integral component of membrane; P:phosphorylation; F:transferase activity; F:kinase activity; P:intracellular signal transduction; P:cyclic nucleotide biosynthetic process | No chromatin function |
| GABP01038341.1 | serine threonine- kinase ULK3 isoform X3 | F:nucleotide binding; F:ATP binding; F:protein kinase activity; P:protein phosphorylation | No chromatin function |
| GABP01038686.1 | cytochrome b5 domain-containing RLF | F:molecular_function | No chromatin function |
| GABP01038721.1 | proteophosphoglycan related | F:molecular_function | No chromatin function |
| GABP01039490.1 | proteasome-associated ECM29 homolog | F:protein complex scaffold; C:membrane; C:integral component of membrane; C:proteasome complex; P:proteasome assembly | No chromatin function |

| | | | |
|---|---|---|---|
| GABP01041886.1 | hypothetical protein GUITHDRAFT_142802 | F:nucleotide binding; F:ATP binding | No chromatin function |
| GABP01042129.1 | ankyrin repeat domain-containing 2-like | F:transferase activity, transferring glycosyl groups; C:motile cilium; P:metabolic process; F:transferase activity; F:NAD+ ADP-ribosyltransferase activity | No chromatin function |
| GABP01047379.1 | laminin subunit alpha | P:cell adhesion | No chromatin function |
| GABP01052477.1 | HEAT repeat-containing | P:lipid transport; F:lipid transporter activity; P:metabolic process; C:membrane; C:integral component of membrane; F:lyase activity; F:hydrolase activity | No chromatin function |
| GABP01058090.1 | enterotoxin | F:molecular_function | No chromatin function |
| GABP01058297.1 | flagellar attachment zone 1-like isoform X3 | | No chromatin function |
| GABP01060270.1 | 26S proteasome non-ATPase regulatory subunit 2 homolog A | C:nucleus; P:catabolic process; C:cytosol; C:protein complex; F:peptidase activity | No chromatin function |
| GABP01060955.1 | Metal homeostasis factor | C:cytoplasm; P:homeostatic process; F:ion binding; P:transport | No chromatin function |
| GABP01061020.1 | ATPase | F:ATP binding; F:calcium ion binding | No chromatin function |
| GABP01062726.1 | AFG1 family ATPase | F:ATP binding | No chromatin function |
| GABP01064398.1 | calmodulin | P:anatomical structure development; F:molecular_function | No chromatin function |
| GABP01065350.1 | peptidase M24 | P:biological_process | No chromatin function |
| GABP01067148.1 | segregation and condensation A | F:nucleotide binding; P:cell division; F:ATP binding; C:cytoplasm; F:catalytic activity; P:metabolic process; P:chromosome segregation; P:cell cycle | No chromatin function |
| GABP01068131.1 | nucleoredoxin | C:cell; C:nucleus; F:thioredoxin-disulfide reductase activity; P:cellular oxidant detoxification; P:cell redox homeostasis; P:negative regulation of Wnt signaling pathway; P:cardiovascular system development; P:negative regulation of protein ubiquitination | No chromatin function |
| GABP01071133.1 | aminoglycoside hydroxyurea antibiotic resistance kinase | P:secondary metabolic process; P:phosphorylation; F:phosphotransferase activity, alcohol group as acceptor; P:protein phosphorylation; F:kinase activity | No chromatin function |
| GABP01073837.1 | 26S proteasome non-ATPase regulatory subunit 3 | P:catabolic process; C:protein complex; C:intracellular | No chromatin function |
| GABP01074434.1 | family | P:transmembrane transport; P:metabolic process; F:lyase activity; F:symporter activity; C:cell outer membrane; P:transport; F:organic anion transmembrane transporter activity; C:plasma membrane; F:aconitate hydratase activity; C:membrane; C:integral component of membrane; F:4 iron, 4 sulfur cluster binding; F:iron-sulfur cluster binding; P:organic anion transport | No chromatin function |
| GABP01075463.1 | GTP-binding YPTC1 | C:cytoplasm; P:signal transduction; F:ion binding; C:organelle; P:transport | No chromatin function |
| GABP01077256.1 | phosphate permease | F:ATP binding; P:transmembrane transport; F:transmembrane transporter activity; C:membrane; C:integral component of membrane; F:transporter activity | No chromatin function |
| GABP01077680.1 | prolactin regulatory element-binding | C:membrane; C:integral component of membrane | No chromatin function |
| GABP01085452.1 | hypothetical protein Pmar_PMAR024059 | F:calcium ion binding; C:membrane; C:integral component of membrane; P:cell adhesion; P:homophilic cell adhesion via plasma membrane adhesion molecules; C:plasma membrane | No chromatin function |
| GABP01091442.1 | AAA family | F:ATPase activity; P:biological_process; F:ion binding | No chromatin function |

| | | | |
|---|---|---|---|
| GABP01091581.1 | PKL CAK Fmp29 kinase subdomain-containing | P:chemotaxis; C:motile cilium; C:bacterial-type flagellum; P:metabolic process; C:membrane; C:integral component of membrane; P:phosphorylation; P:bacterial-type flagellum-dependent cell motility; F:kinase activity; F:motor activity | No chromatin function |
| GABP01098544.1 | Tetratricopeptide TPR_1 repeat-containing | F:ATP binding; F:identical protein binding; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | No chromatin function |
| GABP01106019.1 | TPR repeat-containing | F:ATP binding; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | No chromatin function |
| GABP01106763.1 | Ca2+-binding 1 | F:calcium ion binding; C:membrane; C:integral component of membrane; C:cytosol | No chromatin function |
| GABP01107017.1 | PUB domain-containing | C:cell; P:cell redox homeostasis | No chromatin function |
| GABP01111651.1 | Tetratricopeptide TPR_2 repeat | F:ATP binding; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | No chromatin function |
| GABP01113111.1 | ATP-binding cassette sub-family F member 1 | F:ATPase activity; F:transmembrane transporter activity | No chromatin function |
| GABP01007195.1 | Exportin-1 | P:nucleocytoplasmic transport; F:protein transporter activity; F:transcription factor binding; C:organelle; C:intracellular | Nuclear transport |
| GABP01008608.1 | Nrp1p | F:nucleotide binding; F:nucleic acid binding; P:biological_process; F:zinc ion binding; C:cytoplasmic stress granule; F:RNA binding; C:cytoplasm; F:metal ion binding; C:nucleolus; F:molecular_function | Nuclear transport |
| GABP01012168.1 | transportin-3 isoform X1 | C:intracellular | Nuclear transport |
| GABP01013760.1 | karyopherin beta | C:nucleus; P:nucleocytoplasmic transport; P:protein targeting; F:molecular_function | Nuclear transport |
| GABP01014773.1 | nucleoporin-interacting NIC96 | C:nuclear pore nuclear basket; P:protein transport; F:structural molecule activity; P:nuclear pore complex assembly; P:positive regulation of defense response to virus by host; C:nuclear pore linkers; P:transport; P:xenophagy; C:nuclear pore; P:ribosomal large subunit export from nucleus; C:nucleus; C:nuclear membrane; P:nuclear pore organization; P:protein import into nucleus; C:membrane; F:structural constituent of nuclear pore; P:mitophagy in response to mitochondrial depolarization; C:nuclear periphery; P:mRNA transport; P:poly(A)+ mRNA export from nucleus | Nuclear transport |
| GABP01021402.1 | hypothetical protein PGTG_16976 | C:nuclear envelope; P:nucleocytoplasmic transport; C:nuclear pore outer ring; C:nuclear periphery; F:nucleocytoplasmic transporter activity; P:poly(A)+ mRNA export from nucleus; C:nuclear pore | Nuclear transport |
| GABP01023583.1 | nuclear pore complex Nup85 isoform X1 | | Nuclear transport |
| GABP01024949.1 | exportin-T | C:intracellular | Nuclear transport |

| | | | |
|---|---|---|---|
| GABP01025039.1 | nuclear pore complex Nup133 | P:neural tube development; C:nuclear pore outer ring; P:somite development; P:protein transport; P:paraxial mesoderm development; P:neurogenesis; C:condensed chromosome kinetochore; C:kinetochore; C:chromosome; P:posttranscriptional tethering of RNA polymerase II gene DNA at nuclear periphery; C:chromosome, centromeric region; P:nuclear pore organization; P:mRNA transport; P:regulation of transcription, DNA-templated; C:condensed chromosome outer kinetochore; F:nucleocytoplasmic transporter activity; P:transport; C:nuclear pore; P:mRNA export from nucleus; C:nucleus; C:nuclear membrane; P:RNA export from nucleus; P:protein import into nucleus; C:membrane; C:integral component of membrane; F:structural constituent of nuclear pore | Nuclear transport |
| GABP01025684.1 | Ran-specific GTPase-activating | P:cell division; C:nucleus; P:nucleocytoplasmic transport; P:protein targeting; P:cytoskeleton organization; C:microtubule organizing center; P:catabolic process; C:cytoplasm; P:mitotic nuclear division; F:enzyme regulator activity; F:enzyme binding | Nuclear transport |
| GABP01025685.1 | Ran-specific GTPase-activating | P:cytoskeleton organization; P:catabolic process; F:ion binding; P:mitotic nuclear division; P:biosynthetic process; P:DNA metabolic process; P:cell division; C:nucleus; P:nucleocytoplasmic transport; P:protein targeting; C:microtubule organizing center; C:cytoplasm; F:GTPase activity; F:enzyme regulator activity; F:enzyme binding | Nuclear transport |
| GABP01025686.1 | ran-specific GTPase-activating | P:cytoskeleton organization; P:catabolic process; F:ion binding; P:mitotic nuclear division; P:biosynthetic process; P:DNA metabolic process; P:cell division; C:nucleus; P:nucleocytoplasmic transport; P:protein targeting; C:microtubule organizing center; C:cytoplasm; F:GTPase activity; F:enzyme regulator activity; F:enzyme binding | Nuclear transport |
| GABP01025687.1 | Ran-specific GTPase-activating | P:nucleocytoplasmic transport; C:organelle; C:intracellular; F:molecular_function; P:cell cycle | Nuclear transport |
| GABP01026269.1 | transportin-1 | P:nucleocytoplasmic transport; P:protein targeting; P:anatomical structure development; C:intracellular | Nuclear transport |
| GABP01029845.1 | hrp1p | C:intracellular; F:molecular_function | Nuclear transport |
| GABP01040680.1 | nuclear pore complex NUP88 | P:protein export from nucleus; C:plastid; F:transporter activity; C:nuclear pore; P:ribosomal small subunit export from nucleus; C:nuclear envelope; P:mRNA export from nucleus; P:ribosomal large subunit export from nucleus; C:nucleus; F:protein binding; P:systemic acquired resistance; P:protein import into nucleus; P:innate immune response | Nuclear transport |
| GABP01046815.1 | Rab GDP dissociation inhibitor | C:cell; P:vesicle-mediated transport; F:enzyme regulator activity | Nuclear transport |
| GABP01051913.1 | nuclear pore complex | C:myosin complex; F:zinc ion binding; F:protein disulfide oxidoreductase activity; P:protein complex assembly; F:electron carrier activity; P:metabolic process; P:cell redox homeostasis; P:protein peptidyl-prolyl isomerization; C:ribosome; P:protein folding; C:intracellular; P:translation; F:motor activity; F:nucleotide binding; F:ATP binding; F:peptidyl-prolyl cis-trans isomerase activity; F:structural constituent of ribosome; C:intracellular ribonucleoprotein complex; C:mitochondrion; C:membrane; C:integral component of membrane; P:oxidation-reduction process; F:isomerase activity | Nuclear transport |

| | | | |
|---|---|---|---|
| GABP01054070.1 | nuclear pore complex Nup155 | C:nuclear envelope; P:transport; F:molecular_function | Nuclear transport |
| GABP01054201.1 | nuclear pore complex | P:protein complex assembly; C:Golgi membrane; C:Golgi apparatus; P:protein peptidyl-prolyl isomerization; C:endoplasmic reticulum membrane; F:nucleotide binding; F:translation initiation factor activity; F:ATP binding; C:intracellular ribonucleoprotein complex; F:isomerase activity; F:zinc ion binding; F:aspartic-type endopeptidase activity; C:endoplasmic reticulum; F:peptidase activity; C:ribosome; P:translational initiation; P:protein folding; C:intracellular; P:translation; F:peptidyl-prolyl cis-trans isomerase activity; P:Notch signaling pathway; F:structural constituent of ribosome; C:membrane; C:mitochondrion; P:proteolysis; C:integral component of membrane; P:protein processing; F:hydrolase activity | Nuclear transport |
| GABP01056574.1 | nucleo TPR | P:mitotic spindle assembly checkpoint; P:nucleocytoplasmic transport; P:protein import into nucleus; C:viral nucleocapsid; F:nucleocytoplasmic transporter activity; C:intracellular; C:nuclear pore | Nuclear transport |
| GABP01057107.1 | nuclear pore complex GP210 | C:nucleus | Nuclear transport |
| GABP01070915.1 | nuclear movement nudC | F:peptidyl-prolyl cis-trans isomerase activity; P:protein peptidyl-prolyl isomerization; P:protein folding; F:isomerase activity | Nuclear transport |
| GABP01071110.1 | importin subunit beta-1-like | C:nucleus; P:nucleocytoplasmic transport; P:protein targeting | Nuclear transport |
| GABP01099449.1 | NPL4 family | F:molecular_function | Nuclear transport |
| GABP01106666.1 | nuclear pore complex Nup98-Nup96 | C:nuclear envelope | Nuclear transport |
| GABP01107332.1 | GTP-binding nuclear ran | P:signal transduction; F:ion binding; P:anatomical structure development; P:ribosome biogenesis; P:mitotic nuclear division; C:chromosome; C:protein complex; C:nucleolus; C:extracellular region; C:nucleoplasm; P:cell division; C:nuclear envelope; P:nucleocytoplasmic transport; P:protein targeting; C:endosome; F:RNA binding; C:microtubule organizing center; F:GTPase activity; C:cytoplasmic, membrane-bounded vesicle | Nuclear transport |
| GABP01107334.1 | GTP-binding nuclear ran | P:signal transduction; F:ion binding; P:ribosome biogenesis; P:mitotic nuclear division; C:ribosome; C:plastid; C:nuclear envelope; P:cell division; P:nucleocytoplasmic transport; P:protein targeting; F:GTPase activity; C:cytosol; C:cytoplasmic, membrane-bounded vesicle | Nuclear transport |

| | | | |
|---|---|---|---|
| GABP01005642.1 | DNA-binding SMUBP-2 | F:RNA-dependent ATPase activity; F:ATP-dependent helicase activity; F:helicase activity; C:neuronal cell body; F:protein self-association; C:cell projection; C:cellular_component; C:growth cone; F:nucleotide binding; P:negative regulation of transcription from RNA polymerase II promoter; F:ATP binding; F:transcriptional repressor activity, RNA polymerase II core promoter proximal region sequence-specific binding; F:RNA binding; C:intracellular ribonucleoprotein complex; P:DNA duplex unwinding; C:axon; P:biological_process; F:zinc ion binding; F:DNA-dependent ATPase activity; P:transcription, DNA-templated; P:metabolic process; P:protein homooligomerization; P:spinal cord motor neuron differentiation; P:regulation of transcription, DNA-templated; F:transcription factor binding; F:tRNA binding; P:translation; P:regulation of transcription from RNA polymerase II promoter; F:nucleic acid binding; F:DNA binding; C:nucleus; C:perinuclear region of cytoplasm; F:ATP-dependent 5'-3' DNA helicase activity; F:metal ion binding; C:cytoplasm; F:ribosome binding; C:membrane; F:ATP-dependent 5'-3' RNA helicase activity; C:SMN complex; F:hydrolase activity | Nucleic acid binding |
| GABP01012688.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; F:metal ion binding; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01015490.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01015908.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01016014.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01016016.1 | cold-shock | F:nucleic acid binding; F:DNA binding; P:nucleic acid phosphodiester bond hydrolysis; F:zinc ion binding; C:cytoplasm; F:exonuclease activity; P:transcription, DNA-templated; F:metal ion binding; F:nuclease activity; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01020886.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated; P:response to stress | Nucleic acid binding |
| GABP01021125.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01021233.1 | cold-shock | P:cellular nitrogen compound metabolic process; C:plastid; F:molecular_function | Nucleic acid binding |
| GABP01021403.1 | cold-shock | F:nucleotide binding; F:nucleic acid binding; F:DNA binding; F:zinc ion binding; C:cytoplasm; F:metal ion binding; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01028462.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01031792.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01035390.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |

| | | | |
|---|---|---|---|
| GABP01037338.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01037606.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01043998.1 | cold-shock | F:nucleic acid binding; F:DNA binding; F:zinc ion binding; C:cytoplasm; P:transcription, DNA-templated; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01050277.1 | cold shock domain-containing E1 isoform X1 | C:CRD-mediated mRNA stability complex; F:nucleic acid binding; F:DNA binding; F:RNA binding; C:cytoplasm; P:regulation of transcription, DNA-templated; C:Golgi apparatus; P:nuclear-transcribed mRNA catabolic process, no-go decay; C:mitochondrial inner membrane; C:plasma membrane; F:poly(A) RNA binding | Nucleic acid binding |
| GABP01050611.1 | MULTISPECIES: cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01054204.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01054624.1 | cold-shock | F:nucleic acid binding; F:DNA binding; P:response to cold; F:zinc ion binding; C:cytoplasm; P:transcription, DNA-templated; F:metal ion binding; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01057886.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01064490.1 | cold-shock | F:molecular_function | Nucleic acid binding |
| GABP01075776.1 | transcriptional regulator | F:ATPase activity; C:cell division site; F:phosphorus-oxygen lyase activity; P:protein localization; P:peptidyl-threonine trans-autophosphorylation; P:protein phosphorylation; P:mRNA splicing, via spliceosome; F:nucleotide binding; F:ATP binding; P:phosphorelay signal transduction system; C:cytosol; P:intra-S DNA damage checkpoint; P:phosphorylation; P:intracellular signal transduction; P:cyclic nucleotide biosynthetic process; P:meiotic DNA replication checkpoint; P:metabolic process; P:regulation of transcription, DNA-templated; F:protein kinase activity; F:protein serine/threonine kinase activity; P:nucleobase-containing compound metabolic process; C:intracellular; P:negative regulation of phosphorylation; P:mRNA export from nucleus; P:DNA damage checkpoint; F:DNA binding; C:nucleus; P:cell cycle checkpoint; P:replication fork protection; C:RES complex; C:membrane; C:integral component of membrane; F:transferase activity; F:kinase activity; P:maintenance of RNA location | Nucleic acid binding |
| GABP01078403.1 | cold-shock | F:molecular_function | Nucleic acid binding |
| GABP01087494.1 | cold-shock | F:molecular_function | Nucleic acid binding |
| GABP01090614.1 | cold-shock | F:zinc ion binding; P:transcription, DNA-templated; P:RNA modification; P:regulation of transcription, DNA-templated; F:pseudouridine synthase activity; F:nucleic acid binding; F:DNA binding; P:pseudouridine synthesis; F:RNA binding; C:cytoplasm; F:metal ion binding; C:membrane; C:integral component of membrane | Nucleic acid binding |

| | | | |
|---|---|---|---|
| GABP01090616.1 | cold-shock | F:nucleic acid binding; F:DNA binding; F:zinc ion binding; C:cytoplasm; F:metal ion binding; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01091847.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01092866.1 | NAC domain-containing | C:nucleus; P:transport; F:molecular_function | Nucleic acid binding |
| GABP01092993.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01095262.1 | kh domain-containing | F:molecular_function | Nucleic acid binding |
| GABP01099257.1 | cold-shock | F:molecular_function | Nucleic acid binding |
| GABP01101338.1 | cold-shock | F:molecular_function | Nucleic acid binding |
| GABP01108319.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01108894.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:regulation of transcription, DNA-templated | Nucleic acid binding |
| GABP01109196.1 | cold-shock | F:molecular_function | Nucleic acid binding |
| GABP01113515.1 | cold-shock | F:DNA binding; C:cytoplasm; P:cellular nitrogen compound metabolic process; P:biosynthetic process | Nucleic acid binding |
| GABP01093238.1 | Phosphoserine phosphatase | F:phosphatase activity | Protein modification |
| GABP01002312.1 | peptidyl-prolyl isomerase FKBP12 | F:ion binding; P:cellular protein modification process; C:cytosol; P:ribosome biogenesis; P:cellular nitrogen compound metabolic process; P:protein folding; F:isomerase activity | Protein modification |
| GABP01004963.1 | ubiquitin carboxyl-terminal hydrolase isozyme L5 | P:catabolic process; P:anatomical structure development; F:peptidase activity; C:intracellular | Protein modification |
| GABP01007176.1 | phosphoserine phosphatase | F:phosphatase activity; F:kinase activity | Protein modification |
| GABP01007184.1 | phosphatase 2C domain-containing | P:biological_process; F:molecular_function | Protein modification |
| GABP01007390.1 | E3 SUMO- ligase | F:zinc ion binding; F:metal ion binding; P:protein sumoylation; P:metabolic process; F:ligase activity; F:SUMO transferase activity | Protein modification |
| GABP01008727.1 | serine threonine kinase | P:biological_process; F:ion binding | Protein modification |
| GABP01015112.1 | ubiquitin carboxyl-terminal hydrolase 5 isoform X2 | F:ubiquitin-like protein binding; C:cytoplasm; P:catabolic process; P:cellular protein modification process; F:peptidase activity | Protein modification |
| GABP01015422.1 | ubiquitin-like modifier-activating enzyme 1 | C:nucleus; P:catabolic process; P:cellular protein modification process; C:cytosol; F:ligase activity | Protein modification |
| GABP01015652.1 | Peptidyl-prolyl cis-trans isomerase | P:cellular protein modification process; P:protein folding; F:isomerase activity | Protein modification |
| GABP01017846.1 | Ubiquitin carboxyl-terminal hydrolase | P:protein deubiquitination; P:proteolysis; F:peptidase activity; P:ubiquitin-dependent protein catabolic process; F:thiol-dependent ubiquitinyl hydrolase activity; F:hydrolase activity | Protein modification |
| GABP01018851.1 | SUMO-activating enzyme subunit 2 | P:cellular protein modification process; C:protein complex; F:ligase activity; C:nucleoplasm | Protein modification |
| GABP01021217.1 | tyrosine kinase domain | F:kinase activity | Protein modification |

| | | | |
|---|---|---|---|
| GABP01021367.1 | peptidyl-prolyl cis-trans isomerase | P:cellular protein modification process; P:protein folding; F:isomerase activity | Protein modification |
| GABP01022641.1 | Ess1p | P:cellular protein modification process; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:molecular_function | Protein modification |
| GABP01023104.1 | small ubiquitin-related modifier 2 | C:nucleus; P:catabolic process; P:cellular protein modification process; F:molecular_function | Protein modification |
| GABP01024670.1 | cyclin-dependent kinase-like Serine/Threonine kinase family protein | F:ion binding; P:cellular protein modification process; F:signal transducer activity; F:kinase activity; C:intracellular | Protein modification |
| GABP01026822.1 | ubiquitin partial | C:nucleus; C:cytoplasm; P:catabolic process; F:ion binding; P:cellular protein modification process | Protein modification |
| GABP01027147.1 | peptidyl-prolyl cis-trans isomerase FKBP62-like | P:biological_process; C:cytoplasm; F:molecular_function | Protein modification |
| GABP01027943.1 | E3 SUMO- ligase | F:zinc ion binding; F:metal ion binding; P:protein sumoylation; P:metabolic process; F:ligase activity; F:SUMO transferase activity | Protein modification |
| GABP01032128.1 | E3 ubiquitin- ligase HERC2 | F:DNA binding; F:ATP binding; F:ubiquitin-protein transferase activity; F:zinc ion binding; F:metal ion binding; F:calcium ion binding; P:metabolic process; C:membrane; C:integral component of membrane; F:ligase activity; P:protein ubiquitination; F:hydrolase activity | Protein modification |
| GABP01032222.1 | ubiquitin carboxyl-terminal hydrolase | P:protein deubiquitination; C:cell division site; C:Golgi apparatus; P:endocytosis; P:ubiquitin-dependent protein catabolic process; F:thiol-dependent ubiquitin-specific protease activity; C:peroxisome; C:nucleus; C:cell tip; C:cytoplasm; P:ubiquitin-dependent protein catabolic process via the multivesicular body sorting pathway; F:metal ion binding; F:calcium ion binding; C:cytosol; P:cellular protein localization; F:thiol-dependent ubiquitinyl hydrolase activity; F:hydrolase activity | Protein modification |
| GABP01033562.1 | cell division cycle 48 | F:ATPase activity; P:cell division; P:membrane organization; P:catabolic process; P:autophagy; F:ion binding; C:endoplasmic reticulum; C:protein complex; P:cell cycle; P:response to stress | Protein modification |
| GABP01034393.1 | CAMK CDPK kinase | P:cellular protein modification process; F:kinase activity; C:plasma membrane | Protein modification |
| GABP01035085.1 | E3 ubiquitin- ligase TRIM23 isoform X1 | P:regulation of Rho protein signal transduction; P:protein transport; C:Golgi membrane; C:Golgi apparatus; F:enzyme activator activity; P:protein phosphorylation; C:cellular_component; P:positive regulation of GTPase activity; F:NAD+ ADP-ribosyltransferase activity; F:GTP binding; F:nucleotide binding; F:ATP binding; C:extracellular exosome; F:RNA binding; F:GTPase activity; F:calcium ion binding; P:viral process; F:GDP binding; P:phosphorylation; F:ligase activity; P:biological_process; F:zinc ion binding; F:identical protein binding; P:small GTPase mediated signal transduction; P:metabolic process; F:protein kinase activity; F:Rho guanyl-nucleotide exchange factor activity; F:protein serine/threonine kinase activity; C:intracellular; C:lysosomal membrane; C:lysosome; F:nucleic acid binding; C:nucleus; F:ubiquitin-protein transferase activity; F:protein binding; P:positive regulation of catalytic activity; F:metal ion binding; C:cytoplasm; C:membrane; C:endomembrane system; F:transferase activity; F:kinase activity; P:protein ubiquitination | Protein modification |

| | | | |
|---|---|---|---|
| GABP01041205.1 | peptidyl-prolyl cis trans isomerase | C:nucleus; F:peptidyl-prolyl cis-trans isomerase activity; P:histone lysine methylation; P:metabolic process; F:histone-lysine N-methyltransferase activity; P:regulation of transcription, DNA-templated; C:chromosome; P:protein peptidyl-prolyl isomerization; F:isomerase activity | Protein modification |
| GABP01043318.1 | NEDD4-like E3 ubiquitin- ligase WWP2 isoform X1 | F:ubiquitin protein ligase activity; P:palate development; P:protein peptidyl-prolyl isomerization; P:positive regulation of transcription from RNA polymerase II promoter; F:UDP-N-acetylmuramoylalanyl-D-glutamyl-2,6-diaminopimelate-D-alanyl-D-alanine ligase activity; P:negative regulation of protein transport; P:negative regulation of transcription from RNA polymerase II promoter; C:extracellular exosome; P:negative regulation of transcription, DNA-templated; P:regulation of ion transmembrane transport; F:tubulin-glutamic acid ligase activity; F:transcriptional activator activity, RNA polymerase II transcription factor binding; F:ligase activity; P:protein autoubiquitination; F:isomerase activity; F:protein-glutamic acid ligase activity; F:protein-glycine ligase activity, elongating; F:tubulin-glycine ligase activity; P:metabolic process; P:negative regulation of sequence-specific DNA binding transcription factor activity; F:coenzyme F420-2 alpha-glutamyl ligase activity; F:transcription factor binding; P:regulation of potassium ion transmembrane transporter activity; F:coenzyme F420-0 gamma-glutamyl ligase activity; P:protein ubiquitination involved in ubiquitin-dependent protein catabolic process; P:regulation of intrinsic apoptotic signaling pathway in response to DNA damage; P:negative regulation of gene expression; F:ubiquitin-protein transferase activity; C:nucleus; P:proteasome-mediated ubiquitin-dependent protein catabolic process; F:peptidyl-prolyl cis-trans isomerase activity; P:regulation of membrane potential; F:protein binding; F:RNA polymerase II transcription factor binding; C:cytoplasm; C:membrane; P:negative regulation of transporter activity; P:protein K63-linked ubiquitination; F:ribosomal S6-glutamic acid ligase activity; F:protein-glycine ligase activity; F:protein-glycine ligase activity, initiating; P:protein ubiquitination | Protein modification |
| GABP01044258.1 | probable phosphatase 2C | P:biological_process; F:molecular_function | Protein modification |
| GABP01048640.1 | ubiquitin conjugation factor E4 | F:ubiquitin-protein transferase activity; C:ubiquitin ligase complex; F:ubiquitin-ubiquitin ligase activity; C:membrane; C:integral component of membrane; P:protein ubiquitination; P:ubiquitin-dependent protein catabolic process | Protein modification |
| GABP01050844.1 | Calcium-dependent | F:ion binding; P:cellular protein modification process; F:kinase activity; C:intracellular | Protein modification |
| GABP01050990.1 | UBX domain-containing 6 | F:protein disulfide oxidoreductase activity; P:metabolic process; P:glycerol ether metabolic process; P:cell redox homeostasis; F:protein kinase activity; P:protein phosphorylation; F:nucleotide binding; P:glycoprotein catabolic process; F:nucleic acid binding; F:DNA binding; F:mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase activity; F:ATP binding; C:cell; C:cytoplasm; F:metal ion binding; C:membrane; C:integral component of membrane; P:oxidation-reduction process; F:hydrolase activity | Protein modification |
| GABP01061975.1 | probable phosphatase 2C 59 | P:biological_process; F:molecular_function | Protein modification |
| GABP01063870.1 | ubiquitin- ligase | P:biological_process | Protein modification |

| | | | |
|---|---|---|---|
| GABP01067096.1 | Zn-finger in Ran binding protein | F:nucleotidyltransferase activity; F:DNA-directed DNA polymerase activity; F:zinc ion binding; P:queuosine biosynthetic process; P:metabolic process; P:DNA biosynthetic process; F:transferase activity; P:tRNA modification; F:queuine tRNA-ribosyltransferase activity | Protein modification |
| GABP01067423.1 | calcium-dependent kinase 12-like | P:protein autophosphorylation; F:calmodulin-dependent protein kinase activity; C:extracellular region; P:protein phosphorylation; C:cellular_component; F:molecular_function; F:calcium-dependent protein serine/threonine kinase activity; F:nucleotide binding; F:ATP binding; F:calcium ion binding; P:phosphorylation; P:intracellular signal transduction; P:biological_process; F:transferase activity, transferring glycosyl groups; F:nucleotidyltransferase activity; P:protein ADP-ribosylation; F:protein kinase activity; P:pathogenesis; C:plasma membrane; C:nucleus; P:abscisic acid-activated signaling pathway; F:calmodulin binding; C:cytoplasm; F:NAD(P)+-protein-arginine ADP-ribosyltransferase activity; C:membrane; C:integral component of membrane; F:transferase activity; P:peptidyl-serine phosphorylation; F:kinase activity | Protein modification |
| GABP01068790.1 | peptidylprolyl isomerase | F:translation initiation factor activity; F:peptidyl-prolyl cis-trans isomerase activity; F:zinc ion binding; F:RNA binding; P:protein complex assembly; C:mitochondrion; C:membrane; C:integral component of membrane; P:protein peptidyl-prolyl isomerization; P:translational initiation; P:protein folding; F:isomerase activity | Protein modification |
| GABP01070177.1 | histidine kinase | P:signal transduction; P:signal transduction by protein phosphorylation; F:signal transducer activity; C:intracellular; F:phosphorelay sensor kinase activity; F:nucleotide binding; F:ATP binding; P:phosphorelay signal transduction system; C:membrane; C:integral component of membrane; F:transferase activity, transferring phosphorus-containing groups; P:phosphorylation; F:transferase activity; F:kinase activity | Protein modification |
| GABP01070634.1 | ankyrin repeat | F:DNA binding; F:ubiquitin-protein transferase activity; F:catalytic activity; C:membrane; C:integral component of membrane; P:nucleoside metabolic process; P:protein ubiquitination | Protein modification |
| GABP01072511.1 | dual specificity protein phosphatase | F:phosphatase activity | Protein modification |
| GABP01075103.1 | peptidyl-prolyl cis-trans isomerase dodo isoform X2 | C:nucleus; P:histone lysine methylation; F:histone-lysine N-methyltransferase activity; P:regulation of transcription, DNA-templated; C:chromosome | Protein modification |
| GABP01079708.1 | histone deacetylase | P:metabolic process; F:acetylspermidine deacetylase activity; F:hydrolase activity | Protein modification |
| GABP01080068.1 | peptidyl-prolyl cis-trans isomerase | C:cytoplasm; P:cellular protein modification process; P:protein folding; F:isomerase activity | Protein modification |
| GABP01081896.1 | S-phase kinase-associated protein 1A, putative | P:signal transduction; P:catabolic process; P:cellular protein modification process; C:protein complex; C:extracellular region; P:cell cycle; P:response to stress; P:immune system process; C:nucleoplasm; P:chromosome organization; C:cytosol; F:kinase activity; F:ligase activity | Protein modification |
| GABP01083637.1 | ubiquitin carboxyl-terminal hydrolase 6 | F:peptidase activity | Protein modification |
| GABP01085577.1 | peptidyl-prolyl cis-trans isomerase-like 3 | P:cellular protein modification process; P:protein folding; F:isomerase activity | Protein modification |

| | | | |
|---|---|---|---|
| GABP01086680.1 | methyltransferase | F:methyltransferase activity | Protein modification |
| GABP01087788.1 | E3 ubiquitin- ligase rififylin | F:zinc ion binding; F:metal ion binding; P:metabolic process; F:ligase activity | Protein modification |
| GABP01089469.1 | peptidyl-prolyl cis-trans isomerase NIMA-interacting 1 | P:histone H3-K4 trimethylation; F:identical protein binding; P:negative regulation of histone deacetylation; P:metabolic process; P:regulation of transcription involved in G2/M transition of mitotic cell cycle; F:calcium-dependent cysteine-type endopeptidase activity; P:protein peptidyl-prolyl isomerization; P:termination of RNA polymerase II transcription; P:positive regulation of transcription from RNA polymerase II promoter; C:intracellular; F:RNA polymerase II core binding; P:positive regulation of RNA polymerase II transcriptional preinitiation complex assembly; P:negative regulation of transcription from RNA polymerase II promoter; P:positive regulation of chromatin silencing at rDNA; P:regulation of phosphorylation of RNA polymerase II C-terminal domain; F:peptidyl-prolyl cis-trans isomerase activity; C:cytoplasm; P:regulation of transcription involved in G1/S transition of mitotic cell cycle; P:proteolysis; C:membrane; C:integral component of membrane; P:positive regulation of protein dephosphorylation; P:negative regulation of protein ubiquitination involved in ubiquitin-dependent protein catabolic process; F:isomerase activity | Protein modification |
| GABP01091145.1 | ubiquitin-like modifier-activating enzyme 1 | C:cytoplasm; P:cellular protein modification process; C:organelle; F:ligase activity | Protein modification |
| GABP01091696.1 | probable phosphatase 2C 11 | F:molecular_function | Protein modification |
| GABP01092348.1 | Ubiquitin carboxyl-terminal hydrolase isozyme L3 | F:ubiquitin-like protein binding; C:nucleus; C:cytoplasm; P:catabolic process; P:cell differentiation; P:anatomical structure development; F:peptidase activity; C:extracellular region | Protein modification |
| GABP01092376.1 | cell division cycle 48 | F:ATPase activity; P:cell division; F:ion binding | Protein modification |
| GABP01092377.1 | cell division CDC48CY | F:ATPase activity; P:cell morphogenesis; F:ion binding; C:Golgi apparatus; C:nucleolus; C:ribosome; C:extracellular region; P:transport; C:plasma membrane; P:cell cycle; P:reproduction; P:growth; C:nuclear envelope; P:cell division; C:cytoskeleton; C:cell wall; P:cell differentiation; C:cytosol | Protein modification |
| GABP01092380.1 | cell division cycle 48 | F:ATPase activity; P:cell division; F:ion binding | Protein modification |
| GABP01095525.1 | GNAT family | F:transferase activity, transferring acyl groups; F:acetyltransferase activity; F:calcium ion binding; P:metabolic process; P:N-terminal protein amino acid acetylation; F:transferase activity; F:glucosamine 6-phosphate N-acetyltransferase activity; F:N-acetyltransferase activity | Protein modification |
| GABP01097390.1 | phosphatase | P:biological_process; F:molecular_function | Protein modification |
| GABP01102119.1 | peptidyl-prolyl cis-trans isomerase ESS1 | P:cellular protein modification process; C:cellular_component; F:isomerase activity | Protein modification |
| GABP01102156.1 | serine threonine | P:peptidyl-tyrosine phosphorylation; P:cell communication; F:protein kinase activity; P:protein phosphorylation; C:intracellular; F:nucleotide binding; P:MAPK cascade; F:ATP binding; P:activation of MAPKK activity; F:non-membrane spanning protein tyrosine kinase activity; C:membrane; C:integral component of membrane; F:MAP kinase kinase kinase activity; P:phosphorylation; F:transferase activity; F:carbohydrate binding; F:kinase activity | Protein modification |

| | | | |
|---|---|---|---|
| GABP01102157.1 | serine threonine | P:peptidyl-tyrosine phosphorylation; P:cell communication; F:protein kinase activity; P:protein phosphorylation; C:intracellular; F:nucleotide binding; P:MAPK cascade; F:ATP binding; P:activation of MAPKK activity; F:non-membrane spanning protein tyrosine kinase activity; C:membrane; C:integral component of membrane; F:MAP kinase kinase kinase activity; P:phosphorylation; F:transferase activity; F:carbohydrate binding; F:kinase activity | Protein modification |
| GABP01103223.1 | NEDD8-activating enzyme E1 regulatory subunit-like | P:biological_process; C:intracellular | Protein modification |
| GABP01108809.1 | peptidyl-prolyl cis-trans isomerase FKBP62-like | P:biological_process; C:cytoplasm; F:molecular_function | Protein modification |
| GABP01114183.1 | E3 ubiquitin- ligase DTX1-like isoform X2 | F:zinc ion binding; C:nuclear membrane; P:Notch signaling pathway; F:metal ion binding; C:nucleoplasm | Protein modification |
| GABP01008937.1 | signal recognition particle subunit SRP72-like | C:intracellular | Ribosome related |
| GABP01012378.1 | 40S ribosomal | F:structural constituent of ribosome; C:cytosol; C:ribosome; P:translation | Ribosome related |
| GABP01028703.1 | 60S ribosomal L11 | P:cytoskeleton organization; P:ribonucleoprotein complex assembly; C:nuclear chromosome; P:chromosome segregation; P:ribosome biogenesis; P:mitotic nuclear division; C:nucleolus; C:ribosome; P:translation; F:structural constituent of ribosome; P:chromosome organization; C:cytosol; P:neurological system process | Ribosome related |
| GABP01029097.1 | 40S ribosomal X isoform | P:cell proliferation; P:embryo development; P:membrane organization; P:ribosome biogenesis; C:ribosome; C:extracellular region; P:symbiosis, encompassing mutualism through parasitism; C:nucleoplasm; P:translation; P:protein targeting; F:structural constituent of ribosome; P:nucleobase-containing compound catabolic process; C:cytosol; F:rRNA binding | Ribosome related |
| GABP01030162.1 | 30S ribosomal chloroplastic | F:nucleotide binding; C:chloroplast stroma; F:nucleic acid binding; C:chloroplast; F:flavin adenine dinucleotide binding; F:electron carrier activity; C:chloroplast envelope; P:mRNA 3'-end processing; C:ribosome; C:chloroplast thylakoid membrane; C:thylakoid | Ribosome related |
| GABP01034150.1 | 60S ribosomal | F:structural constituent of ribosome; C:cytosol; C:ribosome; F:rRNA binding; P:translation | Ribosome related |
| GABP01035193.1 | partial | C:nucleus; F:structural constituent of ribosome; C:ribosome; P:translation | Ribosome related |
| GABP01036833.1 | H ACA ribonucleo complex subunit 1 | F:RNA binding; F:transmembrane transporter activity; P:chromosome organization; P:homeostatic process; C:nuclear chromosome; P:ribosome biogenesis; P:biosynthetic process; C:nucleolus; P:DNA metabolic process; C:nucleoplasm | Ribosome related |
| GABP01043097.1 | H ACA ribonucleo complex subunit 2 | F:RNA binding; C:cytoplasm; P:ribosome biogenesis; P:cellular nitrogen compound metabolic process; C:nucleolus | Ribosome related |
| GABP01044597.1 | RNA recognition motif family protein | F:molecular_function | Ribosome related |
| GABP01047378.1 | RING finger | C:nucleus; F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:membrane; C:integral component of membrane; C:viral nucleocapsid | Ribosome related |

| | | | |
|---|---|---|---|
| GABP01049641.1 | midasin type A von Willebrand factor domain-containing | P:biological_process; C:organelle; C:intracellular; F:molecular_function | Ribosome related |
| GABP01059436.1 | RING finger | F:zinc ion binding; C:intracellular ribonucleoprotein complex; F:metal ion binding; C:membrane; C:integral component of membrane; C:viral nucleocapsid | Ribosome related |
| GABP01059450.1 | Ubiquitin-60S ribosomal L40 | P:embryo development; C:cytoskeleton; F:structural constituent of ribosome; P:catabolic process; P:cellular protein modification process; C:cytosol; C:protein complex; C:nucleolus; C:ribosome; P:reproduction; P:translation | Ribosome related |
| GABP01060770.1 | 40S ribosomal S18 | F:RNA binding; F:structural constituent of ribosome; C:cytosol; P:ribosome biogenesis; C:ribosome; P:translation | Ribosome related |
| GABP01069070.1 | RNA-binding protein | F:nucleotide binding; F:nucleic acid binding; C:chloroplast; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid; C:cytosol; C:plastid | Ribosome related |
| GABP01069754.1 | ribosomal RPL4 | F:structural constituent of ribosome; C:cell wall; C:cytosol; C:nucleolus; C:ribosome; C:vacuole; C:plasma membrane; C:plastid; P:translation | Ribosome related |
| GABP01077003.1 | rRNA 2 -O-methyltransferase fibrillarin | F:RNA binding; P:chromosome organization; P:cellular protein modification process; P:tRNA metabolic process; P:ribosome biogenesis; F:methyltransferase activity; C:nucleolus; C:nucleoplasm | Ribosome related |
| GABP01092637.1 | Pseudouridine synthase | P:pseudouridine synthesis; F:RNA binding; P:RNA modification; F:lyase activity; F:pseudouridine synthase activity; F:isomerase activity | Ribosome related |
| GABP01108197.1 | major vault | C:nucleus; C:intracellular ribonucleoprotein complex; C:cytoplasm; C:membrane; P:proteolysis; C:integral component of membrane; F:peptidase activity | Ribosome related |
| GABP01113372.1 | nucleolar 56 | F:RNA binding; C:cytoplasm; P:anatomical structure development; P:ribosome biogenesis; P:cellular nitrogen compound metabolic process; C:nucleolus; C:nucleoplasm; F:enzyme binding | Ribosome related |
| GABP01020885.1 | ELAV 1 isoform X2 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; P:positive regulation of angiogenesis; F:poly(A) RNA binding | RNA binding |
| GABP01034386.1 | KH domain-containing | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01003171.1 | 28 kDa chloroplastic | F:nucleotide binding; F:nucleic acid binding; F:DNA binding; C:nucleus; C:intracellular ribonucleoprotein complex; P:cysteine biosynthetic process; C:viral nucleocapsid; P:RNA processing; C:nucleosome; P:nucleosome assembly | RNA binding |
| GABP01003363.1 | RNA recognition motif-containing | F:molecular_function | RNA binding |
| GABP01004758.1 | RNA-binding Nova-1 isoform X2 | C:nucleus; F:RNA binding; C:cytoplasm | RNA binding |
| GABP01006434.1 | RNA-binding Nova-2 | P:RNA splicing; P:chemical synaptic transmission; P:RNA processing; C:nucleolus; P:mRNA splicing, via spliceosome; F:poly(A) RNA binding; P:spinal cord development; F:nucleic acid binding; C:nucleus; F:protein binding; F:RNA binding; P:locomotory behavior; P:regulation of mRNA processing; C:intracellular membrane-bounded organelle; P:regulation of RNA metabolic process; F:mRNA binding | RNA binding |
| GABP01006897.1 | RNA binding | F:nucleotide binding; F:nucleic acid binding | RNA binding |
| GABP01007489.1 | RNA-binding | F:nucleotide binding; F:nucleic acid binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid | RNA binding |

| | | | |
|---|---|---|---|
| GABP01009936.1 | cell wall integrity scw1 | F:nucleotide binding; F:nucleic acid binding; F:oxidoreductase activity, acting on CH-OH group of donors; F:flavin adenine dinucleotide binding; F:oxidoreductase activity; P:oxidation-reduction process | RNA binding |
| GABP01010407.1 | zinc finger CCHC domain-containing 24-like | F:nucleic acid binding; F:zinc ion binding; P:homoiothermy; C:membrane; C:integral component of membrane; F:ice binding; P:response to freezing; F:poly(A) RNA binding | RNA binding |
| GABP01010590.1 | poly(rC)-binding 4 isoform X1 | F:molecular_function | RNA binding |
| GABP01013816.1 | nucleic acid-binding domain | C:BLOC-1 complex; F:metal ion binding | RNA binding |
| GABP01014523.1 | flowering time control FCA isoform X1 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid | RNA binding |
| GABP01015738.1 | RNA-binding -like | F:nucleotide binding; F:nucleic acid binding; F:transferase activity, transferring acyl groups other than amino-acyl groups; P:metabolic process | RNA binding |
| GABP01016362.1 | flowering time control FCA | F:nucleotide binding; F:nucleic acid binding; F:DNA binding; C:nucleus; F:RNA binding; P:regulation of flower development; C:cytoplasm; P:embryo development ending in seed dormancy; F:abscisic acid binding; P:embryo sac development; C:chromatin; P:chromatin silencing by small RNA | RNA binding |
| GABP01017627.1 | far upstream element-binding 3-like | F:nucleic acid binding; F:RNA binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01019261.1 | far upstream element-binding 1 | F:nucleic acid binding; F:transferase activity, transferring acyl groups; F:RNA binding; P:metabolic process | RNA binding |
| GABP01019640.1 | RNA recognition motif-containing | F:nucleotide binding; F:nucleic acid binding; C:nucleus; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid; P:RNA processing | RNA binding |
| GABP01020446.1 | far upstream element-binding partial | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; P:intracellular mRNA localization; C:cytoplasm; P:mRNA stabilization; P:telomere maintenance via telomerase; F:mRNA binding; C:cytoplasmic mRNA processing body; C:nuclear chromosome, telomeric region | RNA binding |
| GABP01021184.1 | nucleolin 2-like isoform X6 | F:nucleotide binding; F:nucleic acid binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid | RNA binding |
| GABP01022905.1 | ELAV 1 isoform X1 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; P:positive regulation of angiogenesis | RNA binding |
| GABP01024018.1 | RNA binding | F:molecular_function | RNA binding |
| GABP01024384.1 | sex-lethal homolog isoform X1 | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding | RNA binding |
| GABP01024419.1 | RNA binding | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid | RNA binding |
| GABP01025266.1 | far upstream element-binding 3 isoform X2 | F:molecular_function | RNA binding |
| GABP01025484.1 | far upstream element-binding 2-like | F:nucleic acid binding; F:RNA binding | RNA binding |

| | | | |
|---|---|---|---|
| GABP01025501.1 | cold-shock | P:pyridoxal phosphate biosynthetic process; F:transferase activity, transferring acyl groups; F:FMN binding; P:metabolic process; P:regulation of transcription, DNA-templated; C:plasma membrane; F:pyridoxamine-phosphate oxidase activity; F:nucleic acid binding; F:DNA binding; F:[acyl-carrier-protein] S-malonyltransferase activity; C:cytoplasm; F:oxidoreductase activity; P:oxidation-reduction process; F:transferase activity | RNA binding |
| GABP01026239.1 | RNA binding partial | F:nucleic acid binding; F:RNA binding; F:metal ion binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01026248.1 | far upstream element-binding 2 isoform X1 | C:cytoplasmic stress granule; P:mRNA catabolic process; F:poly(A) RNA binding; C:nucleoplasm; P:regulation of miRNA metabolic process; F:nucleic acid binding; C:nucleus; F:protein binding; F:RNA binding; P:positive regulation of gene expression; C:membrane; F:mRNA binding; P:RNA interference | RNA binding |
| GABP01026495.1 | domain K- type RNA binding s family | F:nucleic acid binding; F:RNA binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01026765.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | RNA binding |
| GABP01027441.1 | transcriptional regulator cold shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:regulation of transcription, DNA-templated | RNA binding |
| GABP01027915.1 | RNA binding partial | F:nucleic acid binding; F:transferase activity, transferring glycosyl groups; F:RNA binding; P:protein ADP-ribosylation; F:metal ion binding; F:NAD(P)+-protein-arginine ADP-ribosyltransferase activity; P:metabolic process; C:membrane; C:integral component of membrane; F:transferase activity; F:NAD+ ADP-ribosyltransferase activity | RNA binding |
| GABP01028113.1 | lupus la | C:nuclear body; F:protein disulfide oxidoreductase activity; P:protein transport; P:glycerol ether metabolic process; P:protein phosphorylation; F:GTP binding; F:nucleotide binding; F:ATP binding; C:cell; F:RNA binding; C:intracellular ribonucleoprotein complex; F:calcium ion binding; F:zinc ion binding; P:signal transduction; P:small GTPase mediated signal transduction; P:RNA processing; F:protein kinase activity; P:cell redox homeostasis; P:mRNA processing; P:regulation of heterochromatin domain assembly; C:intracellular; F:nucleic acid binding; C:nucleus; C:nuclear chromatin; C:NURS complex; F:metal ion binding; C:membrane; C:viral nucleocapsid; P:oxidation-reduction process | RNA binding |
| GABP01029051.1 | RNA-binding 19 | F:nucleotide binding; P:regulation of translation; F:nucleic acid binding; F:translation initiation factor activity; C:nucleus; F:RNA binding; P:mRNA metabolic process; C:cytoplasm; P:translational initiation | RNA binding |
| GABP01029420.1 | poly(rC)-binding 3-like isoform X3 | P:biological_process; F:molecular_function | RNA binding |
| GABP01029591.1 | RNA-binding 39 | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding; C:membrane; C:integral component of membrane; P:mRNA processing | RNA binding |
| GABP01029592.1 | flowering time control FCA isoform X2 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding | RNA binding |
| GABP01030617.1 | zinc finger C3H1 type-like 2 | F:molecular_function | RNA binding |
| GABP01032121.1 | CUGBP Elav-like family member 1 isoform X3 | F:molecular_function | RNA binding |

| | | | |
|---|---|---|---|
| GABP01032647.1 | far upstream element-binding protein 1-like | F:nucleic acid binding; C:plastid nucleoid; C:chloroplast; F:RNA binding; C:membrane; C:integral component of membrane; P:oxidation-reduction process; P:chloroplast rRNA processing; F:protein kinase activity; P:protein phosphorylation; F:monooxygenase activity | RNA binding |
| GABP01032692.1 | far upstream element-binding protein 1-like | F:protein complex scaffold; F:RNA binding; P:negative regulation of cardiac muscle hypertrophy; C:postsynaptic membrane; P:regulation of protein kinase A signaling; F:beta-tubulin binding; C:neuromuscular junction; P:phosphorylation; F:protein phosphatase 2B binding; F:protein kinase binding; F:protein domain specific binding; F:protein phosphatase binding; C:mitochondrial crista; C:endoplasmic reticulum; C:mitochondrial outer membrane; C:mitochondrial inner membrane; F:microtubule binding; F:poly(A) RNA binding; C:lipid particle; F:nucleic acid binding; F:protein binding; C:mitochondrial matrix; C:membrane; C:mitochondrion; C:integral component of membrane; F:protein kinase A regulatory subunit binding; P:negative regulation of protein dephosphorylation; F:kinase activity; P:negative regulation of NFAT protein import into nucleus | RNA binding |
| GABP01032705.1 | RNA binding | C:nucleus; F:RNA binding; C:cytoplasm | RNA binding |
| GABP01032960.1 | lupus la | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid; P:RNA processing | RNA binding |
| GABP01034427.1 | RNA-binding | F:ATPase activity; F:poly(A) binding; F:RNA polymerase II core binding; F:regulatory region RNA binding; P:mRNA splicing, via spliceosome; P:positive regulation of transcription elongation from RNA polymerase II promoter; F:nucleotide binding; F:ATP binding; P:translational termination; F:mRNA binding; P:negative regulation of translation; F:identical protein binding; F:ATPase activity, coupled to transmembrane movement of substances; P:transmembrane transport; F:eukaryotic initiation factor 4G binding; P:metabolic process; P:regulation of transcription, DNA-templated; P:transport; P:mRNA export from nucleus; F:nucleic acid binding; F:DNA binding; P:negative regulation of termination of RNA polymerase II transcription, poly(A)-coupled; C:nucleus; C:cytoplasm; C:membrane; C:integral component of membrane | RNA binding |
| GABP01035392.1 | cold-shock | F:tRNA (adenine-N1-)-methyltransferase activity; P:transcription, DNA-templated; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; P:regulation of transcription, DNA-templated; F:nucleic acid binding; F:DNA binding; F:L-ascorbic acid binding; C:tRNA (m1A) methyltransferase complex; C:cytoplasm; F:iron ion binding; F:metal ion binding; F:oxidoreductase activity; P:oxidation-reduction process; P:tRNA methylation | RNA binding |
| GABP01036076.1 | RNA binding partial | F:nucleic acid binding; F:RNA binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01036140.1 | sex-lethal homolog | F:nucleotide binding; F:nucleic acid binding; F:RNA binding | RNA binding |
| GABP01037289.1 | RNA recognition motif-containing | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding; C:intracellular ribonucleoprotein complex; F:U1 snRNP binding; C:viral nucleocapsid; P:mRNA processing | RNA binding |

| | | | |
|---|---|---|---|
| GABP01037323.1 | myelin expression factor 2 | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; C:intracellular ribonucleoprotein complex; C:cytoplasm; C:membrane; C:integral component of membrane; C:viral nucleocapsid | RNA binding |
| GABP01038649.1 | far upstream element-binding 3 isoform X2 | F:nucleic acid binding; F:RNA binding | RNA binding |
| GABP01039819.1 | Far upstream element-binding partial | F:nucleic acid binding; C:nucleus; F:RNA binding; P:positive regulation of gene expression; F:poly(A) RNA binding; C:nucleoplasm | RNA binding |
| GABP01041044.1 | zinc finger CCCH domain-containing 38-like | P:RNA processing; F:helicase activity; F:poly(A) RNA binding; F:ATP-dependent RNA helicase activity; C:nuclear pore; F:nucleotide binding; F:nucleic acid binding; F:ATP binding; C:nucleus; F:ubiquitin-protein transferase activity; F:metal ion binding; C:mitochondrion; P:poly(A)+ mRNA export from nucleus; F:hydrolase activity; P:protein ubiquitination | RNA binding |
| GABP01042032.1 | RNA recognition motif ( or RNP domain) partial | F:nucleotide binding; F:nucleic acid binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01043186.1 | RNA recognition motif-containing protein RRM | F:molecular_function | RNA binding |
| GABP01043295.1 | RNA-binding | F:molecular_function | RNA binding |
| GABP01043598.1 | cutinase negative acting | F:nucleotide binding; F:nucleic acid binding; C:nucleus; C:intracellular ribonucleoprotein complex; P:metabolic process; C:viral nucleocapsid; P:RNA processing; F:ligase activity; F:hydrolase activity | RNA binding |
| GABP01043843.1 | PREDICTED: RNA-binding protein Nova-1 isoform X2 | F:nucleic acid binding; F:RNA binding | RNA binding |
| GABP01044468.1 | RNA binding partial | F:nucleic acid binding; F:RNA binding; F:metal ion binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01044608.1 | RNA-binding | F:nucleotide binding; F:nucleic acid binding | RNA binding |

| | | | |
|---|---|---|---|
| GABP01046187.1 | RNA-binding 14 isoform X2 | P:DNA repair; P:negative regulation of translational initiation; P:negative regulation of translation in response to stress; C:nuclear speck; P:DNA replication; P:positive regulation of transcription from RNA polymerase II promoter; C:nucleolus; F:RNA polymerase II transcription cofactor activity; C:nucleoplasm; F:nucleotide binding; C:transcription factor complex; F:RNA binding; C:intracellular ribonucleoprotein complex; P:response to arsenic-containing substance; P:regulation of alternative mRNA splicing, via spliceosome; C:mediator complex; F:mRNA binding; P:histone deacetylation; P:intracellular estrogen receptor signaling pathway; P:negative regulation of translation; P:positive regulation of muscle cell differentiation; P:stress-activated MAPK cascade; F:miRNA binding; P:glucocorticoid receptor signaling pathway; F:zinc ion binding; C:cytoplasmic stress granule; P:miRNA mediated inhibition of translation; P:cap-independent translational initiation; P:transcription, DNA-templated; P:IRES-dependent translational initiation; P:SMAD protein signal transduction; P:DNA recombination; P:negative regulation of centriole replication; P:regulation of transcription, DNA-templated; P:response to hormone; P:regulation of nucleocytoplasmic transport; F:poly(A) RNA binding; F:nucleic acid binding; C:nucleus; F:pre-mRNA intronic binding; F:protein binding; F:metal ion binding; C:cytoplasm; F:pre-mRNA intronic pyrimidine-rich binding; F:ligand-dependent nuclear receptor transcription coactivator activity; F:protein binding, bridging | RNA binding |
| GABP01046668.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01046670.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01046827.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; F:electron carrier activity; P:regulation of transcription, DNA-templated; F:heme binding | RNA binding |
| GABP01047567.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01047904.1 | cold-shock | F:nucleic acid binding; F:DNA binding; C:cytoplasm; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated | RNA binding |
| GABP01049658.1 | ELAV 3 isoform X11 | F:RNA binding | RNA binding |
| GABP01049965.1 | pumilio homolog 2-like isoform X1 | C:nuclear envelope; C:cytoplasm; P:signal transduction; F:mRNA binding; P:chromosome segregation; P:response to stress; P:immune system process; P:translation | RNA binding |
| GABP01055365.1 | RNA-binding 25 isoform X1 | P:cell death; F:mRNA binding; P:mRNA processing; C:nucleoplasm | RNA binding |
| GABP01055475.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01056035.1 | homologous to Drosophila SQD (squid) | F:nucleotide binding; F:nucleic acid binding; P:reproduction | RNA binding |

| | | | |
|---|---|---|---|
| GABP01062611.1 | myelin expression factor 2 | F:ATPase activity; F:ATPase activity, coupled to transmembrane movement of substances; P:transmembrane transport; P:neuron differentiation; P:transcription, DNA-templated; P:metabolic process; P:transport; C:nucleoplasm; F:poly(A) RNA binding; F:nucleotide binding; F:nucleic acid binding; F:DNA binding; F:ATP binding; C:nucleus; F:protein binding; F:RNA binding; C:cytoplasm; C:membrane; C:integral component of membrane; P:myotube differentiation | RNA binding |
| GABP01062921.1 | RNA-binding squid isoform X1 | P:embryo development; C:nucleus; P:cellular component assembly; P:cell differentiation; P:anatomical structure formation involved in morphogenesis; P:developmental maturation; C:chromosome; P:cellular nitrogen compound metabolic process; F:molecular_function; P:reproduction | RNA binding |
| GABP01064418.1 | RNA-binding Nova- | F:molecular_function | RNA binding |
| GABP01065022.1 | RNA-binding | P:positive regulation of translation; P:multicellular organism development; P:regulation of mRNA stability; P:mRNA splicing, via spliceosome; C:nucleoplasm; F:mRNA 3'-UTR binding; F:nucleotide binding; P:maturation of LSU-rRNA from tricistronic rRNA transcript (SSU-rRNA, 5.8S rRNA, LSU-rRNA); F:RNA binding; F:double-stranded RNA binding; P:regulation of stem cell population maintenance; C:cytosol; F:mRNA binding; F:protein kinase binding; P:3'-UTR-mediated mRNA stabilization; F:protein homodimerization activity; P:RNA processing; F:poly(A) RNA binding; F:nucleic acid binding; F:AU-rich element binding; C:nucleus; F:protein binding; F:mRNA 3'-UTR AU-rich region binding; C:cytoplasm; P:ribosomal large subunit biogenesis; P:mRNA stabilization; C:membrane | RNA binding |
| GABP01067684.1 | rna-binding | F:molecular_function | RNA binding |
| GABP01069066.1 | cold-shock | F:nucleic acid binding; F:DNA binding; F:zinc ion binding; C:cytoplasm; P:transcription, DNA-templated; F:metal ion binding; P:regulation of transcription, DNA-templated | RNA binding |
| GABP01069067.1 | cold-shock | F:nucleic acid binding; F:DNA binding; F:zinc ion binding; C:cytoplasm; P:transcription, DNA-templated; F:metal ion binding; P:regulation of transcription, DNA-templated | RNA binding |
| GABP01069068.1 | cold-shock | F:nucleic acid binding; F:DNA binding; F:zinc ion binding; C:cytoplasm; P:transcription, DNA-templated; F:metal ion binding; P:regulation of transcription, DNA-templated | RNA binding |
| GABP01069615.1 | partial | F:zinc ion binding; C:cytoplasmic stress granule; P:metabolic process; F:hydrolase activity, hydrolyzing O-glycosyl compounds; C:nucleolus; F:nucleotide binding; F:nucleic acid binding; P:carbohydrate metabolic process; F:RNA binding; C:cytoplasm; F:metal ion binding; F:hydrolase activity; F:hydrolase activity, acting on glycosyl bonds | RNA binding |
| GABP01071338.1 | RNA-binding 19 | F:nucleotide binding; F:nucleic acid binding; P:lipid transport; F:lipid transporter activity; P:signal transduction; C:membrane; C:integral component of membrane; F:hedgehog receptor activity; C:nucleolus | RNA binding |
| GABP01071560.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01071563.1 | domain K- type RNA binding s family | F:nucleic acid binding; F:RNA binding; C:membrane; C:integral component of membrane | RNA binding |

| | | | |
|---|---|---|---|
| GABP01071565.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01073018.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01077596.1 | ELAV 2 isoform X2 | F:RNA binding; C:cytoplasm; P:cellular nitrogen compound metabolic process | RNA binding |
| GABP01078020.1 | KH domain-containing | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding | RNA binding |
| GABP01079007.1 | myelin expression factor 2 | F:nucleotide binding; F:nucleic acid binding; C:nucleus; C:cytoplasm; F:poly(A) RNA binding; C:nucleoplasm | RNA binding |
| GABP01079106.1 | RNA-binding with multiple partial | F:nucleotide binding; F:nucleic acid binding; F:RNA binding; F:translation elongation factor activity; F:metal ion binding; P:translational elongation | RNA binding |
| GABP01085339.1 | vigilin | F:nucleic acid binding; C:nucleus; F:transcription factor activity, sequence-specific DNA binding; F:RNA binding; F:double-stranded DNA binding; P:regulation of transcription, DNA-templated | RNA binding |
| GABP01086632.1 | vigilin-like | F:nucleic acid binding; P:biological_process; F:RNA binding; C:cellular_component | RNA binding |
| GABP01087662.1 | RNA recognition motif-containing | F:nucleotide binding; F:nucleic acid binding; C:nucleus; F:RNA binding; P:positive regulation of translation; C:cytoplasm; P:regulation of heart contraction; P:RNA processing; P:mRNA processing; F:poly(A) RNA binding; F:mRNA 3'-UTR binding | RNA binding |
| GABP01089013.1 | RNA-binding | F:nucleotide binding; P:regulation of translation; F:nucleic acid binding; F:RNA binding; C:intracellular ribonucleoprotein complex; C:viral nucleocapsid | RNA binding |
| GABP01089016.1 | RNA-binding | F:nucleotide binding; P:regulation of translation; F:nucleic acid binding; C:nucleus; F:RNA binding; C:intracellular ribonucleoprotein complex; C:cytoplasm; C:viral nucleocapsid | RNA binding |
| GABP01092671.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01092674.1 | kh domain-containing protein | F:molecular_function | RNA binding |
| GABP01093269.1 | RNA-binding Nova-2-like isoform X2 | F:nucleic acid binding; P:regulation by virus of viral protein levels in host cell; F:RNA binding; F:single-stranded RNA binding; C:cytosol; P:response to cytokinin | RNA binding |
| GABP01093435.1 | partial | F:RNA binding; P:cellular nitrogen compound metabolic process | RNA binding |
| GABP01096831.1 | RNA binding (nil per os homologue) | F:nucleotide binding; F:nucleic acid binding; C:membrane; C:integral component of membrane | RNA binding |
| GABP01100740.1 | RNA-binding | F:nucleotide binding; P:regulation of translation; F:nucleic acid binding; F:RNA binding; C:intracellular ribonucleoprotein complex; C:membrane; C:integral component of membrane; C:viral nucleocapsid | RNA binding |
| GABP01101104.1 | far upstream element-binding protein 1 | F:nucleotide binding; F:nucleic acid binding; F:DNA binding; C:nucleus; F:RNA binding; C:cytosol | RNA binding |
| GABP01102453.1 | zinc finger (CCCH type) motif-containing | F:zinc ion binding; F:alanine-tRNA ligase activity; F:metal ion binding; P:metabolic process; F:ligase activity | RNA binding |
| GABP01104797.1 | RNA recognition motif-containing | F:nucleotide binding; F:nucleic acid binding | RNA binding |
| GABP01105354.1 | cold-shock | F:zinc ion binding; P:transmembrane transport; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated; C:cell outer membrane; F:helicase activity; F:receptor activity; P:transport; F:metallopeptidase activity; F:nucleic acid binding; F:DNA binding; C:cytoplasm; F:metal ion binding; C:membrane; P:proteolysis; C:ATP-binding cassette (ABC) transporter complex | RNA binding |

| GABP01109124.1 | insulin-like growth factor 2 mRNA-binding 2 isoform X3 | F:RNA binding; C:intracellular | RNA binding |
|---|---|---|---|
| GABP01110215.1 | Far upstream element-binding 1 | F:nucleic acid binding; F:RNA binding | RNA binding |
| GABP01110582.1 | RNA binding partial | F:molecular_function | RNA binding |
| GABP01112355.1 | cold-shock | P:RNA phosphodiester bond hydrolysis, exonucleolytic; P:nucleic acid phosphodiester bond hydrolysis; P:RNA phosphodiester bond hydrolysis; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated; F:ribonuclease activity; F:nucleic acid binding; F:DNA binding; F:exoribonuclease II activity; P:RNA metabolic process; F:RNA binding; C:cytoplasm; F:exonuclease activity; F:nuclease activity; F:hydrolase activity | RNA binding |
| GABP01017222.1 | ATP-dependent RNA helicase DB10 | F:ATPase activity; C:mitochondrion; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01091872.1 | ATP-dependent RNA helicase DDX1 | P:aromatic compound catabolic process; P:metabolic process; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, NAD(P)H as one donor, and incorporation of two atoms of oxygen into one donor; F:helicase activity; F:dipeptidyl-peptidase activity; F:nucleotide binding; F:dioxygenase activity; F:nucleic acid binding; F:ATP binding; F:iron ion binding; F:oxidoreductase activity; F:choline monooxygenase activity; C:membrane; P:proteolysis; C:integral component of membrane; P:oxidation-reduction process; F:2 iron, 2 sulfur cluster binding; F:hydrolase activity | RNA Helicase |
| GABP01093778.1 | DEAD-box ATP-dependent RNA helicase 56-like | F:ATPase activity; F:DNA binding; P:nucleocytoplasmic transport; F:RNA binding; C:cell wall; F:ion binding; C:cytosol; F:helicase activity; P:mRNA processing; C:nucleolus; P:response to stress | RNA Helicase |
| GABP01006380.1 | ATP-dependent RNA helicase | F:ATPase activity; F:RNA binding; F:ion binding; C:cytosol; F:helicase activity; P:mRNA processing; C:nucleolus; C:plastid | RNA Helicase |
| GABP01008529.1 | ATP-dependent RNA helicase DED1 | C:cytoplasm; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:molecular_function | RNA Helicase |
| GABP01008697.1 | DEAD DEAH box helicase domain-containing | C:cytoplasm; F:molecular_function; P:translation | RNA Helicase |
| GABP01012697.1 | RNA helicase DRH1 | F:ATPase activity; C:mitochondrion; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01017210.1 | ATP-dependent RNA helicase dbp2 | C:cytoplasm; P:cellular nitrogen compound metabolic process; C:organelle; F:molecular_function | RNA Helicase |
| GABP01017225.1 | ATP-dependent RNA helicase DB10 | P:biological_process; F:helicase activity | RNA Helicase |
| GABP01017228.1 | DEAD-box ATP-dependent RNA helicase 40-like isoform X2 | F:ATPase activity; C:mitochondrion; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01017232.1 | ATP-dependent RNA helicase DBP2 | F:ATPase activity; F:RNA binding; C:mitochondrion; P:ribonucleoprotein complex assembly; F:ion binding; P:nucleobase-containing compound catabolic process; C:cytosol; P:ribosome biogenesis; F:helicase activity; C:nucleolus | RNA Helicase |
| GABP01017233.1 | ATP-dependent RNA helicase DBP2 | F:ATPase activity; F:RNA binding; C:mitochondrion; P:ribonucleoprotein complex assembly; F:ion binding; P:nucleobase-containing compound catabolic process; C:cytosol; P:ribosome biogenesis; F:helicase activity; C:nucleolus | RNA Helicase |

| | | | |
|---|---|---|---|
| GABP01019151.1 | DEAD-box ATP-dependent RNA helicase 46 | F:ATPase activity; C:mitochondrion; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01020165.1 | RNA helicase DRH1 | F:ATPase activity; C:nucleus; C:mitochondrion; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01020385.1 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 17 | F:ATPase activity; F:helicase activity; P:mRNA processing; C:intracellular | RNA Helicase |
| GABP01020413.1 | DEAD-box ATP-dependent RNA helicase 20 | F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01022189.1 | ATP-dependent RNA helicase DB10 | F:ATPase activity; C:mitochondrion; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01022953.1 | RNA helicase DRH1 | F:ATPase activity; F:helicase activity; P:cellular nitrogen compound metabolic process; C:organelle; C:intracellular | RNA Helicase |
| GABP01023570.1 | DEAD/DEAH box helicase family protein | P:cell division; C:nucleus; P:chromosome organization; P:cell differentiation; P:anatomical structure development; F:helicase activity; C:protein complex; P:mRNA processing; P:response to stress; P:growth | RNA Helicase |
| GABP01028351.1 | DEAD-family helicase | F:ATPase activity; C:nucleus; F:ion binding; F:helicase activity; P:cellular nitrogen compound metabolic process; C:plastid | RNA Helicase |
| GABP01029917.1 | DEAD-box ATP-dependent RNA helicase 35 | F:ATPase activity; C:nucleus; F:ion binding; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01033001.1 | ATP-dependent RNA helicase DHX36 | F:molecular_function | RNA Helicase |
| GABP01033638.1 | RNA helicase DRH1 | F:ATPase activity; C:nucleus; C:cytoplasm; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01033639.1 | RNA helicase DRH1 | F:ATPase activity; F:helicase activity; P:cellular nitrogen compound metabolic process; C:organelle; C:intracellular | RNA Helicase |
| GABP01033640.1 | RNA helicase DRH1 | F:ATPase activity; C:mitochondrion; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01038277.1 | DEAD-box ATP-dependent RNA helicase 20 | F:ATPase activity; C:cytosol; F:helicase activity; P:cellular nitrogen compound metabolic process; C:organelle | RNA Helicase |
| GABP01042647.1 | ATP-dependent RNA helicase DB10 | F:ATPase activity; F:helicase activity; P:cellular nitrogen compound metabolic process; C:organelle; C:intracellular | RNA Helicase |
| GABP01043763.1 | helicase associated domain (ha2) | F:helicase activity | RNA Helicase |
| GABP01044710.1 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 17 | F:ATPase activity; C:nucleus; F:RNA binding; F:ion binding; P:nucleobase-containing compound catabolic process; C:cytosol; P:ribosome biogenesis; F:helicase activity | RNA Helicase |
| GABP01044711.1 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 17 | F:ATPase activity; F:ion binding; C:cytosol; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |

| | | | |
|---|---|---|---|
| GABP01047237.1 | ATP-dependent RNA helicase DDX1 | P:multicellular organism development; F:poly(A) binding; P:tRNA splicing, via endonucleolytic cleavage and ligation; F:ATP-dependent helicase activity; F:helicase activity; P:response to virus; P:regulation of nucleic acid-templated transcription; C:nucleoplasm; F:ATP-dependent RNA helicase activity; P:double-strand break repair; F:nucleotide binding; F:ATP binding; C:intracellular ribonucleoprotein complex; F:double-stranded RNA binding; P:DNA duplex unwinding; P:response to exogenous dsRNA; F:nuclease activity; F:DNA/RNA helicase activity; C:tRNA-splicing ligase complex; P:nucleic acid phosphodiester bond hydrolysis; C:cytoplasmic stress granule; P:metabolic process; F:chromatin binding; F:poly(A) RNA binding; F:nucleic acid binding; C:nucleus; F:transcription cofactor activity; C:cytoplasm; C:membrane; C:integral component of membrane; P:RNA secondary structure unwinding; C:cleavage body; F:hydrolase activity | RNA Helicase |
| GABP01049826.1 | ATP-dependent RNA helicase dbp9 | F:nucleic acid binding; F:ATP binding; P:metabolic process; C:cytosol; F:helicase activity; F:hydrolase activity | RNA Helicase |
| GABP01053330.1 | helicase associated domain (ha2) | F:nucleotide binding; F:nucleic acid binding; F:hydrolase activity, acting on ester bonds; P:nucleic acid phosphodiester bond hydrolysis; P:DNA catabolic process; F:ATP binding; F:endonuclease activity; P:metabolic process; F:helicase activity; F:hydrolase activity | RNA Helicase |
| GABP01053373.1 | DEAD-box ATP-dependent RNA helicase 46-like isoform X1 | P:biological_process; F:helicase activity | RNA Helicase |
| GABP01053559.1 | Atp-dependent RNA helicase | F:RNA-dependent ATPase activity; P:cytoplasmic translational initiation; F:helicase activity; P:rRNA processing; C:nucleolus; C:acetyl-CoA carboxylase complex; P:mRNA splicing, via spliceosome; F:ATP-dependent RNA helicase activity; F:nucleotide binding; F:translation initiation factor activity; F:ATP binding; F:RNA binding; P:nuclear-transcribed mRNA catabolic process, nonsense-mediated decay; F:acetyl-CoA carboxylase activity; C:cytosol; F:ligase activity; P:metabolic process; P:fatty acid biosynthetic process; P:messenger ribonucleoprotein complex assembly; P:nuclear polyadenylation-dependent mRNA catabolic process; F:nucleic acid binding; C:nucleus; C:membrane; C:mitochondrion; C:integral component of membrane; P:regulation of gene expression; P:RNA secondary structure unwinding; F:hydrolase activity | RNA Helicase |
| GABP01056293.1 | DEAD-box ATP-dependent RNA helicase 14-like | F:ATPase activity; F:helicase activity; P:cellular nitrogen compound metabolic process; C:organelle; C:intracellular | RNA Helicase |
| GABP01056294.1 | DEAD-box ATP-dependent RNA helicase 46 | F:ATPase activity; P:biological_process; F:helicase activity; C:organelle; C:intracellular | RNA Helicase |
| GABP01068954.1 | ATP-dependent RNA | F:ATPase activity; C:cytoplasm; P:ribonucleoprotein complex assembly; F:ion binding; F:translation factor activity, RNA binding; F:helicase activity; C:organelle | RNA Helicase |
| GABP01077605.1 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 17 | F:ATPase activity; F:ion binding; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01086022.1 | DEAD-box ATP-dependent RNA helicase 37-like | F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |
| GABP01091378.1 | ATP-dependent RNA helicase DDX1 | C:nucleus; F:RNA binding; C:cytoplasm; F:helicase activity; P:cellular nitrogen compound metabolic process | RNA Helicase |

| | | | |
|---|---|---|---|
| GABP01093777.1 | ATP-dependent RNA helicase WM6 | F:ATPase activity; C:nucleus; P:nucleocytoplasmic transport; P:chromosome organization; F:ion binding; F:helicase activity; C:protein complex; C:chromosome; P:mRNA processing; P:biosynthetic process; P:DNA metabolic process; P:response to stress | RNA Helicase |
| GABP01093780.1 | DEAD-box ATP-dependent RNA helicase 56 isoform X1 | F:ATPase activity; F:ion binding; F:helicase activity; C:protein complex; C:chromosome; P:mRNA processing; P:biosynthetic process; P:DNA metabolic process; P:response to stress; C:nucleus; P:nucleocytoplasmic transport; F:RNA binding; P:chromosome organization | RNA Helicase |
| GABP01093783.1 | DEAD (Asp-Glu-Ala-Asp) box polypeptide DDX39 | F:ATPase activity; P:catabolic process; F:ion binding; F:helicase activity; C:protein complex; F:peptidase activity; P:mRNA processing; C:nucleolus; P:response to stress; F:DNA binding; P:nucleocytoplasmic transport; F:RNA binding; C:cytosol | RNA Helicase |
| GABP01098012.1 | DEAD (Asp-Glu-Ala-Asp) box polypeptide 17 | F:ATPase activity; C:nucleus; C:cytoplasm; F:ion binding; P:nucleobase-containing compound catabolic process; F:helicase activity | RNA Helicase |
| GABP01101885.1 | ATP-dependent RNA helicase DED1 | F:ATPase activity; C:cytoplasm; F:ion binding; F:helicase activity; P:cellular nitrogen compound metabolic process; C:organelle | RNA Helicase |
| GABP01030241.1 | Luciferase partial | F:oxidoreductase activity; C:cytoplasmic, membrane-bounded vesicle | Scintillon |
| GABP01030236.1 | Luciferase partial | F:oxidoreductase activity; C:cytoplasmic, membrane-bounded vesicle | Scintillon |
| GABP01030239.1 | Luciferase partial | F:oxidoreductase activity; C:cytoplasmic, membrane-bounded vesicle | Scintillon |
| GABP01030243.1 | Luciferase partial | F:oxidoreductase activity; C:cytoplasmic, membrane-bounded vesicle | Scintillon |
| GABP01088786.1 | luciferin binding partial | F:oxidoreductase activity; C:cytoplasmic, membrane-bounded vesicle | Scintillon |
| GABP01000767.1 | DNA-directed RNA polymerase II largest | F:DNA binding; F:nucleotidyltransferase activity; P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:plastid | Transcription |
| GABP01009293.1 | spt5p | P:cellular nitrogen compound metabolic process; P:biosynthetic process; C:organelle; C:intracellular; F:molecular_function | Transcription |
| GABP01018344.1 | transcription elongation factor A 2 isoform X2 | F:molecular_function | Transcription |
| GABP01022305.1 | SNW domain-containing 1 | P:embryo development; C:nucleus; P:signal transduction; P:cell morphogenesis; P:cell differentiation; P:anatomical structure formation involved in morphogenesis; P:cellular nitrogen compound metabolic process; P:biosynthetic process; F:molecular_function | Transcription |
| GABP01039884.1 | prohibitin | P:cell proliferation; C:cellular_component | Transcription |
| GABP01040436.1 | RNA polymerase-associated RTF1 homolog | P:histone H3-K4 trimethylation; P:stem cell population maintenance; P:DNA-templated transcription, initiation; C:nucleolus; P:positive regulation of transcription from RNA polymerase II promoter; F:poly(A) RNA binding; P:transcription elongation from RNA polymerase II promoter; F:single-stranded DNA binding; P:negative regulation of transcription from RNA polymerase II promoter; P:positive regulation of transcription elongation from RNA polymerase II promoter; F:DNA binding; C:nucleus; P:endodermal cell fate commitment; P:histone modification; P:positive regulation of histone H3-K4 methylation; C:Cdc73/Paf1 complex | Transcription |

| | | | |
|---|---|---|---|
| GABP01041516.1 | RNA polymerase II-associated factor 1 homolog | P:histone modification; C:Cdc73/Paf1 complex; P:histone methylation; P:negative regulation of flower development; P:transcription elongation from RNA polymerase II promoter | Transcription |
| GABP01075003.1 | DNA-directed RNA polymerase II subunit RPB2 | P:embryo development; F:DNA binding; F:nucleotidyltransferase activity; C:cytosol; C:protein complex; P:cellular nitrogen compound metabolic process; P:biosynthetic process; P:reproduction; C:nucleoplasm | Transcription |
| GABP01091603.1 | RNA polymerase-associated CTR9 homolog | P:embryo development; P:chromosome organization; P:cellular protein modification process; P:cell differentiation; P:anatomical structure formation involved in morphogenesis; P:cellular nitrogen compound metabolic process; P:biosynthetic process | Transcription |
| GABP01110885.1 | cdc73 domain protein, partial | P:biological_process | Transcription |
| GABP01112174.1 | Transcription elongation factor SPT6 | F:nucleic acid binding; F:DNA binding; P:heart development; F:translation elongation factor activity; P:somitogenesis; P:negative regulation of histone H3-K27 methylation; P:regulation of DNA-templated transcription, elongation; P:embryonic heart tube development; P:otic vesicle development; P:nucleobase-containing compound metabolic process; P:translational elongation; P:regulation of transcription from RNA polymerase II promoter | Transcription |
| GABP01008323.1 | eukaryotic translation initiation factor 3 subunit A | C:cytoplasm | Translation |
| GABP01008644.1 | translation elongation factor IF5A | C:cytoplasm; F:translation factor activity, RNA binding | Translation |
| GABP01014342.1 | eukaryotic translation initiation factor | F:translation initiation factor activity; F:oxidoreductase activity; P:metabolic process; P:oxidation-reduction process; P:translational initiation; P:regulation of translational initiation; F:oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor | Translation |
| GABP01019058.1 | eukaryotic peptide chain release factor subunit 1 isoform X1 | C:nucleus; C:cytoplasm; F:translation factor activity, RNA binding; P:DNA metabolic process | Translation |
| GABP01020481.1 | eukaryotic initiation factor 4A-III | F:ATPase activity; P:embryo development; P:signal transduction; P:protein complex assembly; F:ion binding; P:cellular protein modification process; F:translation factor activity, RNA binding; F:helicase activity; P:developmental maturation; C:protein complex; P:mRNA processing; P:response to stress; P:reproduction; C:nucleus; P:cell differentiation; C:cytosol | Translation |
| GABP01025837.1 | elongation factor 3 | F:nucleotide binding; F:ATPase activity; F:ATP binding; C:nucleus; F:translation elongation factor activity; P:metabolic process; P:translational elongation; F:hydrolase activity | Translation |
| GABP01025838.1 | elongation factor 3 | F:nucleotide binding; F:ATPase activity; F:ATP binding; C:nucleus; F:translation elongation factor activity; P:metabolic process; P:translational elongation; F:hydrolase activity | Translation |
| GABP01025839.1 | elongation factor ef-3 | F:molecular_function | Translation |
| GABP01025840.1 | Elongation factor 3 | F:molecular_function | Translation |
| GABP01025842.1 | elongation factor 3 | F:ATPase activity; C:nucleus; F:ion binding; F:translation factor activity, RNA binding | Translation |
| GABP01036791.1 | atp-dependent rna helicase eif4a | F:ATPase activity; C:nucleus; F:ion binding; C:cytosol; F:mRNA binding; F:translation factor activity, RNA binding; F:helicase activity; C:protein complex | Translation |

| | | | |
|---|---|---|---|
| GABP01040136.1 | translation elongation factor | F:GTPase activity; F:ion binding; F:translation factor activity, RNA binding | Translation |
| GABP01040144.1 | translation elongation factor | P:carbohydrate metabolic process; F:GTPase activity; F:ion binding; F:translation factor activity, RNA binding; C:intracellular; F:hydrolase activity, acting on glycosyl bonds | Translation |
| GABP01040147.1 | translation elongation factor | P:carbohydrate metabolic process; F:GTPase activity; F:ion binding; F:translation factor activity, RNA binding; C:intracellular; F:hydrolase activity, acting on glycosyl bonds | Translation |
| GABP01043468.1 | elongation factor 2 | C:cytoplasm; F:GTPase activity; F:ion binding; F:translation factor activity, RNA binding | Translation |
| GABP01058504.1 | Translation machinery-associated 22 | F:translation initiation factor activity; F:peptidyl-prolyl cis-trans isomerase activity; F:zinc ion binding; F:RNA binding; P:protein complex assembly; C:mitochondrion; C:membrane; C:integral component of membrane; P:protein peptidyl-prolyl isomerization; P:translational initiation; P:protein folding; F:isomerase activity | Translation |
| GABP01066197.1 | Eukaryotic translation initiation factor 2 subunit 3 | C:nucleus; P:transmembrane transport; F:GTPase activity; P:ribonucleoprotein complex assembly; F:ion binding; C:cytosol; F:translation factor activity, RNA binding; C:extracellular region | Translation |
| GABP01085272.1 | Ubiquitin-associated translation elongation factor EF1B | F:ubiquitin-protein transferase activity; F:zinc ion binding; C:plasmodesma; F:translation elongation factor activity; F:metal ion binding; F:oxidoreductase activity; P:oxidation-reduction process; P:translational elongation; P:protein ubiquitination | Translation |
| GABP01103058.1 | eukaryotic translation initiation factor 3 subunit | F:RNA binding; C:cytoplasm; P:translation | Translation |
| GABP01001793.1 | beta-adaptin B | C:nucleus; P:vesicle-mediated transport; C:cytosol; F:protein transporter activity; C:Golgi apparatus; C:cytoplasmic, membrane-bounded vesicle; C:protein complex; C:plasma membrane | Transport |
| GABP01002650.1 | beta-adaptin C | P:vesicle-mediated transport; C:cytoplasm; F:protein transporter activity; C:protein complex; C:organelle; C:plasma membrane | Transport |
| GABP01018411.1 | clathrin heavy | C:cytoplasmic, membrane-bounded vesicle; C:protein complex; P:transport | Transport |
| GABP01023614.1 | vacuolar ATP synthase catalytic subunit A | F:ATPase activity; P:small molecule metabolic process; F:transmembrane transporter activity; F:ion binding; C:protein complex; P:cellular nitrogen compound metabolic process | Transport |
| GABP01024966.1 | vacuolin A | C:membrane; P:proteolysis; C:integral component of membrane; F:peptidase activity | Transport |
| GABP01027328.1 | AP-1 complex subunit mu-1 | P:membrane organization; C:nucleus; C:endosome; P:vesicle-mediated transport; C:microtubule organizing center; P:anatomical structure development; C:cytosol; C:Golgi apparatus; P:vacuolar transport; C:cytoplasmic, membrane-bounded vesicle; C:protein complex; F:molecular_function | Transport |
| GABP01027349.1 | unc-119 homolog B-like | P:cell morphogenesis | Transport |
| GABP01027389.1 | ADP-ribosylation factor | P:signal transduction; F:ion binding; P:transport; C:intracellular | Transport |
| GABP01034904.1 | conserved hypothetical protein | P:transport | Transport |
| GABP01051746.1 | transport Sec31B isoform X2 | C:membrane; C:integral component of membrane; C:vesicle coat | Transport |
| GABP01058496.1 | coatomer epsilon subunit | P:vesicle-mediated transport; C:intracellular | Transport |

| | | | |
|---|---|---|---|
| GABP01063693.1 | AP-4 complex subunit epsilon | P:vesicle-mediated transport; C:cytosol; C:protein complex; F:molecular_function | Transport |
| GABP01068409.1 | clathrin heavy | P:vesicle-mediated transport; F:structural molecule activity; C:cytoplasmic, membrane-bounded vesicle; C:Golgi apparatus; C:protein complex; C:plasma membrane | Transport |
| GABP01071504.1 | ADP-ribosylation factor 3 | P:signal transduction; P:cell morphogenesis; F:ion binding; C:Golgi apparatus; C:extracellular region; P:cell cycle; P:cell division; C:nucleus; P:vesicle-mediated transport; C:microtubule organizing center; P:cytoskeleton-dependent intracellular transport; C:cytoplasm; F:GTPase activity; C:cilium; P:cell differentiation; F:cytoskeletal protein binding | Transport |
| GABP01074882.1 | coatomer subunit gamma | C:nucleus; C:cytosol; C:cytoplasmic, membrane-bounded vesicle; C:Golgi apparatus; C:protein complex; P:transport; C:plasma membrane; C:plastid; F:molecular_function | Transport |
| GABP01074883.1 | coatomer gamma 2-subunit | C:cytoplasm; C:Golgi apparatus; P:transport | Transport |
| GABP01076315.1 | AP-1 complex subunit gamma-1 | P:vesicle-mediated transport; F:oxidoreductase activity; C:Golgi apparatus; C:cytoplasmic, membrane-bounded vesicle; P:vacuolar transport; C:protein complex | Transport |
| GABP01085735.1 | Beta-adaptin A | C:cytoplasm; C:protein complex; C:organelle; P:transport; F:enzyme binding | Transport |
| GABP01087543.1 | coatomer subunit beta | P:vesicle-mediated transport; C:cytoplasmic, membrane-bounded vesicle; C:Golgi apparatus; C:protein complex; F:enzyme binding | Transport |
| GABP01091689.1 | coatomer subunit alpha-1-like | P:vesicle-mediated transport; C:cytoplasm; C:Golgi apparatus; C:protein complex; F:molecular_function | Transport |
| GABP01096640.1 | vesicle-fusing ATPase | P:vesicle-mediated transport; C:Golgi apparatus; F:molecular_function | Transport |
| GABP01105178.1 | dynamin | C:nuclear envelope; C:cytoplasmic, membrane-bounded vesicle; C:extracellular region; C:plasma membrane; F:molecular_function; C:lysosome | Transport |
| GABP01000601.1 | cysteine--tRNA ligase | F:zinc ion binding; P:metabolic process; F:helicase activity; P:translation; P:cysteinyl-tRNA aminoacylation; F:nucleotide binding; F:nucleic acid binding; P:tRNA aminoacylation for protein translation; F:ATP binding; F:aminoacyl-tRNA ligase activity; F:RNA binding; C:cytoplasm; F:metal ion binding; F:cysteine-tRNA ligase activity; F:ligase activity; F:hydrolase activity | tRNA |
| GABP01006089.1 | probable isoleucine--tRNA cytoplasmic | F:ion binding; C:cytosol; P:tRNA metabolic process; P:cellular amino acid metabolic process; F:ligase activity; P:translation | tRNA |
| GABP01010487.1 | methionine-tRNA ligase | C:cytoplasm; P:tRNA metabolic process; P:cellular amino acid metabolic process; F:ligase activity; P:translation | tRNA |
| GABP01010489.1 | tyrosyl or methionyl-tRNA synthetase | P:cell proliferation; P:cell-cell signaling; P:tRNA metabolic process; P:cellular amino acid metabolic process; C:protein complex; C:extracellular region; C:organelle; P:response to stress; P:immune system process; P:translation; P:cell motility; F:RNA binding; C:cytoplasm; F:ligase activity; F:enzyme binding | tRNA |

| | | | |
|---|---|---|---|
| GABP01012351.1 | cysteine--tRNA ligase | P:nucleic acid phosphodiester bond hydrolysis; F:transferase activity, transferring glycosyl groups; F:zinc ion binding; P:nucleoside transmembrane transport; P:transport; P:cysteinyl-tRNA aminoacylation; P:translation; F:nucleotide binding; F:nucleic acid binding; P:tRNA aminoacylation for protein translation; F:ATP binding; F:nucleoside transmembrane transporter activity; F:aminoacyl-tRNA ligase activity; F:RNA binding; C:cytoplasm; F:exonuclease activity; F:metal ion binding; F:oxidoreductase activity; C:membrane; C:integral component of membrane; F:cysteine-tRNA ligase activity; P:oxidation-reduction process; F:ligase activity | tRNA |
| GABP01012352.1 | cysteine--tRNA ligase | P:nucleic acid phosphodiester bond hydrolysis; F:transferase activity, transferring glycosyl groups; F:zinc ion binding; P:nucleoside transmembrane transport; P:transport; P:translation; P:cysteinyl-tRNA aminoacylation; F:nucleotide binding; F:nucleic acid binding; P:tRNA aminoacylation for protein translation; F:ATP binding; F:nucleoside transmembrane transporter activity; F:aminoacyl-tRNA ligase activity; F:RNA binding; C:cytoplasm; F:exonuclease activity; F:metal ion binding; F:oxidoreductase activity; C:membrane; C:integral component of membrane; F:cysteine-tRNA ligase activity; P:oxidation-reduction process; F:ligase activity | tRNA |
| GABP01025655.1 | tRNA-dihydrouridine(47) synthase [NAD(P)(+)]-like | F:RNA binding; F:oxidoreductase activity; P:tRNA metabolic process | tRNA |
| GABP01027998.1 | cysteine--tRNA ligase | F:zinc ion binding; P:translation; P:cysteinyl-tRNA aminoacylation; F:nucleotide binding; F:nucleic acid binding; P:tRNA aminoacylation for protein translation; F:ATP binding; F:aminoacyl-tRNA ligase activity; F:RNA binding; C:cytoplasm; F:metal ion binding; F:cysteine-tRNA ligase activity; F:ligase activity | tRNA |
| GABP01030307.1 | DUF2431 domain | F:zinc ion binding; P:tRNA processing; P:phenylalanyl-tRNA aminoacylation; F:S-adenosylmethionine-dependent methyltransferase activity; C:nucleolus; F:tRNA binding; F:magnesium ion binding; F:rRNA (uridine-N3-)-methyltransferase activity; C:nucleus; F:ATP binding; C:cytoplasm; F:metal ion binding; P:rRNA base methylation; F:phenylalanine-tRNA ligase activity; C:membrane; C:integral component of membrane | tRNA |
| GABP01030310.1 | ferredoxin-fold anticodon-binding domain-containing 1 isoform X1 | F:zinc ion binding; P:tRNA processing; P:phenylalanyl-tRNA aminoacylation; C:nucleolus; F:tRNA binding; F:S-adenosylmethionine-dependent methyltransferase activity; F:magnesium ion binding; F:rRNA (uridine-N3-)-methyltransferase activity; C:nucleus; F:ATP binding; C:cytoplasm; P:rRNA base methylation; F:metal ion binding; F:phenylalanine-tRNA ligase activity; C:membrane; C:integral component of membrane | tRNA |
| GABP01030833.1 | lupus la | C:intracellular; F:molecular_function | tRNA |
| GABP01033307.1 | cysteine--tRNA ligase | P:nucleic acid phosphodiester bond hydrolysis; F:transferase activity, transferring glycosyl groups; F:zinc ion binding; P:cysteinyl-tRNA aminoacylation; P:translation; F:nucleotide binding; F:nucleic acid binding; P:tRNA aminoacylation for protein translation; F:ATP binding; F:aminoacyl-tRNA ligase activity; F:RNA binding; C:cytoplasm; F:exonuclease activity; F:metal ion binding; C:membrane; F:cysteine-tRNA ligase activity; F:ligase activity | tRNA |

| GABP01040241.1 | aspartyl glutamyl-tRNA(Asn Gln) A subunit [ [ | F:nucleotide binding; F:ATP binding; C:membrane; C:integral component of membrane; F:transferase activity; F:carbon-nitrogen ligase activity, with glutamine as amido-N-donor; P:transport; F:ligase activity; F:transporter activity; F:glutaminyl-tRNA synthase (glutamine-hydrolyzing) activity; P:translation | tRNA |
|---|---|---|---|
| GABP01041291.1 | valine-tRNA partial | C:cytoplasm; P:cell differentiation; P:anatomical structure development; P:tRNA metabolic process; P:cellular amino acid metabolic process; F:ligase activity; P:translation | tRNA |
| GABP01041716.1 | alanine--tRNA ligase | F:RNA binding; C:mitochondrion; F:ion binding; C:cytosol; P:tRNA metabolic process; P:cellular amino acid metabolic process; C:plastid; F:ligase activity; P:translation | tRNA |
| GABP01050728.1 | tRNA (guanine(37)-N1)-methyltransferase | C:cytoplasm; F:methyltransferase activity | tRNA |
| GABP01053573.1 | queuine tRNA-ribosyltransferase subunit qtrtd1-like | F:transferase activity, transferring glycosyl groups; P:queuosine biosynthetic process; C:cytoplasm; F:transferase activity; P:tRNA modification; F:queuine tRNA-ribosyltransferase activity | tRNA |
| GABP01057052.1 | aminoacyl tRNA synthase complex-interacting multifunctional 1 | P:cell proliferation; P:cell-cell signaling; P:tRNA metabolic process; P:cellular amino acid metabolic process; C:protein complex; C:extracellular region; C:organelle; P:response to stress; P:immune system process; P:translation; P:cell motility; F:RNA binding; C:cytoplasm; F:enzyme binding | tRNA |
| GABP01057110.1 | D-tyrosyl-tRNA(Tyr) deacylase | F:hydrolase activity, acting on ester bonds; F:D-tyrosyl-tRNA(Tyr) deacylase activity; C:cytoplasm; P:regulation of translational fidelity; F:aminoacyl-tRNA editing activity; P:tRNA metabolic process; P:D-amino acid catabolic process; F:hydrolase activity | tRNA |
| GABP01068015.1 | leucine--tRNA cytoplasmic isoform X1 | C:cytoplasm; P:tRNA metabolic process; P:cellular amino acid metabolic process; F:ligase activity; P:translation | tRNA |
| GABP01071863.1 | aspartyl-tRNA synthetase | C:cytoskeleton; P:cell differentiation; P:anatomical structure development; P:tRNA metabolic process; P:cellular amino acid metabolic process; C:protein complex; F:ligase activity; P:growth; P:translation | tRNA |
| GABP01072521.1 | threonyl-tRNA synthetase family | C:cytoplasm; P:tRNA metabolic process; P:cellular amino acid metabolic process; F:ligase activity; P:translation | tRNA |
| GABP01073479.1 | tRNA-specific adenosine deaminase | P:biological_process; F:ion binding | tRNA |
| GABP01079005.1 | Serine-arginine 55 | F:nucleotide binding; F:nucleic acid binding; P:tRNA aminoacylation for protein translation; P:gluconeogenesis; F:ATP binding; F:glucose-6-phosphate isomerase activity; F:aminoacyl-tRNA ligase activity; C:membrane; P:glycolytic process; C:integral component of membrane; C:cytosol | tRNA |
| GABP01099160.1 | arginyl-tRNA synthetase family | P:tRNA metabolic process; P:cellular amino acid metabolic process; F:ligase activity; P:translation | tRNA |

| | | | |
|---|---|---|---|
| GABP01112799.1 | cysteine--tRNA ligase | F:transferase activity, transferring glycosyl groups; F:zinc ion binding; P:nucleoside transmembrane transport; P:RNA processing; F:cysteine synthase activity; P:transport; P:cysteine biosynthetic process from serine; P:translation; P:cysteinyl-tRNA aminoacylation; F:nucleotide binding; F:nucleic acid binding; P:tRNA aminoacylation for protein translation; F:ATP binding; F:nucleoside transmembrane transporter activity; F:aminoacyl-tRNA ligase activity; F:RNA binding; C:cytoplasm; F:metal ion binding; F:oxidoreductase activity; C:membrane; C:integral component of membrane; F:cysteine-tRNA ligase activity; P:oxidation-reduction process; F:ligase activity | tRNA |
| GABP01033975.1 | hypothetical protein | | Unknown function |
| GABP01045030.1 | ---NA--- | | Unknown function |
| GABP01046140.1 | ---NA--- | | Unknown function |
| GABP01076526.1 | Brf1p family coiled coil | C:cell | Unknown function |
| GABP01000383.1 | hypothetical protein F442_13252 | | Unknown function |
| GABP01001215.1 | ---NA--- | | Unknown function |
| GABP01001919.1 | ---NA--- | | Unknown function |
| GABP01001945.1 | ---NA--- | | Unknown function |
| GABP01001947.1 | ---NA--- | | Unknown function |
| GABP01003403.1 | hypothetical protein | F:copper ion binding; F:oxidoreductase activity; C:membrane; C:integral component of membrane; P:oxidation-reduction process | unknown function |
| GABP01006726.1 | hypothetical protein Pmar_PMAR008308 | F:nucleic acid binding | Unknown function |
| GABP01007166.1 | ---NA--- | | Unknown function |
| GABP01011623.1 | Fe-S cluster assembly ATPase | F:nucleotide binding; F:ATPase activity; F:ATP binding; F:translation elongation factor activity; P:metabolic process; P:transport; P:translational elongation | Unknown function |
| GABP01011834.1 | unknown | F:nucleotide binding; P:RNA methylation; F:ATP binding; P:7-methylguanosine RNA capping; F:methyltransferase activity; F:transferase activity; P:methylation | Unknown function |
| GABP01012262.1 | hypothetical protein | | Unknown function |
| GABP01012608.1 | Hypothetical protein | F:ATP binding; C:membrane; C:integral component of membrane; F:protein kinase activity; F:ADP binding; P:protein phosphorylation | Unknown function |
| GABP01012640.1 | PREDICTED: calmodulin-like protein 4 | F:nucleotide binding; F:ATP binding; F:microtubule motor activity; P:metabolic process; P:microtubule-based movement; F:microtubule binding; C:microtubule | Unknown function |
| GABP01013748.1 | ---NA--- | | Unknown function |
| GABP01013841.1 | ---NA--- | | Unknown function |
| GABP01015476.1 | hypothetical protein Pmar_PMAR005561 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01016206.1 | Hypothetical protein | P:membrane organization; P:response to organic substance; P:Gram-negative-bacterium-type cell outer membrane assembly; C:membrane; C:integral component of membrane; C:cell outer membrane; C:outer membrane; P:lipopolysaccharide transport | Unknown function |

| | | | |
|---|---|---|---|
| GABP01016308.1 | ---NA--- | | Unknown function |
| GABP01017049.1 | unknown | | Unknown function |
| GABP01017373.1 | hypothetical protein Pmar_PMAR024059 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01017417.1 | ---NA--- | | Unknown function |
| GABP01018154.1 | hypothetical protein | | Unknown function |
| GABP01019126.1 | ---NA--- | | Unknown function |
| GABP01019668.1 | translation initiation factor IF-2 | F:nucleotide binding; F:GTP binding; F:translation initiation factor activity; C:cytoplasm; P:signal transduction; F:GTPase activity; C:membrane; P:translational initiation; C:intracellular; P:translation | Unknown function |
| GABP01019816.1 | unknown | C:membrane; C:integral component of membrane | Unknown function |
| GABP01019931.1 | ---NA--- | | Unknown function |
| GABP01020444.1 | hypothetical protein | F:catalytic activity; P:metabolic process; C:membrane; C:integral component of membrane | Unknown function |
| GABP01021625.1 | hypothetical protein | P:transcription, DNA-templated; P:proline metabolic process; P:metabolic process; P:regulation of transcription, DNA-templated; P:proline biosynthetic process; F:1-pyrroline-5-carboxylate dehydrogenase activity; F:DNA binding; F:transcription factor activity, sequence-specific DNA binding; F:proline dehydrogenase activity; P:glutamate biosynthetic process; F:oxidoreductase activity; P:oxidation-reduction process; F:benzaldehyde dehydrogenase (NAD+) activity; F:oxidoreductase activity, acting on the aldehyde or oxo group of donors, NAD or NADP as acceptor; P:proline catabolic process to glutamate | Unknown function |
| GABP01021749.1 | HEPN domain | F:DNA binding | Unknown function |
| GABP01022526.1 | ---NA--- | | Unknown function |
| GABP01022967.1 | ---NA--- | | Unknown function |
| GABP01023035.1 | unnamed protein product | | Unknown function |
| GABP01023209.1 | ---NA--- | | Unknown function |
| GABP01023212.1 | calmodulin-domain kinase 1 | F:nucleotide binding; F:ATP binding; F:metal ion binding; F:calcium ion binding; F:protein kinase activity; P:phosphorylation; P:protein phosphorylation; F:kinase activity | Unknown function |
| GABP01024097.1 | ---NA--- | | Unknown function |
| GABP01024587.1 | hypothetical protein | | Unknown function |
| GABP01024647.1 | RNase III inhibitor | F:pyridoxal phosphate binding; F:zinc ion binding; F:catalytic activity; C:membrane; C:integral component of membrane; P:biosynthetic process | Unknown function |
| GABP01025117.1 | hypothetical protein PHLGIDRAFT_25750 | | Unknown function |
| GABP01025380.1 | peptidase V | | Unknown function |
| GABP01025476.1 | hypothetical protein | | Unknown function |

| | | | |
|---|---|---|---|
| GABP01025477.1 | hypothetical protein | P:purine nucleotide metabolic process; P:GDP metabolic process; P:GMP metabolic process; F:guanylate kinase activity; F:monooxygenase activity; F:acid phosphatase activity; F:N,N-dimethylaniline monooxygenase activity; F:nucleotide binding; F:ATP binding; F:NADP binding; F:flavin adenine dinucleotide binding; F:oxidoreductase activity; C:membrane; C:integral component of membrane; P:oxidation-reduction process; P:phosphorylation; F:transferase activity; P:dephosphorylation; F:kinase activity | Unknown function |
| GABP01025633.1 | ---NA--- | | Unknown function |
| GABP01026256.1 | hypothetical protein | F:DNA binding; F:transcription factor activity, sequence-specific DNA binding; P:transcription, DNA-templated; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated; F:sequence-specific DNA binding | unknown function |
| GABP01026687.1 | ---NA--- | | Unknown function |
| GABP01026775.1 | UPF0609 partial | F:ion binding | Unknown function |
| GABP01026845.1 | acetyltransferase ribosomal N-acetylase | | Unknown function |
| GABP01026846.1 | ATPase | | Unknown function |
| GABP01026941.1 | agenet domain-containing | | Unknown function |
| GABP01027381.1 | domain-containing | | Unknown function |
| GABP01027440.1 | hypothetical protein | | Unknown function |
| GABP01028068.1 | ---NA--- | | Unknown function |
| GABP01028184.1 | ---NA--- | | Unknown function |
| GABP01028897.1 | hypothetical protein TRIADDRAFT_61890 | | Unknown function |
| GABP01029451.1 | ---NA--- | | Unknown function |
| GABP01029531.1 | PREDICTED: uncharacterized protein LOC106730680 isoform X3 | | Unknown function |
| GABP01029538.1 | ---NA--- | | Unknown function |
| GABP01029574.1 | ---NA--- | | Unknown function |
| GABP01029672.1 | ---NA--- | | Unknown function |
| GABP01029753.1 | hypothetical protein | F:DNA binding; F:transcription factor activity, sequence-specific DNA binding; P:transcription, DNA-templated; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated | unknown function |
| GABP01029941.1 | Bacl-2 | F:nucleotide binding; F:ATP binding; C:membrane; C:integral component of membrane; F:protein kinase activity; P:protein phosphorylation | Unknown function |
| GABP01030160.1 | ---NA--- | | Unknown function |
| GABP01030296.1 | hypothetical protein L484_006881 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01030297.1 | multidrug resistance | | Unknown function |

| | | | |
|---|---|---|---|
| GABP01030308.1 | hypothetical protein | F:zinc ion binding; P:tRNA processing; P:phenylalanyl-tRNA aminoacylation; F:S-adenosylmethionine-dependent methyltransferase activity; C:nucleolus; F:tRNA binding; F:magnesium ion binding; F:rRNA (uridine-N3-)-methyltransferase activity; C:nucleus; F:ATP binding; C:cytoplasm; F:metal ion binding; P:rRNA base methylation; F:phenylalanine-tRNA ligase activity; C:membrane; C:integral component of membrane | Unknown function |
| GABP01030312.1 | hypothetical protein | F:DNA binding; C:nucleus; F:zinc ion binding; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated; F:RNA polymerase II transcription factor activity, sequence-specific DNA binding; P:regulation of transcription from RNA polymerase II promoter | unknown function |
| GABP01030430.1 | ---NA--- | | Unknown function |
| GABP01030508.1 | cell envelope biogenesis | | Unknown function |
| GABP01030836.1 | hypothetical protein | P:nucleic acid phosphodiester bond hydrolysis; F:Type I site-specific deoxyribonuclease activity; P:metabolic process; F:helicase activity; P:DNA restriction-modification system; P:DNA modification; F:nucleotide binding; F:DNA binding; F:ATP binding; F:endonuclease activity; P:phosphorylation; F:kinase activity; F:hydrolase activity | Unknown function |
| GABP01030854.1 | hypothetical protein LR48_Vigan05g172300 | F:nucleotide binding; F:nucleic acid binding; C:U2AF; F:RNA binding; F:metal ion binding; P:mRNA splicing, via spliceosome | Unknown function |
| GABP01031166.1 | unnamed protein product | | Unknown function |
| GABP01031244.1 | laminin A chain | C:basement membrane; P:system development; P:basement membrane assembly | Unknown function |
| GABP01031369.1 | ---NA--- | | Unknown function |
| GABP01032115.1 | hypothetical protein | | Unknown function |
| GABP01032263.1 | CBS domain-containing | | Unknown function |
| GABP01032290.1 | hypothetical protein Pmar_PMAR005561 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01032501.1 | like subfamily A member 2 | P:social behavior; F:heat shock protein binding; P:protein import into mitochondrial matrix; P:positive regulation of ATPase activity; C:endoplasmic reticulum; F:unfolded protein binding; P:protein folding; F:ATPase activator activity; C:plasma membrane; C:integral component of mitochondrial outer membrane; C:nucleus; F:ATP binding; P:response to heat; C:extracellular space; C:cytoplasm; F:metal ion binding; C:mitochondrion; C:membrane; C:integral component of membrane | Unknown function |
| GABP01032550.1 | hypothetical protein Pmar_PMAR012722 | | Unknown function |
| GABP01032640.1 | ---NA--- | | Unknown function |
| GABP01032701.1 | ---NA--- | | Unknown function |
| GABP01032845.1 | FHA domain | F:zinc ion binding | Unknown function |
| GABP01032860.1 | unknown | | Unknown function |
| GABP01033090.1 | ---NA--- | | Unknown function |

| | | | |
|---|---|---|---|
| GABP01033743.1 | hypothetical protein BN946_scf184975.g11 | P:protein import into nucleus; C:intracellular | Unknown function |
| GABP01033943.1 | ---NA--- | | Unknown function |
| GABP01034612.1 | hypothetical protein | F:ATP binding; F:protein kinase activity; P:protein phosphorylation | Unknown function |
| GABP01034908.1 | 50S ribosomal L13 | F:structural constituent of ribosome; C:intracellular ribonucleoprotein complex; C:membrane; C:integral component of membrane; C:ribosome; P:translation | Unknown function |
| GABP01035275.1 | RNA polymerase I-specific transcription initiation factor RRN9 | F:translation initiation factor activity; P:transcription of nuclear large rRNA transcript from RNA polymerase I promoter; F:TBP-class protein binding; C:RNA polymerase I upstream activating factor complex; C:nucleolus; P:translational initiation; P:regulation of transcription from RNA polymerase I promoter | Unknown function |
| GABP01035294.1 | hypothetical protein PHLGIDRAFT_33436 | | Unknown function |
| GABP01035356.1 | ---NA--- | | Unknown function |
| GABP01035657.1 | ---NA--- | | Unknown function |
| GABP01036276.1 | hypothetical protein | | Unknown function |
| GABP01036342.1 | ---NA--- | | Unknown function |
| GABP01036611.1 | ---NA--- | | Unknown function |
| GABP01036845.1 | cytochrome P460 | | Unknown function |
| GABP01037014.1 | hypothetical protein | | Unknown function |
| GABP01037577.1 | ---NA--- | | Unknown function |
| GABP01038039.1 | helicase partial | F:helicase activity; P:double-strand break repair | Unknown function |
| GABP01038071.1 | hypothetical protein | F:translation initiation factor activity; F:zinc ion binding; F:RNA binding; P:protein complex assembly; C:mitochondrion; P:translational initiation | Unknown function |
| GABP01038323.1 | unnamed protein product | F:DNA binding; C:nucleus; F:transcription factor activity, sequence-specific DNA binding; F:zinc ion binding; P:transcription, DNA-templated; P:regulation of transcription, DNA-templated; F:RNA polymerase II transcription factor activity, sequence-specific DNA binding; P:regulation of transcription from RNA polymerase II promoter | unknown function |
| GABP01038454.1 | ---NA--- | | Unknown function |
| GABP01038542.1 | unknown | | Unknown function |
| GABP01038586.1 | BRICHOS domain-containing partial | | Unknown function |
| GABP01039439.1 | ---NA--- | | Unknown function |
| GABP01039598.1 | ---NA--- | | Unknown function |
| GABP01040237.1 | hypothetical protein Pmar_PMAR024059 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01040238.1 | hypothetical protein Pmar_PMAR025533 | | Unknown function |
| GABP01040327.1 | Uncharacterized (plasmid) | C:membrane; C:integral component of membrane | Unknown function |

| | | | |
|---|---|---|---|
| GABP01040328.1 | Uncharacterized (plasmid) | C:membrane; C:integral component of membrane | Unknown function |
| GABP01040352.1 | hypothetical protein | F:oxidoreductase activity; P:metabolic process; P:oxidation-reduction process | Unknown function |
| GABP01040562.1 | ---NA--- | | Unknown function |
| GABP01040617.1 | ---NA--- | | Unknown function |
| GABP01040782.1 | TPR repeat-containing domain | | Unknown function |
| GABP01040939.1 | tyrosine phosphatase | P:metabolic process; F:hydrolase activity | Unknown function |
| GABP01041102.1 | ---NA--- | | Unknown function |
| GABP01041714.1 | ---NA--- | | Unknown function |
| GABP01042344.1 | ---NA--- | | Unknown function |
| GABP01042783.1 | another transcription unit | | Unknown function |
| GABP01043079.1 | PREDICTED: uncharacterized protein LOC102709548 isoform X2 | F:DNA binding; F:zinc ion binding; F:transposase activity; C:membrane; C:integral component of membrane; P:transposition, DNA-mediated | unknown function |
| GABP01043498.1 | hypothetical protein Pmar_PMAR012351 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01044069.1 | PREDICTED: multimerin-1 | | Unknown function |
| GABP01044401.1 | ---NA--- | | Unknown function |
| GABP01045630.1 | ---NA--- | | Unknown function |
| GABP01045668.1 | 3-phosphoshikimate 1-carboxyvinyltransferase | P:aromatic amino acid family biosynthetic process; P:DNA recombination; F:ATP-dependent helicase activity; P:chorismate biosynthetic process; P:cellular amino acid biosynthetic process; F:helicase activity; F:nucleotide binding; F:nucleic acid binding; P:response to oxidative stress; F:ATP binding; F:3-phosphoshikimate 1-carboxyvinyltransferase activity; C:cytoplasm; F:catalytic activity; F:transferase activity; F:transferase activity, transferring alkyl or aryl (other than methyl) groups; F:hydrolase activity | Unknown function |
| GABP01045926.1 | ---NA--- | | Unknown function |
| GABP01046137.1 | unknown | | Unknown function |
| GABP01046322.1 | leucine-rich repeat-containing DDB_G0290503 | | Unknown function |
| GABP01047438.1 | ---NA--- | | Unknown function |
| GABP01047599.1 | ---NA--- | | Unknown function |
| GABP01047819.1 | hypothetical protein | | Unknown function |
| GABP01048445.1 | ---NA--- | | Unknown function |
| GABP01049480.1 | 2-polyprenylphenol hydroxylase | F:flavin adenine dinucleotide binding; F:oxidoreductase activity; P:oxidation-reduction process; P:pyrimidine nucleotide biosynthetic process; F:2 iron, 2 sulfur cluster binding; F:ferredoxin-NADP+ reductase activity | Unknown function |
| GABP01050036.1 | unknown | | Unknown function |
| GABP01050227.1 | Zn-finger in Ran binding | F:zinc ion binding | Unknown function |

| | | | |
|---|---|---|---|
| GABP01050417.1 | hypothetical protein ACA1_165800 | | Unknown function |
| GABP01050418.1 | ---NA--- | | Unknown function |
| GABP01050896.1 | ---NA--- | | Unknown function |
| GABP01051129.1 | unnamed product | | Unknown function |
| GABP01051294.1 | ---NA--- | | Unknown function |
| GABP01051417.1 | hypothetical protein Pmar_PMAR012722 | F:nucleotide binding; F:nucleic acid binding; F:ATP binding; P:metabolic process; F:helicase activity; F:hydrolase activity | Unknown function |
| GABP01052068.1 | ---NA--- | | Unknown function |
| GABP01052078.1 | hypothetical protein TGME49_070360 | F:nucleic acid binding; F:methyltransferase activity; P:methylation | Unknown function |
| GABP01052086.1 | WW domain | C:WASH complex; C:intracellular | Unknown function |
| GABP01052107.1 | hypothetical protein A1Q1_03254 | | Unknown function |
| GABP01052138.1 | protecting protein DprA protein | P:DNA mediated transformation | Unknown function |
| GABP01052591.1 | hypothetical protein | F:translation elongation factor activity; P:metabolic process; C:membrane; C:integral component of membrane; P:translational elongation; F:hydrolase activity | Unknown function |
| GABP01053266.1 | secreted containing DUF1598 | F:DNA binding; P:self proteolysis; C:cytoplasm; F:metal ion binding; P:transcription, DNA-templated; P:metabolic process; C:integral component of membrane; P:regulation of transcription, DNA-templated; F:4 iron, 4 sulfur cluster binding; F:iron-sulfur cluster binding; F:dinitrosyl-iron complex binding; F:lysozyme activity | unknown function |
| GABP01054279.1 | ---NA--- | | Unknown function |
| GABP01054488.1 | ---NA--- | | Unknown function |
| GABP01055468.1 | unnamed protein product | | Unknown function |
| GABP01055953.1 | unnamed protein product | | Unknown function |
| GABP01056034.1 | ---NA--- | | Unknown function |
| GABP01056326.1 | hypothetical protein | | Unknown function |
| GABP01056415.1 | MULTISPECIES: hypothetical protein | C:membrane; P:proteolysis; C:integral component of membrane; F:serine-type endopeptidase activity | Unknown function |
| GABP01056479.1 | ---NA--- | | Unknown function |
| GABP01056805.1 | ---NA--- | | Unknown function |
| GABP01057032.1 | hypothetical protein | F:ATPase activity; F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen; P:metabolic process; F:lyase activity; F:heme binding; F:nucleotide binding; F:ATP binding; F:iron ion binding; P:nicotinamide nucleotide metabolic process; P:oxidation-reduction process; P:phosphorylation; F:kinase activity; F:ADP-dependent NAD(P)H-hydrate dehydratase activity | Unknown function |
| GABP01057496.1 | hypothetical protein Pmar_PMAR024059 | C:membrane; C:integral component of membrane | Unknown function |

| | | | |
|---|---|---|---|
| GABP01057751.1 | ---NA--- | | Unknown function |
| GABP01057852.1 | hypothetical protein Pmar_PMAR005561 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01058055.1 | BTB POZ domain-containing partial | | Unknown function |
| GABP01058295.1 | ---NA--- | | Unknown function |
| GABP01058299.1 | unnamed protein product | F:nucleotide binding; F:DNA binding; F:ATP binding; F:nucleosome binding; P:ATP-dependent chromatin remodeling | Unknown function |
| GABP01058437.1 | aspartyl-tRNA partial | | Unknown function |
| GABP01059031.1 | hypothetical protein Pmar_PMAR024059 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01059348.1 | ---NA--- | | Unknown function |
| GABP01060337.1 | heterogeneous nuclear ribonucleo R isoform X1 | F:nucleotide binding; F:nucleic acid binding; C:intracellular ribonucleoprotein complex; C:virion; C:viral nucleocapsid; F:mRNA binding; P:axon extension involved in axon guidance | Unknown function |
| GABP01060915.1 | hypothetical protein Pmar_PMAR012722 | | Unknown function |
| GABP01061538.1 | melanoma-associated antigen 12 | | Unknown function |
| GABP01061979.1 | ---NA--- | | Unknown function |
| GABP01061989.1 | DNA-cytosine methyltransferase | F:DNA (cytosine-5-)-methyltransferase activity; P:C-5 methylation of cytosine; F:methyltransferase activity; F:transferase activity; P:methylation | Unknown function |
| GABP01063710.1 | 4-hydroxyphenylacetate 3- oxygenase component | F:oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen; F:oxidoreductase activity, acting on the CH-CH group of donors; F:flavin adenine dinucleotide binding; P:metabolic process; P:oxidation-reduction process; F:monooxygenase activity; P:phenylacetate catabolic process | Unknown function |
| GABP01063885.1 | site-specific recombinase phage integrase family | | Unknown function |
| GABP01064278.1 | ---NA--- | | Unknown function |
| GABP01064513.1 | ALK4 5 7 receptor S T kinase | F:protein kinase activity; P:transmembrane receptor protein serine/threonine kinase signaling pathway; F:protein serine/threonine kinase activity; P:protein phosphorylation; P:signal transduction by protein phosphorylation; F:transmembrane receptor protein serine/threonine kinase activity; F:nucleotide binding; F:ATP binding; F:receptor signaling protein serine/threonine kinase activity; C:membrane; C:integral component of membrane; P:phosphorylation; F:transferase activity; F:kinase activity | Unknown function |

| | | | |
|---|---|---|---|
| GABP01064610.1 | nucleo TPR-like isoform X1 | F:identical protein binding; P:regulation of translational fidelity; P:SNARE complex assembly; F:aminoacyl-tRNA editing activity; F:protein transporter activity; C:Golgi membrane; P:intracellular protein transport; F:nucleotide binding; P:isoleucyl-tRNA aminoacylation; P:tRNA aminoacylation for protein translation; F:ATP binding; F:aminoacyl-tRNA ligase activity; P:ER to Golgi vesicle-mediated transport; F:isoleucine-tRNA ligase activity; C:cytoplasm; P:Golgi vesicle docking; C:membrane; C:ER to Golgi transport vesicle membrane; C:cytosol; P:Golgi vesicle transport; P:vesicle fusion with Golgi apparatus | Unknown function |
| GABP01065017.1 | unknown | | Unknown function |
| GABP01065371.1 | ---NA--- | | Unknown function |
| GABP01065753.1 | unchacterized apicomplexan-specific serine rich low complexity | | Unknown function |
| GABP01066226.1 | heterogeneous nuclear ribonucleo U 1 isoform X1 | | Unknown function |
| GABP01066292.1 | ---NA--- | | Unknown function |
| GABP01066361.1 | hypothetical protein Z520_05011 | | Unknown function |
| GABP01066362.1 | hypothetical protein Z520_05011 | | Unknown function |
| GABP01066364.1 | hypothetical protein EMIHUDRAFT_228901 | P:carbohydrate metabolic process; P:metabolic process; F:hydrolase activity, hydrolyzing O-glycosyl compounds; F:beta-galactosidase activity; F:hydrolase activity; F:hydrolase activity, acting on glycosyl bonds | Unknown function |
| GABP01066365.1 | Ankyrin repeat and fibronectin type-III domain-containing partial | | Unknown function |
| GABP01066368.1 | ---NA--- | | Unknown function |
| GABP01067929.1 | phosphotransferase enzyme family | P:metabolic process; F:transferase activity | Unknown function |
| GABP01068255.1 | hypothetical protein | F:pyridoxal phosphate binding; P:pyridoxine biosynthetic process; C:cytoplasm; F:O-phospho-L-serine:2-oxoglutarate aminotransferase activity; F:catalytic activity; F:transaminase activity; P:L-serine biosynthetic process; P:cellular amino acid biosynthetic process; F:transferase activity | Unknown function |
| GABP01071135.1 | suppressor of tumorigenicity 14 homolog | | Unknown function |
| GABP01072333.1 | ribosomal N-acetylase | P:metabolic process; F:transferase activity; C:ribosome; F:N-acetyltransferase activity | Unknown function |
| GABP01072542.1 | ---NA--- | | Unknown function |
| GABP01073222.1 | Uncharacterized conserved UCP033563 | | Unknown function |
| GABP01073327.1 | ---NA--- | | Unknown function |
| GABP01073463.1 | outer membrane probably involved in nutrient binding | | Unknown function |

| | | | |
|---|---|---|---|
| GABP01073530.1 | neurogenic locus notch homolog 2 | P:regulation of developmental process; P:Notch signaling pathway; P:multicellular organism development; F:calcium ion binding; C:membrane; C:integral component of membrane; P:cell differentiation; P:regulation of transcription, DNA-templated; F:receptor activity | Unknown function |
| GABP01074350.1 | ---NA--- | | Unknown function |
| GABP01074364.1 | ---NA--- | | Unknown function |
| GABP01074433.1 | AAA family ATPase | | Unknown function |
| GABP01074511.1 | ---NA--- | | Unknown function |
| GABP01074604.1 | lantibiotic dehydratase | | Unknown function |
| GABP01074794.1 | Bacl-2 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01074795.1 | Bacl-2 | | Unknown function |
| GABP01074798.1 | Bacl-2 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01074799.1 | Bacl-2 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01074800.1 | Bacl-2 | F:ATP binding; F:ribokinase activity; C:membrane; C:integral component of membrane; P:glutathione biosynthetic process; P:D-ribose metabolic process; F:glutathione synthase activity; P:carbohydrate phosphorylation | Unknown function |
| GABP01074970.1 | ---NA--- | | Unknown function |
| GABP01075697.1 | regulator of nonsense transcripts 1 ( RENT1) | | Unknown function |
| GABP01075784.1 | hypothetical protein RP20_CCG023621 | F:nucleic acid binding; F:RNA binding | Unknown function |
| GABP01077330.1 | hypothetical protein Pmar_PMAR012722 | | Unknown function |
| GABP01078536.1 | SPRY domain-containing | C:intracellular; F:molecular_function | Unknown function |
| GABP01079690.1 | SPRY domain-containing | C:intracellular; F:molecular_function | Unknown function |
| GABP01080601.1 | ---NA--- | | Unknown function |
| GABP01080645.1 | ---NA--- | | Unknown function |
| GABP01080707.1 | von Willebrand factor D and EGF domain-containing | C:membrane; C:integral component of membrane | Unknown function |
| GABP01081547.1 | ---NA--- | | Unknown function |
| GABP01081720.1 | DUF1640 domain-containing | P:metabolic process; F:serine-type endopeptidase activity; F:lipid binding; F:peptidase activity; P:lipoprotein metabolic process; C:ribosome; C:extracellular region; P:transport; C:intracellular; P:translation; F:nucleotide binding; P:tRNA aminoacylation for protein translation; P:lipid transport; F:ATP binding; F:aminoacyl-tRNA ligase activity; F:structural constituent of ribosome; C:intracellular ribonucleoprotein complex; C:membrane; P:proteolysis; C:integral component of membrane | Unknown function |
| GABP01081812.1 | ---NA--- | | Unknown function |
| GABP01082448.1 | hypothetical protein FAVG1_11046 | | Unknown function |

| | | | |
|---|---|---|---|
| GABP01082450.1 | hypothetical protein FAVG1_11046 | | Unknown function |
| GABP01082452.1 | ---NA--- | | Unknown function |
| GABP01087629.1 | uncharacterized domain | C:membrane; C:integral component of membrane | Unknown function |
| GABP01088445.1 | conserved hypothetical protein | F:metal ion binding | Unknown function |
| GABP01088446.1 | ---NA--- | | Unknown function |
| GABP01088905.1 | ---NA--- | | Unknown function |
| GABP01089094.1 | predicted protein | F:voltage-gated anion channel activity; P:regulation of anion transmembrane transport; P:transmembrane transport; C:mitochondrial outer membrane; P:anion transport; P:regulation of anion transport | Unknown function |
| GABP01091024.1 | ---NA--- | | Unknown function |
| GABP01094009.1 | ---NA--- | | Unknown function |
| GABP01095118.1 | ---NA--- | | Unknown function |
| GABP01095119.1 | ---NA--- | | Unknown function |
| GABP01095149.1 | MULTISPECIES: capsule biosynthesis | | Unknown function |
| GABP01095254.1 | dystonin isoform X16 | | Unknown function |
| GABP01096316.1 | histidinol dehydrogenase | | Unknown function |
| GABP01099011.1 | hypothetical protein AQZ50_00545 | | Unknown function |
| GABP01099095.1 | centriolin isoform X1 | P:cell division; P:mitotic cell cycle | Unknown function |
| GABP01099956.1 | Hypothetical protein SS50377_14564 | | Unknown function |
| GABP01100784.1 | ---NA--- | | Unknown function |
| GABP01100847.1 | ---NA--- | | Unknown function |
| GABP01101975.1 | unknown | | Unknown function |
| GABP01102360.1 | hypothetical protein Pmar_PMAR025533 | | Unknown function |
| GABP01102652.1 | hypothetical protein AXG93_1112s1270 | | Unknown function |
| GABP01102653.1 | hypothetical protein AXG93_1112s1270 | | Unknown function |
| GABP01103390.1 | hypothetical protein | | Unknown function |
| GABP01105470.1 | hypothetical protein EMIHUDRAFT_248901 | C:membrane; C:integral component of membrane | Unknown function |
| GABP01107066.1 | ---NA--- | | Unknown function |
| GABP01107555.1 | NEDD4-binding 2-like 1 | P:biological_process; C:cellular_component; F:molecular_function | Unknown function |
| GABP01108206.1 | ---NA--- | | Unknown function |

| | | | |
|---|---|---|---|
| GABP01108208.1 | hypothetical protein | | Unknown function |
| GABP01108587.1 | ---NA--- | | Unknown function |
| GABP01109111.1 | unknown | P:nucleotide catabolic process; F:hydrolase activity | Unknown function |
| GABP01109810.1 | ---NA--- | | Unknown function |
| GABP01110471.1 | ---NA--- | | Unknown function |
| GABP01111334.1 | hypothetical protein HH1059_810 | C:viral envelope; P:fusion of virus membrane with host plasma membrane; F:host cell surface receptor binding | Unknown function |
| GABP01112125.1 | ---NA--- | | Unknown function |
| GABP01113342.1 | hypothetical protein | F:DNA binding; F:ATP binding; P:transcription, DNA-templated; P:phosphorelay signal transduction system; C:membrane; C:integral component of membrane; P:regulation of transcription, DNA-templated; F:protein kinase activity; F:ADP binding; P:protein phosphorylation; C:intracellular | unknown function |

Tableau 4.S2. Gene ontology terms significantly enriched and depleted in the chromatin enriched samples

| GO-ID | Term | Category | P-Value | #Test | #Ref | #notAnnotTest | #notAnnotRef | ratio test | ratio ref | test/ref | Over/Under |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GO:0005635 | nuclear envelope | C | 8,56E-03 | 7 | 82 | 626 | 23615 | 0,01 | 0,00 | 3,22 | OVER |
| GO:0005654 | nucleoplasm | C | 1,20E-03 | 23 | 410 | 610 | 23287 | 0,04 | 0,02 | 2,14 | OVER |
| GO:0005730 | nucleolus | C | 9,91E-04 | 19 | 301 | 614 | 23396 | 0,03 | 0,01 | 2,41 | OVER |
| GO:0005739 | mitochondrion | C | 9,05E-04 | 43 | 932 | 590 | 22765 | 0,07 | 0,04 | 1,78 | OVER |
| GO:0016023 | cytoplasmic, membrane-bounded vesicle | C | 2,45E-04 | 20 | 295 | 613 | 23402 | 0,03 | 0,01 | 2,59 | OVER |
| GO:0031981 | nuclear lumen | C | 3,51E-05 | 38 | 675 | 595 | 23022 | 0,06 | 0,03 | 2,18 | OVER |
| GO:0043234 | protein complex | C | 7,75E-06 | 75 | 1620 | 558 | 22077 | 0,13 | 0,07 | 1,83 | OVER |
| GO:0044444 | cytoplasmic part | C | 5,19E-06 | 169 | 4533 | 464 | 19164 | 0,36 | 0,24 | 1,54 | OVER |
| GO:0005829 | cytosol | C | 4,60E-08 | 84 | 1656 | 549 | 22041 | 0,15 | 0,08 | 2,04 | OVER |
| GO:0005737 | cytoplasm | C | 4,36E-09 | 254 | 6873 | 379 | 16824 | 0,67 | 0,41 | 1,64 | OVER |
| GO:0043231 | intracellular membrane-bounded organelle | C | 2,37E-11 | 210 | 5096 | 423 | 18601 | 0,50 | 0,27 | 1,81 | OVER |
| GO:0005634 | nucleus | C | 6,85E-17 | 129 | 2194 | 504 | 21503 | 0,26 | 0,10 | 2,51 | OVER |
| GO:0016853 | isomerase activity | F | 4,68E-03 | 23 | 447 | 610 | 23250 | 0,04 | 0,02 | 1,96 | OVER |
| GO:0016787 | hydrolase activity | F | 2,48E-03 | 107 | 2997 | 526 | 20700 | 0,20 | 0,14 | 1,41 | OVER |
| GO:0003729 | mRNA binding | F | 1,81E-03 | 9 | 97 | 624 | 23600 | 0,01 | 0,00 | 3,51 | OVER |
| GO:0016746 | transferase activity, transferring acyl groups | F | 8,60E-04 | 28 | 515 | 605 | 23182 | 0,05 | 0,02 | 2,08 | OVER |
| GO:0008233 | peptidase activity | F | 6,28E-04 | 7 | 789 | 626 | 22908 | 0,01 | 0,03 | 0,32 | UNDER |
| GO:0016301 | kinase activity | F | 4,60E-04 | 30 | 2012 | 603 | 21685 | 0,05 | 0,09 | 0,54 | UNDER |
| GO:0016765 | transferase activity, transferring alkyl or aryl groups | F | 1,40E-04 | 10 | 82 | 623 | 23615 | 0,02 | 0,00 | 4,62 | OVER |
| GO:0022857 | transmembrane transporter activity | F | 1,14E-04 | 10 | 1067 | 623 | 22630 | 0,02 | 0,05 | 0,34 | UNDER |
| GO:0016491 | oxidoreductase activity | F | 9,74E-05 | 81 | 1934 | 552 | 21763 | 0,15 | 0,09 | 1,65 | OVER |
| GO:0003723 | RNA binding | F | 1,00E-11 | 71 | 1067 | 562 | 22630 | 0,13 | 0,05 | 2,68 | OVER |
| GO:0043167 | ion binding | F | 9,29E-13 | 182 | 4049 | 451 | 19648 | 0,40 | 0,21 | 1,96 | OVER |
| GO:0016874 | ligase activity | F | 6,93E-13 | 48 | 516 | 585 | 23181 | 0,08 | 0,02 | 3,69 | OVER |
| GO:0003824 | catalytic activity | F | 3,87E-13 | 339 | 9250 | 294 | 14447 | 1,15 | 0,64 | 1,80 | OVER |
| GO:0017111 | nucleoside-triphosphatase activity | F | 1,74E-16 | 92 | 1301 | 541 | 22396 | 0,17 | 0,06 | 2,93 | OVER |
| GO:0005488 | binding | F | 1,68E-18 | 243 | 5351 | 390 | 18346 | 0,62 | 0,29 | 2,14 | OVER |
| GO:0004386 | helicase activity | F | 3,91E-20 | 46 | 293 | 587 | 23404 | 0,08 | 0,01 | 6,26 | OVER |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| GO:0006605 | protein targeting | P | 9,77E-03 | 13 | 217 | 620 | 23480 | 0,02 | 0,01 | 2,27 | OVER |
| GO:0006457 | protein folding | P | 7,30E-03 | 22 | 443 | 611 | 23254 | 0,04 | 0,02 | 1,89 | OVER |
| GO:0051276 | chromosome organization | P | 7,20E-03 | 22 | 442 | 611 | 23255 | 0,04 | 0,02 | 1,89 | OVER |
| GO:0006996 | organelle organization | P | 7,14E-03 | 37 | 862 | 596 | 22835 | 0,06 | 0,04 | 1,64 | OVER |
| GO:0046907 | intracellular transport | P | 6,27E-03 | 22 | 429 | 611 | 23268 | 0,04 | 0,02 | 1,95 | OVER |
| GO:0006399 | tRNA metabolic process | P | 5,43E-03 | 12 | 184 | 621 | 23513 | 0,02 | 0,01 | 2,47 | OVER |
| GO:0009056 | catabolic process | P | 5,20E-03 | 50 | 1243 | 583 | 22454 | 0,09 | 0,06 | 1,55 | OVER |
| GO:0006464 | cellular protein modification process | P | 2,91E-03 | 36 | 2130 | 597 | 21567 | 0,06 | 0,10 | 0,61 | UNDER |
| GO:0006259 | DNA metabolic process | P | 8,85E-04 | 28 | 518 | 605 | 23179 | 0,05 | 0,02 | 2,07 | OVER |
| GO:0016043 | cellular component organization | P | 8,62E-04 | 67 | 1638 | 566 | 22059 | 0,12 | 0,07 | 1,59 | OVER |
| GO:0006913 | nucleocytoplasmic transport | P | 3,52E-04 | 16 | 213 | 617 | 23484 | 0,03 | 0,01 | 2,86 | OVER |
| GO:0022618 | ribonucleoprotein complex assembly | P | 3,26E-04 | 14 | 169 | 619 | 23528 | 0,02 | 0,01 | 3,15 | OVER |
| GO:0022613 | ribonucleoprotein complex biogenesis | P | 2,84E-04 | 27 | 466 | 606 | 23231 | 0,04 | 0,02 | 2,22 | OVER |
| GO:0006091 | generation of precursor metabolites and energy | P | 1,52E-05 | 37 | 621 | 596 | 23076 | 0,06 | 0,03 | 2,31 | OVER |
| GO:0006790 | sulfur compound metabolic process | P | 1,28E-05 | 22 | 276 | 611 | 23421 | 0,04 | 0,01 | 3,06 | OVER |
| GO:0010467 | gene expression | P | 6,36E-07 | 58 | 1061 | 575 | 22636 | 0,10 | 0,05 | 2,15 | OVER |
| GO:0051186 | cofactor metabolic process | P | 3,24E-07 | 34 | 461 | 599 | 23236 | 0,06 | 0,02 | 2,86 | OVER |
| GO:0071704 | organic substance metabolic process | P | 8,61E-09 | 202 | 5181 | 431 | 18516 | 0,47 | 0,28 | 1,67 | OVER |
| GO:0044238 | primary metabolic process | P | 8,61E-09 | 202 | 5181 | 431 | 18516 | 0,47 | 0,28 | 1,67 | OVER |
| GO:0009058 | biosynthetic process | P | 1,96E-10 | 137 | 2944 | 496 | 20753 | 0,28 | 0,14 | 1,95 | OVER |
| GO:0006520 | cellular amino acid metabolic process | P | 8,21E-12 | 46 | 516 | 587 | 23181 | 0,08 | 0,02 | 3,52 | OVER |
| GO:0006397 | mRNA processing | P | 8,22E-14 | 32 | 215 | 601 | 23482 | 0,05 | 0,01 | 5,82 | OVER |
| GO:0016070 | RNA metabolic process | P | 2,29E-14 | 44 | 395 | 589 | 23302 | 0,07 | 0,02 | 4,41 | OVER |
| GO:0090304 | nucleic acid metabolic process | P | 1,81E-14 | 69 | 886 | 564 | 22811 | 0,12 | 0,04 | 3,15 | OVER |
| GO:0044699 | single-organism process | P | 2,48E-17 | 198 | 4092 | 435 | 19605 | 0,46 | 0,21 | 2,18 | OVER |
| GO:0044237 | cellular metabolic process | P | 1,76E-23 | 274 | 5871 | 359 | 17826 | 0,76 | 0,33 | 2,32 | OVER |
| GO:0044281 | small molecule metabolic process | P | 4,97E-25 | 135 | 1882 | 498 | 21815 | 0,27 | 0,09 | 3,14 | OVER |
| GO:0034641 | cellular nitrogen compound metabolic process | P | 2,63E-30 | 208 | 3427 | 425 | 20270 | 0,49 | 0,17 | 2,89 | OVER |

Tableau 4.S3. Differentially expressed proteins between LD6 and LD18

| | Accession | Sequence Desription | Functional characterization | LD6 Average Peak Area Value | LD18 Average Peak Area Value | Log2 Fold Change | p value |
|---|---|---|---|---|---|---|---|
| Over represented in LD6 | GABP01076222.1 | plastid NAP50 | Chloroplast | 4883334 | 166801 | -3,093060381 | 0,01932 |
| | GABP01008608.1 | Nrp1p | Nuclear transport | 7640334 | 635001 | -2,633372129 | 0,04097 |
| | GABP01072542.1 | ---NA--- | Unknown function | 23036668 | 3175001 | -2,370655159 | 0,03529 |
| | GABP01033001.1 | ATP-dependent RNA helicase DHX36 | RNA Helicase | 3713334 | 513801 | -2,346303596 | 0,04393 |
| | GABP01002608.1 | heterogeneous nuclear ribonucleo U 1 | mRNA processing | 6530001 | 1029801 | -2,330057832 | 0,04047 |
| Over represented in LD18 | GABP01005025.1 | photosynthetic NDH subunit of lumenal location chloroplastic | Chloroplast | 0 | 592234 | 22,4697239 | 0,00003 |
| | GABP01057848.1 | Methylthioribose-1-phosphate isomerase | Enzyme | 0 | 295668 | 22,0103872 | 0,00334 |
| | GABP01107334.1 | GTP-binding nuclear ran | Nuclear transport | 0 | 291001 | 21,11289621 | 0,00730 |
| | GABP01023168.1 | UDP-glucose 4-epimerase | Enzyme | 0 | 69434 | 20,15972685 | 0,01320 |
| | GABP01025042.1 | NAD(P) mitochondrial | Mitochondria | 0 | 191934 | 20,02147722 | 0,01212 |
| | GABP01026845.1 | acetyltransferase ribosomal N-acetylase | Unknown function | 0 | 147934 | 19,79390831 | 0,01267 |
| | GABP01028133.1 | heat shock HSP60 | Chaperone | 0 | 106401 | 19,70359283 | 0,01692 |
| | GABP01058616.1 | cyclopropane fatty acid synthase | Enzyme | 236068 | 17749668 | 9,24043361 | 0,00000 |
| | GABP01007195.1 | Exportin-1 | Nuclear transport | 841668 | 56393334 | 8,920520159 | 0,00000 |
| | GABP01086285.1 | anthranilate synthase | Enzyme | 1300001 | 55400001 | 8,361014502 | 0,00000 |
| | GABP01073222.1 | Uncharacterized conserved UCP033563 | Unknown function | 902668 | 26969001 | 7,654597664 | 0,00001 |
| | GABP01034904.1 | conserved hypothetical protein | Transport | 6820001 | 106003334 | 6,531561398 | 0,00002 |
| | GABP01092866.1 | NAC domain-containing | Nucleic acid binding | 359268 | 4326668 | 6,34239031 | 0,00008 |
| | GABP01061975.1 | probable phosphatase 2C 59 | Protein modification | 276334 | 2835001 | 6,23157197 | 0,00008 |

| | | | | | | |
|---|---|---|---|---|---|---|
| GABP01025839.1 | elongation factor ef-3 | Translation | 1466668 | 10727668 | 5,797392739 | 0,00038 |
| GABP01005642.1 | DNA-binding SMUBP-2 | Nucleic acid binding | 1906668 | 10956668 | 5,400143656 | 0,00015 |
| GABP01071717.1 | polysaccharide deacetylase | Enzyme | 3813334 | 18906668 | 5,205959198 | 0,00022 |
| GABP01091442.1 | AAA family | No chromatin function | 184334 | 1059001 | 5,178999846 | 0,02041 |
| GABP01041716.1 | alanine--tRNA ligase | tRNA | 2350001 | 5360001 | 4,198840611 | 0,00210 |
| GABP01020446.1 | far upstream element-binding partial | RNA binding | 660001 | 2185001 | 4,176359716 | 0,00084 |
| GABP01108587.1 | ---NA--- | Unknown function | 3163334 | 5003334 | 3,847790957 | 0,00565 |
| GABP01024384.1 | sex-lethal homolog isoform X1 | RNA binding | 654668 | 1510334 | 3,720995935 | 0,00306 |
| GABP01003893.1 | heat shock 90 | Chaperone | 5036668 | 14166668 | 3,700043473 | 0,00326 |
| GABP01017373.1 | hypothetical protein Pmar_PMAR024059 | Unknown function | 7376668 | 10440001 | 3,370127508 | 0,00948 |
| GABP01083012.1 | 116 kDa U5 small nuclear ribonucleo component | mRNA processing | 1856668 | 2321001 | 3,29880412 | 0,01381 |
| GABP01012084.1 | acetyl- carboxylase | Mitochondria | 856001 | 1288334 | 2,913042616 | 0,02415 |
| GABP01059031.1 | hypothetical protein Pmar_PMAR024059 | Unknown function | 3426668 | 4230001 | 2,897728572 | 0,01563 |
| GABP01021184.1 | nucleolin 2-like isoform X6 | RNA binding | 508001 | 643001 | 2,835109925 | 0,03543 |
| GABP01046117.1 | Unknown | Unknown function | 1203001 | 1410001 | 2,323963454 | 0,04122 |